The Addison-Wesley Series in Economics

Abel/Bernanke
Macroeconomics

Bade/Parkin
Foundations of Microeconomics

Bade/Parkin
Foundations of Macroeconomics

Bierman/Fernandez
Game Theory with Economic Applications

Binger/Hoffman
Microeconomics with Calculus

Boyer
Principles of Transportation Economics

Branson
Macroeconomic Theory and Policy

Bruce
Public Finance and the American Economy

Byrns/Stone
Economics

Carlton/Perloff
Modern Industrial Organization

Caves/Frankel/Jones
World Trade and Payments: An Introduction

Chapman
Environmental Economics: Theory, Application, and Policy

Cooter/Ulen
Law and Economics

Downs
An Economic Theory of Democracy

Eaton/Mishkin
Online Readings to Accompany The Economics of Money, Banking, and Financial Markets

Ehrenberg/Smith
Modern Labor Economics

Ekelund/Tollison
Economics: Private Markets and Public Choice

Fusfeld
The Age of the Economist

Gerber
International Economics

Ghiara
Learning Economics: A Practical Workbook

Gordon
Macroeconomics

Gregory
Essentials of Economics

Gregory/Stuart
Russian and Soviet Economic Performance and Structure

Hartwick/Olewiler
The Economics of Natural Resource Use

Hubbard
Money, the Financial System, and the Economy

Hughes/Cain
American Economic History

Husted/Melvin
International Economics

Jehle/Reny
Advanced Microeconomic Theory

Klein
Mathematical Methods for Economics

Krugman/Obstfeld
International Economics: Theory and Policy

Laidler
The Demand for Money: Theories, Evidence, and Problems

Leeds/von Allmen
The Economics of Sports

Lipsey/Courant/Ragan
Economics

McCarty
Dollars and Sense: An Introduction to Economics

Melvin
International Money and Finance

Miller
Economics Today

Miller/Benjamin/North
The Economics of Public Issues

Mills/Hamilton
Urban Economics

Mishkin
The Economics of Money, Banking, and Financial Markets

Parkin
Economics

Parkin/Bade
Economics in Action Software

Perloff
Microeconomics

Phelps
Health Economics

Riddell/Shackelford/Stamos/Schneider
Economics: A Tool for Critically Understanding Society

Ritter/Silber/Udell
Principles of Money, Banking, and Financial Markets

Rohlf
Introduction to Economic Reasoning

Ruffin/Gregory
Principles of Economics

Sargent
Rational Expectations and Inflation

Scherer
Industry Structure, Strategy, and Public Policy

Schotter
Microeconomics: A Modern Approach

Stock/Watson
Introduction to Econometrics

Studenmund
Using Econometrics: A Practical Guide

Tietenberg
Environmental and Natural Resource Economics

Tietenberg
Environmental Economics and Policy

Todaro/Smith
Economic Development

Waldman/Jensen
Industrial Organization: Theory and Practice

Williamson
Macroeconomics

The Economics of Money, Banking, and Financial Markets

Seventh Edition

Frederic S. Mishkin
Columbia University

Boston San Francisco New York
London Toronto Sydney Tokyo Singapore Madrid
Mexico City Munich Paris Cape Town Hong Kong Montreal

Editor in Chief: Denise Clinton
Acquisitions Editor: Victoria Warneck
Executive Development Manager: Sylvia Mallory
Development Editor: Jane Tufts
Production Supervisor: Meredith Gertz
Text Design: Studio Montage
Cover Design: Regina Hagen Kolenda and Studio Montage
Composition: Argosy Publishing
Senior Manufacturing Supervisor: Hugh Crawford
Senior Marketing Manager: Barbara LeBuhn
Cover images: © PhotoDisc
Media Producer: Melissa Honig
Supplements Editor: Diana Theriault

Credits to copyrighted material appear on p. C-1, which constitutes a continuation of the copyright page.

Library of Congress Cataloguing-in-Publication Data

Mishkin, Frederic S.
The economics of money, banking, and financial markets / Frederic S. Mishkin.—7th ed.
p. cm. — (The Addison-Wesley series in economics)
Supplemented by a subscription to a companion web site.
Includes bibliographical references and index.
ISBN 0-321-12235-6
1. Finance. 2. Money. 3. Banks and banking. I. Title. II. Series.
HG173.M632 2004
332—dc21

2003041912

Copyright © 2004 by Frederic S. Mishkin. All rights reserved. No part of this publication may be reproduced, stored in a retrieval system, or transmitted, in any form or by any means, electronic, mechanical, photocopying, recording, or otherwise, without the prior written permission of the publisher. Printed in the United States of America.

1 2 3 4 5 6 7 8 9 10—DOW—06050403

To Sally

CONTENTS IN BRIEF

PART I *Introduction 1*

1 Why Study Money, Banking, and Financial Markets? 3
2 An Overview of the Financial System 23
3 What Is Money? 44

PART II *Financial Markets 59*

4 Understanding Interest Rates 61
5 The Behavior of Interest Rates 85
6 The Risk and Term Structure of Interest Rates 120
7 The Stock Market, the Theory of Rational Expectations, and the Efficient Market Hypothesis 141

PART III *Financial Institutions 167*

8 An Economic Analysis of Financial Structure 169
9 Banking and the Management of Financial Institutions 201
10 Banking Industry: Structure and Competition 229
11 Economic Analysis of Banking Regulation 260
12 Nonbank Finance 287
13 Financial Derivatives 309

PART IV *Central Banking and the Conduct of Monetary Policy 333*

14 Structure of Central Banks and the Federal Reserve System 335
15 Multiple Deposit Creation and the Money Supply Process 357
16 Determinants of the Money Supply 374
17 Tools of Monetary Policy 393
18 Conduct of Monetary Policy: Goals and Targets 411

PART V *International Finance and Monetary Policy 433*

19 The Foreign Exchange Market 435
20 The International Financial System 462
21 Monetary Policy Strategy: The International Experience 487

PART VI *Monetary Theory* 515

22 The Demand for Money 517
23 The Keynesian Framework and the *ISLM* Model 536
24 Monetary and Fiscal Policy in the *ISLM* Model 561
25 Aggregate Demand and Supply Analysis 582
26 Transmission Mechanisms of Monetary Policy: The Evidence 603
27 Money and Inflation 632
28 Rational Expectations: Implications for Policy 658

CONTENTS

PART I *Introduction 1*

CHAPTER 1
WHY STUDY MONEY, BANKING, AND FINANCIAL MARKETS? 3

Preview 3
Why Study Financial Markets? 3
- The Bond Market and Interest Rates 3
- The Stock Market 5
- The Foreign Exchange Market 5

Why Study Banking and Financial Institutions? 7
- Structure of the Financial System 7
- Banks and Other Financial Institutions 8
- Financial Innovation 8

Why Study Money and Monetary Policy? 8
- Money and Business Cycles 9
- Money and Inflation 10
- Money and Interest Rates 12
- Conduct of Monetary Policy 12
- Fiscal Policy and Monetary Policy 12

How We Will Study Money, Banking, and Financial Markets 13
- Exploring the Web 14

Concluding Remarks 17
Summary, Key Terms, Questions and Problems, and Web Exercises 17

Appendix to Chapter 1
Defining Aggregate Output, Income, the Price Level, and the Inflation Rate 20

Aggregate Output and Income 20
Real Versus Nominal Magnitudes 20
Aggregate Price Level 21
Growth Rates and the Inflation Rate 22

CHAPTER 2
AN OVERVIEW OF THE FINANCIAL SYSTEM 23

Preview 23
Function of Financial Markets 23
Structure of Financial Markets 25
Debt and Equity Markets 25
Primary and Secondary Markets 26
Exchanges and Over-the-Counter Markets 27
Money and Capital Markets 27
Internationalization of Financial Markets 28
International Bond Market, Eurobonds, and Eurocurrencies 28
World Stock Markets 29
Function of Financial Intermediaries 29
Transaction Costs 29

Following the Financial News Foreign Stock Market Indexes 30

Box 1 Global: The Importance of Financial Intermediaries to Securities Markets: An International Comparison 31

Risk Sharing 31
Asymmetric Information: Adverse Selection and Moral Hazard 32
Financial Intermediaries 34
Depository Institutions 34
Contractual Savings Institutions 35
Investment Intermediaries 37
Regulation of the Financial System 37
Increasing Information Available to Investors 39
Ensuring the Soundness of Financial Intermediaries 39
Financial Regulation Abroad 40
Summary, Key Terms, Questions and Problems, and Web Exercises 41

CHAPTER 3
WHAT IS MONEY? 44

Preview 44
Meaning of Money 44
Functions of Money 45
Medium of Exchange 45
Unit of Account 46
Store of Value 47
Evolution of the Payments System 48
Commodity Money 48
Fiat Money 48
Checks 48

Box 1 Global: Birth of the Euro: Will It Benefit Europe? 49

Electronic Payment 50

Box 2 E-Finance: Why Are Scandinavians So Far Ahead of Americans in Using Electronic Payments? 50

E-Money 51

Measuring Money 51

The Federal Reserve's Monetary Aggregates 51

Box 3 E-Finance: Are We Headed for a Cashless Society? 52

Following the Financial News The Monetary Aggregates 54

How Reliable Are the Money Data? 55

Summary, Key Terms, Questions and Problems, and Web Exercises 56

PART II *Financial Markets* 59

CHAPTER 4
UNDERSTANDING INTEREST RATES 61

Preview 61

Measuring Interest Rates 61

Present Value 61

Four Types of Credit Market Instruments 63

Yield to Maturity 64

Box 1 Global: Negative T-Bill Rates? Japan Shows the Way 69

Other Measures of Interest Rates 69

Current Yield 70

Yield on a Discount Basis 71

Application Reading the *Wall Street Journal*: The Bond Page 72

Following the Financial News Bond Prices and Interest Rates 73

The Distinction Between Interest Rates and Returns 75

Maturity and the Volatility of Bond Returns: Interest-Rate Risk 78

Box 2 Helping Investors to Select Desired Interest-Rate Risk 79

Summary 79

The Distinction Between Real and Nominal Interest Rates 79

Box 3 With TIPS, Real Interest Rates Have Become Observable in the United States 82

Summary, Key Terms, Questions and Problems, and Web Exercises 82

CHAPTER 5
THE BEHAVIOR OF INTEREST RATES 85

Preview 85

Determinants of Asset Demand 85

Wealth 86

Expected Returns 86
Risk 87
Liquidity 87
Theory of Asset Demand 87
Supply and Demand in the Bond Market 87
Demand Curve 88
Supply Curve 90
Market Equilibrium 90
Supply and Demand Analysis 91
Loanable Funds Framework 91
Changes in Equilibrium Interest Rates 93
Shifts in the Demand for Bonds 93
Shifts in the Supply of Bonds 97
Application Changes in the Equilibrium Interest Rate Due to Expected Inflation or Business Cycle Expansions 99
Changes in Expected Inflation: The Fisher Effect 99
Business Cycle Expansion 100
Application Explaining Low Japanese Interest Rates 103
Application Reading the *Wall Street Journal* "Credit Markets" Column 103
Following the Financial News The "Credit Markets" Column 104
Supply and Demand in the Market for Money: The Liquidity Preference Framework 105
Changes in Equilibrium Interest Rates in the Liquidity Reference Framework 107
Shifts in the Demand for Money 107
Shifts in the Supply of Money 108
Application Changes in the Equilibrium Interest Rate Due to Changes in Income, the Price Level, or the Money Supply 108
Changes in Income 108
Changes in the Price Level 108
Changes in the Money Supply 109
Following the Financial News Forecasting Interest Rates 111
Application Money and Interest Rates 112
Does a Higher Rate of Growth of the Money Supply Lower Interest Rates? 114
Summary, Key Terms, Questions and Problems, and Web Exercises 117

CHAPTER 6
THE RISK AND TERM STRUCTURE OF INTEREST RATES 120
Preview 120
Risk Structure of Interest Rates 120
Default Risk 120
Application The Enron Bankruptcy and the Baa-Aaa Spread 124

Liquidity 125
Income Tax Considerations 125
Summary 127

Application Effects of the Bush Tax Cut on Bond Interest Rates 127

Term Structure of Interest Rates 127

Following the Financial News Yield Curves 128

Expectations Theory 129
Segmented Markets Theory 132
Liquidity Premium and Preferred Habitat Theories 133
Evidence on the Term Structure 136
Summary 137

Application Interpreting Yield Curves, 1980–2003 137

Summary, Key Terms, Questions and Problems, and Web Exercises 138

CHAPTER 7
THE STOCK MARKET, THE THEORY OF RATIONAL EXPECTATIONS, AND THE EFFICIENT MARKET HYPOTHESIS 141

Preview 141

Computing the Price of Common Stock 141
The One-Period Valuation Model 142
The Generalized Dividend Valuation Model 143
The Gordon Growth Model 143

How the Market Sets Security Prices 144

Application Monetary Policy and Stock Prices 146

Application The September 11 Terrorist Attacks, the Enron Scandal, and the Stock Market 146

The Theory of Rational Expectations 147
Formal Statement of the Theory 148
Rationale Behind the Theory 149
Implications of the Theory 149

The Efficient Markets Hypothesis: Rational Expectations in Financial Markets 150
Rationale Behind the Hypothesis 151
Stronger Version of the Efficient Market Hypothesis 152

Evidence on the Efficient Market Hypothesis 153
Evidence in Favor of Market Efficiency 153

Application Should Foreign Exchange Rates Follow a Random Walk? 155

Evidence Against Market Efficiency 156
Overview of the Evidence on the Efficient Market Hypothesis 158

Application Practical Guide to Investing in the Stock Market 158

How Valuable Are Published Reports by Investment Advisers? 158

Following the Financial News Stock Prices 159

Box 1 Should You Hire an Ape as Your Investment Adviser? 160

Should You Be Skeptical of Hot Tips? 160
Do Stock Prices Always Rise When There Is Good News? 161
Efficient Market Prescription for the Investor 161

Evidence on Rational Expectations in Other Markets 162

Application What Do the Black Monday Crash of 1987 and the Tech Crash of 2000 Tell Us About Rational Expectations and Efficient Markets? 163

Summary, Key Terms, Questions and Problems, and Web Exercises 164

PART III *Financial Institutions* 167

CHAPTER 8
AN ECONOMIC ANALYSIS OF FINANCIAL STRUCTURE 169

Preview 169

Basic Puzzles About Financial Structure Throughout the World 169

Transaction Costs 173
How Transaction Costs Influence Financial Structure 173
How Financial Intermediaries Reduce Transaction Costs 173

Asymmetric Information: Adverse Selection and Moral Hazard 174

The Lemons Problem: How Adverse Selection Influences Financial Structure 175
Lemons in the Stock and Bond Markets 175
Tools to Help Solve Adverse Selection Problems 176

Box 1 The Enron Implosion and the Arthur Andersen Conviction 178

How Moral Hazard Affects the Choice Between Debt and Equity Contracts 180
Moral Hazard in Equity Contracts: The Principal–Agent Problem 181
Tools to Help Solve the Principal–Agent Problem 182

Box 2 E-Finance: Venture Capitalists and the High-Tech Sector 183

How Moral Hazard Influences Financial Structure in Debt Markets 184
Tools to Help Solve Moral Hazard in Debt Contracts 184
Summary 186

Application Financial Development and Economic Growth 187

Financial Crises and Aggregate Economic Activity 189
Factors Causing Financial Crises 189

Application Financial Crises in the United States 191

Box 3 Case Study of a Financial Crisis: The Great Depression 194

Application Financial Crises in Emerging-Market Countries: Mexico, 1994–1995; East Asia, 1997–1998; and Argentina, 2001–2002 194

Summary, Key Terms, Questions and Problems, and Web Exercises 199

CHAPTER 9
BANKING AND THE MANAGEMENT OF FINANCIAL INSTITUTIONS 201

Preview 201
The Bank Balance Sheet 201
Liabilities 201
Assets 204
Basic Banking 205
General Principles of Bank Management 208
Liquidity Management and the Role of Reserves 208
Asset Management 211
Liability Management 212
Capital Adequacy Management 213
Application Strategies for Managing Bank Capital 215
Application Did the Capital Crunch Cause a Credit Crunch in the Early 1990s? 216
Managing Credit Risk 217
Screening and Monitoring 217
Long-Term Customer Relationships 218
Loan Commitments 219
Collateral and Compensating Balances 219
Credit Rationing 220
Managing Interest-Rate Risk 220
Gap and Duration Analysis 221
Application Strategies for Managing Interest-Rate Risk 222
Off-Balance-Sheet Activities 223
Loan Sales 223
Generation of Fee Income 223
Trading Activities and Risk Management Techniques 224
Box 1 Global: Barings, Daiwa, Sumitomo, and Allied Irish: Rogue Traders and the Principal–Agent Problem 225
Summary, Key Terms, Questions and Problems, and Web Exercises 226

CHAPTER 10
BANKING INDUSTRY: STRUCTURE AND COMPETITION 229

Preview 229
Historical Development of the Banking System 229
Multiple Regulatory Agencies 231
Financial Innovation and the Evolution of the Banking Industry 232
Responses to Changes in Demand Conditions: Interest Rate Volatility 233
Responses to Changes in Supply Conditions: Information Technology 234
Box 1 E-Finance: Will "Clicks" Dominate "Bricks" in the Banking Industry? 236
Avoidance of Existing Regulations 237
Financial Innovation and the Decline of Traditional Banking 239

Structure of the U.S. Commercial Banking Industry 243
Restrictions on Branching 244
Response to Branching Restrictions 245
Bank Consolidation and Nationwide Banking 245
Box 2 E-Finance: Information Technology and Bank Consolidation 247
The Riegle-Neal Interstate Banking and Branching Efficiency Act of 1994 248
What Will the Structure of the U.S. Banking Industry Look Like in the Future? 248
Box 3 Global: Comparison of Banking Structure in the United States and Abroad 249
Are Bank Consolidation and Nationwide Banking Good Things? 249
Separation of the Banking and Other Financial Service Industries 250
Erosion of Glass-Steagall 250
The Gramm-Leach-Bliley Financial Services Modernization Act of 1999: Repeal of Glass-Steagall 251
Implications for Financial Consolidation 251
Separation of Banking and Other Financial Services Industries Throughout the World . . . 251
Thrift Industry: Regulation and Structure 252
Savings and Loan Associations 252
Mutual Savings Banks 253
Credit Unions 253
International Banking 253
Eurodollar Market 254
Box 4 Global: Ironic Birth of the Eurodollar Market 255
Structure of U.S. Banking Overseas 255
Foreign Banks in the United States 256
Summary, Key Terms, Questions and Problems, and Web Exercises 257

CHAPTER 11
ECONOMIC ANALYSIS OF BANKING REGULATION 260

Preview 260
Asymmetric Information and Banking Regulation 260
Government Safety Net: Deposit Insurance and the FDIC 260
Box 1 Global: The Spread of Government Deposit Insurance Throughout the World: Is This a Good Thing? 262
Restrictions on Asset Holdings and Bank Capital Requirements 264
Bank Supervision: Chartering and Examination 265
Box 2 Global: Basel 2: Is It Spinning Out of Control? 265
Assessment of Risk Management 267
Disclosure Requirements 268
Consumer Protection 269
Restrictions on Competition 269
Box 3 E-Finance: Electronic Banking: New Challenges for Bank Regulation 270

International Banking Regulation 272
Problems in Regulating International Banking 272
Summary 272
The 1980s U.S. Banking Crisis: Why? 273
Early Stages of the Crisis 274
Later Stages of the Crisis: Regulatory Forbearance 275
Competitive Equality in Banking Act of 1987 276
Political Economy of the Savings and Loan Crisis 276
The Principal–Agent Problem for Regulators and Politicians 277
Savings and Loan Bailout: The Financial Institutions Reform, Recovery, and Enforcement Act of 1989 278
Federal Deposit Insurance Corporation Improvement Act of 1991 279
Banking Crises Throughout the World 280
Scandinavia 280
Latin America 281
Russia and Eastern Europe 282
Japan 282
East Asia 284
"Déjà Vu All Over Again" 284
Summary, Key Terms, Questions and Problems, and Web Exercises 284

CHAPTER 12
NONBANK FINANCE 287

Preview 287
Insurance 287
Life Insurance 287
Property and Casualty Insurance 288
The Competitive Threat from the Banking Industry 290
Application Insurance Management 290
Screening 291
Risk-Based Premiums 291
Restrictive Provisions 292
Prevention of Fraud 292
Cancellation of Insurance 292
Deductibles 292
Coinsurance 293
Limits on the Amount of Insurance 293
Summary 293
Pension Funds 294
Private Pension Plans 295
Public Pension Plans 295
Box 1 Should Social Security Be Privatized? 296
Finance Companies 296
Mutual Funds 297

Box 2 E-Finance: Mutual Funds and the Internet 298
Money Market Mutual Funds 299
Hedge Funds 299
Box 3 The Long-Term Capital Management Debacle 300
Government Financial Intermediation 301
Federal Credit Agencies 301
Box 4 Are Fannie Mae and Freddie Mac Getting Too Big for Their Britches? 302
Securities Market Operations 302
Investment Banking 303
Following the Financial News New Securities Issues 304
Securities Brokers and Dealers 304
Organized Exchanges 305
Box 5 The Return of the Financial Supermarket? 305
Box 6 E-Finance: The Internet Comes to Wall Street 306
Summary, Key Terms, Questions and Problems, and Web Exercises 306

CHAPTER 13
FINANCIAL DERIVATIVES 309
Preview 309
Hedging 309
Interest-Rate Forward Contracts 310
Application Hedging with Interest-Rate Forward Contracts 310
Pros and Cons of Forward Contracts 311
Financial Futures Contracts and Markets 311
Following the Financial News Financial Futures 312
Application Hedging with Financial Futures 314
Organization of Trading in Financial Futures Markets 315
The Globalization of Financial Futures Markets 317
Explaining the Success of Futures Markets 317
Application Hedging Foreign Exchange Risk 319
Hedging Foreign Exchange Risk with Forward Contracts 319
Hedging Foreign Exchange Risk with Futures Contracts 320
Options 320
Following the Financial News Futures Options 321
Option Contracts 322
Profits and Losses on Option and Futures Contracts 322
Application Hedging with Futures Options 325

Factors Affecting the Prices of Option Premiums 326
Summary 327
Interest-Rate Swaps 328
Interest-Rate Swap Contracts 328
Application Hedging with Interest-Rate Swaps 329
Advantages of Interest-Rate Swaps 329
Disadvantages of Interest-Rate Swaps 330
Financial Intermediaries in Interest-Rate Swaps 330
Summary, Key Terms, Questions and Problems, and Web Exercises 330

PART IV *Central Banking and the Conduct of Monetary Policy* 333

CHAPTER 14
STRUCTURE OF CENTRAL BANKS AND THE FEDERAL RESERVE SYSTEM 335
Preview 335
Origins of the Federal Reserve System 335
Box 1 Inside the Fed: The Political Genius of the Founders of the Federal Reserve System 336
Formal Structure of the Federal Reserve System 336
Federal Reserve Banks 337
Box 2 Inside the Fed: The Special Role of the Federal Reserve Bank of New York 339
Member Banks 340
Board of Governors of the Federal Reserve System 340
Federal Open Market Committee (FOMC) 341
Box 3 Inside the Fed: The Role of the Research Staff 342
The FOMC Meeting 343
Box 4 Inside the Fed: Green, Blue, and Beige: What Do These Colors Mean at the Fed? 344
Informal Structure of the Federal Reserve System 344
Box 5 Inside the Fed: The Role of Member Banks in the Federal Reserve System 346
How Independent Is the Fed? 346
Structure and Independence of Foreign Central Banks 349
Bank of Canada 349
Bank of England 349
Bank of Japan 350
European Central Bank 350
The Trend Toward Greater Independence 351
Explaining Central Bank Behavior 351

Box 6 Inside the Fed: Federal Reserve Transparency 352

Should the Fed Be Independent? 352
The Case for Independence 352
The Case Against Independence 354
Central Bank Independence and Macroeconomic Performance Throughout the World . . .354
Summary, Key Terms, Questions and Problems, and Web Exercises 355

CHAPTER 15
MULTIPLE DEPOSIT CREATION AND THE MONEY SUPPLY PROCESS 357

Preview 357
Four Players in the Money Supply Process 357
The Fed's Balance Sheet 358
Liabilities 358
Assets 359
Control of the Monetary Base 359
Federal Reserve Open Market Operations 359
Shifts from Deposits into Currency 363

Box 1 Global: Foreign Exchange Rate Intervention and the Monetary Base 363

Discount Loans 364
Other Factors That Affect the Monetary Base 365
Overview of the Fed's Ability to Control the Monetary Base 365
Multiple Deposit Creation: A Simple Model 365
Deposit Creation: The Single Bank 366
Deposit Creation: The Banking System 367
Deriving the Formula for Multiple Deposit Creation 370
Critique of the Simple Model 371
Summary, Key Terms, Questions and Problems, and Web Exercises 372

CHAPTER 16
DETERMINANTS OF THE MONEY SUPPLY 374

Preview 374
The Money Supply Model and the Money Multiplier 375
Deriving the Money Multiplier 375
Intuition Behind the Money Multiplier 377
Factors that Determine the Money Multiplier 378
Changes in the Required Reserve Ratio r 378
Changes in the Currency Ratio c 379
Changes in the Excess Reserves Ratio e 379
Additional Factors That Determine the Money Supply 381
Changes in the Nonborrowed Monetary Base MB_n 382
Changes in the Discount Loans DL from the Fed 382
Overview of the Money Supply Process 383

Application Explaining Movements in the Money Supply, 1980–2002 384

Application The Great Depression Bank Panics, 1930–1933 387

Summary, Key Terms, Questions and Problems, and Web Exercises 390

CHAPTER 17
TOOLS OF MONETARY POLICY 393

Preview 393

The Market for Reserves and the Federal Funds Rate 393
- Supply and Demand in the Market for Reserves 394
- How Changes in the Tools of Monetary Policy Affect the Federal Funds Rate 395

Open Market Operations 398
- A Day at the Trading Desk 398
- Advantages of Open Market Operations 400

Discount Policy 400
- Operation of the Discount Window 401
- Lender of Last Resort 402
- Advantages and Disadvantages of Discount Policy 403

Reserve Requirements 403

Box 1 Inside the Fed: Discounting to Prevent a Financial Panic: The Black Monday Stock Market Crash of 1987 and the Terrorist Destruction of the World Trade Center in September 2001 404

- Advantages and Disadvantages of Reserve Requirement Changes 405

Application Why Have Reserve Requirements Been Declining Worldwide? 406

Application The Channel/Corridor System for Setting Interest Rates in Other Countries 406

Summary, Key Terms, Questions and Problems, and Web Exercises 408

CHAPTER 18
CONDUCT OF MONETARY POLICY: GOALS AND TARGETS 411

Preview 411

Goals of Monetary Policy 411
- High Employment 411
- Economic Growth 412
- Price Stability 412

Box 1 Global: The Growing European Commitment to Price Stability 413

- Interest-Rate Stability 413
- Stability of Financial Markets 413
- Stability in Foreign Exchange Markets 414
- Conflict Among Goals 414

Central Bank Strategy: Use of Targets 414

Choosing the Targets 416
- Criteria for Choosing Intermediate Targets 418
- Criteria for Choosing Operating Targets 419

Fed Policy Procedures: Historical Perspective 419
The Early Years: Discount Policy as the Primary Tool 420
Discovery of Open Market Operations 420
The Great Depression 421

Box 2 Inside the Fed: Bank Panics of 1930–1933: Why Did the Fed Let Them Happen? 421

Reserve Requirements as a Policy Tool 422
War Finance and the Pegging of Interest Rates: 1942–1951 422
Targeting Money Market Conditions: The 1950s and 1960s 423
Targeting Monetary Aggregates: The 1970s 424
New Fed Operating Procedures: October 1979–October 1982 425
De-emphasis of Monetary Aggregates: October 1982–Early 1990s 426
Federal Funds Targeting Again: Early 1990s and Beyond 427
International Considerations 427

Box 3 Global: International Policy Coordination: The Plaza Agreement and the Louvre Accord 428

The Taylor Rule, NAIRU, and the Philips Curve 428

Box 4 Fed Watching 430

Summary, Key Terms, Questions and Problems, and Web Exercises 431

PART V *International Finance and Monetary Policy* 433

CHAPTER 19
THE FOREIGN EXCHANGE MARKET 435

Preview 435
Foreign Exchange Market 435
What Are Foreign Exchange Rates? 436

Following the Financial News Foreign Exchange Rates 437

Why Are Exchange Rates Important? 438
How Is Foreign Exchange Traded? 438
Exchange Rates in the Long Run 439
Law of One Price 439
Theory of Purchasing Power Parity 439
Why the Theory of Purchasing Power Parity Cannot Fully Explain Exchange Rates 440
Factors That Affect Rates in the Long Run 441
Exchange Rates in the Short Run 443
Comparing Expected Returns on Domestic and Foreign Deposits 443
Interest Parity Condition 445
Equilibrium in the Foreign Exchange Market 446
Explaining Changes in Exchange Rates 448
Shifts in the Expected-Return Schedule for Foreign Deposits 448
Shifts in the Expected-Return Schedule for Domestic Deposits 450

Application Changes in the Equilibrium Exchange Rate: Two Examples 452
Changes in Interest Rates 452
Changes in the Money Supply 453
Exchange Rate Overshooting 453
Application Why Are Exchange Rates So Volatile? 455
Application The Dollar and Interest Rates, 1973–2002 455
Application The Euro's First Four Years 457
Application Reading the *Wall Street Journal*: The "Currency Trading" Column . 457
Following the Financial News The "Currency Trading" Column 458
Summary, Key Terms, Questions and Problems, and Web Exercises 459

CHAPTER 20
THE INTERNATIONAL FINANCIAL SYSTEM 462
Preview 462
Intervention in the Foreign Exchange Market 462
Foreign Exchange Intervention and the Money Supply 462
Box 1 Inside the Fed: A Day at the Federal Reserve Bank of New York's Foreign Exchange Desk 463
Unsterilized Intervention 465
Sterilized Intervention 466
Balance of Payments 467
Evolution of the International Financial System 468
Gold Standard 469
Bretton Woods System 469
Box 2 Global: The Euro's Challenge to the Dollar 471
Managed Float 473
European Monetary System (EMS) 474
Application The Foreign Exchange Crisis of September 1992 475
Application Recent Foreign Exchange Crises in Emerging Market Countries: Mexico 1994, East Asia 1997, Brazil 1999, and Argentina 2002 477
Capital Controls 478
Controls on Capital Outflows 478
Controls on Capital Inflows 479
The Role of the IMF 479
Should the IMF Be an International Lender of Last Resort? 480
International Considerations and Monetary Policy 482
Direct Effects of the Foreign Exchange Market on the Money Supply 482
Balance-of-Payments Considerations 483
Exchange Rate Considerations 483
Summary, Key Terms, Questions and Problems, and Web Exercises 484

CHAPTER 21
MONETARY POLICY STRATEGY: THE INTERNATIONAL EXPERIENCE 487

Preview 487
The Role of a Nominal Anchor 487
The Time-Consistency Problem 488
Exchange-Rate Targeting 489
Advantages of Exchange-Rate Targeting 489
Disadvantages of Exchange-Rate Targeting 490
When Is Exchange-Rate Targeting Desirable for Industrialized Countries? 492
When Is Exchange-Rate Targeting Desirable for Emerging Market Countries? 492
Currency Boards 492
Dollarization 493
Box 1 Global: Argentina's Currency Board 494
Monetary Targeting 496
Monetary Targeting in Canada, the United Kingdom, Japan, Germany, and Switzerland . . . 496
Box 2 Global: The European Central Bank's Monetary Policy Strategy 498
Advantages of Monetary Targeting 500
Disadvantages of Monetary Targeting 501
Inflation Targeting 501
Inflation Targeting in New Zealand, Canada, and the United Kingdom 501
Advantages of Inflation Targeting 504
Disadvantages of Inflation Targeting 506
Nominal GDP Targeting 508
Monetary Policy with an Implicit Nominal Anchor 509
Advantages of the Fed's Approach 510
Disadvantages of the Fed's Approach 510
Summary, Key Terms, Questions and Problems, and Web Exercises 512

PART VI *Monetary Theory* 515

CHAPTER 22
THE DEMAND FOR MONEY 517

Preview 517
Quantity Theory of Money 517
Velocity of Money and Equation of Exchange 518
Quantity Theory 519
Quantity Theory of Money Demand 519
Is Velocity a Constant? 520
Keynes's Liquidity Preference Theory 521
Transactions Motive 521
Precautionary Motive 522
Speculative Motive 522
Putting the Three Motives Together 523
Further Developments in the Keynesian Approach 524
Transactions Demand 524

Precautionary Demand 527
Speculative Demand 527
Friedman's Modern Quantity Theory of Money 528
Distinguishing Between the Friedman and Keynesian Theories 530
Empirical Evidence on the Demand for Money 532
Interest Rates and Money Demand 533
Stability of Money Demand 533
Summary, Key Terms, Questions and Problems, and Web Exercises 533

CHAPTER 23
THE KEYNESIAN FRAMEWORK AND THE ISLM MODEL 536
Preview 536
Determination of Aggregate Output 536
Consumer Expenditure and the Consumption Function 538
Investment Spending 539
Box 1 Meaning of the Word *Investment* 540
Equilibrium and the Keynesian Cross Diagram 540
Expenditure Multiplier 542
Application The Collapse of Investment Spending and the Great Depression . . 545
Government's Role 545
Role of International Trade 548
Summary of the Determinants of Aggregate Output 548
The *ISLM* Model 551
Equilibrium in the Goods Market: The *IS* Curve 552
Equilibrium in the Market for Money: The *LM* Curve 555
ISLM Approach to Aggregate Output and Interest Rates 557
Summary, Key Terms, Questions and Problems, and Web Exercises 558

CHAPTER 24
MONETARY AND FISCAL POLICY IN THE ISLM MODEL 561
Preview 561
Factors That Cause the *IS* Curve to Shift 561
Factors That Cause the *LM* Curve to Shift 564
Changes in Equilibrium Level of the Interest Rate and Aggregate Output 566
Response to a Change in Monetary Policy 566
Response to a Change in Fiscal Policy 567
Effectiveness of Monetary Versus Fiscal Policy 568
Monetary Policy Versus Fiscal Policy: The Case of Complete Crowding Out 568
Application Targeting Money Supply Versus Interest Rates 571
ISLM Model in the Long Run 575
ISLM Model and the Aggregate Demand Curve 577
Deriving the Aggregate Demand Curve 577
Factors That Cause the Aggregate Demand Curve to Shift 578
Summary, Key Terms, Questions and Problems, and Web Exercises 580

CHAPTER 25
AGGREGATE DEMAND AND SUPPLY ANALYSIS 582
Preview 582
Aggregate Demand 582
Monetarist View of Aggregate Demand 582
Following the Financial News Aggregate Output, Unemployment, and the Price Level 583
Keynesian View of Aggregate Demand 585
The Crowding-Out Debate 586
Aggregate Supply 587
Shifts in the Aggregate Supply Curve 588
Equilibrium in Aggregate Supply and Demand Analysis 588
Equilibrium in the Short Run 589
Equilibrium in the Long Run 589
Shifts in Aggregate Demand 592
Shifts in Aggregate Supply 594
Shifts in the Long-Run Aggregate Supply Curve: Real Business Cycle Theory and Hysteresis 596
Conclusions 597
Application Explaining Past Business Cycle Episodes 598
Vietnam War Buildup, 1964–1970 598
Negative Supply Shocks, 1973–1975 and 1978–1980 598
Favorable Supply Shocks, 1995–1999 599
Summary, Key Terms, Questions and Problems, and Web Exercises 600

CHAPTER 26
TRANSMISSION MECHANISMS OF MONETARY POLICY: THE EVIDENCE 603
Preview 603
Framework for Evaluating Empirical Evidence 603
Structural Model Evidence 604
Reduced-Form Evidence 604
Advantages and Disadvantages of Structural Model Evidence 605
Advantages and Disadvantages of Reduced-Form Evidence 606
Box 1 Perils of Reverse Causation: A Russian Folk Tale 606
Box 2 Perils of Ignoring an Outside Driving Factor: How to Lose a Presidential Election 607
Conclusions 607
Early Keynesian Evidence on the Importance of Money 607
Objections to Early Keynesian Evidence 608
Early Monetarist Evidence on the Importance of Money 611
Timing Evidence 611
Statistical Evidence 613

Historical Evidence 615
Overview of the Monetarist Evidence 615
Box 3 Real Business Cycle Theory and the Debate on Money and Economic Activity 616
Transmission Mechanisms of Monetary Policy 616
Traditional Interest-Rate Channels 617
Other Asset Price Channels 618
Credit View 621
Box 4 Consumers' Balance Sheets and the Great Depression 624
Why Are Credit Channels Likely to Be Important? 625
Application Corporate Scandals and the Slow Recovery from the March 2001 Recession 625
Lessons for Monetary Policy 626
Application Applying the Monetary Policy Lessons to Japan 628
Summary, Key Terms, Questions and Problems, and Web Exercises 629

CHAPTER 27
MONEY AND INFLATION 632
Preview 632
Money and Inflation: Evidence 632
German Hyperinflation, 1921–1923 633
Recent Episodes of Rapid Inflation 633
Meaning of Inflation 634
Views of Inflation 635
Monetarist View 635
Keynesian View 636
Summary 638
Origins of Inflationary Monetary Policy 638
High Employment Targets and Inflation 639
Budget Deficits and Inflation 643
Application Explaining the Rise in U.S. Inflation, 1960–1980 646
Activist/Nonactivist Policy Debate 650
Responses to High Unemployment 650
Activist and Nonactivist Positions 651
Expectations and the Activist/Nonactivist Debate 652
Box 1 Perils of Accommodating Policy: The Terrorism Dilemma 654
Rules Versus Discretion: Conclusions 654
Application Importance of Credibility to Volcker's Victory over Inflation 655
Summary, Key Terms, Questions and Problems, and Web Exercises 655

CHAPTER 28
RATIONAL EXPECTATIONS: IMPLICATIONS FOR POLICY 658
Preview 658
The Lucas Critique of Policy Evaluation 659
Econometric Policy Evaluation 659
Example: The Term Structure of Interest Rates 659
New Classical Macroeconomic Model 660
Effects of Unanticipated and Anticipated Policy 661
Box 1 Proof of the Policy Ineffectiveness Proposition 663
Can an Expansionary Policy Lead to a Decline in Aggregate Output? 663
Implications for Policymakers 664
New Keynesian Model 665
Effects of Unanticipated and Anticipated Policy 666
Implications for Policymakers 666
Comparison of the Two New Models with the Traditional Model 666
Short-Run Output and Price Responses 668
Stabilization Policy 670
Anti-inflation Policies 671
Credibility in Fighting Inflation 673
Box 2 Global: Ending the Bolivian Hyperinflation: Case Study of a Successful Anti-inflation Program 674
Application Credibility and the Reagan Budget Deficits 675
Impact of the Rational Expectations Revolution 676
Summary, Key Terms, Questions and Problems, and Web Exercises 677

GLOSSARY G-1
ANSWERS TO SELECTED QUESTIONS AND PROBLEMS A-1
CREDITS C-1
INDEX I-1

PREFACE

I have continually strived to improve this textbook with each new edition, and the Seventh Edition of *The Economics of Money, Banking, and Financial Markets* is no exception. The text has undergone a major revision, but it retains the basic hallmarks that have made it the best-selling textbook on money and banking in the past six editions:

- A unifying, analytic framework that uses a few basic economic principles to organize students' thinking about the structure of financial markets, the foreign exchange markets, financial institution management, and the role of monetary policy in the economy
- A careful, step-by-step development of models (an approach found in the best principles of economics textbooks), which makes it easier for students to learn
- The complete integration of an international perspective throughout the text
- A thoroughly up-to-date treatment of the latest developments in monetary theory
- Special features called "Following the Financial News" and "Reading the *Wall Street Journal*" to encourage reading of a financial newspaper
- An applications-oriented perspective with numerous applications and special-topic boxes that increase students' interest by showing them how to apply theory to real-world examples

What's New in the Seventh Edition

In addition to the expected updating of all data through the end of 2002 whenever possible, there is major new material in every part of the text. Indeed, this revision is one of the most substantial that I have ever done.

Expanded Coverage of the Stock Market

With the wide swings in the stock prices in recent years, students of money and banking have become increasingly interested in what drives the stock market. As a result, I have expanded the discussion of this market by describing simple valuation methods for stocks and examining recent developments in the stock market and the link between monetary policy and stock prices. I have combined this material with the discussion of the theory of rational expectations and efficient capital markets to create a new Chapter 7, "The Stock Market, the Theory of Rational Expectations, and the Efficient Market Hypothesis."

New Material on Financial Institutions

In light of continuing changes in financial markets and institutions, I have added the following new material to keep the text current:

- Extensive discussion of recent corporate scandals and the collapse of Enron, including their impact on the economy (Chapters 6, 7, 11, and 26)
- Discussion of the role of venture capitalists in the high-tech sector (Chapter 8)
- Examination of how information technology is influencing bank consolidation, and analysis of whether clicks will dominate bricks in the banking industry (Chapter 10)
- New material on the Basel Committee on Bank Supervision and where the Basel Accord is heading (Chapter 11)
- Discussion of the spread of deposit insurance throughout the world (Chapter 11)
- Perspective on the growing concerns about Fannie Mae and Freddie Mac (Chapter 12)
- A new type of special-interest box, the E-Finance box, which relates how changes in technology have affected the conduct of business in banking and financial markets. The placement of these boxes throughout the text helps to demonstrate the impact of technology across a broad range of areas in finance.

Increased International Perspective

The growing importance of the global economy has encouraged me to add more new material with an international perspective:

- Extensive discussion of recent developments in Argentina (Chapters 1, 8, 11, 20, and 21)
- Analysis of how central banks set overnight interest rates in other countries (Chapter 17)
- Discussion of how the euro has fared in its first four years (Chapter 19)
- Additional treatment of recent events in the Japanese economy (Chapters 11 and 26)

New Material on Monetary Theory and Policy

Drawing on my continuing involvement with central banks around the world, I have added new material to keep the discussion of monetary theory and policy current:

- New boxes on Fed watching and Federal Reserve transparency (Chapters 14 and 18)
- Discussion of the changes (implemented in 2003) in the way the Fed administers the discount window (Chapter 17)
- An updated discussion of the market for reserves and how the channel/corridor system for setting interest rates works (Chapter 17)
- Discussion of how the recent corporate scandals have hindered the recovery of the economy from the 2001–2002 recession (Chapter 25)

E-Focus

The incredible advances in electronic (computer and telecommunications) technology in recent years have had a major impact on the financial system. This Seventh Edition reflects these developments by adding many new features with an electronic focus.

Web Enhancement. The Seventh Edition embraces the exploding world of information now available over the World Wide Web. There are few areas where the Internet

has been as valuable as in the realm of money, banking, and financial markets. Data that were once difficult and tedious to collect are now readily available. To help students appreciate what they can access online, I have added a number of new features:

1. **Web Exercises.** This edition adds all-new end-of-chapter Web Exercises. These require that students collect information from online sources or use online resources to enhance their learning experience. The Web Exercises are relatively quick and easy to complete, while still accomplishing the goal of familiarizing students with online sources of data.
2. **Web Sources.** Much of the data used to create the many tables and charts were collected from online sources. Wherever a Web URL is available, it is exactly reported as the source. The interested student or instructor can use this URL to see what has happened since the chart or table was created.
3. **Marginal Web References.** In addition to listing the sources of data used to create the charts and graphs, I have also included in the margin URLs to Web sites that provide information or data that supplement the text. These references include a brief description of what students will find at the site. Interested students can use these sites to extend their study, and instructors can draw from them to supplement their lecture notes. Because the URLs for Web sources and references do sometimes change, the Mishkin Companion Web Site at www.aw.com/mishkin will provide the new URLs when they are needed.

E-Finance Boxes. To illustrate how electronic technology has increasingly permeated financial markets and institutions, I have included the all-new E-Finance boxes, described earlier, to show the ongoing real-world impact of this remarkable development.

Streamlined Coverage and Organization

As textbooks go into later editions, they often grow in length. Over the years, I have resisted this tendency, and in this edition have made even greater efforts to streamline the book. Despite the addition of a lot of new material, the book is substantially shorter. Moreover, at the suggestion of reviewers, I have moved the discussion of rational expectations and efficient markets earlier in the book, to Chapter 7. I have also shifted the material on the foreign exchange market and the determination of exchange rates to Chapter 19 so that it comes immediately before the chapter on the international financial system, allowing this material to be taught together.

Appendices on the Web

The Web site for this book, www.aw.com/mishkin, has allowed me to produce a large amount of new material for the book without lengthening the text, because we have placed this material in appendices on the Web site. The appendices include:

Chapter 2: Financial Market Instruments
Chapter 4: Measuring Interest-Rate Risk: Duration
Chapter 5: Models of Asset Pricing
Chapter 5: Applying the Asset Market Approach to a Commodity Market: The Case of Gold
Chapter 9: Duration Gap Analysis
Chapter 9: Measuring Bank Performance
Chapter 11: Evaluating FDICIA and Other Proposed Reforms of the Bank Regulatory System

Chapter 15: The Fed's Balance Sheet and the Monetary Base
Chapter 16: The M2 Money Multiplier
Chapter 16: Explaining the Behavior of the Currency Ratio
Chapter 22: A Mathematical Treatment of the Baumol-Tobin and Tobin Mean Variance Model
Chapter 22: Empirical Evidence on the Demand for Money
Chapter 24: Algebra of the *ISLM* Model
Chapter 25: Aggregate Supply and the Phillips Curve

Instructors can either use these appendices in class to supplement the material in the textbook, or recommend them to students who want to expand their knowledge of the money and banking field.

Flexibility

In using previous editions, adopters, reviewers, and survey respondents have continually praised this text's flexibility. There are as many ways to teach money, banking, and financial markets as there are instructors. To satisfy the diverse needs of instructors, the text achieves flexibility as follows:

- Core chapters provide the basic analysis used throughout the book, and other chapters or sections of chapters can be used or omitted according to instructor preferences. For example, Chapter 2 introduces the financial system and basic concepts such as transaction costs, adverse selection, and moral hazard. After covering Chapter 2, the instructor may decide to give more detailed coverage of financial structure by assigning Chapter 8, or may choose to skip Chapter 8 and take any of a number of different paths through the book.
- The text also allows instructors to cover the most important issues in monetary theory and policy without having to use the *ISLM* model in Chapters 23 and 24, while more complete treatments of monetary theory make use of the *ISLM* chapters.
- The internationalization of the text through marked international sections within chapters, as well as through complete separate chapters on the foreign exchange market and the international monetary system, is comprehensive yet flexible. Although many instructors will teach all the international material, others will not. Instructors who want less emphasis on international topics can easily skip Chapter 19 on the foreign exchange market and Chapter 20 on the international financial system and monetary policy. The international sections within chapters are self-contained and can be omitted with little loss of continuity.

To illustrate how this book can be used for courses with varying emphases, several course outlines are suggested for a semester teaching schedule. More detailed information about how the text can be used flexibly in your course is available in the Instructor's Manual.

- *General Money and Banking Course:* Chapters 1–5, 9–11, 14, 17, 18, 25, 27, with a choice of 6 of the remaining 15 chapters.
- *General Money and Banking Course with an International Emphasis:* Chapters 1–5, 9–11, 14, 17–20, 25, 27 with a choice of 4 of the remaining 13 chapters.

- *Financial Markets and Institutions Course:* Chapters 1–13, with a choice of 6 of the remaining 15 chapters.
- *Monetary Theory and Policy Course:* Chapters 1–5, 14, 15, 17, 18, 21, 25–28, with a choice of 5 of the remaining 14 chapters.

Pedagogical Aids

In teaching theory or its applications, a textbook must be a solid motivational tool. To this end, I have incorporated a wide variety of pedagogical features to make the material easy to learn:

1. **Previews** at the beginning of each chapter tell students where the chapter is heading, why specific topics are important, and how they relate to other topics in the book.
2. **Applications**, numbering more than 50, demonstrate how the analysis in the book can be used to explain many important real-world situations. A special set of applications, called "Reading the *Wall Street Journal*," shows students how to read daily columns in this leading financial newspaper.
3. **"Following the Financial News" boxes** introduce students to relevant news articles and data that are reported daily in the press, and explain how to read them.
4. **"Inside the Fed" boxes** give students a feel for what is important in the operation and structure of the Federal Reserve System.
5. **Global boxes** include interesting material with an international focus.
6. **E-Finance boxes** relate how changes in technology have affected financial markets or institutions.
7. **Special-interest boxes** highlight dramatic historical episodes, interesting ideas, and intriguing facts related to the subject matter.
8. **Study Guides** are highlighted statements scattered throughout the text that provide hints to the student on how to think about or approach a topic.
9. **Summary tables** provide a useful study aid in reviewing material.
10. **Key statements** are important points set in boldface italic type so that students can easily find them for later reference.
11. **Graphs** with captions, numbering more than 150, help students clearly understand the interrelationship of the variables plotted and the principles of analysis.
12. **Summary** at the end of each chapter lists the main points covered.
13. **Key terms** are important words or phrases, boldfaced when they are defined for the first time and listed by page number at the end of the chapter.
14. **End-of-chapter questions and problems**, numbering more than 400, help students learn the subject matter by applying economic concepts, including a special class of problems that students find particularly relevant, under the heading "Using Economic Analysis to Predict the Future."
15. **Web Exercises** encourage students to collect information from online sources or use online resources to enhance their learning experience.
16. **Web sources** report the Web URL source of the data used to create the many tables and charts.

17. **Marginal Web references** point the student to Web sites that provide information or data that supplement the text material.
18. **Glossary** at the back of the book provides definitions of all the key terms.
19. **Answers section** at the back of the book provides solutions to half of the questions and problems (marked by *).

An Easier Way to Teach Money, Banking, and Financial Markets

The demands for good teaching have increased dramatically in recent years. To meet these demands, I have provided the instructor with supplementary materials, unlike those available with any competing text, that should make teaching this course substantially easier.

This book comes with not only full-color Microsoft PowerPoint electronic transparencies of all the figures and tables but also full-color overhead transparencies. Furthermore, the Instructor's Manual contains transparency masters of the lecture notes, perforated so that they can be easily detached for use in class.

The lecture notes are comprehensive and outline all the major points covered in the text. They have been class-tested successfully—they are in fact the notes that I use in class—and they should help other instructors prepare their lectures as they have helped me. Some instructors might use these lecture notes as their own class notes and prefer to teach with a blackboard. But for those who prefer to teach with visual aids, the PowerPoint presentation and the full-color transparencies of the figures and tables afford the flexibility to take this approach.

I am also aware that many instructors want to make variations in their lectures that depart somewhat from material covered in the text. For their convenience, the entire set of lecture notes has been put on the Instructor's Resource CD-ROM using Microsoft Word. Instructors can modify the lecture notes as they see fit for their own use, for class handouts, or for transparencies to be used with an overhead projector.

The Instructor's Resource CD-ROM also offers the entire contents of the Instructor's Manual, which includes chapter outlines, overviews, and teaching tips; answers to the end-of-chapter problems that are not included in the text. Using this handy feature, instructors can prepare student handouts such as solutions to problem sets made up of end-of-chapter problems, the outline of the lecture that day, or essay discussion questions for homework. I have used handouts of this type in my teaching and have found them to be very effective. Instructors have my permission and are encouraged to photocopy all of the materials on the CD-ROM and use them as they see fit in class.

Supplements Program to Accompany the Seventh Edition

The Economics of Money, Banking, and Financial Markets, Seventh Edition, includes the most comprehensive program of supplements of any money, banking, and financial markets textbook. These items are available to qualified domestic adopters, but in some cases may not be available to international adopters.

For the Professor

1. **Instructor's Resource Manual**, a print supplement prepared by me and offering conventional elements such as sample course outlines, chapter outlines, and

answers to questions and problems in the text. In addition, the manual contains my **Lecture Notes**, numbering more than 300, in transparency master format; these notes comprehensively outline the major points covered in the textbook.
2. **Instructor's Resource CD-ROM**, which conveniently holds the MS Word files to the Instructor's Manual, the Computerized Test Bank, and the MS PowerPoint Lecture Presentation.
3. **Full-Color Transparencies**, numbering more than 150, for *all* of the figures, tables, and summary tables.
4. **PowerPoint Electronic Lecture Presentation**, numbering more than 300 images, which include all the book's figures and tables in full color, plus the lecture notes. Available on the Instructor's Resource CD-ROM.
5. **Printed Test Bank** by James Butkiewicz of the University of Delaware, comprising more than 4,500 multiple-choice and essay test items, many with graphs.
6. **Computerized Test Bank**, allowing the instructor to produce exams efficiently. This product consists of the multiple-choice and essay questions in the printed Test Bank and offers editing capabilities. It is available in Macintosh and Windows versions on the Instructor's Resource CD-ROM.

For the Student

1. **Study Guide and Workbook**, prepared by Erick Eschker of Humboldt State University, John McArthur of Wofford College, and me, which includes chapter synopses and completions, exercises, self-tests, and answers to the exercises and self-tests.
2. **Readings in Money, Banking, and Financial Markets**, edited by James W. Eaton of Bridgewater College and me, updated annually, with over half the articles new each year to enable instructors to keep the content of their course current throughout the life of an edition of the text. The readings are available within MyEconLab (see next section).

Course Management with MyEconLab

Every student who buys a new textbook receives a prepaid subscription to MyEconLab. New to the Seventh Edition of *The Economics of Money, Banking, and Financial Markets*, MyEconLab delivers rich online content and innovative learning tools to your classroom. Instructors who use MyEconLab gain access to powerful communication and assessment tools, and their students receive access to the additional learning resources described next.

Students and MyEconLab

MyEconLab delivers the content and tools your students need to succeed within Addison-Wesley's innovative CourseCompass system. Students whose instructors use MyEconLab gain access to a variety of resources:

- The complete textbook online, in PDF format, with animated graphs that help students master the key concepts
- MathXL for Economics—a powerful tutorial to refresh students on the basics of creating and interpreting graphs; solving applied problems using graphs; calculating ratios and percentages; performing calculations; calculating average, median, and mode; and finding areas
- Research Navigator™—a one-stop research tool, with extensive help on the entire research process, including evaluating sources, drafting, and documentation, as

well as access to a variety of scholarly journals and publications, a complete year of search for full-text articles from the *New York Times*, and a "Best of the Web" Link Library of peer-reviewed Web sites
- eThemes of the Times—thematically related articles from the *New York Times*, accompanied by critical-thinking questions
- *Readings on Money, Banking, and Financial Markets*—edited by James W. Eaton of Bridgewater College and me and updated annually, with a focus on articles from Federal Reserve publications and economics and finance journals
- Additional study resources such as self-testing quizzes for each chapter, a weekly current events feature, online glossary term flashcards, and additional articles and supplemental materials

The Student Access Kit that arrives bundled with all new books walks students step-by-step through the registration process.

Instructors and MyEconLab

With MyEconLab, instructors can customize existing content and add their own. They can manage, create, and assign tests to students, choosing from our Test Bank, or upload tests they've written themselves. MyEconLab also includes advanced tracking features that record students' usage and performance and a Gradebook feature to see students' test results. Please refer to the *Instructor Quick Start Guide* or contact your Addison-Wesley sales representative to set up MyEconLab for your course.

Acknowledgments

As always in so large a project, there are many people to thank. My gratitude goes to Victoria Warneck, economics editor at Addison Wesley; Sylvia Mallory, Executive Development Manager; and Jane Tufts, the best development editor in the business. I also have been assisted by comments from my colleagues at Columbia and from my students.

In addition, I have been guided by the thoughtful commentary of outside reviewers and correspondents, especially Jim Eaton. Their feedback has made this a better book. In particular, I thank the following:

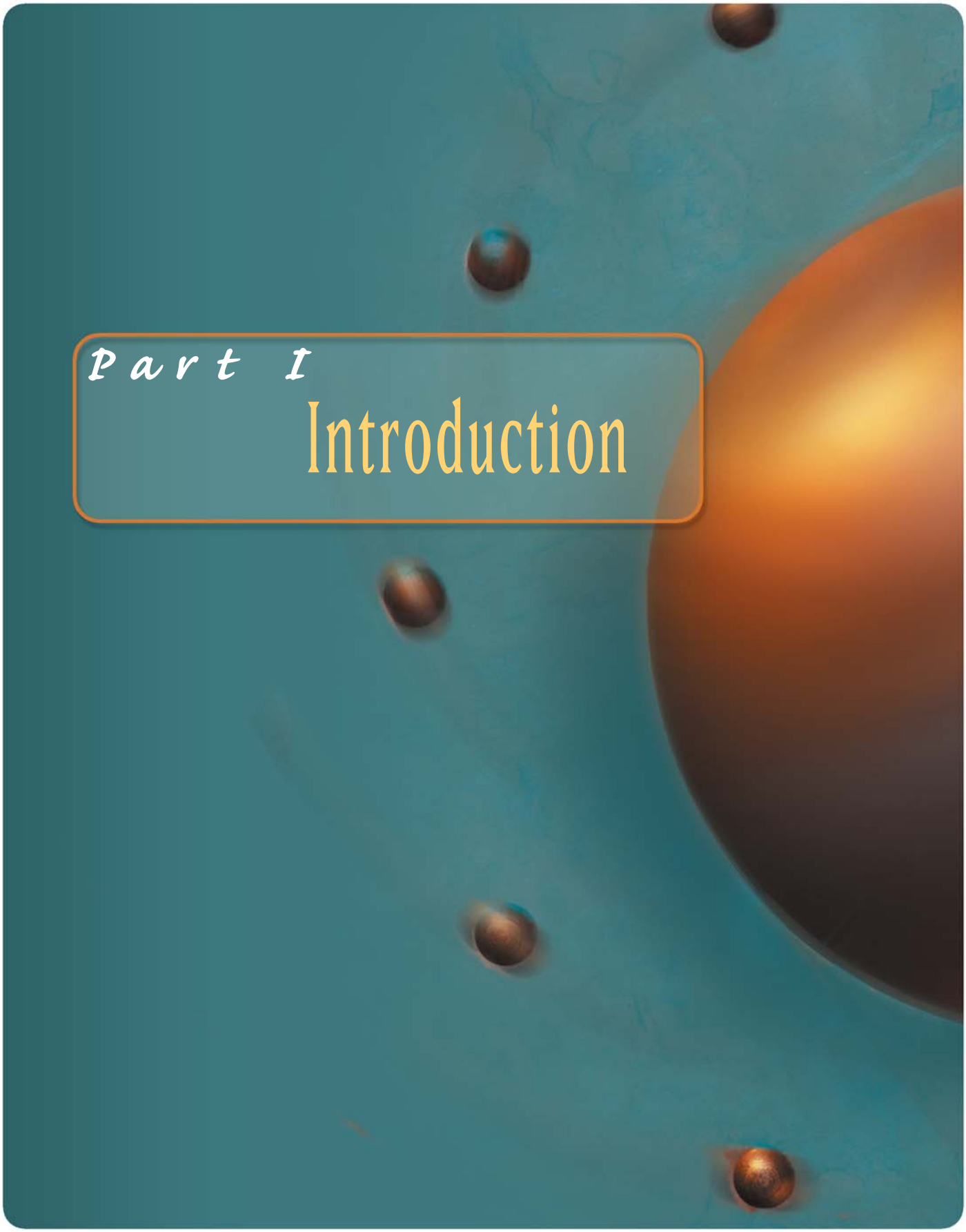
Burton Abrams, University of Delaware
Francis W. Ahking, University of Connecticut
Mohammed Akacem, Metropolitan State College of Denver
Harjit K. Arora, Le Moyne College
Stacie Beck, University of Delaware
Gerry Bialka, University of North Florida
Daniel K. Biederman, University of North Dakota
John Bishop, East Carolina University
Daniel Blake, California State University, Northridge
Robert Boatler, Texas Christian University
Henning Bohn, University of California, Santa Barbara
Michael W. Brandl, University of Texas at Austin
Oscar T. Brookins, Northeastern University
William Walter Brown, California State University, Northridge

James L. Butkiewicz, University of Delaware
Colleen M. Callahan, Lehigh University
Ray Canterbery, Florida State University
Sergio Castello, University of Mobile
Jen-Chi Cheng, Wichita State University
Patrick Crowley, Middlebury College
Sarah E. Culver, University of Alabama, Birmingham
Maria Davis, San Antonio College
Ranjit S. Dighe, State University of New York, Oswego
Richard Douglas, Bowling Green University
Donald H. Dutkowsky, Syracuse University
Richard Eichhorn, Colorado State University
Paul Emberton, Southwest Texas State University
Erick Eschker, Humboldt State University
Robert Eyler, Sonoma State University
L. S. Fan, Colorado State University
Sasan Fayazmanesh, California State University, Fresno
Dennis Fixler, George Washington University
Gary Fleming, Roanoke College
Grant D. Forsyth, Eastern Washington University
James Gale, Michigan Technological University
Stuart M. Glosser, University of Wisconsin, Whitewater
Fred C. Graham, American University
Jo Anna Gray, University of Oregon
David Gulley, Bentley College
Daniel Haak, Stanford University
Larbi Hammami, McGill University
Bassan Harik, Western Michigan University
J. C. Hartline, Rutgers University
Scott E. Hein, Texas Tech University
Robert Stanley Herren, North Dakota State University
Jane Himarios, University of Texas, Arlington
Dar-Yeh Hwang, National Taiwan University
Jayvanth Ishwaran, Stephen F. Austin State University
Jonatan Jelen, Queens College and City College of CUNY
U Jin Jhun, State University of New York, Oswego
Frederick L. Joutz, George Washington University
Bryce Kanago, University of Northern Iowa
Magda Kandil, International Monetary Fund
George G. Kaufman, Loyola University Chicago
Richard H. Keehn, University of Wisconsin, Parkside
Elizabeth Sawyer Kelly, University of Wisconsin, Madison
Jim Lee, Fort Hays State University
Robert Leeson, University of Western Ontario
Tony Lima, California State University, Hayward
Bernard Malamud, University of Nevada, Las Vegas
Marvin Margolis, Millersville University
Stephen McCafferty, Ohio State University
James McCown, Ohio State University

Cheryl McGaughey, Angelo State University
W. Douglas McMillin, Louisiana State University
William Merrill, Iowa State University
Carrie Meyer, George Mason University
Stephen M. Miller, University of Connecticut
Masoud Moghaddam, Saint Cloud State University
Thomas S. Mondschean, DePaul University
Clair Morris, U.S. Naval Academy
Jon Nadenichek, California State University, Northridge
John Nader, Grand Valley State University
Leonce Ndikumana, University of Massachusetts, Amherst
Ray Nelson, Brigham Young University
Inder P. Nijhawan, Fayetteville State University
Nick Noble, Miami University of Ohio
Dennis O'Toole, Virginia Commonwealth University
Mark J. Perry, University of Michigan, Flint
Chung Pham, University of New Mexico
Marvin M. Phaup, George Washington University
Ganga P. Ramdas, Lincoln University
Ronald A. Ratti, University of Missouri, Columbia
Hans Rau, Ball State University
Prosper Raynold, Miami University
Javier Reyes, Texas A&M University
Jack Russ, San Diego State University
Robert S. Rycroft, Mary Washington College
Lynn Schneider, Auburn University, Montgomery
Walter Schwarm, Colorado State University
Harinder Singh, Grand Valley State University
Larry Taylor, Lehigh University
Leigh Tesfatsion, Iowa State University
Frederick D. Thum, University of Texas, Austin
Robert Tokle, Idaho State University
C. Van Marrewijk, Erasmus University
Christopher J. Waller, Indiana University
Maurice Weinrobe, Clark University
James R. Wible, University of New Hampshire
Philip R. Wiest, George Mason University
William Wilkes, Athens State University
Thomas Williams, William Paterson University
Laura Wolff, Southern Illinois University, Edwardsville
Robert Wright, University of Virginia
Ben T. Yu, California State University, Northridge
Ky H. Yuhn, Florida Atlantic University
Jeffrey Zimmerman, Methodist College

Finally, I want to thank my wife, Sally; my son, Matthew; and my daughter, Laura, who provide me with a warm and happy environment that enables me to do my work, and my father, Sydney, now deceased, who a long time ago put me on the path that led to this book.

Frederic S. Mishkin


Part I
Introduction

Chapter 1

Why Study Money, Banking, and Financial Markets?

PREVIEW

On the evening news you have just heard that the Federal Reserve is raising the federal funds rate by $\frac{1}{2}$ of a percentage point. What effect might this have on the interest rate of an automobile loan when you finance your purchase of a sleek new sports car? Does it mean that a house will be more or less affordable in the future? Will it make it easier or harder for you to get a job next year?

This book provides answers to these and other questions by examining how financial markets (such as those for bonds, stocks, and foreign exchange) and financial institutions (banks, insurance companies, mutual funds, and other institutions) work and by exploring the role of money in the economy. Financial markets and institutions not only affect your everyday life but also involve huge flows of funds (trillions of dollars) throughout our economy, which in turn affect business profits, the production of goods and services, and even the economic well-being of countries other than the United States. What happens to financial markets, financial institutions, and money is of great concern to our politicians and can even have a major impact on our elections. The study of money, banking, and financial markets will reward you with an understanding of many exciting issues. In this chapter we provide a road map of the book by outlining these issues and exploring why they are worth studying.

Why Study Financial Markets?

A claim on the borrower's future income that is sold by the borrower to the lender. Also called a financial instrument.

Part II of this book focuses on **financial markets**, markets in which funds are transferred from people who have an excess of available funds to people who have a shortage. Financial markets such as bond and stock markets are crucial to promoting greater economic efficiency by channeling funds from people who do not have a productive use for them to those who do. Indeed, well-functioning financial markets are a key factor in producing high economic growth, and poorly performing financial markets are one reason that many countries in the world remain desperately poor. Activities in financial markets also have direct effects on personal wealth, the behavior of businesses and consumers, and the cyclical performance of the economy.

The Bond Market and Interest Rates

A **security** (also called a *financial instrument*) is a claim on the issuer's future income or **assets** (any financial claim or piece of property that is subject to ownership). A **bond** is a debt security that promises to make payments periodically for a specified

www.federalreserve.gov/releases/

Daily, weekly, monthly, quarterly, and annual releases and historical data for selected interest rates, foreign exchange rates, and so on.

period of time.[1] The bond market is especially important to economic activity because it enables corporations or governments to borrow to finance their activities and because it is where interest rates are determined. An **interest rate** is the cost of borrowing or the price paid for the rental of funds (usually expressed as a percentage of the rental of $100 per year). There are many interest rates in the economy—mortgage interest rates, car loan rates, and interest rates on many different types of bonds.

Interest rates are important on a number of levels. On a personal level, high interest rates could deter you from buying a house or a car because the cost of financing it would be high. Conversely, high interest rates could encourage you to save because you can earn more interest income by putting aside some of your earnings as savings. On a more general level, interest rates have an impact on the overall health of the economy because they affect not only consumers' willingness to spend or save but also businesses' investment decisions. High interest rates, for example, might cause a corporation to postpone building a new plant that would ensure more jobs.

Because changes in interest rates have important effects on individuals, financial institutions, businesses, and the overall economy, it is important to explain fluctuations in interest rates that have been substantial over the past twenty years. For example, the interest rate on three-month Treasury bills peaked at over 16% in 1981. This interest rate then fell to 3% in late 1992 and 1993, rose to above 5% in the mid to late 1990s, and then fell to a low of below 2% in the early 2000s.

Because different interest rates have a tendency to move in unison, economists frequently lump interest rates together and refer to "the" interest rate. As Figure 1

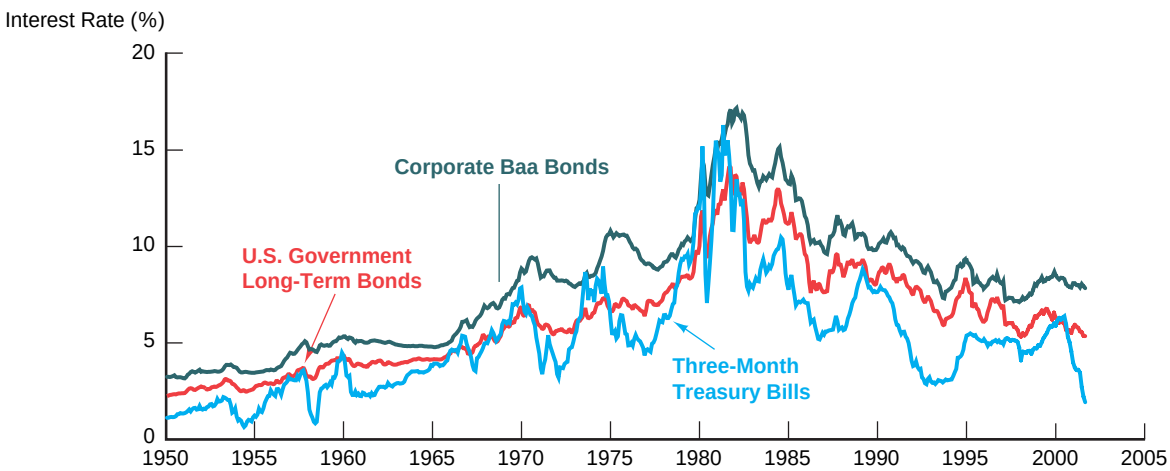


FIGURE 1 Interest Rates on Selected Bonds, 1950–2002

Sources: Federal Reserve *Bulletin*; www.federalreserve.gov/releases/H15/data.htm.

[1]The definition of *bond* used throughout this book is the broad one in common use by academics, which covers short- as well as long-term debt instruments. However, some practitioners in financial markets use the word *bond* to describe only specific long-term debt instruments such as corporate bonds or U.S. Treasury bonds.

shows, however, interest rates on several types of bonds can differ substantially. The interest rate on three-month Treasury bills, for example, fluctuates more than the other interest rates and is lower, on average. The interest rate on Baa (medium-quality) corporate bonds is higher, on average, than the other interest rates, and the spread between it and the other rates became larger in the 1970s.

In Chapter 2 we study the role of bond markets in the economy, and in Chapters 4 through 6 we examine what an interest rate is, how the common movements in interest rates come about, and why the interest rates on different bonds vary.

The Stock Market

http://stockcharts.com/charts/historical/

Historical charts of various stock indexes over differing time periods.

A **common stock** (typically just called a **stock**) represents a share of ownership in a corporation. It is a security that is a claim on the earnings and assets of the corporation. Issuing stock and selling it to the public is a way for corporations to raise funds to finance their activities. The stock market, in which claims on the earnings of corporations (shares of stock) are traded, is the most widely followed financial market in America (that's why it is often called simply "the market"). A big swing in the prices of shares in the stock market is always a big story on the evening news. People often speculate on where the market is heading and get very excited when they can brag about their latest "big killing," but they become depressed when they suffer a big loss. The attention the market receives can probably be best explained by one simple fact: It is a place where people can get rich—or poor—quickly.

As Figure 2 indicates, stock prices have been extremely volatile. After the market rose in the 1980s, on "Black Monday," October 19, 1987, it experienced the worst one-day drop in its entire history, with the Dow Jones Industrial Average (DJIA) falling by 22%. From then until 2000, the stock market experienced one of the great bull markets in its history, with the Dow climbing to a peak of over 11,000. With the collapse of the high-tech bubble in 2000, the stock market fell sharply, dropping by over 30% by 2002. These considerable fluctuations in stock prices affect the size of people's wealth and as a result may affect their willingness to spend.

The stock market is also an important factor in business investment decisions, because the price of shares affects the amount of funds that can be raised by selling newly issued stock to finance investment spending. A higher price for a firm's shares means that it can raise a larger amount of funds, which can be used to buy production facilities and equipment.

In Chapter 2 we examine the role that the stock market plays in the financial system, and we return to the issue of how stock prices behave and respond to information in the marketplace in Chapter 7.

The Foreign Exchange Market

For funds to be transferred from one country to another, they have to be converted from the currency in the country of origin (say, dollars) into the currency of the country they are going to (say, euros). The **foreign exchange market** is where this conversion takes place, and so it is instrumental in moving funds between countries. It is also important because it is where the **foreign exchange rate**, the price of one country's currency in terms of another's, is determined.

Figure 3 shows the exchange rate for the U.S. dollar from 1970 to 2002 (measured as the value of the American dollar in terms of a basket of major foreign currencies). The fluctuations in prices in this market have also been substantial: The dollar weakened considerably from 1971 to 1973, rose slightly in value until 1976, and then reached a low point in the 1978–1980 period. From 1980 to early 1985, the dollar appreciated dramatically in value, but since then it has fallen substantially.

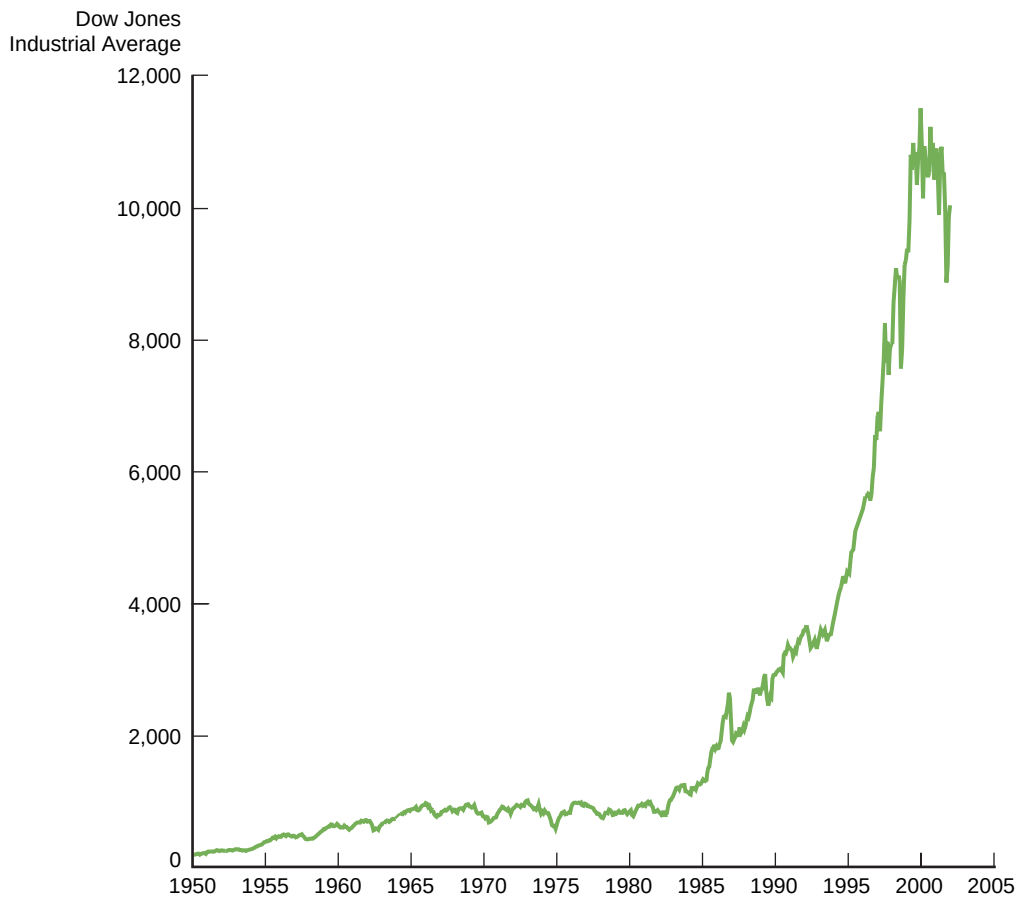


FIGURE 2 Stock Prices as Measured by the Dow Jones Industrial Average, 1950–2002

Source: Dow Jones Indexes: http://finance.yahoo.com/?u.

What have these fluctuations in the exchange rate meant to the American public and businesses? A change in the exchange rate has a direct effect on American consumers because it affects the cost of imports. In 2001 when the euro was worth around 85 cents, 100 euros of European goods (say, French wine) cost $85. When the dollar subsequently weakened, raising the cost of a euro near $1, the same 100 euros of wine now cost $100. Thus a weaker dollar leads to more expensive foreign goods, makes vacationing abroad more expensive, and raises the cost of indulging your desire for imported delicacies. When the value of the dollar drops, Americans will decrease their purchases of foreign goods and increase their consumption of domestic goods (such as travel in the United States or American-made wine).

Conversely, a strong dollar means that U.S. goods exported abroad will cost more in foreign countries, and hence foreigners will buy fewer of them. Exports of steel, for example, declined sharply when the dollar strengthened in the 1980–1985 and

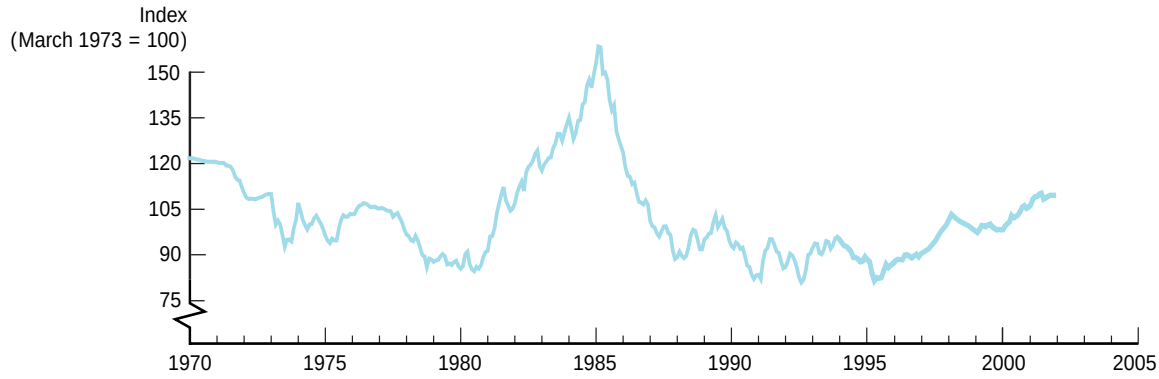


FIGURE 3 Exchange Rate of the U.S. Dollar, 1970–2002

Source: Federal Reserve: www.federalreserve.gov/releases/H10/summary.

1995–2001 periods. A strong dollar benefited American consumers by making foreign goods cheaper but hurt American businesses and eliminated some jobs by cutting both domestic and foreign sales of their products. The decline in the value of the dollar from 1985 to 1995 and 2001 to 2002 had the opposite effect: It made foreign goods more expensive, but made American businesses more competitive. Fluctuations in the foreign exchange markets have major consequences for the American economy.

In Chapter 19 we study how exchange rates are determined in the foreign exchange market in which dollars are bought and sold for foreign currencies.

Why Study Banking and Financial Institutions?

Part III of this book focuses on financial institutions and the business of banking. Banks and other financial institutions are what make financial markets work. Without them, financial markets would not be able to move funds from people who save to people who have productive investment opportunities. They thus also have important effects on the performance of the economy as a whole.

Structure of the Financial System

The financial system is complex, comprising many different types of private sector financial institutions, including banks, insurance companies, mutual funds, finance companies, and investment banks, all of which are heavily regulated by the government. If an individual wanted to make a loan to IBM or General Motors, for example, he or she would not go directly to the president of the company and offer a loan. Instead, he or she would lend to such companies indirectly through **financial intermediaries**, institutions that borrow funds from people who have saved and in turn make loans to others.

Why are financial intermediaries so crucial to well-functioning financial markets? Why do they extend credit to one party but not to another? Why do they usually write

complicated legal documents when they extend loans? Why are they the most heavily regulated businesses in the economy?

We answer these questions in Chapter 8 by developing a coherent framework for analyzing financial structure in the United States and in the rest of the world.

Banks and Other Financial Institutions

Banks are financial institutions that accept deposits and make loans. Included under the term *banks* are firms such as commercial banks, savings and loan associations, mutual savings banks, and credit unions. Banks are the financial intermediaries that the average person interacts with most frequently. A person who needs a loan to buy a house or a car usually obtains it from a local bank. Most Americans keep a large proportion of their financial wealth in banks in the form of checking accounts, savings accounts, or other types of bank deposits. Because banks are the largest financial intermediaries in our economy, they deserve the most careful study. However, banks are not the only important financial institutions. Indeed, in recent years, other financial institutions such as insurance companies, finance companies, pension funds, mutual funds, and investment banks have been growing at the expense of banks, and so we need to study them as well.

In Chapter 9, we examine how banks and other financial institutions manage their assets and liabilities to make profits. In Chapter 10, we extend the economic analysis in Chapter 8 to understand why bank regulation takes the form it does and what can go wrong in the regulatory process. In Chapters 11 and 12, we look at the banking industry and at nonbank financial institutions; we examine how the competitive environment has changed in these industries and learn why some financial institutions have been growing at the expense of others. Because the economic environment for banks and other financial institutions has become increasingly risky, these institutions must find ways to manage risk. How they manage risk with financial derivatives is the topic of Chapter 13.

Financial Innovation

In the good old days, when you took cash out of the bank or wanted to check your account balance, you got to say hello to the friendly human teller. Nowadays you are more likely to interact with an automatic teller machine when withdrawing cash, and you can get your account balance from your home computer. To see why these options have been developed, in Chapter 10 we study why and how financial innovation takes place, with particular emphasis on how the dramatic improvements in information technology have led to new means of delivering financial services electronically, in what has become known as **e-finance**. We also study financial innovation, because it shows us how creative thinking on the part of financial institutions can lead to higher profits. By seeing how and why financial institutions have been creative in the past, we obtain a better grasp of how they may be creative in the future. This knowledge provides us with useful clues about how the financial system may change over time and will help keep our knowledge about banks and other financial institutions from becoming obsolete.

Why Study Money and Monetary Policy?

Money, also referred to as the **money supply**, is defined as anything that is generally accepted in payment for goods or services or in the repayment of debts. Money is linked

to changes in economic variables that affect all of us and are important to the health of the economy. The final two parts of the book examine the role of money in the economy.

Money and Business Cycles

www.federalreserve.gov
General information, monetary policy, banking system, research, and economic data of the Federal Reserve.

The percentage of the labor force not working.

In 1981–1982, total production of goods and services (called **aggregate output**) in the U.S. economy fell and the **unemployment rate** (the percentage of the available labor force unemployed) rose to over 10%. After 1982, the economy began to expand rapidly, and by 1989 the unemployment rate had declined to 5%. In 1990, the eight-year expansion came to an end, with the unemployment rate rising above 7%. The economy bottomed out in 1991, and the subsequent recovery was the longest in U.S. history, with the unemployment rate falling to around 4%. A mild economic downturn then began in March 2001, with unemployment rising to 6%.

Why did the economy expand from 1982 to 1990, contract in 1990 to 1991, boom again from 1991 to 2001, and then contract again in 2001? Evidence suggests that money plays an important role in generating **business cycles**, the upward and downward movement of aggregate output produced in the economy. Business cycles affect all of us in immediate and important ways. When output is rising, for example, it is easier to find a good job; when output is falling, finding a good job might be difficult. Figure 4 shows the movements of the rate of money growth over the 1950–2002 period, with the shaded areas representing **recessions**, periods of declining aggregate output. What we see is that the rate of money growth has declined before every recession. Indeed, every recession since the beginning of the twentieth century has been preceded by a decline in the rate of money growth, indicating that

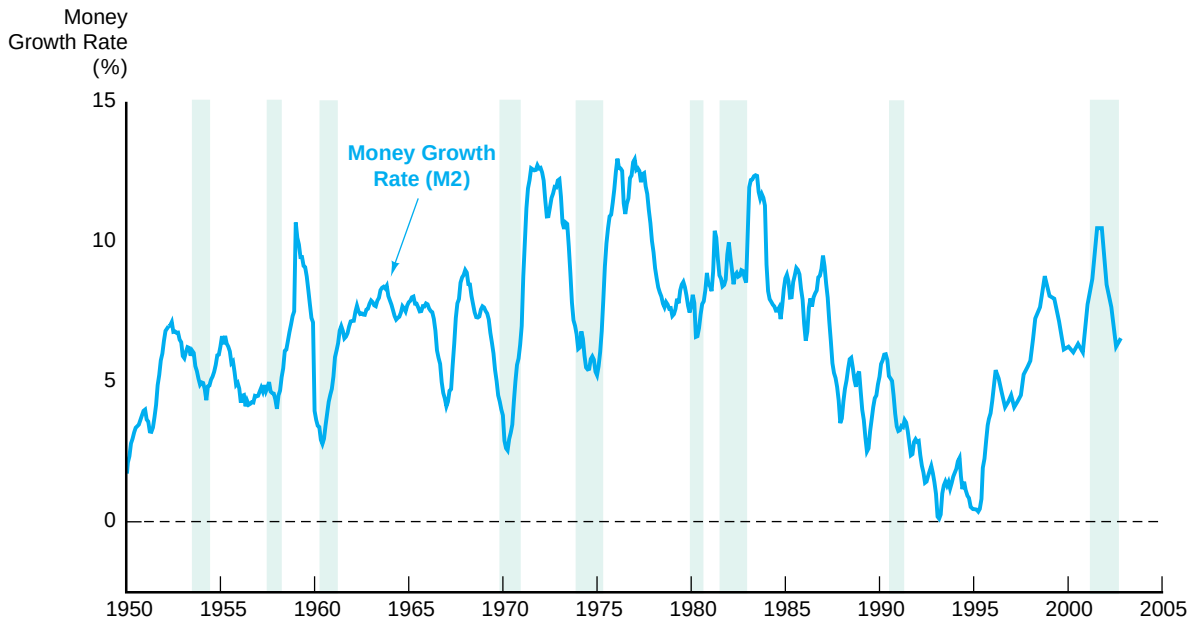


FIGURE 4 Money Growth (M2 Annual Rate) and the Business Cycle in the United States, 1950–2002

Note: Shaded areas represent recessions.
Source: Federal Reserve *Bulletin*, p. A4, Table 1.10; www.federalreserve.gov/releases/h6/hist/h6hist1.txt.

changes in money might also be a driving force behind business cycle fluctuations. However, not every decline in the rate of money growth is followed by a recession.

We explore how money might affect aggregate output in Chapters 22 through 28, where we study **monetary theory**, the theory that relates changes in the quantity of money to changes in aggregate economic activity and the price level.

Money and Inflation

Thirty years ago, the movie you might have paid $9 to see last week would have set you back only a dollar or two. In fact, for $9 you could probably have had dinner, seen the movie, and bought yourself a big bucket of hot buttered popcorn. As shown in Figure 5, which illustrates the movement of average prices in the U.S. economy from 1950 to 2002, the prices of most items are quite a bit higher now than they were then. The average price of goods and services in an economy is called the **aggregate price level**, or, more simply, the *price level* (a more precise definition is found in the appendix to this chapter). From 1950 to 2002, the price level has increased more than sixfold. **Inflation**, a continual increase in the price level, affects individuals, businesses, and the government. Inflation is generally regarded as an important problem to be solved and has often been a primary concern of politicians and policymakers. To solve the inflation problem, we need to know something about its causes.

www.newsengin.com/neFreeTools.nsf/CPIcalc?OpenView

Calculator lets you compute how buying power has changed since 1913.

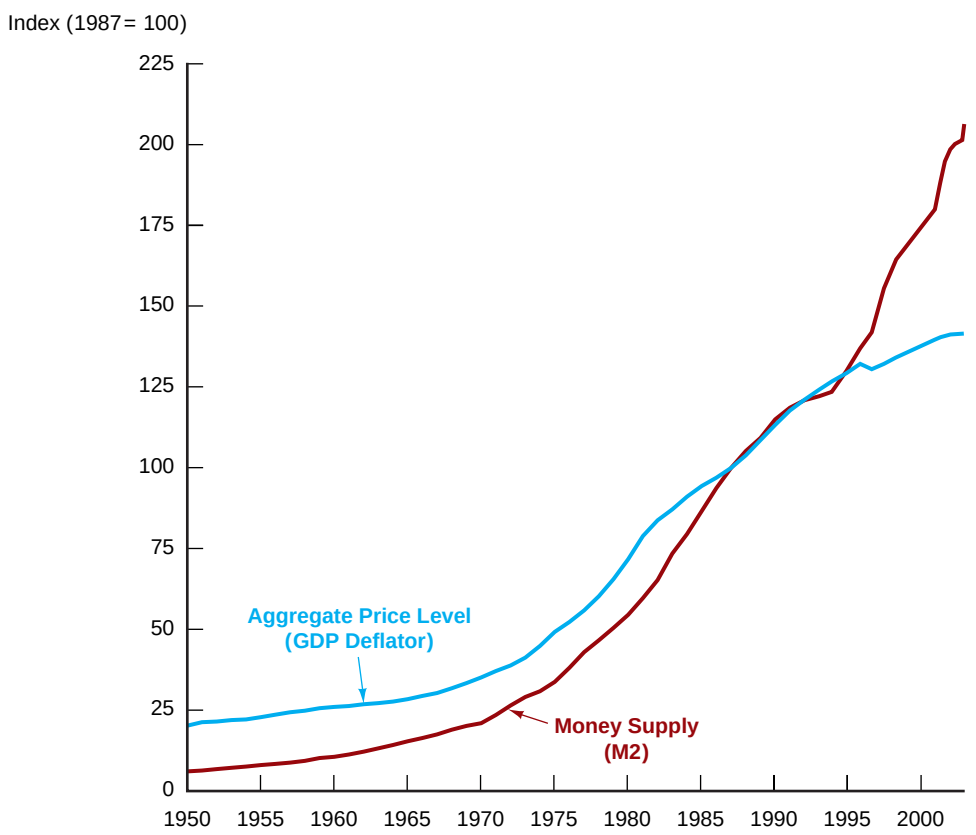


FIGURE 5 Aggregate Price Level and the Money Supply in the United States, 1950–2002

Sources: www.stls.frb.org/fred/data/gdp/gdpdef; www.federalreserve.gov/releases/h6/hist/h6hist10.txt.

What explains inflation? One clue to answering this question is found in Figure 5, which plots the money supply and the price level. As we can see, the price level and the money supply generally move closely together. These data seem to indicate that a continuing increase in the money supply might be an important factor in causing the continuing increase in the price level that we call inflation.

Further evidence that inflation may be tied to continuing increases in the money supply is found in Figure 6. For a number of countries, it plots the average **inflation rate** (the rate of change of the price level, usually measured as a percentage change per year) over the ten-year period 1992–2002 against the average rate of money growth over the same period. As you can see, there is a positive association between inflation and the growth rate of the money supply: The countries with the highest inflation rates are also the ones with the highest money growth rates. Belarus, Brazil, Romania, and Russia, for example, experienced very high inflation during this period, and their rates of money growth were high. By contrast, the United Kingdom and the United States had very low inflation rates over the same period, and their rates of money growth have been low. Such evidence led Milton Friedman, a Nobel laureate in economics, to make the famous statement, "Inflation is always and everywhere a monetary phenomenon."[2] We look at money's role in creating inflation by studying in detail the relationship between changes in the quantity of money and changes in the price level in Chapter 27.

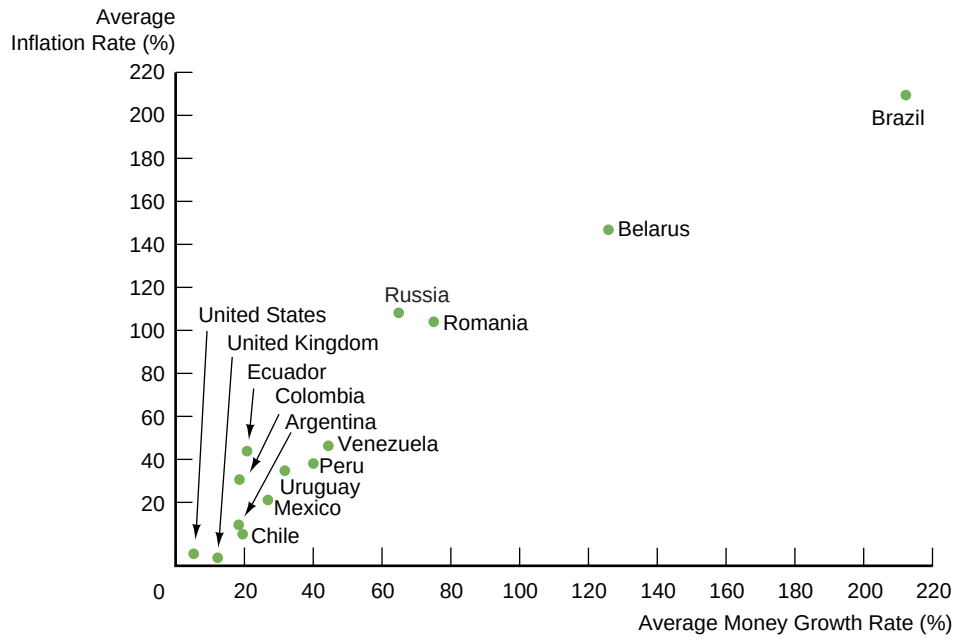


FIGURE 6 Average Inflation Rate Versus Average Rate of Money Growth for Selected Countries, 1992–2002

Source: International Financial Statistics.

[2]Milton Friedman, *Dollars and Deficits* (Upper Saddle River, N.J.: Prentice Hall, 1968), p. 39.

Money and Interest Rates

In addition to other factors, money plays an important role in interest-rate fluctuations, which are of great concern to businesses and consumers. Figure 7 shows the changes in the interest rate on long-term Treasury bonds and the rate of money growth. As the money growth rate rose in the 1960s and 1970s, the long-term bond rate rose with it. However, the relationship between money growth and interest rates has been less clear-cut since 1980. We analyze the relationship between money and interest rates when we examine the behavior of interest rates in Chapter 5.

Conduct of Monetary Policy

Because money can affect many economic variables that are important to the well-being of our economy, politicians and policymakers throughout the world care about the conduct of **monetary policy**, the management of money and interest rates. The organization responsible for the conduct of a nation's monetary policy is the **central bank**. The United States' central bank is the **Federal Reserve System** (also called simply **the Fed**). In Chapters 14–18 and 21, we study how central banks like the Federal Reserve System can affect the quantity of money in the economy and then look at how monetary policy is actually conducted in the United States and elsewhere.

Fiscal Policy and Monetary Policy

Fiscal policy involves decisions about government spending and taxation. A **budget deficit** is the excess of government expenditures over tax revenues for a particular time period, typically a year, while a **budget surplus** arises when tax revenues exceed government expenditures. The government must finance any deficit by borrowing, while a budget surplus leads to a lower government debt burden. As Figure 8 shows, the budget deficit, relative to the size of our economy, peaked in 1983 at 6% of national output (as calculated by the **gross domestic product**, or *GDP*, a measure of

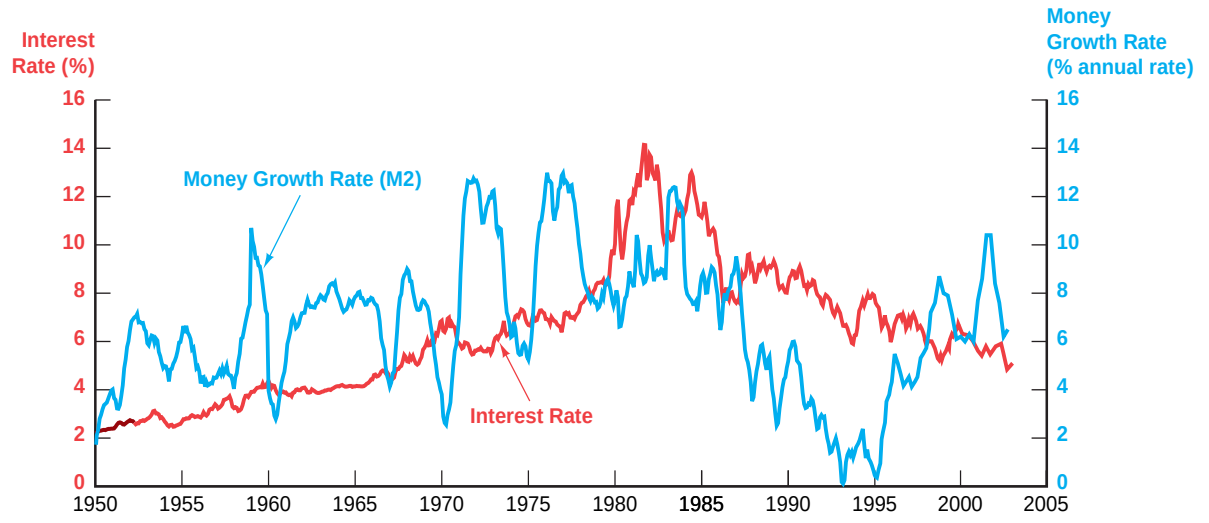


FIGURE 7 Money Growth (M2 Annual Rate) and Interest Rates (Long-Term U.S. Treasury Bonds), 1950–2002

Sources: Federal Reserve *Bulletin*, p. A4, Table 1.10; www.federalreserve.gov/releases/h6/hist/h6hist1.txt.

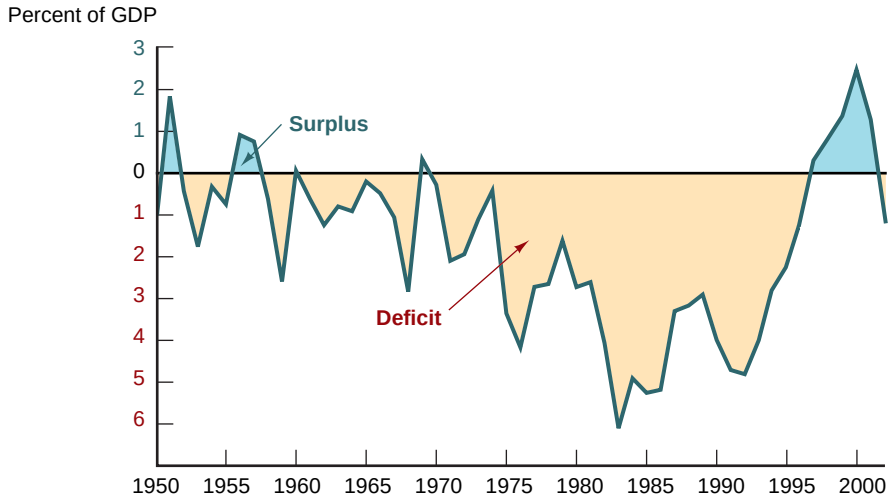


FIGURE 8 Government Budget Surplus or Deficit as a Percentage of Gross Domestic Product, 1950–2002

Source: http://w3.access.gpo.gov/usbudget/fy2003/spreadsheets.html.

www.kowaldesign.com/budget/

This site reports the current Federal budget deficit or surplus and how it has changed since the 1950s. It also reports how the federal budget is spent.

www.brillig.com/debt_clock/

National Debt clock. This site reports the exact national debt at each point in time.

aggregate output described in the appendix to this chapter). Since then, the budget deficit at first declined to less than 3% of GDP, rose again to over 5% by 1989, and fell subsequently, leading to budget surpluses from 1999 to 2001. In the aftermath of the terrorist attacks of September 11, 2001, the budget has swung back again into deficit. What to do about budget deficits and surpluses has been the subject of legislation and bitter battles between the president and Congress in recent years.

You may have heard statements in newspapers or on TV that budget surpluses are a good thing while deficits are undesirable. We explore the accuracy of such claims in Chapters 8 and 21 by seeing how budget deficits might lead to a financial crisis as they did in Argentina in 2001. In Chapter 27, we examine why deficits might result in a higher rate of money growth, a higher rate of inflation, and higher interest rates.

How We Will Study Money, Banking, and Financial Markets

This textbook stresses the economic way of thinking by developing a unifying framework to study money, banking, and financial markets. This analytic framework uses a few basic economic concepts to organize your thinking about the determination of asset prices, the structure of financial markets, bank management, and the role of money in the economy. It encompasses the following basic concepts:

- A simplified approach to the demand for assets
- The concept of equilibrium
- Basic supply and demand to explain behavior in financial markets
- The search for profits

- An approach to financial structure based on transaction costs and asymmetric information
- Aggregate supply and demand analysis

The unifying framework used in this book will keep your knowledge from becoming obsolete and make the material more interesting. It will enable you to learn what *really* matters without having to memorize a mass of dull facts that you will forget soon after the final exam. This framework will also provide you with the tools to understand trends in the financial marketplace and in variables such as interest rates, exchange rates, inflation, and aggregate output.

To help you understand and apply the unifying analytic framework, simple models are constructed in which the variables held constant are carefully delineated, each step in the derivation of the model is clearly and carefully laid out, and the models are then used to explain various phenomena by focusing on changes in one variable at a time, holding all other variables constant.

To reinforce the models' usefulness, this text uses case studies, applications, and special-interest boxes to present evidence that supports or casts doubts on the theories being discussed. This exposure to real-life events and data should dissuade you from thinking that all economists make abstract assumptions and develop theories that have little to do with actual behavior.

To function better in the real world outside the classroom, you must have the tools to follow the financial news that appears in leading financial publications such as the *Wall Street Journal*. To help and encourage you to read the financial section of the newspaper, this book contains two special features. The first is a set of special boxed inserts titled "Following the Financial News" that contain actual columns and data from the *Wall Street Journal* that typically appear daily or periodically. These boxes give you the detailed information and definitions you need to evaluate the data being presented. The second feature is a set of special applications titled "Reading the *Wall Street Journal*" that expand on the "Following the Financial News" boxes. These applications show you how the analytic framework in the book can be used directly to make sense of the daily columns in the United States' leading financial newspaper. In addition to these applications, this book also contains nearly 400 end-of-chapter problems that ask you to apply the analytic concepts you have learned to other real-world issues. Particularly relevant is a special class of problems headed "Predicting the Future." So that you can work on many of these problems on your own, answers to half of them are found at the end of the book. These give you an opportunity to review and apply many of the important financial concepts and tools presented throughout the book.

Exploring the Web

The World Wide Web has become an extremely valuable and convenient resource for financial research. We emphasize the importance of this tool in several ways. First, wherever we utilize the Web to find information to build the charts and tables that appear throughout the text, we include the source site's URL. These sites often contain additional information and are updated frequently. Second, in the margin of the text, we have included the URLs of sites related to the material being discussed. Visit these sites to further explore a topic you find of particular interest. Finally, we have added Web exercises to the end of each chapter. These exercises prompt you to visit sites related to the chapter and to work with real-time data and information.

Web site URLs are subject to frequent change. We have tried to select stable sites, but we realize that even government URLs change. The publisher's web site (www.aw.com/mishkin) will maintain an updated list of current URLs for your reference.

A sample Web exercise has been included in this chapter. This is an especially important example, since it demonstrates how to export data from a web site into Microsoft® Excel for further analysis. We suggest you work through this problem on your own so that you will be able to perform this activity when prompted in subsequent Web exercises.

Web Exercises

You have been hired by Risky Ventures, Inc., as a consultant to help them analyze interest rate trends. They are initially interested in determining the historical relationship between long- and short-term interest rates. The biggest task you must immediately undertake is collecting market interest-rate data. You know the best source of this information is the Web.

1. You decide that your best indicator of long-term interest rates is the 30-year U.S. Treasury note. Your first task is to gather historical data. Go to www.federalreserve.gov/releases/ and click "H.15 Selected Interest Rates, Historical data." The site should look like Figure 9.
   a. Click on "Historical data." Scroll down to "U.S. Government securities/Treasury constant maturities/30 year." Scroll over to the right and click on "annual."

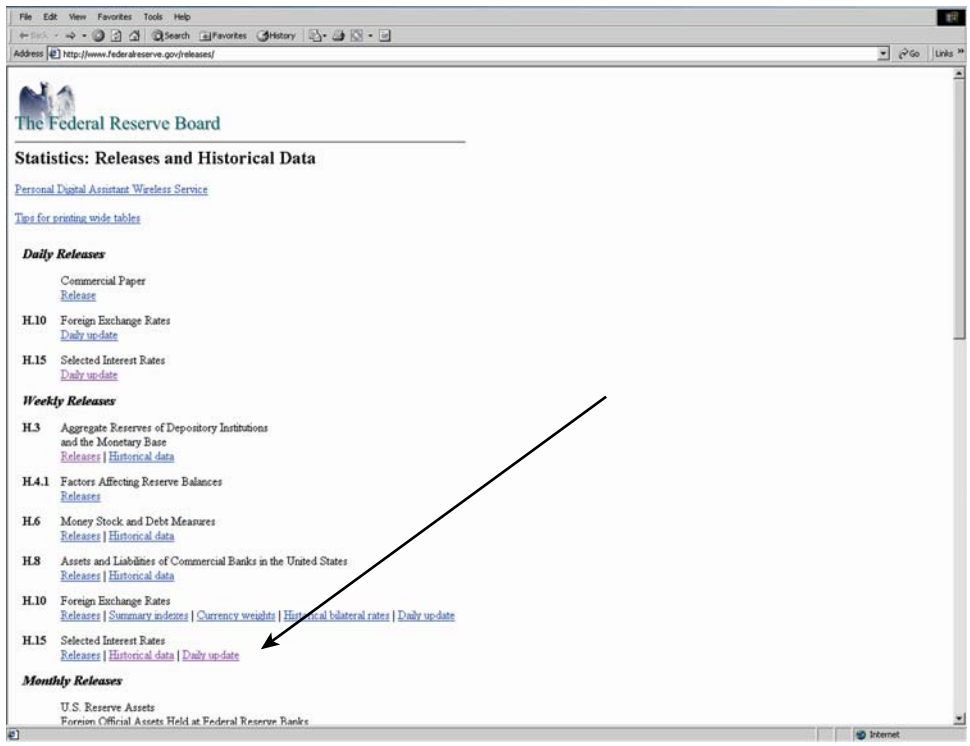


FIGURE 9 Federal Reserve Board Web Site

b. While you have located an accurate source of historical interest rate data, getting it onto a spreadsheet will be very tedious. You recall that Excel will let you convert text data into columns. Begin by highlighting the two columns of data (the year and rate). Right-click on the mouse and choose COPY. Now open Excel and put the cursor in a cell. Click PASTE. Now choose DATA from the tool bar and click on TEXT TO COLUMNS. Follow the wizard (Figure 10), checking the fixed-width option. The list of interest rates should now have the year in one column and the interest rate in the next column. Label your columns.

   Repeat the above steps to collect the 1-year interest rate series. Put it in the column next to the 30-year series. Be sure to line up the years correctly and delete any years that are not included in both series.

c. You now want to analyze the interest rates by graphing them. Again highlight the two columns of data you just created in Excel. Click on the charts icon on the tool bar (or INSERT/CHART). Select scatter diagram and choose any type of scatter diagram that connects the dots. Let the Excel wizard take you through the steps of completing the graph. (See Figure 11.)

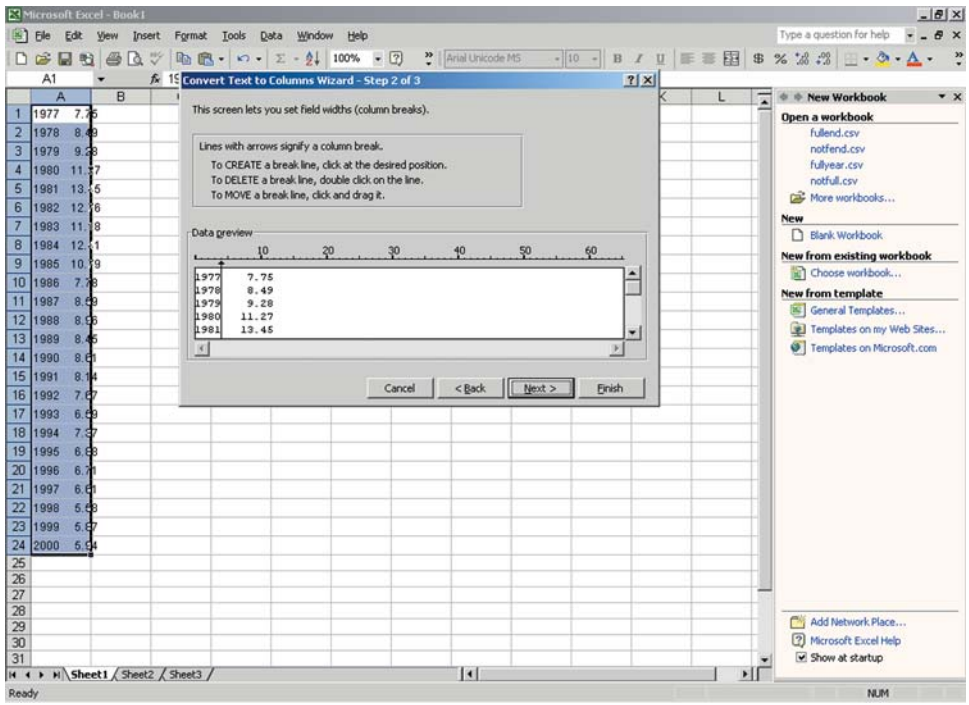


FIGURE 10 Excel Spreadsheet with Interest Rate Data

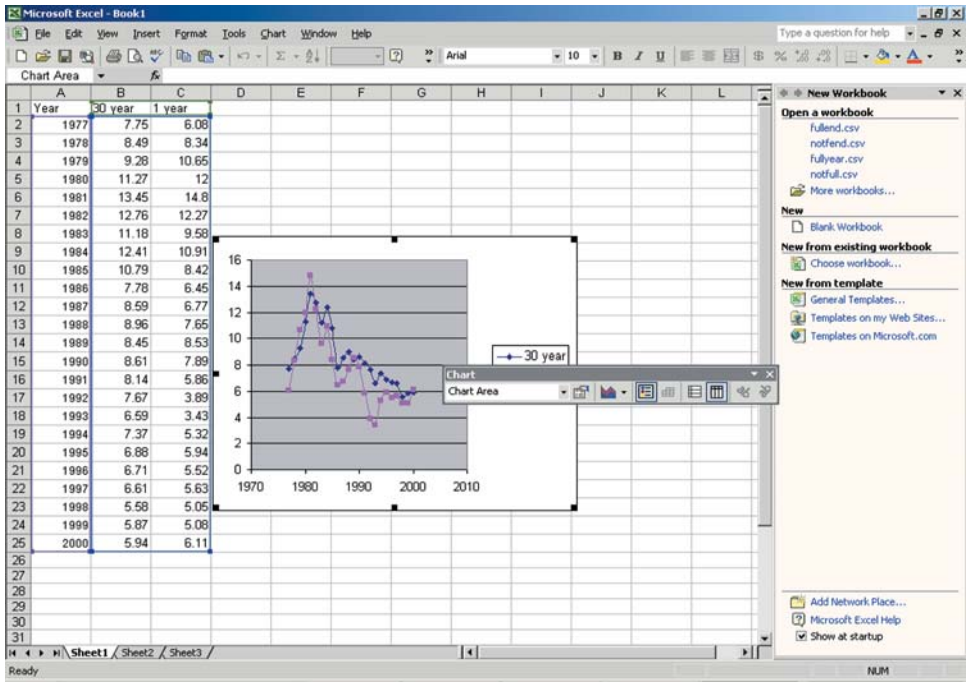


FIGURE 11 Excel Graph of Interest Rate Data

Concluding Remarks

The topic of money, banking, and financial markets is an exciting field that directly affects your life—interest rates influence earnings on your savings and the payments on loans you may seek on a car or a house, and monetary policy may affect your job prospects and the prices of goods in the future. Your study of money, banking, and financial markets will introduce you to many of the controversies about the conduct of economic policy that are currently the subject of hot debate in the political arena and will help you gain a clearer understanding of economic phenomena you frequently hear about in the news media. The knowledge you gain will stay with you and benefit you long after the course is done.

Summary

1. Activities in financial markets have direct effects on individuals' wealth, the behavior of businesses, and the efficiency of our economy. Three financial markets deserve particular attention: the bond market (where interest rates are determined), the stock market (which has a major effect on people's wealth and on firms' investment decisions), and the foreign exchange market (because fluctuations in the foreign exchange rate have major consequences for the American economy).

2. Banks and other financial institutions channel funds from people who might not put them to productive use to people who can do so and thus play a crucial role in improving the efficiency of the economy.
3. Money appears to be a major influence on inflation, business cycles, and interest rates. Because these economic variables are so important to the health of the economy, we need to understand how monetary policy is and should be conducted. We also need to study government fiscal policy because it can be an influential factor in the conduct of monetary policy.
4. This textbook stresses the economic way of thinking by developing a unifying analytic framework for the study of money, banking, and financial markets using a few basic economic principles. This textbook also emphasizes the interaction of theoretical analysis and empirical data.

Key Terms

aggregate income (appendix), p. 20
aggregate output, p. 9
aggregate price level, p. 10
asset, p. 3
banks, p. 8
bond, p. 3
budget deficit, p. 12
budget surplus, p. 12
business cycles, p. 9
central bank, p. 12
common stock, p. 5
e-finance, p. 8
Federal Reserve System (the Fed), p. 12
financial intermediaries, p. 7
financial markets, p. 3
fiscal policy, p. 12
foreign exchange market, p. 5
foreign exchange rate, p. 5
gross domestic product (appendix), p. 12, 20
inflation, p. 10
inflation rate, p. 11
interest rate, p. 4
monetary policy, p. 12
monetary theory, p. 10
money (money supply), p. 8
recession, p. 9
security, p. 3
stock, p. 5
unemployment rate, p. 9

Questions and Problems

Questions marked with an asterisk are answered at the end of the book in an appendix, "Answers to Selected Questions and Problems."

1. Has the inflation rate in the United States increased or decreased in the past few years? What about interest rates?
*2. If history repeats itself and we see a decline in the rate of money growth, what might you expect to happen to:
   a. real output
   b. the inflation rate, and
   c. interest rates?
3. When was the most recent recession?
*4. When interest rates fall, how might you change your economic behavior?
5. Can you think of any financial innovation in the past ten years that has affected you personally? Has it made you better off or worse off? Why?
*6. Is everybody worse off when interest rates rise?
7. What is the basic activity of banks?
*8. Why are financial markets important to the health of the economy?
9. What is the typical relationship between interest rates on three-month Treasury bills, long-term Treasury bonds, and Baa corporate bonds?
*10. What effect might a fall in stock prices have on business investment?
11. What effect might a rise in stock prices have on consumers' decisions to spend?

*12. How does a fall in the value of the pound sterling affect British consumers?

13. How does an increase in the value of the pound sterling affect American businesses?

*14. Looking at Figure 3, in what years would you have chosen to visit the Grand Canyon in Arizona rather than the Tower of London?

15. When the dollar is worth more in relation to currencies of other countries, are you more likely to buy American-made or foreign-made jeans? Are U.S. companies that manufacture jeans happier when the dollar is strong or when it is weak? What about an American company that is in the business of importing jeans into the United States?

Web Exercises

1. In this exercise we are going to practice collecting data from the Web and graphing it using Excel. Use the example in the text as a guide. Go to www.forecasts.org/data/index.htm, click on "stock indices" at the top of the page then choose the U.S. Stock indices – monthly option. Finally, choose the Dow Jones Industrial Average option.
   a. Using the method presented in this chapter, move the data into an Excel spreadsheet.
   b. Using the data from a, prepare a graph. Use the graphing wizard to properly label your axes.
2. In Web Exercise 1 you collected and graphed the Dow Jones Industrial Average. This same site reports forecast values of the DJIA. Go to www.forecasts.org/data/index.htm and click on "FFC Home" at the top of the page. Click on the Dow Jones Industrial link under Forecasts in the far left column.
   a. What is the Dow forecast to be in 3 months?
   b. What percentage increase is forecast for the next three months?

appendix to chapter I

Defining Aggregate Output, Income, the Price Level, and the Inflation Rate

Because these terms are used so frequently throughout the text, we need to have a clear understanding of the definitions of *aggregate output*, *income*, the *price level*, and the *inflation rate*.

Aggregate Output and Income

The most commonly reported measure of aggregate output, the **gross domestic product (GDP)**, is the market value of all final goods and services produced in a country during the course of the year. This measure excludes two sets of items that at first glance you might think would be included. Purchases of goods that have been produced in the past, whether a Rembrandt painting or a house built 20 years ago, are not counted as part of GDP, nor are purchases of stocks or bonds. None of these enter into GDP because they are not goods and services produced during the course of the year. Intermediate goods, which are used up in producing final goods and services, such as the sugar in a candy bar or the energy used to produce steel, are also not counted separately as part of GDP. Because the value of the final goods already includes the value of the intermediate goods, to count them separately would be to count them twice.

Aggregate income, the total income of *factors of production* (land, labor, and capital) from producing goods and services in the economy during the course of the year, is best thought of as being equal to aggregate output. Because the payments for final goods and services must eventually flow back to the owners of the factors of production as income, income payments must equal payments for final goods and services. For example, if the economy has an aggregate output of $10 trillion, total income payments in the economy (aggregate income) are also $10 trillion.

Real Versus Nominal Magnitudes

When the total value of final goods and services is calculated using current prices, the resulting GDP measure is referred to as *nominal GDP*. The word *nominal* indicates that values are measured using current prices. If all prices doubled but actual production of goods and services remained the same, nominal GDP would double even though

people would not enjoy the benefits of twice as many goods and services. As a result, nominal variables can be misleading measures of economic well-being.

A more reliable measure of economic well-being expresses values in terms of prices for an arbitrary base year, currently 1996. GDP measured with constant prices is referred to as *real GDP*, the word *real* indicating that values are measured in terms of fixed prices. Real variables thus measure the quantities of goods and services and do not change because prices have changed, but rather only if actual quantities have changed.

A brief example will make the distinction clearer. Suppose that you have a nominal income of $30,000 in 2004 and that your nominal income was $15,000 in 1996. If all prices doubled between 1996 and 2004, are you better off? The answer is no: Although your income has doubled, your $30,000 buys you only the same amount of goods because prices have also doubled. A real income measure indicates that your income in terms of the goods it can buy is the same. Measured in 1996 prices, the $30,000 of nominal income in 2004 turns out to be only $15,000 of real income. Because your real income is actually the same in the two years, you are no better or worse off in 2004 than you were in 1996.

Because real variables measure quantities in terms of real goods and services, they are typically of more interest than nominal variables. In this text, discussion of aggregate output or aggregate income always refers to real measures (such as real GDP).

Aggregate Price Level

In this chapter, we defined the aggregate price level as a measure of average prices in the economy. Three measures of the aggregate price level are commonly encountered in economic data. The first is the *GDP deflator*, which is defined as nominal GDP divided by real GDP. Thus if 2004 nominal GDP is $10 trillion but 2004 real GDP in 1996 prices is $9 trillion,

$$\text{GDP deflator} = \frac{\$10 \text{ trillion}}{\$9 \text{ trillion}} = 1.11$$

The GDP deflator equation indicates that, on average, prices have risen 11 percent since 1996. Typically, measures of the price level are presented in the form of a price index, which expresses the price level for the base year (in our example, 1996) as 100. Thus the GDP deflator for 2004 would be 111.

Another popular measure of the aggregate price level (which officials in the Fed frequently focus on) is the *PCE deflator*, which is similar to the GDP deflator and is defined as nominal personal consumption expenditures (PCE) divided by real PCE. The measure of the aggregate price level that is most frequently reported in the press is the *consumer price index (CPI)*. The CPI is measured by pricing a "basket" list of goods and services bought by a typical urban household. If over the course of the year, the cost of this basket of goods and services rises from $500 to $600, the CPI has risen by 20 percent. The CPI is also expressed as a price index with the base year equal to 100.

The CPI, the PCE deflator, and the GDP deflator measures of the price level can be used to convert or deflate a nominal magnitude into a real magnitude. This is accomplished by dividing the nominal magnitude by the price index. In our example,

in which the GDP deflator for 2004 is 1.11 (expressed as an index value of 111), real GDP for 2004 equals

$$\frac{\$10 \text{ trillion}}{1.11} = \$9 \text{ trillion in 1996 prices}$$

which corresponds to the real GDP figure for 2004 mentioned earlier.

Growth Rates and the Inflation Rate

The media often talk about the economy's growth rate, and particularly the growth rate of real GDP. A growth rate is defined as the percentage change in a variable,[1] i.e.,

$$\text{growth rate} = \frac{x_t - x_{t-1}}{x_{t-1}} \times 100$$

where t indicates today and $t - 1$ a year earlier.

For example, if real GDP grew from \$9 trillion in 2004 to \$9.5 trillion in 2005, then the GDP growth rate for 2005 would be 5.6%:

$$\text{GDP growth rate} = \frac{\$9.5 \text{ trillion} - \$9 \text{ trillion}}{\$9 \text{ trillion}} \times 100 = 5.6\%$$

The inflation rate is defined as the growth rate of the aggregate price level. Thus if the GDP deflator rose from 111 in 2004 to 113 in 2005, the inflation rate using the GDP deflator would be 1.8%:

$$\text{inflation rate} = \frac{113 - 111}{111} \times 100 = 1.8\%$$

[1] If the growth rate is for a period less than one year, it is usually reported on an annualized basis; that is, it is converted to the growth rate over a year's time, assuming that the growth rate remains constant. For GDP, which is reported quarterly, the annualized growth rate would be approximately four times the percentage change in GDP from the previous quarter. For example, if GDP rose $\frac{1}{2}$% from the first quarter of 2004 to the second quarter of 2004, then the annualized GDP growth rate for the second quarter of 2004 would be reported as 2% ($= 4 \times \frac{1}{2}$%). (A more accurate calculation would be 2.02%, because a precise quarterly growth rate should be compounded on a quarterly basis.)

Chapter 2

An Overview of the Financial System

PREVIEW

Inez the Inventor has designed a low-cost robot that cleans house (even does windows), washes the car, and mows the lawn, but she has no funds to put her wonderful invention into production. Walter the Widower has plenty of savings, which he and his wife accumulated over the years. If we could get Inez and Walter together so that Walter could provide funds to Inez, Inez's robot would see the light of day, and the economy would be better off: We would have cleaner houses, shinier cars, and more beautiful lawns.

Financial markets (bond and stock markets) and financial intermediaries (banks, insurance companies, pension funds) have the basic function of getting people like Inez and Walter together by moving funds from those who have a surplus of funds (Walter) to those who have a shortage of funds (Inez). More realistically, when IBM invents a better computer, it may need funds to bring it to market. Similarly, when a local government needs to build a road or a school, it may need more funds than local property taxes provide. Well-functioning financial markets and financial intermediaries are crucial to economic health.

To study the effects of financial markets and financial intermediaries on the economy, we need to acquire an understanding of their general structure and operation. In this chapter, we learn about the major financial intermediaries and the instruments that are traded in financial markets as well as how these markets are regulated.

This chapter presents an overview of the fascinating study of financial markets and institutions. We return to a more detailed treatment of the regulation, structure, and evolution of the financial system in Chapters 8 through 13.

Function of Financial Markets

Financial markets perform the essential economic function of channeling funds from households, firms, and governments that have saved surplus funds by spending less than their income to those that have a shortage of funds because they wish to spend more than their income. This function is shown schematically in Figure 1. Those who have saved and are lending funds, the lender-savers, are at the left, and those who must borrow funds to finance their spending, the borrower-spenders, are at the right. The principal lender-savers are households, but business enterprises and the government (particularly state and local government), as well as foreigners and their governments, sometimes also find themselves with excess funds and so lend them out.

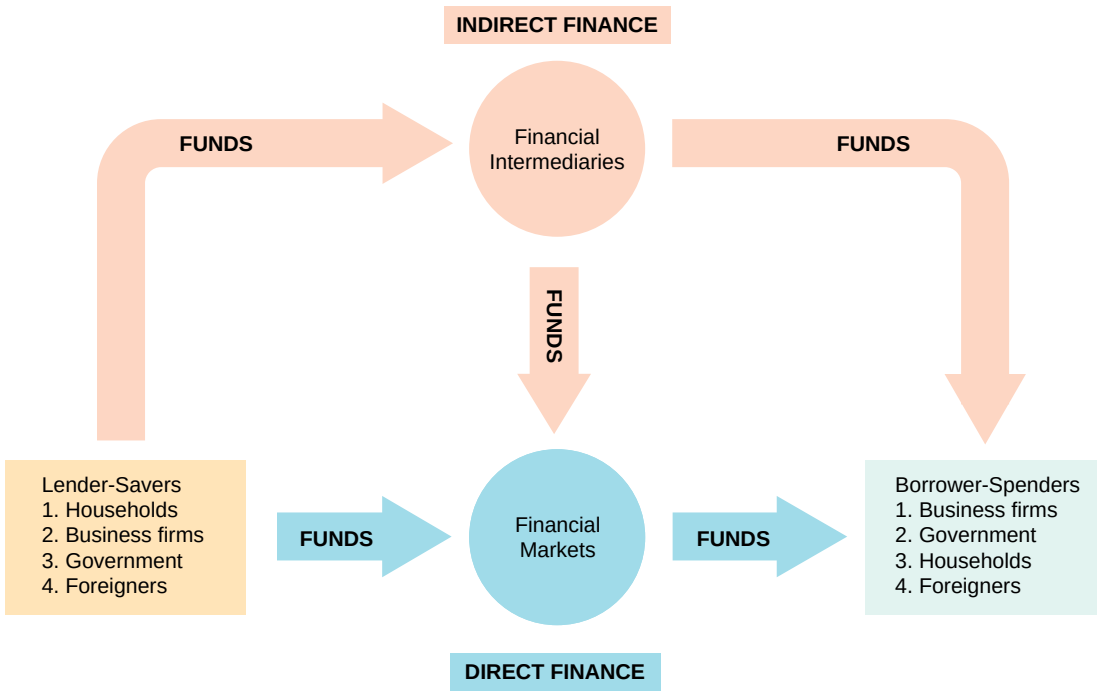


FIGURE 1 Flows of Funds Through the Financial System

The most important borrower-spenders are businesses and the government (particularly the federal government), but households and foreigners also borrow to finance their purchases of cars, furniture, and houses. The arrows show that funds flow from lender-savers to borrower-spenders via two routes.

In *direct finance* (the route at the bottom of Figure 1), borrowers borrow funds directly from lenders in financial markets by selling them *securities* (also called *financial instruments*), which are claims on the borrower's future income or assets. Securities are assets for the person who buys them but **liabilities** (IOUs or debts) for the individual or firm that sells (issues) them. For example, if General Motors needs to borrow funds to pay for a new factory to manufacture electric cars, it might borrow the funds from savers by selling them *bonds*, debt securities that promise to make payments periodically for a specified period of time.

Why is this channeling of funds from savers to spenders so important to the economy? The answer is that the people who save are frequently not the same people who have profitable investment opportunities available to them, the entrepreneurs. Let's first think about this on a personal level. Suppose that you have saved $1,000 this year, but no borrowing or lending is possible because there are no financial markets. If you do not have an investment opportunity that will permit you to earn income with your savings, you will just hold on to the $1,000 and will earn no interest. However, Carl the Carpenter has a productive use for your $1,000: He can use it to

purchase a new tool that will shorten the time it takes him to build a house, thereby earning an extra $200 per year. If you could get in touch with Carl, you could lend him the $1,000 at a rental fee (interest) of $100 per year, and both of you would be better off. You would earn $100 per year on your $1,000, instead of the zero amount that you would earn otherwise, while Carl would earn $100 more income per year (the $200 extra earnings per year minus the $100 rental fee for the use of the funds).

In the absence of financial markets, you and Carl the Carpenter might never get together. Without financial markets, it is hard to transfer funds from a person who has no investment opportunities to one who has them; you would both be stuck with the status quo, and both of you would be worse off. Financial markets are thus essential to promoting economic efficiency.

The existence of financial markets is also beneficial even if someone borrows for a purpose other than increasing production in a business. Say that you are recently married, have a good job, and want to buy a house. You earn a good salary, but because you have just started to work, you have not yet saved much. Over time, you would have no problem saving enough to buy the house of your dreams, but by then you would be too old to get full enjoyment from it. Without financial markets, you are stuck; you cannot buy the house and must continue to live in your tiny apartment.

If a financial market were set up so that people who had built up savings could lend you the funds to buy the house, you would be more than happy to pay them some interest in order to own a home while you are still young enough to enjoy it. Then, over time, you would pay back your loan. The overall outcome would be such that you would be better off, as would the persons who made you the loan. They would now earn some interest, whereas they would not if the financial market did not exist.

Now we can see why financial markets have such an important function in the economy. They allow funds to move from people who lack productive investment opportunities to people who have such opportunities. Thus financial markets are critical for producing an efficient allocation of capital, which contributes to higher production and efficiency for the overall economy. Indeed, as we will explore in Chapter 8, when financial markets break down during financial crises, as they have in Mexico, East Asia, and Argentina in recent years, severe economic hardship results, which can even lead to dangerous political instability.

Well-functioning financial markets also directly improve the well-being of consumers by allowing them to time their purchases better. They provide funds to young people to buy what they need and can eventually afford without forcing them to wait until they have saved up the entire purchase price. Financial markets that are operating efficiently improve the economic welfare of everyone in the society.

Structure of Financial Markets

Now that we understand the basic function of financial markets, let's look at their structure. The following descriptions of several categorizations of financial markets illustrate essential features of these markets.

Debt and Equity Markets

A firm or an individual can obtain funds in a financial market in two ways. The most common method is to issue a debt instrument, such as a bond or a mortgage, which is a contractual agreement by the borrower to pay the holder of the instrument fixed

dollar amounts at regular intervals (interest and principal payments) until a specified date (the maturity date), when a final payment is made. The **maturity** of a debt instrument is the number of years (term) until that instrument's expiration date. A debt instrument is **short-term** if its maturity is less than a year and **long-term** if its maturity is ten years or longer. Debt instruments with a maturity between one and ten years are said to be **intermediate-term**.

http://stockcharts.com/def/servlet/Favorites.CServlet?obj=msummary&cmd=show&disp=SXA

This site contains historical stock market index charts for many countries around the world.

The second method of raising funds is by issuing **equities**, such as common stock, which are claims to share in the net income (income after expenses and taxes) and the assets of a business. If you own one share of common stock in a company that has issued one million shares, you are entitled to 1 one-millionth of the firm's net income and 1 one-millionth of the firm's assets. Equities often make periodic payments (**dividends**) to their holders and are considered long-term securities because they have no maturity date. In addition, owning stock means that you own a portion of the firm and thus have the right to vote on issues important to the firm and to elect its directors.

The main disadvantage of owning a corporation's equities rather than its debt is that an equity holder is a *residual claimant*; that is, the corporation must pay all its debt holders before it pays its equity holders. The advantage of holding equities is that equity holders benefit directly from any increases in the corporation's profitability or asset value because equities confer ownership rights on the equity holders. Debt holders do not share in this benefit, because their dollar payments are fixed. We examine the pros and cons of debt versus equity instruments in more detail in Chapter 8, which provides an economic analysis of financial structure.

The total value of equities in the United States has typically fluctuated between $1 and $20 trillion since the early 1970s, depending on the prices of shares. Although the average person is more aware of the stock market than any other financial market, the size of the debt market is often larger than the size of the equities market: The value of debt instruments was $20 trillion at the end of 2002 while the value of equities was $11 trillion at the end of 2002.

Primary and Secondary Markets

A **primary market** is a financial market in which new issues of a security, such as a bond or a stock, are sold to initial buyers by the corporation or government agency borrowing the funds. A **secondary market** is a financial market in which securities that have been previously issued (and are thus secondhand) can be resold.

The primary markets for securities are not well known to the public because the selling of securities to initial buyers often takes place behind closed doors. An important financial institution that assists in the initial sale of securities in the primary market is the **investment bank**. It does this by **underwriting** securities: It guarantees a price for a corporation's securities and then sells them to the public.

www.nyse.com

New York Stock Exchange. Find listed companies, quotes, company historical data, real-time market indices, and more.

The New York and American stock exchanges and NASDAQ, in which previously issued stocks are traded, are the best-known examples of secondary markets, although the bond markets, in which previously issued bonds of major corporations and the U.S. government are bought and sold, actually have a larger trading volume. Other examples of secondary markets are foreign exchange markets, futures markets, and options markets. Securities brokers and dealers are crucial to a well-functioning secondary market. **Brokers** are agents of investors who match buyers with sellers of securities; **dealers** link buyers and sellers by buying and selling securities at stated prices.

When an individual buys a security in the secondary market, the person who has sold the security receives money in exchange for the security, but the corporation that

issued the security acquires no new funds. A corporation acquires new funds only when its securities are first sold in the primary market. Nonetheless, secondary markets serve two important functions. First, they make it easier and quicker to sell these financial instruments to raise cash; that is, they make the financial instruments more **liquid**. The increased liquidity of these instruments then makes them more desirable and thus easier for the issuing firm to sell in the primary market. Second, they determine the price of the security that the issuing firm sells in the primary market. The investors that buy securities in the primary market will pay the issuing corporation no more than the price they think the secondary market will set for this security. The higher the security's price in the secondary market, the higher will be the price that the issuing firm will receive for a new security in the primary market, and hence the greater the amount of financial capital it can raise. Conditions in the secondary market are therefore the most relevant to corporations issuing securities. It is for this reason that books like this one, that deal with financial markets, focus on the behavior of secondary markets rather than primary markets.

Secondary markets in which buyers and sellers of securities (or their agents or brokers) meet in one central location to

Exchanges and Over-the-Counter Markets

Secondary markets can be organized in two ways. One is to organize **exchanges**, where buyers and sellers of securities (or their agents or brokers) meet in one central location to conduct trades. The New York and American stock exchanges for stocks and the Chicago Board of Trade for commodities (wheat, corn, silver, and other raw materials) are examples of organized exchanges.

www.nasdaq.com

Detailed market and security information for the NASDAQ OTC stock exchange.

The other method of organizing a secondary market is to have an **over-the-counter (OTC) market**, in which dealers at different locations who have an inventory of securities stand ready to buy and sell securities "over the counter" to anyone who comes to them and is willing to accept their prices. Because over-the-counter dealers are in computer contact and know the prices set by one another, the OTC market is very competitive and not very different from a market with an organized exchange.

Many common stocks are traded over-the-counter, although a majority of the largest corporations have their shares traded at organized stock exchanges such as the New York Stock Exchange. The U.S. government bond market, with a larger trading volume than the New York Stock Exchange, is set up as an over-the-counter market. Forty or so dealers establish a "market" in these securities by standing ready to buy and sell U.S. government bonds. Other over-the-counter markets include those that trade other types of financial instruments such as negotiable certificates of deposit, federal funds, banker's acceptances, and foreign exchange.

Money and Capital Markets

Another way of distinguishing between markets is on the basis of the maturity of the securities traded in each market. The **money market** is a financial market in which only short-term debt instruments (generally those with original maturity of less than one year) are traded; the **capital market** is the market in which longer-term debt (generally those with original maturity of one year or greater) and equity instruments are traded. Money market securities are usually more widely traded than longer-term securities and so tend to be more liquid. In addition, as we will see in Chapter 4, short-term securities have smaller fluctuations in prices than long-term securities, making them safer investments. As a result, corporations and banks actively use the money market to earn interest on surplus funds that they expect to have only temporarily. Capital market securities, such as stocks and long-term bonds, are often

held by financial intermediaries such as insurance companies and pension funds, which have little uncertainty about the amount of funds they will have available in the future.[1]

Internationalization of Financial Markets

The growing internationalization of financial markets has become an important trend. Before the 1980s, U.S. financial markets were much larger than financial markets outside the United States, but in recent years the dominance of U.S. markets has been disappearing. The extraordinary growth of foreign financial markets has been the result of both large increases in the pool of savings in foreign countries such as Japan and the deregulation of foreign financial markets, which has enabled them to expand their activities. American corporations and banks are now more likely to tap international capital markets to raise needed funds, and American investors often seek investment opportunities abroad. Similarly, foreign corporations and banks raise funds from Americans, and foreigners have become important investors in the United States. A look at international bond markets and world stock markets will give us a picture of how this globalization of financial markets is taking place.

International Bond Market, Eurobonds, and Eurocurrencies

The traditional instruments in the international bond market are known as **foreign bonds**. Foreign bonds are sold in a foreign country and are denominated in that country's currency. For example, if the German automaker Porsche sells a bond in the United States denominated in U.S. dollars, it is classified as a foreign bond. Foreign bonds have been an important instrument in the international capital market for centuries. In fact, a large percentage of U.S. railroads built in the nineteenth century were financed by sales of foreign bonds in Britain.

A more recent innovation in the international bond market is the **Eurobond**, a bond denominated in a currency other than that of the country in which it is sold—for example, a bond denominated in U.S. dollars sold in London. Currently, over 80 percent of the new issues in the international bond market are Eurobonds, and the market for these securities has grown very rapidly. As a result, the Eurobond market is now larger than the U.S. corporate bond market.

A variant of the Eurobond is **Eurocurrencies**, which are foreign currencies deposited in banks outside the home country. The most important of the Eurocurrencies are **Eurodollars**, which are U.S. dollars deposited in foreign banks outside the United States or in foreign branches of U.S. banks. Because these short-term deposits earn interest, they are similar to short-term Eurobonds. American banks borrow Eurodollar deposits from other banks or from their own foreign branches, and Eurodollars are now an important source of funds for American banks (over $190 billion outstanding).

Note that the new currency, the euro, can create some confusion about the terms Eurobond, Eurocurrencies, and Eurodollars. A bond denominated in euros is called a

[1] If you would like more detail about the different types of money and capital market instruments, you can find this information in an appendix to this chapter, which can be found on this book's web site at www.aw.com/mishkin.

Eurobond only *if it is sold outside the countries that have adopted the euro.* In fact, most Eurobonds are not denominated in euros but are instead denominated in U.S. dollars. Similarly, Eurodollars have nothing to do with euros, but are instead U.S. dollars deposited in banks outside the United States.

World Stock Markets

Until recently, the U.S. stock market was by far the largest in the world, but foreign stock markets have been growing in importance. Now the United States is not always number one: In the mid-1980s, the value of stocks traded in Japan at times exceeded the value of stocks traded in the United States. The increased interest in foreign stocks has prompted the development in the United States of mutual funds specializing in trading in foreign stock markets. American investors now pay attention not only to the Dow Jones Industrial Average but also to stock price indexes for foreign stock markets such as the Nikkei 225 Average (Tokyo) and the Financial Times–Stock Exchange 100-Share Index (London).

http://quote.yahoo.com/m2?u
Major world stock indices, with charts, news, and components.

The internationalization of financial markets is having profound effects on the United States. Foreigners, particularly the Japanese, are not only providing funds to corporations in the United States, but are also helping finance the federal government. Without these foreign funds, the U.S. economy would have grown far less rapidly in the last twenty years. The internationalization of financial markets is also leading the way to a more integrated world economy in which flows of goods and technology between countries are more commonplace. In later chapters, we will encounter many examples of the important roles that international factors play in our economy.

Function of Financial Intermediaries

As shown in Figure 1 (p. 24), funds can move from lenders to borrowers by a second route, called *indirect finance* because it involves a financial intermediary that stands between the lender-savers and the borrower-spenders and helps transfer funds from one to the other. A financial intermediary does this by borrowing funds from the lender-savers and then using these funds to make loans to borrower-spenders. For example, a bank might acquire funds by issuing a liability to the public (an asset for the public) in the form of savings deposits. It might then use the funds to acquire an asset by making a loan to General Motors or by buying a GM bond in the financial market. The ultimate result is that funds have been transferred from the public (the lender-savers) to GM (the borrower-spender) with the help of the financial intermediary (the bank).

The process of indirect finance using financial intermediaries, called **financial intermediation**, is the primary route for moving funds from lenders to borrowers. Indeed, although the media focus much of their attention on securities markets, particularly the stock market, financial intermediaries are a far more important source of financing for corporations than securities markets are. This is true not only for the United States but for other industrialized countries as well (see Box 1). Why are financial intermediaries and indirect finance so important in financial markets? To answer this question, we need to understand the role of transaction costs, risk sharing, and information costs in financial markets.

Transaction Costs

Transaction costs, the time and money spent in carrying out financial transactions, are a major problem for people who have excess funds to lend. As we have seen, Carl the Carpenter needs $1,000 for his new tool, and you know that it is an excellent

Following the Financial News

Foreign Stock Market Indexes

Foreign stock market indexes are published daily in the *Wall Street Journal* next to the "World Markets" column, which reports developments in foreign stock markets.

The first column identifies the country of the foreign stock exchange followed by the market index; for example, the circled entry is for the Nikkei 225 Average in Japan. The second column, "CLOSE," gives the closing value of the index, which was 8558.82 for the Nikkei 225 Average on January 20, 2003. The "NET CHG" column indicates the change in the index from the previous trading day, −131.43, and the "% CHG" column indicates the percentage change in the index, −1.51%. The "YTD NET CHG" column indicates the change in the index from the beginning of the year (year to date), −20.13, and the "YTD % CHG" column indicates the percentage change in the index from the beginning of the year, −0.23%.

International Stock Market Indexes

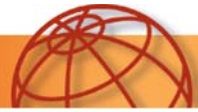
COUNTRY	INDEX	1/20/03 CLOSE	NET CHG	% CHG	YTD NET CHG	YTD % CHG
Argentina	Merval	575.74	− 1.46	− 0.25	+ 50.79	+ 9.68
Australia	All Ordinaries	3028.20	+ 3.50	+ 0.12	+ 52.70	+ 1.77
Belgium	Bel-20	1944.77	− 14.75	− 0.75	− 80.27	− 3.96
Brazil	Sao Paulo Bovespa	11648.38	− 27.32	− 0.23	+ 379.91	+ 3.37
Canada	Toronto 300 Composite	6740.37	− 15.55	− 0.23	+ 125.83	+ 1.90
Chile	Santiago IPSA	1017.96	+ 5.05	+ 0.50	+ 17.96	+ 1.80
China	Dow Jones China 88	127.54	+ 0.09	+ 0.07	+ 9.33	+ 7.89
China	Dow Jones Shanghai	181.24	+ 0.26	+ 0.14	+ 13.30	+ 7.92
China	Dow Jones Shenzhen	170.61	+ 0.68	+ 0.40	+ 13.16	+ 8.36
Europe	DJ STOXX (Euro)	198.30	− 2.63	− 1.31	− 3.42	− 1.70
Europe	DJ STOXX 50	2337.76	− 44.76	− 1.88	− 69.75	− 2.90
Euro Zone	DJ Euro STOXX	205.29	− 2.39	− 1.15	− 0.65	− 0.32
Euro Zone	DJ Euro STOXX 50	2352.81	− 37.55	− 1.57	− 33.60	− 1.41
France	Paris CAC 40	3020.07	− 36.86	− 1.21	− 43.84	− 1.43
Germany	Frankfurt Xetra DAX	2893.55	− 25.27	− 0.87	+ 0.92	+ 0.03
Hong Kong	Hang Seng	9552.02	− 62.57	− 0.65	+ 230.73	+ 2.48
India	Bombay Sensex	3341.89	− 28.50	− 0.85	− 35.39	− 1.05
Israel	Tel Aviv 25	311.62	− 1.87	− 0.60	− 22.29	− 6.68
Italy	Milan MIBtel	17339.00	− 199.00	− 1.13	− 146.00	− 0.84
Japan	Tokyo Nikkei 225	8558.82	− 131.43	− 1.51	− 20.13	− 0.23
Japan	Tokyo Nikkei 300	166.81	− 1.58	− 0.94	+ 1.36	+ 0.82
Japan	Tokyo Topix Index	853.90	− 5.35	− 0.62	+ 10.61	+ 1.26
Mexico	I.P.C. All-Share	6161.12	− 43.34	− 0.70	+ 34.03	+ 0.56
Netherlands	Amsterdam AEX	313.04	− 5.55	− 1.74	− 9.69	− 3.00
Singapore	Straits Times	1363.19	− 3.64	− 0.27	+ 22.16	+ 1.65
South Africa	Johannesburg All Share	9485.48	− 2.94	− 0.03	+ 208.26	+ 2.24
South Korea	KOSPI	634.50	− 1.96	− 0.31	+ 6.95	+ 1.11
Spain	IBEX 35	6390.80	− 67.40	− 1.04	+ 353.90	+ 5.86
Sweden	SX All Share	155.40	+ 1.56	+ 1.01	+ 5.83	+ 3.90
Switzerland	Zurich Swiss Market	4679.70	− 73.90	− 1.55	+ 48.90	+ 1.06
Taiwan	Weighted	4951.03	+ 43.25	+ 0.88	+ 498.58	+ 11.20
U.K.	London FTSE 100-share	3778.60	− 42.00	− 1.10	− 161.80	− 4.11
U.K.	London FTSE 250-share	4312.50	− 9.20	− 0.21	− 6.80	− 0.16

Source: Wall Street Journal, Tuesday, January 21, 2003, p. C6.

investment opportunity. You have the cash and would like to lend him the money, but to protect your investment, you have to hire a lawyer to write up the loan contract that specifies how much interest Carl will pay you, when he will make these interest payments, and when he will repay you the $1,000. Obtaining the contract will cost you $500. When you figure in this transaction cost for making the loan, you realize that you can't earn enough from the deal (you spend $500 to make perhaps $100) and reluctantly tell Carl that he will have to look elsewhere.

This example illustrates that small savers like you or potential borrowers like Carl might be frozen out of financial markets and thus be unable to benefit from them. Can anyone come to the rescue? Financial intermediaries can.

Financial intermediaries can substantially reduce transaction costs because they have developed expertise in lowering them; because their large size allows them to take advantage of **economies of scale**, the reduction in transaction costs per dollar of transactions as the size (scale) of transactions increases. For example, a bank knows



Box 1: Global

The Importance of Financial Intermediaries to Securities Markets: An International Comparison

Patterns of financing corporations differ across countries, but one key fact emerges. Studies of the major developed countries, including the United States, Canada, Great Britain, Japan, Italy, Germany, and France, show that when businesses go looking for funds to finance their activities, they usually obtain them indirectly through financial intermediaries and not directly from securities markets.* Even in the United States and Canada, which have the most developed securities markets in the world, loans from financial intermediaries are far more important for corporate finance than securities markets are. The countries that have made the least use of securities markets are Germany and Japan; in these two countries, financing from financial intermediaries has been almost ten times greater than that from securities markets. However, with the deregulation of Japanese securities markets in recent years, the share of corporate financing by financial intermediaries has been declining relative to the use of securities markets.

Although the dominance of financial intermediaries over securities markets is clear in all countries, the relative importance of bond versus stock markets differs widely across countries. In the United States, the bond market is far more important as a source of corporate finance: On average, the amount of new financing raised using bonds is ten times the amount using stocks. By contrast, countries such as France and Italy make use of equities markets more than the bond market to raise capital.

*See, for example, Colin Mayer, "Financial Systems, Corporate Finance, and Economic Development," in *Asymmetric Information, Corporate Finance, and Investment*, ed. R. Glenn Hubbard (Chicago: University of Chicago Press, 1990), pp. 307–332.

how to find a good lawyer to produce an airtight loan contract, and this contract can be used over and over again in its loan transactions, thus lowering the legal cost per transaction. Instead of a loan contract (which may not be all that well written) costing $500, a bank can hire a topflight lawyer for $5,000 to draw up an airtight loan contract that can be used for 2,000 loans at a cost of $2.50 per loan. At a cost of $2.50 per loan, it now becomes profitable for the financial intermediary to lend Carl the $1,000.

Because financial intermediaries are able to reduce transaction costs substantially, they make it possible for you to provide funds indirectly to people like Carl with productive investment opportunities. In addition, a financial intermediary's low transaction costs mean that it can provide its customers with **liquidity services**, services that make it easier for customers to conduct transactions. For example, banks provide depositors with checking accounts that enable them to pay their bills easily. In addition, depositors can earn interest on checking and savings accounts and yet still convert them into goods and services whenever necessary.

Risk Sharing

Another benefit made possible by the low transaction costs of financial institutions is that they can help reduce the exposure of investors to **risk**; that is, uncertainty about the returns investors will earn on assets. Financial intermediaries do this through the process known as **risk sharing**: they create and sell assets with risk characteristics that people are comfortable with, and the intermediaries then use the funds they acquire by selling these assets to purchase other assets that may have far more risk.

Low transaction costs allow financial intermediaries to do risk sharing at low cost, enabling them to earn a profit on the spread between the returns they earn on risky assets and the payments they make on the assets they have sold. This process of risk sharing is also sometimes referred to as **asset transformation**, because in a sense, risky assets are turned into safer assets for investors.

Financial intermediaries also promote risk sharing by helping individuals to diversify and thereby lower the amount of risk to which they are exposed. **Diversification** entails investing in a collection (**portfolio**) of assets whose returns do not always move together, with the result that overall risk is lower than for individual assets. (Diversification is just another name for the old adage that "you shouldn't put all your eggs in one basket.") Low transaction costs allow financial intermediaries to do this by pooling a collection of assets into a new asset and then selling it to individuals.

Asymmetric Information: Adverse Selection and Moral Hazard

The presence of transaction costs in financial markets explains, in part, why financial intermediaries and indirect finance play such an important role in financial markets. An additional reason is that in financial markets, one party often does not know enough about the other party to make accurate decisions. This inequality is called **asymmetric information**. For example, a borrower who takes out a loan usually has better information about the potential returns and risk associated with the investment projects for which the funds are earmarked than the lender does. Lack of information creates problems in the financial system on two fronts: before the transaction is entered into and after.[2]

Adverse selection is the problem created by asymmetric information *before* the transaction occurs. Adverse selection in financial markets occurs when the potential borrowers who are the most likely to produce an undesirable (*adverse*) outcome—the bad credit risks—are the ones who most actively seek out a loan and are thus most likely to be selected. Because adverse selection makes it more likely that loans might be made to bad credit risks, lenders may decide not to make any loans even though there are good credit risks in the marketplace.

To understand why adverse selection occurs, suppose that you have two aunts to whom you might make a loan—Aunt Louise and Aunt Sheila. Aunt Louise is a conservative type who borrows only when she has an investment she is quite sure will pay off. Aunt Sheila, by contrast, is an inveterate gambler who has just come across a get-rich-quick scheme that will make her a millionaire if she can just borrow $1,000 to invest in it. Unfortunately, as with most get-rich-quick schemes, there is a high probability that the investment won't pay off and that Aunt Sheila will lose the $1,000.

Which of your aunts is more likely to call you to ask for a loan? Aunt Sheila, of course, because she has so much to gain if the investment pays off. You, however, would not want to make a loan to her because there is a high probability that her investment will turn sour and she will be unable to pay you back.

If you knew both your aunts very well—that is, if your information were not asymmetric—you wouldn't have a problem, because you would know that Aunt Sheila is a bad risk and so you would not lend to her. Suppose, though, that you don't

[2]Asymmetric information and the adverse selection and moral hazard concepts are also crucial problems for the insurance industry (see Chapter 12).

know your aunts well. You are more likely to lend to Aunt Sheila than to Aunt Louise because Aunt Sheila would be hounding you for the loan. Because of the possibility of adverse selection, you might decide not to lend to either of your aunts, even though there are times when Aunt Louise, who is an excellent credit risk, might need a loan for a worthwhile investment.

Moral hazard is the problem created by asymmetric information *after* the transaction occurs. Moral hazard in financial markets is the risk (*hazard*) that the borrower might engage in activities that are undesirable (*immoral*) from the lender's point of view, because they make it less likely that the loan will be paid back. Because moral hazard lowers the probability that the loan will be repaid, lenders may decide that they would rather not make a loan.

As an example of moral hazard, suppose that you made a $1,000 loan to another relative, Uncle Melvin, who needs the money to purchase a word processor so he can set up a business typing students' term papers. Once you have made the loan, however, Uncle Melvin is more likely to slip off to the track and play the horses. If he bets on a 20-to-1 long shot and wins with your money, he is able to pay you back your $1,000 and live high off the hog with the remaining $19,000. But if he loses, as is likely, you don't get paid back, and all he has lost is his reputation as a reliable, upstanding uncle. Uncle Melvin therefore has an incentive to go to the track because his gains ($19,000) if he bets correctly are much greater than the cost to him (his reputation) if he bets incorrectly. If you knew what Uncle Melvin was up to, you would prevent him from going to the track, and he would not be able to increase the moral hazard. However, because it is hard for you to keep informed about his whereabouts—that is, because information is asymmetric—there is a good chance that Uncle Melvin will go to the track and you will not get paid back. The risk of moral hazard might therefore discourage you from making the $1,000 loan to Uncle Melvin, even if you were sure that you would be paid back if he used it to set up his business.

Study Guide

Because the concepts of adverse selection and moral hazard are extremely useful in understanding the behavior we examine in this and many of the later chapters (and in life in general), you must understand them fully. One way to distinguish between them is to remember that adverse selection is a problem of asymmetric information *before* entering into a transaction, whereas moral hazard is a problem of asymmetric information *after* the transaction has occurred. A helpful way to nail down these concepts is to think of other examples, for financial or other types of transactions, in which adverse selection or moral hazard plays a role. Several problems at the end of the chapter provide additional examples of situations involving adverse selection and moral hazard.

The problems created by adverse selection and moral hazard are an important impediment to well-functioning financial markets. Again, financial intermediaries can alleviate these problems.

With financial intermediaries in the economy, small savers can provide their funds to the financial markets by lending these funds to a trustworthy intermediary—say, the Honest John Bank—which in turn lends the funds out either by making loans or by buying securities such as stocks or bonds. Successful financial intermediaries have higher earnings on their investments than small savers, because they are better

equipped than individuals to screen out bad credit risks from good ones, thereby reducing losses due to adverse selection. In addition, financial intermediaries have high earnings because they develop expertise in monitoring the parties they lend to, thus reducing losses due to moral hazard. The result is that financial intermediaries can afford to pay lender-savers interest or provide substantial services and still earn a profit.

As we have seen, financial intermediaries play an important role in the economy because they provide liquidity services, promote risk sharing, and solve information problems. The success of financial intermediaries in performing this role is evidenced by the fact that most Americans invest their savings with them and obtain loans from them. Financial intermediaries play a key role in improving economic efficiency because they help financial markets channel funds from lender-savers to people with productive investment opportunities. Without a well-functioning set of financial intermediaries, it is very hard for an economy to reach its full potential. We will explore further the role of financial intermediaries in the economy in Part III.

Financial Intermediaries

We have seen why financial intermediaries play such an important role in the economy. Now we look at the principal financial intermediaries themselves and how they perform the intermediation function. They fall into three categories: depository institutions (banks), contractual savings institutions, and investment intermediaries. Table 1 provides a guide to the discussion of the financial intermediaries that fit into these three categories by describing their primary liabilities (sources of funds) and assets (uses of funds). The relative size of these intermediaries in the United States is indicated in Table 2, which lists the amount of their assets at the end of 1970, 1980, 1990, and 2002.

Depository Institutions

Depository institutions (for simplicity, we refer to these as *banks* throughout this text) are financial intermediaries that accept deposits from individuals and institutions and make loans. The study of money and banking focuses special attention on this group of financial institutions, because they are involved in the creation of deposits, an important component of the money supply. These institutions include commercial banks and the so-called **thrift institutions (thrifts)**: savings and loan associations, mutual savings banks, and credit unions.

Commercial Banks. These financial intermediaries raise funds primarily by issuing checkable deposits (deposits on which checks can be written), savings deposits (deposits that are payable on demand but do not allow their owner to write checks), and time deposits (deposits with fixed terms to maturity). They then use these funds to make commercial, consumer, and mortgage loans and to buy U.S. government securities and municipal bonds. There are slightly fewer than 8,000 commercial banks in the United States, and as a group, they are the largest financial intermediary and have the most diversified portfolios (collections) of assets.

Savings and Loan Associations (S&Ls) and Mutual Savings Banks. These depository institutions, of which there are approximately 1,500, obtain funds primarily through savings deposits (often called *shares*) and time and checkable deposits. In the past, these insti-

Table 1 Primary Assets and Liabilities of Financial Intermediaries

Type of Intermediary	Primary Liabilities (Sources of Funds)	Primary Assets (Uses of Funds)
Depository institutions (banks)		
Commercial banks	Deposits	Business and consumer loans, mortgages, U.S. government securities and municipal bonds
Savings and loan associations	Deposits	Mortgages
Mutual savings banks	Deposits	Mortgages
Credit unions	Deposits	Consumer loans
Contractual savings institutions		
Life insurance companies	Premiums from policies	Corporate bonds and mortgages
Fire and casualty insurance companies	Premiums from policies	Municipal bonds, corporate bonds and stock, U.S. government securities
Pension funds, government retirement funds	Employer and employee contributions	Corporate bonds and stock
Investment intermediaries		
Finance companies	Commercial paper, stocks, bonds	Consumer and business loans
Mutual funds	Shares	Stocks, bonds
Money market mutual funds	Shares	Money market instruments

tutions were constrained in their activities and mostly made mortgage loans for residential housing. Over time, these restrictions have been loosened so that the distinction between these depository institutions and commercial banks has blurred. These intermediaries have become more alike and are now more competitive with each other.

Credit Unions. These financial institutions, numbering about 9,500, are very small cooperative lending institutions organized around a particular group: union members, employees of a particular firm, and so forth. They acquire funds from deposits called *shares* and primarily make consumer loans.

Contractual Savings Institutions

Contractual savings institutions, such as insurance companies and pension funds, are financial intermediaries that acquire funds at periodic intervals on a contractual basis. Because they can predict with reasonable accuracy how much they will have to pay

Table 2 Principal Financial Intermediaries and Value of Their Assets

Type of Intermediary	Value of Assets ($ billions, end of year) 1970	1980	1990	2002
Depository institutions (banks)				
Commercial banks	517	1,481	3,334	7,161
Savings and loan associations and mutual savings banks	250	792	1,365	1,338
Credit unions	18	67	215	553
Contractual savings institutions				
Life insurance companies	201	464	1,367	3,269
Fire and casualty insurance companies	50	182	533	894
Pension funds (private)	112	504	1,629	3,531
State and local government retirement funds	60	197	737	1,895
Investment intermediaries				
Finance companies	64	205	610	1,165
Mutual funds	47	70	654	3,419
Money market mutual funds	0	76	498	2,106

Source: Federal Reserve Flow of Funds Accounts: www.federalreserve.gov/releases/Z1/LevelTables.

out in benefits in the coming years, they do not have to worry as much as depository institutions about losing funds. As a result, the liquidity of assets is not as important a consideration for them as it is for depository institutions, and they tend to invest their funds primarily in long-term securities such as corporate bonds, stocks, and mortgages.

Life Insurance Companies. Life insurance companies insure people against financial hazards following a death and sell annuities (annual income payments upon retirement). They acquire funds from the premiums that people pay to keep their policies in force and use them mainly to buy corporate bonds and mortgages. They also purchase stocks, but are restricted in the amount that they can hold. Currently, with $3.3 trillion in assets, they are among the largest of the contractual savings institutions.

Fire and Casualty Insurance Companies. These companies insure their policyholders against loss from theft, fire, and accidents. They are very much like life insurance companies, receiving funds through premiums for their policies, but they have a greater possibility of loss of funds if major disasters occur. For this reason, they use their funds to buy more liquid assets than life insurance companies do. Their largest holding of assets is municipal bonds; they also hold corporate bonds and stocks and U.S. government securities.

Pension Funds and Government Retirement Funds. Private pension funds and state and local retirement funds provide retirement income in the form of annuities to employees who are covered by a pension plan. Funds are acquired by contributions from employers or from employees, who either have a contribution automatically deducted from their paychecks or contribute voluntarily. The largest asset holdings of pension funds are corporate bonds and stocks. The establishment of pension funds has been actively encouraged by the federal government, both through legislation requiring pension plans and through tax incentives to encourage contributions.

Investment Intermediaries

This category of financial intermediaries includes finance companies, mutual funds, and money market mutual funds.

Finance Companies. Finance companies raise funds by selling commercial paper (a short-term debt instrument) and by issuing stocks and bonds. They lend these funds to consumers, who make purchases of such items as furniture, automobiles, and home improvements, and to small businesses. Some finance companies are organized by a parent corporation to help sell its product. For example, Ford Motor Credit Company makes loans to consumers who purchase Ford automobiles.

Mutual Funds. These financial intermediaries acquire funds by selling shares to many individuals and use the proceeds to purchase diversified portfolios of stocks and bonds. Mutual funds allow shareholders to pool their resources so that they can take advantage of lower transaction costs when buying large blocks of stocks or bonds. In addition, mutual funds allow shareholders to hold more diversified portfolios than they otherwise would. Shareholders can sell (redeem) shares at any time, but the value of these shares will be determined by the value of the mutual fund's holdings of securities. Because these fluctuate greatly, the value of mutual fund shares will too; therefore, investments in mutual funds can be risky.

Money Market Mutual Funds. These relatively new financial institutions have the characteristics of a mutual fund but also function to some extent as a depository institution because they offer deposit-type accounts. Like most mutual funds, they sell shares to acquire funds that are then used to buy money market instruments that are both safe and very liquid. The interest on these assets is then paid out to the shareholders.

A key feature of these funds is that shareholders can write checks against the value of their shareholdings. In effect, shares in a money market mutual fund function like checking account deposits that pay interest. Money market mutual funds have experienced extraordinary growth since 1971, when they first appeared. By 2002, their assets had climbed to nearly $1 trillion.

Regulation of the Financial System

The financial system is among the most heavily regulated sectors of the American economy. The government regulates financial markets for two main reasons: to increase the information available to investors and to ensure the soundness of the financial system. We will examine how these two reasons have led to the present regulatory environment. As a study aid, the principal regulatory agencies of the U.S. financial system are listed in Table 3.

Table 3 Principal Regulatory Agencies of the U.S. Financial System

Regulatory Agency	Subject of Regulation	Nature of Regulations
Securities and Exchange Commission (SEC)	Organized exchanges and financial markets	Requires disclosure of information, restricts insider trading
Commodities Futures Trading Commission (CFTC)	Futures market exchanges	Regulates procedures for trading in futures markets
Office of the Comptroller of the Currency	Federally chartered commercial banks	Charters and examines the books of federally chartered commercial banks and imposes restrictions on assets they can hold
National Credit Union Administration (NCUA)	Federally chartered credit unions	Charters and examines the books of federally chartered credit unions and imposes restrictions on assets they can hold
State banking and insurance commissions	State-chartered depository institutions	Charters and examines the books of state-chartered banks and insurance companies, imposes restrictions on assets they can hold, and imposes restrictions on branching
Federal Deposit Insurance Corporation (FDIC)	Commercial banks, mutual savings banks, savings and loan associations	Provides insurance of up to $100,000 for each depositor at a bank, examines the books of insured banks, and imposes restrictions on assets they can hold
Federal Reserve System	All depository institutions	Examines the books of commercial banks that are members of the system, sets reserve requirements for all banks
Office of Thrift Supervision	Savings and loan associations	Examines the books of savings and loan associations, imposes restrictions on assets they can hold

Increasing Information Available to Investors

Asymmetric information in financial markets means that investors may be subject to adverse selection and moral hazard problems that may hinder the efficient operation of financial markets. Risky firms or outright crooks may be the most eager to sell securities to unwary investors, and the resulting adverse selection problem may keep investors out of financial markets. Furthermore, once an investor has bought a security, thereby lending money to a firm, the borrower may have incentives to engage in risky activities or to commit outright fraud. The presence of this moral hazard problem may also keep investors away from financial markets. Government regulation can reduce adverse selection and moral hazard problems in financial markets and increase their efficiency by increasing the amount of information available to investors.

As a result of the stock market crash in 1929 and revelations of widespread fraud in the aftermath, political demands for regulation culminated in the Securities Act of 1933 and the establishment of the Securities and Exchange Commission (SEC). The SEC requires corporations issuing securities to disclose certain information about their sales, assets, and earnings to the public and restricts trading by the largest stockholders (known as *insiders*) in the corporation. By requiring disclosure of this information and by discouraging insider trading, which could be used to manipulate security prices, the SEC hopes that investors will be better informed and be protected from some of the abuses in financial markets that occurred before 1933. Indeed, in recent years, the SEC has been particularly active in prosecuting people involved in insider trading.

www.sec.gov

The United States Securities and Exchange Commission home page. It contains vast SEC resources, laws and regulations, investor information, and litigation.

Ensuring the Soundness of Financial Intermediaries

Asymmetric information can also lead to widespread collapse of financial intermediaries, referred to as a **financial panic**. Because providers of funds to financial intermediaries may not be able to assess whether the institutions holding their funds are sound, if they have doubts about the overall health of financial intermediaries, they may want to pull their funds out of both sound and unsound institutions. The possible outcome is a financial panic that produces large losses for the public and causes serious damage to the economy. To protect the public and the economy from financial panics, the government has implemented six types of regulations.

Restrictions on Entry. State banking and insurance commissions, as well as the Office of the Comptroller of the Currency (an agency of the federal government), have created very tight regulations governing who is allowed to set up a financial intermediary. Individuals or groups that want to establish a financial intermediary, such as a bank or an insurance company, must obtain a charter from the state or the federal government. Only if they are upstanding citizens with impeccable credentials and a large amount of initial funds will they be given a charter.

Disclosure. There are stringent reporting requirements for financial intermediaries. Their bookkeeping must follow certain strict principles, their books are subject to periodic inspection, and they must make certain information available to the public.

Restrictions on Assets and Activities. There are restrictions on what financial intermediaries are allowed to do and what assets they can hold. Before you put your funds into a bank or some other such institution, you would want to know that your funds are safe and that the bank or other financial intermediary will be able to meet its obligations to you. One way of doing this is to restrict the financial intermediary from engaging in certain risky activities. Legislation passed in 1933 (repealed in 1999) separated commercial banking from the securities industry so that banks could not engage in risky ventures associated with this industry. Another way is to restrict financial

intermediaries from holding certain risky assets, or at least from holding a greater quantity of these risky assets than is prudent. For example, commercial banks and other depository institutions are not allowed to hold common stock because stock prices experience substantial fluctuations. Insurance companies are allowed to hold common stock, but their holdings cannot exceed a certain fraction of their total assets.

Deposit Insurance. The government can insure people's deposits so that they do not suffer any financial loss if the financial intermediary that holds these deposits should fail. The most important government agency that provides this type of insurance is the Federal Deposit Insurance Corporation (FDIC), which insures each depositor at a commercial bank or mutual savings bank up to a loss of $100,000 per account. All commercial and mutual savings banks, with a few minor exceptions, are contributers to the FDIC's Bank Insurance Fund, which is used to pay off depositors in the case of a bank's failure. The FDIC was created in 1934 after the massive bank failures of 1930–1933, in which the savings of many depositors at commercial banks were wiped out. Similar government agencies exist for other depository institutions: The Savings Association Insurance Fund (part of the FDIC) provides deposit insurance for savings and loan associations, and the National Credit Union Share Insurance Fund (NCUSIF) does the same for credit unions.

Limits on Competition. Politicians have often declared that unbridled competition among financial intermediaries promotes failures that will harm the public. Although the evidence that competition does this is extremely weak, it has not stopped the state and federal governments from imposing many restrictive regulations. First are the restrictions on the opening of additional locations (branches). In the past, banks were not allowed to open up branches in other states, and in some states, banks were restricted from opening additional locations.

Restrictions on Interest Rates. Competition has also been inhibited by regulations that impose restrictions on interest rates that can be paid on deposits. For decades after 1933, banks were prohibited from paying interest on checking accounts. In addition, until 1986, the Federal Reserve System had the power under *Regulation Q* to set maximum interest rates that banks could pay on savings deposits. These regulations were instituted because of the widespread belief that unrestricted interest-rate competition helped encourage bank failures during the Great Depression. Later evidence does not seem to support this view, and restrictions like Regulation Q have been abolished.

In later chapters, we will look more closely at government regulation of financial markets and will see whether it has improved the functioning of financial markets.

Financial Regulation Abroad

Not surprisingly, given the similarity of the economic system here and in Japan, Canada, and the nations of Western Europe, financial regulation in these countries is similar to financial regulation in the United States. The provision of information is improved by requiring corporations issuing securities to report details about assets and liabilities, earnings, and sales of stock, and by prohibiting insider trading. The soundness of intermediaries is ensured by licensing, periodic inspection of financial intermediaries' books, and the provision of deposit insurance (although its coverage is smaller than in the United States and its existence is often intentionally not advertised).

The major differences between financial regulation in the United States and abroad relate to bank regulation. In the past, the United States was the only industrialized country to subject banks to restrictions on branching, which limited banks' size

and restricted them to certain geographic regions. (These restrictions were abolished by legislation in 1994.) U.S. banks are also the most restricted in the range of assets they may hold. Banks abroad frequently hold shares in commercial firms; in Japan and Germany, those stakes can be sizable.

Summary

1. The basic function of financial markets is to channel funds from savers who have an excess of funds to spenders who have a shortage of funds. Financial markets can do this either through direct finance, in which borrowers borrow funds directly from lenders by selling them securities, or through indirect finance, which involves a financial intermediary that stands between the lender-savers and the borrower-spenders and helps transfer funds from one to the other. This channeling of funds improves the economic welfare of everyone in the society, because it allows funds to move from people who have no productive investment opportunities to those who have such opportunities, thereby contributing to increased efficiency in the economy. In addition, channeling of funds directly benefits consumers by allowing them to make purchases when they need them most.
2. Financial markets can be classified as debt and equity markets, primary and secondary markets, exchanges and over-the-counter markets, and money and capital markets.
3. An important trend in recent years is the growing internationalization of financial markets. Eurobonds, which are denominated in a currency other than that of the country in which they are sold, are now the dominant security in the international bond market and have surpassed U.S. corporate bonds as a source of new funds. Eurodollars, which are U.S. dollars deposited in foreign banks, are an important source of funds for American banks.
4. Financial intermediaries are financial institutions that acquire funds by issuing liabilities and in turn use those funds to acquire assets by purchasing securities or making loans. Financial intermediaries play an important role in the financial system, because they reduce transaction costs, allow risk sharing, and solve problems created by adverse selection and moral hazard. As a result, financial intermediaries allow small savers and borrowers to benefit from the existence of financial markets, thereby increasing the efficiency of the economy.
5. The principal financial intermediaries fall into three categories: (a) banks—commercial banks, savings and loan associations, mutual savings banks, and credit unions; (b) contractual savings institutions—life insurance companies, fire and casualty insurance companies, and pension funds; and (c) investment intermediaries—finance companies, mutual funds, and money market mutual funds.
6. The government regulates financial markets and financial intermediaries for two main reasons: to increase the information available to investors and to ensure the soundness of the financial system. Regulations include requiring disclosure of information to the public, restrictions on who can set up a financial intermediary, restrictions on what assets financial intermediaries can hold, the provision of deposit insurance, reserve requirements, and the setting of maximum interest rates that can be paid on checking accounts and savings deposits.

Key Terms

asset transformation, p. 32
adverse selection, p. 32
asymmetric information, p. 32
brokers, p. 26
capital market, p. 27
dealers, p. 26
diversification, p. 32
dividends, p. 26
economies of scale, p. 30

equities, p. 26
Eurobond, p. 28
Eurocurrencies, p. 28
Eurodollars, p. 28
exchanges, p. 27
financial intermediation, p. 29
financial panic, p. 39
foreign bonds, p. 28
intermediate-term, p. 26
investment bank, p. 26
liabilities, p. 24
liquid, p. 27
liquidity services, p. 31
long-term, p. 26
maturity, p. 26
money market, p. 27
moral hazard, p. 33
over-the-counter (OTC) market, p. 27
portfolio, p. 32
primary market, p. 26
risk, p. 31
risk sharing, p. 31
secondary market, p. 26
short-term, p. 26
thrift institutions (thrifts), p. 34
transaction costs, p. 29
underwriting, p. 26

Questions and Problems

Questions marked with an asterisk are answered at the end of the book in an appendix, "Answers to Selected Questions and Problems."

*1. Why is a share of IBM common stock an asset for its owner and a liability for IBM?

2. If I can buy a car today for $5,000 and it is worth $10,000 in extra income next year to me because it enables me to get a job as a traveling anvil seller, should I take out a loan from Larry the Loan Shark at a 90% interest rate if no one else will give me a loan? Will I be better or worse off as a result of taking out this loan? Can you make a case for legalizing loan-sharking?

*3. Some economists suspect that one of the reasons that economies in developing countries grow so slowly is that they do not have well-developed financial markets. Does this argument make sense?

4. The U.S. economy borrowed heavily from the British in the nineteenth century to build a railroad system. What was the principal debt instrument used? Why did this make both countries better off?

*5. "Because corporations do not actually raise any funds in secondary markets, they are less important to the economy than primary markets." Comment.

6. If you suspect that a company will go bankrupt next year, which would you rather hold, bonds issued by the company or equities issued by the company? Why?

*7. How can the adverse selection problem explain why you are more likely to make a loan to a family member than to a stranger?

8. Think of one example in which you have had to deal with the adverse selection problem.

*9. Why do loan sharks worry less about moral hazard in connection with their borrowers than some other lenders do?

10. If you are an employer, what kinds of moral hazard problems might you worry about with your employees?

*11. If there were no asymmetry in the information that a borrower and a lender had, could there still be a moral hazard problem?

12. "In a world without information and transaction costs, financial intermediaries would not exist." Is this statement true, false, or uncertain? Explain your answer.

*13. Why might you be willing to make a loan to your neighbor by putting funds in a savings account earning a 5% interest rate at the bank and having the bank lend her the funds at a 10% interest rate rather than lend her the funds yourself?

14. How does risk sharing benefit both financial intermediaries and private investors?

*15. Discuss some of the manifestations of the globalization of world capital markets.

Web Exercises

1. One of the single best sources of information about financial institutions is the U.S. Flow of Funds report produced by the Federal Reserve. This document contains data on most financial intermediaries. Go to www.federalreserve.gov/releases/Z1/. Go to the most current release. You may have to load Acrobat Reader if your computer does not already have it. The site has a link for a free patch. Go to the Level Tables and answer the following.
   a. What percent of assets do commercial banks hold in loans? What percent of assets are held in mortgage loans?
   b. What percent of assets do Savings and Loans hold in mortgage loans?
   c. What percent of assets do credit unions hold in mortgage loans and in consumer loans?

2. The most famous financial market in the world is the New York Stock Exchange. Go to www.nyse.com.
   a. What is the mission of the NYSE?
   b. Firms must pay a fee to list their shares for sale on the NYSE. What would be the fee for a firm with 5 million shares common outstanding?

appendix to chapter 2

Financial Market Instruments

Here we examine the securities (instruments) traded in financial markets. We first focus on the instruments traded in the money market and then turn to those traded in the capital market.

Money Market Instruments

Because of their short terms to maturity, the debt instruments traded in the money market undergo the least price fluctuations and so are the least risky investments. The money market has undergone great changes in the past three decades, with the amount of some financial instruments growing at a far more rapid rate than others.

The principal money market instruments are listed in Table 1 along with the amount outstanding at the end of 1970, 1980, 1990, and 2002.

United States Treasury Bills. These short-term debt instruments of the U.S. government are issued in 3-, 6-, and 12-month maturities to finance the federal government. They pay a set amount at maturity and have no interest payments, but they effectively pay interest by initially selling at a discount, that is, at a price lower than the set amount paid at maturity. For instance, you might pay $9,000 in May 2004 for a one-year Treasury Bill that can be redeemed in May 2005 for $10,000.

U.S. Treasury bills are the most liquid of all the money market instruments, because they are the most actively traded. They are also the safest of all money market instruments, because there is almost no possibility of *default*, a situation in which the party issuing the debt instrument (the federal government, in this case) is unable to make interest payments or pay off the amount owed when the instrument matures. The federal government is always able to meet its debt obligations, because it can raise taxes or issue *currency* (paper money or coins) to pay off its debts. Treasury bills are held mainly by banks, although small amounts are held by households, corporations, and other financial intermediaries.

Negotiable Bank Certificates of Deposit. A *certificate of deposit* (*CD*) is a debt instrument, sold by a bank to depositors, that pays annual interest of a given amount and at maturity, pays back the original purchase price. Before 1961, CDs were nonnegotiable; that is, they could not be sold to someone else and could not be redeemed from the bank before maturity without paying a substantial penalty. In 1961, to make CDs more liquid and more attractive to investors, Citibank introduced the first negotiable CD in large denominations (over $100,000) that could be resold in a secondary market. This instrument is now issued by almost all the major commercial banks and has been extremely successful, with the amount outstanding currently around $1.2 trillion. CDs

Table 1 Principal Money Market Instruments

Type of Instrument	Amount Outstanding ($ billions, end of year) 1970	1980	1990	2002
U.S. Treasury bills	81	216	527	888
Negotiable bank certificates of deposit (large denominations)	55	317	543	1,177
Commercial paper	33	122	557	1,321
Banker's acceptances	7	42	52	9
Repurchase agreements	3	57	144	470
Federal funds*	16	18	61	29
Eurodollars	2	55	92	213

*Figures after 1970 are for large banks only.

Sources: Federal Reserve Flow of Funds Accounts; Federal Reserve *Bulletin*; *Banking and Monetary Statistics, 1945–1970*; *Annual Statistical Digest, 1971–1975*; *Economic Report of the President.* www.federalreserve.gov/releases/z1

are an extremely important source of funds for commercial banks, from corporations, money market mutual funds, charitable institutions, and government agencies.

Commercial Paper. *Commercial paper* is a short-term debt instrument issued by large banks and well-known corporations, such as General Motors and AT&T. Before the 1960s, corporations usually borrowed their short-term funds from banks, but since then they have come to rely more heavily on selling commercial paper to other financial intermediaries and corporations for their immediate borrowing needs; in other words, they engage in direct finance. Growth of the commercial paper market has been substantial: The amount of commercial paper outstanding has increased by over 3,900% (from $33 billion to $1.3 trillion) in the period 1970–2002. We discuss why the commercial paper market has had such tremendous growth in Chapter 10.

Banker's Acceptances. These money market instruments are created in the course of carrying out international trade and have been in use for hundreds of years. A *banker's acceptance* is a bank draft (a promise of payment similar to a check) issued by a firm, payable at some future date, and guaranteed for a fee by the bank that stamps it "accepted." The firm issuing the instrument is required to deposit the required funds into its account to cover the draft. If the firm fails to do so, the bank's guarantee means that it is obligated to make good on the draft. The advantage to the firm is that the draft is more likely to be accepted when purchasing goods abroad, because the foreign exporter knows that even if the company purchasing the goods goes bankrupt, the bank draft will still be paid off. These "accepted" drafts are often resold in a secondary market at a discount and are therefore similar in function to Treasury bills. Typically, they are held by many of the same parties that hold Treasury bills, and the amount outstanding has experienced limited growth, rising by 28% ($7 billion to $9 billion) from 1970 to 2002.

Following the Financial News

Money Market Rates

The *Wall Street Journal* publishes daily a listing of interest rates on many different financial instruments in its "Money Rates" column. (See "Today's Contents" on page 1 of the *Journal* for the location.)

The four interest rates in the "Money Rates" column that are discussed most frequently in the media are these:

- *Prime rate:* The base interest rate on corporate bank loans, an indicator of the cost of business borrowing from banks
- *Federal funds rate:* The interest rate charged on overnight loans in the federal funds market, a sensitive indicator of the cost to banks of borrowing funds from other banks and the stance of monetary policy
- *Treasury bill rate:* The interest rate on U.S. Treasury bills, an indicator of general interest-rate movements
- *Federal Home Loan Mortgage Corporation rates:* Interest rates on "Freddie Mac"–guaranteed mortgages, an indicator of the cost of financing residential housing purchases

MONEY RATES

Wednesday, June 3, 2003

The key U.S., and foreign annual interest rates below are a guide to general levels but don't always represent actual transactions.

PRIME RATE: 4.25% (effective 11/07/02).

DISCOUNT RATE: 2.25% (effective 01/09/03).

FEDERAL FUNDS: 1.250% high, 1.000% low, 1.125% near closing bid, 1.188% offered. Effective rate: 1.22%. Source: Prebon Yamane (USA) Inc. Federal-funds target rate: 1.250% (effective 11/06/02).

CALL MONEY: 3.00% (effective 11/07/02).

COMMERCIAL PAPER: Placed directly by General Electric Capital Corp.: 1.05% 30 to 35 days; 1.24% 36 to 43 days; 1.23% 44 to 70 days; 1.21% 71 to 99 days; 1.19% 100 to 113 days; 1.05% 114 to 122 days; 1.19% 123 to 143 days; 1.17% 144 to 270 days.

EURO COMMERCIAL PAPER: Placed directly by General Electric Capital Corp.: 2.25% 30 days; 2.20% two months; 2.19% three months; 2.15% four months; 2.14% five months; 2.13% six months.

DEALER COMMERCIAL PAPER: High-grade unsecured notes sold through dealers by major corporations: 1.21% 30 days; 1.20% 60 days; 1.19% 90 days.

CERTIFICATES OF DEPOSIT: 1.26% one month; 1.21% three months; 1.18% six months.

BANKERS ACCEPTANCE: 1.25% 30 days; 1.22% 60 days; 1.19% 90 days; 1.17% 120 days; 1.16% 150 days; 1.14% 180 days; Source: Prebon Yamane (USA) Inc.

LONDON INTERBANK OFFERED RATES (LIBOR): 1.31875% one month; 1.2800% three months; 1.2300% six months; 1.2300% one year. Effective rate for contracts entered into two days from date appearing at top of this column.

EURO INTERBANK OFFERED RATES (EURIBOR): 2.319% one month; 2.235% three months; 2.179% six months; 2.122% one year. Source: Reuters.

FOREIGN PRIME RATES: Canada 5.00%; European Central Bank 2.50%; Japan 1.375%; Switzerland 2.25%; Britain 3.75%

TREASURY BILLS: Results of the Monday, June 2, 2003, auction of short-term U.S. government bills, sold at a discount from face value in units of $1,000 to $1 million: 1.110% 13 weeks; 1.095% 26 weeks. Tuesday, June 3, 2003 auction: 1.140% 4 weeks.

OVERNIGHT REPURCHASE RATE: 1.22%. Source: Garban Intercapital

FREDDIE MAC: Posted yields on 30-year mortgage commitments. Delivery within 30 days 4.68%, 60 days 4.80%, standard conventional fixed-rate mortgages: 2.875%, 2% rate capped one-year adjustable rate mortgages.

FANNIE MAE: Posted yields on 30 year mortgage commitments (priced at par) for delivery within 30 days 4.78%, 60 days 4.87% standard conventional fixed-rate mortgages; 3.00% 6/2 rate capped one-year adjustable rate mortgages. Constant Maturity Debt Index: 1.193% three months; 1.119% six months; 1.187% one year

MERRILL LYNCH READY ASSETS TRUST: 0.78%.

CONSUMER PRICE INDEX: April 183.8, up 2.2% from a year ago. Bureau of Labor Statistics.

Source: Wall Street Journal, Wednesday, June 4, 2003, p. C14.

Repurchase Agreements. *Repurchase agreements*, or *repos*, are effectively short-term loans (usually with a maturity of less than two weeks) in which Treasury bills serve as *collateral*, an asset that the lender receives if the borrower does not pay back the loan. Repos are made as follows: A large corporation, such as General Motors, may have some idle funds in its bank account, say $1 million, which it would like to lend for a week. GM uses this excess $1 million to buy Treasury bills from a bank, which agrees

to repurchase them the next week at a price slightly above GM's purchase price. The effect of this agreement is that GM makes a loan of $1 million to the bank and holds $1 million of the bank's Treasury bills until the bank repurchases the bills to pay off the loan. Repurchase agreements are a fairly recent innovation in financial markets, having been introduced in 1969. They are now an important source of bank funds (over $400 billion). The most important lenders in this market are large corporations.

Federal (Fed) Funds. These are typically overnight loans between banks of their deposits at the Federal Reserve. The *federal funds* designation is somewhat confusing, because these loans are not made by the federal government or by the Federal Reserve, but rather by banks to other banks. One reason why a bank might borrow in the federal funds market is that it might find it does not have enough deposits at the Fed to meet the amount required by regulators. It can then borrow these deposits from another bank, which transfers them to the borrowing bank using the Fed's wire transfer system. This market is very sensitive to the credit needs of the banks, so the interest rate on these loans, called the federal funds rate, is a closely watched barometer of the tightness of credit market conditions in the banking system and the stance of monetary policy; when it is high, it indicates that the banks are strapped for funds, whereas when it is low, banks' credit needs are low.

Capital Market Instruments

Capital market instruments are debt and equity instruments with maturities of greater than one year. They have far wider price fluctuations than money market instruments and are considered to be fairly risky investments. The principal capital market instruments are listed in Table 2, which shows the amount outstanding at the end of 1970, 1980, 1990, and 2002.

Table 2 Principal Capital Market Instruments

Type of Instrument	Amount Outstanding ($ billions, end of year) 1970	1980	1990	2002
Corporate stocks (market value)	906	1,601	4,146	11,734
Residential mortgages	355	1,106	2,886	6,930
Corporate bonds	167	366	1,008	2,699
U.S. government securities (marketable long-term)	160	407	1,653	2,169
U.S. government agency securities	51	193	435	2,305
State and local government bonds	146	310	870	1,442
Bank commercial loans	152	459	818	1,345
Consumer loans	134	355	813	1,757
Commercial and farm mortgages	116	352	829	1,461

Sources: Federal Reserve Flow of Funds Accounts; Federal Reserve *Bulletin*; *Banking and Monetary Statistics, 1941–1970*. http://www.federalreserve.gov/releases/z1

Stocks. *Stocks* are equity claims on the net income and assets of a corporation. Their value of $11 trillion at the end of 2002 exceeds that of any other type of security in the capital market. The amount of new stock issues in any given year is typically quite small—less than 1% of the total value of shares outstanding. Individuals hold around half of the value of stocks; the rest are held by pension funds, mutual funds, and insurance companies.

Mortgages. *Mortgages* are loans to households or firms to purchase housing, land, or other real structures, where the structure or land itself serves as collateral for the loans. The mortgage market is the largest debt market in the United States, with the amount of residential mortgages (used to purchase residential housing) outstanding more than quadruple the amount of commercial and farm mortgages. Savings and loan associations and mutual savings banks have been the primary lenders in the residential mortgage market, although commercial banks have started to enter this market more aggressively. The majority of commercial and farm mortgages are made by commercial banks and life insurance companies. The federal government plays an active role in the mortgage market via the three government agencies—the Federal National Mortgage Association (FNMA, "Fannie Mae"), the Government National Mortgage Association (GNMA, "Ginnie Mae"), and the Federal Home Loan Mortgage Corporation (FHLMC, "Freddie Mac")—that provide funds to the mortgage market by selling bonds and using the proceeds to buy mortgages. An important development in the residential mortgage market in recent years is the mortgage-backed security (see Box 1).

Box 1

Mortgage-Backed Securities

A major change in the residential mortgage market in recent years has been the creation of an active secondary market for mortgages. Because mortgages have different terms and interest rates, they were not sufficiently liquid to trade as securities on secondary markets. To stimulate mortgage lending, in 1970 the Government National Mortgage Association (GNMA, called "Ginnie Mae") developed the concept of a pass-through *mortgage-backed security* when it began a program in which it guaranteed interest and principal payments on bundles of standardized mortgages. Under this program, private financial institutions such as savings and loans and commercial banks were now able to gather a group of GNMA-guaranteed mortgages into a bundle of, say, $1 million and then sell this bundle as a security to a third party (usually a large institutional investor such as a pension fund). When individuals make their mortgage payments on the GNMA-guaranteed mortgage to the financial institution, the financial institution passes the payments through to the owner of the security by sending a check for the total of all the payments. Because GNMA guarantees the payments, these pass-through securities have a very low default risk and are very popular, with amounts outstanding exceeding $500 billion.

Mortgage-backed securities are issued not only by the government agencies, but also by private financial institutions. Indeed, mortgage-backed securities have been so successful that they have completely transformed the residential mortgage market. Throughout the 1970s, over 80% of residential mortgages were owned outright by savings and loans, mutual savings banks, and commercial banks. Now only one-third are owned outright by these institutions, with two-thirds held as mortgage-backed securities.

Corporate Bonds. These are long-term bonds issued by corporations with very strong credit ratings. The typical *corporate bond* sends the holder an interest payment twice a year and pays off the face value when the bond matures. Some corporate bonds, called *convertible bonds*, have the additional feature of allowing the holder to convert them into a specified number of shares of stock at any time up to the maturity date. This feature makes these convertible bonds more desirable to prospective purchasers than bonds without it, and allows the corporation to reduce its interest payments, because these bonds can increase in value if the price of the stock appreciates sufficiently. Because the outstanding amount of both convertible and nonconvertible bonds for any given corporation is small, they are not nearly as liquid as other securities such as U.S. government bonds.

Although the size of the corporate bond market is substantially smaller than that of the stock market, with the amount of corporate bonds outstanding less than one-fourth that of stocks, the volume of new corporate bonds issued each year is substantially greater than the volume of new stock issues. Thus the behavior of the corporate bond market is probably far more important to a firm's financing decisions than the behavior of the stock market. The principal buyers of corporate bonds are life insurance companies; pension funds and households are other large holders.

U.S. Government Securities. These long-term debt instruments are issued by the U.S. Treasury to finance the deficits of the federal government. Because they are the most widely traded bonds in the United States (the volume of transactions on average exceeds $100 billion daily), they are the most liquid security traded in the capital market. They are held by the Federal Reserve, banks, households, and foreigners.

U.S. Government Agency Securities. These are long-term bonds issued by various government agencies such as Ginnie Mae, the Federal Farm Credit Bank, and the Tennessee Valley Authority to finance such items as mortgages, farm loans, or power-generating equipment. Many of these securities are guaranteed by the federal government. They function much like U.S. government bonds and are held by similar parties.

State and Local Government Bonds. State and local bonds, also called *municipal bonds*, are long-term debt instruments issued by state and local governments to finance expenditures on schools, roads, and other large programs. An important feature of these bonds is that their interest payments are exempt from federal income tax and generally from state taxes in the issuing state. Commercial banks, with their high income tax rate, are the biggest buyers of these securities, owning over half the total amount outstanding. The next biggest group of holders consists of wealthy individuals in high income brackets, followed by insurance companies.

Consumer and Bank Commercial Loans. These are loans to consumers and businesses made principally by banks, but—in the case of consumer loans—also by finance companies. There are often no secondary markets in these loans, which makes them the least liquid of the capital market instruments listed in Table 2. However, secondary markets have been rapidly developing.

Chapter 3

What Is Money?

PREVIEW

If you had lived in America before the Revolutionary War, your money might have consisted primarily of Spanish doubloons (silver coins that were also called *pieces of eight*). Before the Civil War, the principal forms of money in the United States were not only gold and silver coins but also paper notes, called *banknotes*, issued by private banks. Today, you use not only coins and dollar bills issued by the government as money, but also checks written on accounts held at banks. Money has been different things at different times; however, it has *always* been important to people and to the economy.

To understand the effects of money on the economy, we must understand exactly what money is. In this chapter, we develop precise definitions by exploring the functions of money, looking at why and how it promotes economic efficiency, tracing how its forms have evolved over time, and examining how money is currently measured.

Meaning of Money

As the word *money* is used in everyday conversation, it can mean many things, but to economists, it has a very specific meaning. To avoid confusion, we must clarify how economists' use of the word *money* differs from conventional usage.

Economists define *money* (also referred to as the *money supply*) as anything that is generally accepted in payment for goods or services or in the repayment of debts. Currency, consisting of dollar bills and coins, clearly fits this definition and is one type of money. When most people talk about money, they're talking about **currency** (paper money and coins). If, for example, someone comes up to you and says, "Your money or your life," you should quickly hand over all your currency rather than ask, "What exactly do you mean by 'money'?"

To define money merely as currency is much too narrow for economists. Because checks are also accepted as payment for purchases, checking account deposits are considered money as well. An even broader definition of money is often needed, because other items such as savings deposits can in effect function as money if they can be quickly and easily converted into currency or checking account deposits. As you can see, there is no single, precise definition of money or the money supply, even for economists.

To complicate matters further, the word *money* is frequently used synonymously with *wealth*. When people say, "Joe is rich—he has an awful lot of money," they probably mean that Joe has not only a lot of currency and a high balance in his checking account but has also stocks, bonds, four cars, three houses, and a yacht. Thus while "currency" is too narrow a definition of money, this other popular usage is much too broad. Economists make a distinction between money in the form of currency, demand deposits, and other items that are used to make purchases and **wealth**, the total collection of pieces of property that serve to store value. Wealth includes not only money but also other assets such as bonds, common stock, art, land, furniture, cars, and houses.

People also use the word *money* to describe what economists call *income*, as in the sentence "Sheila would be a wonderful catch; she has a good job and earns a lot of money." **Income** is a *flow* of earnings per unit of time. Money, by contrast, is a *stock*: It is a certain amount at a given point in time. If someone tells you that he has an income of $1,000, you cannot tell whether he earned a lot or a little without knowing whether this $1,000 is earned per year, per month, or even per day. But if someone tells you that she has $1,000 in her pocket, you know exactly how much this is.

Keep in mind that the money discussed in this book refers to anything that is generally accepted in payment for goods and services or in the repayment of debts and is distinct from income and wealth.

Functions of Money

Whether money is shells or rocks or gold or paper, it has three primary functions in any economy: as a medium of exchange, as a unit of account, and as a store of value. Of the three functions, its function as a medium of exchange is what distinguishes money from other assets such as stocks, bonds, and houses.

Medium of Exchange

In almost all market transactions in our economy, money in the form of currency or checks is a **medium of exchange**; it is used to pay for goods and services. The use of money as a medium of exchange promotes economic efficiency by minimizing the time spent in exchanging goods and services. To see why, let's look at a *barter economy*, one without money, in which goods and services are exchanged directly for other goods and services.

Take the case of Ellen the Economics Professor, who can do just one thing well: give brilliant economics lectures. In a barter economy, if Ellen wants to eat, she must find a farmer who not only produces the food she likes but also wants to learn economics. As you might expect, this search will be difficult and time-consuming, and Ellen might spend more time looking for such an economics-hungry farmer than she will teaching. It is even possible that she will have to quit lecturing and go into farming herself. Even so, she may still starve to death.

The time spent trying to exchange goods or services is called a *transaction cost*. In a barter economy, transaction costs are high because people have to satisfy a "double coincidence of wants"—they have to find someone who has a good or service they want and who also wants the good or service they have to offer.

Let's see what happens if we introduce money into Ellen the Economics Professor's world. Ellen can teach anyone who is willing to pay money to hear her lecture. She can then go to any farmer (or his representative at the supermarket) and buy the food she needs with the money she has been paid. The problem of the double coincidence of wants is avoided, and Ellen saves a lot of time, which she may spend doing what she does best: teaching.

As this example shows, money promotes economic efficiency by eliminating much of the time spent exchanging goods and services. It also promotes efficiency by allowing people to specialize in what they do best. Money is therefore essential in an economy: It is a lubricant that allows the economy to run more smoothly by lowering transaction costs, thereby encouraging specialization and the division of labor.

The need for money is so strong that almost every society beyond the most primitive invents it. For a commodity to function effectively as money, it has to meet several criteria: (1) It must be easily standardized, making it simple to ascertain its value; (2) it must be widely accepted; (3) it must be divisible, so that it is easy to "make change"; (4) it must be easy to carry; and (5) it must not deteriorate quickly. Forms of money that have satisfied these criteria have taken many unusual forms throughout human history, ranging from wampum (strings of beads) used by Native Americans, to tobacco and whiskey, used by the early American colonists, to cigarettes, used in prisoner-of-war camps during World War II.[1] The diversity of forms of money that have been developed over the years is as much a testament to the inventiveness of the human race as the development of tools and language.

Unit of Account

The second role of money is to provide a **unit of account**; that is, it is used to measure value in the economy. We measure the value of goods and services in terms of money, just as we measure weight in terms of pounds or distance in terms of miles. To see why this function is important, let's look again at a barter economy where money does not perform this function. If the economy has only three goods—say, peaches, economics lectures, and movies—then we need to know only three prices to tell us how to exchange one for another: the price of peaches in terms of economics lectures (that is, how many economics lectures you have to pay for a peach), the price of peaches in terms of movies, and the price of economics lectures in terms of movies. If there were ten goods, we would need to know 45 prices in order to exchange one good for another; with 100 goods, we would need 4,950 prices; and with 1,000 goods, 499,500 prices.[2]

Imagine how hard it would be in a barter economy to shop at a supermarket with 1,000 different items on its shelves, having to decide whether chicken or fish is a better buy if the price of a pound of chicken were quoted as 4 pounds of butter and the price of a pound of fish as 8 pounds of tomatoes. To make it possible to compare

---

[1] An extremely entertaining article on the development of money in a prisoner-of-war camp during World War II is R. A. Radford, "The Economic Organization of a P.O.W. Camp," *Economica* 12 (November 1945): 189–201.

[2] The formula for telling us the number of prices we need when we have N goods is the same formula that tells us the number of pairs when there are N items. It is

$$\frac{N(N-1)}{2}$$

In the case of ten goods, for example, we would need

$$\frac{10(10-1)}{2} = \frac{90}{2} = 45$$

prices, the tag on each item would have to list up to 999 different prices, and the time spent reading them would result in very high transaction costs.

The solution to the problem is to introduce money into the economy and have all prices quoted in terms of units of that money, enabling us to quote the price of economics lectures, peaches, and movies in terms of, say, dollars. If there were only three goods in the economy, this would not be a great advantage over the barter system, because we would still need three prices to conduct transactions. But for ten goods we now need only ten prices; for 100 goods, 100 prices; and so on. At the 1,000-good supermarket, there are now only 1,000 prices to look at, not 499,500!

We can see that using money as a unit of account reduces transaction costs in an economy by reducing the number of prices that need to be considered. The benefits of this function of money grow as the economy becomes more complex.

Store of Value

Money also functions as a **store of value**; it is a repository of purchasing power over time. A store of value is used to save purchasing power from the time income is received until the time it is spent. This function of money is useful, because most of us do not want to spend our income immediately upon receiving it, but rather prefer to wait until we have the time or the desire to shop.

Money is not unique as a store of value; any asset—whether money, stocks, bonds, land, houses, art, or jewelry—can be used to store wealth. Many such assets have advantages over money as a store of value: They often pay the owner a higher interest rate than money, experience price appreciation, and deliver services such as providing a roof over one's head. If these assets are a more desirable store of value than money, why do people hold money at all?

The answer to this question relates to the important economic concept of **liquidity**, the relative ease and speed with which an asset can be converted into a medium of exchange. Liquidity is highly desirable. Money is the most liquid asset of all because it *is* the medium of exchange; it does not have to be converted into anything else in order to make purchases. Other assets involve transaction costs when they are converted into money. When you sell your house, for example, you have to pay a brokerage commission (usually 5% to 7% of the sales price), and if you need cash immediately to pay some pressing bills, you might have to settle for a lower price in order to sell the house quickly. Because money is the most liquid asset, people are willing to hold it even if it is not the most attractive store of value.

How good a store of value money is depends on the price level, because its value is fixed in terms of the price level. A doubling of all prices, for example, means that the value of money has dropped by half; conversely, a halving of all prices means that the value of money has doubled. During inflation, when the price level is increasing rapidly, money loses value rapidly, and people will be more reluctant to hold their wealth in this form. This is especially true during periods of extreme inflation, known as **hyperinflation**, in which the inflation rate exceeds 50% per month.

Hyperinflation occurred in Germany after World War I, with inflation rates sometimes exceeding 1,000% per month. By the end of the hyperinflation in 1923, the price level had risen to more than 30 billion times what it had been just two years before. The quantity of money needed to purchase even the most basic items became excessive. There are stories, for example, that near the end of the hyperinflation, a wheelbarrow of cash would be required to pay for a loaf of bread. Money was losing its value so rapidly that workers were paid and given time off several times during the day to spend their wages before the money became worthless. No one wanted to hold

on to money, and so the use of money to carry out transactions declined and barter became more and more dominant. Transaction costs skyrocketed, and as we would expect, output in the economy fell sharply.

Evolution of the Payments System

www.federalreserve.gov/paymentsys.htm
This site reports on the Federal Reserve's policies regarding payments systems.

We can obtain a better picture of the functions of money and the forms it has taken over time by looking at the evolution of the **payments system**, the method of conducting transactions in the economy. The payments system has been evolving over centuries, and with it the form of money. At one point, precious metals such as gold were used as the principal means of payment and were the main form of money. Later, paper assets such as checks and currency began to be used in the payments system and viewed as money. Where the payments system is heading has an important bearing on how money will be defined in the future.

Commodity Money

To obtain perspective on where the payments system is heading, it is worth exploring how it has evolved. For any object to function as money, it must be universally acceptable; everyone must be willing to take it in payment for goods and services. An object that clearly has value to everyone is a likely candidate to serve as money, and a natural choice is a precious metal such as gold or silver. Money made up of precious metals or another valuable commodity is called **commodity money**, and from ancient times until several hundred years ago, commodity money functioned as the medium of exchange in all but the most primitive societies. The problem with a payments system based exclusively on precious metals is that such a form of money is very heavy and is hard to transport from one place to another. Imagine the holes you'd wear in your pockets if you had to buy things only with coins! Indeed, for large purchases such as a house, you'd have to rent a truck to transport the money payment.

Fiat Money

The next development in the payments system was *paper currency* (pieces of paper that function as a medium of exchange). Initially, paper currency carried a guarantee that it was convertible into coins or into a quantity of precious metal. However, currency has evolved into **fiat money**, paper currency decreed by governments as legal tender (meaning that legally it must be accepted as payment for debts) but not convertible into coins or precious metal. Paper currency has the advantage of being much lighter than coins or precious metal, but it can be accepted as a medium of exchange only if there is some trust in the authorities who issue it and if printing has reached a sufficiently advanced stage that counterfeiting is extremely difficult. Because paper currency has evolved into a legal arrangement, countries can change the currency that they use at will. Indeed, this is currently a hot topic of debate in Europe, which has adopted a unified currency (see Box 1).

Major drawbacks of paper currency and coins are that they are easily stolen and can be expensive to transport in large amounts because of their bulk. To combat this problem, another step in the evolution of the payments system occurred with the development of modern banking: the invention of *checks*.

Checks

A check is an instruction from you to your bank to transfer money from your account to someone else's account when she deposits the check. Checks allow transactions to

Box 1: Global

Birth of the Euro: Will It Benefit Europe?

As part of the December 1991 Maastricht Treaty on European Union, the European Economic Commission outlined a plan to achieve the creation of a single European currency starting in 1999. Despite concerns, the new common currency—the euro—came into existence right on schedule in January 1999, with 11 of the 15 European Union countries participating in the monetary union: Austria, Belgium, Finland, France, Germany, Italy, Ireland, Luxembourg, the Netherlands, Portugal, and Spain. Denmark, Sweden, and the United Kingdom chose not to participate initially, and Greece failed to meet the economic criteria specified by the Maastricht Treaty (such as having a budget deficit less than 3% of GDP and total government debt less than 60% of GDP) but was able to join later.

Starting January 1, 1999, the exchange rates of countries entering the monetary union were fixed permanently to the euro (which became a unit of account), the European Central Bank took over monetary policy from the individual national central banks, and the governments of the member countries began to issue debt in euros. In early 2002, euro notes and coins began to circulate and by June 2002, the old national currencies were phased out completely, so that only euros could be used in the member countries.

Advocates of monetary union point out the advantages that the single currency has in eliminating the transaction costs incurred in exchanging one currency for another. In addition, the use of a single currency may promote further integration of the European economies and enhance competition. Skeptics who think that monetary union may be bad for Europe suggest that because labor will not be very mobile across national boundaries and because fiscal transfers (i.e., tax income from one region being spent on another) from better-performing regions to worse-performing regions will not take place as occurs in the United States, a single currency may lead to some regions of Europe being depressed for substantial periods of time while other regions are booming.

Whether the euro will be good for the economies of Europe and increase their GDP is an open question. However, the motive behind monetary union was probably more political than economic. European monetary union may encourage political union, producing a unified Europe that can play a stronger economic and political role on the world stage.

take place without the need to carry around large amounts of currency. The introduction of checks was a major innovation that improved the efficiency of the payments system. Frequently, payments made back and forth cancel each other; without checks, this would involve the movement of a lot of currency. With checks, payments that cancel each other can be settled by canceling the checks, and no currency need be moved. The use of checks thus reduces the transportation costs associated with the payments system and improves economic efficiency. Another advantage of checks is that they can be written for any amount up to the balance in the account, making transactions for large amounts much easier. Checks are also advantageous in that loss from theft is greatly reduced, and because they provide convenient receipts for purchases.

There are, however, two problems with a payments system based on checks. First, it takes time to get checks from one place to another, a particularly serious problem if you are paying someone in a different location who needs to be paid quickly. In addition, if you have a checking account, you know that it usually takes several business days before a bank will allow you to make use of the funds from a check you have deposited. If your need for cash is urgent, this feature of paying by check can be

frustrating. Second, all the paper shuffling required to process checks is costly; it is estimated that it currently costs over $10 billion per year to process all the checks written in the United States.

Electronic Payment

The development of inexpensive computers and the spread of the Internet now make it cheap to pay bills electronically. In the past, you had to pay your bills by mailing a check, but now banks provide a web site in which you just log on, make a few clicks, and thereby transmit your payment electronically. Not only do you save the cost of the stamp, but paying bills becomes (almost) a pleasure, requiring little effort. Electronic payment systems provided by banks now even spare you the step of logging on to pay the bill. Instead, recurring bills can be automatically deducted from your bank account. Estimated cost savings when a bill is paid electronically rather than by a check exceed one dollar. Electronic payment is thus becoming far more common in the United States, but Americans lag considerably behind Europeans, particularly Scandinavians, in their use of electronic payments (see Box 2).

Box 2: E-Finance

Why Are Scandinavians So Far Ahead of Americans in Using Electronic Payments?

Americans are the biggest users of checks in the world. Close to 100 billion checks are written every year in the United States, and over three-quarters of noncash transactions are conducted with paper. In contrast, in most countries of Europe, more than two-thirds of noncash transactions are electronic, with Finland and Sweden having the greatest proportion of online banking customers of any countries in the world. Indeed, if you were Finnish or Swedish, instead of writing a check, you would be far more likely to pay your bills online, using a personal computer or even a mobile phone. Why do Europeans and especially Scandinavians so far outpace Americans in the use of electronic payments?

First, Europeans got used to making payments without checks even before the advent of the personal computer. Europeans have long made use of so-called *giro* payments, in which banks and post offices transfer funds for customers to pay bills. Second, Europeans—and particularly Scandinavians—are much greater users of mobile phones and the Internet than are Americans. Finland has the highest per capita use of mobile phones in the world, and Finland and Sweden lead the world in the percentage of the population that accesses the Internet. Maybe these usage patterns stem from the low population densities of these countries and the cold and dark winters that keep Scandinavians inside at their PCs. For their part, Scandinavians would rather take the view that their high-tech culture is the product of their good education systems and the resulting high degree of computer literacy, the presence of top technology companies such as Finland's Nokia and Sweden's Ericsson, and government policies promoting the increased use of personal computers, such as Sweden's tax incentives for companies to provide their employees with home computers. The wired populations of Finland and Sweden are (percentage-wise) the biggest users of online banking in the world.

Americans are clearly behind the curve in their use of electronic payments, which has imposed a high cost on the U.S. economy. Switching from checks to electronic payments might save the U.S. economy tens of billions of dollars per year, according to some estimates. Indeed, the U.S. federal government is trying to switch all its payments to electronic ones by directly depositing them into bank accounts, in order to reduce its expenses. Can Americans be weaned from paper checks and fully embrace the world of high-tech electronic payments?

E-Money

Electronic payments technology can not only substitute for checks, but can substitute for cash, as well, in the form of **electronic money** (or **e-money**), money that exists only in electronic form. The first form of e-money was the *debit card*. Debit cards, which look like credit cards, enable consumers to purchase goods and services by electronically transferring funds directly from their bank accounts to a merchant's account. Debit cards are used in many of the same places that accept credit cards and are now often becoming faster to use than cash. At most supermarkets, for example, you can swipe your debit card through the card reader at the checkout station, press a button, and the amount of your purchases is deducted from your bank account. Most banks and companies such as Visa and MasterCard issue debit cards, and your ATM card typically can function as a debit card.

A more advanced form of e-money is the *stored-value card*. The simplest form of stored-value card is purchased for a preset dollar amount that the consumer pays up front, like a prepaid phone card. The more sophisticated stored-value card is known as a **smart card**. It contains a computer chip that allows it to be loaded with digital cash from the owner's bank account whenever needed. Smart cards can be loaded from ATM machines, personal computers with a smart card reader, or specially equipped telephones.

A third form of electronic money is often referred to as **e-cash**, which is used on the Internet to purchase goods or services. A consumer gets e-cash by setting up an account with a bank that has links to the Internet and then has the e-cash transferred to her PC. When she wants to buy something with e-cash, she surfs to a store on the Web and clicks the "buy" option for a particular item, whereupon the e-cash is automatically transferred from her computer to the merchant's computer. The merchant can then have the funds transferred from the consumer's bank account to his before the goods are shipped.

Given the convenience of e-money, you might think that we would move quickly to the cashless society in which all payments were made electronically. However, this hasn't happened, as discussed in Box 3.

Measuring Money

The definition of money as anything that is generally accepted in payment for goods and services tells us that money is defined by people's behavior. What makes an asset money is that people believe it will be accepted by others when making payment. As we have seen, many different assets have performed this role over the centuries, ranging from gold to paper currency to checking accounts. For that reason, this behavioral definition does not tell us exactly what assets in our economy should be considered money. To measure money, we need a precise definition that tells us exactly what assets should be included.

The Federal Reserve's Monetary Aggregates

The Federal Reserve System (the Fed), the central banking authority responsible for monetary policy in the United States, has conducted many studies on how to measure money. The problem of measuring money has recently become especially crucial because extensive financial innovation has produced new types of assets that might properly belong in a measure of money. Since 1980, the Fed has modified its measures of money several times and has settled on the following measures of the money

Box 3: E-Finance

Are We Headed for a Cashless Society?

Predictions of a cashless society have been around for decades, but they have not come to fruition. For example, *Business Week* predicted in 1975 that electronic means of payment "would soon revolutionize the very concept of money itself," only to reverse itself several years later. Pilot projects in recent years with smart cards to convert consumers to the use of e-money have not been a success. Mondex, one of the widely touted, early stored-value cards that was launched in Britain in 1995, is only used on a few British university campuses. In Germany and Belgium, millions of people carry bank cards with computer chips embedded in them that enable them to make use of e-money, but very few use them. Why has the movement to a cashless society been so slow in coming?

Although e-money might be more convenient and may be more efficient than a payments system based on paper, several factors work against the disappearance of the paper system. First, it is very expensive to set up the computer, card reader, and telecommunications networks necessary to make electronic money the dominant form of payment. Second, electronic means of payment raise security and privacy concerns. We often hear media reports that an unauthorized hacker has been able to access a computer database and to alter information stored there. Because this is not an uncommon occurrence, unscrupulous persons might be able to access bank accounts in electronic payments systems and steal funds by moving them from someone else's accounts into their own. The prevention of this type of fraud is no easy task, and a whole new field of computer science has developed to cope with security issues. A further concern is that the use of electronic means of payment leaves an electronic trail that contains a large amount of personal data on buying habits. There are worries that government, employers, and marketers might be able to access these data, thereby encroaching on our privacy.

The conclusion from this discussion is that although the use of e-money will surely increase in the future, to paraphrase Mark Twain, "the reports of cash's death are greatly exaggerated."

www.federalreserve.gov/releases/h6/Current/

The Federal Reserve reports the current levels of M1, M2, and M3 on its web site.

supply, which are also referred to as **monetary aggregates** (see Table 1 and the Following the Financial News box).

The narrowest measure of money that the Fed reports is **M1**, which includes currency, checking account deposits, and traveler's checks. These assets are clearly money, because they can be used directly as a medium of exchange. Until the mid-1970s, only commercial banks were permitted to establish checking accounts, and they were not allowed to pay interest on them. With the financial innovation that has occurred (discussed more extensively in Chapter 9), regulations have changed so that other types of banks, such as savings and loan associations, mutual savings banks, and credit unions, can also offer checking accounts. In addition, banking institutions can offer other checkable deposits, such as NOW (negotiated order of withdrawal) accounts and ATS (automatic transfer from savings) accounts, that do pay interest on their balances. Table 1 lists the assets included in the measures of the monetary aggregates; both demand deposits (checking accounts that pay no interest) and these other checkable deposits are included in the M1 measure.

The **M2** monetary aggregate adds to M1 other assets that have check-writing features (money market deposit accounts and money market mutual fund shares) and other assets (savings deposits, small-denomination time deposits and repurchase agreements) that are extremely liquid, because they can be turned into cash quickly at very little cost.

Table 1 Measures of the Monetary Aggregates

	Value as of December 2002 ($billions)
M1 = Currency	626.5
+ Traveler's checks	7.7
+ Demand deposits	290.7
+ Other checkable deposits	281.2
Total M1	1,206.1
M2 = M1	
+ Small-denomination time deposits and repurchase agreements	1,332.3
+ Savings deposits and money market deposit accounts	2,340.4
+ Money market mutual fund shares (noninstitutional)	923.7
Total M2	5,802.5
M3 = M2	
+ Large-denomination time deposits and repurchase agreements	1,105.2
+ Money market mutual fund shares (institutional)	767.7
+ Repurchase agreements	511.7
+ Eurodollars	341.1
Total M3	8,528.2

Source: www.federalreserve.gov/releases/h6/hist.

Note: The *Travelers checks* item includes only traveler's checks issued by non-banks, while traveler's checks issued by banks are included in the *Demand deposits* item, which also includes checkable deposits to businesses and which also do not pay interest.

The **M3** monetary aggregate adds to M2 somewhat less liquid assets such as large-denomination time deposits and repurchase agreements, Eurodollars, and institutional money market mutual fund shares.

Because we cannot be sure which of the monetary aggregates is the true measure of money, it is logical to wonder if their movements closely parallel one another. If they do, then using one monetary aggregate to predict future economic performance and to conduct policy will be the same as using another, and it does not matter much that we are not sure of the appropriate definition of money for a given policy decision. However, if the monetary aggregates do not move together, then what one monetary aggregate tells us is happening to the money supply might be quite different from what another monetary aggregate would tell us. The conflicting stories might present a confusing picture that would make it hard for policymakers to decide on the right course of action.

Figure 1 plots the growth rates M1, M2, and M3 from 1960 to 2002. The growth rates of these three monetary aggregates do tend to move together; the timing of their rise and fall is roughly similar until the 1990s, and they all show a higher growth rate on average in the 1970s than in the 1960s.

Yet some glaring discrepancies exist in the movements of these aggregates. According to M1, the growth rate of money did not accelerate between 1968, when it

Following the Financial News

The Monetary Aggregates

Data for the Federal Reserve's monetary aggregates (M1, M2, and M3) are published every Friday. In the *Wall Street Journal*, the data are found in the "Federal Reserve Data" column, an example of which is presented here.

The third entry indicates that the money supply (M2) averaged $5,822.7 billion for the week ending December 23, 2002. The notation "sa" for this entry indicates that the data are seasonally adjusted; that is, seasonal movements, such as those associated with Christmas shopping, have been removed from the data. The notation "nsa" indicates that the data have not been seasonally adjusted.

FEDERAL RESERVE DATA

MONETARY AGGREGATES
(daily average in billions)

	1 Week Ended:	
	Dec. 23	Dec. 16
Money supply (M1) sa	1227.1	1210.1
Money supply (M1) nsa	1256.0	1214.9
Money supply (M2) sa	5822.7	5811.3
Money supply (M2) nsa	5834.5	5853.9
Money supply (M3) sa	8542.8	8549.2
Money supply (M3) nsa	8572.6	8623.0

	4 Weeks Ended:	
	Dec. 23	Nov. 25
Money supply (M1) sa	1218.3	1197.5
Money supply (M1) nsa	1230.9	1195.9
Money supply (M2) sa	5815.5	5795.8
Money supply (M2) nsa	5835.7	5780.7
Money supply (M3) sa	8543.4	8465.4
Money supply (M3) nsa	8578.1	8440.5

	Month	
	Nov.	Oct.
Money supply (M1) sa	1200.7	1199.6
Money supply (M2) sa	5800.7	5753.8
Money supply (M3) sa	8485.2	8348.4

nsa-Not seasonally adjusted
sa-Seasonally adjusted.

Source: Wall Street Journal, Friday, January 3, 2003, p. C10.

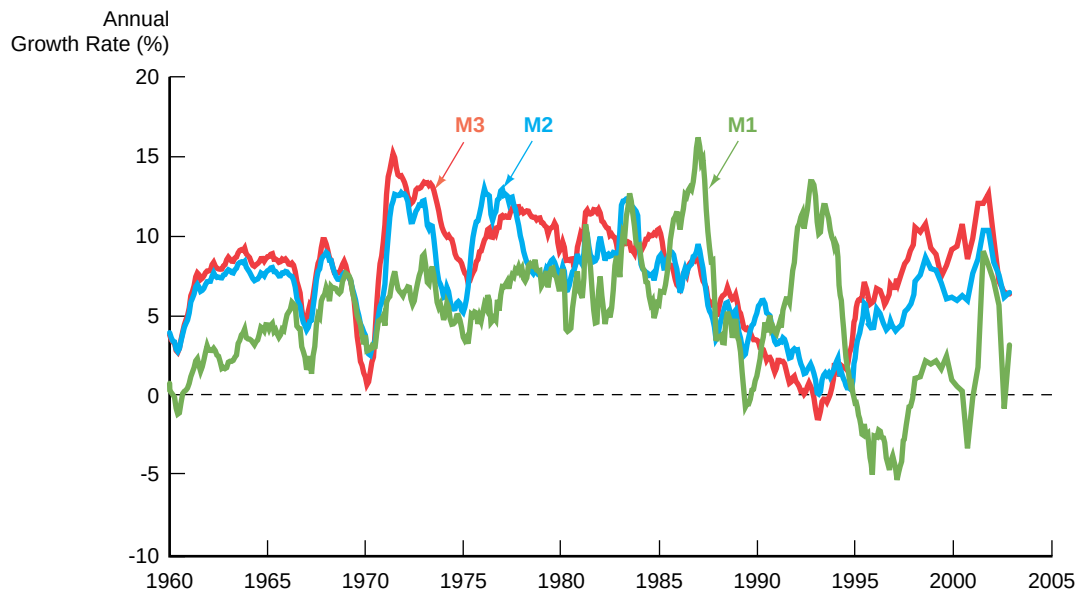


FIGURE 1 Growth Rates of the Three Money Aggregates, 1960–2002

Sources: Federal Reserve *Bulletin*, p. A4, Table 1.10, various issues; Citibase databank; www.federalreserve.gov/releases/h6/hist/h6hist1.txt.

was in the 6–7% range, and 1971, when it was at a similar level. In the same period, the M2 and M3 measures tell a different story; they show a marked acceleration from the 8–10% range to the 12–15% range. Similarly, while the growth rate of M1 actually increased from 1989 to 1992, the growth rates of M2 and M3 in this same period instead showed a downward trend. Furthermore, from 1992 to 1998, the growth rate of M1 fell sharply while the growth rates of M2 and M3 rose substantially; from 1998 to 2002, M1 growth generally remained well below M2 and M3 growth. Thus, the different measures of money tell a very different story about the course of monetary policy in recent years.

From the data in Figure 1, you can see that obtaining a single precise, correct measure of money does seem to matter and that it does make a difference which monetary aggregate policymakers and economists choose as the true measure of money.

How Reliable Are the Money Data?

The difficulties of measuring money arise not only because it is hard to decide what is the best definition of money, but also because the Fed frequently later revises earlier estimates of the monetary aggregates by large amounts. There are two reasons why the Fed revises its figures. First, because small depository institutions need to report the amounts of their deposits only infrequently, the Fed has to estimate these amounts until these institutions provide the actual figures at some future date. Second, the adjustment of the data for seasonal variation is revised substantially as more data become available. To see why this happens, let's look at an example of the seasonal variation of the money data around Christmas-time. The monetary aggregates always rise around Christmas because of increased spending during the holiday season; the rise is greater in some years than in others. This means that the factor that adjusts the data for the seasonal variation due to Christmas must be estimated from several years of data, and the estimates of this seasonal factor become more precise only as more data become available. When the data on the monetary aggregates are revised, the seasonal adjustments often change dramatically from the initial calculation.

Table 2 shows how severe a problem these data revisions can be. It provides the rates of money growth from one-month periods calculated from initial estimates of the M2 monetary aggregate, along with the rates of money growth calculated from a major revision of the M2 numbers published in February 2003. As the table shows, for one-month periods the initial versus the revised data can give a different picture of what is happening to monetary policy. For January 2003, for example, the initial data indicated that the growth rate of M2 at an annual rate was 2.2%, whereas the revised data indicate a much higher growth rate of 5.4%.

A distinctive characteristic shown in Table 2 is that the differences between the initial and revised M2 series tend to cancel out. You can see this by looking at the last row of the table, which shows the average rate of M2 growth for the two series and the average difference between them. The average M2 growth for the initial calculation of M2 is 6.5%, and the revised number is 6.5%, a difference of 0.0%. The conclusion we can draw is that the initial data on the monetary aggregates reported by the Fed are not a reliable guide to what is happening to short-run movements in the money supply, such as the one-month growth rates. However, the initial money data are reasonably reliable for longer periods, such as a year. The moral is that ***we probably should not pay much attention to short-run movements in the money supply numbers, but should be concerned only with longer-run movements.***

Table 2 Growth Rate of M2: Initial and Revised Series, 2002 (percent, compounded annual rate)

Period	Initial Rate	Revised Rate	Difference (Revised Rate – Initial Rate)
January	2.2	5.4	3.2
February	6.8	8.7	1.9
March	–1.4	0.2	1.6
April	–4.0	–2.6	1.4
May	14.8	15.4	0.6
June	7.6	7.1	–0.5
July	13.6	11.0	–2.6
August	9.9	8.6	–1.3
September	5.1	5.7	0.6
October	10.9	8.3	–2.6
November	10.2	8.0	–2.2
December	2.8	2.8	0.0
Average	6.5	6.5	0.0

Source: Federal Reserve Statistical Release H.6: www.federalreserve.gov/releases/h6.

Summary

1. To economists, the word *money* has a different meaning from *income* or *wealth*. Money is anything that is generally accepted as payment for goods or services or in the repayment of debts.
2. Money serves three primary functions: as a medium of exchange, as a unit of account, and as a store of value. Money as a medium of exchange avoids the problem of double coincidence of wants that arises in a barter economy by lowering transaction costs and encouraging specialization and the division of labor. Money as a unit of account reduces the number of prices needed in the economy, which also reduces transaction costs. Money also functions as a store of value, but performs this role poorly if it is rapidly losing value due to inflation.
3. The payments system has evolved over time. Until several hundred years ago, the payments system in all but the most primitive societies was based primarily on precious metals. The introduction of paper currency lowered the cost of transporting money. The next major advance was the introduction of checks, which lowered transaction costs still further. We are currently moving toward an electronic payments system in which paper is eliminated and all transactions are handled by computers. Despite the potential efficiency of such a system, obstacles are slowing the movement to the checkless society and the development of new forms of electronic money.
4. The Federal Reserve System has defined three different measures of the money supply—M1, M2, and M3. These measures are not equivalent and do not always move together, so they cannot be used interchangeably by policymakers. Obtaining the precise, correct measure of money does seem to matter and has implications for the conduct of monetary policy.
5. Another problem in the measurement of money is that the data are not always as accurate as we would like.

Substantial revisions in the data do occur; they indicate that initially released money data are not a reliable guide to short-run (say, month-to-month) movements in the money supply, although they are more reliable over longer periods of time, such as a year.

Key Terms

commodity money, p. 48
currency, p. 44
e-cash, p. 51
electronic money (e-money), p. 51
fiat money, p. 48
hyperinflation, p. 47
income, p. 45
liquidity, p. 47
M1, p. 52
M2, p. 52
M3, p. 53
medium of exchange, p. 45
monetary aggregates, p. 52
payments system, p. 48
smart card, p. 51
store of value, p. 47
unit of account, p. 46
wealth, p. 45

Questions and Problems

Questions marked with an asterisk are answered at the end of the book in an appendix, "Answers to Selected Questions and Problems."

1. Which of the following three expressions uses the economists' definition of money?
   a. "How much money did you earn last week?"
   b. "When I go to the store, I always make sure that I have enough money."
   c. "The love of money is the root of all evil."

*2. There are three goods produced in an economy by three individuals:

Good	*Producer*
Apples	Orchard owner
Bananas	Banana grower
Chocolate	Chocolatier

If the orchard owner likes only bananas, the banana grower likes only chocolate, and the chocolatier likes only apples, will any trade between these three persons take place in a barter economy? How will introducing money into the economy benefit these three producers?

3. Why did cavemen not need money?

*4. Why were people in the United States in the nineteenth century sometimes willing to be paid by check rather than with gold, even though they knew that there was a possibility that the check might bounce?

5. In ancient Greece, why was gold a more likely candidate for use as money than wine was?

*6. Was money a better store of value in the United States in the 1950s than it was in the 1970s? Why or why not? In which period would you have been more willing to hold money?

7. Would you be willing to give up your checkbook and instead use an electronic means of payment if it were made available? Why or why not?

8. Rank the following assets from most liquid to least liquid:
   a. Checking account deposits
   b. Houses
   c. Currency
   d. Washing machines
   e. Savings deposits
   f. Common stock

*9. Why have some economists described money during a hyperinflation as a "hot potato" that is quickly passed from one person to another?

10. In Brazil, a country that was undergoing a rapid inflation before 1994, many transactions were conducted in dollars rather than in reals, the domestic currency. Why?

*11. Suppose that a researcher discovers that a measure of the total amount of debt in the U.S. economy over the past 20 years was a better predictor of inflation and the business cycle than M1, M2, or M3. Does this discovery mean that we should define money as equal to the total amount of debt in the economy?

12. Look up the M1, M2, and M3 numbers in the Federal Reserve *Bulletin* for the most recent one-year period. Have their growth rates been similar? What implications do their growth rates have for the conduct of monetary policy?

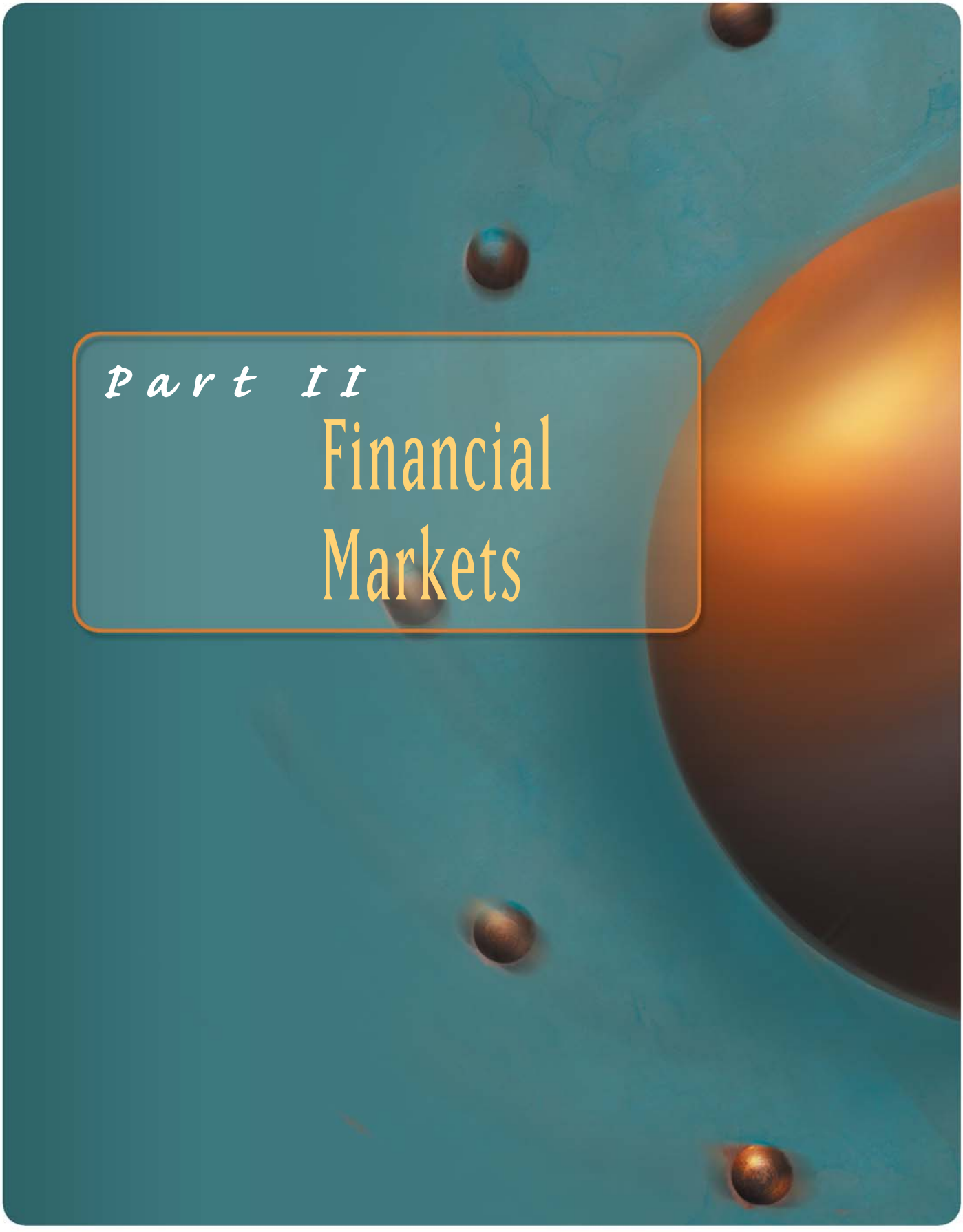
*13. Which of the Federal Reserve's measures of the monetary aggregates, M1, M2, or M3, is composed of the most liquid assets? Which is the largest measure?

14. For each of the following assets, indicate which of the monetary aggregates (M1, M2, M3) includes them:
    a. Currency
    b. Money market mutual funds
    c. Eurodollars
    d. Small-denomination time deposits
    e. Large-denomination repurchase agreements
    f. Checkable deposits

*15. Why are revisions of monetary aggregates less of a problem for measuring long-run movements of the money supply than they are for measuring short-run movements?

Web Exercises

1. Go to www.federalreserve.gov/releases/h6/Current/.
    a. What has been the growth rate in M1, M2, and M3 over the last 12 months?
    b. From what you know about the state of the economy, does this seem expansionary or restrictive?
2. Go to www.federalreserve.gov/paymentsys.htm and select one topic on which the Federal Reserve has a written policy. Write a one-paragraph summary of this policy.


Part II
Financial
Markets

Chapter 4

Understanding Interest Rates

PREVIEW

Interest rates are among the most closely watched variables in the economy. Their movements are reported almost daily by the news media, because they directly affect our everyday lives and have important consequences for the health of the economy. They affect personal decisions such as whether to consume or save, whether to buy a house, and whether to purchase bonds or put funds into a savings account. Interest rates also affect the economic decisions of businesses and households, such as whether to use their funds to invest in new equipment for factories or to save their money in a bank.

Before we can go on with the study of money, banking, and financial markets, we must understand exactly what the phrase *interest rates* means. In this chapter, we see that a concept known as the *yield to maturity* is the most accurate measure of interest rates; the yield to maturity is what economists mean when they use the term *interest rate*. We discuss how the yield to maturity is measured and examine alternative (but less accurate) ways in which interest rates are quoted. We'll also see that a bond's interest rate does not necessarily indicate how good an investment the bond is because what it earns (its rate of return) does not necessarily equal its interest rate. Finally, we explore the distinction between real interest rates, which are adjusted for inflation, and nominal interest rates, which are not.

Although learning definitions is not always the most exciting of pursuits, it is important to read carefully and understand the concepts presented in this chapter. Not only are they continually used throughout the remainder of this text, but a firm grasp of these terms will give you a clearer understanding of the role that interest rates play in your life as well as in the general economy.

www.bloomberg.com/markets/

Under "Rates & Bonds," you can access information on key interest rates, U.S. Treasuries, Government bonds, and municipal bonds.

Measuring Interest Rates

Different debt instruments have very different streams of payment with very different timing. Thus we first need to understand how we can compare the value of one kind of debt instrument with another before we see how interest rates are measured. To do this, we make use of the concept of *present value*.

Present Value

The concept of **present value** (or **present discounted value**) is based on the common-sense notion that a dollar paid to you one year from now is less valuable to you than a dollar paid to you today: This notion is true because you can deposit a dollar in a

A credit market instrument providing the borrower with an amount of funds that must be repaid to the lender at the maturity date along with an additional payment (interest).

savings account that earns interest and have more than a dollar in one year. Economists use a more formal definition, as explained in this section.

Let's look at the simplest kind of debt instrument, which we will call a **simple loan**. In this loan, the lender provides the borrower with an amount of funds (called the *principal*) that must be repaid to the lender at the *maturity date*, along with an additional payment for the interest. For example, if you made your friend, Jane, a simple loan of $100 for one year, you would require her to repay the principal of $100 in one year's time along with an additional payment for interest; say, $10. In the case of a simple loan like this one, the interest payment divided by the amount of the loan is a natural and sensible way to measure the interest rate. This measure of the so-called *simple interest rate*, i, is:

$$i = \frac{\$10}{\$100} = 0.10 = 10\%$$

If you make this $100 loan, at the end of the year you would have $110, which can be rewritten as:

$$\$100 \times (1 + 0.10) = \$110$$

If you then lent out the $110, at the end of the second year you would have:

$$\$110 \times (1 + 0.10) = \$121$$

or, equivalently,

$$\$100 \times (1 + 0.10) \times (1 + 0.10) = \$100 \times (1 + 0.10)^2 = \$121$$

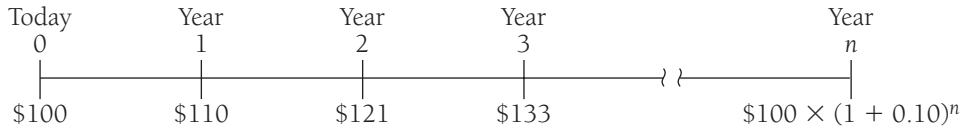
Continuing with the loan again, you would have at the end of the third year:

$$\$121 \times (1 + 0.10) = \$100 \times (1 + 0.10)^3 = \$133$$

Generalizing, we can see that at the end of n years, your $100 would turn into:

$$\$100 \times (1 + i)^n$$

The amounts you would have at the end of each year by making the $100 loan today can be seen in the following timeline:



This timeline immediately tells you that you are just as happy having $100 today as having $110 a year from now (of course, as long as you are sure that Jane will pay you back). Or that you are just as happy having $100 today as having $121 two years from now, or $133 three years from now or $100 × (1 + 0.10)^n, n years from now. The timeline tells us that we can also work backward from future amounts to the present: for example, $133 = $100 × (1 + 0.10)^3 three years from now is worth $100 today, so that:

$$\$100 = \frac{\$133}{(1 + 0.10)^3}$$

The process of calculating today's value of dollars received in the future, as we have done above, is called *discounting the future*. We can generalize this process by writing

today's (present) value of $100 as *PV*, the future value of $133 as *FV*, and replacing 0.10 (the 10% interest rate) by *i*. This leads to the following formula:

$$PV = \frac{FV}{(1 + i)^n} \quad (1)$$

Intuitively, what Equation 1 tells us is that if you are promised $1 for certain ten years from now, this dollar would not be as valuable to you as $1 is today because if you had the $1 today, you could invest it and end up with more than $1 in ten years.

The concept of present value is extremely useful, because it allows us to figure out today's value (price) of a credit market instrument at a given simple interest rate *i* by just adding up the individual present values of all the future payments received. This information allows us to compare the value of two instruments with very different timing of their payments.

As an example of how the present value concept can be used, let's assume that you just hit the $20 million jackpot in the New York State Lottery, which promises you a payment of $1 million for the next twenty years. You are clearly excited, but have you really won $20 million? No, not in the present value sense. In today's dollars, that $20 million is worth a lot less. If we assume an interest rate of 10% as in the earlier examples, the first payment of $1 million is clearly worth $1 million today, but the next payment next year is only worth $1 million/$(1 + 0.10) = \$909{,}090$, a lot less than $1 million. The following year the payment is worth $1 million/$(1 + 0.10)^2 = \$826{,}446$ in today's dollars, and so on. When you add all these up, they come to $9.4 million. You are still pretty excited (who wouldn't be?), but because you understand the concept of present value, you recognize that you are the victim of false advertising. You didn't really win $20 million, but instead won less than half as much.

Four Types of Credit Market Instruments

In terms of the timing of their payments, there are four basic types of credit market instruments.

1. A simple loan, which we have already discussed, in which the lender provides the borrower with an amount of funds, which must be repaid to the lender at the maturity date along with an additional payment for the interest. Many money market instruments are of this type: for example, commercial loans to businesses.

2. A **fixed-payment loan** (which is also called a **fully amortized loan**) in which the lender provides the borrower with an amount of funds, which must be repaid by making the same payment every period (such as a month), consisting of part of the principal and interest for a set number of years. For example, if you borrowed $1,000, a fixed-payment loan might require you to pay $126 every year for 25 years. Installment loans (such as auto loans) and mortgages are frequently of the fixed-payment type.

3. A **coupon bond** pays the owner of the bond a fixed interest payment (coupon payment) every year until the maturity date, when a specified final amount (**face value** or **par value**) is repaid. The coupon payment is so named because the bondholder used to obtain payment by clipping a coupon off the bond and sending it to the bond issuer, who then sent the payment to the holder. Nowadays, it is no longer necessary to send in coupons to receive these payments. A coupon bond with $1,000 face value, for example, might pay you a coupon payment of $100 per year for ten years, and at the maturity date repay you the face value amount of $1,000. (The face value of a bond is usually in $1,000 increments.)

A coupon bond is identified by three pieces of information. First is the corporation or government agency that issues the bond. Second is the maturity date of the

bond. Third is the bond's **coupon rate**, the dollar amount of the yearly coupon payment expressed as a percentage of the face value of the bond. In our example, the coupon bond has a yearly coupon payment of $100 and a face value of $1,000. The coupon rate is then $100/$1,000 = 0.10, or 10%. Capital market instruments such as U.S. Treasury bonds and notes and corporate bonds are examples of coupon bonds.

4. A **discount bond** (also called a **zero-coupon bond**) is bought at a price below its face value (at a discount), and the face value is repaid at the maturity date. Unlike a coupon bond, a discount bond does not make any interest payments; it just pays off the face value. For example, a discount bond with a face value of $1,000 might be bought for $900; in a year's time the owner would be repaid the face value of $1,000. U.S. Treasury bills, U.S. savings bonds, and long-term zero-coupon bonds are examples of discount bonds.

These four types of instruments require payments at different times: Simple loans and discount bonds make payment only at their maturity dates, whereas fixed-payment loans and coupon bonds have payments periodically until maturity. How would you decide which of these instruments provides you with more income? They all seem so different because they make payments at different times. To solve this problem, we use the concept of present value, explained earlier, to provide us with a procedure for measuring interest rates on these different types of instruments.

Yield to Maturity

Of the several common ways of calculating interest rates, the most important is the **yield to maturity**, the interest rate that equates the present value of payments received from a debt instrument with its value today.[1] Because the concept behind the calculation of the yield to maturity makes good economic sense, economists consider it the most accurate measure of interest rates.

To understand the yield to maturity better, we now look at how it is calculated for the four types of credit market instruments.

Simple Loan. Using the concept of present value, the yield to maturity on a simple loan is easy to calculate. For the one-year loan we discussed, today's value is $100, and the payments in one year's time would be $110 (the repayment of $100 plus the interest payment of $10). We can use this information to solve for the yield to maturity i by recognizing that the present value of the future payments must equal today's value of a loan. Making today's value of the loan ($100) equal to the present value of the $110 payment in a year (using Equation 1) gives us:

$$\$100 = \frac{\$110}{1 + i}$$

Solving for i,

$$i = \frac{\$110 - \$100}{\$100} = \frac{\$10}{\$100} = 0.10 = 10\%$$

This calculation of the yield to maturity should look familiar, because it equals the interest payment of $10 divided by the loan amount of $100; that is, it equals the simple interest rate on the loan. An important point to recognize is that *for simple loans, the simple interest rate equals the yield to maturity*. Hence the same term i is used to denote both the yield to maturity and the simple interest rate.

[1] In other contexts, it is also called the *internal rate of return*.

Study Guide

The key to understanding the calculation of the yield to maturity is equating today's value of the debt instrument with the present value of all of its future payments. The best way to learn this principle is to apply it to other specific examples of the four types of credit market instruments in addition to those we discuss here. See if you can develop the equations that would allow you to solve for the yield to maturity in each case.

Fixed-Payment Loan. Recall that this type of loan has the same payment every period throughout the life of the loan. On a fixed-rate mortgage, for example, the borrower makes the same payment to the bank every month until the maturity date, when the loan will be completely paid off. To calculate the yield to maturity for a fixed-payment loan, we follow the same strategy we used for the simple loan—we equate today's value of the loan with its present value. Because the fixed-payment loan involves more than one payment, the present value of the fixed-payment loan is calculated as the sum of the present values of all payments (using Equation 1).

In the case of our earlier example, the loan is \$1,000 and the yearly payment is \$126 for the next 25 years. The present value is calculated as follows: At the end of one year, there is a \$126 payment with a *PV* of \$126/$(1 + i)$; at the end of two years, there is another \$126 payment with a *PV* of \$126/$(1 + i)^2$; and so on until at the end of the twenty-fifth year, the last payment of \$126 with a *PV* of \$126/$(1 + i)^{25}$ is made. Making today's value of the loan (\$1,000) equal to the sum of the present values of all the yearly payments gives us:

$$\$1{,}000 = \frac{\$126}{1+i} + \frac{\$126}{(1+i)^2} + \frac{\$126}{(1+i)^3} + \cdots + \frac{\$126}{(1+i)^{25}}$$

More generally, for any fixed-payment loan,

$$LV = \frac{FP}{1+i} + \frac{FP}{(1+i)^2} + \frac{FP}{(1+i)^3} + \cdots + \frac{FP}{(1+i)^n} \qquad (2)$$

where LV = loan value
FP = fixed yearly payment
n = number of years until maturity

For a fixed-payment loan amount, the fixed yearly payment and the number of years until maturity are known quantities, and only the yield to maturity is not. So we can solve this equation for the yield to maturity i. Because this calculation is not easy, many pocket calculators have programs that allow you to find i given the loan's numbers for LV, FP, and n. For example, in the case of the 25-year loan with yearly payments of \$126, the yield to maturity that solves Equation 2 is 12%. Real estate brokers always have a pocket calculator that can solve such equations so that they can immediately tell the prospective house buyer exactly what the yearly (or monthly) payments will be if the house purchase is financed by taking out a mortgage.[2]

Coupon Bond. To calculate the yield to maturity for a coupon bond, follow the same strategy used for the fixed-payment loan: Equate today's value of the bond with its present value. Because coupon bonds also have more than one payment, the present

[2]The calculation with a pocket calculator programmed for this purpose requires simply that you enter the value of the loan *LV,* the number of years to maturity *n,* and the interest rate *i* and then run the program.

value of the bond is calculated as the sum of the present values of all the coupon payments plus the present value of the final payment of the face value of the bond.

The present value of a $1,000-face-value bond with ten years to maturity and yearly coupon payments of $100 (a 10% coupon rate) can be calculated as follows: At the end of one year, there is a $100 coupon payment with a *PV* of $100/(1 + *i*); at the end of the second year, there is another $100 coupon payment with a *PV* of $100/(1 + *i*)^2; and so on until at maturity, there is a $100 coupon payment with a *PV* of $100/(1 + *i*)^{10} plus the repayment of the $1,000 face value with a *PV* of $1,000/(1 + *i*)^{10}. Setting today's value of the bond (its current price, denoted by *P*) equal to the sum of the present values of all the payments for this bond gives:

$$P = \frac{\$100}{1+i} + \frac{\$100}{(1+i)^2} + \frac{\$100}{(1+i)^3} + \cdots + \frac{\$100}{(1+i)^{10}} + \frac{\$1{,}000}{(1+i)^{10}}$$

More generally, for any coupon bond,[3]

$$P = \frac{C}{1+i} + \frac{C}{(1+i)^2} + \frac{C}{(1+i)^3} + \cdots + \frac{C}{(1+i)^n} + \frac{F}{(1+i)^n} \qquad (3)$$

where
P = price of coupon bond
C = yearly coupon payment
F = face value of the bond
n = years to maturity date

In Equation 3, the coupon payment, the face value, the years to maturity, and the price of the bond are known quantities, and only the yield to maturity is not. Hence we can solve this equation for the yield to maturity *i*. Just as in the case of the fixed-payment loan, this calculation is not easy, so business-oriented pocket calculators have built-in programs that solve this equation for you.[4]

Let's look at some examples of the solution for the yield to maturity on our 10%-coupon-rate bond that matures in ten years. If the purchase price of the bond is $1,000, then either using a pocket calculator with the built-in program or looking at a bond table, we will find that the yield to maturity is 10 percent. If the price is $900, we find that the yield to maturity is 11.75%. Table 1 shows the yields to maturity calculated for several bond prices.

Table 1 Yields to Maturity on a 10%-Coupon-Rate Bond Maturing in Ten Years (Face Value = $1,000)

Price of Bond ($)	Yield to Maturity (%)
1,200	7.13
1,100	8.48
1,000	10.00
900	11.75
800	13.81

[3]Most coupon bonds actually make coupon payments on a semiannual basis rather than once a year as assumed here. The effect on the calculations is only very slight and will be ignored here.

[4]The calculation of a bond's yield to maturity with the programmed pocket calculator requires simply that you enter the amount of the yearly coupon payment *C*, the face value *F*, the number of years to maturity *n*, and the price of the bond *P* and then run the program.

Three interesting facts are illustrated by Table 1:

1. When the coupon bond is priced at its face value, the yield to maturity equals the coupon rate.
2. The price of a coupon bond and the yield to maturity are negatively related; that is, as the yield to maturity rises, the price of the bond falls. As the yield to maturity falls, the price of the bond rises.
3. The yield to maturity is greater than the coupon rate when the bond price is below its face value.

These three facts are true for any coupon bond and are really not surprising if you think about the reasoning behind the calculation of the yield to maturity. When you put $1,000 in a bank account with an interest rate of 10%, you can take out $100 every year and you will be left with the $1,000 at the end of ten years. This is similar to buying the $1,000 bond with a 10% coupon rate analyzed in Table 1, which pays a $100 coupon payment every year and then repays $1,000 at the end of ten years. If the bond is purchased at the par value of $1,000, its yield to maturity must equal 10%, which is also equal to the coupon rate of 10%. The same reasoning applied to any coupon bond demonstrates that if the coupon bond is purchased at its par value, the yield to maturity and the coupon rate must be equal.

It is straightforward to show that the bond price and the yield to maturity are negatively related. As i, the yield to maturity, rises, all denominators in the bond price formula must necessarily rise. Hence a rise in the interest rate as measured by the yield to maturity means that the price of the bond must fall. Another way to explain why the bond price falls when the interest rises is that a higher interest rate implies that the future coupon payments and final payment are worth less when discounted back to the present; hence the price of the bond must be lower.

There is one special case of a coupon bond that is worth discussing because its yield to maturity is particularly easy to calculate. This bond is called a **consol** or a **perpetuity**; it is a perpetual bond with no maturity date and no repayment of principal that makes fixed coupon payments of $\$C$ forever. Consols were first sold by the British Treasury during the Napoleonic Wars and are still traded today; they are quite rare, however, in American capital markets. The formula in Equation 3 for the price of the consol P simplifies to the following:[5]

$$P = \frac{C}{i} \tag{4}$$

---

[5]The bond price formula for a consol is:

$$P = \frac{C}{1+i} + \frac{C}{(1+i)^2} + \frac{C}{(1+i)^3} + \cdots$$

which can be written as:

$$P = C\,(x + x^2 + x^3 + \cdots)$$

in which $x = 1/(1+i)$. The formula for an infinite sum is:

$$1 + x + x^2 + x^3 + \cdots = \frac{1}{1-x} \quad \text{for} \quad x < 1$$

and so:

$$P = C\left(\frac{1}{1-x} - 1\right) = C\left[\frac{1}{1 - 1/(1+i)} - 1\right]$$

which by suitable algebraic manipulation becomes:

$$P = C\left(\frac{1+i}{i} - \frac{i}{i}\right) = \frac{C}{i}$$

where P = price of the consol
C = yearly payment

One nice feature of consols is that you can immediately see that as i goes up, the price of the bond falls. For example, if a consol pays \$100 per year forever and the interest rate is 10%, its price will be \$1,000 = \$100/0.10. If the interest rate rises to 20%, its price will fall to \$500 = \$100/0.20. We can also rewrite this formula as

$$i = \frac{C}{P} \tag{5}$$

We see then that it is also easy to calculate the yield to maturity for the consol (despite the fact that it never matures). For example, with a consol that pays \$100 yearly and has a price of \$2,000, the yield to maturity is easily calculated to be 5% (= \$100/\$2,000).

Discount Bond. The yield-to-maturity calculation for a discount bond is similar to that for the simple loan. Let us consider a discount bond such as a one-year U.S. Treasury bill, which pays off a face value of \$1,000 in one year's time. If the current purchase price of this bill is \$900, then equating this price to the present value of the \$1,000 received in one year, using Equation 1, gives:

$$\$900 = \frac{\$1{,}000}{1 + i}$$

and solving for i,

$$\begin{aligned}
(1 + i) \times \$900 &= \$1{,}000 \\
\$900 + \$900i &= \$1{,}000 \\
\$900i &= \$1{,}000 - \$900 \\
i &= \frac{\$1{,}000 - \$900}{\$900} = 0.111 = 11.1\%
\end{aligned}$$

More generally, for any one-year discount bond, the yield to maturity can be written as:

$$i = \frac{F - P}{P} \tag{6}$$

where F = face value of the discount bond
P = current price of the discount bond

In other words, the yield to maturity equals the increase in price over the year $F - P$ divided by the initial price P. In normal circumstances, investors earn positive returns from holding these securities and so they sell at a discount, meaning that the current price of the bond is below the face value. Therefore, $F - P$ should be positive, and the yield to maturity should be positive as well. However, this is not always the case, as recent extraordinary events in Japan indicate (see Box 1).

An important feature of this equation is that it indicates that for a discount bond, the yield to maturity is negatively related to the current bond price. This is the same conclusion that we reached for a coupon bond. For example, Equation 6 shows that a rise in the bond price from \$900 to \$950 means that the bond will have a smaller

Box 1: Global

Negative T-Bill Rates? Japan Shows the Way

We normally assume that interest rates must always be positive. Negative interest rates would imply that you are willing to pay more for a bond today than you will receive for it in the future (as our formula for yield to maturity on a discount bond demonstrates). Negative interest rates therefore seem like an impossibility because you would do better by holding cash that has the same value in the future as it does today.

The Japanese have demonstrated that this reasoning is not quite correct. In November 1998, interest rates on Japanese six-month Treasury bills became negative, yielding an interest rate of –0.004%, with investors paying more for the bills than their face value. This is an extremely unusual event—no other country in the world has seen negative interest rates during the last fifty years. How could this happen?

As we will see in Chapter 5, the weakness of the Japanese economy and a negative inflation rate drove Japanese interest rates to low levels, but these two factors can't explain the negative rates. The answer is that large investors found it more convenient to hold these six-month bills as a store of value rather than holding cash because the bills are denominated in larger amounts and can be stored electronically. For that reason, some investors were willing to hold them, despite their negative rates, even though in monetary terms the investors would be better off holding cash. Clearly, the convenience of T-bills goes only so far, and thus their interest rates can go only a little bit below zero.

increase in its price at maturity, and the yield to maturity falls from 11.1 to 5.3%. Similarly, a fall in the yield to maturity means that the price of the discount bond has risen.

Summary. The concept of present value tells you that a dollar in the future is not as valuable to you as a dollar today because you can earn interest on this dollar. Specifically, a dollar received n years from now is worth only $\$1/(1 + i)^n$ today. The present value of a set of future payments on a debt instrument equals the sum of the present values of each of the future payments. The yield to maturity for an instrument is the interest rate that equates the present value of the future payments on that instrument to its value today. Because the procedure for calculating the yield to maturity is based on sound economic principles, this is the measure that economists think most accurately describes the interest rate.

Our calculations of the yield to maturity for a variety of bonds reveal the important fact that ***current bond prices and interest rates are negatively related: When the interest rate rises, the price of the bond falls, and vice versa***.

www.teachmefinance.com

A review of the key financial concepts: time value of money, annuities, perpetuities, and so on.

Other Measures of Interest Rates

The yield to maturity is the most accurate measure of interest rates; this is what economists mean when they use the term *interest rate*. Unless otherwise specified, the terms *interest rate* and *yield to maturity* are used synonymously in this book. However, because the yield to maturity is sometimes difficult to calculate, other, less accurate

measures of interest rates have come into common use in bond markets. You will frequently encounter two of these measures—the *current yield* and the *yield on a discount basis*—when reading the newspaper, and it is important for you to understand what they mean and how they differ from the more accurate measure of interest rates, the yield to maturity.

Current Yield

The **current yield** is an approximation of the yield to maturity on coupon bonds that is often reported, because in contrast to the yield to maturity, it is easily calculated. It is defined as the yearly coupon payment divided by the price of the security,

$$i_c = \frac{C}{P} \tag{7}$$

where i_c = current yield
P = price of the coupon bond
C = yearly coupon payment

This formula is identical to the formula in Equation 5, which describes the calculation of the yield to maturity for a consol. Hence, for a consol, the current yield is an exact measure of the yield to maturity. When a coupon bond has a long term to maturity (say, 20 years or more), it is very much like a consol, which pays coupon payments forever. Thus you would expect the current yield to be a rather close approximation of the yield to maturity for a long-term coupon bond, and you can safely use the current-yield calculation instead of calculating the yield to maturity with a financial calculator. However, as the time to maturity of the coupon bond shortens (say, it becomes less than five years), it behaves less and less like a consol and so the approximation afforded by the current yield becomes worse and worse.

We have also seen that when the bond price equals the par value of the bond, the yield to maturity is equal to the coupon rate (the coupon payment divided by the par value of the bond). Because the current yield equals the coupon payment divided by the bond price, the current yield is also equal to the coupon rate when the bond price is at par. This logic leads us to the conclusion that when the bond price is at par, the current yield equals the yield to maturity. This means that the closer the bond price is to the bond's par value, the better the current yield will approximate the yield to maturity.

The current yield is negatively related to the price of the bond. In the case of our 10%-coupon-rate bond, when the price rises from \$1,000 to \$1,100, the current yield falls from 10% (= \$100/\$1,000) to 9.09% (= \$100/\$1,100). As Table 1 indicates, the yield to maturity is also negatively related to the price of the bond; when the price rises from \$1,000 to \$1,100, the yield to maturity falls from 10 to 8.48%. In this we see an important fact: The current yield and the yield to maturity always move together; a rise in the current yield always signals that the yield to maturity has also risen.

The general characteristics of the current yield (the yearly coupon payment divided by the bond price) can be summarized as follows: The current yield better approximates the yield to maturity when the bond's price is nearer to the bond's par value and the maturity of the bond is longer. It becomes a worse approximation when the bond's price is further from the bond's par value and the bond's maturity is shorter. Regardless of whether the current yield is a good approximation of the yield to maturity, a change in the current yield *always* signals a change in the same direction of the yield to maturity.

Yield on a Discount Basis

Before the advent of calculators and computers, dealers in U.S. Treasury bills found it difficult to calculate interest rates as a yield to maturity. Instead, they quoted the interest rate on bills as a **yield on a discount basis** (or **discount yield**), and they still do so today. Formally, the yield on a discount basis is defined by the following formula:

$$i_{db} = \frac{F - P}{F} \times \frac{360}{\text{days to maturity}} \tag{8}$$

where i_{db} = yield on a discount basis
F = face value of the discount bond
P = purchase price of the discount bond

This method for calculating interest rates has two peculiarities. First, it uses the percentage gain on the face value of the bill $(F - P)/F$ rather than the percentage gain on the purchase price of the bill $(F - P)/P$ used in calculating the yield to maturity. Second, it puts the yield on an annual basis by considering the year to be 360 days long rather than 365 days.

Because of these peculiarities, the discount yield understates the interest rate on bills as measured by the yield to maturity. On our one-year bill, which is selling for \$900 and has a face value of \$1,000, the yield on a discount basis would be as follows:

$$i_{db} = \frac{\$1{,}000 - \$900}{\$1{,}000} \times \frac{360}{365} = 0.099 = 9.9\%$$

whereas the yield to maturity for this bill, which we calculated before, is 11.1%. The discount yield understates the yield to maturity by a factor of over 10%. A little more than 1% ([365 − 360]/360 = 0.014 = 1.4%) can be attributed to the understatement of the length of the year: When the bill has one year to maturity, the second term on the right-hand side of the formula is 360/365 = 0.986 rather than 1.0, as it should be.

The more serious source of the understatement, however, is the use of the percentage gain on the face value rather than on the purchase price. Because, by definition, the purchase price of a discount bond is always less than the face value, the percentage gain on the face value is necessarily smaller than the percentage gain on the purchase price. The greater the difference between the purchase price and the face value of the discount bond, the more the discount yield understates the yield to maturity. Because the difference between the purchase price and the face value gets larger as maturity gets longer, we can draw the following conclusion about the relationship of the yield on a discount basis to the yield to maturity: The yield on a discount basis always understates the yield to maturity, and this understatement becomes more severe the longer the maturity of the discount bond.

Another important feature of the discount yield is that, like the yield to maturity, it is negatively related to the price of the bond. For example, when the price of the bond rises from \$900 to \$950, the formula indicates that the yield on a discount basis declines from 9.9 to 4.9%. At the same time, the yield to maturity declines from 11.1 to 5.3%. Here we see another important factor about the relationship of yield on a discount basis to yield to maturity: They always move together. That is, a rise in the discount yield always means that the yield to maturity has risen, and a decline in the discount yield means that the yield to maturity has declined as well.

The characteristics of the yield on a discount basis can be summarized as follows: Yield on a discount basis understates the more accurate measure of the interest rate, the yield to maturity; and the longer the maturity of the discount bond, the greater

this understatement becomes. Even though the discount yield is a somewhat misleading measure of the interest rates, a change in the discount yield always indicates a change in the same direction for the yield to maturity.

Application

Reading the *Wall Street Journal*: The Bond Page

Now that we understand the different interest-rate definitions, let's apply our knowledge and take a look at what kind of information appears on the bond page of a typical newspaper, in this case the *Wall Street Journal*. The "Following the Financial News" box contains the *Journal*'s listing for three different types of bonds on Wednesday, January 23, 2003. Panel (a) contains the information on U.S. Treasury bonds and notes. Both are coupon bonds, the only difference being their time to maturity from when they were originally issued: Notes have a time to maturity of less than ten years; bonds have a time to maturity of more than ten years.

The information found in the "Rate" and "Maturity" columns identifies the bond by coupon rate and maturity date. For example, T-bond 1 has a coupon rate of 4.75%, indicating that it pays out $47.50 per year on a $1,000-face-value bond and matures in January 2003. In bond market parlance, it is referred to as the Treasury's $4\frac{3}{4}$s of 2003. The next three columns tell us about the bond's price. By convention, all prices in the bond market are quoted per $100 of face value. Furthermore, the numbers after the colon represent thirty-seconds ($^{x}/_{32}$, or 32nds). In the case of T-bond 1, the first price of 100:02 represents $100\frac{2}{32} = 100.0625$, or an actual price of $1000.62 for a $1,000-face-value bond. The bid price tells you what price you will receive if you sell the bond, and the asked price tells you what you must pay for the bond. (You might want to think of the bid price as the "wholesale" price and the asked price as the "retail" price.) The "Chg." column indicates how much the bid price has changed in 32nds (in this case, no change) from the previous trading day.

Notice that for all the bonds and notes, the asked price is more than the bid price. Can you guess why this is so? The difference between the two (the *spread*) provides the bond dealer who trades these securities with a profit. For T-bond 1, the dealer who buys it at $100\frac{2}{32}$, and sells it for $100\frac{3}{32}$, makes a profit of $\frac{1}{32}$. This profit is what enables the dealer to make a living and provide the service of allowing you to buy and sell bonds at will.

The "Ask Yld." column provides the yield to maturity, which is 0.43% for T-bond 1. It is calculated with the method described earlier in this chapter using the asked price as the price of the bond. The asked price is used in the calculation because the yield to maturity is most relevant to a person who is going to buy and hold the security and thus earn the yield. The person selling the security is not going to be holding it and hence is less concerned with the yield.

The figure for the current yield is not usually included in the newspaper's quotations for Treasury securities, but it has been added in panel (a) to give you some real-world examples of how well the current yield approximates

Following the Financial News

Bond Prices and Interest Rates

Bond prices and interest rates are published daily. In the *Wall Street Journal*, they can be found in the "NYSE/AMEX Bonds" and "Treasury/Agency Issues" section of the paper. Three basic formats for quoting bond prices and yields are illustrated here.

TREASURY BONDS, NOTES AND BILLS

January 22, 2003

Representative Over-the-Counter quotation based on transactions of $1 million or more.

Treasury bond, note and bill quotes are as of mid-afternoon. Colons in bid-and-asked quotes represent 32nds; 101:01 means 101 1/32. Net changes in 32nds. n-Treasury note. i-Inflation-Indexed issue. Treasury bill quotes in hundredths, quoted on terms of a rate of discount. Days to maturity calculated from settlement date. All yields are to maturity and based on the asked quote. Latest 13-week and 26-week bills are boldfaced. For bonds callable prior to maturity, yields are computed to the earliest call date for issues quoted above par and to the maturity date for issues below par. *When issued.

Source: eSpeed/Cantor Fitzgerald

U.S. Treasury strips as of 3 p.m. Eastern time, also based on transactions of $1 million or more. Colons in bid and asked quotes represent 32nds; 99:01 means 99 1/32. Net changes in 32nds. Yields calculated on the asked quotation. ci-stripped coupon interest. bp-Treasury bond, stripped principal. np-Treasury note, stripped principal. For bonds callable prior to maturity, yields are computed to the earliest call date for issues quoted above par and to the maturity date for issues below par.

Source: Bear, Stearns & Co. via Street Software Technology, Inc.

(a) Treasury bonds and notes

GOVT. BONDS & NOTES

	Rate	Maturity Mo/Yr	Bid	Asked	Chg.	Ask Yld.	
T-bond 1	4.750	Jan 03n	100:02	100:03	. . .	0.43	Current Yield = 4.75%
	5.500	Jan 03n	100:02	100:03	–1	0.46	
	5.750	Aug 03n	102:17	102:18	. . .	0.16	
T-bond 2	11.125	Aug 03	105:16	105:17	–1	1.22	Current Yield = 10.55%
T-bond 3	5.250	Feb 29	103:17	103:18	23	5.00	Current Yield = 5.07%
	3.875	Apr 29i	122:03	122:04	2	2.69	
	6.125	Aug 29	116:10	116:11	24	5.00	
T-bond 4	5.375	Feb 31	107:27	107:28	24	4.86	Current Yield = 4.98%

(b) Treasury bills

TREASURY BILLS

Maturity	Days to Mat.	Bid	Asked	Chg.	Ask Yld.	Maturity	Days to Mat.	Bid	Asked	Chg.	Ask Yld.
Jan 30 03	7	1.15	1.14	-0.01	1.16	May 01 03	98	1.14	1.13	-0.02	1.15
Feb 06 03	14	1.14	1.13	-0.01	1.15	May 08 03	105	1.14	1.13	-0.03	1.15
Feb 13 03	21	1.14	1.13	-0.01	1.15	May 15 03	112	1.15	1.14	-0.02	1.16
Feb 20 03	28	1.14	1.13	. . .	1.15	May 22 03	119	1.15	1.14	-0.02	1.16
Feb 27 03	35	1.13	1.12	-0.01	1.14	May 29 03	126	1.15	1.14	-0.01	1.16
Mar 06 03	42	1.13	1.12	. . .	1.14	Jun 05 03	133	1.15	1.14	-0.02	1.16
Mar 13 03	49	1.13	1.12	-0.01	1.14	Jun 12 03	140	1.16	1.15	-0.01	1.17
Mar 20 03	56	1.12	1.11	-0.01	1.13	Jun 19 03	147	1.15	1.14	-0.02	1.16
Mar 27 03	63	1.13	1.12	-0.01	1.14	Jun 26 03	154	1.15	1.14	-0.01	1.16
Apr 03 03	70	1.13	1.12	-0.01	1.14	Jul 03 03	161	1.15	1.14	-0.02	1.16
Apr 10 03	77	1.12	1.11	-0.03	1.13	Jul 10 03	168	1.16	1.15	-0.02	1.17
Apr 17 03	84	1.14	1.13	-0.01	1.15	Jul 17 03	175	1.16	1.15	-0.03	1.17
Apr 24 03	91	1.15	1.14	. . .	1.16	Jul 24 03	182	1.17	1.16	. . .	1.18

(c) New York Stock Exchange bonds

NEW YORK BONDS
CORPORATION BONDS

	Bonds	Cur Yld	Vol	Close	Net Chg.	
Bond 1	AT&T $5^5/_8$04	5.5	238	101.63	. . .	Yield to Maturity = 3.68%
	AT&T $6^3/_8$04	6.2	60	102.63	–0.13	
	AT&T $7^1/_2$04	7.2	101	103.63	–0.13	
	AT&T $8^1/_8$24	8.0	109	101	0.38	
	ATT 8.35s25	8.3	60	101	0.50	
	AT&T $6^1/_2$29	7.5	190	87.25	0.13	
Bond 2	AT&T $8^5/_8$31	8.4	138	102.75	0.88	Yield to Maturity = 8.40%

Source: Wall Street Journal, Thursday, January 23, 2003, p. C11.

the yield to maturity. Our previous discussion provided us with some rules for deciding when the current yield is likely to be a good approximation and when it is not.

T-bonds 3 and 4 mature in around 30 years, meaning that their characteristics are like those of a consol. The current yields should then be a good approximation of the yields to maturity, and they are: The current yields are within two-tenths of a percentage point of the values for the yields to maturity. This approximation is reasonable even for T-bond 4, which has a price about 7% above its face value.

Now let's take a look at T-bonds 1 and 2, which have a much shorter time to maturity. The price of T-bond 1 differs by less than 1% from the par value, and look how poor an approximation the current yield is for the yield to maturity; it overstates the yield to maturity by more than 4 percentage points. The approximation for T-bond 2 is even worse, with the overstatement over 9 percentage points. This bears out what we learned earlier about the current yield: It can be a very misleading guide to the value of the yield to maturity for a short-term bond if the bond price is not extremely close to par.

Two other categories of bonds are reported much like the Treasury bonds and notes in the newspaper. Government agency and miscellaneous securities include securities issued by U.S. government agencies such as the Government National Mortgage Association, which makes loans to savings and loan institutions, and international agencies such as the World Bank. Tax-exempt bonds are the other category reported in a manner similar to panel (a), except that yield-to-maturity calculations are not usually provided. Tax-exempt bonds include bonds issued by local government and public authorities whose interest payments are exempt from federal income taxes.

Panel (b) quotes yields on U.S. Treasury bills, which, as we have seen, are discount bonds. Since there is no coupon, these securities are identified solely by their maturity dates, which you can see in the first column. The next column, "Days to Mat.," provides the number of days to maturity of the bill. Dealers in these markets always refer to prices by quoting the yield on a discount basis. The "Bid" column gives the discount yield for people selling the bills to dealers, and the "Asked" column gives the discount yield for people buying the bills from dealers. As with bonds and notes, the dealers' profits are made by the asked price being higher than the bid price, leading to the asked discount yield being lower than the bid discount yield.

The "Chg." column indicates how much the asked discount yield changed from the previous day. When financial analysts talk about changes in the yield, they frequently describe the changes in terms of **basis points**, which are hundredths of a percentage point. For example, a financial analyst would describe the −0.01 change in the asked discount yield for the February 13, 2003, T-bill by saying that it had fallen by 1 basis point.

As we learned earlier, the yield on a discount basis understates the yield to maturity, which is reported in the column of panel (b) headed "Ask Yld." This is evident from a comparison of the "Ask Yld." and "Asked" columns. As we would also expect from our discussion of the calculation of yields on a discount basis, the understatement grows as the maturity of the bill lengthens.

Panel (c) has quotations for corporate bonds traded on the New York Stock Exchange. Corporate bonds traded on the American Stock Exchange are reported in like manner. The first column identifies the bond by indicating the corporation that issued it. The bonds we are looking at have all been issued by American Telephone and Telegraph (AT&T). The next column tells the coupon rate and the maturity date ($5\frac{5}{8}$ and 2004 for Bond 1). The "Cur. Yld." column reports the current yield (5.5), and "Vol." gives the volume of trading in that bond (238 bonds of $1,000 face value traded that day). The "Close" price is the last traded price that day per $100 of face value. The price of 101.63 represents $1016.30 for a $1,000-face-value bond. The "Net Chg." is the change in the closing price from the previous trading day.

The yield to maturity is also given for two bonds. This information is not usually provided in the newspaper, but it is included here because it shows how misleading the current yield can be for a bond with a short maturity such as the $5\frac{5}{8}$s, of 2004. The current yield of 5.5% is a misleading measure of the interest rate because the yield to maturity is actually 3.68 percent. By contrast, for the $8\frac{5}{8}$s, of 2031, with nearly 30 years to maturity, the current yield and the yield to maturity are exactly equal.

The Distinction Between Interest Rates and Returns

Many people think that the interest rate on a bond tells them all they need to know about how well off they are as a result of owning it. If Irving the Investor thinks he is better off when he owns a long-term bond yielding a 10% interest rate and the interest rate rises to 20%, he will have a rude awakening: As we will shortly see, if he has to sell the bond, Irving has lost his shirt! How well a person does by holding a bond or any other security over a particular time period is accurately measured by the **return**, or, in more precise terminology, the **rate of return**. For any security, the rate of return is defined as the payments to the owner plus the change in its value, expressed as a fraction of its purchase price. To make this definition clearer, let us see what the return would look like for a $1,000-face-value coupon bond with a coupon rate of 10% that is bought for $1,000, held for one year, and then sold for $1,200. The payments to the owner are the yearly coupon payments of $100, and the change in its value is $1,200 − $1,000 = $200. Adding these together and expressing them as a fraction of the purchase price of $1,000 gives us the one-year holding-period return for this bond:

$$\frac{\$100 + \$200}{\$1{,}000} = \frac{\$300}{\$1{,}000} = 0.30 = 30\%$$

You may have noticed something quite surprising about the return that we have just calculated: It equals 30%, yet as Table 1 indicates, initially the yield to maturity was only 10 percent. This demonstrates that ***the return on a bond will not necessarily equal the interest rate on that bond***. We now see that the distinction between interest rate and return can be important, although for many securities the two may be closely related.

Study Guide

The concept of *return* discussed here is extremely important because it is used continually throughout the book. Make sure that you understand how a return is calculated and why it can differ from the interest rate. This understanding will make the material presented later in the book easier to follow.

More generally, the return on a bond held from time t to time $t + 1$ can be written as:

$$RET = \frac{C + P_{t+1} - P_t}{P_t} \tag{9}$$

where RET = return from holding the bond from time t to time $t + 1$
P_t = price of the bond at time t
P_{t+1} = price of the bond at time $t + 1$
C = coupon payment

A convenient way to rewrite the return formula in Equation 9 is to recognize that it can be split into two separate terms:

$$RET = \frac{C}{P_t} + \frac{P_{t+1} - P_t}{P_t}$$

The first term is the current yield i_c (the coupon payment over the purchase price):

$$\frac{C}{P_t} = i_c$$

The second term is the **rate of capital gain**, or the change in the bond's price relative to the initial purchase price:

$$\frac{P_{t+1} - P_t}{P_t} = g$$

where g = rate of capital gain. Equation 9 can then be rewritten as:

$$RET = i_c + g \tag{10}$$

which shows that the return on a bond is the current yield i_c plus the rate of capital gain g. This rewritten formula illustrates the point we just discovered. Even for a bond for which the current yield i_c is an accurate measure of the yield to maturity, the return can differ substantially from the interest rate. Returns will differ from the interest rate, especially if there are sizable fluctuations in the price of the bond that produce substantial capital gains or losses.

To explore this point even further, let's look at what happens to the returns on bonds of different maturities when interest rates rise. Table 2 calculates the one-year return on several 10%-coupon-rate bonds all purchased at par when interest rates on

Table 2 One-Year Returns on Different-Maturity 10%-Coupon-Rate Bonds When Interest Rates Rise from 10% to 20%

(1) Years to Maturity When Bond Is Purchased	(2) Initial Current Yield (%)	(3) Initial Price ($)	(4) Price Next Year* ($)	(5) Rate of Capital Gain (%)	(6) Rate of Return (2 + 5) (%)
30	10	1,000	503	−49.7	−39.7
20	10	1,000	516	−48.4	−38.4
10	10	1,000	597	−40.3	−30.3
5	10	1,000	741	−25.9	−15.9
2	10	1,000	917	−8.3	+1.7
1	10	1,000	1,000	0.0	+10.0

*Calculated using Equation 3.

all these bonds rise from 10 to 20%. Several key findings in this table are generally true of all bonds:

- The only bond whose return equals the initial yield to maturity is one whose time to maturity is the same as the holding period (see the last bond in Table 2).
- A rise in interest rates is associated with a fall in bond prices, resulting in capital losses on bonds whose terms to maturity are longer than the holding period.
- The more distant a bond's maturity, the greater the size of the percentage price change associated with an interest-rate change.
- The more distant a bond's maturity, the lower the rate of return that occurs as a result of the increase in the interest rate.
- Even though a bond has a substantial initial interest rate, its return can turn out to be negative if interest rates rise.

At first it frequently puzzles students (as it puzzles poor Irving the Investor) that a rise in interest rates can mean that a bond has been a poor investment. The trick to understanding this is to recognize that a rise in the interest rate means that the price of a bond has fallen. A rise in interest rates therefore means that a capital loss has occurred, and if this loss is large enough, the bond can be a poor investment indeed. For example, we see in Table 2 that the bond that has 30 years to maturity when purchased has a capital loss of 49.7% when the interest rate rises from 10 to 20%. This loss is so large that it exceeds the current yield of 10%, resulting in a negative return (loss) of −39.7%. If Irving does not sell the bond, his capital loss is often referred to as a "paper loss." This is a loss nonetheless because if he had not bought this bond and had instead put his money in the bank, he would now be able to buy more bonds at their lower price than he presently owns.

Maturity and the Volatility of Bond Returns: Interest-Rate Risk

The finding that the prices of longer-maturity bonds respond more dramatically to changes in interest rates helps explain an important fact about the behavior of bond markets: ***Prices and returns for long-term bonds are more volatile than those for shorter-term bonds***. Price changes of +20% and −20% within a year, with corresponding variations in returns, are common for bonds more than 20 years away from maturity.

We now see that changes in interest rates make investments in long-term bonds quite risky. Indeed, the riskiness of an asset's return that results from interest-rate changes is so important that it has been given a special name, **interest-rate risk**.[6] Dealing with interest-rate risk is a major concern of managers of financial institutions and investors, as we will see in later chapters (see also Box 2).

Although long-term debt instruments have substantial interest-rate risk, short-term debt instruments do not. Indeed, bonds with a maturity that is as short as the holding period have no interest-rate risk.[7] We see this for the coupon bond at the bottom of Table 2, which has no uncertainty about the rate of return because it equals the yield to maturity, which is known at the time the bond is purchased. The key to understanding why there is no interest-rate risk for *any* bond whose time to maturity matches the holding period is to recognize that (in this case) the price at the end of the holding period is already fixed at the face value. The change in interest rates can then have no effect on the price at the end of the holding period for these bonds, and the return will therefore be equal to the yield to maturity known at the time the bond is purchased.[8]

---

[6]Interest-rate risk can be quantitatively measured using the concept of *duration*. This concept and how it is calculated is discussed in an appendix to this chapter, which can be found on this book's web site at www.aw.com/mishkin.

[7]The statement that there is no interest-rate risk for any bond whose time to maturity matches the holding period is literally true only for discount bonds and zero-coupon bonds that make no intermediate cash payments before the holding period is over. A coupon bond that makes an intermediate cash payment before the holding period is over requires that this payment be reinvested. Because the interest rate at which this payment can be reinvested is uncertain, there is some uncertainty about the return on this coupon bond even when the time to maturity equals the holding period. However, the riskiness of the return on a coupon bond from reinvesting the coupon payments is typically quite small, and so the basic point that a coupon bond with a time to maturity equaling the holding period has very little risk still holds true.

[8]In the text, we are assuming that all holding periods are short and equal to the maturity on short-term bonds and are thus not subject to interest-rate risk. However, if an investor's holding period is longer than the term to maturity of the bond, the investor is exposed to a type of interest-rate risk called *reinvestment risk*. Reinvestment risk occurs because the proceeds from the short-term bond need to be reinvested at a future interest rate that is uncertain.

To understand reinvestment risk, suppose that Irving the Investor has a holding period of two years and decides to purchase a $1,000 one-year bond at face value and will then purchase another one at the end of the first year. If the initial interest rate is 10%, Irving will have $1,100 at the end of the year. If the interest rate rises to 20%, as in Table 2, Irving will find that buying $1,100 worth of another one-year bond will leave him at the end of the second year with $1,100 × (1 + 0.20) = $1,320. Thus Irving's two-year return will be ($1,320 − $1,000)/1,000 = 0.32 = 32%, which equals 14.9% at an annual rate. In this case, Irving has earned more by buying the one-year bonds than if he had initially purchased the two-year bond with an interest rate of 10%. Thus when Irving has a holding period that is longer than the term to maturity of the bonds he purchases, he benefits from a rise in interest rates. Conversely, if interest rates fall to 5%, Irving will have only $1,155 at the end of two years: $1,100 × (1 + 0.05). Thus his two-year return will be ($1,155 − $1,000)/1,000 = 0.155 = 15.5%, which is 7.2 percent at an annual rate. With a holding period greater than the term to maturity of the bond, Irving now loses from a fall in interest rates.

We have thus seen that when the holding period is longer than the term to maturity of a bond, the return is uncertain because the future interest rate when reinvestment occurs is also uncertain—in short, there is reinvestment risk. We also see that if the holding period is longer than the term to maturity of the bond, the investor benefits from a rise in interest rates and is hurt by a fall in interest rates.

Box 2

Helping Investors to Select Desired Interest-Rate Risk

Because many investors want to know how much interest-rate risk they are exposed to, some mutual fund companies try to educate investors about the perils of interest-rate risk, as well as to offer investment alternatives that match their investors' preferences.

Vanguard Group, for example, offers eight separate high-grade bond mutual funds. In its prospectus, Vanguard separates the funds by the average maturity of the bonds they hold and demonstrates the effect of interest-rate changes by computing the percentage change in bond value resulting from a 1% increase and decrease in interest rates. Three of the bond funds invest in bonds with average maturities of one to three years, which Vanguard rates as having the lowest interest-rate risk. Three other funds hold bonds with average maturities of five to ten years, which Vanguard rates as having medium interest-rate risk. Two funds hold long-term bonds with maturities of 15 to 30 years, which Vanguard rates as having high interest-rate risk.

By providing this information, Vanguard hopes to increase its market share in the sales of bond funds. Not surprisingly, Vanguard is one of the most successful mutual fund companies in the business.

Summary

The return on a bond, which tells you how good an investment it has been over the holding period, is equal to the yield to maturity in only one special case: when the holding period and the maturity of the bond are identical. Bonds whose term to maturity is longer than the holding period are subject to interest-rate risk: Changes in interest rates lead to capital gains and losses that produce substantial differences between the return and the yield to maturity known at the time the bond is purchased. Interest-rate risk is especially important for long-term bonds, where the capital gains and losses can be substantial. This is why long-term bonds are not considered to be safe assets with a sure return over short holding periods.

The Distinction Between Real and Nominal Interest Rates

www.martincapital.com/charts.htm

Go to charts of real versus nominal rates to view 30 years of nominal interest rates compared to real rates for the 30-year T-bond and 90-day T-bill.

So far in our discussion of interest rates, we have ignored the effects of inflation on the cost of borrowing. What we have up to now been calling the interest rate makes no allowance for inflation, and it is more precisely referred to as the **nominal interest rate**, which is distinguished from the **real interest rate**, the interest rate that is adjusted by subtracting expected changes in the price level (inflation) so that it more accurately reflects the true cost of borrowing.[9] The real interest rate is more accurately defined by the *Fisher equation*, named for Irving Fisher, one of the great monetary economists of the

[9]The real interest rate defined in the text is more precisely referred to as the *ex ante real interest rate* because it is adjusted for *expected* changes in the price level. This is the real interest rate that is most important to economic decisions, and typically it is what economists mean when they make reference to the "real" interest rate. The interest rate that is adjusted for *actual* changes in the price level is called the *ex post real interest rate*. It describes how well a lender has done in real terms *after the fact*.

twentieth century. The Fisher equation states that the nominal interest rate i equals the real interest rate i_r plus the expected rate of inflation π^e:[10]

$$i = i_r + \pi^e \tag{11}$$

Rearranging terms, we find that the real interest rate equals the nominal interest rate minus the expected inflation rate:

$$i_r = i - \pi^e \tag{12}$$

To see why this definition makes sense, let us first consider a situation in which you have made a one-year simple loan with a 5% interest rate ($i = 5\%$) and you expect the price level to rise by 3% over the course of the year ($\pi^e = 3\%$). As a result of making the loan, at the end of the year you will have 2% more in **real terms**, that is, in terms of real goods and services you can buy. In this case, the interest rate you have earned in terms of real goods and services is 2%; that is,

$$i_r = 5\% - 3\% = 2\%$$

as indicated by the Fisher definition.

Now what if the interest rate rises to 8%, but you expect the inflation rate to be 10% over the course of the year? Although you will have 8% more dollars at the end of the year, you will be paying 10% more for goods; the result is that you will be able to buy 2% fewer goods at the end of the year and you are 2% worse off *in real terms.* This is also exactly what the Fisher definition tells us, because:

$$i_r = 8\% - 10\% = -2\%$$

As a lender, you are clearly less eager to make a loan in this case, because in terms of real goods and services you have actually earned a negative interest rate of 2%. By contrast, as the borrower, you fare quite well because at the end of the year, the amounts you will have to pay back will be worth 2% less in terms of goods and services—you as the borrower will be ahead by 2% in real terms. ***When the real interest rate is low, there are greater incentives to borrow and fewer incentives to lend.***

A similar distinction can be made between nominal returns and real returns. Nominal returns, which do not allow for inflation, are what we have been referring to as simply "returns." When inflation is subtracted from a nominal return, we have the real return, which indicates the amount of extra goods and services that can be purchased as a result of holding the security.

The distinction between real and nominal interest rates is important because the real interest rate, which reflects the real cost of borrowing, is likely to be a better indicator of the incentives to borrow and lend. It appears to be a better guide to how peo-

---

[10] A more precise formulation of the Fisher equation is:

$$i = i_r + \pi^e + (i_r \times \pi^e)$$

because:

$$1 + i = (1 + i_r)(1 + \pi^e) = 1 + i_r + \pi^e + (i_r \times \pi^e)$$

and subtracting 1 from both sides gives us the first equation. For small values of i_r and π^e, the term $i_r \times \pi^e$ is so small that we ignore it, as in the text.

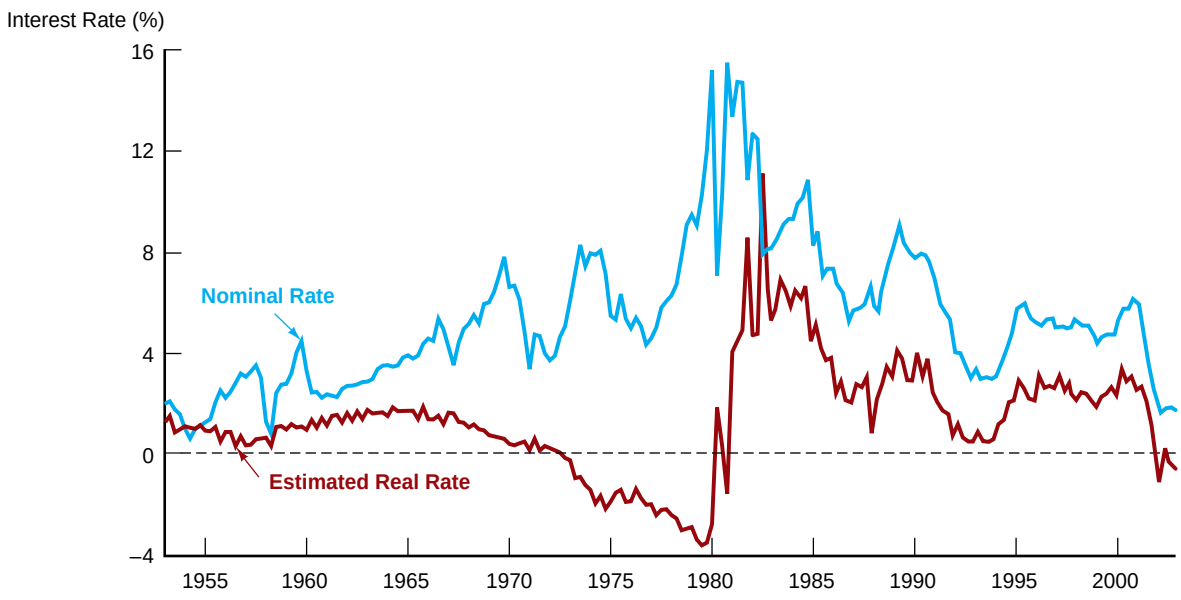


FIGURE 1 Real and Nominal Interest Rates (Three-Month Treasury Bill), 1953–2002

Sources: Nominal rates from www.federalreserve.gov/releases/H15. The real rate is constructed using the procedure outlined in Frederic S. Mishkin, "The Real Interest Rate: An Empirical Investigation," *Carnegie-Rochester Conference Series on Public Policy* 15 (1981): 151–200. This procedure involves estimating expected inflation as a function of past interest rates, inflation, and time trends and then subtracting the expected inflation measure from the nominal interest rate.

ple will be affected by what is happening in credit markets. Figure 1, which presents estimates from 1953 to 2002 of the real and nominal interest rates on three-month U.S. Treasury bills, shows us that nominal and real rates often do not move together. (This is also true for nominal and real interest rates in the rest of the world.) In particular, when nominal rates in the United States were high in the 1970s, real rates were actually extremely low—often negative. By the standard of nominal interest rates, you would have thought that credit market conditions were tight in this period, because it was expensive to borrow. However, the estimates of the real rates indicate that you would have been mistaken. In real terms, the cost of borrowing was actually quite low.[11]

[11]Because most interest income in the United States is subject to federal income taxes, the true earnings in real terms from holding a debt instrument are not reflected by the real interest rate defined by the Fisher equation but rather by the *after-tax real interest rate,* which equals the nominal interest rate *after income tax payments have been subtracted,* minus the expected inflation rate. For a person facing a 30% tax rate, the after-tax interest rate earned on a bond yielding 10% is only 7% because 30% of the interest income must be paid to the Internal Revenue Service. Thus the after-tax real interest rate on this bond when expected inflation is 5% equals 2% (= 7% − 5%). More generally, the after-tax real interest rate can be expressed as:

$$i(1 - \tau) - \pi^e$$

where τ = the income tax rate.

This formula for the after-tax real interest rate also provides a better measure of the effective cost of borrowing for many corporations and homeowners in the United States because in calculating income taxes, they can deduct

Box 3

With TIPS, Real Interest Rates Have Become Observable in the United States

When the U.S. Treasury decided to issue TIPS (Treasury Inflation Protection Securities), in January 1997, a version of indexed Treasury coupon bonds, it was somewhat late in the game. Other countries such as the United Kingdom, Canada, Australia, and Sweden had already beaten the United States to the punch. (In September 1998, the U.S. Treasury also began issuing the Series I savings bond, which provides inflation protection for small investors.)

These indexed securities have successfully acquired a niche in the bond market, enabling governments to raise more funds. In addition, because their interest and principal payments are adjusted for changes in the price level, the interest rate on these bonds provides a direct measure of a real interest rate. These indexed bonds are very useful to policymakers, especially monetary policymakers, because by subtracting their interest rate from a nominal interest rate on a nonindexed bond, they generate more insight into expected inflation, a valuable piece of information. For example, on January 22, 2003, the interest rate on the ten-year Treasury bond was 3.84%, while that on the ten-year TIPS was 2.19%. Thus, the implied expected inflation rate for the next ten years, derived from the difference between these two rates, was 1.65%. The private sector finds the information provided by TIPS very useful: Many commercial and investment banks routinely publish the expected U.S. inflation rates derived from these bonds.

Until recently, real interest rates in the United States were not observable; only nominal rates were reported. This all changed when, in January 1997, the U.S. Treasury began to issue **indexed bonds**, whose interest and principal payments are adjusted for changes in the price level (see Box 3).

interest payments on loans from their income. Thus if you face a 30% tax rate and take out a mortgage loan with a 10% interest rate, you are able to deduct the 10% interest payment and thus lower your taxes by 30% of this amount. Your after-tax nominal cost of borrowing is then 7% (10% minus 30% of the 10% interest payment), and when the expected inflation rate is 5%, the effective cost of borrowing in real terms is again 2% (= 7% − 5%).

As the example (and the formula) indicates, after-tax real interest rates are always below the real interest rate defined by the Fisher equation. For a further discussion of measures of after-tax real interest rates, see Frederic S. Mishkin, "The Real Interest Rate: An Empirical Investigation," *Carnegie-Rochester Conference Series on Public Policy* 15 (1981): 151–200.

Summary

1. The yield to maturity, which is the measure that most accurately reflects the interest rate, is the interest rate that equates the present value of future payments of a debt instrument with its value today. Application of this principle reveals that bond prices and interest rates are negatively related: When the interest rate rises, the price of the bond must fall, and vice versa.
2. Two less accurate measures of interest rates are commonly used to quote interest rates on coupon and discount bonds. The current yield, which equals the coupon payment divided by the price of a coupon bond, is a less accurate measure of the yield to maturity the shorter the maturity of the bond and the greater the gap between the price and the par value. The yield on a discount basis (also called the discount yield) understates the yield to maturity on a discount bond, and the understatement worsens with the distance from maturity of the discount security. Even though these

measures are misleading guides to the size of the interest rate, a change in them always signals a change in the same direction for the yield to maturity.

3. The return on a security, which tells you how well you have done by holding this security over a stated period of time, can differ substantially from the interest rate as measured by the yield to maturity. Long-term bond prices have substantial fluctuations when interest rates change and thus bear interest-rate risk. The resulting capital gains and losses can be large, which is why long-term bonds are not considered to be safe assets with a sure return.
4. The real interest rate is defined as the nominal interest rate minus the expected rate of inflation. It is a better measure of the incentives to borrow and lend than the nominal interest rate, and it is a more accurate indicator of the tightness of credit market conditions than the nominal interest rate.

Key Terms

basis point, p. 74
consol or perpetuity, p. 67
coupon bond, p. 63
coupon rate, p. 64
current yield, p. 70
discount bond (zero-coupon bond), p. 64
face value (par value), p. 63
fixed-payment loan (fully amortized loan), p. 63
indexed bond, p. 82
interest-rate risk, p. 78
nominal interest rate, p. 79
present discounted value, p. 61
present value, p. 61
rate of capital gain, p. 76
real interest rate, p. 79
real terms, p. 80
return (rate of return), p. 75
simple loan, p. 62
yield on a discount basis (discount yield), p. 71
yield to maturity, p. 64

Questions and Problems

Questions marked with an asterisk are answered at the end of the book in an appendix, "Answers to Selected Questions and Problems."

*1. Would a dollar tomorrow be worth more to you today when the interest rate is 20% or when it is 10%?
2. You have just won $20 million in the state lottery, which promises to pay you $1 million (tax free) every year for the next 20 years. Have you really won $20 million?
*3. If the interest rate is 10%, what is the present value of a security that pays you $1,100 next year, $1,210 the year after, and $1,331 the year after that?
4. If the security in Problem 3 sold for $3,500, is the yield to maturity greater or less than 10%? Why?
*5. Write down the formula that is used to calculate the yield to maturity on a 20-year 10% coupon bond with $1,000 face value that sells for $2,000.
6. What is the yield to maturity on a $1,000-face-value discount bond maturing in one year that sells for $800?
*7. What is the yield to maturity on a simple loan for $1 million that requires a repayment of $2 million in five years' time?
8. To pay for college, you have just taken out a $1,000 government loan that makes you pay $126 per year for 25 years. However, you don't have to start making these payments until you graduate from college two years from now. Why is the yield to maturity necessarily less than 12%, the yield to maturity on a normal $1,000 fixed-payment loan in which you pay $126 per year for 25 years?
*9. Which $1,000 bond has the higher yield to maturity, a 20-year bond selling for $800 with a current yield of 15% or a one-year bond selling for $800 with a current yield of 5%?

10. Pick five U.S. Treasury bonds from the bond page of the newspaper, and calculate the current yield. Note when the current yield is a good approximation of the yield to maturity.

***11.** You are offered two bonds, a one-year U.S. Treasury bond with a yield to maturity of 9% and a one-year U.S. Treasury bill with a yield on a discount basis of 8.9%. Which would you rather own?

12. If there is a decline in interest rates, which would you rather be holding, long-term bonds or short-term bonds? Why? Which type of bond has the greater interest-rate risk?

***13.** Francine the Financial Adviser has just given you the following advice: "Long-term bonds are a great investment because their interest rate is over 20%." Is Francine necessarily right?

14. If mortgage rates rise from 5% to 10% but the expected rate of increase in housing prices rises from 2% to 9%, are people more or less likely to buy houses?

***15.** Interest rates were lower in the mid-1980s than they were in the late 1970s, yet many economists have commented that real interest rates were actually much higher in the mid-1980s than in the late 1970s. Does this make sense? Do you think that these economists are right?

Web Exercises

1. Investigate the data available from the Federal Reserve at www.federalreserve.gov/releases/. Answer the following questions:
   a. What is the difference in the interest rates on commercial paper for financial firms when compared to nonfinancial firms?
   b. What was the interest rate on the one-month Eurodollar at the end of 2002?
   c. What is the most recent interest rate report for the 30-year Treasury note?

2. Figure 1 in the text shows the estimated real and nominal rates for three-month treasury bills. Go to www.martincapital.com/charts.htm and click on "interest rates and yields," then on "real interest rates."
   a. Compare the three-month real rate to the long-term real rate. Which is greater?
   b. Compare the short-term nominal rate to the long-term nominal rate. Which appears most volatile?

3. In this chapter we have discussed long-term bonds as if there were only one type, coupon bonds. In fact there are also long-term discount bonds. A discount bond is sold at a low price and the whole return comes in the form of a price appreciation. You can easily compute the current price of a discount bond using the financial calculator at http://app.ny.frb.org/sbr/.

   To compute the redemption values for savings bonds, fill in the information at the site and click on the Compute Values button. A maximum of five years of data will be displayed for each computation.

appendix to chapter 4

Measuring Interest-Rate Risk: Duration

In our discussion of interest-rate risk, we saw that when interest rates change, a bond with a longer term to maturity has a larger change in its price and hence more interest-rate risk than a bond with a shorter term to maturity. Although this is a useful general fact, in order to measure interest-rate risk, the manager of a financial institution needs more precise information on the actual capital gain or loss that occurs when the interest rate changes by a certain amount. To do this, the manager needs to make use of the concept of duration, the average lifetime of a debt security's stream of payments.

The fact that two bonds have the same term to maturity does not mean that they have the same interest-rate risk. A long-term discount bond with ten years to maturity, a so-called *zero-coupon bond*, makes all of its payments at the end of the ten years, whereas a 10% coupon bond with ten years to maturity makes substantial cash payments before the maturity date. Since the coupon bond makes payments earlier than the zero-coupon bond, we might intuitively guess that the coupon bond's *effective maturity*, the term to maturity that accurately measures interest-rate risk, is shorter than it is for the zero-coupon discount bond.

Indeed, this is exactly what we find in example 1.

EXAMPLE 1: Rate of Capital Gain

Calculate the rate of capital gain or loss on a ten-year zero-coupon bond for which the interest rate has increased from 10% to 20%. The bond has a face value of $1,000.

Solution

The rate of capital gain or loss is −49.7%.

$$g = \frac{P_{t+1} - P_t}{P_t}$$

where

P_{t+1} = price of the bond one year from now $= \dfrac{\$1{,}000}{(1 + 0.20)^9} = \$193.81$

P_t = price of the bond today $= \dfrac{\$1{,}000}{(1 + 0.10)^{10}} = \$385.54$

Thus:

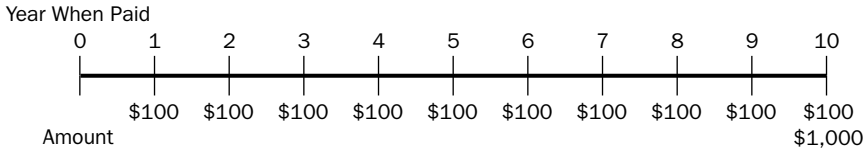
$$g = \frac{\$193.81 - \$385.54}{\$385.54}$$

$$g = -0.497 = -49.7\%$$

But as we have already calculated in Table 2 in Chapter 4, the capital gain on the 10% ten-year coupon bond is −40.3%. We see that interest-rate risk for the ten-year coupon bond is less than for the ten-year zero-coupon bond, so the effective maturity on the coupon bond (which measures interest-rate risk) is, as expected, shorter than the effective maturity on the zero-coupon bond.

Calculating Duration

To calculate the duration or effective maturity on any debt security, Frederick Macaulay, a researcher at the National Bureau of Economic Research, invented the concept of duration more than half a century ago. Because a zero-coupon bond makes no cash payments before the bond matures, it makes sense to define its effective maturity as equal to its actual term to maturity. Macaulay then realized that he could measure the effective maturity of a coupon bond by recognizing that a coupon bond is equivalent to a set of zero-coupon discount bonds. A ten-year 10% coupon bond with a face value of $1,000 has cash payments identical to the following set of zero-coupon bonds: a $100 one-year zero-coupon bond (which pays the equivalent of the $100 coupon payment made by the $1,000 ten-year 10% coupon bond at the end of one year), a $100 two-year zero-coupon bond (which pays the equivalent of the $100 coupon payment at the end of two years), … , a $100 ten-year zero-coupon bond (which pays the equivalent of the $100 coupon payment at the end of ten years), and a $1,000 ten-year zero-coupon bond (which pays back the equivalent of the coupon bond's $1,000 face value). This set of coupon bonds is shown in the following time line:



This same set of coupon bonds is listed in column (2) of Table 1, which calculates the duration on the ten-year coupon bond when its interest rate is 10%.

To get the effective maturity of this set of zero-coupon bonds, we would want to sum up the effective maturity of each zero-coupon bond, weighting it by the percentage of the total value of all the bonds that it represents. In other words, the duration of this set of zero-coupon bonds is the weighted average of the effective maturities of the individual zero-coupon bonds, with the weights equaling the proportion of the total value represented by each zero-coupon bond. We do this in several steps in Table 1. First we calculate the present value of each of the zero-coupon bonds when the interest rate is 10% in column (3). Then in column (4) we divide each of these present values by $1,000, the total present value of the set of zero-coupon bonds, to get the percentage of the total value of all the bonds that each bond represents. Note that the sum of the weights in column (4) must total 100%, as shown at the bottom of the column.

Table 1 Calculating Duration on a $1,000 Ten-Year 10% Coupon Bond When Its Interest Rate Is 10%

(1) Year	(2) Cash Payments (Zero-Coupon Bonds) ($)	(3) Present Value (*PV*) of Cash Payments (i = 10%) ($)	(4) Weights (% of total *PV* = *PV*/$1,000) (%)	(5) Weighted Maturity (1 × 4)/100 (years)
1	100	90.91	9.091	0.09091
2	100	82.64	8.264	0.16528
3	100	75.13	7.513	0.22539
4	100	68.30	6.830	0.27320
5	100	62.09	6.209	0.31045
6	100	56.44	5.644	0.33864
7	100	51.32	5.132	0.35924
8	100	46.65	4.665	0.37320
9	100	42.41	4.241	0.38169
10	100	38.55	3.855	0.38550
10	1,000	385.54	38.554	3.85500
Total		1,000.00	100.000	6.75850

To get the effective maturity of the set of zero-coupon bonds, we add up the weighted maturities in column (5) and obtain the figure of 6.76 years. This figure for the effective maturity of the set of zero-coupon bonds is the duration of the 10% ten-year coupon bond because the bond is equivalent to this set of zero-coupon bonds. In short, we see that ***duration is a weighted average of the maturities of the cash payments***.

The duration calculation done in Table 1 can be written as follows:

$$DUR = \sum_{t=1}^{n} t\frac{CP_t}{(1+i)^t} \Bigg/ \sum_{t=1}^{n} \frac{CP_t}{(1+i)^t} \tag{1}$$

where DUR = duration
t = years until cash payment is made
CP_t = cash payment (interest plus principal) at time t
i = interest rate
n = years to maturity of the security

This formula is not as intuitive as the calculation done in Table 1, but it does have the advantage that it can easily be programmed into a calculator or computer, making duration calculations very easy.

If we calculate the duration for an 11-year 10% coupon bond when the interest rate is again 10%, we find that it equals 7.14 years, which is greater than the 6.76 years for the ten-year bond. Thus we have reached the expected conclusion: ***All else being equal, the longer the term to maturity of a bond, the longer its duration.***

You might think that knowing the maturity of a coupon bond is enough to tell you what its duration is. However, that is not the case. To see this and to give you more practice in calculating duration, in Table 2 we again calculate the duration for the ten-year 10% coupon bond, but when the current interest rate is 20%, rather than 10% as in Table 1. The calculation in Table 2 reveals that the duration of the coupon bond at this higher interest rate has fallen from 6.76 years to 5.72 years. The explanation is fairly straightforward. When the interest rate is higher, the cash payments in the future are discounted more heavily and become less important in present-value terms relative to the total present value of all the payments. The relative weight for these cash payments drops as we see in Table 2, and so the effective maturity of the bond falls. We have come to an important conclusion: ***All else being equal, when interest rates rise, the duration of a coupon bond falls.***

The duration of a coupon bond is also affected by its coupon rate. For example, consider a ten-year 20% coupon bond when the interest rate is 10%. Using the same procedure, we find that its duration at the higher 20% coupon rate is 5.98 years versus 6.76 years when the coupon rate is 10%. The explanation is that a higher coupon rate means that a relatively greater amount of the cash payments are made earlier in the life of the bond, and so the effective maturity of the bond must fall. We have thus established a third fact about duration: ***All else being equal, the higher the coupon rate on the bond, the shorter the bond's duration.***

Table 2 Calculating Duration on a $1,000 Ten-Year 10% Coupon Bond When Its Interest Rate Is 20%

(1) Year	(2) Cash Payments (Zero-Coupon Bonds) ($)	(3) Present Value (*PV*) of Cash Payments (i = 20%) ($)	(4) Weights (% of total *PV* = *PV*/$580.76) (%)	(5) Weighted Maturity (1 × 4)/100 (years)
1	100	83.33	14.348	0.14348
2	100	69.44	11.957	0.23914
3	100	57.87	9.965	0.29895
4	100	48.23	8.305	0.33220
5	100	40.19	6.920	0.34600
6	100	33.49	5.767	0.34602
7	100	27.91	4.806	0.33642
8	100	23.26	4.005	0.32040
9	100	19.38	3.337	0.30033
10	100	16.15	2.781	0.27810
10	$1,000	161.51	27.808	2.78100
Total		580.76	100.000	5.72204

Study Guide

To make certain that you understand how to calculate duration, practice doing the calculations in Tables 1 and 2. Try to produce the tables for calculating duration in the case of an 11-year 10% coupon bond and also for the 10-year 20% coupon bond mentioned in the text when the current interest rate is 10%. Make sure your calculations produce the same results found in this appendix.

One additional fact about duration makes this concept useful when applied to a portfolio of securities. Our examples have shown that duration is equal to the weighted average of the durations of the cash payments (the effective maturities of the corresponding zero-coupon bonds). So if we calculate the duration for two different securities, it should be easy to see that the duration of a portfolio of the two securities is just the weighted average of the durations of the two securities, with the weights reflecting the proportion of the portfolio invested in each.

EXAMPLE 2: Duration

A manager of a financial institution is holding 25% of a portfolio in a bond with a five-year duration and 75% in a bond with a ten-year duration. What is the duration of the portfolio?

Solution

The duration of the portfolio is 8.75 years.

$$(0.25 \times 5) + (0.75 \times 10) = 1.25 + 7.5 = 8.75 \text{ years}$$

We now see that ***the duration of a portfolio of securities is the weighted average of the durations of the individual securities, with the weights reflecting the proportion of the portfolio invested in each***. This fact about duration is often referred to as the *additive property of duration*, and it is extremely useful, because it means that the duration of a portfolio of securities is easy to calculate from the durations of the individual securities.

To summarize, our calculations of duration for coupon bonds have revealed four facts:

1. The longer the term to maturity of a bond, everything else being equal, the greater its duration.
2. When interest rates rise, everything else being equal, the duration of a coupon bond falls.
3. The higher the coupon rate on the bond, everything else being equal, the shorter the bond's duration.
4. Duration is additive: The duration of a portfolio of securities is the weighted average of the durations of the individual securities, with the weights reflecting the proportion of the portfolio invested in each.

Duration and Interest-Rate Risk

Now that we understand how duration is calculated, we want to see how it can be used by the practicing financial institution manager to measure interest-rate risk. Duration is a particularly useful concept, because it provides a good approximation, particularly when interest-rate changes are small, for how much the security price changes for a given change in interest rates, as the following formula indicates:

$$\%\Delta P \approx -DUR \times \frac{\Delta i}{1 + i} \tag{2}$$

where $\%\Delta P = (P_{t+1} - P_t)/P_t$ = percent change in the price of the security from t to $t + 1$ = rate of capital gain
DUR = duration
i = interest rate

EXAMPLE 3: Duration and Interest-Rate Risk

A pension fund manager is holding a ten-year 10% coupon bond in the fund's portfolio and the interest rate is currently 10%. What loss would the fund be exposed to if the interest rate rises to 11% tomorrow?

Solution

The approximate percentage change in the price of the bond is −6.15%.

As the calculation in Table 1 shows, the duration of a ten-year 10% coupon bond is 6.76 years.

$$\%\Delta P \approx -DUR \times \frac{\Delta i}{1 + i}$$

where

DUR = duration = 6.76
Δi = change in interest rate = 0.11 − 0.10 = 0.01
i = current interest rate = 0.10

Thus:

$$\%\Delta P \approx -6.76 \times \frac{0.01}{1 + 0.10}$$
$$\%\Delta P \approx -0.0615 = -6.15\%$$

EXAMPLE 4: Duration and Interest-Rate Risk

Now the pension manager has the option to hold a ten-year coupon bond with a coupon rate of 20% instead of 10%. As mentioned earlier, the duration for this 20% coupon bond is 5.98 years when the interest rate is 10%. Find the approximate change in the bond price when the interest rate increases from 10% to 11%.

Solution

This time the approximate change in bond price is −5.4%. This change in bond price is much smaller than for the higher-duration coupon bond:

$$\%\Delta P \approx -DUR \times \frac{\Delta i}{1 + i}$$

where

$$DUR = \text{duration} = 5.98$$
$$\Delta i = \text{change in interest rate} = 0.11 - 0.10 = 0.01$$
$$i = \text{current interest rate} = 0.10$$

Thus:

$$\%\Delta P \approx -5.98 \times \frac{0.01}{1 + 0.10}$$
$$\%\Delta P \approx -0.054 = -5.4\%$$

The pension fund manager realizes that the interest-rate risk on the 20% coupon bond is less than on the 10% coupon, so he switches the fund out of the 10% coupon bond and into the 20% coupon bond.

Examples 3 and 4 have led the pension fund manager to an important conclusion about the relationship of duration and interest-rate risk: ***The greater the duration of a security, the greater the percentage change in the market value of the security for a given change in interest rates. Therefore, the greater the duration of a security, the greater its interest-rate risk.***

This reasoning applies equally to a portfolio of securities. So by calculating the duration of the fund's portfolio of securities using the methods outlined here, a pension fund manager can easily ascertain the amount of interest-rate risk the entire fund is exposed to. As we will see in Chapter 9, duration is a highly useful concept for the management of interest-rate risk that is widely used by managers of banks and other financial institutions.

Chapter 5

The Behavior of Interest Rates

PREVIEW

In the early 1950s, nominal interest rates on three-month Treasury bills were about 1% at an annual rate; by 1981, they had reached over 15%, then fell to 3% in 1993, rose to above 5% by the mid-1990s, and fell below 2% in the early 2000s. What explains these substantial fluctuations in interest rates? One reason why we study money, banking, and financial markets is to provide some answers to this question.

In this chapter, we examine how the overall level of *nominal* interest rates (which we refer to as simply "interest rates") is determined and which factors influence their behavior. We learned in Chapter 4 that interest rates are negatively related to the price of bonds, so if we can explain why bond prices change, we can also explain why interest rates fluctuate. To do this, we make use of supply and demand analysis for bond markets and money markets to examine how interest rates change.

In order to derive a demand curve for assets like money or bonds, the first step in our analysis, we must first understand what determines the demand for these assets. We do this by examining an economic theory known as the *theory of asset demand,* which outlines criteria that are important when deciding how much of an asset to buy. Armed with this theory, we can then go on to derive the demand curve for bonds or money. After deriving supply curves for these assets, we develop the concept of *market equilibrium*, the point at which the quantity supplied equals the quantity demanded. Then we use this model to explain changes in equilibrium interest rates.

Because interest rates on different securities tend to move together, in this chapter we will proceed as if there were only one type of security and a single interest rate in the entire economy. In the following chapter, we expand our analysis to look at why interest rates on different types of securities differ.

Determinants of Asset Demand

Before going on to our supply and demand analysis of the bond market and the market for money, we must first understand what determines the quantity demanded of an asset. Recall that an asset is a piece of property that is a store of value. Items such as money, bonds, stocks, art, land, houses, farm equipment, and manufacturing machinery are all assets. Facing the question of whether to buy and hold an asset or

whether to buy one asset rather than another, an individual must consider the following factors:

1. **Wealth**, the total resources owned by the individual, including all assets
2. **Expected return** (the return expected over the next period) on one asset relative to alternative assets
3. **Risk** (the degree of uncertainty associated with the return) on one asset relative to alternative assets
4. **Liquidity** (the ease and speed with which an asset can be turned into cash) relative to alternative assets

Study Guide

As we discuss each factor that influences asset demand, remember that we are always holding all the other factors constant. Also, think of additional examples of how changes in each factor would influence your decision to purchase a particular asset: say, a house or a share of common stock. This intuitive approach will help you understand how the theory works in practice.

Wealth

When we find that our wealth has increased, we have more resources available with which to purchase assets, and so, not surprisingly, the quantity of assets we demand increases. Therefore, the effect of changes in wealth on the quantity demanded of an asset can be summarized as follows: ***Holding everything else constant, an increase in wealth raises the quantity demanded of an asset.***

Expected Returns

In Chapter 4, we saw that the return on an asset (such as a bond) measures how much we gain from holding that asset. When we make a decision to buy an asset, we are influenced by what we expect the return on that asset to be. If a Mobil Oil Corporation bond, for example, has a return of 15% half the time and 5% the other half of the time, its expected return (which you can think of as the average return) is 10% ($= 0.5 \times 15\% + 0.5 \times 5\%$).[1] If the expected return on the Mobil Oil bond rises relative to expected returns on alternative assets, holding everything else constant, then it becomes more desirable to purchase it, and the quantity demanded increases. This can occur in either of two ways: (1) when the expected return on the Mobil Oil bond rises while the return on an alternative asset—say, stock in IBM—remains unchanged or (2) when the return on the alternative asset, the IBM stock, falls while the return on the Mobil Oil bond remains unchanged. To summarize, ***an increase in an asset's expected return relative to that of an alternative asset, holding everything else unchanged, raises the quantity demanded of the asset.***

[1] If you are interested in finding out more information on how to calculate expected returns, as well as standard deviations of returns that measure risk, you can look at an appendix to this chapter describing models of asset pricing that is on this book's web site at www.aw.com/mishkin. This appendix also describes how diversification lowers the overall risk of a portfolio and has a discussion of systematic risk and basic asset pricing models such as the capital asset pricing model and arbitrage pricing theory.

Risk

The degree of risk or uncertainty of an asset's returns also affects the demand for the asset. Consider two assets, stock in Fly-by-Night Airlines and stock in Feet-on-the-Ground Bus Company. Suppose that Fly-by-Night stock has a return of 15% half the time and 5% the other half of the time, making its expected return 10%, while stock in Feet-on-the-Ground has a fixed return of 10%. Fly-by-Night stock has uncertainty associated with its returns and so has greater risk than stock in Feet-on-the-Ground, whose return is a sure thing.

A *risk-averse* person prefers stock in Feet-on-the-Ground (the sure thing) to Fly-by-Night stock (the riskier asset), even though the stocks have the same expected return, 10%. By contrast, a person who prefers risk is a *risk preferrer* or *risk lover.* Most people are risk-averse, especially in their financial decisions: Everything else being equal, they prefer to hold the less risky asset. Hence, ***holding everything else constant, if an asset's risk rises relative to that of alternative assets, its quantity demanded will fall.***

Liquidity

Another factor that affects the demand for an asset is how quickly it can be converted into cash at low costs—its liquidity. An asset is liquid if the market in which it is traded has depth and breadth; that is, if the market has many buyers and sellers. A house is not a very liquid asset, because it may be hard to find a buyer quickly; if a house must be sold to pay off bills, it might have to be sold for a much lower price. And the transaction costs in selling a house (broker's commissions, lawyer's fees, and so on) are substantial. A U.S. Treasury bill, by contrast, is a highly liquid asset. It can be sold in a well-organized market where there are many buyers, so it can be sold quickly at low cost. ***The more liquid an asset is relative to alternative assets, holding everything else unchanged, the more desirable it is, and the greater will be the quantity demanded.***

Theory of Asset Demand

All the determining factors we have just discussed can be assembled into the **theory of asset demand**, which states that, holding all of the other factors constant:

1. The quantity demanded of an asset is positively related to wealth.
2. The quantity demanded of an asset is positively related to its expected return relative to alternative assets.
3. The quantity demanded of an asset is negatively related to the risk of its returns relative to alternative assets.
4. The quantity demanded of an asset is positively related to its liquidity relative to alternative assets.

These results are summarized in Table 1.

Supply and Demand in the Bond Market

Our first approach to the analysis of interest-rate determination looks at supply and demand in the bond market. The first step in the analysis is to obtain a bond **demand curve**, which shows the relationship between the quantity demanded and the price when all other economic variables are held constant (that is, values of other variables are taken as given). You may recall from previous economics courses that the

SUMMARY **Table 1 Response of the Quantity of an Asset Demanded to Changes in Wealth, Expected Returns, Risk, and Liquidity**

Variable	Change in Variable	Change in Quantity Demanded
Wealth	↑	↑
Expected return relative to other assets	↑	↑
Risk relative to other assets	↑	↓
Liquidity relative to other assets	↑	↑

Note: Only increases in the variables are shown. The effect of decreases in the variables on the change in demand would be the opposite of those indicated in the rightmost column.

assumption that all other economic variables are held constant is called *ceteris paribus*, which means "other things being equal" in Latin.

Demand Curve

To clarify our analysis, let us consider the demand for one-year discount bonds, which make no coupon payments but pay the owner the $1,000 face value in a year. If the holding period is one year, then as we have seen in Chapter 4, the return on the bonds is known absolutely and is equal to the interest rate as measured by the yield to maturity. This means that the expected return on this bond is equal to the interest rate i, which, using Equation 6 in Chapter 4, is:

$$i = RET^e = \frac{F - P}{P}$$

where i = interest rate = yield to maturity
RET^e = expected return
F = face value of the discount bond
P = initial purchase price of the discount bond

This formula shows that a particular value of the interest rate corresponds to each bond price. If the bond sells for $950, the interest rate and expected return is:

$$\frac{\$1{,}000 - \$950}{\$950} = 0.053 = 5.3\%$$

At this 5.3% interest rate and expected return corresponding to a bond price of $950, let us assume that the quantity of bonds demanded is $100 billion, which is plotted as point A in Figure 1. To display both the bond price and the corresponding interest rate, Figure 1 has two vertical axes. The left vertical axis shows the bond price, with the price of bonds increasing from $750 near the bottom of the axis toward $1,000 at the top. The right vertical axis shows the interest rate, which increases in the *opposite* direction from 0% at the top of the axis to 33% near the bottom. The right and left vertical axes run in opposite directions because, as we learned in Chapter 4, bond

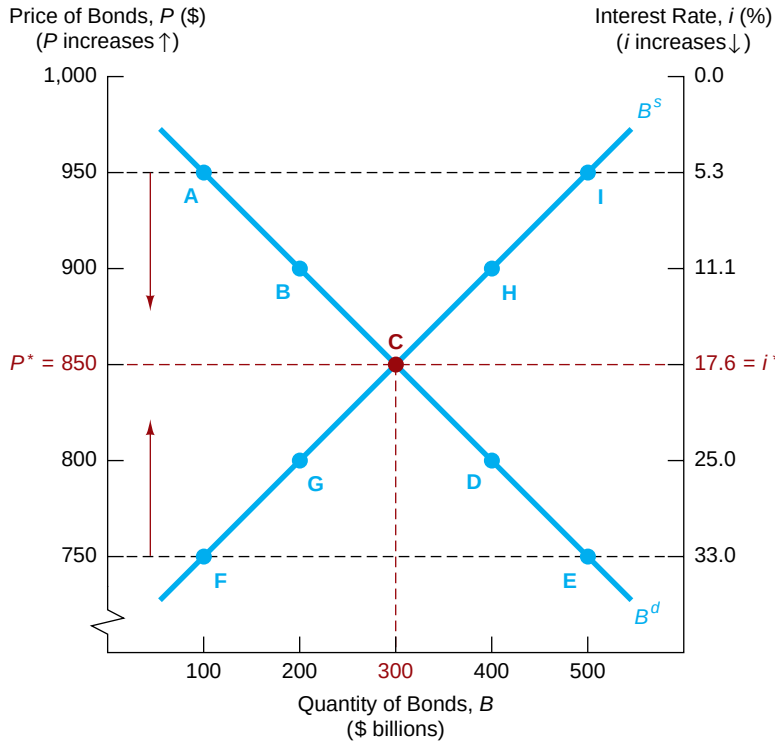


FIGURE 1 Supply and Demand for Bonds

Equilibrium in the bond market occurs at point C, the intersection of the demand curve B^d and the bond supply curve B^s. The equilibrium price is $P^* = \$850$, and the equilibrium interest rate is $i^* = 17.6\%$. (*Note:* P and i increase in opposite directions. P on the left vertical axis increases as we go up the axis from \$750 near the bottom to \$1,000 at the top, while i on the right vertical axis increases as we go down the axis from 0% at the top to 33% near the bottom.)

price and interest rate are always negatively related: As the price of the bond rises, the interest rate on the bond necessarily falls.

At a price of \$900, the interest rate and expected return equals:

$$\frac{\$1{,}000 - \$900}{\$900} = 0.111 = 11.1\%$$

Because the expected return on these bonds is higher, with all other economic variables (such as income, expected returns on other assets, risk, and liquidity) held constant, the quantity demanded of bonds will be higher as predicted by the theory of asset demand. Point B in Figure 1 shows that the quantity of bonds demanded at the price of \$900 has risen to \$200 billion. Continuing with this reasoning, if the bond price is \$850 (interest rate and expected return = 17.6%), the quantity of bonds demanded (point C) will be greater than at point B. Similarly, at the lower prices of \$800 (interest rate = 25%) and \$750 (interest rate = 33.3%), the quantity of bonds demanded will be even higher (points D and E). The curve B^d, which connects these points, is the demand curve for bonds. It has the usual downward slope, indicating that at lower prices of the bond (everything else being equal), the quantity demanded is higher.[2]

[2]Note that although our analysis indicates that the demand curve is downward-sloping, it does not imply that the curve is a straight line. For ease of exposition, however, we will draw demand curves and supply curves as straight lines.

Supply Curve

An important assumption behind the demand curve for bonds in Figure 1 is that all other economic variables besides the bond's price and interest rate are held constant. We use the same assumption in deriving a **supply curve**, which shows the relationship between the quantity supplied and the price when all other economic variables are held constant.

When the price of the bonds is \$750 (interest rate = 33.3%), point F shows that the quantity of bonds supplied is \$100 billion for the example we are considering. If the price is \$800, the interest rate is the lower rate of 25%. Because at this interest rate it is now less costly to borrow by issuing bonds, firms will be willing to borrow more through bond issues, and the quantity of bonds supplied is at the higher level of \$200 billion (point G). An even higher price of \$850, corresponding to a lower interest rate of 17.6%, results in a larger quantity of bonds supplied of \$300 billion (point C). Higher prices of \$900 and \$950 result in even greater quantities of bonds supplied (points H and I). The B^s curve, which connects these points, is the supply curve for bonds. It has the usual upward slope found in supply curves, indicating that as the price increases (everything else being equal), the quantity supplied increases.

Market Equilibrium

In economics, **market equilibrium** occurs when the amount that people are willing to buy (*demand*) equals the amount that people are willing to sell (*supply*) at a given price. In the bond market, this is achieved when the quantity of bonds demanded equals the quantity of bonds supplied:

$$B^d = B^s \tag{1}$$

In Figure 1, equilibrium occurs at point C, where the demand and supply curves intersect at a bond price of \$850 (interest rate of 17.6%) and a quantity of bonds of \$300 billion. The price of $P^* = 850$, where the quantity demanded equals the quantity supplied, is called the *equilibrium* or *market-clearing* price. Similarly, the interest rate of $i^* = 17.6\%$ that corresponds to this price is called the equilibrium or market-clearing interest rate.

The concepts of market equilibrium and equilibrium price or interest rate are useful, because there is a tendency for the market to head toward them. We can see that it does in Figure 1 by first looking at what happens when we have a bond price that is above the equilibrium price. When the price of bonds is set too high, at, say, \$950, the quantity of bonds supplied at point I is greater than the quantity of bonds demanded at point A. A situation like this, in which the quantity of bonds supplied exceeds the quantity of bonds demanded, is called a condition of **excess supply**. Because people want to sell more bonds than others want to buy, the price of the bonds will fall, and this is why the downward arrow is drawn in the figure at the bond price of \$950. As long as the bond price remains above the equilibrium price, there will continue to be an excess supply of bonds, and the price will continue to fall. This will stop only when the price has reached the equilibrium price of \$850, where the excess supply of bonds has been eliminated.

Now let's look at what happens when the price of bonds is below the equilibrium price. If the price of the bonds is set too low, at, say, \$750, the quantity demanded at point E is greater than the quantity supplied at point F. This is called a condition of **excess demand**. People now want to buy more bonds than others are willing to sell, and so the price of bonds will be driven up. This is illustrated by the upward arrow drawn in the figure at the bond price of \$750. Only when the excess demand for

bonds is eliminated by the price rising to the equilibrium level of $850 is there no further tendency for the price to rise.

We can see that the concept of equilibrium price is a useful one because it indicates where the market will settle. Because each price on the left vertical axis of Figure 1 corresponds to a value of the interest rate on the right vertical axis, the same diagram also shows that the interest rate will head toward the equilibrium interest rate of 17.6%. When the interest rate is below the equilibrium interest rate, as it is when it is at 5.3%, the price of the bond is above the equilibrium price, and there will be an excess supply of bonds. The price of the bond then falls, leading to a rise in the interest rate toward the equilibrium level. Similarly, when the interest rate is above the equilibrium level, as it is when it is at 33.3%, there is excess demand for bonds, and the bond price will rise, driving the interest rate back down to the equilibrium level of 17.6%.

Supply and Demand Analysis

Our Figure 1 is a conventional supply and demand diagram with price on the left vertical axis and quantity on the horizontal axis. Because the interest rate that corresponds to each bond price is also marked on the right vertical axis, this diagram allows us to read the equilibrium interest rate, giving us a model that describes the determination of interest rates. It is important to recognize that a supply and demand diagram like Figure 1 can be drawn for *any* type of bond because the interest rate and price of a bond are *always* negatively related for any type of bond, whether a discount bond or a coupon bond.

Loanable Funds Framework

Throughout this book we will use diagrams like Figure 1 and analyze interest rate behavior in terms of the supply and demand for bonds. However, the analysis of the bond market that we have developed here has another interpretation with a different terminology. Here we discuss this other terminology, which is couched in terms of the supply and demand for loanable funds used by some economists. We include this discussion in case you come across this other terminology, but you will not need to make use of it to understand how interest rates are determined.

One disadvantage of the diagram in Figure 1 is that interest rates run in an unusual direction on the right vertical axis: As we go up the right axis, interest rates fall. Because economists are typically more concerned with the value of interest rates than with the price of bonds, we could plot the supply of and demand for bonds on a diagram that has only a left vertical axis that provides the values of the interest rates running in the usual direction, rising as we go up the axis. Figure 2 is such a diagram, in which points A through I match the corresponding points in Figure 1.

However, making interest rates run in the "usual" direction on the vertical axis presents us with a problem. Our demand curve for bonds, points A through E, now looks peculiar because it has an upward slope. This upward slope is, however, completely consistent with our usual demand analysis, which produces a negative relationship between price and quantity. The inverse relationship between bond prices and interest rates means that in moving from point A to point B to point C, bond prices are falling and, consistent with usual demand analysis, the quantity demanded is rising. Similarly, our supply curve for bonds, points F through I, has an unusual-looking downward slope but is completely consistent with the usual view that price and the quantity supplied are positively related.

One way to give the demand curve the usual downward slope and the supply curve the usual upward slope is to rename the horizontal axis and the demand and

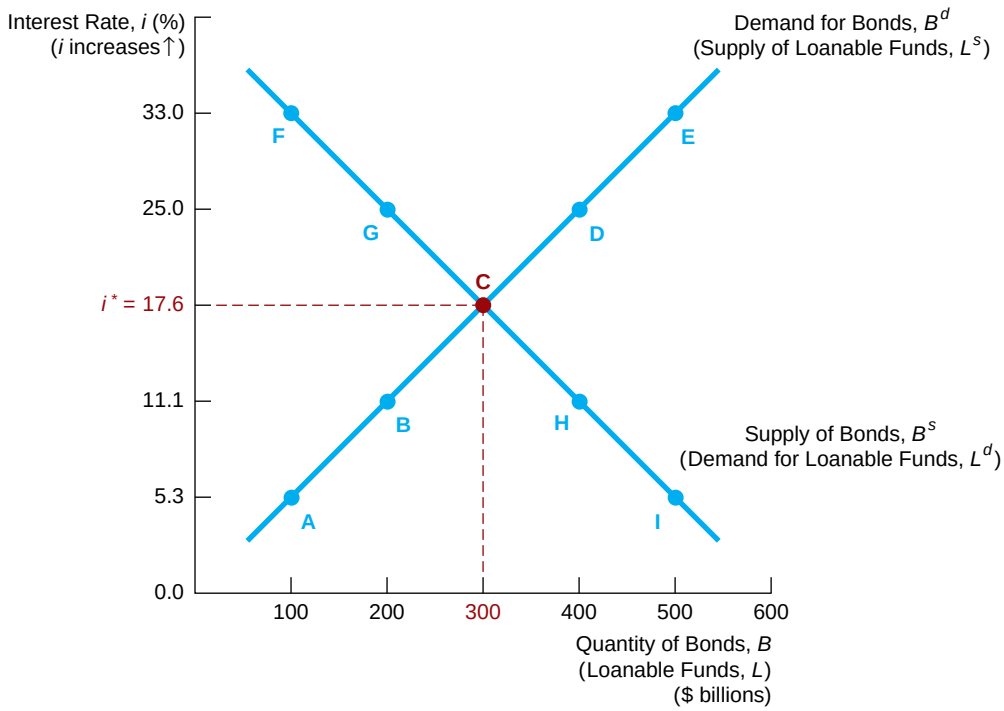


FIGURE 2 A Comparison of Terminology: Loanable Funds and Supply and Demand for Bonds
The demand for bonds is equivalent to the supply of loanable funds, and the supply of bonds is equivalent to the demand for loanable funds. (*Note:* i increases as we go up the vertical axis, in contrast to Figure 1, in which the opposite occurs.)

supply curves. Because a firm supplying bonds is in fact taking out a loan from a person buying a bond, "supplying a bond" is equivalent to "demanding a loan." Thus the supply curve for bonds can be reinterpreted as indicating the *quantity of loans demanded* for each value of the interest rate. If we rename the horizontal axis **loanable funds**, defined as the quantity of loans, the supply of bonds can be reinterpreted as the *demand for loanable funds*. Similarly, the demand curve for bonds can be reidentified as the *supply of loanable funds* because buying (demanding) a bond is equivalent to supplying a loan. Figure 2 relabels the curves and the horizontal axis using the loanable funds terminology in parentheses, and now the renamed loanable funds demand curve has the usual downward slope and the renamed loanable funds supply curve the usual upward slope.

Because supply and demand diagrams that explain how interest rates are determined in the bond market often use the loanable funds terminology, this analysis is frequently referred to as the **loanable funds framework**. However, because in later chapters describing the conduct of monetary policy we focus on how the demand for and supply of bonds is affected, we will continue to conduct supply and demand analysis in terms of bonds, as in Figure 1, rather than loanable funds. Whether the analysis is done in terms of loanable funds or in terms of the demand for and supply

of bonds, the results are the same: The two ways of analyzing the determination of interest rates are equivalent.

An important feature of the analysis here is that supply and demand are always in terms of *stocks* (amounts at a given point in time) of assets, not in terms of *flows*. This approach is somewhat different from certain loanable funds analyses, which are conducted in terms of flows (loans per year). The **asset market approach** for understanding behavior in financial markets—which emphasizes stocks of assets rather than flows in determining asset prices—is now the dominant methodology used by economists, because correctly conducting analyses in terms of flows is very tricky, especially when we encounter inflation.[3]

Changes in Equilibrium Interest Rates

We will now use the supply and demand framework for bonds to analyze why interest rates change. To avoid confusion, it is important to make the distinction between *movements along* a demand (or supply) curve and *shifts in* a demand (or supply) curve. When quantity demanded (or supplied) changes as a result of a change in the price of the bond (or, equivalently, a change in the interest rate), we have a *movement along* the demand (or supply) curve. The change in the quantity demanded when we move from point A to B to C in Figure 1, for example, is a movement along a demand curve. A *shift in* the demand (or supply) curve, by contrast, occurs when the quantity demanded (or supplied) changes *at each given price (or interest rate)* of the bond in response to a change in some other factor besides the bond's price or interest rate. When one of these factors changes, causing a shift in the demand or supply curve, there will be a new equilibrium value for the interest rate.

In the following pages, we will look at how the supply and demand curves shift in response to changes in variables, such as expected inflation and wealth, and what effects these changes have on the equilibrium value of interest rates.

Shifts in the Demand for Bonds

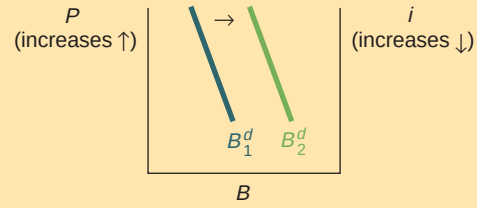
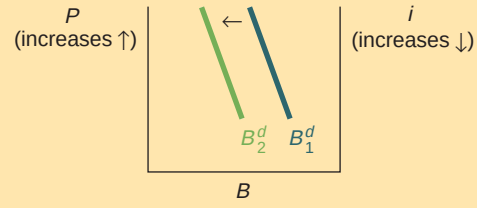
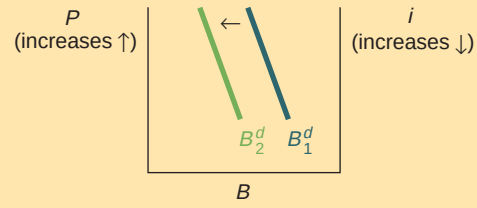
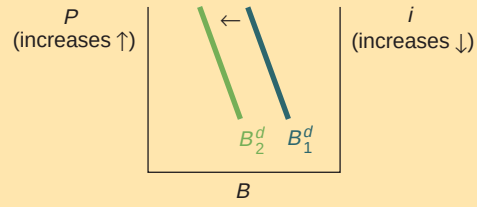
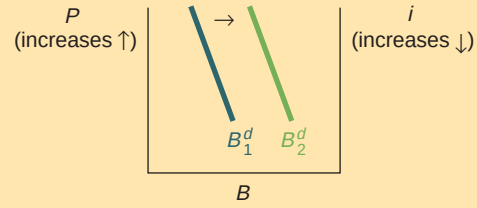
The theory of asset demand demonstrated at the beginning of the chapter provides a framework for deciding what factors cause the demand curve for bonds to shift. These factors include changes in four parameters:

1. Wealth
2. Expected returns on bonds relative to alternative assets
3. Risk of bonds relative to alternative assets
4. Liquidity of bonds relative to alternative assets

To see how a change in each of these factors (holding all other factors constant) can shift the demand curve, let us look at some examples. (As a study aid, Table 2 summarizes the effects of changes in these factors on the bond demand curve.)

[3]The asset market approach developed in the text is useful in understanding not only how interest rates behave but also how any asset price is determined. A second appendix to this chapter, which is on this book's web site at www.aw.com/mishkin, shows how the asset market approach can be applied to understanding the behavior of commodity markets; in particular, the gold market.

SUMMARY Table 2 Factors That Shift the Demand Curve for Bonds

Variable	Change in Variable	Change in Quantity Demanded	Shift in Demand Curve
Wealth	↑	↑	
Expected interest rate	↑	↓	
Expected inflation	↑	↓	
Riskiness of bonds relative to other assets	↑	↓	
Liquidity of bonds relative to other assets	↑	↑	

Note: *P* and *i* increase in opposite directions: *P* on the left vertical axis increases as we go up the axis, while *i* on the right vertical axis increases as we go down the axis. Only increases in the variables are shown. The effect of decreases in the variables on the change in demand would be the opposite of those indicated in the remaining columns.

Wealth. When the economy is growing rapidly in a business cycle expansion and wealth is increasing, the quantity of bonds demanded at each bond price (or interest rate) increases as shown in Figure 3. To see how this works, consider point B on the initial demand curve for bonds B_1^d. It tells us that at a bond price of $900 and an interest rate of 11.1%, the quantity of bonds demanded is $200 billion. With higher wealth, the quantity of bonds demanded at the same interest rate must rise, say, to $400 billion (point B′). Similarly, the higher wealth causes the quantity demanded at a bond price of $800 and an interest rate of 25% to rise from $400 billion to $600 billion (point D to D′). Continuing with this reasoning for every point on the initial demand curve B_1^d, we can see that the demand curve shifts to the right from B_1^d to B_2^d as is indicated by the arrows.

The conclusion we have reached is that ***in a business cycle expansion with growing wealth, the demand for bonds rises and the demand curve for bonds shifts to the right***. Using the same reasoning, ***in a recession, when income and wealth are falling, the demand for bonds falls, and the demand curve shifts to the left***.

Another factor that affects wealth is the public's propensity to save. If households save more, wealth increases and, as we have seen, the demand for bonds rises and the demand curve for bonds shifts to the right. Conversely, if people save less, wealth and the demand for bonds will fall and the demand curve shifts to the left.

Expected Returns. For a one-year discount bond and a one-year holding period, the expected return and the interest rate are identical, so nothing besides today's interest rate affects the expected return.

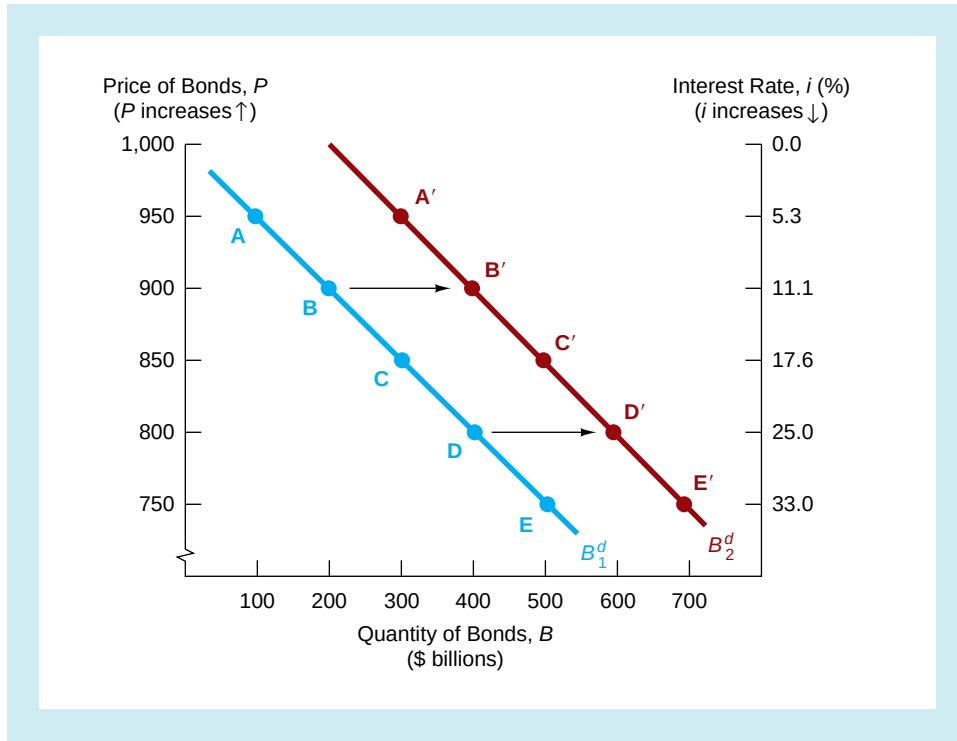


FIGURE 3 Shift in the Demand Curve for Bonds

When the demand for bonds increases, the demand curve shifts to the right as shown. (*Note: P* and *i* increase in opposite directions. *P* on the left vertical axis increases as we go up the axis, while *i* on the right vertical axis increases as we go down the axis.)

For bonds with maturities of greater than one year, the expected return may differ from the interest rate. For example, we saw in Chapter 4, Table 2, that a rise in the interest rate on a long-term bond from 10 to 20% would lead to a sharp decline in price and a very negative return. Hence if people begin to think that interest rates will be higher next year than they had originally anticipated, the expected return today on long-term bonds would fall, and the quantity demanded would fall at each interest rate. ***Higher expected interest rates in the future lower the expected return for long-term bonds, decrease the demand, and shift the demand curve to the left.***

By contrast, a revision downward of expectations of future interest rates would mean that long-term bond prices would be expected to rise more than originally anticipated, and the resulting higher expected return today would raise the quantity demanded at each bond price and interest rate. ***Lower expected interest rates in the future increase the demand for long-term bonds and shift the demand curve to the right*** (as in Figure 3).

Changes in expected returns on other assets can also shift the demand curve for bonds. If people suddenly became more optimistic about the stock market and began to expect higher stock prices in the future, both expected capital gains and expected returns on stocks would rise. With the expected return on bonds held constant, the expected return on bonds today relative to stocks would fall, lowering the demand for bonds and shifting the demand curve to the left.

A change in expected inflation is likely to alter expected returns on physical assets (also called *real assets*) such as automobiles and houses, which affect the demand for bonds. An increase in expected inflation, say, from 5 to 10%, will lead to higher prices on cars and houses in the future and hence higher nominal capital gains. The resulting rise in the expected returns today on these real assets will lead to a fall in the expected return on bonds relative to the expected return on real assets today and thus cause the demand for bonds to fall. Alternatively, we can think of the rise in expected inflation as lowering the real interest rate on bonds, and the resulting decline in the relative expected return on bonds causes the demand for bonds to fall. ***An increase in the expected rate of inflation lowers the expected return for bonds, causing their demand to decline and the demand curve to shift to the left.***

Risk. If prices in the bond market become more volatile, the risk associated with bonds increases, and bonds become a less attractive asset. ***An increase in the riskiness of bonds causes the demand for bonds to fall and the demand curve to shift to the left.***

Conversely, an increase in the volatility of prices in another asset market, such as the stock market, would make bonds more attractive. ***An increase in the riskiness of alternative assets causes the demand for bonds to rise and the demand curve to shift to the right*** (as in Figure 3).

Liquidity. If more people started trading in the bond market, and as a result it became easier to sell bonds quickly, the increase in their liquidity would cause the quantity of bonds demanded at each interest rate to rise. ***Increased liquidity of bonds results in an increased demand for bonds, and the demand curve shifts to the right*** (see Figure 3). ***Similarly, increased liquidity of alternative assets lowers the demand for bonds and shifts the demand curve to the left.*** The reduction of brokerage commissions for trading common stocks that occurred when the fixed-rate commission structure was

abolished in 1975, for example, increased the liquidity of stocks relative to bonds, and the resulting lower demand for bonds shifted the demand curve to the left.

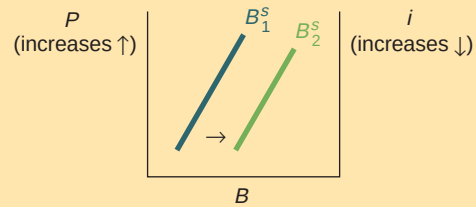
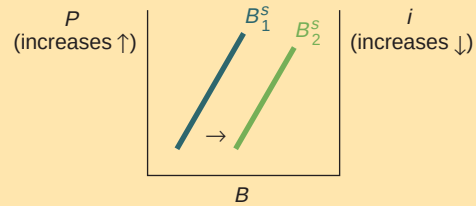
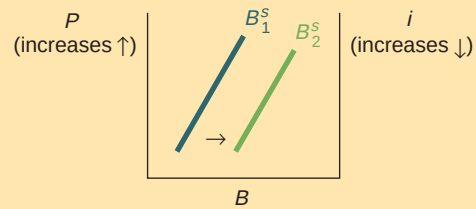
Shifts in the Supply of Bonds

Certain factors can cause the supply curve for bonds to shift, among them these:

1. Expected profitability of investment opportunities
2. Expected inflation
3. Government activities

We will look at how the supply curve shifts when each of these factors changes (all others remaining constant). (As a study aid, Table 3 summarizes the effects of changes in these factors on the bond supply curve.)

SUMMARY **Table 3 Factors That Shift the Supply of Bonds**

Variable	Change in Variable	Change in Quantity Supplied	Shift in Supply Curve
Profitability of investments	↑	↑	
Expected inflation	↑	↑	
Government deficit	↑	↑	

Note: *P* and *i* increase in opposite directions: *P* on the left vertical axis increases as we go up the axis, while *i* on the right vertical axis increases as we go down the axis. Only increases in the variables are shown. The effect of decreases in the variables on the change in supply would be the opposite of those indicated in the remaining columns.

Expected Profitability of Investment Opportunities. The more profitable plant and equipment investments that a firm expects it can make, the more willing it will be to borrow in order to finance these investments. When the economy is growing rapidly, as in a business cycle expansion, investment opportunities that are expected to be profitable abound, and the quantity of bonds supplied at any given bond price and interest rate will increase (see Figure 4). ***Therefore, in a business cycle expansion, the supply of bonds increases, and the supply curve shifts to the right. Likewise, in a recession, when there are far fewer expected profitable investment opportunities, the supply of bonds falls, and the supply curve shifts to the left.***

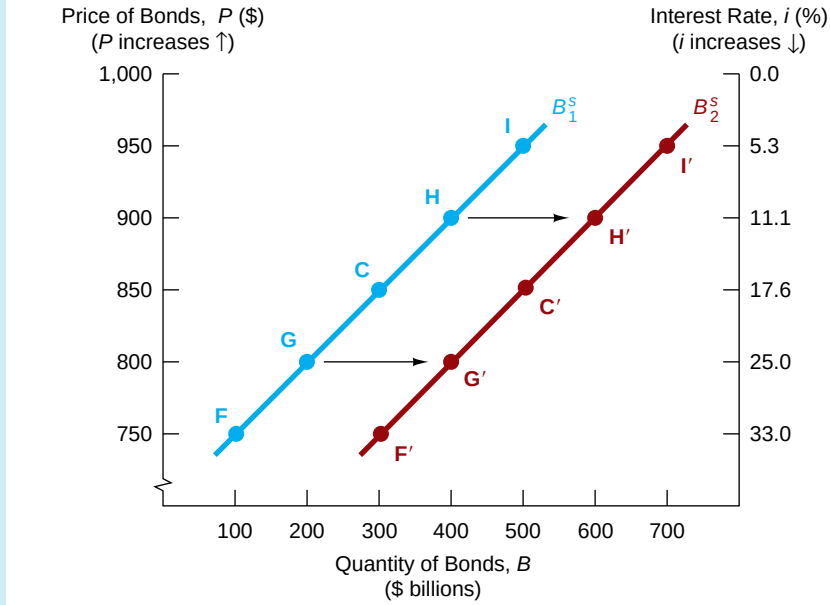
Expected Inflation. As we saw in Chapter 4, the real cost of borrowing is more accurately measured by the real interest rate, which equals the (nominal) interest rate minus the expected inflation rate. For a given interest rate, when expected inflation increases, the real cost of borrowing falls; hence the quantity of bonds supplied increases at any given bond price and interest rate. ***An increase in expected inflation causes the supply of bonds to increase and the supply curve to shift to the right*** (see Figure 4).

ftp://ftp.bls.gov/pub/special.requests/cpi/cpiai.txt
Contains historical information about inflation.

Government Activities. The activities of the government can influence the supply of bonds in several ways. The U.S. Treasury issues bonds to finance government deficits, the gap between the government's expenditures and its revenues. When these deficits are large, the Treasury sells more bonds, and the quantity of bonds supplied at each bond price and interest rate increases. ***Higher government deficits increase the supply of bonds and shift the supply curve to the right*** (see Figure 4). ***On the other hand,***

FIGURE 4 Shift in the Supply Curve for Bonds

When the supply of bonds increases, the supply curve shifts to the right. (*Note:* P and i increase in opposite directions. P on the left vertical axis increases as we go up the axis, while i on the right vertical axis increases as we go down the axis.)



government surpluses, as occurred in the late 1990s, decrease the supply of bonds and shift the supply curve to the left.

State and local governments and other government agencies also issue bonds to finance their expenditures, and this can also affect the supply of bonds. We will see in later chapters that the conduct of monetary policy involves the purchase and sale of bonds, which in turn influences the supply of bonds.

Application

Changes in the Equilibrium Interest Rate Due to Expected Inflation or Business Cycle Expansions

We now can use our knowledge of how supply and demand curves shift to analyze how the equilibrium interest rate can change. The best way to do this is to pursue several applications that are particularly relevant to our understanding of how monetary policy affects interest rates.

Study Guide

Supply and demand analysis for the bond market is best learned by practicing applications. When there is an application in the text and we look at how the interest rate changes because some economic variable increases, see if you can draw the appropriate shifts in the supply and demand curves when this same economic variable decreases. While you are practicing applications, keep two things in mind:

1. When you examine the effect of a variable change, remember that we are assuming that all other variables are unchanged; that is, we are making use of the ceteris paribus assumption.
2. Remember that the interest rate is negatively related to the bond price, so when the equilibrium bond price rises, the equilibrium interest rate falls. Conversely, if the equilibrium bond price moves downward, the equilibrium interest rate rises.

Changes in Expected Inflation: The Fisher Effect

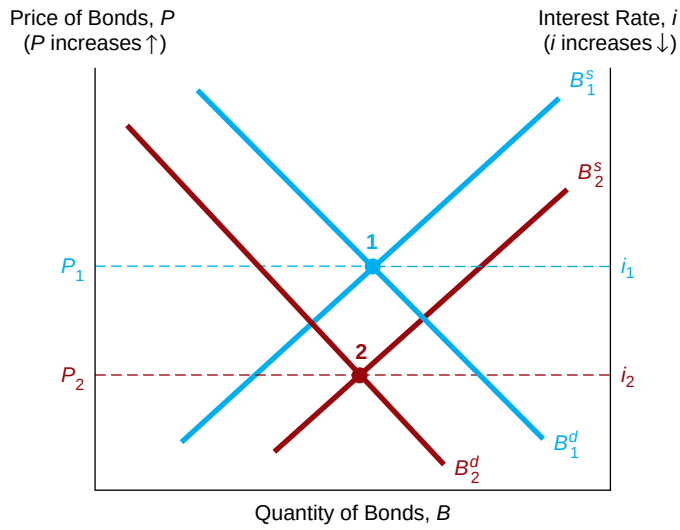
We have already done most of the work to evaluate how a change in expected inflation affects the nominal interest rate, in that we have already analyzed how a change in expected inflation shifts the supply and demand curves. Figure 5 shows the effect on the equilibrium interest rate of an increase in expected inflation.

Suppose that expected inflation is initially 5% and the initial supply and demand curves B^s_1 and B^d_1 intersect at point 1, where the equilibrium bond price is P_1 and the equilibrium interest rate is i_1. If expected inflation rises to 10%, the expected return on bonds relative to real assets falls for any given bond price and interest rate. As a result, the demand for bonds falls, and the demand curve shifts to the left from B^d_1 to B^d_2. The rise in expected inflation also shifts the supply curve. At any given bond price and interest rate, the real cost of borrowing has declined, causing the quantity of bonds supplied to increase, and the supply curve shifts to the right, from B^s_1 to B^s_2.

When the demand and supply curves shift in response to the change in expected inflation, the equilibrium moves from point 1 to point 2, the intersection

FIGURE 5 Response to a Change in Expected Inflation

When expected inflation rises, the supply curve shifts from B_1^s to B_2^s, and the demand curve shifts from B_1^d to B_2^d. The equilibrium moves from point 1 to point 2, with the result that the equilibrium bond price (left axis) falls from P_1 to P_2 and the equilibrium interest rate (right axis) rises from i_1 to i_2. (*Note:* P and i increase in opposite directions. P on the left vertical axis increases as we go up the axis, while i on the right vertical axis increases as we go down the axis.)



of B_2^d and B_2^s. The equilibrium bond price has fallen from P_1 to P_2, and because the bond price is negatively related to the interest rate (as is indicated by the interest rate rising as we go down the right vertical axis), this means that the interest rate has risen from i_1 to i_2. Note that Figure 5 has been drawn so that the equilibrium quantity of bonds remains the same for both point 1 and point 2. However, depending on the size of the shifts in the supply and demand curves, the equilibrium quantity of bonds could either rise or fall when expected inflation rises.

Our supply and demand analysis has led us to an important observation: ***When expected inflation rises, interest rates will rise.*** This result has been named the **Fisher effect**, after Irving Fisher, the economist who first pointed out the relationship of expected inflation to interest rates. The accuracy of this prediction is shown in Figure 6. The interest rate on three-month Treasury bills has usually moved along with the expected inflation rate. Consequently, it is understandable that many economists recommend that inflation must be kept low if we want to keep interest rates low.

Business Cycle Expansion

Figure 7 analyzes the effects of a business cycle expansion on interest rates. In a business cycle expansion, the amounts of goods and services being produced in the economy rise, so national income increases. When this occurs, businesses will be more willing to borrow, because they are likely to have many profitable investment opportunities for which they need financing. Hence at a given bond price and interest rate, the quantity of bonds that firms want to sell (that is, the supply of bonds) will increase. This means that in a business cycle expansion, the supply curve for bonds shifts to the right (see Figure 7) from B_1^s to B_2^s.

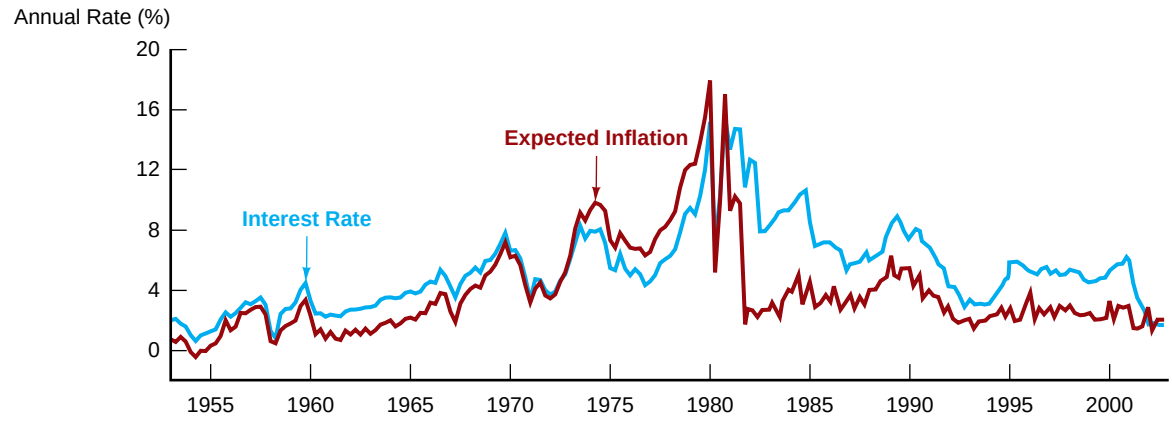
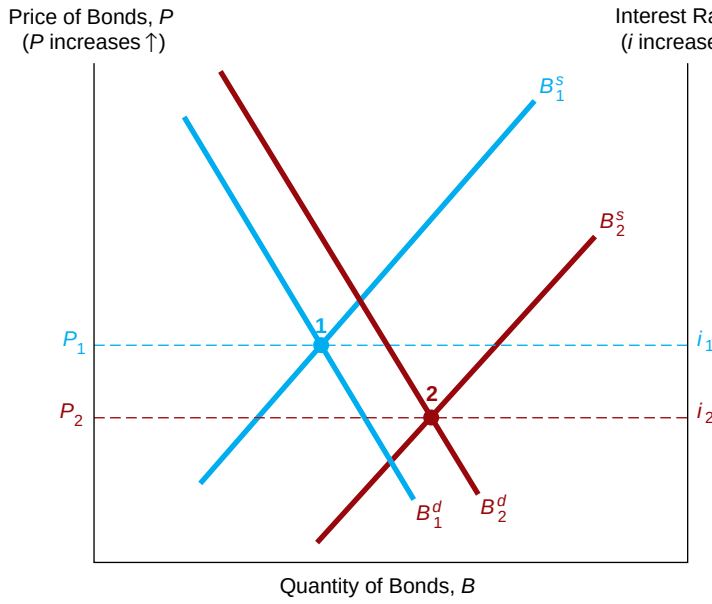


FIGURE 6 Expected Inflation and Interest Rates (Three-Month Treasury Bills), 1953–2002

Source: Expected inflation calculated using procedures outlined in Frederic S. Mishkin, "The Real Interest Rate: An Empirical Investigation," *Carnegie-Rochester Conference Series on Public Policy* 15 (1981): 151–200. These procedures involve estimating expected inflation as a function of past interest rates, inflation, and time trends.

FIGURE 7 Response to a Business Cycle Expansion

In a business cycle expansion, when income and wealth are rising, the demand curve shifts rightward from B_1^d to B_2^d, and the supply curve shifts rightward from B_1^s to B_2^s. If the supply curve shifts to the right more than the demand curve, as in this figure, the equilibrium bond price (left axis) moves down from P_1 to P_2, and the equilibrium interest rate (right axis) rises from i_1 to i_2. (*Note:* P and i increase in opposite directions. P on the left vertical axis increases as we go up the axis, while i on the right vertical axis increases as we go down the axis.)



Expansion in the economy will also affect the demand for bonds. As the business cycle expands, wealth is likely to increase, and then the theory of asset demand tells us that the demand for bonds will rise as well. We see this in Figure 7, where the demand curve has shifted to the right, from B^d_1 to B^d_2.

Given that both the supply and demand curves have shifted to the right, we know that the new equilibrium reached at the intersection of B^d_2 and B^s_2 must also move to the right. However, depending on whether the supply curve shifts more than the demand curve or vice versa, the new equilibrium interest rate can either rise or fall.

The supply and demand analysis used here gives us an ambiguous answer to the question of what will happen to interest rates in a business cycle expansion. The figure has been drawn so that the shift in the supply curve is greater than the shift in the demand curve, causing the equilibrium bond price to fall to P_2, leading to a rise in the equilibrium interest rate to i_2. The reason the figure has been drawn so that a business cycle expansion and a rise in income lead to a higher interest rate is that this is the outcome we actually see in the data. Figure 8 plots the movement of the interest rate on three-month U.S. Treasury bills from 1951 to 2002 and indicates when the business cycle is undergoing recessions (shaded areas). As you can see, the interest rate rises during business cycle expansions and falls during recessions, which is what the supply and demand diagram indicates.

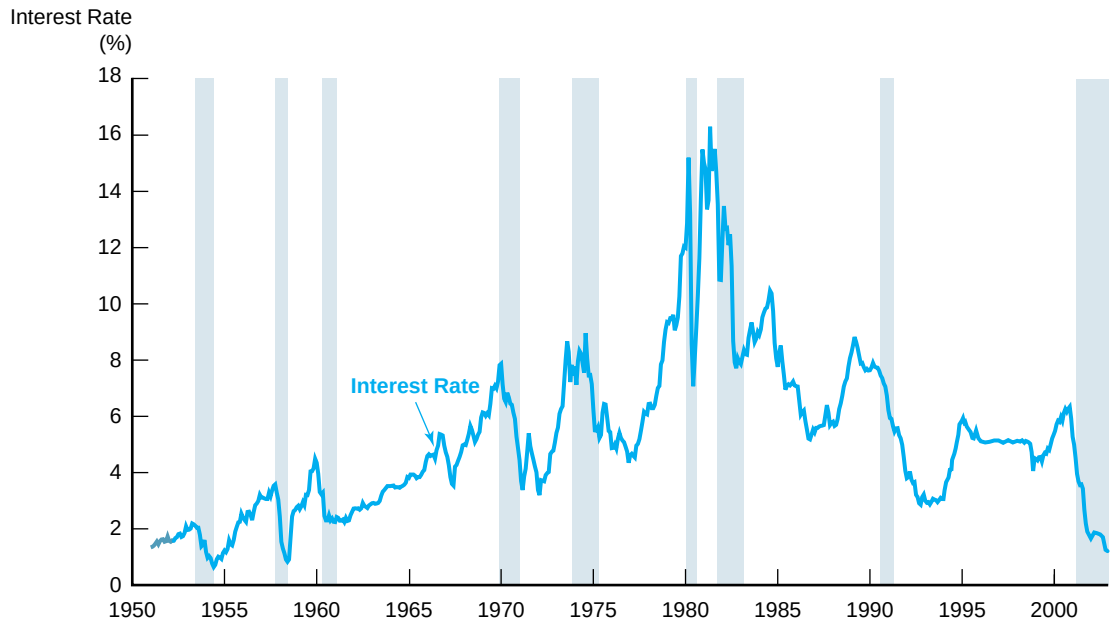


FIGURE 8 Business Cycle and Interest Rates (Three-Month Treasury Bills), 1951–2002

Shaded areas indicate periods of recession. The figure shows that interest rates rise during business cycle expansions and fall during contractions, which is what Figure 7 suggests would happen.

Source: Federal Reserve: www.federalreserve.gov/releases/H15/data.htm.

Application

Explaining Low Japanese Interest Rates

In the 1990s and early 2000s, Japanese interest rates became the lowest in the world. Indeed, in November 1998, an extraordinary event occurred: Interest rates on Japanese six-month Treasury bills turned slightly negative (see Chapter 4). Why did Japanese rates drop to such low levels?

In the late 1990s and early 2000s, Japan experienced a prolonged recession, which was accompanied by deflation, a negative inflation rate. Using these facts, analysis similar to that used in the preceding application explains the low Japanese interest rates.

Negative inflation caused the demand for bonds to rise because the expected return on real assets fell, thereby raising the relative expected return on bonds and in turn causing the demand curve to shift to the right. The negative inflation also raised the real interest rate and therefore the real cost of borrowing for any given nominal rate, thereby causing the supply of bonds to contract and the supply curve to shift to the left. The outcome was then exactly the opposite of that graphed in Figure 5: The rightward shift of the demand curve and leftward shift of the supply curve led to a rise in the bond price and a fall in interest rates.

The business cycle contraction and the resulting lack of investment opportunities in Japan also led to lower interest rates, by decreasing the supply of bonds and shifting the supply curve to the left. Although the demand curve also would shift to the left because wealth decreased during the business cycle contraction, we have seen in the preceding application that the demand curve would shift less than the supply curve. Thus, the bond price rose and interest rates fell (the opposite outcome to that in Figure 7).

Usually, we think that low interest rates are a good thing, because they make it cheap to borrow. But the Japanese example shows that just as there is a fallacy in the adage, "You can never be too rich or too thin": (maybe you can't be too rich, but you can certainly be too thin and do damage to your health), there is a fallacy in always thinking that lower interest rates are better. In Japan, the low and even negative interest rates were a sign that the Japanese economy was in real trouble, with falling prices and a contracting economy. Only when the Japanese economy returns to health will interest rates rise back to more normal levels.

Application

Reading the *Wall Street Journal* "Credit Markets" Column

Now that we have an understanding of how supply and demand determine prices and interest rates in the bond market, we can use our analysis to understand discussions about bond prices and interest rates appearing in the financial press. Every day, the *Wall Street Journal* reports on developments in the bond market on the previous business day in its "Credit Markets" column, an example of which is found in the "Following the Financial News" box. Let's see how statements in the "Credit Markets" column can be explained using our supply and demand framework.

The column describes how the coming announcement of the Bush stimulus package, which was larger than expected, has led to a decline in the prices of Treasury bonds. This is exactly what our supply and demand analysis predicts would happen.

The larger than expected stimulus package has raised concerns about rising future issuance of government bonds, as is mentioned in the second paragraph. The increased supply of bonds in the future will thus shift the supply curve to the right, thereby lowering the price of these bonds in the future by more than expected. The resulting decline in the expected return on these bonds because of their higher future price will lead to an immediate rightward shift in the demand for these bonds today. The outcome is thus a fall in their equilibrium price and a rise in their interest rates.

Our analysis thus demonstrates why, even though the Bush plan has not increased the supply of bonds today, the price of these bonds falls immediately.

Following the Financial News

The "Credit Markets" Column

The "Credit Markets" column appears daily in the *Wall Street Journal*; an example is presented here. It is found in the third section, "Money and Investing."

CREDIT MARKETS

Treasurys Drop Ahead of Bush Stimulus Package
Selloff Is Fueled by Reports Of More Extensive Plan Than Investors Expected

By Michael Mackenzie
Dow Jones Newswires

NEW YORK—Already buckling amid signs of improvement in the economy and a departure of investors seeking better returns in corporate bonds and equities, Treasurys face another bearish element when President Bush outlines his fiscal-stimulus package today.

Reports that the package could total about $600 billion over 10 years, much larger than expected by bond investors, contributed to a further selloff yesterday amid concerns about rising future issuance of government bonds.

After closing 2002 around 2.73% and 3.81%, respectively, five-year and 10-year Treasury yields have risen sharply in the new year. Yesterday, five-year and 10-year yields ended at 3.04% and 4.06%, respectively, up from 2.98% and 4.03% Friday.

The benchmark 10-year note's price, which moves inversely to its yield, at 4 p.m. was down 11/32 point, or $3.44 per $1,000 face value, at 99 15/32.

The 30-year bond's price was down 14/32 point at 105 27/32 to yield 4.984%, up from 4.949% Friday.

The selloff was concentrated in shorter-maturity Treasurys, as investors sold those issues while buying long-dated Treasurys in so-called curve-flattening trades. Later, hedging related to nongovernment bond issues helped lift prices from lows but failed to spark any real rally.

Although uncertainty about geopolitical issues continued to lend some support to Treasurys, the proposed Bush stimulus package "is front and center for the Treasurys market at the moment," said Michael Kastner, head of taxable fixed income for Deutsche Private Banking, New York. "Details are leaking out, and Treasurys are selling off."

The prospect of rising government spending means more Treasury issuance, concentrated in the five-and 10-year areas, analysts said. Lehman Brothers forecast "net supply" of Treasurys would increase about $300 billion this year.

"The Treasury market already reflects the assumption that a large stimulus package will be unveiled," said Joseph Shatz, government-securities strategist at Merrill Lynch. However, he noted that key questions for the market are "what elements of stimulus will be passed, and the time frame of stimulus objectives."

Indeed, there are some factors that mitigate the package's short-term impact on the economy and the market, some added. Analysts at Wrightson ICAP in Jersey City, N.J., said roughly half of a $500 billion to $600 billion stimulus package "will be longer-term supply-side tax reform measures spread evenly over the period, while the other half would be more quick-focused fixes for the business cycle."

The proposal to eliminate taxes individuals pay on dividends would boost stocks, likely at the expense of bonds, analysts said.

They also noted that the Bush proposals have to muster congressional support, which could take some time.

Yet, most added, there is no escaping the sense that the stars are aligned against the Treasury market this year, with a hefty stimulus package another bleak factor clouding the outlook for government bonds.

"Treasury yields are currently too low," said Deutsche's Mr. Kastner. "Uncertainty over Iraq is maintaining some support for Treasurys, but we are starting to sense that the mood of the market is one of selling the rallies."

Source: Wall Street Journal, Tuesday, January 7, 2003, p. C14.

Supply and Demand in the Market for Money: The Liquidity Preference Framework

Whereas the loanable funds framework determines the equilibrium interest rate using the supply of and demand for bonds, an alternative model developed by John Maynard Keynes, known as the **liquidity preference framework**, determines the equilibrium interest rate in terms of the supply of and demand for money. Although the two frameworks look different, the liquidity preference analysis of the market for money is closely related to the loanable funds framework of the bond market.[4]

The starting point of Keynes's analysis is his assumption that there are two main categories of assets that people use to store their wealth: money and bonds. Therefore, total wealth in the economy must equal the total quantity of bonds plus money in the economy, which equals the quantity of bonds supplied (B^s) plus the quantity of money supplied (M^s). The quantity of bonds (B^d) and money (M^d) that people want to hold and thus demand must also equal the total amount of wealth, because people cannot purchase more assets than their available resources allow. The conclusion is that the quantity of bonds and money supplied must equal the quantity of bonds and money demanded:

$$B^s + M^s = B^d + M^d \tag{2}$$

Collecting the bond terms on one side of the equation and the money terms on the other, this equation can be rewritten as:

$$B^s - B^d = M^d - M^s \tag{3}$$

The rewritten equation tells us that if the market for money is in equilibrium ($M^s = M^d$), the right-hand side of Equation 3 equals zero, implying that $B^s = B^d$, meaning that the bond market is also in equilibrium.

Thus it is the same to think about determining the equilibrium interest rate by equating the supply and demand for bonds or by equating the supply and demand for money. In this sense, the liquidity preference framework, which analyzes the market for money, is equivalent to the loanable funds framework, which analyzes the bond market. In practice, the approaches differ, because by assuming that there are only two kinds of assets, money and bonds, the liquidity preference approach implicitly ignores any effects on interest rates that arise from changes in the expected returns on real assets such as automobiles and houses. In most instances, however, both frameworks yield the same predictions.

The reason that we approach the determination of interest rates with both frameworks is that the loanable funds framework is easier to use when analyzing the effects from changes in expected inflation, whereas the liquidity preference framework provides a simpler analysis of the effects from changes in income, the price level, and the supply of money.

Because the definition of money that Keynes used includes currency (which earns no interest) and checking account deposits (which in his time typically earned little

[4]Note that the term *market for money* refers to the market for the medium of exchange, money. This market differs from the *money market* referred to by finance practitioners, which is the financial market in which short-term debt instruments are traded.

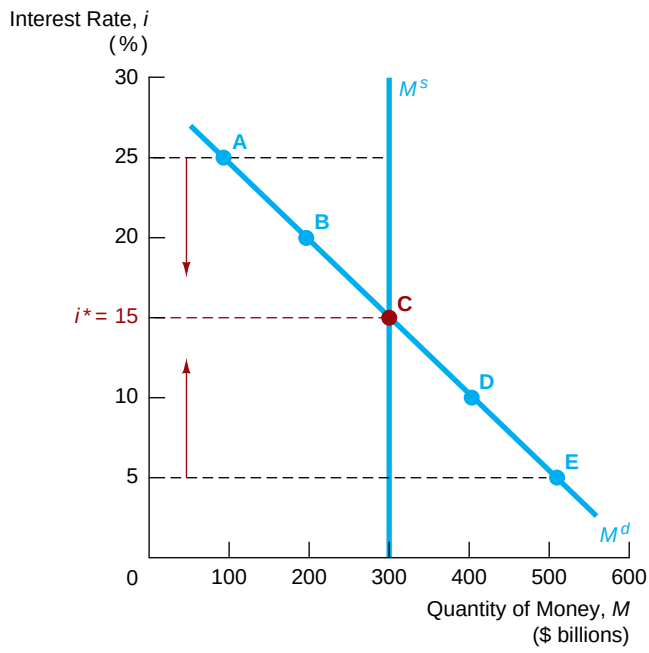
or no interest), he assumed that money has a zero rate of return. Bonds, the only alternative asset to money in Keynes's framework, have an expected return equal to the interest rate i.[5] As this interest rate rises (holding everything else unchanged), the expected return on money falls relative to the expected return on bonds, and as the theory of asset demand tells us, this causes the demand for money to fall.

We can also see that the demand for money and the interest rate should be negatively related by using the concept of **opportunity cost**, the amount of interest (expected return) sacrificed by not holding the alternative asset—in this case, a bond. As the interest rate on bonds, i, rises, the opportunity cost of holding money rises, and so money is less desirable and the quantity of money demanded must fall.

Figure 9 shows the quantity of money demanded at a number of interest rates, with all other economic variables, such as income and the price level, held constant. At an interest rate of 25%, point A shows that the quantity of money demanded is $100 billion. If the interest rate is at the lower rate of 20%, the opportunity cost of money is lower, and the quantity of money demanded rises to $200 billion, as indicated by the move from point A to point B. If the interest rate is even lower, the quantity of money demanded is even higher, as is indicated by points C, D, and E. The curve M^d connecting these points is the demand curve for money, and it slopes downward.

At this point in our analysis, we will assume that a central bank controls the amount of money supplied at a fixed quantity of $300 billion, so the supply curve for

FIGURE 9 Equilibrium in the Market for Money



[5]Keynes did not actually assume that the expected returns on bonds equaled the interest rate but rather argued that they were closely related (see Chapter 24). This distinction makes no appreciable difference in our analysis.

money M^s in the figure is a vertical line at $300 billion. The equilibrium where the quantity of money demanded equals the quantity of money supplied occurs at the intersection of the supply and demand curves at point C, where

$$M^d = M^s \tag{4}$$

The resulting equilibrium interest rate is at $i^* = 15\%$.

We can again see that there is a tendency to approach this equilibrium by first looking at the relationship of money demand and supply when the interest rate is above the equilibrium interest rate. When the interest rate is 25%, the quantity of money demanded at point A is $100 billion, yet the quantity of money supplied is $300 billion. The excess supply of money means that people are holding more money than they desire, so they will try to get rid of their excess money balances by trying to buy bonds. Accordingly, they will bid up the price of bonds, and as the bond price rises, the interest rate will fall toward the equilibrium interest rate of 15%. This tendency is shown by the downward arrow drawn at the interest rate of 25%.

Likewise, if the interest rate is 5%, the quantity of money demanded at point E is $500 billion, but the quantity of money supplied is only $300 billion. There is now an excess demand for money because people want to hold more money than they currently have. To try to obtain more money, they will sell their only other asset—bonds—and the price will fall. As the price of bonds falls, the interest rate will rise toward the equilibrium rate of 15%. Only when the interest rate is at its equilibrium value will there be no tendency for it to move further, and the interest rate will settle to its equilibrium value.

Changes in Equilibrium Interest Rates in the Liquidity Preference Framework

Analyzing how the equilibrium interest rate changes using the liquidity preference framework requires that we understand what causes the demand and supply curves for money to shift.

Study Guide

Learning the liquidity preference framework also requires practicing applications. When there is an application in the text to examine how the interest rate changes because some economic variable increases, see if you can draw the appropriate shifts in the supply and demand curves when this same economic variable decreases. And remember to use the ceteris paribus assumption: When examining the effect of a change in one variable, hold all other variables constant.

Shifts in the Demand for Money

In Keynes's liquidity preference analysis, two factors cause the demand curve for money to shift: income and the price level.

Income Effect. In Keynes's view, there were two reasons why income would affect the demand for money. First, as an economy expands and income rises, wealth increases and people will want to hold more money as a store of value. Second, as the economy

expands and income rises, people will want to carry out more transactions using money, with the result that they will also want to hold more money. The conclusion is that ***a higher level of income causes the demand for money to increase and the demand curve to shift to the right***.

Price-Level Effect. Keynes took the view that people care about the amount of money they hold in real terms; that is, in terms of the goods and services that it can buy. When the price level rises, the same nominal quantity of money is no longer as valuable; it cannot be used to purchase as many real goods or services. To restore their holdings of money in real terms to its former level, people will want to hold a greater nominal quantity of money, so ***a rise in the price level causes the demand for money to increase and the demand curve to shift to the right***.

Shifts in the Supply of Money

We will assume that the supply of money is completely controlled by the central bank, which in the United States is the Federal Reserve. (Actually, the process that determines the money supply is substantially more complicated, involving banks, depositors, and borrowers from banks. We will study it in more detail later in the book.) For now, all we need to know is that ***an increase in the money supply engineered by the Federal Reserve will shift the supply curve for money to the right***.

Application

Changes in the Equilibrium Interest Rate Due to Changes in Income, the Price Level, or the Money Supply

To see how the liquidity preference framework can be used to analyze the movement of interest rates, we will again look at several applications that will be useful in evaluating the effect of monetary policy on interest rates. (As a study aid, Table 4 summarizes the shifts in the demand and supply curves for money.)

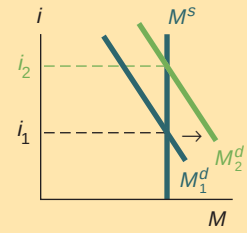
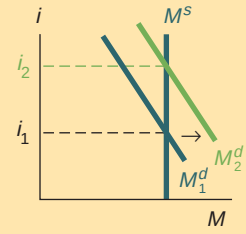
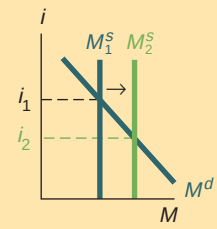
Changes in Income

When income is rising during a business cycle expansion, we have seen that the demand for money will rise, shown in Figure 10 by the shift rightward in the demand curve from M_1^d to M_2^d. The new equilibrium is reached at point 2 at the intersection of the M_2^d curve with the money supply curve M^s. As you can see, the equilibrium interest rate rises from i_1 to i_2. The liquidity preference framework thus generates the conclusion that ***when income is rising during a business cycle expansion (holding other economic variables constant), interest rates will rise***. This conclusion is unambiguous when contrasted to the conclusion reached about the effects of a change in income on interest rates using the loanable funds framework.

Changes in the Price Level

When the price level rises, the value of money in terms of what it can purchase is lower. To restore their purchasing power in real terms to its former level, people will want to hold a greater nominal quantity of money. A higher price level shifts the demand curve for money to the right from M_1^d to M_2^d (see Figure 10). The equilibrium moves from point 1 to point 2, where the equilibrium interest rate has risen from i_1 to i_2, illustrating that ***when the price level increases, with the supply of money and other economic variables held constant, interest rates will rise***.

SUMMARY **Table 4 Factors That Shift the Demand for and Supply of Money**

Variable	Change in Variable	Change in Money Demand (M^d) or Supply (M^s)	Change in Interest Rate	
Income	↑	M^d ↑	↑	
Price level	↑	M^d ↑	↑	
Money supply	↑	M^s ↑	↓	

Note: Only increases in the variables are shown. The effect of decreases in the variables on the change in demand would be the opposite of those indicated in the remaining columns.

Changes in the Money Supply

www.federalreserve.gov/releases/H6/Current

The Federal Reserve reports money supply data at 4:30 p.m. every Thursday.

An increase in the money supply due to expansionary monetary policy by the Federal Reserve implies that the supply curve for money shifts to the right. As is shown in Figure 11 by the movement of the supply curve from M_1^s to M_2^s, the equilibrium moves from point 1 down to point 2, where the M_2^s supply curve intersects with the demand curve M^d and the equilibrium interest rate has fallen from i_1 to i_2. ***When the money supply increases (everything else remaining equal), interest rates will decline.***[6]

[6]This same result can be generated using the loanable funds framework. As we will see in Chapters 15 and 16, the primary way that a central bank produces an increase in the money supply is by buying bonds and thereby decreasing the supply of bonds to the public. The resulting shift to the left of the supply curve for bonds will lead to a decline in the equilibrium interest rate.

FIGURE 10 Response to a Change in Income or the Price Level

In a business cycle expansion, when income is rising, or when the price level rises, the demand curve shifts from M_1^d to M_2^d. The supply curve is fixed at $M^s = \overline{M}$. The equilibrium interest rate rises from i_1 to i_2.

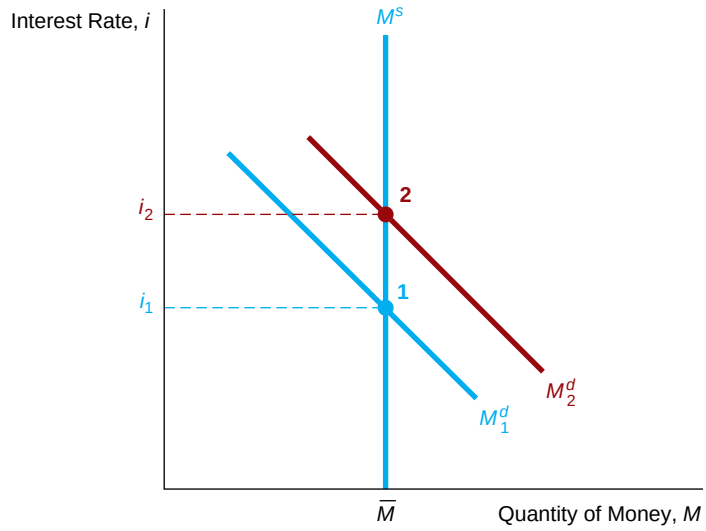
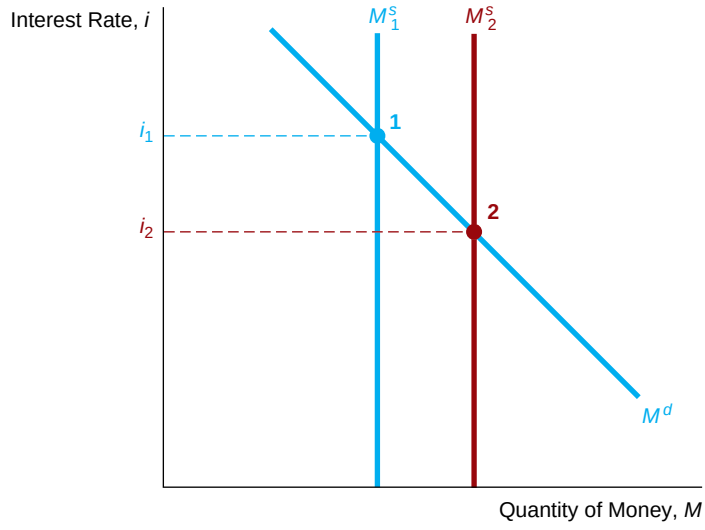


FIGURE 11 Response to a Change in the Money Supply

When the money supply increases, the supply curve shifts from M_1^s to M_2^s, and the equilibrium interest rate falls from i_1 to i_2.



Following the Financial News

Forecasting Interest Rates

Forecasting interest rates is a time-honored profession. Economists are hired (sometimes at very high salaries) to forecast interest rates, because businesses need to know what the rates will be in order to plan their future spending, and banks and investors require interest-rate forecasts in order to decide which assets to buy. Interest-rate forecasters predict what will happen to the factors that affect the supply and demand for bonds and for money—factors such as the strength of the economy, the profitability of investment opportunities, the expected inflation rate, and the size of government budget deficits and borrowing. They then use the supply and demand analysis we have outlined in this chapter to come up with their interest-rate forecasts.

The *Wall Street Journal* reports interest-rate forecasts by leading prognosticators twice a year (early January and July) in its "Economy" column or in its "Credit Markets" column, which surveys developments in the bond market daily. Forecasting interest rates is a perilous business. To their embarrassment, even the top experts are frequently far off in their forecasts.

The *Wall Street Journal* Forecasting Survey for 2003

In percent except for dollar vs. yen and dollar vs. euro

	JULY 2002 SURVEY						NEW FORECASTS FOR 2003									
	3-MO. TREASURY BILL-a Dec.	10-YR. TREASURY NOTE Dec.	GDP-b Q1–Q3 2002	CPI-c Nov.	$U.S. vs. YEN Dec.	UNEMPL. Nov.	3-MO. TREASURY BILLS-a June	10-YR. TREASURY NOTE June	GDP-b Q1 2003	GDP-b Q2 2003	GDP-b Q3 2003	GDP-b Q4 2003	CPI-c May	$U.S. vs. YEN June	$U.S. vs. EURO June	UNEMPL. May
Susan M. Sterne, Economic Analysis	2.50	5.30	2.9	2.1	125	5.8	2.25	5.50	**4.6**	4.0	4.2	4.3	2.5	115	1.10	5.6
Gail Fosler, The Conference Board	2.30	5.35	2.2	2.7	132	6.1	1.50	5.10	**4.2**	3.1	4.1	5.2	2.5	131	0.87	5.9
Stephen Gallagher, Societe Generale	2.15	5.60	2.4	2.7	120	5.9	1.25	4.50	**4.0**	3.0	3.5	3.5	2.3	125	1.00	5.7
Ian Shepherdson, High Frequency Economics	2.00	5.25	3.3	N.A.	135	6.5	1.25	4.75	**4.0**	4.0	5.0	5.0	2.2	N.A.	N.A.	6.5
James F. Smith, University of North Carolina	2.45	4.30	4.8	1.8	143	5.2	1.48	4.00	**3.8**	4.3	3.2	2.8	1.5	137	0.89	5.6
Lawrence Kudlow, Kudlow & Co. LLC	1.90	5.30	3.7	2.2	130	5.8	1.50	5.00	**3.6**	4.5	4.5	5.0	2.1	130	1.00	5.8
D. Malpass/J. Ryding, Bear Stearns	2.00	5.10	1.6	1.9	132	6.1	1.60	4.80	**3.6**	3.9	4.1	4.1	1.8	130	0.95	5.9
Michael K. Evans, Evans Carroll & Assoc.	1.75	5.00	1.0	2.1	130	6.1	1.30	4.30	**3.5**	2.0	3.5	2.5	3.0	135	1.00	6.2
Tracy Herrick, Jefferies & Company Inc.	2.60	5.40	2.3	1.8	118	5.4	1.20	4.00	**3.5**	3.0	3.5	4.0	2.7	125	1.05	5.7
David L. Littman, Comerica Bank	2.42	5.50	3.4	1.2	117	6.0	1.60	4.80	**3.5**	4.0	4.0	4.0	2.2	130	0.97	5.7
Paul McCulley, PIMCO	1.70	5.20	2.3	1.5	130	6.0	1.20	4.15	**3.5**	2.5	3.0	2.5	2.3	125	1.03	5.8
Henry Willmore, Barclays Capital	2.40	5.60	4.2	2.3	135	6.2	1.40	4.40	**3.5**	4.5	4.5	2.0	2.3	135	0.92	5.9
J. Meil/A. Raha, Eaton Corp.	2.80	5.40	N.A.	2.8	130	5.6	1.50	4.30	**3.3**	3.5	3.4	3.4	2.2	123	1.00	6.0
A. Hodge/W. Mak, Global Insight	2.30	5.60	1.2	2.0	125	6.0	1.30	5.60	**3.2**	3.3	3.9	4.5	2.2	123	0.98	6.1
Kurt Karl, Swiss Re	2.70	5.40	0.4	2.5	120	6.0	1.60	5.00	**3.2**	3.6	4.2	3.9	1.9	128	1.00	5.8
Richard D. Rippe, Prudential Securities	2.25	5.20	3.3	2.4	115	6.0	1.30	4.50	**3.2**	3.5	3.9	4.3	2.3	120	1.05	6.1
Daniel Laufenberg, American Express	2.30	5.15	3.1	2.5	125	5.6	1.50	4.50	**3.1**	3.0	4.1	3.8	2.0	120	1.04	5.7
John D. Mueller, LBMC LLC	2.50	5.60	4.7	1.6	115	5.5	1.50	4.90	**3.1**	5.5	6.7	6.0	0.9	118	0.96	5.6
Diane C. Swonk, Bank One, NA	3.00	5.10	2.5	2.0	123	5.7	1.44	4.00	**3.1**	2.9	3.3	3.3	2.9	123	0.98	6.1
David Wyss, Standard and Poor's	1.80	4.90	2.1	2.4	115	5.7	1.20	4.30	**3.1**	3.1	4.5	3.8	2.2	130	1.02	6.4
James W. Coons, Huntington National Bank	1.85	5.00	1.5	2.5	120	5.7	1.50	4.35	**3.0**	3.5	3.5	3.5	2.4	130	1.00	5.8
Richard DeKaser, National City Corporation	2.69	5.05	2.5	2.3	135	5.7	1.27	4.53	**3.0**	3.2	4.3	4.4	2.4	118	1.04	5.8
Neal Soss, CSFB	1.70	4.25	2.0	2.6	122	5.5	1.25	3.40	**3.0**	2.7	2.9	2.7	2.4	112	1.07	6.1
Brian S. Wesbury, Griffin Kubik Steph. & Thomp.	2.32	5.80	1.8	2.1	126	5.4	1.25	4.50	**3.0**	3.0	4.8	5.2	2.6	125	1.05	5.8
Stuart Hoffman, PNC Financial Services Group	2.00	5.30	1.3	2.3	125	5.7	1.25	4.05	**2.8**	3.0	3.5	3.5	2.4	125	1.02	5.8
John Lonski, Moody's Investors Service	2.30	5.40	2.4	1.8	125	5.5	1.53	4.40	**2.8**	2.9	3.5	3.7	2.3	122	1.04	5.7
R. T. McGee/T.W. Synnott, US Trust Co.	2.00	5.40	1.8	2.2	125	5.7	1.20	4.30	**2.8**	2.9	3.5	3.8	2.2	125	1.05	5.8
David Lereah, National Association of Realtors	N.A.	N.A.	N.A.	N.A.	N.A.	N.A.	1.80	4.40	**2.7**	3.0	3.6	3.2	2.4	125	0.98	5.7
Maria Fiorini Ramirez, MFR Inc.	2.00	5.00	1.7	1.7	123	5.9	1.25	4.00	**2.6**	2.0	2.5	2.7	2.1	128	1.05	6.2
J. Prakken/C. Varvares, Macroeconomic Adv.	N.A.	N.A.	N.A.	N.A.	N.A.	N.A.	1.20	4.32	**2.6**	3.4	3.7	3.7	2.0	121	1.01	5.7
David W. Berson, Fannie Mae	2.00	5.20	2.2	2.6	121	5.5	1.30	4.40	**2.5**	3.5	3.7	3.6	2.0	135	1.05	5.8

(continued)

N.A. Not Available; a Treasury bill rates are on a bond-equivalent basis; b Real gross domestic product, average annualized rate for first three quarters, based on January and July surveys; c Year-to-year change in the consumer price index; d David Rosenberg replaces Bruce Steinberg at Merrill Lynch; e Averages are for analysts polled at time of survey

Source: Wall Street Journal, Thursday, January 2, 2003, p. A2.

Following the Financial News

The *Wall Street Journal* Forecasting Survey for 2003 (*continued*)

In percent except for dollar vs. yen and dollar vs. euro

	JULY 2002 SURVEY						NEW FORECASTS FOR 2003									
	3-MO. TREASURY BILL-a Dec.	10-YR. TREASURY NOTE Dec.	GDP-b Q1–Q3 2002	CPI-c Nov.	$U.S. vs. YEN Dec.	UNEMPL. Nov.	3-MO. TREASURY BILLS-a June	10-YR. TREASURY NOTE June	GDP-b Q1 2003	GDP-b Q2 2003	GDP-b Q3 2003	GDP-b Q4 2003	CPI-c May	$U.S. vs. YEN June	$U.S. vs. EURO June	UNEMPL. May
Maury Harris, UBS Warburg	2.00	5.00	1.7	2.0	120	5.9	1.60	4.60	**2.5**	4.5	3.5	3.5	2.3	115	1.05	5.7
William B. Hummer, Wayne Hummer Invest.	2.21	5.05	2.2	2.4	119	5.5	1.31	4.14	**2.5**	3.1	3.6	3.8	2.1	125	1.05	5.8
R. Shrouds/R. Fry, DuPont Co.	1.80	5.00	2.9	2.1	110	5.8	1.30	4.50	**2.5**	3.0	3.5	3.5	1.9	128	1.05	6.0
Allen Sinai, Decision Economics Inc.	1.82	4.94	0.8	1.9	123	6.0	1.27	4.17	**2.5**	2.2	2.9	3.2	2.2	135	1.06	6.5
Sung Won Sohn, Wells Fargo & Co.	2.05	5.20	1.8	3.0	115	5.7	1.30	4.40	**2.5**	3.7	3.8	3.8	1.5	125	0.99	5.8
Gary Thayer, A.G. Edwards	2.20	5.60	2.0	1.8	120	5.5	1.40	4.50	**2.5**	3.5	3.0	4.5	2.1	119	1.06	5.7
Mark Zandi, Economy.com	2.20	5.25	1.1	2.2	125	6.0	1.70	4.50	**2.4**	2.7	3.2	3.8	2.2	125	1.00	6.3
R. Berner/D. Greenlaw, Morgan Stanley	2.00	5.30	1.9	2.6	124	5.8	1.50	4.50	**2.3**	3.8	3.9	3.5	1.9	120	1.05	5.9
David Resler, Nomura Securities International	1.90	5.10	3.2	2.4	120	5.9	1.25	4.25	**2.3**	3.0	3.5	3.8	1.8	125	1.04	6.0
Edward Leamer, UCLA Anderson Forecast	N.A.	N.A.	N.A.	N.A.	N.A.	N.A.	2.03	4.00	**2.2**	2.3	2.7	3.2	2.4	N.A.	1.10	6.1
David Rosenberg, Merrill Lynch[d]	2.25	5.25	2.3	2.0	125	5.8	1.20	4.00	**2.2**	3.3	3.0	3.5	2.4	125	1.07	6.5
Saul Hymans, RSQE, University of Michigan	2.30	5.30	1.9	2.7	122	5.6	1.72	4.03	**2.1**	4.3	4.2	4.2	2.4	N.A.	N.A.	6.1
Nicholas S. Perna, Perna Associates	2.62	5.53	1.3	2.1	122	5.7	1.47	4.53	**2.1**	3.0	2.9	3.1	2.3	114	1.03	5.5
Richard Yamarone, Argus Research	3.00	5.65	3.1	2.8	128	4.7	1.70	4.60	**2.1**	2.5	3.0	2.5	3.3	128	1.00	5.6
Ram Bhagavatula, The Royal Bank of Scotland	2.45	5.35	3.8	2.4	121	5.4	1.10	3.75	**2.0**	2.8	4.1	4.3	21	127	1.06	6.1
J. Dewey Daane, Vanderbilt University	2.00	5.00	0.8	2.0	121	6.0	1.60	4.50	**2.0**	2.2	2.4	2.6	2.0	118	1.00	5.9
Peter Hooper, Deutsche Bank Securities	2.25	5.40	2.7	2.3	130	5.7	1.75	4.50	**2.0**	4.0	4.0	3.9	1.4	130	1.05	6.1
William T. Wilson, Ernst & Young	2.50	6.25	2.0	2.5	115	5.5	1.60	5.40	**2.0**	5.0	4.0	4.1	1.4	120	1.00	5.6
Robert DiClemente, Citibank SSB	1.90	5.30	2.6	1.9	125	5.8	1.30	4.60	**1.8**	2.7	3.3	4.1	1.9	132	0.93	5.9
Mike Cosgrove, Econoclast	2.00	5.30	2.1	2.3	130	6.0	1.30	4.30	**1.6**	2.9	3.5	4.0	2.4	125	1.00	6.0
William C. Dudley, Goldman Sachs	2.00	5.00	1.5	2.4	132	6.0	1.00	4.20	**1.5**	2.5	3.0	3.5	2.1	120	1.08	6.4
Ethan S. Harris, Lehman Brothers	N.A.	4.85	N.A.	2.2	116	6.1	1.20	4.20	**1.5**	3.0	3.5	4.0	2.3	124	1.07	6.2
Donald H. Straszheim, Straszheim Global Adv.	2.25	5.15	N.A.	1.9	114	5.9	1.40	4.40	**1.0**	2.0	4.0	4.0	2.0	127	1.04	6.1
A. Gary Shilling, A. Gary Shilling & Co.	1.50	4.00	–1.1	0.5	130	6.4	0.75	3.50	**–2.0**	–2.0	2.0	3.0	1.2	130	0.94	7.3
AVERAGE [e]	**2.20**	**5.20**	**2.3**	**2.2**	**122**	**5.8**	**1.41**	**4.42**	**2.7**	**3.2**	**3.7**	**3.7**	**2.2**	**125**	**1.02**	**6.0**
ACTUAL NUMBERS as of Dec. 31, 2002	**1.21**	**3.82**	**3.4**	**2.2**	**119**	**6.0**										

N.A. Not Available; a Treasury bill rates are on a bond-equivalent basis; b Real gross domestic product, average annualized rate for first three quarters, based on January and July surveys; c Year-to-year change in the consumer price index; d David Rosenberg replaces Bruce Steinberg at Merrill Lynch; e Averages are for analysts polled at time of survey

Source: Wall Street Journal, Thursday, January 2, 2003, p. A2.

Application

Money and Interest Rates

The liquidity preference analysis in Figure 11 seems to lead to the conclusion that an increase in the money supply will lower interest rates. This conclusion has important policy implications because it has frequently caused politicians to call for a more rapid growth of the money supply in order to drive down interest rates.

But is this conclusion that money and interest rates should be negatively related correct? Might there be other important factors left out of the liquidity preference analysis in Figure 11 that would reverse this conclusion? We will provide answers to these questions by applying the supply and demand analysis we have used in this chapter to obtain a deeper understanding of the relationship between money and interest rates.

An important criticism of the conclusion that a rise in the money supply lowers interest rates has been raised by Milton Friedman, a Nobel laureate in economics. He acknowledges that the liquidity preference analysis is correct and calls the result—that an increase in the money supply (*everything else*

remaining equal) lowers interest rates—the *liquidity effect*. However, he views the liquidity effect as merely part of the story: An increase in the money supply might not leave "everything else equal" and will have other effects on the economy that may make interest rates rise. If these effects are substantial, it is entirely possible that when the money supply rises, interest rates too may rise.

We have already laid the groundwork to discuss these other effects because we have shown how changes in income, the price level, and expected inflation affect the equilibrium interest rate.

Study Guide

To get further practice with the loanable funds and liquidity preference frameworks, show how the effects discussed here work by drawing the supply and demand diagrams that explain each effect. This exercise will also improve your understanding of the effect of money on interest rates.

1. *Income Effect*. Because an increasing money supply is an expansionary influence on the economy, it should raise national income and wealth. Both the liquidity preference and loanable funds frameworks indicate that interest rates will then rise (see Figures 7 and 10). Thus ***the income effect of an increase in the money supply is a rise in interest rates in response to the higher level of income***.

2. *Price-Level Effect*. An increase in the money supply can also cause the overall price level in the economy to rise. The liquidity preference framework predicts that this will lead to a rise in interest rates. So ***the price-level effect from an increase in the money supply is a rise in interest rates in response to the rise in the price level***.

3. *Expected-Inflation Effect*. The higher inflation rate that results from an increase in the money supply also affects interest rates by affecting the expected inflation rate. Specifically, an increase in the money supply may lead people to expect a higher price level in the future—hence the expected inflation rate will be higher. The loanable funds framework has shown us that this increase in expected inflation will lead to a higher level of interest rates. Therefore, ***the expected-inflation effect of an increase in the money supply is a rise in interest rates in response to the rise in the expected inflation rate***.

At first glance it might appear that the price-level effect and the expected-inflation effect are the same thing. They both indicate that increases in the price level induced by an increase in the money supply will raise interest rates. However, there is a subtle difference between the two, and this is why they are discussed as two separate effects.

Suppose that there is a onetime increase in the money supply today that leads to a rise in prices to a permanently higher level by next year. As the price level rises over the course of this year, the interest rate will rise via the price-level effect. Only at the end of the year, when the price level has risen to its peak, will the price-level effect be at a maximum.

The rising price level will also raise interest rates via the expected-inflation effect, because people will expect that inflation will be higher over the course of the year. However, when the price level stops rising next year, inflation and the expected inflation rate will return to zero. Any rise in interest rates as a result of the earlier rise in expected inflation will then be

reversed. We thus see that in contrast to the price-level effect, which reaches its greatest impact next year, the expected-inflation effect will have its smallest impact (zero impact) next year. The basic difference between the two effects, then, is that the price-level effect remains even after prices have stopped rising, whereas the expected-inflation effect disappears.

An important point is that the expected-inflation effect will persist only as long as the price level continues to rise. As we will see in our discussion of monetary theory in subsequent chapters, a onetime increase in the money supply will not produce a continually rising price level; only a higher rate of money supply growth will. Thus a higher rate of money supply growth is needed if the expected-inflation effect is to persist.

Does a Higher Rate of Growth of the Money Supply Lower Interest Rates?

We can now put together all the effects we have discussed to help us decide whether our analysis supports the politicians who advocate a greater rate of growth of the money supply when they feel that interest rates are too high. Of all the effects, only the liquidity effect indicates that a higher rate of money growth will cause a decline in interest rates. In contrast, the income, price-level, and expected-inflation effects indicate that interest rates will rise when money growth is higher. Which of these effects are largest, and how quickly do they take effect? The answers are critical in determining whether interest rates will rise or fall when money supply growth is increased.

Generally, the liquidity effect from the greater money growth takes effect immediately, because the rising money supply leads to an immediate decline in the equilibrium interest rate. The income and price-level effects take time to work, because it takes time for the increasing money supply to raise the price level and income, which in turn raise interest rates. The expected-inflation effect, which also raises interest rates, can be slow or fast, depending on whether people adjust their expectations of inflation slowly or quickly when the money growth rate is increased.

Three possibilities are outlined in Figure 12; each shows how interest rates respond over time to an increased rate of money supply growth starting at time T. Panel (a) shows a case in which the liquidity effect dominates the other effects so that the interest rate falls from i_1 at time T to a final level of i_2. The liquidity effect operates quickly to lower the interest rate, but as time goes by, the other effects start to reverse some of the decline. Because the liquidity effect is larger than the others, however, the interest rate never rises back to its initial level.

Panel (b) has a smaller liquidity effect than the other effects, with the expected-inflation effect operating slowly because expectations of inflation are slow to adjust upward. Initially, the liquidity effect drives down the interest rate. Then the income, price-level, and expected-inflation effects begin to raise it. Because these effects are dominant, the interest rate eventually rises above its initial level to i_2. In the short run, lower interest rates result from increased money growth, but eventually they end up climbing above the initial level.

Panel (c) has the expected-inflation effect dominating as well as operating rapidly because people quickly raise their expectations of inflation when the rate of money growth increases. The expected-inflation effect begins immediately to overpower the liquidity effect, and the interest rate immediately starts

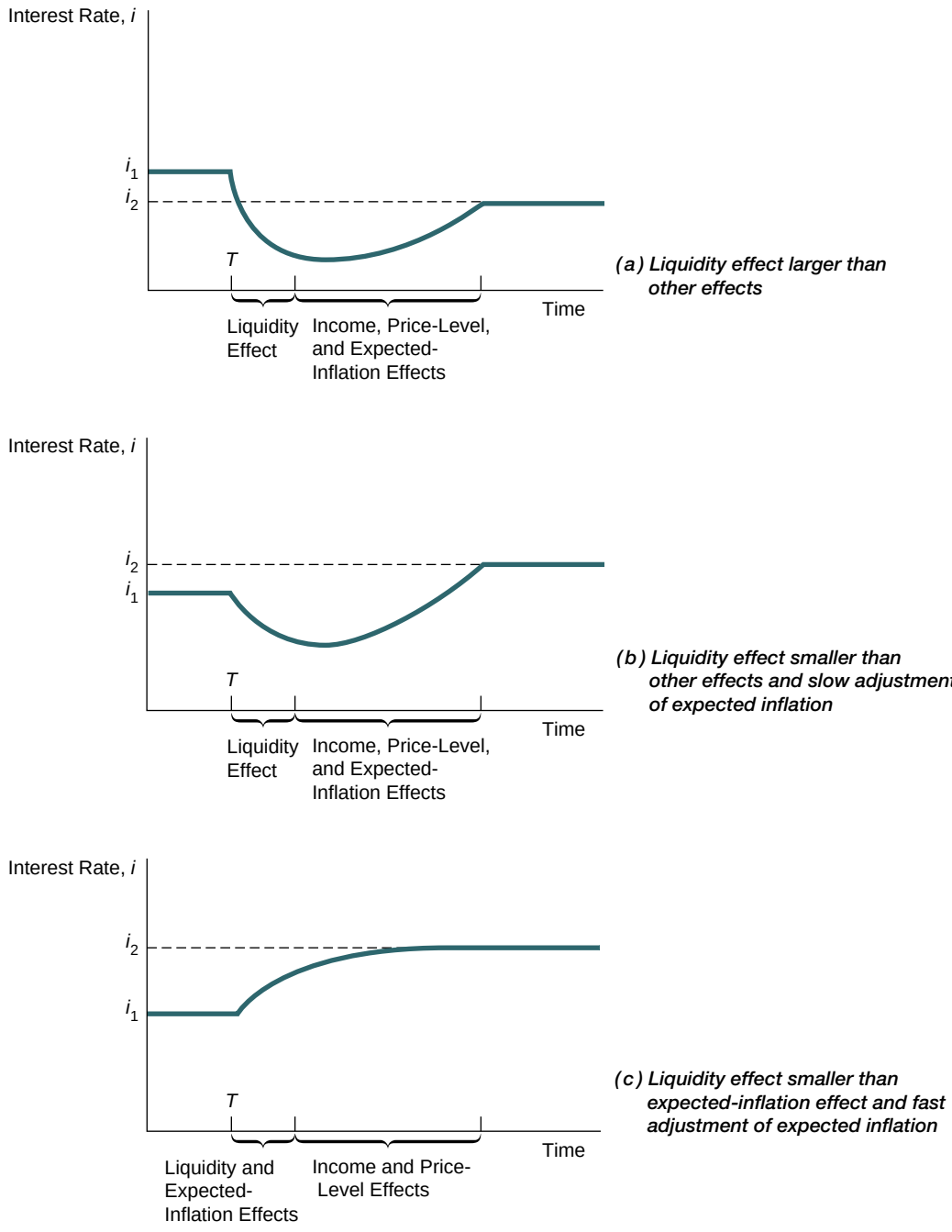


FIGURE 12 Response Over Time to an Increase in Money Supply Growth

to climb. Over time, as the income and price-level effects start to take hold, the interest rate rises even higher, and the eventual outcome is an interest rate that is substantially above the initial interest rate. The result shows clearly that increasing money supply growth is not the answer to reducing interest rates; rather, money growth should be reduced in order to lower interest rates!

An important issue for economic policymakers is which of these three scenarios is closest to reality. If a decline in interest rates is desired, then an increase in money supply growth is called for when the liquidity effect dominates the other effects, as in panel (a). A decrease in money growth is appropriate if the other effects dominate the liquidity effect and expectations of inflation adjust rapidly, as in panel (c). If the other effects dominate the liquidity effect but expectations of inflation adjust only slowly, as in panel (b), then whether you want to increase or decrease money growth depends on whether you care more about what happens in the short run or the long run.

Which scenario is supported by the evidence? The relationship of interest rates and money growth from 1950 to 2002 is plotted in Figure 13. When the rate of money supply growth began to climb in the mid-1960s, interest rates rose, indicating that the liquidity effect was dominated by the price-level, income, and expected-inflation effects. By the 1970s, interest rates reached

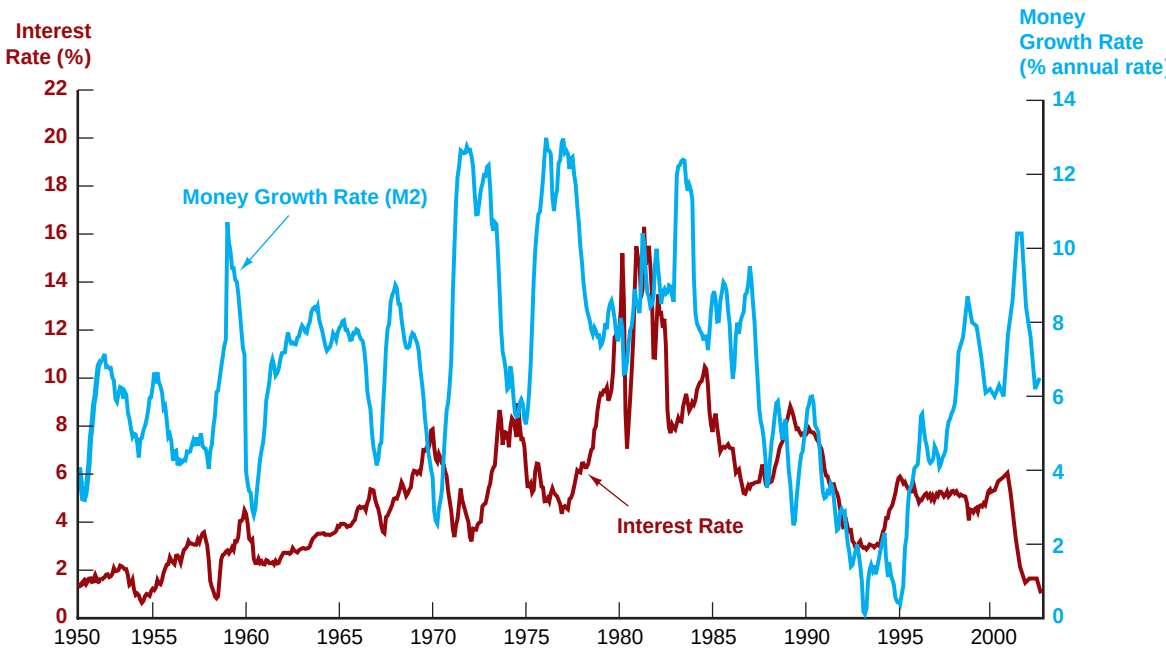


FIGURE 13 Money Growth (M2, Annual Rate) and Interest Rates (Three-Month Treasury Bills), 1950–2002

Sources: Federal Reserve: www.federalreserve.gov/releases/h6/hist/h6hist1.txt.

levels unprecedented in the post-World War II period, as did the rate of money supply growth.

The scenario depicted in panel (a) of Figure 12 seems doubtful, and the case for lowering interest rates by raising the rate of money growth is much weakened. Looking back at Figure 6, which shows the relationship between interest rates and expected inflation, you should not find this too surprising. The rise in the rate of money supply growth in the 1960s and 1970s is matched by a large rise in expected inflation, which would lead us to predict that the expected-inflation effect would be dominant. It is the most plausible explanation for why interest rates rose in the face of higher money growth. However, Figure 13 does not really tell us which one of the two scenarios, panel (b) or panel (c) of Figure 12, is more accurate. It depends critically on how fast people's expectations about inflation adjust. However, recent research using more sophisticated methods than just looking at a graph like Figure 13 do indicate that increased money growth temporarily lowers short-term interest rates.[7]

[7]See Lawrence J. Christiano and Martin Eichenbaum, "Identification and the Liquidity Effect of a Monetary Policy Shock," in *Business Cycles, Growth, and Political Economy,* ed. Alex Cukierman, Zvi Hercowitz, and Leonardo Leiderman (Cambridge, Mass.: MIT Press, 1992), pp. 335–370; Eric M. Leeper and David B. Gordon, "In Search of the Liquidity Effect," *Journal of Monetary Economics* 29 (1992): 341–370; Steven Strongin, "The Identification of Monetary Policy Disturbances: Explaining the Liquidity Puzzle," *Journal of Monetary Economics* 35 (1995): 463–497; Adrian Pagan and John C. Robertson, "Resolving the Liquidity Effect," *Federal Reserve Bank of St. Louis Review* 77 (May-June 1995): 33–54; and Ben S. Bernanke and Ilian Mihov, "Measuring Monetary Policy," *Quarterly Journal of Economics* 113, 3 (August 1998), pp. 869–902.

Summary

1. The theory of asset demand tells us that the quantity demanded of an asset is (a) positively related to wealth, (b) positively related to the expected return on the asset relative to alternative assets, (c) negatively related to the riskiness of the asset relative to alternative assets, and (d) positively related to the liquidity of the asset relative to alternative assets.
2. The supply and demand analysis for bonds, frequently referred to as the loanable funds framework, provides one theory of how interest rates are determined. It predicts that interest rates will change when there is a change in demand because of changes in income (or wealth), expected returns, risk, or liquidity or when there is a change in supply because of changes in the attractiveness of investment opportunities, the real cost of borrowing, or government activities.
3. An alternative theory of how interest rates are determined is provided by the liquidity preference framework, which analyzes the supply of and demand for money. It shows that interest rates will change when there is a change in the demand for money because of changes in income or the price level or when there is a change in the supply of money.
4. There are four possible effects of an increase in the money supply on interest rates: the liquidity effect, the income effect, the price-level effect, and the expected-inflation effect. The liquidity effect indicates that a rise in money supply growth will lead to a decline in interest rates; the other effects work in the opposite direction. The evidence seems to indicate that the income, price-level, and expected-inflation effects dominate the liquidity effect such that an increase in money supply growth leads to higher rather than lower interest rates.

Key Terms

asset market approach, p. 93
demand curve, p. 87
expected return, p. 86
excess demand, p. 90
excess supply, p. 90
Fisher effect, p. 100
liquidity, p. 86
liquidity preference framework, p. 105
loanable funds, p. 92
loanable funds framework, p. 92
market equilibrium, p. 90
opportunity cost, p. 106
risk, p. 86
supply curve, p. 90
theory of asset demand, p. 87
wealth, p. 86

Questions and Problems

Questions marked with an asterisk are answered at the end of the book in an appendix, "Answers to Selected Questions and Problems."

1. Explain why you would be more or less willing to buy a share of Microsoft stock in the following situations:
   a. Your wealth falls.
   b. You expect the stock to appreciate in value.
   c. The bond market becomes more liquid.
   d. You expect gold to appreciate in value.
   e. Prices in the bond market become more volatile.

*2. Explain why you would be more or less willing to buy a house under the following circumstances:
   a. You just inherited $100,000.
   b. Real estate commissions fall from 6% of the sales price to 5% of the sales price.
   c. You expect Microsoft stock to double in value next year.
   d. Prices in the stock market become more volatile.
   e. You expect housing prices to fall.

3. Explain why you would be more or less willing to buy gold under the following circumstances:
   a. Gold again becomes acceptable as a medium of exchange.
   b. Prices in the gold market become more volatile.
   c. You expect inflation to rise, and gold prices tend to move with the aggregate price level.
   d. You expect interest rates to rise.

*4. Explain why you would be more or less willing to buy long-term AT&T bonds under the following circumstances:
   a. Trading in these bonds increases, making them easier to sell.
   b. You expect a bear market in stocks (stock prices are expected to decline).
   c. Brokerage commissions on stocks fall.
   d. You expect interest rates to rise.
   e. Brokerage commissions on bonds fall.

5. What would happen to the demand for Rembrandts if the stock market undergoes a boom? Why?

Answer each question by drawing the appropriate supply and demand diagrams.

*6. An important way in which the Federal Reserve decreases the money supply is by selling bonds to the public. Using a supply and demand analysis for bonds, show what effect this action has on interest rates. Is your answer consistent with what you would expect to find with the liquidity preference framework?

7. Using both the liquidity preference framework and supply and demand for bonds framework, show why interest rates are procyclical (rising when the economy is expanding and falling during recessions).

*8. Why should a rise in the price level (but not in expected inflation) cause interest rates to rise when the nominal money supply is fixed?

9. Find the "Credit Markets" column in the *Wall Street Journal*. Underline the statements in the column that explain bond price movements, and draw the appropriate supply and demand diagrams that support these statements.

10. What effect will a sudden increase in the volatility of gold prices have on interest rates?

*11. How might a sudden increase in people's expectations of future real estate prices affect interest rates?

12. Explain what effect a large federal deficit might have on interest rates.

*13. Using both the supply and demand for bonds and liquidity preference frameworks, show what the effect is on interest rates when the riskiness of bonds rises. Are the results the same in the two frameworks?

14. If the price level falls next year, remaining fixed thereafter, and the money supply is fixed, what is likely to happen to interest rates over the next two years? (*Hint:* Take account of both the price-level effect and the expected-inflation effect.)

*15. Will there be an effect on interest rates if brokerage commissions on stocks fall? Explain your answer.

Using Economic Analysis to Predict the Future

16. The president of the United States announces in a press conference that he will fight the higher inflation rate with a new anti-inflation program. Predict what will happen to interest rates if the public believes him.

*17. The chairman of the Fed announces that interest rates will rise sharply next year, and the market believes him. What will happen to today's interest rate on AT&T bonds, such as the $8\frac{1}{8}$s of 2022?

18. Predict what will happen to interest rates if the public suddenly expects a large increase in stock prices.

*19. Predict what will happen to interest rates if prices in the bond market become more volatile.

20. If the next chair of the Federal Reserve Board has a reputation for advocating an even slower rate of money growth than the current chair, what will happen to interest rates? Discuss the possible resulting situations.

Web Exercises

1. One of the largest single influences on the level of interest rates is inflation. There are a number of sites that report inflation over time. Go to ftp://ftp.bls.gov/pub/special.requests/cpi/cpiai.txt and review the data available. Note that the last columns report various averages. Move this data into a spreadsheet using the method discussed in the Web exploration at the end of Chapter 1. What has the average rate of inflation been since 1950, 1960, 1970, 1980, and 1990? What year had the lowest level of inflation? What year had the highest level of inflation?

2. Increasing prices erodes the purchasing power of the dollar. It is interesting to compute what goods would have cost at some point in the past after adjusting for inflation. Go to www.interest.com/hugh/calc/cpi.cgi. What would a car that cost $22,000 today have cost the year that you were born?

3. One of the points made in this chapter is that inflation erodes investment returns. Go to www.src-net.com/InvestmentMultiplier/iminflation.htm and review how changes in inflation alter your real return. What happens to the difference between the adjusted value of an investment compared to its inflation-adjusted value as:
   a. Inflation increases?
   b. The investment horizon lengthens?
   c. Expected returns increase?

appendix 1 to chapter 5

Models of Asset Pricing

In Chapter 4, we saw that the return on an asset (such as a bond) measures how much we gain from holding that asset. When we make a decision to buy an asset, we are influenced by what we expect the return on that asset to be and its risk. Here we show how to calculate expected return and risk, which is measured by the standard deviation.

Expected Return

If a Mobil Oil Corporation bond, for example, has a return of 15% half of the time and 5% the other half of the time, its expected return (which you can think of as the average return) is 10%. More formally, the expected return on an asset is the weighted average of all possible returns, where the weights are the probabilities of occurrence of that return:

$$R^e = p_1R_1 + p_2R_2 + \ldots + p_nR_n \tag{1}$$

where R^e = expected return
n = number of possible outcomes (states of nature)
R_i = return in the ith state of nature
p_i = probability of occurrence of the return R_i

EXAMPLE 1: Expected Return

What is the expected return on the Mobil Oil bond if the return is 12% two-thirds of the time and 8% one-third of the time?

Solution

The expected return is 10.68%:

$$R^e = p_1R_1 + p_2R_2$$

where

p_1 = probability of occurrence of return 1 $= \frac{2}{3} = 0.67$
R_1 = return in state 1 $= 12\% = 0.12$

p_2 = probability of occurrence return 2 $= \frac{1}{3} = .33$
R_2 = return in state 2 $= 8\% = 0.08$

Thus:

$$R^e = (0.67)(0.12) + (0.33)(0.08) = 0.1068 = 10.68\%$$

Calculating Standard Deviation of Returns

The degree of risk or uncertainty of an asset's returns also affects the demand for the asset. Consider two assets, stock in Fly-by-Night Airlines and stock in Feet-on-the-Ground Bus Company. Suppose that Fly-by-Night stock has a return of 15% half of the time and 5% the other half of the time, making its expected return 10%, while stock in Feet-on-the-Ground has a fixed return of 10%. Fly-by-Night stock has uncertainty associated with its returns and so has greater risk than stock in Feet-on-the-Ground, whose return is a sure thing.

To see this more formally, we can use a measure of risk called the standard deviation. The standard deviation of returns on an asset is calculated as follows. First calculate the expected return, R^e; then subtract the expected return from each return to get a deviation; then square each deviation and multiply it by the probability of occurrence of that outcome; finally, add up all these weighted squared deviations and take the square root. The formula for the standard deviation, σ, is thus:

$$\sigma = \sqrt{p_1(R_1 - R^e)^2 + p_2(R_2 - R^e)^2 + \ldots + p_n(R_n - R^e)^2} \qquad (2)$$

The higher the standard deviation, σ, the greater the risk of an asset.

EXAMPLE 2: Standard Deviation

What is the standard deviation of the returns on the Fly-by-Night Airlines stock and Feet-on-the-Ground Bus Company, with the same return outcomes and probabilities described above? Of these two stocks, which is riskier?

Solution

Fly-by-Night Airlines has a standard deviation of returns of 5%.

$$\sigma = \sqrt{p_1(R_1 - R^e)^2 + p_2(R_2 - R^e)^2}$$

$$R^e = p_1R_1 + p_2R_2$$

where

p_1 = probability of occurrence of return 1 $= \frac{1}{2} = 0.50$
R_1 = return in state 1 $= 15\% = 0.15$
p_2 = probability of occurrence of return 2 $= \frac{1}{2} = 0.50$
R_2 = return in state 2 $= 5\% = 0.05$
R^e = expected return $= (0.50)(0.15) + (0.50)(0.05) = 0.10$

Thus:

$$\sigma = \sqrt{(0.50)(0.15 - 0.10)^2 + (0.50)(0.05 - 0.10)^2}$$

$$\sigma = \sqrt{(0.50)(0.0025) + (0.50)(0.0025)} = \sqrt{0.0025} = 0.05 = 5\%$$

Feet-on-the-Ground Bus Company has a standard deviation of returns of 0%.

$$\sigma = \sqrt{p_1(R_1 - R^e)^2}$$

$$R^e = p_1 R_1$$

where

p_1 = probability of occurrence of return 1 = 1.0
R_1 = return in state 1 = 10% = 0.10
R^e = expected return = (1.0)(0.10) = 0.10

Thus:

$$\sigma = \sqrt{(1.0)(0.10 - 0.10)^2}$$

$$= \sqrt{0} = 0 = 0\%$$

Clearly, Fly-by-Night Airlines is a riskier stock, because its standard deviation of returns of 5% is higher than the zero standard deviation of returns for Feet-on-the-Ground Bus Company, which has a certain return.

Benefits of Diversification

Our discussion of the theory of asset demand indicates that most people like to avoid risk; that is, they are risk-averse. Why, then, do many investors hold many risky assets rather than just one? Doesn't holding many risky assets expose the investor to more risk?

The old warning about not putting all your eggs in one basket holds the key to the answer: Because holding many risky assets (called *diversification*) reduces the overall risk an investor faces, diversification is beneficial. To see why this is so, let's look at some specific examples of how an investor fares on his investments when he is holding two risky securities.

Consider two assets: common stock of Frivolous Luxuries, Inc., and common stock of Bad Times Products, Unlimited. When the economy is strong, which we'll assume is one-half of the time, Frivolous Luxuries has high sales and the return on the stock is 15%; when the economy is weak, the other half of the time, sales are low and the return on the stock is 5%. On the other hand, suppose that Bad Times Products thrives when the economy is weak, so that its stock has a return of 15%, but it earns less when the economy is strong and has a return on the stock of 5%. Since both these stocks have an expected return of 15% half the time and 5% the other half of the time, both have an expected return of 10%. However, both stocks carry a fair amount of risk, because there is uncertainty about their actual returns.

Suppose, however, that instead of buying one stock or the other, Irving the Investor puts half his savings in Frivolous Luxuries stock and the other half in Bad

Times Products stock. When the economy is strong, Frivolous Luxuries stock has a return of 15%, while Bad Times Products has a return of 5%. The result is that Irving earns a return of 10% (the average of 5% and 15%) on his holdings of the two stocks. When the economy is weak, Frivolous Luxuries has a return of only 5% and Bad Times Products has a return of 15%, so Irving still earns a return of 10% regardless of whether the economy is strong or weak. Irving is better off from this strategy of diversification because his expected return is 10%, the same as from holding either Frivolous Luxuries or Bad Times Products alone, and yet he is not exposed to *any* risk.

Although the case we have described demonstrates the benefits of diversification, it is somewhat unrealistic. It is quite hard to find two securities with the characteristic that when the return of one is high, the return of the other is always low.[1] In the real world, we are more likely to find at best returns on securities that are independent of each other; that is, when one is high, the other is just as likely to be high as to be low.

Suppose that both securities have an expected return of 10%, with a return of 5% half the time and 15% the other half of the time. Sometimes both securities will earn the higher return and sometimes both will earn the lower return. In this case if Irving holds equal amounts of each security, he will on average earn the same return as if he had just put all his savings into one of these securities. However, because the returns on these two securities are independent, it is just as likely that when one earns the high 15% return, the other earns the low 5% return and vice versa, giving Irving a return of 10% (equal to the expected return). Because Irving is more likely to earn what he expected to earn when he holds both securities instead of just one, we can see that Irving has again reduced his risk through diversification.[2]

The one case in which Irving will not benefit from diversification occurs when the returns on the two securities move perfectly together. In this case, when the first security has a return of 15%, the other also has a return of 15% and holding both securities results in a return of 15%. When the first security has a return of 5%, the other has a return of 5% and holding both results in a return of 5%. The result of diversifying by holding both securities is a return of 15% half of the time and 5% the other half of the time, which is exactly the same set of returns that are earned by holding only one of the securities. Consequently, diversification in this case does not lead to any reduction of risk.

The examples we have just examined illustrate the following important points about diversification:

1. Diversification is almost always beneficial to the risk-averse investor since it reduces risk unless returns on securities move perfectly together (which is an extremely rare occurrence).
2. The less the returns on two securities move together, the more benefit (risk reduction) there is from diversification.

---

[1] Such a case is described by saying that the returns on the two securities are perfectly *negatively* correlated.

[2] We can also see that diversification in the example above leads to lower risk by examining the standard deviation of returns when Irving diversifies and when he doesn't. The standard deviation of returns if Irving holds only one of the two securities is $\sqrt{0.5 \times (15\% - 10\%)^2 + 0.5 \times (5\% - 10\%)^2} = 5\%$. When Irving holds equal amounts of each security, there is a probability of $^1/_4$ that he will earn 5% on both (for a total return of 5%), a probability of $^1/_4$ that he will earn 15% on both (for a total return of 15%), and a probability of $^1/_2$ that he will earn 15% on one and 5% on the other (for a total return of 10%). The standard deviation of returns when Irving diversifies is thus $\sqrt{0.25 \times (15\% - 10\%)^2 + 0.25 \times (5\% - 10\%)^2 + 0.5 \times (10\% - 10\%)^2} = 3.5\%$. Since the standard deviation of returns when Irving diversifies is lower than when he holds only one security, we can see that diversification has reduced risk.

Diversification and Beta

In the previous section, we demonstrated the benefits of diversification. Here, we examine diversification and the relationship between risk and returns in more detail. As a result, we obtain an understanding of two basic theories of asset pricing: the capital asset pricing model (CAPM) and arbitrage pricing theory (APT).

We start our analysis by considering a portfolio of n assets whose return is:

$$R_p = x_1R_1 + x_2R_2 + \cdots + x_nR_n \tag{3}$$

where R_p = the return on the portfolio of n assets
R_i = the return on asset i
x_i = the proportion of the portfolio held in asset i

The expected return on this portfolio, $E(R_p)$, equals

$$\begin{aligned} E(R_p) &= E(x_1R_1) + E(x_2R_2) + \cdots + E(x_nR_n) \\ &= x_1E(R_1) + x_2E(R_2) + \cdots + x_nE(R_n) \end{aligned} \tag{4}$$

An appropriate measure of the risk for this portfolio is the standard deviation of the portfolio's return (σ_p) or its squared value, the variance of the portfolio's return (σ_p^2), which can be written as:

$$\begin{aligned} \sigma_p^2 = E[R_p - E(R_p)]^2 &= E[\{x_1R_1 + \cdots + x_nR_n\} - \{x_1E(R_1) + \cdots + x_nE(R_n)\}]^2 \\ &= E[x_1\{R_1 - E(R_1)\} + \cdots + x_n\{R_n - E(R_n)\}]^2 \end{aligned}$$

This expression can be rewritten as:

$$\begin{aligned} &\sigma_p^2 = E[\{x_1[R_1 - E(R_1)] + \cdots + x_n[R_n - E(R_n)]\} \times \{R_p - E(R_p)\}] \\ &= x_1E[\{R_1 - E(R_1)\} \times \{R_p - E(R_p)\}] + \cdots + x_nE[\{R_n - E(R_n)\} \times \{R_p - E(R_p)\}] \end{aligned}$$

This gives us the following expression for the variance for the portfolio's return:

$$\sigma_p^2 = x_1\sigma_{1p} + x_2\sigma_{2p} + x_n\sigma_{np} \tag{5}$$

where

$$\begin{aligned} \sigma_{ip} &= \text{the covariance of the return on asset } i \\ &\text{with the portfolio's return} = E[\{R_i - E(R_i)\} \times \{R_p - E(R_p)\}] \end{aligned}$$

Equation 5 tells us that the contribution to risk of asset i to the portfolio is $x_i\sigma_{ip}$. By dividing this contribution to risk by the total portfolio risk (σ_p^2), we have the proportionate contribution of asset i to the portfolio risk:

$$x_i\sigma_{ip}/\sigma_p^2$$

The ratio σ_{ip}/σ_p^2 tells us about the sensitivity of asset i's return to the portfolio's return. The higher the ratio is, the more the value of the asset moves with changes in the

value of the portfolio, and the more asset *i* contributes to portfolio risk. Our algebraic manipulations have thus led to the following important conclusion: ***The marginal contribution of an asset to the risk of a portfolio depends not on the risk of the asset in isolation, but rather on the sensitivity of that asset's return to changes in the value of the portfolio.***

If the total of all risky assets in the market is included in the portfolio, then it is called the *market portfolio*. If we suppose that the portfolio, *p*, is the market portfolio, *m*, then the ratio σ_{im}/σ_m^2 is called the asset *i*'s *beta*, that is:

$$\beta_i = \sigma_{im}/\sigma_m^2 \tag{6}$$

where

$$\beta_i = \text{the beta of asset } i$$

An asset's beta then is a measure of the asset's marginal contribution to the risk of the market portfolio. A higher beta means that an asset's return is more sensitive to changes in the value of the market portfolio and that the asset contributes more to the risk of the portfolio.

Another way to understand beta is to recognize that the return on asset *i* can be considered as being made up of two components—one that moves with the market's return (R_m) and the other a random factor with an expected value of zero that is unique to the asset (ϵ_i) and so is uncorrelated with the market return:

$$R_i = \alpha_i + \beta_i R_m + \epsilon_i \tag{7}$$

The expected return of asset *i* can then be written as:

$$E(R_i) = \alpha_i + \beta_i E(R_m)$$

It is easy to show that β_i in the above expression is the beta of asset *i* we defined before by calculating the covariance of asset *i*'s return with the market return using the two equations above:

$$\sigma_{im} = E[\{R_i - E(R_i)\} \times \{R_m - E(R_m)\}] = E[\{\beta_i[R_m - E(R_m)] + \epsilon_i\} \times \{R_m - E(R_m)\}]$$

However, since ϵ_i is uncorrelated with R_m, $E[\{\epsilon_i\} \times \{R_m - E(R_m)\}] = 0$. Therefore,

$$\sigma_{im} = \beta_i \sigma_m^2$$

Dividing through by σ_m^2 gives us the following expression for β_i:

$$\beta_i = \sigma_{im}/\sigma_m^2$$

which is the same definition for beta we found in Equation 6.

The reason for demonstrating that the β_i in Equation 7 is the same as the one we defined before is that Equation 7 provides better intuition about how an asset's beta measures its sensitivity to changes in the market return. Equation 7 tells us that when

the beta of an asset is 1.0, it's return on average increases by 1 percentage point when the market return increases by 1 percentage point; when the beta is 2.0, the asset's return increases by 2 percentage points when the market return increases by 1 percentage point; and when the beta is 0.5, the asset's return only increases by 0.5 percentage point on average when the market return increases by 1 percentage point.

Equation 7 also tells us that we can get estimates of beta by comparing the average return on an asset with the average market return. For those of you who know a little econometrics, this estimate of beta is just an ordinary least squares regression of the asset's return on the market return. Indeed, the formula for the ordinary least squares estimate of $\beta_i = \sigma_{im}/\sigma_m^2$ is exactly the same as the definition of β_i earlier.

Systematic and Nonsystematic Risk

We can derive another important idea about the riskiness of an asset using Equation 7. The variance of asset *i*'s return can be calculated from Equation 7 as:

$$\sigma_i^2 = E[R_i - E(R_i)]^2 = E\{\beta_i[R_m - E(R_m)\} + \epsilon_i]^2$$

and since ϵ_i is uncorrelated with market return:

$$\sigma_i^2 = \beta_i^2\sigma_m^2 + \sigma_\epsilon^2$$

The total variance of the asset's return can thus be broken up into a component that is related to market risk, $\beta_i^2\sigma_m^2$, and a component that is unique to the asset, σ_ϵ^2. The $\beta_i^2\sigma_m^2$ component related to market risk is referred to as *systematic risk* and the σ_ϵ^2 component unique to the asset is called *nonsystematic risk*. We can thus write the total risk of an asset as being made up of systematic risk and nonsystematic risk:

$$\text{Total Asset Risk} = \text{Systematic Risk} + \text{Nonsystematic Risk} \quad (8)$$

Systematic and nonsystematic risk each have another feature that makes the distinction between these two types of risk important. Systematic risk is the part of an asset's risk that cannot be eliminated by holding the asset as part of a diversified portfolio, whereas nonsystematic risk is the part of an asset's risk that can be eliminated in a diversified portfolio. Understanding these features of systematic and nonsystematic risk leads to the following important conclusion: ***The risk of a well-diversified portfolio depends only on the systematic risk of the assets in the portfolio.***

We can see that this conclusion is true by considering a portfolio of *n* assets, each of which has the same weight on the portfolio of (1/*n*). Using Equation 7, the return on this portfolio is:

$$R_p = (1/n)\sum_{i=1}^{n}\alpha_i + (1/n)\sum_{i=1}^{n}\beta_i R_m + (1/n)\sum_{i=1}^{n}\epsilon_i$$

which can be rewritten as:

$$R_p = \overline{\alpha} + \overline{\beta}R_m + 1/n)\sum_{i=1}^{n}\epsilon_i$$

where

$$\bar{\alpha} = \text{the average of the } \alpha_i\text{'s} = (1/n)\sum_{i=1}^{n}\alpha_i$$

$$\bar{\beta} = \text{the average of the } \beta_i\text{'s} = (1/n)\sum_{i=1}^{n}\alpha_i$$

If the portfolio is well diversified so that the ϵ_i's are uncorrelated with each other, then using this fact and the fact that all the ϵ_i's are uncorrelated with the market return, the variance of the portfolio's return is calculated as:

$$\sigma_p^2 = \bar{\beta}^2\sigma_m^2 + (1/n)(\text{average varience of } \epsilon_i)$$

As n gets large the second term, $(1/n)$(average variance of ϵ_i), becomes very small, so that a well-diversified portfolio has a risk of $\bar{\beta}^2\sigma_m^2$, which is only related to systematic risk. As the previous conclusion indicated, nonsystematic risk can be eliminated in a well-diversified portfolio. This reasoning also tells us that the risk of a well-diversified portfolio is greater than the risk of the market portfolio if the average beta of the assets in the portfolio is greater than one; however, the portfolio's risk is less than the market portfolio if the average beta of the assets is less than one.

The Capital Asset Pricing Model (CAPM)

We can now use the ideas we developed about systematic and nonsystematic risk and betas to derive one of the most widely used models of asset pricing—the capital asset pricing model (CAPM) developed by William Sharpe, John Litner, and Jack Treynor.

Each cross in Figure 1 shows the standard deviation and expected return for each risky asset. By putting different proportions of these assets into portfolios, we can generate a standard deviation and expected return for each of the portfolios using Equations 4 and 5. The shaded area in the figure shows these combinations of standard deviation and expected return for these portfolios. Since risk-averse investors always prefer to have higher expected return and lower standard deviation of the return, the most attractive standard deviation-expected return combinations are the ones that lie along the heavy line, which is called the *efficient portfolio frontier*. These are the standard deviation-expected return combinations risk-averse investors would always prefer.

The capital asset pricing model assumes that investors can borrow and lend as much as they want at a risk-free rate of interest, R_f. By lending at the risk-free rate, the investor earns an expected return of R_f and his investment has a zero standard deviation because it is risk-free. The standard deviation-expected return combination for this risk-free investment is marked as point A in Figure 1. Suppose an investor decides to put half of his total wealth in the risk-free loan and the other half in the portfolio on the efficient portfolio frontier with a standard deviation-expected return combination marked as point M in the figure. Using Equation 4, you should be able to verify that the expected return on this new portfolio is halfway between R_f and $E(R_m)$; that is, $[R_f + E(R_m)]/2$. Similarly, because the covariance between the risk-free return and the return on portfolio M must necessarily be zero, since there is no uncertainty about the

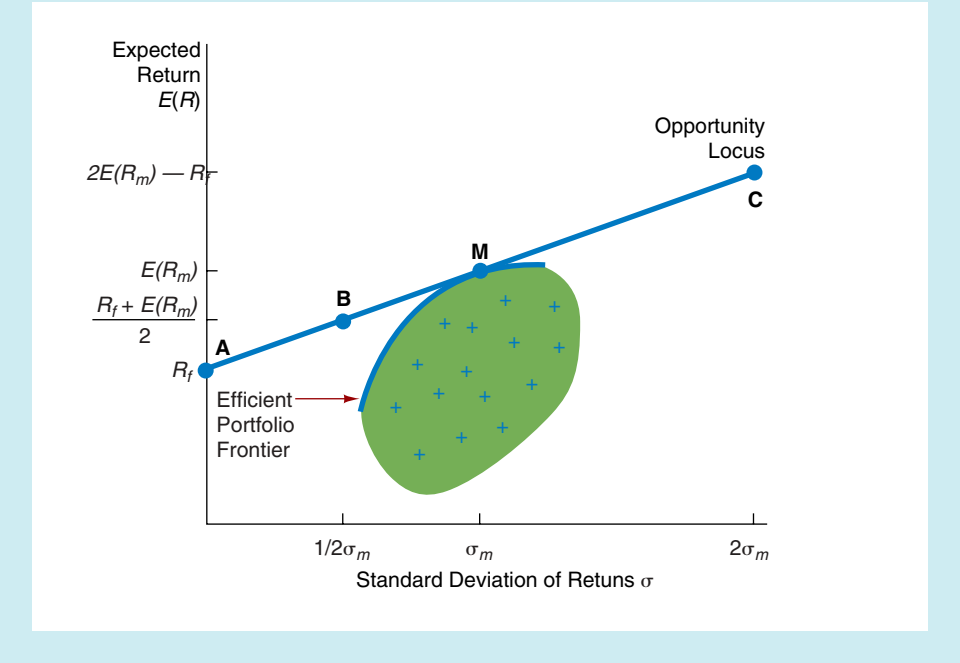


FIGURE 1 Risk Expected Return Trade-off

The crosses show the combination of standard deviation and expected return for each risky asset. The efficient portfolio frontier indicates the most preferable standard deviation-expected return combinations that can be achieved by putting risky assets into portfolios. By borrowing and lending at the risk-free rate and investing in portfolio M, the investor can obtain standard deviation-expected return combinations that lie along the line connecting A, B, M, and C. This line, the *opportunity locus*, contains the best combinations of standard deviations and expected returns available to the investor; hence the opportunity locus shows the trade-off between expected returns and risk for the investor.

return on the risk-free loan, you should also be able to verify, using Equation 5, that the standard deviation of the return on the new portfolio is halfway between zero and $\boldsymbol{\sigma}_m$, that is, $(1/2)\boldsymbol{\sigma}_m$. The standard deviation-expected return combination for this new portfolio is marked as point B in the figure, and as you can see it lies on the line between point A and point M. Similarly, if an investor borrows the total amount of her wealth at the risk-free rate R_f and invests the proceeds plus her wealth (that is, twice her wealth) in portfolio M, then the standard deviation of this new portfolio will be twice the standard deviation of return on portfolio M, $2\boldsymbol{\sigma}_m$. On the other hand, using Equation 4, the expected return on this new portfolio is $E(R_m)$ plus $E(R_m) - R_f$, which equals $2E(R_m) - R_f$. This standard deviation-expected return combination is plotted as point C in the figure.

You should now be able to see that both point B and point C are on the line connecting point A and point M. Indeed, by choosing different amounts of borrowing and lending, an investor can form a portfolio with a standard deviation-expected return combination that lies anywhere on the line connecting points A and M. You may have noticed that point M has been chosen so that the line connecting points A and M is tangent to the efficient portfolio frontier. The reason for choosing point M in this way is that it leads to standard deviation-expected return combinations along the line that are the most desirable for a risk-averse investor. This line can be thought of as the *opportunity locus*, which shows the best combinations of standard deviations and expected returns available to the investor.

The capital asset pricing model makes another assumption: All investors have the same assessment of the expected returns and standard deviations of all assets. In this case, portfolio M is the same for all investors. Thus when all investors' holdings of portfolio M are added together, they must equal all of the risky assets in the market,

which is just the market portfolio. The assumption that all investors have the same assessment of risk and return for all assets thus means that portfolio M is the market portfolio.Therefore, the R_m and σ_m in Figure 1 are identical to the market return, R_m, and the standard deviation of this return, σ_m, referred to earlier in this appendix.

The conclusion that the market portfolio and portfolio M are one and the same means that the opportunity locus in Figure 1 can be thought of as showing the trade-off between expected returns and increased risk for the investor. This trade-off is given by the slope of the opportunity locus, $E(R_m) - R_f$, and it tells us that when an investor is willing to increase the risk of his portfolio by σ_m, then he can earn an additional expected return of $E(R_m) - R_f$. The market price of a unit of market risk, σ_m, is $E(R_m) - R_f$. $E(R_m) - R_f$ is therefore referred to as the *market price of risk*.

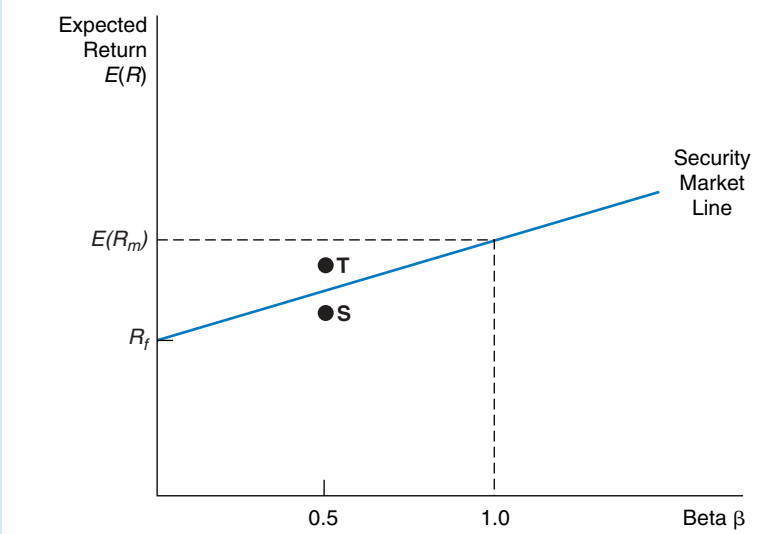
We now know that market price of risk is $E(R_m) - R_f$ and we also have learned that an asset's beta tells us about systematic risk, because it is the marginal contribution of that asset to a portfolio's risk. Therefore the amount an asset's expected return exceeds the risk-free rate, $E(R_i) - R_f$, should equal the market price of risk times the marginal contribution of that asset to portfolio risk, $[E(R_m) - R_f]\beta_i$. This reasoning yields the CAPM asset pricing relationship:

$$E(R_i) = R_f + \beta_i[E(R_m) - R_f] \tag{9}$$

This CAPM asset pricing equation is represented by the upward sloping line in Figure 2, which is called the *security market line*. It tells us the expected return that the market sets for a security given its beta. For example, it tells us that if a security has a beta of 1.0 so that its marginal contribution to a portfolio's risk is the same as the market portfolio, then it should be priced to have the same expected return as the market portfolio, $E(R_m)$.

FIGURE 2 Security Market Line

The security market line derived from the capital asset pricing model describes the relationship between an asset's beta and its expected return.



To see that securities should be priced so that their expected return-beta combination should lie on the security market line, consider a security like S in Figure 2, which is below the security market line. If an investor makes an investment in which half is put into the market portfolio and half into a risk-free loan, then the beta of this investment will be 0.5, the same as security S. However, this investment will have an expected return on the security market line, which is greater than that for security S. Hence investors will not want to hold security S and its current price will fall, thus raising its expected return until it equals the amount indicated on the security market line. On the other hand, suppose there is a security like T which has a beta of 0.5 but whose expected return is above the security market line. By including this security in a well-diversified portfolio with other assets with a beta of 0.5, none of which can have an expected return less than that indicated by the security line (as we have shown), investors can obtain a portfolio with a higher expected return than that obtained by putting half into a risk-free loan and half into the market portfolio. This would mean that all investors would want to hold more of security T, and so its price would rise, thus lowering its expected return until it equaled the amount indicated on the security market line.

The capital asset pricing model formalizes the following important idea: ***An asset should be priced so that is has a higher expected return not when it has a greater risk in isolation, but rather when its systematic risk is greater.***

Arbitrage Pricing Theory

Although the capital asset pricing model has proved to be very useful in practice, deriving it does require the adoption of some unrealistic assumptions; for example, the assumption that investors can borrow and lend freely at the risk-free rate, or the assumption that all investors have the same assessment of expected returns and standard deviations of returns for all assets. An important alternative to the capital asset pricing model is the arbitrage pricing theory (APT) developed by Stephen Ross of M.I.T.

In contrast to CAPM, which has only one source of systematic risk, the market return, APT takes the view that there can be several sources of systematic risk in the economy that cannot be eliminated through diversification. These sources of risk can be thought of as factors that may be related to such items as inflation, aggregate output, default risk premiums, and/or the term structure of interest rates. The return on an asset i can thus be written as being made up of components that move with these factors and a random component that is unique to the asset (ϵ_i):

$$R_i = \beta_i^1(\text{factor } 1) + \beta_i^2(\text{factor } 2) + \cdots + \beta_i^k(\text{factor } k) + \epsilon_i \qquad (10)$$

Since there are k factors, this model is called a k-factor model. The $\beta_i^1, \cdots, \beta_i^k$ describe the sensitivity of the asset i's return to each of these factors.

Just as in the capital asset pricing model, these systematic sources of risk should be priced. The market price for each factor j can be thought of as $E(R_{\text{factor } j}) - R_f$, and hence the expected return on a security can be written as:

$$E(R_i) = R_f + \beta_i^1[E(R_{\text{factor } 1}) - R_f] + \cdots + \beta_i^k[E(R_{\text{factor } k}) - R_f] \qquad (11)$$

This asset pricing equation indicates that all the securities should have the same market price for the risk contributed by each factor. If the expected return for a security were above the amount indicated by the APT pricing equation, then it would provide a higher expected return than a portfolio of other securities with the same average sensitivity to each factor. Hence investors would want to hold more of this security and its price would rise until the expected return fell to the value indicated by the APT pricing equation. On the other hand, if the security's expected return were less than the amount indicated by the APT pricing equation, then no one would want to hold this security, because a higher expected return could be obtained with a portfolio of securities with the same average sensitivity to each factor. As a result, the price of the security would fall until its expected return rose to the value indicated by the APT equation.

As this brief outline of arbitrage pricing theory indicates, the theory supports a basic conclusion from the capital asset pricing model: An asset should be priced so that it has a higher expected return not when it has a greater risk in isolation, but rather when its systematic risk is greater. There is still substantial controversy about whether a variant of the capital asset pricing model or the arbitrage pricing theory is a better description of reality. At the present time, both frameworks are considered valuable tools for understanding how risk affects the prices of assets.

appendix 2 to chapter 5

Applying the Asset Market Approach to a Commodity Market: The Case of Gold

Both models of interest-rate determination in Chapter 4 make use of an asset market approach in which supply and demand are always considered in terms of stocks of assets (amounts at a given point in time). The asset market approach is useful in understanding not only why interest rates fluctuate but also how any asset's price is determined.

One asset that has fascinated people for thousands of years is gold. It has been a driving force in history: The conquest of the Americas by Europeans was to a great extent the result of the quest for gold, to cite just one example. The fascination with gold continues to the present day, and developments in the gold market are followed closely by financial analysts and the media. This appendix shows how the asset market approach can be applied to understanding the behavior of commodity markets, in particular the gold market. (The analysis in this appendix can also be used to understand behavior in many other asset markets.)

Supply and Demand in the Gold Market

The analysis of a commodity market, such as the gold market, proceeds in a similar fashion to the analysis of the bond market by examining the supply of and demand for the commodity. We again use our analysis of the determinants of asset demand to obtain a demand curve for gold, which shows the relationship between the quantity of gold demanded and the price when all other economic variables are held constant.

Demand Curve

To derive the relationship between the quantity of gold demanded and its price, we again recognize that an important determinant of the quantity demanded is its expected return:

$$R^e = \frac{P^e_{t+1} - P_t}{P_t} = g^e$$

where
R^e = expected return
P_t = price of gold today
P^e_{t+1} = expected price of gold next year
g^e = expected capital gain

In deriving the demand curve, we hold all other variables constant, particularly the expected price of gold next year P^e_{t+1}. With a given value of the expected price of gold next year P^e_{t+1}, a lower price of gold today P_t means that there will be a greater

appreciation in the price of gold over the coming year. The result is that a lower price of gold today implies a higher expected capital gain over the coming year and hence a higher expected return: $R^e = (P^e_{t+1} - P_t)/P_t$. Thus because the price of gold today (which for simplicity we will denote as P) is lower, the expected return on gold is higher, and the quantity demanded is higher. Consequently, the demand curve G^d_1 slopes downward in Figure 1.

Supply Curve

To derive the supply curve, expressing the relationship between the quantity supplied and the price, we again assume that all other economic variables are held constant. A higher price of gold will induce producers to mine for extra gold and also possibly induce governments to sell some of their gold stocks to the public, thus increasing the quantity supplied. Hence the supply curve G^s_1 in Figure 1 slopes upward. Notice that the supply curve in the figure is drawn to be very steep. The reason for this is that the actual amount of gold produced in any year is only a tiny fraction of the outstanding stock of gold that has been accumulated over hundreds of years. Thus the increase in the quantity of the gold supplied in response to a higher price is only a small fraction of the stock of gold, resulting in a very steep supply curve.

Market Equilibrium

Market equilibrium in the gold market occurs when the quantity of gold demanded equals the quantity of gold supplied:

$$G^d = G^s$$

With the initial demand and supply curves of G^d_1 and G^s_1, equilibrium occurs at point 1, where these curves intersect at a gold price of P_1. At a price above this

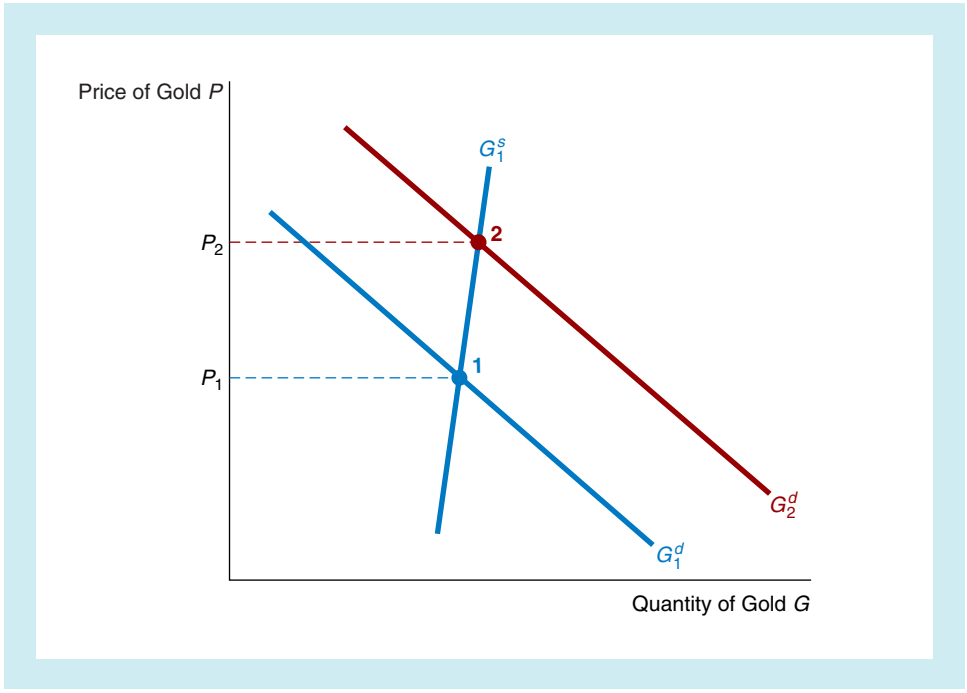


FIGURE 1 A Change in the Equilibrium Price of Gold

When the demand curve shifts rightward from G^d_1 to G^d_2—say, because expected inflation rises—equilibrium moves from point 1 to point 2, and the equilibrium price of gold rises from P_1 to P_2.

equilibrium, the amount of gold supplied exceeds the amount demanded, and this condition of excess supply leads to a decline in the gold price until it reaches P_1, the equilibrium price. Similarly, if the price is below P_1, there is excess demand for gold, which drives the price upward until it settles at the equilibrium price P_1.

Changes in the Equilibrium Price of Gold

Changes in the equilibrium price of gold occur when there is a shift in either the supply curve or the demand curve; that is, when the quantity demanded or supplied changes at each given price of gold in response to a change in some factor other than today's gold price.

Shift in the Demand Curve for Gold

Our analysis of the determinants of asset demand in the chapter provides the factors that shift the demand curve for gold: wealth, expected return on gold relative to alternative assets, riskiness of gold relative to alternative assets, and liquidity of gold relative to alternative assets. The analysis of how changes in each of these factors shift the demand curve for gold is the same as that found in the chapter.

When wealth rises, at a given price of gold, the quantity demanded increases, and the demand curve shifts to the right, as in Figure 1. When the expected return on gold relative to other assets rises—either because speculators think that the future price of gold will be higher or because the expected return on other assets declines—gold becomes more desirable; the quantity demanded therefore increases at any given price of gold, and the demand curve shifts to the right, as in Figure 1. When the relative riskiness of gold declines, either because gold prices become less volatile or because returns on other assets become more volatile, gold becomes more desirable, the quantity demanded at every given price rises, and the demand curve again shifts to the right. When the gold market becomes relatively more liquid and gold therefore becomes more desirable, the quantity demanded at any given price rises, and the demand curve also shifts to the right, as in Figure 1.

Shifts in the Supply Curve for Gold

The supply curve for gold shifts when there are changes in technology that make gold mining more efficient or when governments at any given price of gold decide to increase sales of their holdings of gold. In these cases, the quantity of gold supplied at any given price increases, and the supply curve shifts to the right.

Study Guide

To give yourself practice with supply and demand analysis in the gold market, see if you can analyze what happens to the price of gold for the following situations, remembering that all other things are held constant: 1) Interest rates rise, 2) the gold market becomes more liquid, 3) the volatility of gold prices increases, 4) the stock market is expected to turn bullish in the near future, 5) investors suddenly become fearful that there will be a collapse in real estate prices, and 6) Russia sells a lot of gold in the open market to raise hard currency to finance expenditures.

Application

Changes in the Equilibrium Price of Gold Due to a Rise in Expected Inflation

To illustrate how changes in the equilibrium price of gold occur when supply and demand curves shift, let's look at what happens when there is a change in expected inflation.

Suppose that expected inflation is 5% and the initial supply and demand curves are at G_1^s and G_1^d so that the equilibrium price of gold is at P_1 in Figure 1. If expected inflation now rises to 10%, prices of goods and commodities next year will be expected to be higher than they otherwise would have been, and the price of gold next year P_{t+1}^e will also be expected to be higher than otherwise. Now at any given price of gold today, gold is expected to have a greater rate of appreciation over the coming year and hence a higher expected capital gain and return. The greater expected return means that the quantity of gold demanded increases at any given price, thus shifting the demand curve from G_1^d to G_2^d. Equilibrium therefore moves from point 1 to point 2, and the price of gold rises from P_1 to P_2.

By using a supply and demand diagram like that in Figure 1, you should be able to see that if the expected rate of inflation falls, the price of gold today will also fall. We thus reach the following conclusion: ***The price of gold should be positively related to the expected inflation rate.***

Because the gold market responds immediately to any changes in expected inflation, it is considered a good barometer of the trend of inflation in the future. Indeed, Alan Greenspan, the chairman of the Board of Governors of the Federal Reserve System, at one point advocated using the price of gold as an indicator of inflationary pressures in the economy. Not surprisingly, then, the gold market is followed closely by financial analysts and monetary policymakers.

Chapter 6

The Risk and Term Structure of Interest Rates

PREVIEW

In our supply and demand analysis of interest-rate behavior in Chapter 5, we examined the determination of just one interest rate. Yet we saw earlier that there are enormous numbers of bonds on which the interest rates can and do differ. In this chapter, we complete the interest-rate picture by examining the relationship of the various interest rates to one another. Understanding why they differ from bond to bond can help businesses, banks, insurance companies, and private investors decide which bonds to purchase as investments and which ones to sell.

We first look at why bonds with the same term to maturity have different interest rates. The relationship among these interest rates is called the **risk structure of interest rates**, although risk, liquidity, and income tax rules all play a role in determining the risk structure. A bond's term to maturity also affects its interest rate, and the relationship among interest rates on bonds with different terms to maturity is called the **term structure of interest rates**. In this chapter, we examine the sources and causes of fluctuations in interest rates relative to one another and look at a number of theories that explain these fluctuations.

Risk Structure of Interest Rates

Figure 1 shows the yields to maturity for several categories of long-term bonds from 1919 to 2002. It shows us two important features of interest-rate behavior for bonds of the same maturity: Interest rates on different categories of bonds differ from one another in any given year, and the spread (or difference) between the interest rates varies over time. The interest rates on municipal bonds, for example, are above those on U.S. government (Treasury) bonds in the late 1930s but lower thereafter. In addition, the spread between the interest rates on Baa corporate bonds (riskier than Aaa corporate bonds) and U.S. government bonds is very large during the Great Depression years 1930–1933, is smaller during the 1940s–1960s, and then widens again afterwards. What factors are responsible for these phenomena?

Default Risk

One attribute of a bond that influences its interest rate is its risk of **default**, which occurs when the issuer of the bond is unable or unwilling to make interest payments when promised or pay off the face value when the bond matures. A corporation suffering big losses, such as Chrysler Corporation did in the 1970s, might be more likely

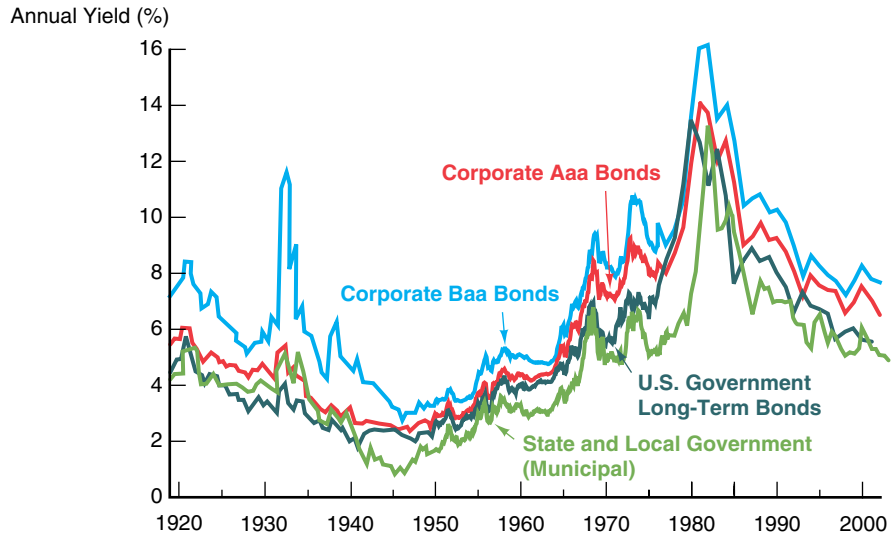


FIGURE 1 Long-Term Bond Yields, 1919–2002

Sources: Board of Governors of the Federal Reserve System, *Banking and Monetary Statistics, 1941–1970;* Federal Reserve: www.federalreserve.gov/releases/h15/data/.

www.federalreserve.gov/Releases/h15/update/

The Federal Reserve reports the returns on different quality bonds. Look at the bottom of the listing of interest rates for AAA and BBB rated bonds.

to suspend interest payments on its bonds.[1] The default risk on its bonds would therefore be quite high. By contrast, U.S. Treasury bonds have usually been considered to have no default risk because the federal government can always increase taxes to pay off its obligations. Bonds like these with no default risk are called **default-free bonds**. (However, during the budget negotiations in Congress in 1995 and 1996, the Republicans threatened to let Treasury bonds default, and this had an impact on the bond market, as one application following this section indicates.) The spread between the interest rates on bonds with default risk and default-free bonds, called the **risk premium**, indicates how much additional interest people must earn in order to be willing to hold that risky bond. Our supply and demand analysis of the bond market in Chapter 5 can be used to explain why a bond with default risk always has a positive risk premium and why the higher the default risk is, the larger the risk premium will be.

To examine the effect of default risk on interest rates, let us look at the supply and demand diagrams for the default-free (U.S. Treasury) and corporate long-term bond markets in Figure 2. To make the diagrams somewhat easier to read, let's assume that initially corporate bonds have the same default risk as U.S. Treasury bonds. In this case, these two bonds have the same attributes (identical risk and maturity); their equilibrium prices and interest rates will initially be equal ($P^c_1 = P^T_1$ and $i^c_1 = i^T_1$), and the risk premium on corporate bonds ($i^c_1 - i^T_1$) will be zero.

[1] Chrysler did not default on its loans in this period, but it would have were it not for a government bailout plan intended to preserve jobs, which in effect provided Chrysler with funds that were used to pay off creditors.

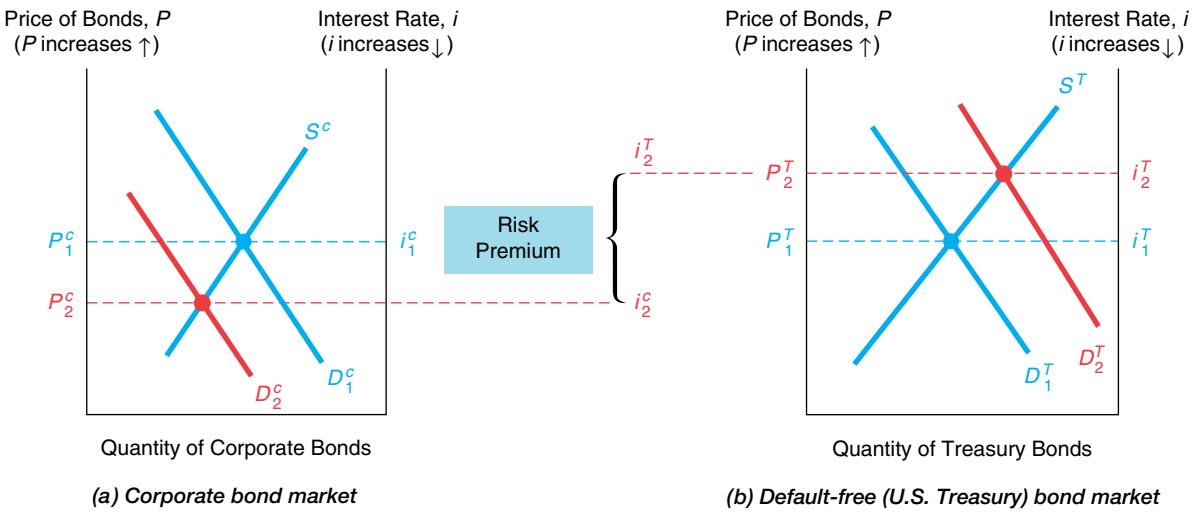


FIGURE 2 Response to an Increase in Default Risk on Corporate Bonds

An increase in default risk on corporate bonds shifts the demand curve from D_1^c to D_2^c. Simultaneously, it shifts the demand curve for Treasury bonds from D_1^T to D_2^T. The equilibrium price for corporate bonds (left axis) falls from P_1^c to P_2^c, and the equilibrium interest rate on corporate bonds (right axis) rises from i_1^c to i_2^c. In the Treasury market, the equilibrium bond price rises from P_1^T to P_2^T, and the equilibrium interest rate falls from i_1^T to i_2^T. The brace indicates the difference between i_2^c and i_2^T, the risk premium on corporate bonds. (*Note:* P and i increase in opposite directions. P on the left vertical axis increases as we go up the axis, while i on the right vertical axis increases as we go down the axis.)

Study Guide

Two exercises will help you gain a better understanding of the risk structure:

1. Put yourself in the shoes of an investor—see how your purchase decision would be affected by changes in risk and liquidity.
2. Practice drawing the appropriate shifts in the supply and demand curves when risk and liquidity change. For example, see if you can draw the appropriate shifts in the supply and demand curves when, in contrast to the examples in the text, a corporate bond has a decline in default risk or an improvement in its liquidity.

If the possibility of a default increases because a corporation begins to suffer large losses, the default risk on corporate bonds will increase, and the expected return on these bonds will decrease. In addition, the corporate bond's return will be more uncertain as well. The theory of asset demand predicts that because the expected return on the corporate bond falls relative to the expected return on the default-free Treasury bond while its relative riskiness rises, the corporate bond is less desirable (holding everything else equal), and demand for it will fall. The demand curve for corporate bonds in panel (a) of Figure 2 then shifts to the left, from D_1^c to D_2^c.

At the same time, the expected return on default-free Treasury bonds increases relative to the expected return on corporate bonds, while their relative riskiness

declines. The Treasury bonds thus become more desirable, and demand rises, as shown in panel (b) by the rightward shift in the demand curve for these bonds from D^T_1 to D^T_2.

As we can see in Figure 2, the equilibrium price for corporate bonds (left axis) falls from P^c_1 to P^c_2, and since the bond price is negatively related to the interest rate, the equilibrium interest rate on corporate bonds (right axis) rises from i^c_1 to i^c_2. At the same time, however, the equilibrium price for the Treasury bonds rises from P^T_1 to P^T_2, and the equilibrium interest rate falls from i^T_1 to i^T_2. The spread between the interest rates on corporate and default-free bonds—that is, the risk premium on corporate bonds—has risen from zero to $i^c_2 - i^T_2$. We can now conclude that ***a bond with default risk will always have a positive risk premium, and an increase in its default risk will raise the risk premium***.

Because default risk is so important to the size of the risk premium, purchasers of bonds need to know whether a corporation is likely to default on its bonds. Two major investment advisory firms, Moody's Investors Service and Standard and Poor's Corporation, provide default risk information by rating the quality of corporate and municipal bonds in terms of the probability of default. The ratings and their description are contained in Table 1. Bonds with relatively low risk of default are called *investment-grade* securities and have a rating of Baa (or BBB) and above. Bonds with

Table 1 Bond Ratings by Moody's and Standard and Poor's

Rating			
Moody's	Standard and Poor's	Descriptions	Examples of Corporations with Bonds Outstanding in 2003
Aaa	AAA	Highest quality (lowest default risk)	General Electric, Pfizer Inc., North Carolina State, Mobil Oil
Aa	AA	High quality	Wal-Mart, McDonald's, Credit Suisse First Boston
A	A	Upper medium grade	Hewlett-Packard, Anheuser-Busch, Ford, Household Finance
Baa	BBB	Medium grade	Motorola, Albertson's, Pennzoil, Weyerhaeuser Co., Tommy Hilfiger
Ba	BB	Lower medium grade	Royal Caribbean, Levi Strauss
B	B	Speculative	Rite Aid, Northwest Airlines Inc., Six Flags
Caa	CCC, CC	Poor (high default risk)	Revlon, United Airlines
Ca	C	Highly speculative	US Airways, Polaroid
C	D	Lowest grade	Enron, Oakwood Homes

ratings below Baa (or BBB) have higher default risk and have been aptly dubbed speculative-grade or **junk bonds**. Because these bonds always have higher interest rates than investment-grade securities, they are also referred to as high-yield bonds.

Next let's look back at Figure 1 and see if we can explain the relationship between interest rates on corporate and U.S. Treasury bonds. Corporate bonds always have higher interest rates than U.S. Treasury bonds because they always have some risk of default, whereas U.S. Treasury bonds do not. Because Baa-rated corporate bonds have a greater default risk than the higher-rated Aaa bonds, their risk premium is greater, and the Baa rate therefore always exceeds the Aaa rate. We can use the same analysis to explain the huge jump in the risk premium on Baa corporate bond rates during the Great Depression years 1930–1933 and the rise in the risk premium after 1970 (see Figure 1). The depression period saw a very high rate of business failures and defaults. As we would expect, these factors led to a substantial increase in default risk for bonds issued by vulnerable corporations, and the risk premium for Baa bonds reached unprecedentedly high levels. Since 1970, we have again seen higher levels of business failures and defaults, although they were still well below Great Depression levels. Again, as expected, default risks and risk premiums for corporate bonds rose, widening the spread between interest rates on corporate bonds and Treasury bonds.

Application

The Enron Bankruptcy and the Baa-Aaa Spread

In December 2001, the Enron Corporation, a firm specializing in trading in the energy market, and once the seventh-largest corporation in the United States, was forced to declare bankruptcy after it became clear that it had used shady accounting to hide its financial problems. (The Enron bankruptcy, the largest ever in the United States, will be discussed further in Chapter 8.) Because of the scale of the bankruptcy and the questions it raised about the quality of the information in accounting statements, the Enron collapse had a major impact on the corporate bond market. Let's see how our supply and demand analysis explains the behavior of the spread between interest rates on lower quality (Baa-rated) and highest quality (Aaa-rated) corporate bonds in the aftermath of the Enron failure.

As a consequence of the Enron bankruptcy, many investors began to doubt the financial health of corporations with lower credit ratings such as Baa. The increase in default risk for Baa bonds made them less desirable at any given interest rate, decreased the quantity demanded, and shifted the demand curve for Baa bonds to the left. As shown in panel (a) of Figure 2, the interest rate on Baa bonds should have risen, which is indeed what happened. Interest rates on Baa bonds rose by 24 basis points (0.24 percentage points) from 7.81% in November 2001 to 8.05% in December 2001. But the increase in the perceived default risk for Baa bonds after the Enron bankruptcy made the highest quality (Aaa) bonds relatively more attractive and shifted the demand curve for these securities to the right—an outcome described by some analysts as a "flight to quality." Just as our analysis predicts in Figure 2, interest rates on Aaa bonds fell by 20 basis points, from 6.97% in November to 6.77% in December. The overall outcome was that the spread between interest rates on Baa and Aaa bonds rose by 44 basis points from 0.84% before the bankruptcy to 1.28% afterward.

Liquidity

Another attribute of a bond that influences its interest rate is its liquidity. As we learned in Chapter 4, a liquid asset is one that can be quickly and cheaply converted into cash if the need arises. The more liquid an asset is, the more desirable it is (holding everything else constant). U.S. Treasury bonds are the most liquid of all long-term bonds, because they are so widely traded that they are the easiest to sell quickly and the cost of selling them is low. Corporate bonds are not as liquid, because fewer bonds for any one corporation are traded; thus it can be costly to sell these bonds in an emergency, because it might be hard to find buyers quickly.

How does the reduced liquidity of the corporate bonds affect their interest rates relative to the interest rate on Treasury bonds? We can use supply and demand analysis with the same figure that was used to analyze the effect of default risk, Figure 2, to show that the lower liquidity of corporate bonds relative to Treasury bonds increases the spread between the interest rates on these two bonds. Let us start the analysis by assuming that initially corporate and Treasury bonds are equally liquid and all their other attributes are the same. As shown in Figure 2, their equilibrium prices and interest rates will initially be equal: $P_1^C = P_1^T$ and $i_1^C = i_1^T$. If the corporate bond becomes less liquid than the Treasury bond because it is less widely traded, then (as the theory of asset demand indicates) its demand will fall, shifting its demand curve from D_1^C to D_2^C as in panel (a). The Treasury bond now becomes relatively more liquid in comparison with the corporate bond, so its demand curve shifts rightward from D_1^T to D_2^T as in panel (b). The shifts in the curves in Figure 2 show that the price of the less liquid corporate bond falls and its interest rate rises, while the price of the more liquid Treasury bond rises and its interest rate falls.

The result is that the spread between the interest rates on the two bond types has risen. Therefore, the differences between interest rates on corporate bonds and Treasury bonds (that is, the risk premiums) reflect not only the corporate bond's default risk but its liquidity, too. This is why a risk premium is more accurately a "risk and liquidity premium," but convention dictates that it is called a *risk premium*.

Income Tax Considerations

Returning to Figure 1, we are still left with one puzzle—the behavior of municipal bond rates. Municipal bonds are certainly not default-free: State and local governments have defaulted on the municipal bonds they have issued in the past, particularly during the Great Depression and even more recently in the case of Orange County, California, in 1994 (more on this in Chapter 13). Also, municipal bonds are not as liquid as U.S. Treasury bonds.

Why is it, then, that these bonds have had lower interest rates than U.S. Treasury bonds for at least 40 years, as indicated in Figure 1? The explanation lies in the fact that interest payments on municipal bonds are exempt from federal income taxes, a factor that has the same effect on the demand for municipal bonds as an increase in their expected return.

Let us imagine that you have a high enough income to put you in the 35% income tax bracket, where for every extra dollar of income you have to pay 35 cents to the government. If you own a $1,000-face-value U.S. Treasury bond that sells for $1,000 and has a coupon payment of $100, you get to keep only $65 of the payment after taxes. Although the bond has a 10% interest rate, you actually earn only 6.5% after taxes.

Suppose, however, that you put your savings into a $1,000-face-value municipal bond that sells for $1,000 and pays only $80 in coupon payments. Its interest rate is only 8%, but because it is a tax-exempt security, you pay no taxes on the $80 coupon payment, so you earn 8% after taxes. Clearly, you earn more on the municipal bond

after taxes, so you are willing to hold the riskier and less liquid municipal bond even though it has a lower interest rate than the U.S. Treasury bond. (This was not true before World War II, when the tax-exempt status of municipal bonds did not convey much of an advantage because income tax rates were extremely low.)

Another way of understanding why municipal bonds have lower interest rates than Treasury bonds is to use the supply and demand analysis displayed in Figure 3. We assume that municipal and Treasury bonds have identical attributes and so have the same bond prices and interest rates as drawn in the figure: $P^m_1 = P^T_1$ and $i^m_1 = i^T_1$. Once the municipal bonds are given a tax advantage that raises their after-tax expected return relative to Treasury bonds and makes them more desirable, demand for them rises, and their demand curve shifts to the right, from D^m_1 to D^m_2. The result is that their equilibrium bond price rises from P^m_1 to P^m_2, and their equilibrium interest rate falls from i^m_1 to i^m_2. By contrast, Treasury bonds have now become less desirable relative to municipal bonds; demand for Treasury bonds decreases, and D^T_1 shifts to D^T_2. The Treasury bond price falls from P^T_1 to P^T_2, and the interest rate rises from i^T_1 to i^T_2. The resulting lower interest rates for municipal bonds and higher interest rates for Treasury bonds explains why municipal bonds can have interest rates below those of Treasury bonds.[2]

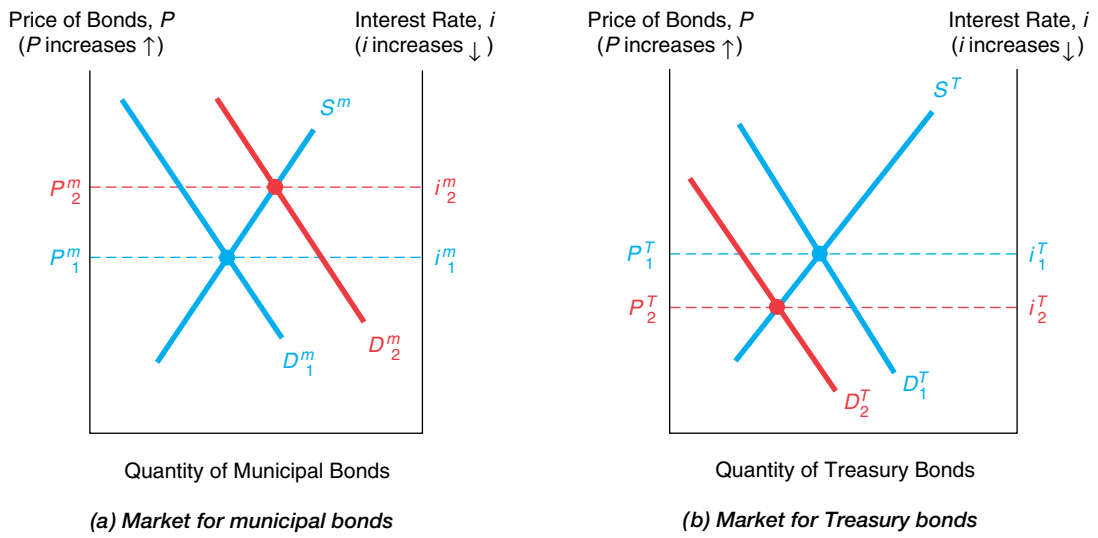


FIGURE 3 Interest Rates on Municipal and Treasury Bonds

When the municipal bond is given tax-free status, demand for the municipal bond shifts rightward from D^m_1 to D^m_2 and demand for the Treasury bond shifts leftward from D^T_1 to D^T_2. The equilibrium price of the municipal bond (left axis) rises from P^m_1 to P^m_2, so its interest rate (right axis) falls from i^m_1 to i^m_2, while the equilibrium price of the Treasury bond falls from P^T_1 to P^T_2 and its interest rate rises from i^T_1 to i^T_2. The result is that municipal bonds end up with lower interest rates than those on Treasury bonds. (*Note:* P and i increase in opposite directions. P on the left vertical axis increases as we go up the axis, while i on the right vertical axis increases as we go down the axis.)

[2] In contrast to corporate bonds, Treasury bonds are exempt from state and local income taxes. Using the analysis in the text, you should be able to show that this feature of Treasury bonds provides an additional reason why interest rates on corporate bonds are higher than those on Treasury bonds.

Summary

The risk structure of interest rates (the relationship among interest rates on bonds with the same maturity) is explained by three factors: default risk, liquidity, and the income tax treatment of the bond's interest payments. As a bond's default risk increases, the risk premium on that bond (the spread between its interest rate and the interest rate on a default-free Treasury bond) rises. The greater liquidity of Treasury bonds also explains why their interest rates are lower than interest rates on less liquid bonds. If a bond has a favorable tax treatment, as do municipal bonds, whose interest payments are exempt from federal income taxes, its interest rate will be lower.

Application

Effects of the Bush Tax Cut on Bond Interest Rates

The Bush tax cut passed in 2001 scheduled a reduction of the top income tax bracket from 39% to 35% over a ten-year period. What is the effect of this income tax decrease on interest rates in the municipal bond market relative to those in the Treasury bond market?

Our supply and demand analysis provides the answer. A decreased income tax rate for rich people means that the after-tax expected return on tax-free municipal bonds relative to that on Treasury bonds is lower, because the interest on Treasury bonds is now taxed at a lower rate. Because municipal bonds now become less desirable, their demand decreases, shifting the demand curve to the left, which lowers their price and raises their interest rate. Conversely, the lower income tax rate makes Treasury bonds more desirable; this change shifts their demand curve to the right, raises their price, and lowers their interest rates.

Our analysis thus shows that the Bush tax cut raises the interest rates on municipal bonds relative to interest rates on Treasury bonds.

Term Structure of Interest Rates

We have seen how risk, liquidity, and tax considerations (collectively embedded in the risk structure) can influence interest rates. Another factor that influences the interest rate on a bond is its term to maturity: Bonds with identical risk, liquidity, and tax characteristics may have different interest rates because the time remaining to maturity is different. A plot of the yields on bonds with differing terms to maturity but the same risk, liquidity, and tax considerations is called a **yield curve**, and it describes the term structure of interest rates for particular types of bonds, such as government bonds. The "Following the Financial News" box shows several yield curves for Treasury securities that were published in the *Wall Street Journal*. Yield curves can be classified as upward-sloping, flat, and downward-sloping (the last sort is often referred to as an **inverted yield curve**). When yield curves slope upward, as in the "Following the Financial News" box, the long-term interest rates are above the short-term interest rates; when yield curves are flat, short- and long-term interest rates are the same; and when yield curves are inverted, long-term interest rates are below short-term interest rates. Yield curves can also have more complicated shapes in which they first slope up and then down, or vice versa. Why do we usually see

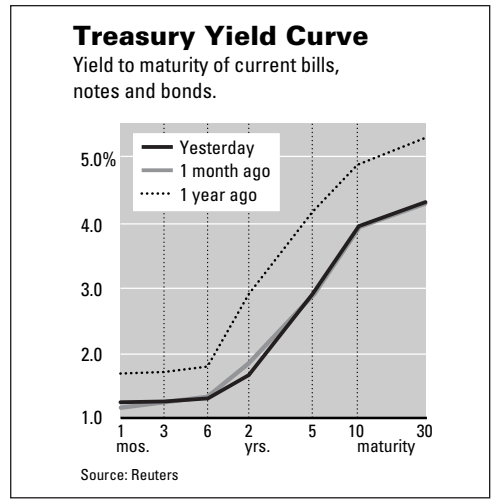
Following the Financial News

Yield Curves

The *Wall Street Journal* publishes a daily plot of the yield curves for Treasury securities, an example of which is presented here. It is typically found on page 2 of the "Money and Investing" section.

The numbers on the vertical axis indicate the interest rate for the Treasury security, with the maturity given by the numbers on the horizontal axis. For example, the yield curve marked "Yesterday" indicates that the interest rate on the three-month Treasury bill yesterday was 1.25%, while the one-year bill had an interest rate of 1.35% and the ten-year bond had an interest rate of 4.0%. As you can see, the yield curves in the plot have the typical upward slope.

Source: Wall Street Journal, Wednesday, January 22, 2003, p. C2.



www.ratecurve.com/yc2.html
Check out today's yield curve.

upward slopes of the yield curve as in the "Following the Financial News" box but sometimes other shapes?

Besides explaining why yield curves take on different shapes at different times, a good theory of the term structure of interest rates must explain the following three important empirical facts:

1. As we see in Figure 4, interest rates on bonds of different maturities move together over time.
2. When short-term interest rates are low, yield curves are more likely to have an upward slope; when short-term interest rates are high, yield curves are more likely to slope downward and be inverted.
3. Yield curves almost always slope upward, as in the "Following the Financial News" box.

Three theories have been put forward to explain the term structure of interest rates; that is, the relationship among interest rates on bonds of different maturities reflected in yield curve patterns: (1) the expectations theory, (2) the segmented markets theory, and (3) the liquidity premium theory, each of which is described in the following sections. The expectations theory does a good job of explaining the first two facts on our list, but not the third. The segmented markets theory can explain fact 3 but not the other two facts, which are well explained by the expectations theory. Because each theory explains facts that the other cannot, a natural way to seek a better understanding of the term structure is to combine features of both theories, which leads us to the liquidity premium theory, which can explain all three facts.

If the liquidity premium theory does a better job of explaining the facts and is hence the most widely accepted theory, why do we spend time discussing the other two theories? There are two reasons. First, the ideas in these two theories provide the

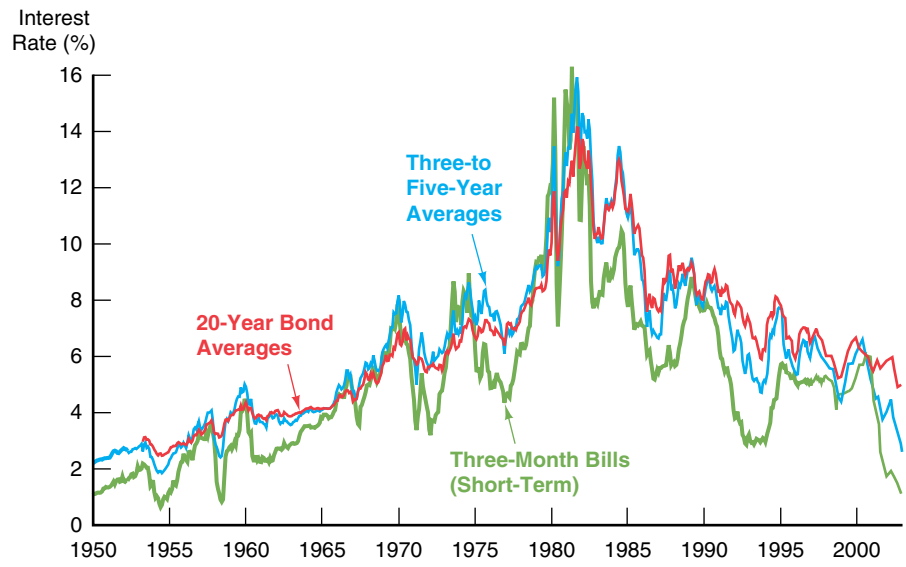


FIGURE 4 Movements over Time of Interest Rates on U.S. Government Bonds with Different Maturities

Sources: Board of Governors of the Federal Reserve System, *Banking and Monetary Statistics, 1941–1970;* Federal Reserve: www.federalreserve.gov/releases/h15/data.htm#top.

groundwork for the liquidity premium theory. Second, it is important to see how economists modify theories to improve them when they find that the predicted results are inconsistent with the empirical evidence.

Expectations Theory

The **expectations theory** of the term structure states the following commonsense proposition: The interest rate on a long-term bond will equal an average of short-term interest rates that people expect to occur over the life of the long-term bond. For example, if people expect that short-term interest rates will be 10% on average over the coming five years, the expectations theory predicts that the interest rate on bonds with five years to maturity will be 10% too. If short-term interest rates were expected to rise even higher after this five-year period so that the average short-term interest rate over the coming 20 years is 11%, then the interest rate on 20-year bonds would equal 11% and would be higher than the interest rate on five-year bonds. We can see that the explanation provided by the expectations theory for why interest rates on bonds of different maturities differ is that short-term interest rates are expected to have different values at future dates.

The key assumption behind this theory is that buyers of bonds do not prefer bonds of one maturity over another, so they will not hold any quantity of a bond if its expected return is less than that of another bond with a different maturity. Bonds that have this characteristic are said to be *perfect substitutes*. What this means in practice is that if bonds with different maturities are perfect substitutes, the expected return on these bonds must be equal.

To see how the assumption that bonds with different maturities are perfect substitutes leads to the expectations theory, let us consider the following two investment strategies:

1. Purchase a one-year bond, and when it matures in one year, purchase another one-year bond.
2. Purchase a two-year bond and hold it until maturity.

Because both strategies must have the same expected return if people are holding both one- and two-year bonds, the interest rate on the two-year bond must equal the average of the two one-year interest rates. For example, let's say that the current interest rate on the one-year bond is 9% and you expect the interest rate on the one-year bond next year to be 11%. If you pursue the first strategy of buying the two one-year bonds, the expected return over the two years will average out to be (9% + 11%)/2 = 10% per year. You will be willing to hold both the one- and two-year bonds only if the expected return per year of the two-year bond equals this. Therefore, the interest rate on the two-year bond must equal 10%, the average interest rate on the two one-year bonds.

We can make this argument more general. For an investment of $1, consider the choice of holding, for two periods, a two-period bond or two one-period bonds. Using the definitions

i_t = today's (time t) interest rate on a one-period bond
i^e_{t+1} = interest rate on a one-period bond expected for next period (time $t + 1$)
i_{2t} = today's (time t) interest rate on the two-period bond

the expected return over the two periods from investing $1 in the two-period bond and holding it for the two periods can be calculated as:

$$(1 + i_{2t})(1 + i_{2t}) - 1 = 1 + 2i_{2t} + (i_{2t})^2 - 1 = 2i_{2t} + (i_{2t})^2$$

After the second period, the $1 investment is worth $(1 + i_{2t})(1 + i_{2t})$. Subtracting the $1 initial investment from this amount and dividing by the initial $1 investment gives the rate of return calculated in the previous equation. Because $(i_{2t})^2$ is extremely small—if $i_{2t} = 10\% = 0.10$, then $(i_{2t})^2 = 0.01$—we can simplify the expected return for holding the two-period bond for the two periods to

$$2i_{2t}$$

With the other strategy, in which one-period bonds are bought, the expected return on the $1 investment over the two periods is:

$$(1 + i_t)(1 + i^e_{t+1}) - 1 = 1 + i_t + i^e_{t+1} + i_t(i^e_{t+1}) - 1 = i_t + i^e_t + i_t(i^e_{t+1})$$

This calculation is derived by recognizing that after the first period, the $1 investment becomes $1 + i_t$, and this is reinvested in the one-period bond for the next period, yielding an amount $(1 + i_t)(1 + i^e_{t+1})$. Then subtracting the $1 initial investment from this amount and dividing by the initial investment of $1 gives the expected return for the strategy of holding one-period bonds for the two periods. Because $i_t(i^e_{t+1})$ is also extremely small—if $i_t = i^e_{t+1} = 0.10$, then $i_t(i^e_{t+1}) = 0.01$—we can simplify this to:

$$i_t + i^e_{t+1}$$

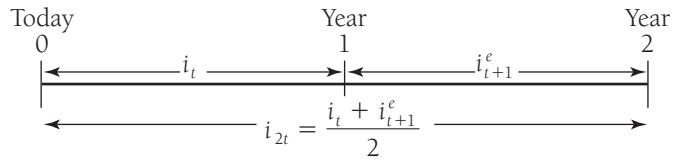
Both bonds will be held only if these expected returns are equal; that is, when:

$$2i_{2t} = i_t + i^e_{t+1}$$

Solving for i_{2t} in terms of the one-period rates, we have:

$$i_{2t} = \frac{i_t + i^e_{t+1}}{2} \tag{1}$$

which tells us that the two-period rate must equal the average of the two one-period rates. Graphically, this can be shown as:



We can conduct the same steps for bonds with a longer maturity so that we can examine the whole term structure of interest rates. Doing so, we will find that the interest rate of i_{nt} on an n-period bond must equal:

$$i_{nt} = \frac{i_t + i^e_{t+1} + i^e_{t+2} + \cdots + i^e_{t+(n-1)}}{n} \tag{2}$$

Equation 2 states that the n-period interest rate equals the average of the one-period interest rates expected to occur over the n-period life of the bond. This is a restatement of the expectations theory in more precise terms.[3]

A simple numerical example might clarify what the expectations theory in Equation 2 is saying. If the one-year interest rate over the next five years is expected to be 5, 6, 7, 8, and 9%, Equation 2 indicates that the interest rate on the two-year bond would be:

$$\frac{5\% + 6\%}{2} = 5.5\%$$

while for the five-year bond it would be:

$$\frac{5\% + 6\% + 7\% + 8\% + 9\%}{5} = 7\%$$

Doing a similar calculation for the one-, three-, and four-year interest rates, you should be able to verify that the one- to five-year interest rates are 5.0, 5.5, 6.0, 6.5, and 7.0%, respectively. Thus we see that the rising trend in expected short-term interest rates produces an upward-sloping yield curve along which interest rates rise as maturity lengthens.

The expectations theory is an elegant theory that provides an explanation of why the term structure of interest rates (as represented by yield curves) changes at different times. When the yield curve is upward-sloping, the expectations theory suggests that short-term interest rates are expected to rise in the future, as we have seen in our numerical example. In this situation, in which the long-term rate is currently above the short-term rate, the average of future short-term rates is expected to be higher than the current short-term rate, which can occur only if short-term interest rates are expected to rise. This is what we see in our numerical example. When the yield curve is inverted (slopes downward), the average of future short-term interest rates is

[3]The analysis here has been conducted for discount bonds. Formulas for interest rates on coupon bonds would differ slightly from those used here, but would convey the same principle.

expected to be below the current short-term rate, implying that short-term interest rates are expected to fall, on average, in the future. Only when the yield curve is flat does the expectations theory suggest that short-term interest rates are not expected to change, on average, in the future.

The expectations theory also explains fact 1 that interest rates on bonds with different maturities move together over time. Historically, short-term interest rates have had the characteristic that if they increase today, they will tend to be higher in the future. Hence a rise in short-term rates will raise people's expectations of future short-term rates. Because long-term rates are the average of expected future short-term rates, a rise in short-term rates will also raise long-term rates, causing short- and long-term rates to move together.

The expectations theory also explains fact 2 that yield curves tend to have an upward slope when short-term interest rates are low and are inverted when short-term rates are high. When short-term rates are low, people generally expect them to rise to some normal level in the future, and the average of future expected short-term rates is high relative to the current short-term rate. Therefore, long-term interest rates will be substantially above current short-term rates, and the yield curve would then have an upward slope. Conversely, if short-term rates are high, people usually expect them to come back down. Long-term rates would then drop below short-term rates because the average of expected future short-term rates would be below current short-term rates and the yield curve would slope downward and become inverted.[4]

The expectations theory is an attractive theory because it provides a simple explanation of the behavior of the term structure, but unfortunately it has a major shortcoming: It cannot explain fact 3, which says that yield curves usually slope upward. The typical upward slope of yield curves implies that short-term interest rates are usually expected to rise in the future. In practice, short-term interest rates are just as likely to fall as they are to rise, and so the expectations theory suggests that the typical yield curve should be flat rather than upward-sloping.

Segmented Markets Theory

As the name suggests, the **segmented markets theory** of the term structure sees markets for different-maturity bonds as completely separate and segmented. The interest rate for each bond with a different maturity is then determined by the supply of and demand for that bond with no effects from expected returns on other bonds with other maturities.

The key assumption in the segmented markets theory is that bonds of different maturities are not substitutes at all, so the expected return from holding a bond of one maturity has no effect on the demand for a bond of another maturity. This theory of the term structure is at the opposite extreme to the expectations theory, which assumes that bonds of different maturities are perfect substitutes.

The argument for why bonds of different maturities are not substitutes is that investors have strong preferences for bonds of one maturity but not for another, so they will be concerned with the expected returns only for bonds of the maturity they prefer. This might occur because they have a particular holding period in mind, and

[4]The expectations theory explains another important fact about the relationship between short-term and long-term interest rates. As you can see in Figure 4, short-term interest rates are more volatile than long-term rates. If interest rates are *mean-reverting*—that is, if they tend to head back down after they are at unusually high levels or go back up when they are at unusually low levels—then an average of these short-term rates must necessarily have lower volatility than the short-term rates themselves. Because the expectations theory suggests that the long-term rate will be an average of future short-term rates, it implies that the long-term rate will have lower volatility than short-term rates.

if they match the maturity of the bond to the desired holding period, they can obtain a certain return with no risk at all.[5] (We have seen in Chapter 4 that if the term to maturity equals the holding period, the return is known for certain because it equals the yield exactly, and there is no interest-rate risk.) For example, people who have a short holding period would prefer to hold short-term bonds. Conversely, if you were putting funds away for your young child to go to college, your desired holding period might be much longer, and you would want to hold longer-term bonds.

In the segmented markets theory, differing yield curve patterns are accounted for by supply and demand differences associated with bonds of different maturities. If, as seems sensible, investors have short desired holding periods and generally prefer bonds with shorter maturities that have less interest-rate risk, the segmented markets theory can explain fact 3 that yield curves typically slope upward. Because in the typical situation the demand for long-term bonds is relatively lower than that for short-term bonds, long-term bonds will have lower prices and higher interest rates, and hence the yield curve will typically slope upward.

Although the segmented markets theory can explain why yield curves usually tend to slope upward, it has a major flaw in that it cannot explain facts 1 and 2. Because it views the market for bonds of different maturities as completely segmented, there is no reason for a rise in interest rates on a bond of one maturity to affect the interest rate on a bond of another maturity. Therefore, it cannot explain why interest rates on bonds of different maturities tend to move together (fact 1). Second, because it is not clear how demand and supply for short- versus long-term bonds change with the level of short-term interest rates, the theory cannot explain why yield curves tend to slope upward when short-term interest rates are low and to be inverted when short-term interest rates are high (fact 2).

Because each of our two theories explains empirical facts that the other cannot, a logical step is to combine the theories, which leads us to the liquidity premium theory.

Liquidity Premium and Preferred Habitat Theories

The **liquidity premium theory** of the term structure states that the interest rate on a long-term bond will equal an average of short-term interest rates expected to occur over the life of the long-term bond plus a liquidity premium (also referred to as a term premium) that responds to supply and demand conditions for that bond.

The liquidity premium theory's key assumption is that bonds of different maturities are substitutes, which means that the expected return on one bond *does* influence the expected return on a bond of a different maturity, but it allows investors to prefer one bond maturity over another. In other words, bonds of different maturities are assumed to be substitutes but not perfect substitutes. Investors tend to prefer shorter-term bonds because these bonds bear less interest-rate risk. For these reasons, investors must be offered a positive liquidity premium to induce them to hold longer-term bonds. Such an outcome would modify the expectations theory by adding a positive liquidity premium to the equation that describes the relationship between long- and short-term interest rates. The liquidity premium theory is thus written as:

$$i_{nt} = \frac{i_t + i^e_{t+1} + i^e_{t+2} + \cdots + i^e_{t+(n-1)}}{n} + l_{nt} \tag{3}$$

[5]The statement that there is no uncertainty about the return if the term to maturity equals the holding period is literally true only for a discount bond. For a coupon bond with a long holding period, there is some risk because coupon payments must be reinvested before the bond matures. Our analysis here is thus being conducted for discount bonds. However, the gist of the analysis remains the same for coupon bonds because the amount of this risk from reinvestment is small when coupon bonds have the same term to maturity as the holding period.

where l_{nt} = the liquidity (term) premium for the n-period bond at time t, which is always positive and rises with the term to maturity of the bond, n.

Closely related to the liquidity premium theory is the **preferred habitat theory**, which takes a somewhat less direct approach to modifying the expectations hypothesis but comes up with a similar conclusion. It assumes that investors have a preference for bonds of one maturity over another, a particular bond maturity (preferred habitat) in which they prefer to invest. Because they prefer bonds of one maturity over another they will be willing to buy bonds that do not have the preferred maturity only if they earn a somewhat higher expected return. Because investors are likely to prefer the habitat of short-term bonds over that of longer-term bonds, they are willing to hold long-term bonds only if they have higher expected returns. This reasoning leads to the same Equation 3 implied by the liquidity premium theory with a term premium that typically rises with maturity.

http://stockcharts.com/charts/YieldCurve.html
This site lets you look at the dynamic yield curve at any point in time since 1995.

The relationship between the expectations theory and the liquidity premiums and preferred habitat theories is shown in Figure 5. There we see that because the liquidity premium is always positive and typically grows as the term to maturity increases, the yield curve implied by the liquidity premium theory is always above the yield curve implied by the expectations theory and generally has a steeper slope.

A simple numerical example similar to the one we used for the expectations hypothesis further clarifies what the liquidity premium and preferred habitat theories in Equation 3 are saying. Again suppose that the one-year interest rate over the next five years is expected to be 5, 6, 7, 8, and 9%, while investors' preferences for holding short-term bonds means that the liquidity premiums for one- to five-year bonds are 0, 0.25, 0.5, 0.75, and 1.0%, respectively. Equation 3 then indicates that the interest rate on the two-year bond would be:

$$\frac{5\% + 6\%}{2} + 0.25\% = 5.75\%$$

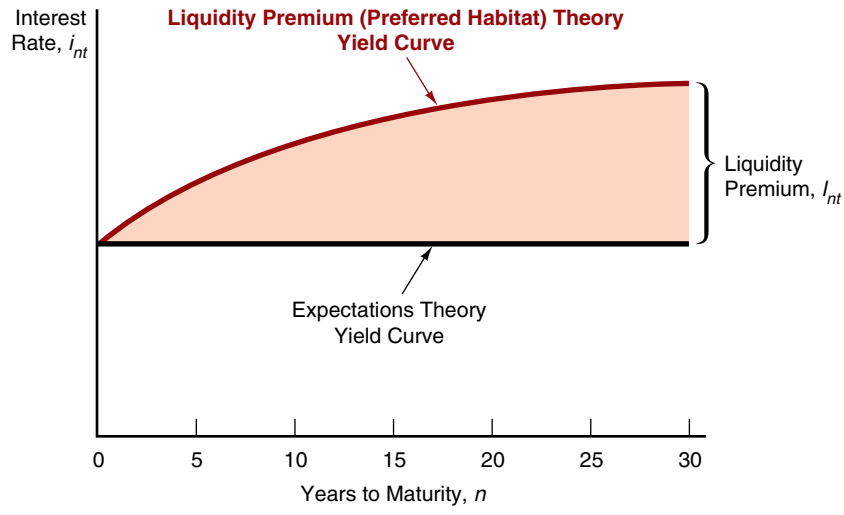


FIGURE 5 The Relationship Between the Liquidity Premium (Preferred Habitat) and Expectations Theory

Because the liquidity premium is always positive and grows as the term to maturity increases, the yield curve implied by the liquidity premium and preferred habitat theories is always above the yield curve implied by the expectations theory and has a steeper slope. Note that the yield curve implied by the expectations theory is drawn under the scenario of unchanging future one-year interest rates.

while for the five-year bond it would be:

$$\frac{5\% + 6\% + 7\% + 8\% + 9\%}{5} + 1\% = 8\%$$

Doing a similar calculation for the one-, three-, and four-year interest rates, you should be able to verify that the one- to five-year interest rates are 5.0, 5.75, 6.5, 7.25, and 8.0%, respectively. Comparing these findings with those for the expectations theory, we see that the liquidity premium and preferred habitat theories produce yield curves that slope more steeply upward because of investors' preferences for short-term bonds.

Let's see if the liquidity premium and preferred habitat theories are consistent with all three empirical facts we have discussed. They explain fact 1 that interest rates on different-maturity bonds move together over time: A rise in short-term interest rates indicates that short-term interest rates will, on average, be higher in the future, and the first term in Equation 3 then implies that long-term interest rates will rise along with them.

They also explain why yield curves tend to have an especially steep upward slope when short-term interest rates are low and to be inverted when short-term rates are high (fact 2). Because investors generally expect short-term interest rates to rise to some normal level when they are low, the average of future expected short-term rates will be high relative to the current short-term rate. With the additional boost of a positive liquidity premium, long-term interest rates will be substantially above current short-term rates, and the yield curve would then have a steep upward slope. Conversely, if short-term rates are high, people usually expect them to come back down. Long-term rates would then drop below short-term rates because the average of expected future short-term rates would be so far below current short-term rates that despite positive liquidity premiums, the yield curve would slope downward.

The liquidity premium and preferred habitat theories explain fact 3 that yield curves typically slope upward by recognizing that the liquidity premium rises with a bond's maturity because of investors' preferences for short-term bonds. Even if short-term interest rates are expected to stay the same on average in the future, long-term interest rates will be above short-term interest rates, and yield curves will typically slope upward.

How can the liquidity premium and preferred habitat theories explain the occasional appearance of inverted yield curves if the liquidity premium is positive? It must be that at times short-term interest rates are expected to fall so much in the future that the average of the expected short-term rates is well below the current short-term rate. Even when the positive liquidity premium is added to this average, the resulting long-term rate will still be below the current short-term interest rate.

As our discussion indicates, a particularly attractive feature of the liquidity premium and preferred habitat theories is that they tell you what the market is predicting about future short-term interest rates just from the slope of the yield curve. A steeply rising yield curve, as in panel (a) of Figure 6, indicates that short-term interest rates are expected to rise in the future. A moderately steep yield curve, as in panel (b), indicates that short-term interest rates are not expected to rise or fall much in the future. A flat yield curve, as in panel (c), indicates that short-term rates are expected to fall moderately in the future. Finally, an inverted yield curve, as in panel (d), indicates that short-term interest rates are expected to fall sharply in the future.

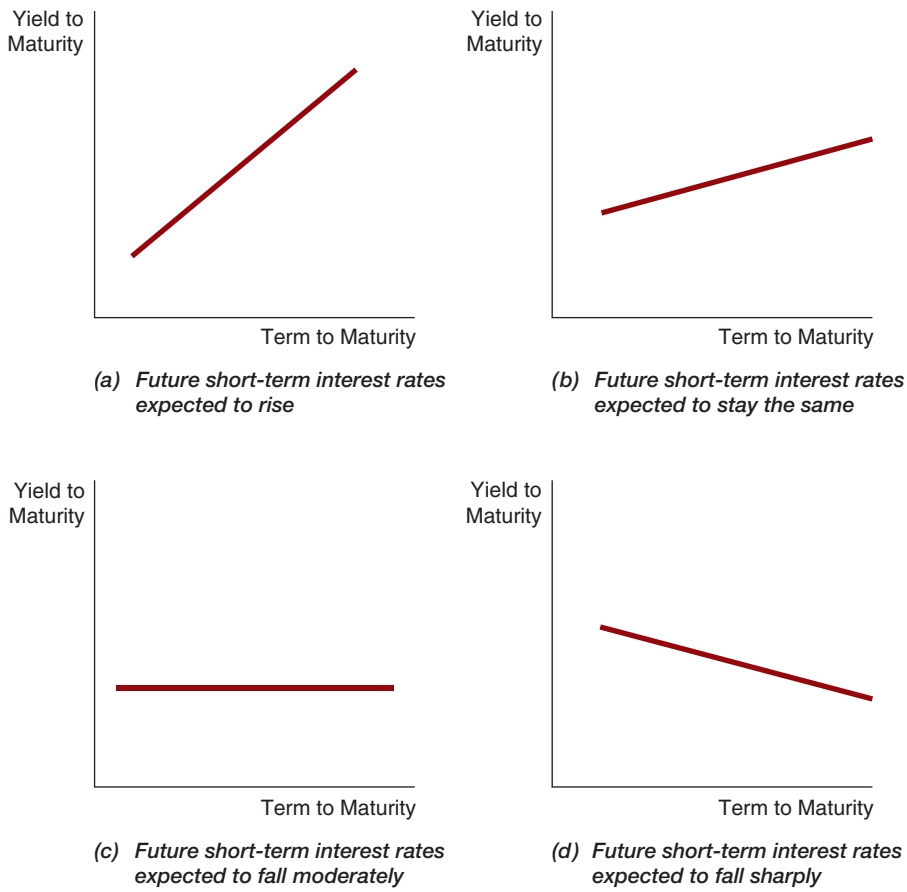


FIGURE 6 Yield Curves and the Market's Expectations of Future Short-Term Interest Rates According to the Liquidity Premium Theory

Evidence on the Term Structure

In the 1980s, researchers examining the term structure of interest rates questioned whether the slope of the yield curve provides information about movements of future short-term interest rates.[6] They found that the spread between long- and short-term interest rates does not always help predict future short-term interest rates, a finding that may stem from substantial fluctuations in the liquidity (term) premium for long-term bonds. More recent research using more discriminating tests now favors a different view. It shows that the term structure contains quite a bit of information for the very short run (over the next several months) and the long run (over several years)

[6]Robert J. Shiller, John Y. Campbell, and Kermit L. Schoenholtz, "Forward Rates and Future Policy: Interpreting the Term Structure of Interest Rates," *Brookings Papers on Economic Activity* 1 (1983): 173–217; N. Gregory Mankiw and Lawrence H. Summers, "Do Long-Term Interest Rates Overreact to Short-Term Interest Rates?" *Brookings Papers on Economic Activity* 1 (1984): 223–242.

but is unreliable at predicting movements in interest rates over the intermediate term (the time in between).[7]

Summary

The liquidity premium and preferred habitat theories are the most widely accepted theories of the term structure of interest rates because they explain the major empirical facts about the term structure so well. They combine the features of both the expectations theory and the segmented markets theory by asserting that a long-term interest rate will be the sum of a liquidity (term) premium and the average of the short-term interest rates that are expected to occur over the life of the bond.

The liquidity premium and preferred habitat theories explain the following facts: (1) Interest rates on bonds of different maturities tend to move together over time, (2) yield curves usually slope upward, and (3) when short-term interest rates are low, yield curves are more likely to have a steep upward slope, whereas when short-term interest rates are high, yield curves are more likely to be inverted.

The theories also help us predict the movement of short-term interest rates in the future. A steep upward slope of the yield curve means that short-term rates are expected to rise, a mild upward slope means that short-term rates are expected to remain the same, a flat slope means that short-term rates are expected to fall moderately, and an inverted yield curve means that short-term rates are expected to fall sharply.

Application

Interpreting Yield Curves, 1980–2003

Figure 7 illustrates several yield curves that have appeared for U.S. government bonds in recent years. What do these yield curves tell us about the public's expectations of future movements of short-term interest rates?

Study Guide

Try to answer the preceding question before reading further in the text. If you have trouble answering it with the liquidity premium and preferred habitat theories, first try answering it with the expectations theory (which is simpler because you don't have to worry about the liquidity premium). When you understand what the expectations of future interest rates are in this case, modify your analysis by taking the liquidity premium into account.

The steep inverted yield curve that occurred on January 15, 1981, indicated that short-term interest rates were expected to decline sharply in the future. In order for longer-term interest rates with their positive liquidity premium to be well below the short-term interest rate, short-term interest rates must be expected to decline so sharply that their average is far below the current short-term rate. Indeed, the public's expectations of sharply lower short-term interest rates evident in the yield curve were realized soon after January 15; by March, three-month Treasury bill rates had declined from the 16% level to 13%.

[7]Eugene Fama, "The Information in the Term Structure," *Journal of Financial Economics* 13 (1984): 509–528; Eugene Fama and Robert Bliss, "The Information in Long-Maturity Forward Rates," *American Economic Review* 77 (1987): 680–692; John Y. Campbell and Robert J. Shiller, "Cointegration and Tests of the Present Value Models," *Journal of Political Economy* 95 (1987): 1062–1088; John Y. Campbell and Robert J. Shiller, "Yield Spreads and Interest Rate Movements: A Bird's Eye View," *Review of Economic Studies* 58 (1991): 495–514.

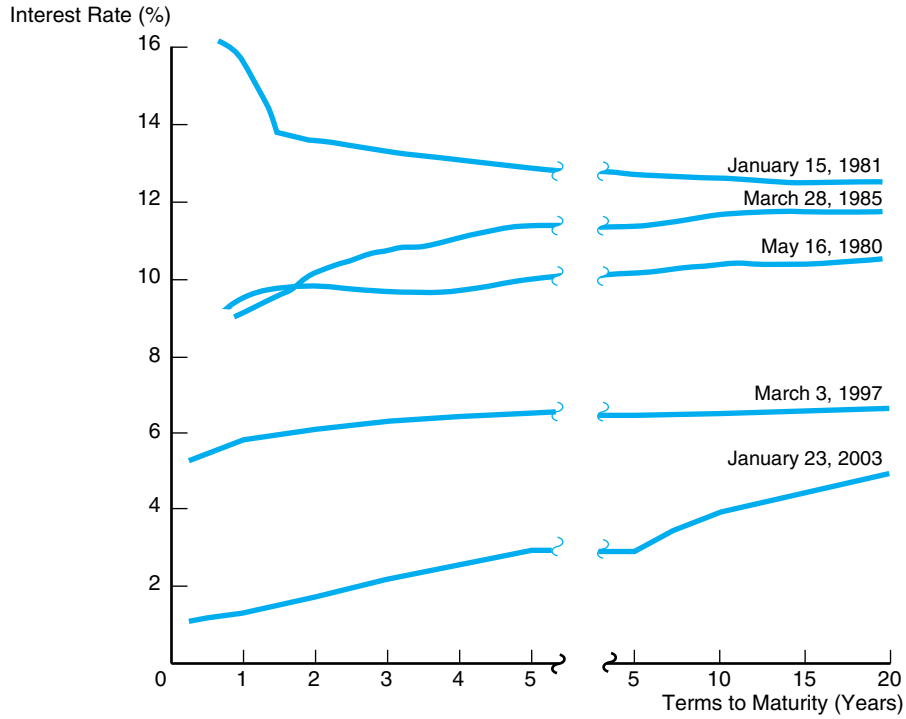


FIGURE 7 Yield Curves for U.S. Government Bonds

Sources: Federal Reserve Bank of St. Louis; *U.S. Financial Data,* various issues; *Wall Street Journal,* various dates.

The steep upward-sloping yield curves on March 28, 1985, and January 23, 2003, indicated that short-term interest rates would climb in the future. The long-term interest rate is above the short-term interest rate when short-term interest rates are expected to rise because their average plus the liquidity premium will be above the current short-term rate. The moderately upward-sloping yield curves on May 16, 1980, and March 3, 1997, indicated that short-term interest rates were expected neither to rise nor to fall in the near future. In this case, their average remains the same as the current short-term rate, and the positive liquidity premium for longer-term bonds explains the moderate upward slope of the yield curve.

Summary

1. Bonds with the same maturity will have different interest rates because of three factors: default risk, liquidity, and tax considerations. The greater a bond's default risk, the higher its interest rate relative to other bonds; the greater a bond's liquidity, the lower its interest rate; and bonds with tax-exempt status will have lower interest rates than they otherwise would. The relationship among interest rates on bonds with the same maturity that arise because of these three factors is known as the *risk structure of interest rates*.

2. Four theories of the term structure provide explanations of how interest rates on bonds with different terms to maturity are related. The expectations theory views long-term interest rates as equaling the average of future short-term interest rates expected to occur over the life of the bond; by contrast, the segmented markets theory treats the determination of interest rates for each bond's maturity as the outcome of supply and demand in that market only. Neither of these theories by itself can explain the fact that interest rates on bonds of different maturities move together over time and that yield curves usually slope upward.
3. The liquidity premium and preferred habitat theories combine the features of the other two theories, and by so doing are able to explain the facts just mentioned. They view long-term interest rates as equaling the average of future short-term interest rates expected to occur over the life of the bond plus a liquidity premium. These theories allow us to infer the market's expectations about the movement of future short-term interest rates from the yield curve. A steeply upward-sloping curve indicates that future short-term rates are expected to rise, a mildly upward-sloping curve indicates that short-term rates are expected to stay the same, a flat curve indicates that short-term rates are expected to decline slightly, and an inverted yield curve indicates that a substantial decline in short-term rates is expected in the future.

Key Terms

default, p. 120
default-free bonds, p. 121
expectations theory, p. 129
inverted yield curve, p. 127
junk bonds, p. 124
liquidity premium theory, p. 133
preferred habitat theory, p. 134
risk premium, p. 121
risk structure of interest rates, p. 120
segmented markets theory, p. 132
term structure of interest rates, p. 120
yield curve, p. 127

Questions and Problems

Questions marked with an asterisk are answered at the end of the book in an appendix, "Answers to Selected Questions and Problems."

1. Which should have the higher risk premium on its interest rates, a corporate bond with a Moody's Baa rating or a corporate bond with a C rating? Why?

*2. Why do U.S. Treasury bills have lower interest rates than large-denomination negotiable bank CDs?

3. Risk premiums on corporate bonds are usually *anticyclical*; that is, they decrease during business cycle expansions and increase during recessions. Why is this so?

*4. "If bonds of different maturities are close substitutes, their interest rates are more likely to move together." Is this statement true, false, or uncertain? Explain your answer.

5. If yield curves, on average, were flat, what would this say about the liquidity (term) premiums in the term structure? Would you be more or less willing to accept the expectations theory?

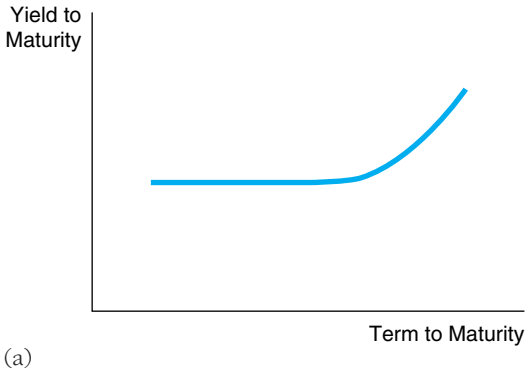
*6. Assuming that the expectations theory is the correct theory of the term structure, calculate the interest rates in the term structure for maturities of one to five years, and plot the resulting yield curves for the following series of one-year interest rates over the next five years:
   (a) 5%, 7%, 7%, 7%, 7%
   (b) 5%, 4%, 4%, 4%, 4%

   How would your yield curves change if people preferred shorter-term bonds over longer-term bonds?

7. Assuming that the expectations theory is the correct theory of the term structure, calculate the interest rates in the term structure for maturities of one to five years, and plot the resulting yield curves for the following path of one-year interest rates over the next five years:
   (a) 5%, 6%, 7%, 6%, 5%
   (b) 5%, 4%, 3%, 4%, 5%

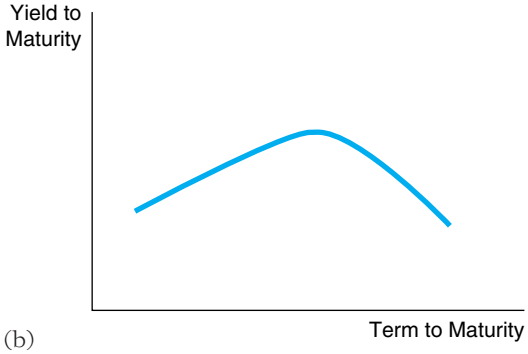
   How would your yield curves change if people preferred shorter-term bonds over longer-term bonds?

*8. If a yield curve looks like the one shown in figure (a) in this section, what is the market predicting about the movement of future short-term interest rates? What might the yield curve indicate about the market's predictions about the inflation rate in the future?



(a)

9. If a yield curve looks like the one shown in (b), what is the market predicting about the movement of future short-term interest rates? What might the yield curve indicate about the market's predictions about the inflation rate in the future?



(b)

*10. What effect would reducing income tax rates have on the interest rates of municipal bonds? Would interest rates of Treasury securities be affected, and if so, how?

Using Economic Analysis to Predict the Future

11. Predict what will happen to interest rates on a corporation's bonds if the federal government guarantees today that it will pay creditors if the corporation goes bankrupt in the future. What will happen to the interest rates on Treasury securities?

*12. Predict what would happen to the risk premiums on corporate bonds if brokerage commissions were lowered in the corporate bond market.

13. If the income tax exemption on municipal bonds were abolished, what would happen to the interest rates on these bonds? What effect would the change have on interest rates on U.S. Treasury securities?

*14. If the yield curve suddenly becomes steeper, how would you revise your predictions of interest rates in the future?

15. If expectations of future short-term interest rates suddenly fall, what would happen to the slope of the yield curve?

Web Exercises

1. The amount of additional interest investors receive due to the various premiums changes over time. Sometimes the risk premiums are much larger than at other times. For example, the default risk premium was very small in the late 1990s when the economy was so healthy business failures were rare. This risk premium increases during recessions.

   Go to www.federalreserve.gov/releases/releases/h15 (historical data) and find the interest rate listings for AAA and Baa rated bonds at three points in time, the most recent, June 1, 1995, and June 1, 1992. Prepare a graph that shows these three time periods (see Figure 1 for an example). Are the risk premiums stable or do they change over time?

2. Figure 7 shows a number of yield curves at various points in time. Go to www.bloomberg.com, and click on "Markets" at the top of the page. Find the Treasury yield curve. Does the current yield curve fall above or below the most recent one listed in Figure 7? Is the current yield curve flatter or steeper than the most recent one reported in Figure 7?

3. Investment companies attempt to explain to investors the nature of the risk the investor incurs when buying shares in their mutual funds. For example, Vanguard carefully explains interest rate risk and offers alternative funds with different interest rate risks. Go to http://flagship5.vanguard.com/VGApp/hnw/FundsStocksOverview.
   a. Select the bond fund you would recommend to an investor who has very low tolerance for risk and a short investment horizon. Justify your answer.
   b. Select the bond fund you would recommend to an investor who has very high tolerance for risk and a long investment horizon. Justify your answer.

Chapter 7

The Stock Market, the Theory of Rational Expectations, and the Efficient Market Hypothesis

PREVIEW

Rarely does a day go by that the stock market isn't a major news item. We have witnessed huge swings in the stock market in recent years. The 1990s were an extraordinary decade for stocks: the Dow Jones and S&P 500 indexes increased more than 400%, while the tech-laden NASDAQ index rose more than 1,000%. By early 2000, both indexes had reached record highs. Unfortunately, the good times did not last, and many investors lost their shirts. Starting in early 2000, the stock market began to decline: the NASDAQ crashed, falling by over 50%, while the Dow Jones and S&P 500 indexes fell by 30% through January 2003.

Because so many people invest in the stock market and the price of stocks affects the ability of people to retire comfortably, the market for stocks is undoubtedly the financial market that receives the most attention and scrutiny. In this chapter, we look at how this important market works.

We begin by discussing the fundamental theories that underlie the valuation of stocks. These theories are critical to understanding the forces that cause the value of stocks to rise and fall minute by minute and day by day. Once we have learned the methods for stock valuation, we need to explore how expectations about the market affect its behavior. We do so by examining the *theory of rational expectations*. When this theory is applied to financial markets, the outcome is the *efficient market hypothesis*. The theory of rational expectations is also central to debates about the conduct of monetary policy, to be discussed in Chapter 28.

Theoretically, the theory of rational expectations should be a powerful tool for analyzing behavior. But to establish that it is *in reality* a useful tool, we must compare the outcomes predicted by the theory with empirical evidence. Although the evidence is mixed and controversial, it indicates that for many purposes, the theory of rational expectations is a good starting point for analyzing expectations.

Computing the Price of Common Stock

Common stock is the principal way that corporations raise equity capital. Holders of common stock own an interest in the corporation consistent with the percentage of outstanding shares owned. This ownership interest gives **stockholders**—those who hold stock in a corporation—a bundle of rights. The most important are the right to vote and to be the **residual claimant** of all funds flowing into the firm (known as **cash flows**), meaning that the stockholder receives whatever remains after all other

claims against the firm's assets have been satisfied. Stockholders are paid dividends from the net earnings of the corporation. **Dividends** are payments made periodically, usually every quarter, to stockholders. The board of directors of the firm sets the level of the dividend, usually upon the recommendation of management. In addition, the stockholder has the right to sell the stock.

http://stocks.tradingcharts.com
Access detailed stock quotes, charts, and historical stock data.

One basic principle of finance is that the value of any investment is found by computing the value today of all cash flows the investment will generate over its life. For example, a commercial building will sell for a price that reflects the net cash flows (rents – expenses) it is projected to have over its useful life. Similarly, we value common stock as the value in today's dollars of all future cash flows. The cash flows a stockholder might earn from stock are dividends, the sales price, or both.

To develop the theory of stock valuation, we begin with the simplest possible scenario: You buy the stock, hold it for one period to get a dividend, then sell the stock. We call this the *one-period valuation model*.

The One-Period Valuation Model

Suppose that you have some extra money to invest for one year. After a year, you will need to sell your investment to pay tuition. After watching *CNBC* or *Wall Street Week* on TV, you decide that you want to buy Intel Corp. stock. You call your broker and find that Intel is currently selling for $50 per share and pays $0.16 per year in dividends. The analyst on *Wall Street Week* predicts that the stock will be selling for $60 in one year. Should you buy this stock?

To answer this question, you need to determine whether the current price accurately reflects the analyst's forecast. To value the stock today, you need to find the present discounted value of the expected cash flows (future payments) using the formula in Equation 1 of Chapter 4. Note that in this equation, the discount factor used to discount the cash flows is the required return on investments in equity rather than the interest rate. The cash flows consist of one dividend payment plus a final sales price. When these cash flows are discounted back to the present, the following equation computes the current price of the stock:

$$P_0 = \frac{Div_1}{(1 + k_e)} + \frac{P_1}{(1 + k_e)} \tag{1}$$

where
P_0 = the current price of the stock. The zero subscript refers to time period zero, or the present.
Div_1 = the dividend paid at the end of year 1.
k_e = the required return on investments in equity.
P_1 = the price at the end of the first period; the assumed sales price of the stock.

To see how Equation 1 works, let's compute the price of the Intel stock if, after careful consideration, you decide that you would be satisfied to earn a 12% return on the investment. If you have decided that $k_e = 0.12$, are told that Intel pays $0.16 per year in dividends ($Div_1 = 0.16$), and forecast the share price of $60 for next year ($P_1 = \$60$), you get the following from Equation 1:

$$P_0 = \frac{0.16}{1 + 0.12} + \frac{\$60}{1 + 0.12} = \$0.14 + \$53.57 = \$53.71$$

Based on your analysis, you find that the present value of all cash flows from the stock is $53.71. Because the stock is currently priced at $50 per share, you would choose to buy it. However, you should be aware that the stock may be selling for less than $53.71, because other investors place a different risk on the cash flows or estimate the cash flows to be less than you do.

The Generalized Dividend Valuation Model

Using the same concept, the one-period dividend valuation model can be extended to any number of periods: The value of stock is the present value of all future cash flows. The only cash flows that an investor will receive are dividends and a final sales price when the stock is ultimately sold in period n. The generalized multi-period formula for stock valuation can be written as:

$$P_0 = \frac{D_1}{(1 + k_e)^1} + \frac{D_2}{(1 + k_e)^2} + \ldots + \frac{D_n}{(1 + k_e)^n} + \frac{P_n}{(1 + k_e)^n} \qquad (2)$$

If you tried to use Equation 2 to find the value of a share of stock, you would soon realize that you must first estimate the value the stock will have at some point in the future before you can estimate its value today. In other words, you must find P_n in order to find P_0. However, if P_n is far in the future, it will not affect P_0. For example, the present value of a share of stock that sells for $50 seventy-five years from now using a 12% discount rate is just one cent [$\$50/(1.12^{75}) = \$0.01$]. This reasoning implies that the current value of a share of stock can be calculated as simply the present value of the future dividend stream. The **generalized dividend model** is rewritten in Equation 3 without the final sales price:

$$P_0 = \sum_{t=1}^{\infty} \frac{D_t}{(1 + k_e)^t} \qquad (3)$$

Consider the implications of Equation 3 for a moment. The generalized dividend model says that the price of stock is determined only by the present value of the dividends and that nothing else matters. Many stocks do not pay dividends, so how is it that these stocks have value? *Buyers of the stock expect that the firm will pay dividends someday.* Most of the time a firm institutes dividends as soon as it has completed the rapid growth phase of its life cycle.

The generalized dividend valuation model requires that we compute the present value of an infinite stream of dividends, a process that could be difficult, to say the least. Therefore, simplified models have been developed to make the calculations easier. One such model is the **Gordon growth model**, which assumes constant dividend growth.

The Gordon Growth Model

Many firms strive to increase their dividends at a constant rate each year. Equation 4 rewrites Equation 3 to reflect this constant growth in dividends:

$$P_0 = \frac{D_0 \times (1 + g)^1}{(1 + k_e)^1} + \frac{D_0 \times (1 + g)^2}{(1 + k_e)^2} + \cdots + \frac{D_0 \times (1 + g)^\infty}{(1 + k_e)^\infty} \qquad (4)$$

where
D_0 = the most recent dividend paid
g = the expected constant growth rate in dividends
k_e = the required return on an investment in equity

Equation 4 has been simplified using algebra to obtain Equation 5.[1]

$$P_0 = \frac{D_0 \times (1 + g)}{(k_e - g)} = \frac{D_1}{(k_e - g)} \quad (5)$$

This model is useful for finding the value of stock, given a few assumptions:

1. *Dividends are assumed to continue growing at a constant rate forever.* Actually, as long as they are expected to grow at a constant rate for an extended period of time, the model should yield reasonable results. This is because errors about distant cash flows become small when discounted to the present.
2. *The growth rate is assumed to be less than the required return on equity, k_e.* Myron Gordon, in his development of the model, demonstrated that this is a reasonable assumption. In theory, if the growth rate were faster than the rate demanded by holders of the firm's equity, in the long run the firm would grow impossibly large.

How the Market Sets Security Prices

Suppose you went to an auto auction. The cars are available for inspection before the auction begins, and you find a little Mazda Miata that you like. You test-drive it in the parking lot and notice that it makes a few strange noises, but you decide that you would still like the car. You decide $5,000 would be a fair price that would allow you to pay some repair bills should the noises turn out to be serious. You see that the auction is ready to begin, so you go in and wait for the Miata to enter.

Suppose there is another buyer who also spots the Miata. He test-drives the car and recognizes that the noises are simply the result of worn brake pads that he can fix himself at a nominal cost. He decides that the car is worth $7,000. He also goes in and waits for the Miata to enter.

Who will buy the car and for how much? Suppose only the two of you are interested in the Miata. You begin the bidding at $4,000. He ups your bid to $4,500. You

[1] To generate Equation 5 from Equation 4, first multiply both sides of Equation 4 by $(1 + k_e)/(1 + g)$ and subtract Equation 4 from the result. This yields:

$$\frac{P_0 \times (1 + k_e)}{(1 + g)} - P_0 = D_0 - \frac{D_0 \times (1 + g)^{\infty}}{(1 + k_e)^{\infty}}$$

Assuming that k_e is greater than g, the term on the far right will approach zero and can be dropped. Thus, after factoring P_0 out of the left-hand side:

$$P_0 \times \left[\frac{1 + k_e}{1 + g} - 1\right] = D_0$$

Next, simplify by combining terms to:

$$P_0 \times \frac{(1 + k_e) - (1 + g)}{1 + g} = D_0$$

$$P_0 = \frac{D_0 \times (1 + g)}{k_e - g} = \frac{D_1}{k_e - g}$$

bid your top price of $5,000. He counters with $5,100. The price is now higher than you are willing to pay, so you stop bidding. The car is sold to the more informed buyer for $5,100.

This simple example raises a number of points. First, the price is set by the buyer willing to pay the highest price. The price is not necessarily the highest price the asset could fetch, but it is incrementally greater than what any other buyer is willing to pay.

Second, the market price will be set by the buyer who can take best advantage of the asset. The buyer who purchased the car knew that he could fix the noise easily and cheaply. Because of this he was willing to pay more for the car than you were. The same concept holds for other assets. For example, a piece of property or a building will sell to the buyer who can put the asset to the most productive use.

Finally, the example shows the role played by information in asset pricing. Superior information about an asset can increase its value by reducing its risk. When you consider buying a stock, there are many unknowns about the future cash flows. The buyer who has the best information about these cash flows will discount them at a lower interest rate than will a buyer who is very uncertain.

Now let us apply these ideas to stock valuation. Suppose that you are considering the purchase of stock expected to pay a $2 dividend next year. Market analysts expect the firm to grow at 3% indefinitely. You are *uncertain* about both the constancy of the dividend stream and the accuracy of the estimated growth rate. To compensate yourself for this uncertainty (risk), you require a return of 15%.

Now suppose Jennifer, another investor, has spoken with industry insiders and feels more confident about the projected cash flows. Jennifer requires only a 12% return because her perceived risk is lower than yours. Bud, on the other hand, is dating the CEO of the company. He knows with more certainty what the future of the firm actually is, and thus requires only a 10% return.

What are the values each investor will give to the stock? Applying the Gordon growth model yields the following stock prices:

Investor	Discount Rate	Stock Price
You	15%	$16.67
Jennifer	12%	$22.22
Bud	10%	$28.57

You are willing to pay $16.67 for the stock. Jennifer would pay up to $22.22, and Bud would pay $28.57. The investor with the lowest perceived risk is willing to pay the most for the stock. If there were no other traders but these three, the market price would be between $22.22 and $28.57. If you already held the stock, you would sell it to Bud.

We thus see that the players in the market, bidding against each other, establish the market price. When new information is released about a firm, expectations change and with them, prices change. New information can cause changes in expectations about the level of future dividends or the risk of those dividends. Since market participants are constantly receiving new information and revising their expectations, it is reasonable that stock prices are constantly changing as well.

Application

Monetary Policy and Stock Prices

Stock market analysts tend to hang on every word that the Chairman of the Federal Reserve utters because they know that an important determinant of stock prices is monetary policy. But how does monetary policy affect stock prices?

The Gordon growth model in Equation 5 tells us how. Monetary policy can affect stock prices in two ways. First, when the Fed lowers interest rates, the return on bonds (an alternative asset to stocks) declines, and investors are likely to accept a lower required rate of return on an investment in equity (k_e). The resulting decline in k_e would lower the denominator in the Gordon growth model (Equation 5), lead to a higher value of P_0, and raise stock prices. Furthermore, a lowering of interest rates is likely to stimulate the economy, so that the growth rate in dividends, g, is likely to be somewhat higher. This rise in g also causes the denominator in Equation 5 to fall, which also leads to a higher P_0 and a rise in stock prices.

As we will see in Chapter 26, the impact of monetary policy on stock prices is one of the key ways in which monetary policy affects the economy.

Application

The September 11 Terrorist Attacks, the Enron Scandal, and the Stock Market

In 2001, two big shocks hit the stock market: the September 11 terrorist attacks and the Enron scandal. Our analysis of stock price evaluation, again using the Gordon growth model, can help us understand how these events affected stock prices.

The September 11 terrorist attacks raised the possibility that terrorism against the United States would paralyze the country. These fears led to a downward revision of the growth prospects for U.S. companies, thus lowering the dividend growth rate (g) in the Gordon model. The resulting rise in the denominator in Equation 5 would lead to a decline in P_0 and hence a decline in stock prices.

Increased uncertainty for the U.S. economy would also raise the required return on investment in equity. A higher k_e also leads to a rise in the denominator in Equation 5, a decline in P_0, and a general fall in stock prices. As the Gordon model predicts, the stock market fell by over 10% immediately after September 11.

Subsequently, the U.S. successes against the Taliban in Afghanistan and the absence of further terrorist attacks reduced market fears and uncertainty, causing g to recover and k_e to fall. The denominator in Equation 5 then fell, leading to a recovery in P_0 and a rebound in the stock market in October and November. However, by the beginning of 2002, the Enron scandal and disclosures that many companies had overstated their earnings caused many investors to doubt the formerly rosy forecast of earnings and dividend growth

for corporations. The resulting revision of g downward, and the rise in k_e because of increased uncertainty about the quality of accounting information, would lead to a rise in the denominator in the Gordon Equation 5, thereby lowering P_0 for many companies and hence the overall stock market. As predicted by our analysis, this is exactly what happened. The stock market recovery was aborted, and the market began a downward slide.

The Theory of Rational Expectations

The analysis of stock price evaluation we have outlined in the previous section depends on people's expectations—especially of cash flows. Indeed, it is difficult to think of any sector in the economy in which expectations are not crucial; this is why it is important to examine how expectations are formed. We do so by outlining the *theory of rational expectations*, currently the most widely used theory to describe the formation of business and consumer expectations.

In the 1950s and 1960s, economists regularly viewed expectations as formed from past experience only. Expectations of inflation, for example, were typically viewed as being an average of past inflation rates. This view of expectation formation, called **adaptive expectations**, suggests that changes in expectations will occur slowly over time as past data change.[2] So if inflation had formerly been steady at a 5% rate, expectations of future inflation would be 5% too. If inflation rose to a steady rate of 10%, expectations of future inflation would rise toward 10%, but slowly: In the first year, expected inflation might rise only to 6%; in the second year, to 7%; and so on.

Adaptive expectations have been faulted on the grounds that people use more information than just past data on a single variable to form their expectations of that variable. Their expectations of inflation will almost surely be affected by their predictions of future monetary policy as well as by current and past monetary policy. In addition, people often change their expectations quickly in the light of new information. To meet these objections to adaptive expectations, John Muth developed an alternative theory of expectations, called **rational expectations**, which can be stated as follows: ***Expectations will be identical to optimal forecasts (the best guess of the future) using all available information.***[3]

What exactly does this mean? To explain it more clearly, let's use the theory of rational expectations to examine how expectations are formed in a situation that most of us encounter at some point in our lifetime: our drive to work. Suppose that when Joe Commuter travels when it is not rush hour, it takes an average of 30 minutes for

---

[2]More specifically, adaptive expectations—say, of inflation—are written as a weighted average of past inflation rates:

$$\pi_t^e = (1 - \lambda)\sum_{j=0}^{\infty} \lambda^j \pi_{t-j}$$

where π_t^e = adaptive expectation of inflation at time t
π_{t-j} = inflation at time $t - j$
λ = a constant between the values of 0 and 1

[3]John Muth, "Rational Expectations and the Theory of Price Movements," *Econometrica* 29 (1961): 315–335.

his trip. Sometimes it takes him 35 minutes, other times 25 minutes, but the average non-rush-hour driving time is 30 minutes. If, however, Joe leaves for work during the rush hour, it takes him, on average, an additional 10 minutes to get to work. Given that he leaves for work during the rush hour, the best guess of the driving time—the **optimal forecast**—is 40 minutes.

If the only information available to Joe before he leaves for work that would have a potential effect on his driving time is that he is leaving during the rush hour, what does rational expectations theory allow you to predict about Joe's expectations of his driving time? Since the best guess of his driving time using all available information is 40 minutes, Joe's expectation should also be the same. Clearly, an expectation of 35 minutes would not be rational, because it is not equal to the optimal forecast, the best guess of the driving time.

Suppose that the next day, given the same conditions and the same expectations, it takes Joe 45 minutes to drive because he hits an abnormally large number of red lights, and the day after that he hits all the lights right and it takes him only 35 minutes. Do these variations mean that Joe's 40-minute expectation is irrational? No, an expectation of 40 minutes' driving time is still a rational expectation. In both cases, the forecast is off by 5 minutes, so the expectation has not been perfectly accurate. However, the forecast does not have to be perfectly accurate to be rational—it need only be the *best possible* given the available information; that is, it has to be correct *on average,* and the 40-minute expectation meets this requirement. Since there is bound to be some randomness in Joe's driving time regardless of driving conditions, an optimal forecast will never be completely accurate.

The example makes the following important point about rational expectations: ***Even though a rational expectation equals the optimal forecast using all available information, a prediction based on it may not always be perfectly accurate.***

What if an item of information relevant to predicting driving time is unavailable or ignored? Suppose that on Joe's usual route to work there is an accident that causes a two-hour traffic jam. If Joe has no way of ascertaining this information, his rush-hour expectation of 40 minutes' driving time is still rational, because the accident information is not available to him for incorporation into his optimal forecast. However, if there was a radio or TV traffic report about the accident that Joe did not bother to listen to or heard but ignored, his 40-minute expectation is no longer rational. In light of the availability of this information, Joe's optimal forecast should have been two hours and 40 minutes.

Accordingly, there are two reasons why an expectation may fail to be rational:

1. People might be aware of all available information but find it takes too much effort to make their expectation the best guess possible.
2. People might be unaware of some available relevant information, so their best guess of the future will not be accurate.

Nonetheless, it is important to recognize that if an additional factor is important but information about it is not available, an expectation that does not take account of it can still be rational.

Formal Statement of the Theory

We can state the theory of rational expectations somewhat more formally. If X stands for the variable that is being forecast (in our example, Joe Commuter's driving time), X^e for the expectation of this variable (Joe's expectation of his driving time), and X^{of}

for the optimal forecast of X using all available information (the best guess possible of his driving time), the theory of rational expectations then simply says:

$$X^{e} = X^{of} \tag{6}$$

That is, the expectation of X equals the optimal forecast using all available information.

Rationale Behind the Theory

Why do people try to make their expectations match their best possible guess of the future using all available information? The simplest explanation is that it is costly for people not to do so. Joe Commuter has a strong incentive to make his expectation of the time it takes him to drive to work as accurate as possible. If he underpredicts his driving time, he will often be late to work and risk being fired. If he overpredicts, he will, on average, get to work too early and will have given up sleep or leisure time unnecessarily. Accurate expectations are desirable, and there are strong incentives for people to try to make them equal to optimal forecasts by using all available information.

The same principle applies to businesses. Suppose that an appliance manufacturer—say, General Electric—knows that interest-rate movements are important to the sales of appliances. If GE makes poor forecasts of interest rates, it will earn less profit, because it might produce either too many appliances or too few. There are strong incentives for GE to acquire all available information to help it forecast interest rates and use the information to make the best possible guess of future interest-rate movements.

The incentives for equating expectations with optimal forecasts are especially strong in financial markets. In these markets, people with better forecasts of the future get rich. The application of the theory of rational expectations to financial markets (where it is called the **efficient market hypothesis** or the **theory of efficient capital markets**) is thus particularly useful.

Implications of the Theory

Rational expectations theory leads to two commonsense implications for the forming of expectations that are important in the analysis of the aggregate economy:

1. ***If there is a change in the way a variable moves, the way in which expectations of this variable are formed will change as well.*** This tenet of rational expectations theory can be most easily understood through a concrete example. Suppose that interest rates move in such a way that they tend to return to a "normal" level in the future. If today's interest rate is high relative to the normal level, an optimal forecast of the interest rate in the future is that it will decline to the normal level. Rational expectations theory would imply that when today's interest rate is high, the expectation is that it will fall in the future.

Suppose now that the way in which the interest rate moves changes so that when the interest rate is high, it stays high. In this case, when today's interest rate is high, the optimal forecast of the future interest rate, and hence the rational expectation, is that it will stay high. Expectations of the future interest rate will no longer indicate that the interest rate will fall. The change in the way the interest-rate variable moves has therefore led to a change in the way that expectations of future interest rates are formed. The rational expectations analysis here is generalizable to expectations of any variable. Hence when there is a change in the way any variable moves, the way in which expectations of this variable are formed will change too.

2. ***The forecast errors of expectations will on average be zero and cannot be predicted ahead of time.*** The forecast error of an expectation is $X - X^e$, the difference between the realization of a variable X and the expectation of the variable; that is, if Joe Commuter's driving time on a particular day is 45 minutes and his expectation of the driving time is 40 minutes, the forecast error is 5 minutes.

Suppose that in violation of the rational expectations tenet, Joe's forecast error is not, on average, equal to zero; instead, it equals 5 minutes. The forecast error is now predictable ahead of time because Joe will soon notice that he is, on average, 5 minutes late for work and can improve his forecast by increasing it by 5 minutes. Rational expectations theory implies that this is exactly what Joe will do because he will want his forecast to be the best guess possible. When Joe has revised his forecast upward by 5 minutes, on average, the forecast error will equal zero so that it cannot be predicted ahead of time. Rational expectations theory implies that forecast errors of expectations cannot be predicted.

The Efficient Market Hypothesis: Rational Expectations in Financial Markets

www.investorhome.com/emh.htm

Learn more about the efficient market hypothesis.

While the theory of rational expectations was being developed by monetary economists, financial economists were developing a parallel theory of expectation formation in financial markets. It led them to the same conclusion as that of the rational expectations theorists: Expectations in financial markets are equal to optimal forecasts using all available information.[4] Although financial economists gave their theory another name, calling it *the efficient market hypothesis,* in fact their theory is just an application of rational expectations to the pricing of securities.

The efficient market hypothesis is based on the assumption that prices of securities in financial markets fully reflect all available information. You may recall from Chapter 4 that the rate of return from holding a security equals the sum of the capital gain on the security (the change in the price), plus any cash payments, divided by the initial purchase price of the security:

$$R = \frac{P_{t+1} - P_t + C}{P_t} \tag{7}$$

where R = rate of return on the security held from time t to $t + 1$ (say, the end of 2000 to the end of 2001)
P_{t+1} = price of the security at time $t + 1$, the end of the holding period
P_t = price of the security at time t, the beginning of the holding period
C = cash payment (coupon or dividend payments) made in the period t to $t + 1$

Let's look at the expectation of this return at time t, the beginning of the holding period. Because the current price P_t and the cash payment C are known at the beginning, the only variable in the definition of the return that is uncertain is the price next

[4]The development of the efficient market hypothesis was not wholly independent of the development of rational expectations theory, in that financial economists were aware of Muth's work.

period, P_{t+1}.[5] Denoting the expectation of the security's price at the end of the holding period as P^{e}_{t+1}, the expected return R^{e} is:

$$R^{e} = \frac{P^{e}_{t+1} - P_t + C}{P_t}$$

The efficient market hypothesis also views expectations of future prices as equal to optimal forecasts using all currently available information. In other words, the market's expectations of future securities prices are rational, so that:

$$P^{e}_{t+1} = P^{of}_{t+1}$$

which in turn implies that the expected return on the security will equal the optimal forecast of the return:

$$R^{e} = R^{of} \tag{8}$$

Unfortunately, we cannot observe either R^{e} or P^{e}_{t+1}, so the rational expectations equations by themselves do not tell us much about how the financial market behaves. However, if we can devise some way to measure the value of R^{e}, these equations will have important implications for how prices of securities change in financial markets.

The supply and demand analysis of the bond market developed in Chapter 5 shows us that the expected return on a security (the interest rate, in the case of the bond examined) will have a tendency to head toward the equilibrium return that equates the quantity demanded to the quantity supplied. Supply and demand analysis enables us to determine the expected return on a security with the following equilibrium condition: The expected return on a security R^{e} equals the equilibrium return R^{*}, which equates the quantity of the security demanded to the quantity supplied; that is,

$$R^{e} = R^{*} \tag{9}$$

The academic field of finance explores the factors (risk and liquidity, for example) that influence the equilibrium returns on securities. For our purposes, it is sufficient to know that we can determine the equilibrium return and thus determine the expected return with the equilibrium condition.

We can derive an equation to describe pricing behavior in an efficient market by using the equilibrium condition to replace R^{e} with R^{*} in the rational expectations equation (Equation 8). In this way, we obtain:

$$R^{of} = R^{*} \tag{10}$$

This equation tells us that ***current prices in a financial market will be set so that the optimal forecast of a security's return using all available information equals the security's equilibrium return***. Financial economists state it more simply: In an efficient market, a security's price fully reflects all available information.

Rationale Behind the Hypothesis

Let's see what the efficient markets condition means in practice and why it is a sensible characterization of pricing behavior. Suppose that the equilibrium return on a security—say, Exxon common stock—is 10% at an annual rate, and its current price

[5] There are cases where C might not be known at the beginning of the period, but that does not make a substantial difference to the analysis. We would in that case assume that not only price expectations but also the expectations of C are optimal forecasts using all available information.

P_t is lower than the optimal forecast of tomorrow's price P^{of}_{t+1} so that the optimal forecast of the return at an annual rate is 50%, which is greater than the equilibrium return of 10%. We are now able to predict that, on average, Exxon's return would be abnormally high. This situation is called an **unexploited profit opportunity** because, on average, people would be earning more than they should, given the characteristics of that security. Knowing that, on average, you can earn such an abnormally high rate of return on Exxon because $R^{of} > R^*$, you would buy more, which would in turn drive up its current price P_t relative to the expected future price P^{of}_{t+1}, thereby lowering R^{of}. When the current price had risen sufficiently so that R^{of} equals R^* and the efficient markets condition (Equation 10) is satisfied, the buying of Exxon will stop, and the unexploited profit opportunity will have disappeared.

Similarly, a security for which the optimal forecast of the return is −5% and the equilibrium return is 10% ($R^{of} < R^*$) would be a poor investment, because, on average, it earns less than the equilibrium return. In such a case, you would sell the security and drive down its current price relative to the expected future price until R^{of} rose to the level of R^* and the efficient markets condition is again satisfied. What we have shown can be summarized as follows:

$$\begin{gathered} R^{of} > R^* \rightarrow P_t\uparrow \rightarrow R^{of}\downarrow \\ R^{of} < R^* \rightarrow P_t\downarrow \rightarrow R^{of}\uparrow \\ \text{until} \\ R^{of} = R^* \end{gathered}$$

Another way to state the efficient markets condition is this: ***In an efficient market, all unexploited profit opportunities will be eliminated.***

An extremely important factor in this reasoning is that ***not everyone in a financial market must be well informed about a security or have rational expectations for its price to be driven to the point at which the efficient markets condition holds***. Financial markets are structured so that many participants can play. As long as a few keep their eyes open for unexploited profit opportunities, they will eliminate the profit opportunities that appear, because in so doing, they make a profit. The efficient market hypothesis makes sense, because it does not require everyone in a market to be cognizant of what is happening to every security.

Stronger Version of the Efficient Market Hypothesis

Many financial economists take the efficient market hypothesis one step further in their analysis of financial markets. Not only do they define efficient markets as those in which expectations are rational—that is, equal to optimal forecasts using all available information—but they also add the condition that an efficient market is one in which prices reflect the true fundamental (intrinsic) value of the securities. Thus in an efficient market, all prices are always correct and reflect **market fundamentals** (items that have a direct impact on future income streams of the securities). This stronger view of market efficiency has several important implications in the academic field of finance. First, it implies that in an efficient capital market, one investment is as good as any other because the securities' prices are correct. Second, it implies that a security's price reflects all available information about the intrinsic value of the security. Third, it implies that security prices can be used by managers of both financial and nonfinancial firms to assess their cost of capital (cost of financing their investments) accurately and hence that security prices can be used to help them make the correct decisions about whether a specific investment is worth making or not. The stronger version of market efficiency is a basic tenet of much analysis in the finance field.

Evidence on the Efficient Market Hypothesis

Early evidence on the efficient market hypothesis was quite favorable to it, but in recent years, deeper analysis of the evidence suggests that the hypothesis may not always be entirely correct. Let's first look at the earlier evidence in favor of the hypothesis and then examine some of the more recent evidence that casts some doubt on it.

Evidence in Favor of Market Efficiency

Evidence in favor of market efficiency has examined the performance of investment analysts and mutual funds, whether stock prices reflect publicly available information, the random-walk behavior of stock prices, and the success of so-called technical analysis.

Performance of Investment Analysts and Mutual Funds. We have seen that one implication of the efficient market hypothesis is that when purchasing a security, you cannot expect to earn an abnormally high return, a return greater than the equilibrium return. This implies that it is impossible to beat the market. Many studies shed light on whether investment advisers and mutual funds (some of which charge steep sales commissions to people who purchase them) beat the market. One common test that has been performed is to take buy and sell recommendations from a group of advisers or mutual funds and compare the performance of the resulting selection of stocks with the market as a whole. Sometimes the advisers' choices have even been compared to a group of stocks chosen by throwing darts at a copy of the financial page of the newspaper tacked to a dartboard. The *Wall Street Journal*, for example, has a regular feature called "Investment Dartboard" that compares how well stocks picked by investment advisers do relative to stocks picked by throwing darts. Do the advisers win? To their embarrassment, the dartboard beats them as often as they beat the dartboard. Furthermore, even when the comparison includes only advisers who have been successful in the past in predicting the stock market, the advisers still don't regularly beat the dartboard.

Consistent with the efficient market hypothesis, mutual funds also do not beat the market. Not only do mutual funds not outperform the market on average, but when they are separated into groups according to whether they had the highest or lowest profits in a chosen period, the mutual funds that did well in the first period do not beat the market in the second period.[6]

The conclusion from the study of investment advisers and mutual fund performance is this: ***Having performed well in the past does not indicate that an investment adviser or a mutual fund will perform well in the future.*** This is not pleasing news to investment advisers, but it is exactly what the efficient market hypothesis predicts. It says that some advisers will be lucky and some will be unlucky. Being lucky does not mean that a forecaster actually has the ability to beat the market.

[6]An early study that found that mutual funds do not outperform the market is Michael C. Jensen, "The Performance of Mutual Funds in the Period 1945–64," *Journal of Finance* 23 (1968): 389–416. Further studies on mutual fund performance are Mark Grimblatt and Sheridan Titman, "Mutual Fund Performance: An Analysis of Quarterly Portfolio Holdings," *Journal of Business* 62 (1989): 393–416; R. A. Ippolito, "Efficiency with Costly Information: A Study of Mutual Fund Performance, 1965–84," *Quarterly Journal of Economics* 104 (1989): 1–23; J. Lakonishok, A. Shleifer, and R. Vishny, "The Structure and Performance of the Money Management Industry," *Brookings Papers on Economic Activity, Microeconomics* (1992); and B. Malkiel, "Returns from Investing in Equity Mutual Funds, 1971–1991," *Journal of Finance* 50 (1995): 549–72.

Do Stock Prices Reflect Publicly Available Information? The efficient market hypothesis predicts that stock prices will reflect all publicly available information. Thus if information is already publicly available, a positive announcement about a company will not, on average, raise the price of its stock because this information is already reflected in the stock price. Early empirical evidence also confirmed this conjecture from the efficient market hypothesis: Favorable earnings announcements or announcements of stock splits (a division of a share of stock into multiple shares, which is usually followed by higher earnings) do not, on average, cause stock prices to rise.[7]

Random-Walk Behavior of Stock Prices. The term **random walk** describes the movements of a variable whose future changes cannot be predicted (are random) because, given today's value, the variable is just as likely to fall as to rise. An important implication of the efficient market hypothesis is that stock prices should approximately follow a random walk; that is, ***future changes in stock prices should, for all practical purposes, be unpredictable***. The random-walk implication of the efficient market hypothesis is the one most commonly mentioned in the press, because it is the most readily comprehensible to the public. In fact, when people mention the "random-walk theory of stock prices," they are in reality referring to the efficient market hypothesis.

The case for random-walk stock prices can be demonstrated. Suppose that people could predict that the price of Happy Feet Corporation (HFC) stock would rise 1% in the coming week. The predicted rate of capital gains and rate of return on HFC stock would then be over 50% at an annual rate. Since this is very likely to be far higher than the equilibrium rate of return on HFC stock ($R^{of} > R^*$), the efficient markets hypothesis indicates that people would immediately buy this stock and bid up its current price. The action would stop only when the predictable change in the price dropped to near zero so that $R^{of} = R^*$.

Similarly, if people could predict that the price of HFC stock would fall by 1%, the predicted rate of return would be negative ($R^{of} < R^*$), and people would immediately sell. The current price would fall until the predictable change in the price rose back to near zero, where the efficient market condition again holds. The efficient market hypothesis suggests that the predictable change in stock prices will be near zero, leading to the conclusion that stock prices will generally follow a random walk.[8]

Financial economists have used two types of tests to explore the hypothesis that stock prices follow a random walk. In the first, they examine stock market records to see if changes in stock prices are systematically related to past changes and hence could have been predicted on that basis. The second type of test examines the data to see if publicly available information other than past stock prices could have been used to predict changes. These tests are somewhat more stringent because additional information (money supply growth, government spending, interest rates, corporate profits) might be used to help forecast stock returns. Early results from both types of tests

---

[7]Ray Ball and Philip Brown, "An Empirical Evaluation of Accounting Income Numbers," *Journal of Accounting Research* 6 (1968):159–178, and Eugene F. Fama, Lawrence Fisher, Michael C. Jensen, and Richard Roll, "The Adjustment of Stock Prices to New Information," *International Economic Review* 10 (1969): 1–21.

[8]Note that the random-walk behavior of stock prices is only an *approximation* derived from the efficient market hypothesis. It would hold exactly only for a stock for which an unchanged price leads to its having the equilibrium return. Then, when the predictable change in the stock price is exactly zero, $R^{of} = R^*$.

generally confirmed the efficient market view that stock prices are not predictable and follow a random walk.[9]

Technical Analysis. A popular technique used to predict stock prices, called *technical analysis*, is to study past stock price data and search for patterns such as trends and regular cycles. Rules for when to buy and sell stocks are then established on the basis of the patterns that emerge. The efficient market hypothesis suggests that technical analysis is a waste of time. The simplest way to understand why is to use the random-walk result derived from the efficient market hypothesis that holds that past stock price data cannot help predict changes. Therefore, technical analysis, which relies on such data to produce its forecasts, cannot successfully predict changes in stock prices.

Two types of tests bear directly on the value of technical analysis. The first performs the empirical analysis described earlier to evaluate the performance of any financial analyst, technical or otherwise. The results are exactly what the efficient market hypothesis predicts: Technical analysts fare no better than other financial analysts; on average, they do not outperform the market, and successful past forecasting does not imply that their forecasts will outperform the market in the future. The second type of test (first performed by Sidney Alexander) takes the rules developed in technical analysis for when to buy and sell stocks and applies them to new data.[10] The performance of these rules is then evaluated by the profits that would have been made using them. These tests also discredit technical analysis: It does not outperform the overall market.

Application

Should Foreign Exchange Rates Follow a Random Walk?

Although the efficient market hypothesis is usually applied to the stock market, it can also be used to show that foreign exchange rates, like stock prices, should generally follow a random walk. To see why this is the case, consider what would happen if people could predict that a currency would appreciate

[9]The first type of test, using only stock market data, is referred to as a test of *weak-form efficiency*, because the information that can be used to predict stock prices is restricted to past price data. The second type of test is referred to as a test of *semistrong-form efficiency*, because the information set is expanded to include all publicly available information, not just past stock prices. A third type of test is called a test of *strong-form efficiency*, because the information set includes insider information, known only to the managers (directors) of the corporation, as when they plan to declare a high dividend. Strong-form tests do sometimes indicate that insider information can be used to predict changes in stock prices. This finding does not contradict the efficient market hypothesis, because the information is not available to the market and hence cannot be reflected in market prices. In fact, there are strict laws against using insider information to trade in financial markets. For an early survey on the three forms of tests, see Eugene F. Fama, "Efficient Capital Markets: A Review of Theory and Empirical Work," *Journal of Finance* 25 (1970): 383–416.

[10]Sidney Alexander, "Price Movements in Speculative Markets: Trends or Random Walks?" *Industrial Management Review,* May 1961, pp. 7–26, and Sidney Alexander, "Price Movements in Speculative Markets: Trends or Random Walks? No. 2," in *The Random Character of Stock Prices,* ed. Paul Cootner (Cambridge, Mass.: MIT Press, 1964), pp. 338–372. More recent evidence also seems to discredit technichal analysis; for example, F. Allen and R. Karjalainen, "Using Genetic Algorithms to Find Technical Trading Rules," *Journal of Financial Economics* 51 (1999): 245–271. However, some other research is more favorable to technical analysis: e.g., R. Sullivan, A. Timmerman, and H. White, "Data-Snooping, Technical Trading Rule Performance and the Bootstrap," Centre for Economic Policy Research Discussion Paper No. 1976, 1998.

by 1% in the coming week. By buying this currency, they could earn a greater than 50% return at an annual rate, which is likely to be far above the equilibrium return for holding a currency. As a result, people would immediately buy the currency and bid up its current price, thereby reducing the expected return. The process would stop only when the predictable change in the exchange rate dropped to near zero so that the optimal forecast of the return no longer differed from the equilibrium return. Likewise, if people could predict that the currency would depreciate by 1% in the coming week, they would sell it until the predictable change in the exchange rate was again near zero. The efficient market hypothesis therefore implies that future changes in exchange rates should, for all practical purposes, be unpredictable; in other words, exchange rates should follow random walks. This is exactly what empirical evidence finds.[11]

Evidence Against Market Efficiency

All the early evidence supporting the efficient market hypothesis appeared to be overwhelming, causing Eugene Fama, a prominent financial economist, to state in his famous 1970 survey of the empirical evidence on the efficient market hypothesis, "The evidence in support of the efficient markets model is extensive, and (somewhat uniquely in economics) contradictory evidence is sparse."[12] However, in recent years, the hypothesis has begun to show a few cracks, referred to as *anomalies*, and empirical evidence indicates that the efficient market hypothesis may not always be generally applicable.

Small-Firm Effect. One of the earliest reported anomalies in which the stock market did not appear to be efficient is called the *small-firm effect*. Many empirical studies have shown that small firms have earned abnormally high returns over long periods of time, even when the greater risk for these firms has been taken into account.[13] The small-firm effect seems to have diminished in recent years, but is still a challenge to the efficient market hypothesis. Various theories have been developed to explain the small-firm effect, suggesting that it may be due to rebalancing of portfolios by institutional investors, tax issues, low liquidity of small-firm stocks, large information costs in evaluating small firms, or an inappropriate measurement of risk for small-firm stocks.

January Effect. Over long periods of time, stock prices have tended to experience an abnormal price rise from December to January that is predictable and hence inconsistent with random-walk behavior. This so-called **January effect** seems to have diminished in recent years for shares of large companies but still occurs for shares of small companies.[14] Some financial economists argue that the January effect is due to

[11]See Richard A. Meese and Kenneth Rogoff, "Empirical Exchange Rate Models of the Seventies: Do They Fit Out of Sample?" *Journal of International Economics* 14 (1983): 3–24.

[12]Eugene F. Fama, "Efficient Capital Markets: A Review of Theory and Empirical Work," *Journal of Finance* 25 (1970): 383–416.

[13]For example, see Marc R. Reinganum, "The Anomalous Stock Market Behavior of Small Firms in January: Empirical Tests of Tax Loss Selling Effects," *Journal of Financial Economics* 12 (1983): 89–104; Jay R. Ritter, "The Buying and Selling Behavior of Individual Investors at the Turn of the Year," *Journal of Finance* 43 (1988): 701–717; and Richard Roll, "Vas Ist Das? The Turn-of-the-Year Effect: Anomaly or Risk Mismeasurement?" *Journal of Portfolio Management* 9 (1988): 18–28.

[14]For example, see Donald B. Keim, "The CAPM and Equity Return Regularities," *Financial Analysts Journal* 42 (May–June 1986): 19–34.

tax issues. Investors have an incentive to sell stocks before the end of the year in December, because they can then take capital losses on their tax return and reduce their tax liability. Then when the new year starts in January, they can repurchase the stocks, driving up their prices and producing abnormally high returns. Although this explanation seems sensible, it does not explain why institutional investors such as private pension funds, which are not subject to income taxes, do not take advantage of the abnormal returns in January and buy stocks in December, thus bidding up their price and eliminating the abnormal returns.[15]

Market Overreaction. Recent research suggests that stock prices may overreact to news announcements and that the pricing errors are corrected only slowly.[16] When corporations announce a major change in earnings—say, a large decline—the stock price may overshoot, and after an initial large decline, it may rise back to more normal levels over a period of several weeks. This violates the efficient market hypothesis, because an investor could earn abnormally high returns, on average, by buying a stock immediately after a poor earnings announcement and then selling it after a couple of weeks when it has risen back to normal levels.

Excessive Volatility. A phenomenon closely related to market overreaction is that the stock market appears to display *excessive volatility*; that is, fluctuations in stock prices may be much greater than is warranted by fluctuations in their fundamental value. In an important paper, Robert Shiller of Yale University found that fluctuations in the S&P 500 stock index could not be justified by the subsequent fluctuations in the dividends of the stocks making up this index. There has been much subsequent technical work criticizing these results, but Shiller's work, along with research finding that there are smaller fluctuations in stock prices when stock markets are closed, has produced a consensus that stock market prices appear to be driven by factors other than fundamentals.[17]

Mean Reversion. Some researchers have also found that stock returns display **mean reversion**: Stocks with low returns today tend to have high returns in the future, and vice versa. Hence stocks that have done poorly in the past are more likely to do well in the future, because mean reversion indicates that there will be a predictable positive change in the future price, suggesting that stock prices are not a random walk. Other researchers have found that mean reversion is not nearly as strong in data after World

[15]Another anomaly that makes the stock market seem less than efficient is that the *Value Line Survey*, one of the most prominent investment advice newsletters, has produced stock recommendations that have yielded abnormally high returns on average. See Fischer Black, "Yes, Virginia, There Is Hope: Tests of the Value Line Ranking System," *Financial Analysts Journal* 29 (September–October 1973): 10–14, and Gur Huberman and Shmuel Kandel, "Market Efficiency and Value Line's Record," *Journal of Business* 63 (1990): 187–216. Whether the excellent performance of the *Value Line Survey* will continue in the future is, of course, a question mark.

[16]Werner De Bondt and Richard Thaler, "Further Evidence on Investor Overreaction and Stock Market Seasonality," *Journal of Finance* 62 (1987): 557–580.

[17]Robert Shiller, "Do Stock Prices Move Too Much to Be Justified by Subsequent Changes in Dividends?" *American Economic Review* 71 (1981): 421–436, and Kenneth R. French and Richard Roll, "Stock Return Variances: The Arrival of Information and the Reaction of Traders," *Journal of Financial Economics* 17 (1986): 5–26.

War II and so have raised doubts about whether it is currently an important phenomenon. The evidence on mean reversion remains controversial.[18]

New Information Is Not Always Immediately Incorporated into Stock Prices. Although it is generally found that stock prices adjust rapidly to new information, as is suggested by the efficient market hypothesis, recent evidence suggests that, inconsistent with the efficient market hypothesis, stock prices do not instantaneously adjust to profit announcements. Instead, on average stock prices continue to rise for some time after the announcement of unexpectedly high profits, and they continue to fall after surprisingly low profit announcments.[19]

Overview of the Evidence on the Efficient Market Hypothesis

As you can see, the debate on the efficient market hypothesis is far from over. The evidence seems to suggest that the efficient market hypothesis may be a reasonable starting point for evaluating behavior in financial markets. However, there do seem to be important violations of market efficiency that suggest that the efficient market hypothesis may not be the whole story and so may not be generalizable to all behavior in financial markets.

Application

Practical Guide to Investing in the Stock Market

The efficient market hypothesis has numerous applications to the real world. It is especially valuable because it can be applied directly to an issue that concerns many of us: how to get rich (or at least not get poor) in the stock market. (The "Following the Financial News" box shows how stock prices are reported daily.) A practical guide to investing in the stock market, which we develop here, provides a better understanding of the use and implications of the efficient market hypothesis.

How Valuable Are Published Reports by Investment Advisers?

Suppose you have just read in the "Heard on the Street" column of the *Wall Street Journal* that investment advisers are predicting a boom in oil stocks because an oil shortage is developing. Should you proceed to withdraw all your hard-earned savings from the bank and invest it in oil stocks?

---

[18]Evidence for mean reversion has been reported by James M. Poterba and Lawrence H. Summers, "Mean Reversion in Stock Prices: Evidence and Implications," *Journal of Financial Economics* 22 (1988): 27–59; Eugene F. Fama and Kenneth R. French, "Permanent and Temporary Components of Stock Prices," *Journal of Political Economy* 96 (1988): 246–273; and Andrew W. Lo and A. Craig MacKinlay, "Stock Market Prices Do Not Follow Random Walks: Evidence from a Simple Specification Test," *Review of Financial Studies* 1 (1988): 41–66. However, Myung Jig Kim, Charles R. Nelson, and Richard Startz, in "Mean Reversion in Stock Prices? A Reappraisal of the Evidence," *Review of Economic Studies* 58 (1991): 515–528, question whether some of these findings are valid. For an excellent summary of this evidence, see Charles Engel and Charles S. Morris, "Challenges to Stock Market Efficiency: Evidence from Mean Reversion Studies," *Federal Reserve Bank of Kansas City Economic Review,* September–October 1991, pp. 21–35. See also N. Jegadeesh and Sheridan Titman, "Returns to Buying Winners and Selling Losers: Implications for Stock Market Efficiency," *Journal of Finance* 48 (1993): 65–92, which shows that mean reversion also occurs for individual stocks.

[19]For example, see R. Ball and P. Brown, "An Empirical Evaluation of Accounting Income Numbers," *Journal of Accounting Research* 6 (1968): 159–178, L. Chan, N. Jegadeesh, and J. Lakonishok, "Momentum Strategies," *Journal of Finance* 51 (1996): 1681–1713, and Eugene Fama, "Market Efficiency, Long-Term Returns and Behavioral Finance," *Journal of Financial Economics* 49 (1998): 283–306.

Following the Financial News

Stock Prices

Stock prices are published daily, and in the *Wall Street Journal*, they are reported in the sections "NYSE—Composite Transactions," "Amex—Composite Transactions," and "NASDAQ National Market Issues." Stock prices are quoted in the following format:

YTD % Chg.	52-Week Hi	52-Week Lo	Stock (Sym.)	Div.	Yld. %	PE	Vol. 100s	Close	Net Chg.
0.6	23.85	15.50♣	IntAlum **IAL**	1.20	6.9	88	21	17.39	0.10
4.0	126.39	54.01	IBM **IBM**	.60	.7	29	76523	80.57	3.07
1.9	37.45	26.05	IntFlavor **IFF**	.60	1.7	21	5952	35.78	0.68
2.9	80.10	47.75♣	IntGameTch **IGT**		...	24	9427	78.15	2.23

Source: Wall Street Journal, January 3, 2003, p. C4.

The following information is included in each column. International Business Machines (IBM) common stock is used as an example.

YTD % Chg: The stock price percentage change for the calendar year to date, adjusted for stock splits and dividends over 10%

52 Weeks Hi: Highest price of a share in the past 52 weeks: 126.39 for IBM stock

52 Weeks Lo: Lowest price of a share in the past 52 weeks: 54.01 for IBM stock

Stock: Company name: IBM for International Business Machines

Sym: Symbol that identifies company: IBM

Div: Annual dividends: $0.60 for IBM

Yld %: Yield for stock expressed as annual dividends divided by today's closing price: 0.7% (= 0.6 ÷ 80.57) for IBM stock

PE: Price-earnings ratio; the stock price divided by the annual earnings per share: 29

Vol 100s: Number of shares (in hundreds) traded that day: 7,652,300 shares for IBM

Close: Closing price (last price) that day: 80.57

Net Chg: Change in the closing price from the previous day: 3.07

Prices quoted for shares traded over-the-counter (through dealers rather than on an organized exchange) are sometimes quoted with the same information, but in many cases only the bid price (the price the dealer is willing to pay for the stock) and the asked price (the price the dealer is willing to sell the stock for) are quoted.

The efficient market hypothesis tells us that when purchasing a security, we cannot expect to earn an abnormally high return, a return greater than the equilibrium return. Information in newspapers and in the published reports of investment advisers is readily available to many market participants and is already reflected in market prices. So acting on this information will not yield abnormally high returns, on average. As we have seen, the empirical evidence for the most part confirms that recommendations from investment advisers cannot help us outperform the general market. Indeed, as Box 1 suggests, human investment advisers in San Francisco do not on average even outperform an orangutan!

Probably no other conclusion is met with more skepticism by students than this one when they first hear it. We all know or have heard of somebody who has been successful in the stock market for a period of many years. We wonder, "How could someone be so consistently successful if he or she did not really know how to predict when returns would be abnormally high?"

Box 1

Should You Hire an Ape as Your Investment Adviser?

The *San Francisco Chronicle* came up with an amusing way of evaluating how successful investment advisers are at picking stocks. They asked eight analysts to pick five stocks at the beginning of the year and then compared the performance of their stock picks to those chosen by Jolyn, an orangutan living at Marine World/Africa USA in Vallejo, California. Consistent with the results found in the "Investment Dartboard" feature of the *Wall Street Journal*, Jolyn beat the investment advisers as often as they beat her. Given this result, you might be just as well off hiring an orangutan as your investment adviser as you would hiring a human being!

The following story, reported in the press, illustrates why such anecdotal evidence is not reliable.

A get-rich-quick artist invented a clever scam. Every week, he wrote two letters. In letter A, he would pick team A to win a particular football game, and in letter B, he would pick the opponent, team B. A mailing list would then be separated into two groups, and he would send letter A to the people in one group and letter B to the people in the other. The following week he would do the same thing but would send these letters only to the group who had received the first letter with the correct prediction. After doing this for ten games, he had a small cluster of people who had received letters predicting the correct winning team for every game. He then mailed a final letter to them, declaring that since he was obviously an expert predictor of the outcome of football games (he had picked winners ten weeks in a row) and since his predictions were profitable for the recipients who bet on the games, he would continue to send his predictions only if he were paid a substantial amount of money. When one of his clients figured out what he was up to, the con man was prosecuted and thrown in jail!

What is the lesson of the story? Even if no forecaster is an accurate predictor of the market, there will always be a group of consistent winners. A person who has done well regularly in the past cannot guarantee that he or she will do well in the future. Note that there will also be a group of persistent losers, but you rarely hear about them because no one brags about a poor forecasting record.

Should You Be Skeptical of Hot Tips?

Suppose your broker phones you with a hot tip to buy stock in the Happy Feet Corporation (HFC) because it has just developed a product that is completely effective in curing athlete's foot. The stock price is sure to go up. Should you follow this advice and buy HFC stock?

The efficient market hypothesis indicates that you should be skeptical of such news. If the stock market is efficient, it has already priced HFC stock so that its expected return will equal the equilibrium return. The hot tip is not particularly valuable and will not enable you to earn an abnormally high return.

You might wonder, though, if the hot tip is based on new information and would give you an edge on the rest of the market. If other market participants

have gotten this information before you, the answer is no. As soon as the information hits the street, the unexploited profit opportunity it creates will be quickly eliminated. The stock's price will already reflect the information, and you should expect to realize only the equilibrium return. But if you are one of the first to gain the new information, it can do you some good. Only then can you be one of the lucky ones who, on average, will earn an abnormally high return by helping eliminate the profit opportunity by buying HFC stock.

Do Stock Prices Always Rise When There Is Good News?

If you follow the stock market, you might have noticed a puzzling phenomenon: When good news about a stock, such as a particularly favorable earnings report, is announced, the price of the stock frequently does not rise. The efficient market hypothesis and the random-walk behavior of stock prices explain this phenomenon.

Because changes in stock prices are unpredictable, when information is announced that has already been expected by the market, the stock price will remain unchanged. The announcement does not contain any new information that should lead to a change in stock prices. If this were not the case and the announcement led to a change in stock prices, it would mean that the change was predictable. Because that is ruled out in an efficient market, ***stock prices will respond to announcements only when the information being announced is new and unexpected***. If the news is expected, there will be no stock price response. This is exactly what the evidence we described earlier, which shows that stock prices reflect publicly available information, suggests will occur.

Sometimes an individual stock price declines when good news is announced. Although this seems somewhat peculiar, it is completely consistent with the workings of an efficient market. Suppose that although the announced news is good, it is not as good as expected. HFC's earnings may have risen 15%, but if the market expected earnings to rise by 20%, the new information is actually unfavorable, and the stock price declines.

Efficient Market Prescription for the Investor

What does the efficient market hypothesis recommend for investing in the stock market? It tells us that hot tips, investment advisers' published recommendations, and technical analysis—all of which make use of publicly available information—cannot help an investor outperform the market. Indeed, it indicates that anyone without better information than other market participants cannot expect to beat the market. So what is an investor to do?

The efficient market hypothesis leads to the conclusion that such an investor (and almost all of us fit into this category) should not try to outguess the market by constantly buying and selling securities. This process does nothing but boost the income of brokers, who earn commissions on each trade.[20] Instead, the investor should pursue a "buy and hold" strategy—purchase stocks and hold them for long periods of time. This will lead to the same returns, on average, but the investor's net profits will be higher, because fewer brokerage commissions will have to be paid.

[20]The investor may also have to pay Uncle Sam capital gains taxes on any profits that are realized when a security is sold—an additional reason why continual buying and selling does not make sense.

It is frequently a sensible strategy for a small investor, whose costs of managing a portfolio may be high relative to its size, to buy into a mutual fund rather than individual stocks. Because the efficient market hypothesis indicates that no mutual fund can consistently outperform the market, an investor should not buy into one that has high management fees or that pays sales commissions to brokers, but rather should purchase a no-load (commission-free) mutual fund that has low management fees.

As we have seen, the evidence indicates that it will not be easy to beat the prescription suggested here, although some of the anomalies to the efficient market hypothesis suggest that an extremely clever investor (which rules out most of us) may be able to outperform a buy-and-hold strategy.

Evidence on Rational Expectations in Other Markets

Evidence in other financial markets also supports the efficient market hypothesis and hence the rationality of expectations. For example, there is little evidence that financial analysts are able to outperform the bond market.[21] The returns on bonds appear to conform to the efficient markets condition of Equation 10.

Rationality of expectations is, however, much harder to test in markets other than financial markets, because price data that reflect expectations are not as readily available. The most common tests of rational expectations in these markets make use of survey data on the forecasts of market participants. For example, one well-known study by James Pesando used a survey of inflation expectations collected from prominent economists and inflation forecasters.[22] In that survey, these people were asked what they predicted the inflation rate would be over the next six months and over the next year. Because rational expectations theory implies that forecast errors should on average be zero and cannot be predicted, tests of the theory involve asking whether the forecast errors in a survey could be predicted ahead of time using publicly available information. The evidence from Pesando's and subsequent studies is mixed. Sometimes the forecast errors cannot be predicted, and at other times they can. The evidence is not as supportive of rational expectations theory as the evidence from financial markets.

Does the fact that forecast errors from surveys are often predictable suggest that we should reject rational expectations theory in these other markets? The answer is: not necessarily. One problem with this evidence is that the expectations data are obtained from surveys rather than from actual economic decisions of market participants. That is a serious criticism of this evidence. Survey responses are not always reliable, because there is little incentive for participants to tell the truth. For example, when people are asked in surveys how much television they watch, responses greatly underestimate the actual time spent. Neither are people very truthful about the shows

[21] See the discussion in Frederic S. Mishkin, "Efficient Markets Theory: Implications for Monetary Policy," *Brookings Papers on Economic Activity* 3 (1978): 707–768, of the results in Michael J. Prell, "How Well Do the Experts Forecast Interest Rates?" *Federal Reserve Bank of Kansas City Monthly Review*, September–October 1973, pp. 3–15.

[22] James Pesando, "A Note on the Rationality of the Livingston Price Expectations," *Journal of Political Economy* 83 (1975): 845–858.

they watch. They may say they watch ballet on public television, but we know they are actually watching Vanna White light up the letters on *Wheel of Fortune* instead, because it, not ballet, gets high Nielsen ratings. How many people will admit to being regular watchers of *Wheel of Fortune*?

A second problem with survey evidence is that a market's behavior may not be equally influenced by the expectations of all the survey participants, making survey evidence a poor guide to market behavior. For example, we have already seen that prices in financial markets often *behave* as if expectations are rational even though many of the market participants do not have rational expectations.[23]

Proof is not yet conclusive on the validity of rational expectations theory in markets other than financial markets. One important conclusion, however, that is supported by the survey evidence is that ***if there is a change in the way a variable moves, there will be a change in the way expectations of this variable are formed as well***.

Application

What Do the Black Monday Crash of 1987 and the Tech Crash of 2000 Tell Us About Rational Expectations and Efficient Markets?

On October 19, 1987, dubbed "Black Monday," the Dow Jones Industrial Average declined more than 20%, the largest one-day decline in U.S. history. The collapse of the high-tech companies' share prices from their peaks in March 2000 caused the heavily tech-laden NASDAQ index to fall from around 5,000 in March 2000 to around 1,500 in 2001 and 2002, for a decline of well over 60%. These two crashes have caused many economists to question the validity of efficient markets and rational expectations. They do not believe that a rational marketplace could have produced such a massive swing in share prices. To what degree should these stock market crashes make us doubt the validity of rational expectations and the efficient market hypothesis?

Nothing in rational expectations theory rules out large changes in stock prices. A large change in stock prices can result from new information that produces a dramatic decline in optimal forecasts of the future valuation of firms. However, economists are hard pressed to come up with fundamental changes in the economy that can explain the Black Monday and tech crashes. One lesson from these crashes is that factors other than market fundamentals probably have an effect on stock prices. Hence these crashes have convinced many economists that the stronger version of the efficient market hypothesis, which states that asset prices reflect the true fundamental (intrinsic) value of securities, is incorrect. They attribute a large role in determination of stock prices to market psychology and to the institutional structure of the marketplace. However, nothing in this view contradicts the basic reasoning behind rational expectations or the efficient market hypothesis—that market participants eliminate unexploited profit opportunities. Even though stock market prices may not always solely reflect

[23]There is some fairly strong evidence for this proposition. For example, Frederic S. Mishkin, "Are Market Forecasts Rational?" *American Economic Review* 71 (1981): 295–306, finds that although survey forecasts of short-term interest rates are not rational, the bond market *behaves* as if the expectations of these interest rates are rational.

market fundamentals, this does not mean that rational expectations do not hold. As long as stock market crashes are unpredictable, the basic lessons of the theory of rational expectations hold.

Some economists have come up with theories of what they call *rational bubbles* to explain stock market crashes. A **bubble** is a situation in which the price of an asset differs from its fundamental market value. In a rational bubble, investors can have rational expectations that a bubble is occurring because the asset price is above its fundamental value but continue to hold the asset anyway. They might do this because they believe that someone else will buy the asset for a higher price in the future. In a rational bubble, asset prices can therefore deviate from their fundamental value for a long time because the bursting of the bubble cannot be predicted and so there are no unexploited profit opportunities.

However, other economists believe that the Black Monday crash of 1987 and the tech crash of 2000 suggest that there may be unexploited profit opportunities and that the theory of rational expectations and the efficient market hypothesis might be fundamentally flawed. The controversy over whether capital markets are efficient or expectations are rational continues.

Summary

1. Stocks are valued as the present value of future dividends. Unfortunately, we do not know very precisely what these dividends will be. This introduces a great deal of error to the valuation process. The Gordon growth model is a simplified method of computing stock value that depends on the assumption that the dividends are growing at a constant rate forever. Given our uncertainty regarding future dividends, this assumption is often the best we can do.
2. The interaction among traders in the market is what actually sets prices on a day-to-day basis. The trader that values the security the most (either because of less uncertainty about the cash flows or because of greater estimated cash flows) will be willing to pay the most. As new information is released, investors will revise their estimates of the true value of the security and will either buy or sell it depending upon how the market price compares to their estimated valuation. Because small changes in estimated growth rates or required return result in large changes in price, it is not surprising that the markets are often volatile.
3. The efficient market hypothesis states that current security prices will fully reflect all available information, because in an efficient market, all unexploited profit opportunities are eliminated. The elimination of unexploited profit opportunities necessary for a financial market to be efficient does not require that all market participants be well informed.
4. The evidence on the efficient market hypothesis is quite mixed. Early evidence on the performance of investment analysts and mutual funds, whether stock prices reflect publicly available information, the random-walk behavior of stock prices, and the success of so-called technical analysis was quite favorable to the efficient market hypothesis. However, in recent years, evidence on the small-firm effect, the January effect, market overreaction, excessive volatility, mean reversion, and new information is not always incorporated into stock prices, suggesting that the hypothesis may not always be entirely correct. The evidence seems to suggest that the efficient market hypothesis may be a reasonable starting point for evaluating behavior in financial markets but may not be generalizable to all behavior in financial markets.
5. The efficient market hypothesis indicates that hot tips, investment advisers' published recommendations, and technical analysis cannot help an investor out-perform the market. The prescription for investors is to pursue a

buy-and-hold strategy—purchase stocks and hold them for long periods of time. Empirical evidence generally supports these implications of the efficient market hypothesis in the stock market.

6. The stock market crash of 1987 and the tech crash of 2000 have convinced many financial economists that the stronger version of the efficient market hypothesis, which states that asset prices reflect the true fundamental (intrinsic) value of securities, is not correct. It is less clear that these crashes shows that the weaker version of the efficient market hypothesis is wrong. Even if the stock market was driven by factors other than fundamentals, these crashes do not clearly demonstrate that many of the basic lessons of the efficient market hypothesis are no longer valid, as long as these crashes could not have been predicted.

Key Terms

adaptive expectations, p. 147
bubble, p. 164
cash flows, p. 141
dividends, p. 142
efficient market hypothesis, p. 149
generalized dividend model, p. 143
Gordon growth model, p. 143
January effect, p. 156
market fundamentals, p. 152
mean reversion, p. 157
optimal forecast, p. 148
random walk, p. 154
rational expectations, p. 147
residual claimant, p. 141
stockholders, p. 141
theory of efficient capital markets, p. 149
unexploited profit opportunity, p. 152

Questions and Problems

Questions marked with an asterisk are answered at the end of the book in an appendix, "Answers to Selected Questions and Problems."

1. What basic principle of finance can be applied to the valuation of any investment asset?

*2. Identify the cash flows available to an investor in stock. How reliably can these cash flows be estimated? Compare the problem of estimating stock cash flows to estimating bond cash flows. Which security would you predict to be more volatile?

3. Compute the price of a share of stock that pays a $1 per year dividend and that you expect to be able to sell in one year for $20, assuming you require a 15% return.

*4. After careful analysis, you have determined that a firm's dividends should grow at 7% on average in the foreseeable future. Its last dividend was $3. Compute the current price of this stock, assuming the required return is 18%.

5. Some economists think that the central banks should try to prick bubbles in the stock market before they get out of hand and cause later damage when they burst. How can monetary policy be used to prick a bubble? Explain how it can do this using the Gordon growth model.

*6. "Forecasters' predictions of inflation are notoriously inaccurate, so their expectations of inflation cannot be rational." Is this statement true, false, or uncertain? Explain your answer.

7. "Whenever it is snowing when Joe Commuter gets up in the morning, he misjudges how long it will take him to drive to work. Otherwise, his expectations of the driving time are perfectly accurate. Considering that it snows only once every ten years where Joe lives, Joe's expectations are almost always perfectly accurate." Are Joe's expectations rational? Why or why not?

*8. If a forecaster spends hours every day studying data to forecast interest rates but his expectations are not as accurate as predicting that tomorrow's interest rates will be identical to today's interest rate, are his expectations rational?

9. "If stock prices did not follow a random walk, there would be unexploited profit opportunities in the market." Is this statement true, false, or uncertain? Explain your answer.

*10. Suppose that increases in the money supply lead to a rise in stock prices. Does this mean that when you see that the money supply has had a sharp rise in the past week, you should go out and buy stocks? Why or why not?

11. If the public expects a corporation to lose $5 a share this quarter and it actually loses $4, which is still the largest loss in the history of the company, what does the efficient market hypothesis say will happen to the price of the stock when the $4 loss is announced?

*12. If I read in the *Wall Street Journal* that the "smart money" on Wall Street expects stock prices to fall, should I follow that lead and sell all my stocks?

13. If my broker has been right in her five previous buy and sell recommendations, should I continue listening to her advice?

*14. Can a person with rational expectations expect the price of IBM to rise by 10% in the next month?

15. "If most participants in the stock market do not follow what is happening to the monetary aggregates, prices of common stocks will not fully reflect information about them." Is this statement true, false, or uncertain? Explain your answer.

*16. "An efficient market is one in which no one ever profits from having better information than the rest." Is this statement true, false, or uncertain? Explain your answer.

17. If higher money growth is associated with higher future inflation and if announced money growth turns out to be extremely high but is still less than the market expected, what do you think would happen to long-term bond prices?

*18. "Foreign exchange rates, like stock prices, should follow a random walk." Is this statement true, false, or uncertain? Explain your answer.

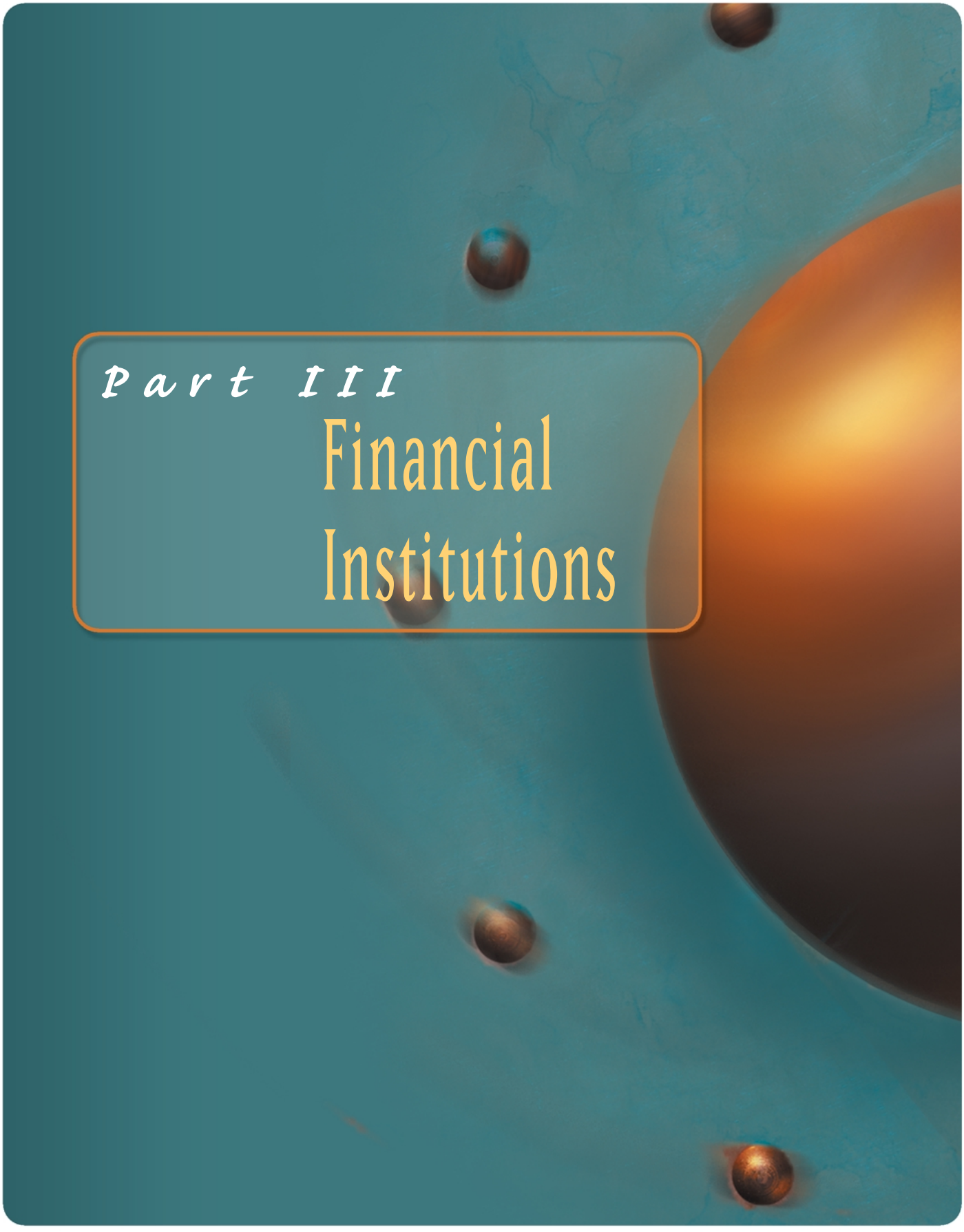
19. Can we expect the value of the dollar to rise by 2% next week if our expectations are rational?

*20. "Human fear is the source of stock market crashes, so these crashes indicate that expectations in the stock market cannot be rational." Is this statement true, false, or uncertain? Explain your answer.

Web Exercises

1. Visit www.forecasts.org/data/index.htm. Click on "Stock Index" at the very top of the page. Now choose "U.S. Stock Indices-monthly." Review the indices for the DJIA, the S&P 500, and the NASDAQ composite. Which index appears most volatile? In which index would you have rather invested in 1985 if the investment had been allowed to compound until now?

2. The Internet is a great source of information on stock prices and stock price movements. There are many sites that provide up-to-the minute data on stock market indices. One of the best is found at http://finance.lycos.com/home/livecharts. This site provides free real-time streaming of stock market data. Click on the $indu to have the chart display the Dow Jones Industrial Average. Look at the stock trend over various intervals by adjusting the update frequency (click on "INT" at the top of the chart). Have stock prices been going up or down over the last day, week, month, and year?


Part III
Financial Institutions

Chapter

8 An Economic Analysis of Financial Structure

PREVIEW

A healthy and vibrant economy requires a financial system that moves funds from people who save to people who have productive investment opportunities. But how does the financial system make sure that your hard-earned savings get channeled to Paula the Productive Investor rather than to Benny the Bum?

This chapter answers that question by providing an economic analysis of how our financial structure is designed to promote economic efficiency. The analysis focuses on a few simple but powerful economic concepts that enable us to explain features of our financial system, such as why financial contracts are written as they are and why financial intermediaries are more important than securities markets for getting funds to borrowers. The analysis also demonstrates the important link between the financial system and the performance of the aggregate economy, which is the subject of Part V of the book. The economic analysis of financial structure explains how the performance of the financial sector affects economic growth and why financial crises occur and have such severe consequences for aggregate economic activity.

Basic Puzzles About Financial Structure Throughout the World

The financial system is complex in structure and function throughout the world. It includes many different types of institutions: banks, insurance companies, mutual funds, stock and bond markets, and so on—all of which are regulated by government. The financial system channels billions of dollars per year from savers to people with productive investment opportunities. If we take a close look at financial structure all over the world, we find eight basic puzzles that we need to solve in order to understand how the financial system works.

The pie chart in Figure 1 indicates how American businesses financed their activities using external funds (those obtained from outside the business itself) in the period 1970–1996. The *Bank Loans* category is made up primarily of bank loans; *Nonbank Loans* is composed primarily of loans by other financial intermediaries; the *Bonds* category includes marketable debt securities such as corporate bonds and commercial paper; and *Stock* consists of new issues of new equity (stock market shares). Figure 2 uses the same classifications as Figure 1 and compares the U.S. data to those of Germany and Japan.

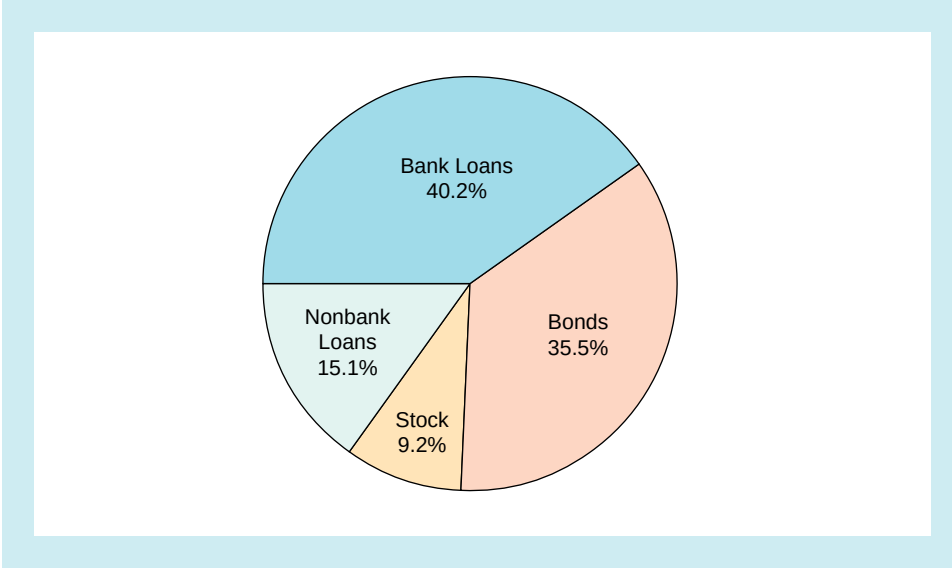


FIGURE 1 Sources of External Funds for Nonfinancial Businesses in the United States

Source: Reinhard H. Schmidt, "Differences Between Financial Systems in European Countries: Consequences for EMU," in Deutsche Bundesbank, ed., *The Monetary Transmission Process: Recent Developments and Lessons for Europe* (Hampshire: Palgrave Publishers, 2001), p. 222.

Now let us explore the eight puzzles.

1. ***Stocks are not the most important source of external financing for businesses.*** Because so much attention in the media is focused on the stock market, many people have the impression that stocks are the most important sources of financing for American corporations. However, as we can see from the pie chart in Figure 1, the stock market accounted for only a small fraction of the external financing of American businesses in the 1970–1996 period: 9.2%.[1] (In fact, in the mid- to late 1980s, American corporations generally stopped issuing shares to finance their activities; instead they purchased large numbers of shares, meaning that the stock market was actually a *negative* source of corporate finance in those years.) Similarly small figures apply in the other countries presented in Figure 2 as well. Why is the stock market less important than other sources of financing in the United States and other countries?

2. ***Issuing marketable debt and equity securities is not the primary way in which businesses finance their operations.*** Figure 1 shows that bonds are a far more important source of financing than stocks in the United States (35.5% versus 9.2%). However, stocks and bonds combined (44.7%), which make up the total share of marketable securities, still supply less than one-half of the external funds corporations need to finance their activities. The fact that issuing marketable securities is not the most important source of financing is true elsewhere in the world as well. Indeed, as

[1]The 9.2% figure for the percentage of external financing provided by stocks is based on the flows of external funds to corporations. However, this flow figure is somewhat misleading, because when a share of stock is issued, it raises funds permanently; whereas when a bond is issued, it raises funds only temporarily until they are paid back at maturity. To see this, suppose that a firm raises $1,000 by selling a share of stock and another $1,000 by selling a $1,000 one-year bond. In the case of the stock issue, the firm can hold on to the $1,000 it raised this way, but to hold on to the $1,000 it raised through debt, it has to issue a new $1,000 bond every year. If we look at the flow of funds to corporations over a 26-year period, as in Figure 1, the firm will have raised $1,000 with a stock issue only once in the 26-year period, while it will have raised $1,000 with debt 26 times, once in each of the 26 years. Thus it will look as though debt is 26 times more important than stocks in raising funds, even though our example indicates that they are actually equally important for the firm.

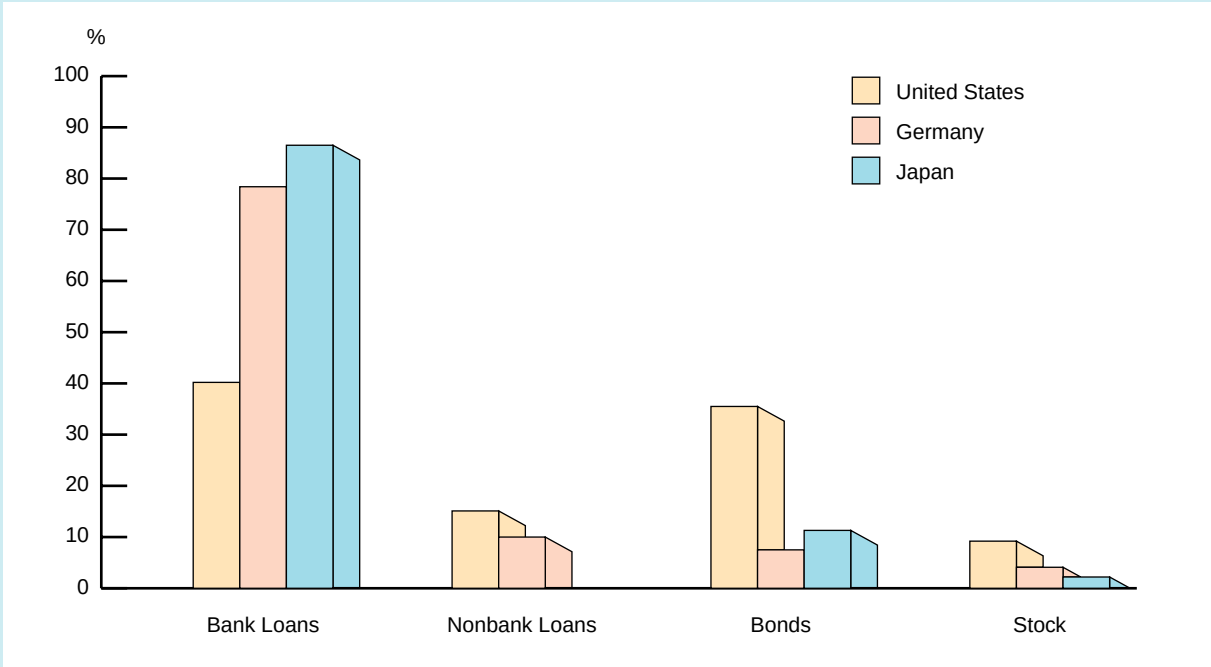


FIGURE 2 Sources of External Funds for Nonfinancial Businesses: A Comparison of the United States with Germany and Japan
The categories of external funds are the same as in Figure 1 and the data are for the period 1970–1996.

Source: Reinhard H. Schmidt, "Differences Between Financial Systems in European Countries: Consequences for EMU," in Deutsche Bundesbank, ed., *The Monetary Transmission Process: Recent Developments and Lessons for Europe* (Hampshire: Palgrave Publishers, 2001), p. 222.

we see in Figure 2, other countries have a much smaller share of external financing supplied by marketable securities than the United States. Why don't businesses use marketable securities more extensively to finance their activities?

3. ***Indirect finance, which involves the activities of financial intermediaries, is many times more important than direct finance, in which businesses raise funds directly from lenders in financial markets.*** Direct finance involves the sale to households of marketable securities such as stocks and bonds. The 44.7% share of stocks and bonds as a source of external financing for American businesses actually greatly overstates the importance of direct finance in our financial system. Since 1970, less than 5% of newly issued corporate bonds and commercial paper and around 50% of stocks have been sold directly to American households. The rest of these securities have been bought primarily by financial intermediaries such as insurance companies, pension funds, and mutual funds. These figures indicate that direct finance is used in less than 10% of the external funding of American business. Because in most countries marketable securities are an even less important source of finance than in the United States, direct finance is also far less important than indirect finance in the rest of the world. Why are financial intermediaries and indirect finance so important in financial markets? In recent years, indirect finance has been declining in importance. Why is this happening?

4. ***Banks are the most important source of external funds used to finance businesses.*** As we can see in Figures 1 and 2, the primary sources of external funds for

businesses throughout the world are loans (55.3% in the United States). Most of these loans are bank loans, so the data suggest that banks have the most important role in financing business activities. An extraordinary fact that surprises most people is that in an average year in the United States, more than four times more funds are raised with bank loans than with stocks. Banks are even more important in countries such as Germany and Japan than they are in the United States, and in developing countries banks play an even more important role in the financial system than they do in the industrialized countries. What makes banks so important to the workings of the financial system? Although banks remain important, their share of external funds for businesses has been declining in recent years. What is driving their decline?

5. ***The financial system is among the most heavily regulated sectors of the economy.*** You learned in Chapter 2 that the financial system is heavily regulated, not only in the United States but in all other developed countries as well. Governments regulate financial markets primarily to promote the provision of information, in part, to protect consumers, and to ensure the soundness (stability) of the financial system. Why are financial markets so extensively regulated throughout the world?

6. ***Only large, well-established corporations have easy access to securities markets to finance their activities.*** Individuals and smaller businesses that are not well established are less likely to raise funds by issuing marketable securities. Instead, they most often obtain their financing from banks. Why do only large, well-known corporations find it easier to raise funds in securities markets?

7. ***Collateral is a prevalent feature of debt contracts for both households and businesses.*** **Collateral** is property that is pledged to the lender to guarantee payment in the event that the borrower is unable to make debt payments. Collateralized debt (also known as **secured debt** to contrast it with **unsecured debt**, such as credit card debt, which is not collateralized) is the predominant form of household debt and is widely used in business borrowing as well. The majority of household debt in the United States consists of collateralized loans: Your automobile is collateral for your auto loan, and your house is collateral for your mortgage. Commercial and farm mortgages, for which property is pledged as collateral, make up one-quarter of borrowing by nonfinancial businesses; corporate bonds and other bank loans also often involve pledges of collateral. Why is collateral such an important feature of debt contracts?

8. ***Debt contracts typically are extremely complicated legal documents that place substantial restrictions on the behavior of the borrower.*** Many students think of a debt contract as a simple IOU that can be written on a single piece of paper. The reality of debt contracts is far different, however. In all countries, bond or loan contracts typically are long legal documents with provisions (called **restrictive covenants**) that restrict and specify certain activities that the borrower can engage in. Restrictive covenants are not just a feature of debt contracts for businesses; for example, personal automobile loan and home mortgage contracts have covenants that require the borrower to maintain sufficient insurance on the automobile or house purchased with the loan. Why are debt contracts so complex and restrictive?

As you may recall from Chapter 2, an important feature of financial markets is that they have substantial transaction and information costs. An economic analysis of how these costs affect financial markets provides us with solutions to the eight puzzles, which in turn provide us with a much deeper understanding of how our financial system works. In the next section, we examine the impact of transaction costs on the structure of our financial system. Then we turn to the effect of information costs on financial structure.

Transaction Costs

Transaction costs are a major problem in financial markets. An example will make this clear.

How Transaction Costs Influence Financial Structure

Say you have $5,000 you would like to invest, and you think about investing in the stock market. Because you have only $5,000, you can buy only a small number of shares. The stockbroker tells you that your purchase is so small that the brokerage commission for buying the stock you picked will be a large percentage of the purchase price of the shares. If instead you decide to buy a bond, the problem is even worse, because the smallest denomination for some bonds you might want to buy is as much as $10,000, and you do not have that much to invest. Indeed, the broker may not be interested in your business at all, because the small size of your account doesn't make spending time on it worthwhile. You are disappointed and realize that you will not be able to use financial markets to earn a return on your hard-earned savings. You can take some consolation, however, in the fact that you are not alone in being stymied by high transaction costs. This is a fact of life for many of us: Only around one-half of American households own any securities.

You also face another problem because of transaction costs. Because you have only a small amount of funds available, you can make only a restricted number of investments. That is, you have to put all your eggs in one basket, and your inability to diversify will subject you to a lot of risk.

How Financial Intermediaries Reduce Transaction Costs

This example of the problems posed by transaction costs and the example outlined in Chapter 2 when legal costs kept you from making a loan to Carl the Carpenter illustrate that small savers like you are frozen out of financial markets and are unable to benefit from them. Fortunately, financial intermediaries, an important part of the financial structure, have evolved to reduce transaction costs and allow small savers and borrowers to benefit from the existence of financial markets.

Economies of Scale. One solution to the problem of high transaction costs is to bundle the funds of many investors together so that they can take advantage of *economies of scale*, the reduction in transaction costs per dollar of investment as the size (scale) of transactions increases. By bundling investors' funds together, transaction costs for each individual investor are far smaller. Economies of scale exist because the total cost of carrying out a transaction in financial markets increases only a little as the size of the transaction grows. For example, the cost of arranging a purchase of 10,000 shares of stock is not much greater than the cost of arranging a purchase of 50 shares of stock.

The presence of economies of scale in financial markets helps explain why financial intermediaries developed and have become such an important part of our financial structure. The clearest example of a financial intermediary that arose because of economies of scale is a mutual fund. A *mutual fund* is a financial intermediary that sells shares to individuals and then invests the proceeds in bonds or stocks. Because it buys large blocks of stocks or bonds, a mutual fund can take advantage of lower transaction costs. These cost savings are then passed on to individual investors after the mutual fund has taken its cut in the form of management fees for administering their accounts. An additional benefit for individual investors is that a mutual fund is large enough to purchase a widely diversified portfolio of securities. The increased diversification for individual investors reduces their risk, making them better off.

Economies of scale are also important in lowering the costs of things such as computer technology that financial institutions need to accomplish their tasks. Once a large mutual fund has invested a lot of money in setting up a telecommunications system, for example, the system can be used for a huge number of transactions at a low cost per transaction.

Expertise. Financial intermediaries are also better able to develop expertise to lower transaction costs. Their expertise in computer technology enables them to offer customers convenient services like being able to call a toll-free number for information on how well their investments are doing and to write checks on their accounts.

An important outcome of a financial intermediary's low transaction costs is the ability to provide its customers with *liquidity services*, services that make it easier for customers to conduct transactions. Money market mutual funds, for example, not only pay shareholders high interest rates, but also allow them to write checks for convenient bill-paying.

Asymmetric Information: Adverse Selection and Moral Hazard

The presence of transaction costs in financial markets explains in part why financial intermediaries and indirect finance play such an important role in financial markets (puzzle 3). To understand financial structure more fully, however, we turn to the role of information in financial markets.[2]

Asymmetric information—one party's insufficient knowledge about the other party involved in a transaction to make accurate decisions—is an important aspect of financial markets. For example, managers of a corporation know whether they are honest or have better information about how well their business is doing than the stockholders do. The presence of asymmetric information leads to adverse selection and moral hazard problems, which were introduced in Chapter 2.

Adverse selection is an asymmetric information problem that occurs *before* the transaction occurs: Potential bad credit risks are the ones who most actively seek out loans. Thus the parties who are the most likely to produce an undesirable outcome are the ones most likely to want to engage in the transaction. For example, big risk takers or outright crooks might be the most eager to take out a loan because they know that they are unlikely to pay it back. Because adverse selection increases the chances that a loan might be made to a bad credit risk, lenders might decide not to make any loans, even though there are good credit risks in the marketplace.

Moral hazard arises *after* the transaction occurs: The lender runs the risk that the borrower will engage in activities that are undesirable from the lender's point of view because they make it less likely that the loan will be paid back. For example, once borrowers have obtained a loan, they may take on big risks (which have possible high returns but also run a greater risk of default) because they are playing with someone else's money. Because moral hazard lowers the probability that the loan will be repaid, lenders may decide that they would rather not make a loan.

[2]An excellent survey of the literature on information and financial structure that expands on the topics discussed in the rest of this chapter is contained in Mark Gertler, "Financial Structure and Aggregate Economic Activity: An Overview," *Journal of Money, Credit and Banking* 20 (1988): 559–588.

The analysis of how asymmetric information problems affect economic behavior is called **agency theory**. We will apply this theory here to explain why financial structure takes the form it does, thereby solving the puzzles described at the beginning of the chapter.

The Lemons Problem: How Adverse Selection Influences Financial Structure

www.nobel.se/economics/laureates/2001/public.html
A complete discussion of the lemons problem on a site dedicated to Nobel prize winners.

A particular characterization of the adverse selection problem and how it interferes with the efficient functioning of a market was outlined in a famous article by Nobel prize winner George Akerlof. It is referred to as the "lemons problem," because it resembles the problem created by lemons in the used-car market.[3] Potential buyers of used cars are frequently unable to assess the quality of the car; that is, they can't tell whether a particular used car is a good car that will run well or a lemon that will continually give them grief. The price that a buyer pays must therefore reflect the *average* quality of the cars in the market, somewhere between the low value of a lemon and the high value of a good car.

The owner of a used car, by contrast, is more likely to know whether the car is a peach or a lemon. If the car is a lemon, the owner is more than happy to sell it at the price the buyer is willing to pay, which, being somewhere between the value of a lemon and a good car, is greater than the lemon's value. However, if the car is a peach, the owner knows that the car is undervalued by the price the buyer is willing to pay, and so the owner may not want to sell it. As a result of this adverse selection, very few good used cars will come to the market. Because the average quality of a used car available in the market will be low and because very few people want to buy a lemon, there will be few sales. The used-car market will then function poorly, if at all.

Lemons in the Stock and Bond Markets

A similar lemons problem arises in securities markets, that is, the debt (bond) and equity (stock) markets. Suppose that our friend Irving the Investor, a potential buyer of securities such as common stock, can't distinguish between good firms with high expected profits and low risk and bad firms with low expected profits and high risk. In this situation, Irving will be willing to pay only a price that reflects the *average* quality of firms issuing securities—a price that lies between the value of securities from bad firms and the value of those from good firms. If the owners or managers of a good firm have better information than Irving and *know* that they are a good firm, they know that their securities are undervalued and will not want to sell them to Irving at the price he is willing to pay. The only firms willing to sell Irving securities will be bad firms (because the price is higher than the securities are worth). Our friend Irving is not stupid; he does not want to hold securities in bad firms, and hence he will decide not to purchase securities in the market. In an outcome similar to that

[3]George Akerlof, "The Market for 'Lemons': Quality, Uncertainty and the Market Mechanism," *Quarterly Journal of Economics* 84 (1970): 488–500. Two important papers that have applied the lemons problem analysis to financial markets are Stewart Myers and N. S. Majluf, "Corporate Financing and Investment Decisions When Firms Have Information That Investors Do Not Have," *Journal of Financial Economics* 13 (1984): 187–221, and Bruce Greenwald, Joseph E. Stiglitz, and Andrew Weiss, "Information Imperfections in the Capital Market and Macroeconomic Fluctuations," *American Economic Review* 74 (1984): 194–199.

in the used-car market, this securities market will not work very well because few firms will sell securities in it to raise capital.

The analysis is similar if Irving considers purchasing a corporate debt instrument in the bond market rather than an equity share. Irving will buy a bond only if its interest rate is high enough to compensate him for the average default risk of the good and bad firms trying to sell the debt. The knowledgeable owners of a good firm realize that they will be paying a higher interest rate than they should, and so they are unlikely to want to borrow in this market. Only the bad firms will be willing to borrow, and because investors like Irving are not eager to buy bonds issued by bad firms, they will probably not buy any bonds at all. Few bonds are likely to sell in this market, and so it will not be a good source of financing.

The analysis we have just conducted explains puzzle 2—why marketable securities are not the primary source of financing for businesses in any country in the world. It also partly explains puzzle 1—why stocks are not the most important source of financing for American businesses. The presence of the lemons problem keeps securities markets such as the stock and bond markets from being effective in channeling funds from savers to borrowers.

Tools to Help Solve Adverse Selection Problems

In the absence of asymmetric information, the lemons problem goes away. If buyers know as much about the quality of used cars as sellers, so that all involved can tell a good car from a bad one, buyers will be willing to pay full value for good used cars. Because the owners of good used cars can now get a fair price, they will be willing to sell them in the market. The market will have many transactions and will do its intended job of channeling good cars to people who want them.

Similarly, if purchasers of securities can distinguish good firms from bad, they will pay the full value of securities issued by good firms, and good firms will sell their securities in the market. The securities market will then be able to move funds to the good firms that have the most productive investment opportunities.

Private Production and Sale of Information. The solution to the adverse selection problem in financial markets is to eliminate asymmetric information by furnishing people supplying funds with full details about the individuals or firms seeking to finance their investment activities. One way to get this material to saver-lenders is to have private companies collect and produce information that distinguishes good from bad firms and then sell it. In the United States, companies such as Standard and Poor's, Moody's, and Value Line gather information on firms' balance sheet positions and investment activities, publish these data, and sell them to subscribers (individuals, libraries, and financial intermediaries involved in purchasing securities).

The system of private production and sale of information does not completely solve the adverse selection problem in securities markets, however, because of the so-called **free-rider problem**. The free-rider problem occurs when people who do not pay for information take advantage of the information that other people have paid for. The free-rider problem suggests that the private sale of information will be only a partial solution to the lemons problem. To see why, suppose that you have just purchased information that tells you which firms are good and which are bad. You believe that this purchase is worthwhile because you can make up the cost of acquiring this information, and then some, by purchasing the securities of good firms that are undervalued. However, when our savvy (free-riding) investor Irving sees you buying certain securities, he buys right along with you, even though he has not paid for any infor-

mation. If many other investors act as Irving does, the increased demand for the undervalued good securities will cause their low price to be bid up immediately to reflect the securities' true value. Because of all these free riders, you can no longer buy the securities for less than their true value. Now because you will not gain any profits from purchasing the information, you realize that you never should have paid for this information in the first place. If other investors come to the same realization, private firms and individuals may not be able to sell enough of this information to make it worth their while to gather and produce it. The weakened ability of private firms to profit from selling information will mean that less information is produced in the marketplace, and so adverse selection (the lemons problem) will still interfere with the efficient functioning of securities markets.

Government Regulation to Increase Information. The free-rider problem prevents the private market from producing enough information to eliminate all the asymmetric information that leads to adverse selection. Could financial markets benefit from government intervention? The government could, for instance, produce information to help investors distinguish good from bad firms and provide it to the public free of charge. This solution, however, would involve the government in releasing negative information about firms, a practice that might be politically difficult. A second possibility (and one followed by the United States and most governments throughout the world) is for the government to regulate securities markets in a way that encourages firms to reveal honest information about themselves so that investors can determine how good or bad the firms are. In the United States, the Securities and Exchange Commission (SEC) is the government agency that requires firms selling their securities in public markets to adhere to standard accounting principles and to disclose information about their sales, assets, and earnings. Similar regulations are found in other countries. However, disclosure requirements do not always work well, as the recent collapse of Enron and accounting scandals at other corporations (WorldCom, etc.) suggest (Box 1).

The asymmetric information problem of adverse selection in financial markets helps explain why financial markets are among the most heavily regulated sectors in the economy (puzzle 5). Government regulation to increase information for investors is needed to reduce the adverse selection problem, which interferes with the efficient functioning of securities (stock and bond) markets.

Although government regulation lessens the adverse selection problem, it does not eliminate it. Even when firms provide information to the public about their sales, assets, or earnings, they still have more information than investors: There is a lot more to knowing the quality of a firm than statistics can provide. Furthermore, bad firms have an incentive to make themselves look like good firms, because this would enable them to fetch a higher price for their securities. Bad firms will slant the information they are required to transmit to the public, thus making it harder for investors to sort out the good firms from the bad.

Financial Intermediation. So far we have seen that private production of information and government regulation to encourage provision of information lessen, but do not eliminate, the adverse selection problem in financial markets. How, then, can the financial structure help promote the flow of funds to people with productive investment opportunities when there is asymmetric information? A clue is provided by the structure of the used-car market.

Box 1

The Enron Implosion and the Arthur Andersen Conviction

Until 2001, Enron Corporation, a firm that specialized in trading in the energy market, appeared to be spectacularly successful. It had a quarter of the energy-trading market and was valued as high as $77 billion in August 2000 (just a little over a year before its collapse), making it the seventh-largest corporation in the United States at that time. However, toward the end of 2001, Enron came crashing down. In October 2001, Enron announced a third-quarter loss of $618 million and disclosed accounting "mistakes." The SEC then engaged in a formal investigation of Enron's financial dealings with partnerships led by its former finance chief. It became clear that Enron was engaged in a complex set of transactions by which it was keeping substantial amounts of debt and financial contracts off of its balance sheet. These transactions enabled Enron to hide its financial difficulties. Despite securing as much as $1.5 billion of new financing from J. P. Morgan Chase and Citigroup, the company was forced to declare bankruptcy in December 2001, the largest bankruptcy in U.S. history up to then.

Arthur Andersen, Enron's accounting firm, and one of the so-called Big Five accounting firms, was then indicted and finally convicted in June 2002 for obstruction of justice for impeding the SEC's investigation of the Enron collapse. This conviction—the first ever against a major accounting firm—meant that Andersen could no longer conduct audits of publicly traded firms, a development leading to its demise.

Enron's incredibly rapid collapse, combined with revelations of faulty accounting information from other publicly traded firms (e.g., WorldCom, which overstated its earnings by nearly $4 billion in 2001 and 2002), has raised concerns that disclosure and accounting regulations may be inadequate for firms that are involved in complicated financial transactions, and that accounting firms may not have the proper incentives to make sure that the accounting numbers are accurate. The scandals at Enron, Arthur Andersen, and other corporations resulted in the passage of legislation that is intended to make future Enrons less likely. The law established an independent oversight board for the accounting profession, prohibited auditors from offering certain consulting services to their clients, increased criminal penalties for corporate fraud, and required corporate chief executive officers and chief financial officers to certify financial reports.

The Enron collapse illustrates that government regulation can lessen asymmetric information problems, but cannot eliminate them. Managers have tremendous incentives to hide their companies' problems, making it hard for investors to know the true value of the firm.

An important feature of the used-car market is that most used cars are not sold directly by one individual to another. An individual considering buying a used car might pay for privately produced information by subscribing to a magazine like *Consumer Reports* to find out if a particular make of car has a good repair record. Nevertheless, reading *Consumer Reports* does not solve the adverse selection problem, because even if a particular make of car has a good reputation, the specific car someone is trying to sell could be a lemon. The prospective buyer might also bring the used car to a mechanic for a once-over. But what if the prospective buyer doesn't know a mechanic who can be trusted or if the mechanic would charge a high fee to evaluate the car?

Because these roadblocks make it hard for individuals to acquire enough information about used cars, most used cars are not sold directly by one individual to another. Instead, they are sold by an intermediary, a used-car dealer who purchases used cars from individuals and resells them to other individuals. Used-car dealers produce information in the market by becoming experts in determining whether a car is a peach or a lemon. Once they know that a car is good, they can sell it with some

form of a guarantee: either a guarantee that is explicit, such as a warranty, or an implicit guarantee in which they stand by their reputation for honesty. People are more likely to purchase a used car because of a dealer's guarantee, and the dealer is able to make a profit on the production of information about automobile quality by being able to sell the used car at a higher price than the dealer paid for it. If dealers purchase and then resell cars on which they have produced information, they avoid the problem of other people free-riding on the information they produced.

Just as used-car dealers help solve adverse selection problems in the automobile market, financial intermediaries play a similar role in financial markets. A financial intermediary, such as a bank, becomes an expert in the production of information about firms, so that it can sort out good credit risks from bad ones. Then it can acquire funds from depositors and lend them to the good firms. Because the bank is able to lend mostly to good firms, it is able to earn a higher return on its loans than the interest it has to pay to its depositors. The resulting profit that the bank earns allows it to engage in this information production activity.

An important element in the ability of the bank to profit from the information it produces is that it avoids the free-rider problem by primarily making private loans rather than by purchasing securities that are traded in the open market. Because a private loan is not traded, other investors cannot watch what the bank is doing and bid up the loan's price to the point that the bank receives no compensation for the information it has produced. The bank's role as an intermediary that holds mostly nontraded loans is the key to its success in reducing asymmetric information in financial markets.

Our analysis of adverse selection indicates that financial intermediaries in general—and banks in particular, because they hold a large fraction of nontraded loans—should play a greater role in moving funds to corporations than securities markets do. Our analysis thus explains puzzles 3 and 4: why indirect finance is so much more important than direct finance and why banks are the most important source of external funds for financing businesses.

Another important fact that is explained by the analysis here is the greater importance of banks in the financial systems of developing countries. As we have seen, when the quality of information about firms is better, asymmetric information problems will be less severe, and it will be easier for firms to issue securities. Information about private firms is harder to collect in developing countries than in industrialized countries; therefore, the smaller role played by securities markets leaves a greater role for financial intermediaries such as banks. A corollary of this analysis is that as information about firms becomes easier to acquire, the role of banks should decline. A major development in the past 20 years in the United States has been huge improvements in information technology. Thus the analysis here suggests that the lending role of financial institutions such as banks in the United States should have declined, and this is exactly what has occurred (see Chapter 10).

Our analysis of adverse selection also explains puzzle 6, which questions why large firms are more likely to obtain funds from securities markets, a direct route, rather than from banks and financial intermediaries, an indirect route. The better known a corporation is, the more information about its activities is available in the marketplace. Thus it is easier for investors to evaluate the quality of the corporation and determine whether it is a good firm or a bad one. Because investors have fewer worries about adverse selection with well-known corporations, they will be willing to invest directly in their securities. Our adverse selection analysis thus suggests that there should be a pecking order for firms that can issue securities. The larger and more established a corporation is, the more likely it will be to issue securities to raise

funds, a view that is known as the **pecking order hypothesis**. This hypothesis is supported in the data, and is described in puzzle 6.

Collateral and Net Worth. Adverse selection interferes with the functioning of financial markets only if a lender suffers a loss when a borrower is unable to make loan payments and thereby defaults. *Collateral*, property promised to the lender if the borrower defaults, reduces the consequences of adverse selection because it reduces the lender's losses in the event of a default. If a borrower defaults on a loan, the lender can sell the collateral and use the proceeds to make up for the losses on the loan. For example, if you fail to make your mortgage payments, the lender can take title to your house, auction it off, and use the receipts to pay off the loan. Lenders are thus more willing to make loans secured by collateral, and borrowers are willing to supply collateral because the reduced risk for the lender makes it more likely they will get the loan in the first place and perhaps at a better loan rate. The presence of adverse selection in credit markets thus provides an explanation for why collateral is an important feature of debt contracts (puzzle 7).

Net worth (also called **equity capital**), the difference between a firm's assets (what it owns or is owed) and its liabilities (what it owes), can perform a similar role to collateral. If a firm has a high net worth, then even if it engages in investments that cause it to have negative profits and so defaults on its debt payments, the lender can take title to the firm's net worth, sell it off, and use the proceeds to recoup some of the losses from the loan. In addition, the more net worth a firm has in the first place, the less likely it is to default, because the firm has a cushion of assets that it can use to pay off its loans. Hence when firms seeking credit have high net worth, the consequences of adverse selection are less important and lenders are more willing to make loans. This analysis lies behind the often-heard lament, "Only the people who don't need money can borrow it!"

Summary. So far we have used the concept of adverse selection to explain seven of the eight puzzles about financial structure introduced earlier: The first four emphasize the importance of financial intermediaries and the relative unimportance of securities markets for the financing of corporations; the fifth, that financial markets are among the most heavily regulated sectors of the economy; the sixth, that only large, well-established corporations have access to securities markets; and the seventh, that collateral is an important feature of debt contracts. In the next section, we will see that the other asymmetric information concept of moral hazard provides additional reasons for the importance of financial intermediaries and the relative unimportance of securities markets for the financing of corporations, the prevalence of government regulation, and the importance of collateral in debt contracts. In addition, the concept of moral hazard can be used to explain our final puzzle (puzzle 8) of why debt contracts are complicated legal documents that place substantial restrictions on the behavior of the borrower.

How Moral Hazard Affects the Choice Between Debt and Equity Contracts

Moral hazard is the asymmetric information problem that occurs after the financial transaction takes place, when the seller of a security may have incentives to hide information and engage in activities that are undesirable for the purchaser of the secu-

rity. Moral hazard has important consequences for whether a firm finds it easier to raise funds with debt than with equity contracts.

Moral Hazard in Equity Contracts: The Principal–Agent Problem

Equity contracts, such as common stock, are claims to a share in the profits and assets of a business. Equity contracts are subject to a particular type of moral hazard called the **principal–agent problem**. When managers own only a small fraction of the firm they work for, the stockholders who own most of the firm's equity (called the *principals*) are not the same people as the managers of the firm, who are the *agents* of the owners. This separation of ownership and control involves moral hazard, in that the managers in control (the agents) may act in their own interest rather than in the interest of the stockholder-owners (the principals) because the managers have less incentive to maximize profits than the stockholder-owners do.

To understand the principal–agent problem more fully, suppose that your friend Steve asks you to become a silent partner in his ice-cream store. The store requires an investment of $10,000 to set up and Steve has only $1,000. So you purchase an equity stake (stock shares) for $9,000, which entitles you to 90% of the ownership of the firm, while Steve owns only 10%. If Steve works hard to make tasty ice cream, keeps the store clean, smiles at all the customers, and hustles to wait on tables quickly, after all expenses (including Steve's salary), the store will have $50,000 in profits per year, of which Steve receives 10% ($5,000) and you receive 90% ($45,000).

But if Steve doesn't provide quick and friendly service to his customers, uses the $50,000 in income to buy artwork for his office, and even sneaks off to the beach while he should be at the store, the store will not earn any profit. Steve can earn the additional $5,000 (his 10% share of the profits) over his salary only if he works hard and forgoes unproductive investments (such as art for his office). Steve might decide that the extra $5,000 just isn't enough to make him expend the effort to be a good manager; he might decide that it would be worth his while only if he earned an extra $10,000. If Steve feels this way, he does not have enough incentive to be a good manager and will end up with a beautiful office, a good tan, and a store that doesn't show any profits. Because the store won't show any profits, Steve's decision not to act in your interest will cost you $45,000 (your 90% of the profits if he had chosen to be a good manager instead).

The moral hazard arising from the principal–agent problem might be even worse if Steve were not totally honest. Because his ice-cream store is a cash business, Steve has the incentive to pocket $50,000 in cash and tell you that the profits were zero. He now gets a return of $50,000, but you get nothing.

Further indications that the principal–agent problem created by equity contracts can be severe are provided by recent corporate scandals in corporations such as Enron and Tyco International, in which managers have been accused of diverting funds for their own personal use. Besides pursuing personal benefits, managers might also pursue corporate strategies (such as the acquisition of other firms) that enhance their personal power but do not increase the corporation's profitability

The principal–agent problem would not arise if the owners of a firm had complete information about what the managers were up to and could prevent wasteful expenditures or fraud. The principal–agent problem, which is an example of moral hazard, arises only because a manager, like Steve, has more information about his activities than the stockholder does—that is, there is asymmetric information. The principal–agent problem would also not arise if Steve alone owned the store and there were no separation of ownership and control. If this were the case, Steve's hard work

and avoidance of unproductive investments would yield him a profit (and extra income) of $50,000, an amount that would make it worth his while to be a good manager.

Tools to Help Solve the Principal–Agent Problem

Production of Information: Monitoring. You have seen that the principal–agent problem arises because managers have more information about their activities and actual profits than stockholders do. One way for stockholders to reduce this moral hazard problem is for them to engage in a particular type of information production, the monitoring of the firm's activities: auditing the firm frequently and checking on what the management is doing. The problem is that the monitoring process can be expensive in terms of time and money, as reflected in the name economists give it, **costly state verification**. Costly state verification makes the equity contract less desirable, and it explains, in part, why equity is not a more important element in our financial structure.

As with adverse selection, the free-rider problem decreases the amount of information production that would reduce the moral hazard (principal–agent) problem. In this example, the free-rider problem decreases monitoring. If you know that other stockholders are paying to monitor the activities of the company you hold shares in, you can take a free ride on their activities. Then you can use the money you save by not engaging in monitoring to vacation on a Caribbean island. If you can do this, though, so can other stockholders. Perhaps all the stockholders will go to the islands, and no one will spend any resources on monitoring the firm. The moral hazard problem for shares of common stock will then be severe, making it hard for firms to issue them to raise capital (providing an additional explanation for puzzle 1).

Government Regulation to Increase Information. As with adverse selection, the government has an incentive to try to reduce the moral hazard problem created by asymmetric information, which provides another reason why the financial system is so heavily regulated (puzzle 5). Governments everywhere have laws to force firms to adhere to standard accounting principles that make profit verification easier. They also pass laws to impose stiff criminal penalties on people who commit the fraud of hiding and stealing profits. However, these measures can be only partly effective. Catching this kind of fraud is not easy; fraudulent managers have the incentive to make it very hard for government agencies to find or prove fraud.

Financial Intermediation. Financial intermediaries have the ability to avoid the free-rider problem in the face of moral hazard, and this is another reason why indirect finance is so important (puzzle 3). One financial intermediary that helps reduce the moral hazard arising from the principal–agent problem is the **venture capital firm**. Venture capital firms pool the resources of their partners and use the funds to help budding entrepreneurs start new businesses. In exchange for the use of the venture capital, the firm receives an equity share in the new business. Because verification of earnings and profits is so important in eliminating moral hazard, venture capital firms usually insist on having several of their own people participate as members of the managing body of the firm, the board of directors, so that they can keep a close watch on the firm's activities. When a venture capital firm supplies start-up funds, the equity in the firm is not marketable to anyone *but* the venture capital firm. Thus other investors are unable to take a free ride on the venture capital firm's verification activities. As a result of this arrangement, the venture capital firm is able to garner the full benefits of its verification activities and is given the appropriate incentives to reduce

the moral hazard problem. Venture capital firms have been important in the development of the high-tech sector in the United States, which has resulted in job creation, economic growth, and increased international competitiveness. However, these firms have made mistakes, as Box 2 indicates.

Debt Contracts. Moral hazard arises with an equity contract, which is a claim on profits in all situations, whether the firm is making or losing money. If a contract could be structured so that moral hazard would exist only in certain situations, there would be a reduced need to monitor managers, and the contract would be more attractive than the equity contract. The debt contract has exactly these attributes because it is a contractual agreement by the borrower to pay the lender *fixed* dollar amounts at periodic intervals. When the firm has high profits, the lender receives the contractual payments and does not need to know the exact profits of the firm. If the managers are hiding profits or are pursuing activities that are personally beneficial but don't increase profitability, the lender doesn't care as long as these activities do not interfere with the ability of the firm to make its debt payments on time. Only when the firm cannot meet its debt payments, thereby being in a state of default, is there a need for the lender to verify the state of the firm's profits. Only in this situation do lenders involved in debt contracts need to act more like equity holders; now they need to know how much income the firm has in order to get their fair share.

Box 2: E-Finance

Venture Capitalists and the High-Tech Sector

Over the last half century, venture capital firms have nurtured the growth of America's high technology sector. Venture capitalists backed many of the most successful high-technology companies during the 1980s and 1990s, including Apple Computer, Cisco Systems, Genetech, Microsoft, Netscape, and Sun Microsystems.

Venture capital firms experienced explosive growth during the last half of the 1990s, with investments growing from $5.6 billion in 1995 to more than $103 billion by 2000, increasingly focused on investing in Internet "dot-com" companies. These two developments led to large losses for venture capitalists, for the following reasons.

First, it is likely that there are relatively few projects worthy of financing at any one time. When too much money chases too few deals, firms that would be rejected at other times will obtain financing. Second, the surge of money into venture capital funds reduced the ability of partners of venture capital firms to provide quality monitoring. Third, the infatuation with dot-com firms, many of which did not have adequately developed business plans, meant that too much investment was directed to this sector. Consequently, in the late 1990s, venture capital firms made many poor investments, which led to large losses by the early 2000s.

Consider the case of Webvan, an Internet grocer that received more than $1 billion in venture capital financing. Even though it was backed by a group of experienced financiers, including Goldman Sachs and Sequoia Capital, its business plan was fundamentally flawed. In its short life, Webvan spent more than $1 billion building automated warehouses and pricey tech gear. The resulting high overhead made it impossible to compete in the grocery business. Had the venture capitalists been actively monitoring the activities of Webvan, they might have balked at Webvan's plan to develop an infrastructure that requried 4,000 orders per day per warehouse just to break even. Not surprisingly, Webvan declared bankruptcy in July 2001.

The advantage of a less frequent need to monitor the firm, and thus a lower cost of state verification, helps explain why debt contracts are used more frequently than equity contracts to raise capital. The concept of moral hazard thus helps explain puzzle 1, why stocks are not the most important source of financing for businesses.[4]

How Moral Hazard Influences Financial Structure in Debt Markets

Even with the advantages just described, debt contracts are still subject to moral hazard. Because a debt contract requires the borrowers to pay out a fixed amount and lets them keep any profits above this amount, the borrowers have an incentive to take on investment projects that are riskier than the lenders would like.

For example, suppose that because you are concerned about the problem of verifying the profits of Steve's ice-cream store, you decide not to become an equity partner. Instead, you lend Steve the $9,000 he needs to set up his business and have a debt contract that pays you an interest rate of 10%. As far as you are concerned, this is a surefire investment because there is a strong and steady demand for ice cream in your neighborhood. However, once you give Steve the funds, he might use them for purposes other than you intended. Instead of opening up the ice-cream store, Steve might use your $9,000 loan to invest in chemical research equipment because he thinks he has a 1-in-10 chance of inventing a diet ice cream that tastes every bit as good as the premium brands but has no fat or calories.

Obviously, this is a very risky investment, but if Steve is successful, he will become a multimillionaire. He has a strong incentive to undertake the riskier investment with your money, because the gains to him would be so large if he succeeded. You would clearly be very unhappy if Steve used your loan for the riskier investment, because if he were unsuccessful, which is highly likely, you would lose most, if not all, of the money you gave him. And if he were successful, you wouldn't share in his success—you would still get only a 10% return on the loan because the principal and interest payments are fixed. Because of the potential moral hazard (that Steve might use your money to finance a very risky venture), you would probably not make the loan to Steve, even though an ice-cream store in the neighborhood is a good investment that would provide benefits for everyone.

Tools to Help Solve Moral Hazard in Debt Contracts

Net Worth. When borrowers have more at stake because their *net worth* (the difference between their assets and their liabilities) is high, the risk of moral hazard—the temptation to act in a manner that lenders find objectionable—will be greatly reduced because the borrowers themselves have a lot to lose. Let's return to Steve and his ice-cream business. Suppose that the cost of setting up either the ice-cream store or the research equipment is $100,000 instead of $10,000. So Steve needs to put $91,000 of his own money into the business (instead of $1,000) in addition to the $9,000 supplied by your loan. Now if Steve is unsuccessful in inventing the no-calorie nonfat ice cream, he has a lot to lose—the $91,000 of net worth ($100,000 in assets minus the $9,000 loan from you). He will think twice about undertaking the riskier investment

[4]Another factor that encourages the use of debt contracts rather than equity contracts in the United States is our tax code. Debt interest payments are a deductible expense for American firms, whereas dividend payments to equity shareholders are not.

and is more likely to invest in the ice-cream store, which is more of a sure thing. Hence when Steve has more of his own money (net worth) in the business, you are more likely to make him the loan.

One way of describing the solution that high net worth provides to the moral hazard problem is to say that it makes the debt contract **incentive-compatible**; that is, it aligns the incentives of the borrower with those of the lender. The greater the borrower's net worth, the greater the borrower's incentive to behave in the way that the lender expects and desires, the smaller the moral hazard problem in the debt contract is, and the easier it is for the firm to borrow. Conversely, when the borrower's net worth is lower, the moral hazard problem is greater, and it is harder for the firm to borrow.

Monitoring and Enforcement of Restrictive Covenants. As the example of Steve and his ice-cream store shows, if you could make sure that Steve doesn't invest in anything riskier than the ice-cream store, it would be worth your while to make him the loan. You can ensure that Steve uses your money for the purpose *you* want it to be used for by writing provisions (restrictive covenants) into the debt contract that restrict his firm's activities. By monitoring Steve's activities to see whether he is complying with the restrictive covenants and enforcing the covenants if he is not, you can make sure that he will not take on risks at your expense. Restrictive covenants are directed at reducing moral hazard either by ruling out undesirable behavior or by encouraging desirable behavior. There are four types of restrictive covenants that achieve this objective:

1. ***Covenants to discourage undesirable behavior.*** Covenants can be designed to lower moral hazard by keeping the borrower from engaging in the undesirable behavior of undertaking risky investment projects. Some such covenants mandate that a loan can be used only to finance specific activities, such as the purchase of particular equipment or inventories. Others restrict the borrowing firm from engaging in certain risky business activities, such as purchasing other businesses.
2. ***Covenants to encourage desirable behavior.*** Restrictive covenants can encourage the borrower to engage in desirable activities that make it more likely that the loan will be paid off. One restrictive covenant of this type requires the breadwinner in a household to carry life insurance that pays off the mortgage upon that person's death. Restrictive covenants of this type for businesses focus on encouraging the borrowing firm to keep its net worth high because higher borrower net worth reduces moral hazard and makes it less likely that the lender will suffer losses. These restrictive covenants typically specify that the firm must maintain minimum holdings of certain assets relative to the firm's size.
3. ***Covenants to keep collateral valuable.*** Because collateral is an important protection for the lender, restrictive covenants can encourage the borrower to keep the collateral in good condition and make sure that it stays in the possession of the borrower. This is the type of covenant ordinary people encounter most often. Automobile loan contracts, for example, require the car owner to maintain a minimum amount of collision and theft insurance and prevent the sale of the car unless the loan is paid off. Similarly, the recipient of a home mortgage must have adequate insurance on the home and must pay off the mortgage when the property is sold.
4. ***Covenants to provide information.*** Restrictive covenants also require a borrowing firm to provide information about its activities periodically in the form of quarterly accounting and income reports, thereby making it easier for the lender to monitor the firm and reduce moral hazard. This type of covenant may also stipulate that the lender has the right to audit and inspect the firm's books at any time.

We now see why debt contracts are often complicated legal documents with numerous restrictions on the borrower's behavior (puzzle 8): Debt contracts require complicated restrictive covenants to lower moral hazard.

Financial Intermediation. Although restrictive covenants help reduce the moral hazard problem, they do not eliminate it completely. It is almost impossible to write covenants that rule out *every* risky activity. Furthermore, borrowers may be clever enough to find loopholes in restrictive covenants that make them ineffective.

Another problem with restrictive covenants is that they must be monitored and enforced. A restrictive covenant is meaningless if the borrower can violate it knowing that the lender won't check up or is unwilling to pay for legal recourse. Because monitoring and enforcement of restrictive covenants are costly, the free-rider problem arises in the debt securities (bond) market just as it does in the stock market. If you know that other bondholders are monitoring and enforcing the restrictive covenants, you can free-ride on their monitoring and enforcement. But other bondholders can do the same thing, so the likely outcome is that not enough resources are devoted to monitoring and enforcing the restrictive covenants. Moral hazard therefore continues to be a severe problem for marketable debt.

As we have seen before, financial intermediaries—particularly banks—have the ability to avoid the free-rider problem as long as they make primarily private loans. Private loans are not traded, so no one else can free-ride on the intermediary's monitoring and enforcement of the restrictive covenants. The intermediary making private loans thus receives the benefits of monitoring and enforcement and will work to shrink the moral hazard problem inherent in debt contracts. The concept of moral hazard has provided us with additional reasons why financial intermediaries play a more important role in channeling funds from savers to borrowers than marketable securities do, as described in puzzles 3 and 4.

Summary

The presence of asymmetric information in financial markets leads to adverse selection and moral hazard problems that interfere with the efficient functioning of those markets. Tools to help solve these problems involve the private production and sale of information, government regulation to increase information in financial markets, the importance of collateral and net worth to debt contracts, and the use of monitoring and restrictive covenants. A key finding from our analysis is that the existence of the free-rider problem for traded securities such as stocks and bonds indicates that financial intermediaries—particularly banks—should play a greater role than securities markets in financing the activities of businesses. Economic analysis of the consequences of adverse selection and moral hazard has helped explain the basic features of our financial system and has provided solutions to the eight puzzles about our financial structure outlined at the beginning of this chapter.

Study Guide

To help you keep track of all the tools that help solve asymmetric information problems, summary Table 1 provides a listing of the asymmetric information problems and what tools can help solve them. In addition, it lists how these tools and asymmetric information problems explain the eight puzzles of financial structure described at the beginning of the chapter.

SUMMARY **Table 1 Asymmetric Information Problems and Tools to Solve Them**

Asymmetric Information Problem	Tools to Solve It	Explains Puzzle No.
Adverse Selection	Private Production and Sale of Information	1, 2
	Government Regulation to Increase Information	5
	Financial Intermediation	3, 4, 6
	Collateral and Net Worth	7
Moral Hazard in Equity Contracts (Principal–Agent Problem)	Production of Information: Monitoring	1
	Government Regulation to Increase Information	5
	Financial Intermediation	3
	Debt Contracts	1
Moral Hazard in Debt Contracts	Net Worth	
	Monitoring and Enforcement of Restrictive Covenants	8
	Financial Intermediation	3, 4

Note: List of puzzles:

1. Stocks are not the most important source of external financing.
2. Marketable securities are not the primary source of finance.
3. Indirect finance is more important than direct finance.
4. Banks are the most important source of external funds.
5. The financial system is heavily regulated.
6. Only large, well-established firms have access to securities markets.
7. Collateral is prevalent in debt contracts.
8. Debt contracts have numerous restrictive covenants.

Application

Financial Development and Economic Growth

Recent research has found that an important reason why many developing countries or ex-communist countries like Russia (which are referred to as *transition countries*) experience very low rates of growth is that their financial systems are underdeveloped (a situation referred to as *financial repression*).[5] The economic analysis of financial structure helps explain how an underdeveloped financial system leads to a low state of economic development and economic growth.

The financial systems in developing and transition countries face several difficulties that keep them from operating efficiently. As we have seen, two important tools used to help solve adverse selection and moral hazard problems in credit markets are collateral and restrictive covenants. In many developing countries, the legal system functions poorly, making it hard to make

[5]See World Bank, *Finance for Growth: Policy Choices in a Volatile World* (World Bank and Oxford University Press, 2001) for a survey of this literature and a list of additional references.

effective use of these two tools. In these countries, bankruptcy procedures are often extremely slow and cumbersome. For example, in many countries, **creditors** (holders of debt) must first sue the defaulting debtor for payment, which can take several years, and then, once a favorable judgment has been obtained, the creditor has to sue again to obtain title to the collateral. The process can take in excess of five years, and by the time the lender acquires the collateral, it well may have been neglected and thus have little value. In addition, governments often block lenders from foreclosing on borrowers in politically powerful sectors such as agriculture. Where the market is unable to use collateral effectively, the adverse selection problem will be worse, because the lender will need even more information about the quality of the borrower in order to screen out a good loan from a bad one. The result is that it will be harder for lenders to channel funds to borrowers with the most productive investment opportunities, thereby leading to less productive investment, and hence a slower-growing economy. Similarly, a poorly developed legal system may make it extremely difficult for borrowers to enforce restrictive covenants. Thus they may have a much more limited ability to reduce moral hazard on the part of borrowers and so will be less willing to lend. Again the outcome will be less productive investment and a lower growth rate for the economy.

Governments in developing and transition countries have also often decided to use their financial systems to direct credit to themselves or to favored sectors of the economy by setting interest rates at artificially low levels for certain types of loans, by creating so-called development finance institutions to make specific types of loans, or by directing existing institutions to lend to certain entities. As we have seen, private institutions have an incentive to solve adverse selection and moral hazard problems and lend to borrowers with the most productive investment opportunities. Governments have less incentive to do so because they are not driven by the profit motive and so their directed credit programs may not channel funds to sectors that will produce high growth for the economy. The outcome is again likely to result in less efficient investment and slower growth.

In addition, banks in many developing and transition countries have been nationalized by their governments. Again, because of the absence of the profit motive, these nationalized banks have little incentive to allocate their capital to the most productive uses. Indeed, the primary loan customer of these nationalized banks is often the government, which does not always use the funds wisely.

We have seen that government regulation can increase the amount of information in financial markets to make them work more efficiently. Many developing and transition countries have an underdeveloped regulatory apparatus that retards the provision of adequate information to the marketplace. For example, these countries often have weak accounting standards, making it very hard to ascertain the quality of a borrower's balance sheet. As a result, asymmetric information problems are more severe, and the financial system is severely hampered in channeling funds to the most productive uses.

The institutional environment of a poor legal system, weak accounting standards, inadequate government regulation, and government intervention through directed credit programs and nationalization of banks all help explain why many countries stay poor while others grow richer.

Financial Crises and Aggregate Economic Activity

Agency theory, our economic analysis of the effects of adverse selection and moral hazard, can help us understand **financial crises**, major disruptions in financial markets that are characterized by sharp declines in asset prices and the failures of many financial and nonfinancial firms. Financial crises have been common in most countries throughout modern history. The United States experienced major financial crises in 1819, 1837, 1857, 1873, 1884, 1893, 1907, and 1930–1933 but has not had a full-scale financial crisis since then.[6] Studying financial crises is worthwhile because they have led to severe economic downturns in the past and have the potential for doing so in the future.

Financial crises occur when there is a disruption in the financial system that causes such a sharp increase in adverse selection and moral hazard problems in financial markets that the markets are unable to channel funds efficiently from savers to people with productive investment opportunities. As a result of this inability of financial markets to function efficiently, economic activity contracts sharply.

Factors Causing Financial Crises

To understand why banking and financial crises occur and, more specifically, how they lead to contractions in economic activity, we need to examine the factors that cause them. Five categories of factors can trigger financial crises: increases in interest rates, increases in uncertainty, asset market effects on balance sheets, problems in the banking sector, and government fiscal imbalances.

Increases in Interest Rates. As we saw earlier, individuals and firms with the riskiest investment projects are exactly those who are willing to pay the highest interest rates. If market interest rates are driven up sufficiently because of increased demand for credit or because of a decline in the money supply, good credit risks are less likely to want to borrow while bad credit risks are still willing to borrow. Because of the resulting increase in adverse selection, lenders will no longer want to make loans. The substantial decline in lending will lead to a substantial decline in investment and aggregate economic activity.

Increases in Uncertainty. A dramatic increase in uncertainty in financial markets, due perhaps to the failure of a prominent financial or nonfinancial institution, a recession, or a stock market crash, makes it harder for lenders to screen good from bad credit risks. The resulting inability of lenders to solve the adverse selection problem makes them less willing to lend, which leads to a decline in lending, investment, and aggregate economic activity.

Asset Market Effects on Balance Sheets. The state of firms' balance sheets has important implications for the severity of asymmetric information problems in the financial system. A sharp decline in the stock market is one factor that can cause a serious deterioration in firms' balance sheets that can increase adverse selection and moral hazard

[6] Although we in the United States have not experienced any financial crises since the Great Depression, we have had several close calls—the October 1987 stock market crash, for example. An important reason why we have escaped financial crises is the timely action of the Federal Reserve to prevent them during episodes like that of October 1987. We look at the issue of the Fed's role in preventing financial crises in Chapter 17.

problems in financial markets and provoke a financial crisis. A decline in the stock market means that the net worth of corporations has fallen, because share prices are the valuation of a corporation's net worth. The decline in net worth as a result of a stock market decline makes lenders less willing to lend because, as we have seen, the net worth of a firm plays a role similar to that of collateral. When the value of collateral declines, it provides less protection to lenders, meaning that losses on loans are likely to be more severe. Because lenders are now less protected against the consequences of adverse selection, they decrease their lending, which in turn causes investment and aggregate output to decline. In addition, the decline in corporate net worth as a result of a stock market decline increases moral hazard by providing incentives for borrowing firms to make risky investments, as they now have less to lose if their investments go sour. The resulting increase in moral hazard makes lending less attractive—another reason why a stock market decline and resultant decline in net worth leads to decreased lending and economic activity.

In economies in which inflation has been moderate, which characterizes most industrialized countries, many debt contracts are typically of fairly long maturity with fixed interest rates. In this institutional environment, unanticipated declines in the aggregate price level also decrease the net worth of firms. Because debt payments are contractually fixed in nominal terms, an unanticipated decline in the price level raises the value of firms' liabilities in *real* terms (increases the burden of the debt) but does not raise the real value of firms' assets. The result is that net worth in *real* terms (the difference between assets and liabilities in *real* terms) declines. A sharp drop in the price level therefore causes a substantial decline in real net worth and an increase in adverse selection and moral hazard problems facing lenders. An unanticipated decline in the aggregate price level thus leads to a drop in lending and economic activity.

Because of uncertainty about the future value of the domestic currency in developing countries (and in some industrialized countries), many nonfinancial firms, banks, and governments in these countries find it easier to issue debt denominated in foreign currencies. This can lead to a financial crisis in a similar fashion to an unanticipated decline in the price level. With debt contracts denominated in foreign currency, when there is an unanticipated decline in the value of the domestic currency, the debt burden of domestic firms increases. Since assets are typically denominated in domestic currency, there is a resulting deterioration in firms' balance sheets and a decline in net worth, which then increases adverse selection and moral hazard problems along the lines just described. The increase in asymmetric information problems leads to a decline in investment and economic activity.

Although we have seen that increases in interest rates have a direct effect on increasing adverse selection problems, increases in interest rates also play a role in promoting a financial crisis through their effect on both firms' and households' balance sheets. A rise in interest rates and therefore in households' and firms' interest payments decreases firms' **cash flow**, the difference between cash receipts and cash expenditures. The decline in cash flow causes a deterioration in the balance sheet because it decreases the liquidity of the household or firm and thus makes it harder for lenders to know whether the firm or household will be able to pay its bills. As a result, adverse selection and moral hazard problems become more severe for potential lenders to these firms and households, leading to a decline in lending and economic activity. There is thus an additional reason why sharp increases in interest rates can be an important factor leading to financial crises.

Problems in the Banking Sector. Banks play a major role in financial markets because they are well positioned to engage in information-producing activities that facilitate productive investment for the economy. The state of banks' balance sheets has an important effect on bank lending. If banks suffer a deterioration in their balance sheets and so have a substantial contraction in their capital, they will have fewer resources to lend, and bank lending will decline. The contraction in lending then leads to a decline in investment spending, which slows economic activity.

If the deterioration in bank balance sheets is severe enough, banks will start to fail, and fear can spread from one bank to another, causing even healthy banks to go under. The multiple bank failures that result are known as a **bank panic**. The source of the contagion is again asymmetric information. In a panic, depositors, fearing for the safety of their deposits (in the absence of deposit insurance) and not knowing the quality of banks' loan portfolios, withdraw their deposits to the point that the banks fail. The failure of a large number of banks in a short period of time means that there is a loss of information production in financial markets and hence a direct loss of financial intermediation by the banking sector. The decrease in bank lending during a financial crisis also decreases the supply of funds to borrowers, which leads to higher interest rates. The outcome of a bank panic is an increase in adverse selection and moral hazard problems in credit markets: These problems produce an even sharper decline in lending to facilitate productive investments that leads to an even more severe contraction in economic activity.

Government Fiscal Imbalances. In emerging market countries (Argentina, Brazil, and Turkey are recent examples), government fiscal imbalances may create fears of default on the government debt. As a result, the government may have trouble getting people to buy its bonds and so it might force banks to purchase them. If the debt then declines in price—which, as we have seen in Chapter 6, will occur if a government default is likely—this can substantially weaken bank balance sheets and lead to a contraction in lending for the reasons described earlier. Fears of default on the government debt can also spark a foreign exchange crisis in which the value of the domestic currency falls sharply because investors pull their money out of the country. The decline in the domestic currency's value will then lead to the destruction of the balance sheets of firms with large amounts of debt denominated in foreign currency. These balance sheet problems lead to an increase in adverse selection and moral hazard problems, a decline in lending, and a contraction of economic activity.

Application

Financial Crises in the United States

As mentioned, the United States has a long history of banking and financial crises, such crises having occurred every 20 years or so in the nineteenth and early twentieth centuries—in 1819, 1837, 1857, 1873, 1884, 1893, 1907, and 1930–1933. Our analysis of the factors that lead to a financial crisis can explain why these crises took place and why they were so damaging to the U.S. economy.

Study Guide

To understand fully what took place in a U.S. financial crisis, make sure that you can state the reasons why each of the factors—increases in interest rates, increases in uncertainty, asset market effects on balance sheets, and problems in the banking sector—increases adverse selection and moral hazard problems, which in turn lead to a decline in economic activity. To help you understand these crises, you might want to refer to Figure 3, a diagram that traces the sequence of events in a U.S. financial crisis.

www.amatecon.com/gd/gdtimeline.html
A time line of the Great Depression.

As shown in Figure 3, most financial crises in the United States have begun with a deterioration in banks' balance sheets, a sharp rise in interest rates (frequently stemming from increases in interest rates abroad), a steep stock market decline, and an increase in uncertainty resulting from a failure of major financial or nonfinancial firms (the Ohio Life Insurance & Trust Company in 1857, the Northern Pacific Railroad and Jay Cooke & Company in 1873, Grant & Ward in 1884, the National Cordage Company in 1893, the Knickerbocker Trust Company in 1907, and the Bank of United States in 1930). During these crises, deterioration in banks' balance sheets, the increase in uncertainty, the rise in interest rates, and the stock market decline increased the severity of adverse selection problems in credit markets; the stock market decline, the deterioration in banks' balance sheets, and the rise in interest rates, which decreased firms' cash flow, also increased moral hazard problems. The rise in adverse selection and moral hazard problems then made it less attractive for lenders to lend and led to a decline in investment and aggregate economic activity.

Because of the worsening business conditions and uncertainty about their bank's health (perhaps banks would go broke), depositors began to withdraw their funds from banks, which led to bank panics. The resulting decline in the number of banks raised interest rates even further and decreased the amount of financial intermediation by banks. Worsening of the problems created by adverse selection and moral hazard led to further economic contraction.

Finally, there was a sorting out of firms that were **insolvent** (had a negative net worth and hence were bankrupt) from healthy firms by bankruptcy proceedings. The same process occurred for banks, often with the help of public and private authorities. Once this sorting out was complete, uncertainty in financial markets declined, the stock market underwent a recovery, and interest rates fell. The overall result was that adverse selection and moral hazard problems diminished and the financial crisis subsided. With the financial markets able to operate well again, the stage was set for the recovery of the economy.

If, however, the economic downturn led to a sharp decline in prices, the recovery process was short-circuited. In this situation, shown in Figure 3, a process called **debt deflation** occurred, in which a substantial decline in the price level set in, leading to a further deterioration in firms' net worth because of the increased burden of indebtedness. When debt deflation set in,

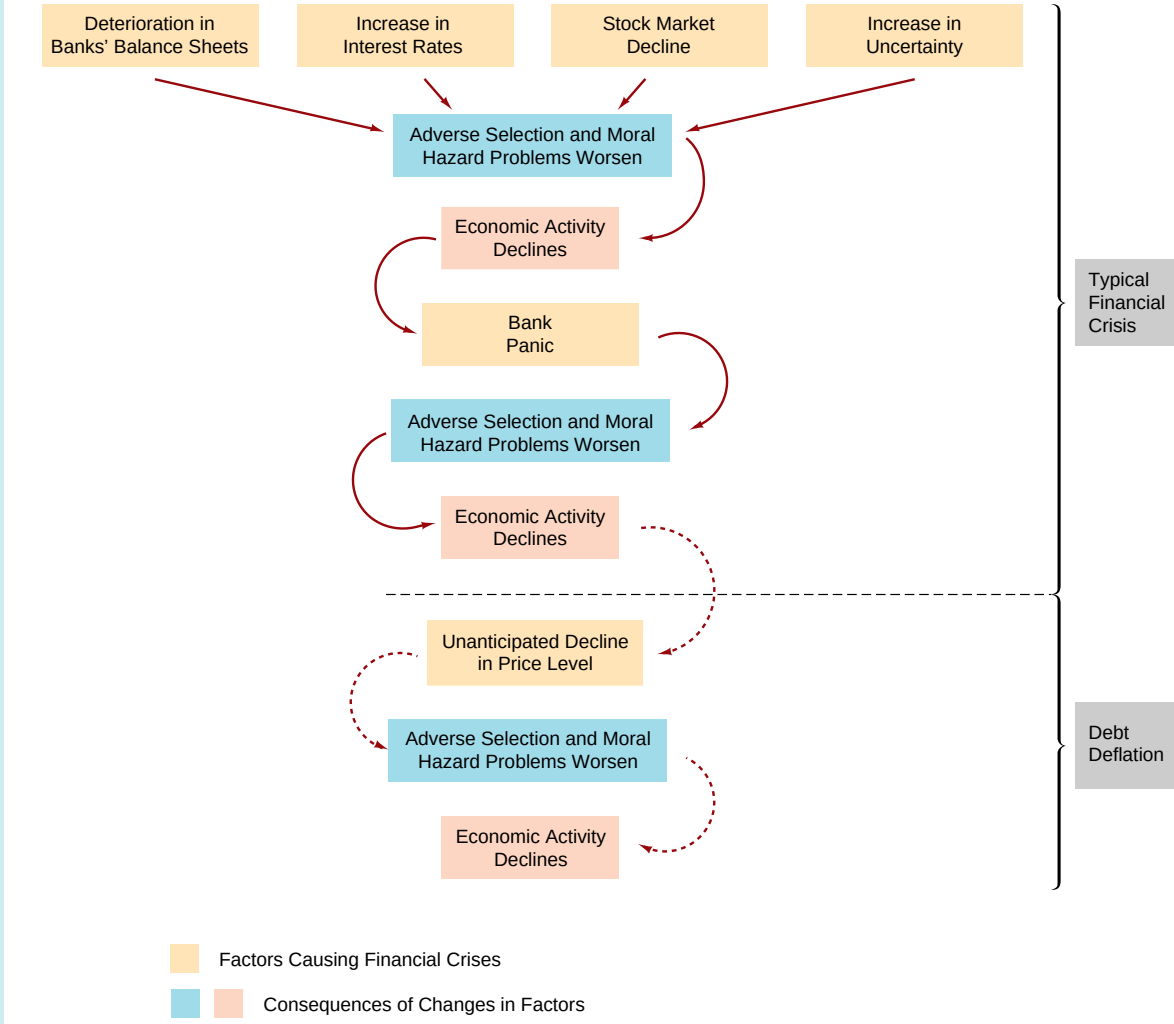


FIGURE 3 Sequence of Events in U.S. Financial Crises

The solid arrows trace the sequence of events in a typical financial crisis; the dotted arrows show the additional set of events that occur if the crisis develops into a debt deflation.

the adverse selection and moral hazard problems continued to increase so that lending, investment spending, and aggregate economic activity remained depressed for a long time. The most significant financial crisis that included debt deflation was the Great Depression, the worst economic contraction in U.S. history (see Box 3).

Box 3

Case Study of a Financial Crisis

The Great Depression. Federal Reserve officials viewed the stock market boom of 1928 and 1929, during which stock prices doubled, as excessive speculation. To curb it, they pursued a tight monetary policy to raise interest rates. The Fed got more than it bargained for when the stock market crashed in October 1929.

Although the 1929 crash had a great impact on the minds of a whole generation, most people forget that by the middle of 1930, more than half of the stock market decline had been reversed. What might have been a normal recession turned into something far different, however, with adverse shocks to the agricultural sector, a continuing decline in the stock market after the middle of 1930, and a sequence of bank collapses from October 1930 until March 1933 in which over one-third of the banks in the United States went out of business (events described in more detail in Chapter 18).

The continuing decline in stock prices after mid-1930 (by mid-1932 stocks had declined to 10% of their value at the 1929 peak) and the increase in uncertainty from the unsettled business conditions created by the economic contraction made adverse selection and moral hazard problems worse in the credit markets. The loss of one-third of the banks reduced the amount of financial intermediation. This intensified adverse selection and moral hazard problems, thereby decreasing the ability of financial markets to channel funds to firms with productive investment opportunities. As our analysis predicts, the amount of outstanding commercial loans fell by half from 1929 to 1933, and investment spending collapsed, declining by 90% from its 1929 level.

The short-circuiting of the process that kept the economy from recovering quickly, which it does in most recessions, occurred because of a fall in the price level by 25% in the 1930–1933 period. This huge decline in prices triggered a debt deflation in which net worth fell because of the increased burden of indebtedness borne by firms. The decline in net worth and the resulting increase in adverse selection and moral hazard problems in the credit markets led to a prolonged economic contraction in which unemployment rose to 25% of the labor force. The financial crisis in the Great Depression was the worst ever experienced in the United States, and it explains why this economic contraction was also the most severe one ever experienced by the nation.*

*See Ben Bernanke, "Nonmonetary Effects of the Financial Crisis in the Propagation of the Great Depression," *American Economic Review* 73 (1983): 257–276, for a discussion of the role of asymmetric information problems in the Great Depression period.

Application

Financial Crises in Emerging-Market Countries: Mexico, 1994–1995; East Asia, 1997–1998; and Argentina, 2001–2002

In recent years, many emerging-market countries have experienced financial crises, the most dramatic of which were the Mexican crisis, which started in December 1994; the East Asian crisis, which started in July 1997; and the Argentine crisis, which started in 2001. An important puzzle is how a developing country can shift dramatically from a path of high growth before a financial crisis—as was true for Mexico and particularly the East Asian countries of Thailand, Malaysia, Indonesia, the Philippines, and South Korea—to a sharp decline in economic activity. We can apply our asymmetric information

analysis of financial crises to explain this puzzle and to understand the Mexican, East Asian, and Argentine financial situations.[7]

Because of the different institutional features of emerging-market countries' debt markets, the sequence of events in the Mexican, East Asian, and Argentine crises is different from that occurring in the United States in the nineteenth and twentieth centuries. Figure 4 diagrams the sequence of events that occurred in Mexico, East Asia, and Argentina.

An important factor leading up to the financial crises in Mexico and East Asia was the deterioration in banks' balance sheets because of increasing loan losses. When financial markets in these countries were deregulated in the early 1990s, a lending boom ensued in which bank credit to the private nonfinancial business sector accelerated sharply. Because of weak supervision by bank regulators and a lack of expertise in screening and monitoring borrowers at banking institutions, losses on the loans began to mount, causing an erosion of banks' net worth (capital). As a result of this erosion, banks had fewer resources to lend, and this lack of lending eventually led to a contraction in economic activity.

Argentina also experienced a deterioration in bank balance sheets leading up to its crisis, but the source of this deterioration was quite different. In contrast to Mexico and the East Asian crisis countries, Argentina had a well-supervised banking system, and a lending boom did not occur before the crisis. On the other hand, in 1998 Argentina entered a recession (you can find out more on why this occurred in Chapter 20) that led to some loan losses. However, it was the fiscal problems of the Argentine government that led to severe weakening of bank balance sheets. Again in contrast to Mexico and the East Asian countries before their crises, Argentina was running substantial budget deficits that could not be financed by foreign borrowing. To solve its fiscal problems, the Argentine government coerced banks into absorbing large amounts of government debt. When investors lost confidence in the ability of the Argentine government to repay this debt, the price of this debt plummeted, leaving big holes in commercial banks' balance sheets. This weakening in bank balance sheets, as in Mexico and East Asia, helped lead to a contraction of economic activity.

Consistent with the U.S. experience in the nineteenth and early twentieth centuries, another precipitating factor in the Mexican and Argentine (but not East Asian) financial crises was a rise in interest rates abroad. Before the Mexican crisis, in February 1994, and before the Argentine crisis, in mid-1999, the Federal Reserve began a cycle of raising the federal funds rate to head off inflationary pressures. Although the monetary policy moves by the Fed were quite successful in keeping inflation in check in the United States, they put upward pressure on interest rates in both Mexico and Argentina.

[7]This chapter does not examine two other recent crises, those in Brazil and Russia. Russia's financial crisis in August 1998 can also be explained with the asymmetric information story here, but it is more appropriate to view it as a symptom of a wider breakdown in the economy—and this is why we do not focus on it here. The Brazilian crisis in January 1999 has features of a more traditional balance-of-payments crisis (see Chapter 20), rather than a financial crisis.

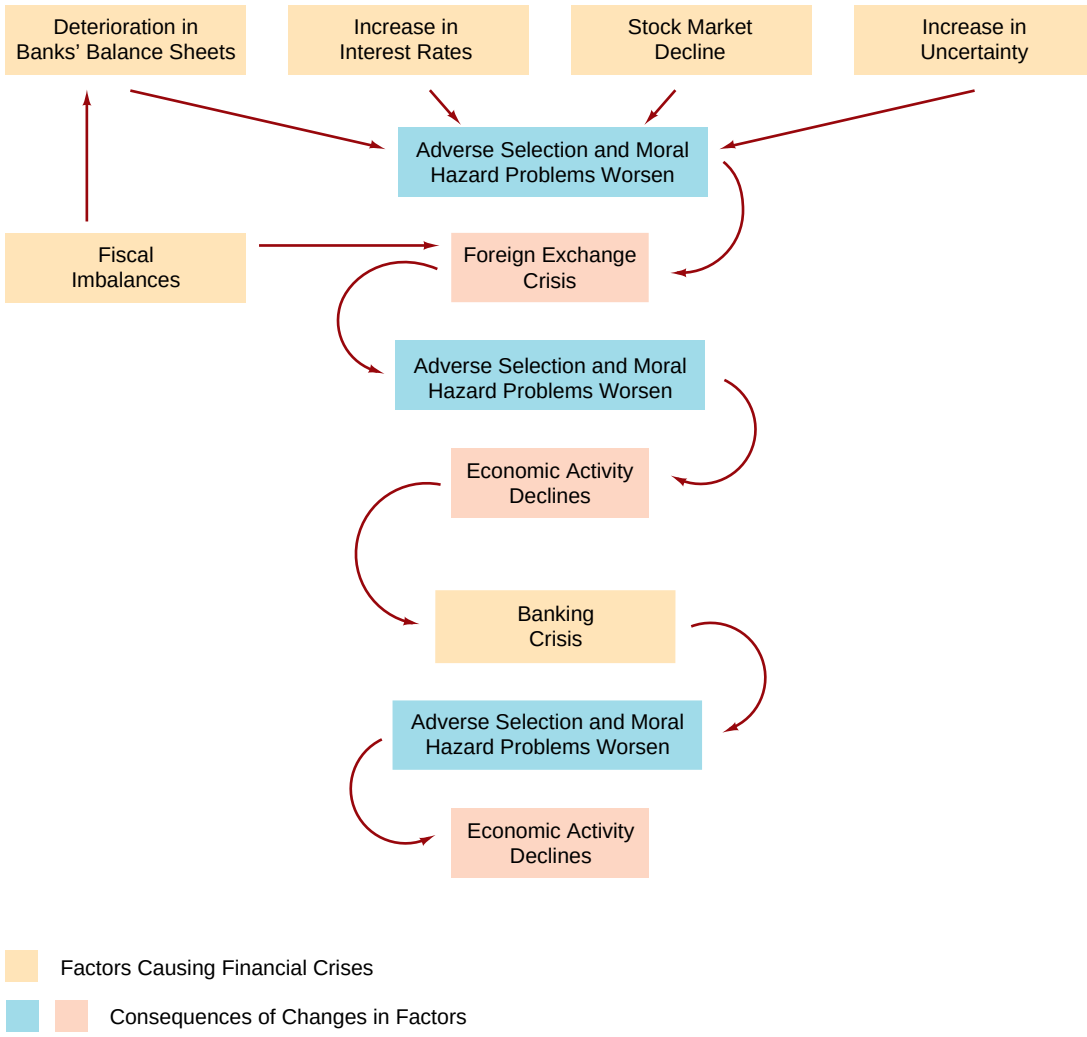


FIGURE 4 Sequence of Events in the Mexican, East Asian, and Argentine Financial Crises
The arrows trace the sequence of events during the financial crisis.

The rise in interest rates in Mexico and Argentina directly added to increased adverse selection in their financial markets because, as discussed earlier, it was more likely that the parties willing to take on the most risk would seek loans.

Also consistent with the U.S. experience in the nineteenth and early twentieth centuries, stock market declines and increases in uncertainty occurred prior to and contributed to full-blown crises in Mexico, Thailand,

South Korea, and Argentina. (The stock market declines in Malaysia, Indonesia, and the Philippines, on the other hand, occurred simultaneously with the onset of the crisis.) The Mexican economy was hit by political shocks in 1994 (specifically, the assassination of the ruling party's presidential candidate and an uprising in the southern state of Chiapas) that created uncertainty, while the ongoing recession increased uncertainty in Argentina. Right before their crises, Thailand and Korea experienced major failures of financial and nonfinancial firms that increased general uncertainty in financial markets

As we have seen, an increase in uncertainty and a decrease in net worth as a result of a stock market decline increase asymmetric information problems. It becomes harder to screen out good from bad borrowers, and the decline in net worth decreases the value of firms' collateral and increases their incentives to make risky investments because there is less equity to lose if the investments are unsuccessful. The increase in uncertainty and stock market declines that occurred before the crisis, along with the deterioration in banks' balance sheets, worsened adverse selection and moral hazard problems (shown at the top of the diagram in Figure 4) and made the economies ripe for a serious financial crisis.

At this point, full-blown speculative attacks developed in the foreign exchange market, plunging these countries into a full-scale crisis. With the Colosio assassination, the Chiapas uprising, and the growing weakness in the banking sector, the Mexican peso came under attack. Even though the Mexican central bank intervened in the foreign exchange market and raised interest rates sharply, it was unable to stem the attack and was forced to devalue the peso on December 20, 1994. In the case of Thailand, concerns about the large current account deficit and weakness in the Thai financial system, culminating with the failure of a major finance company, Finance One, led to a successful speculative attack that forced the Thai central bank to allow the baht to float downward in July 1997. Soon thereafter, speculative attacks developed against the other countries in the region, leading to the collapse of the Philippine peso, the Indonesian rupiah, the Malaysian ringgit, and the South Korean won. In Argentina, a full-scale banking panic began in October–November 2001. This, along with realization that the government was going to default on its debt, also led to a speculative attack on the Argentine peso, resulting in its collapse on January 6, 2002.

The institutional structure of debt markets in Mexico and East Asia now interacted with the currency devaluations to propel the economies into full-fledged financial crises. Because so many firms in these countries had debt denominated in foreign currencies like the dollar and the yen, depreciation of their currencies resulted in increases in their indebtedness in domestic currency terms, even though the value of their assets remained unchanged. When the peso lost half its value by March 1995 and the Thai, Philippine, Malaysian, and South Korean currencies lost between a third and half of their value by the beginning of 1998, firms' balance sheets took a big negative hit, causing a dramatic increase in adverse selection and moral hazard problems. This negative shock was especially severe for Indonesia and Argentina, which saw the value of their currencies fall by over 70%, resulting in insolvency for firms with substantial amounts of debt denominated in foreign currencies.

The collapse of currencies also led to a rise in actual and expected inflation in these countries, and market interest rates rose sky-high (to around 100% in Mexico and Argentina). The resulting increase in interest payments caused reductions in households' and firms' cash flow, which led to further deterioration in their balance sheets. A feature of debt markets in emerging-market countries, like those in Mexico, East Asia, and Argentina is that debt contracts have very short durations, typically less than one month. Thus the rise in short-term interest rates in these countries meant that the effect on cash flow and hence on balance sheets was substantial. As our asymmetric information analysis suggests, this deterioration in households' and firms' balance sheets increased adverse selection and moral hazard problems in the credit markets, making domestic and foreign lenders even less willing to lend.

Consistent with the theory of financial crises outlined in this chapter, the sharp decline in lending helped lead to a collapse of economic activity, with real GDP growth falling sharply.

As shown in Figure 4, further deterioration in the economy occurred because the collapse in economic activity and the deterioration in the cash flow and balance sheets of both firms and households led to worsening banking crises. The problems of firms and households meant that many of them were no longer able to pay off their debts, resulting in substantial losses for the banks. Even more problematic for the banks was that they had many short-term liabilities denominated in foreign currencies, and the sharp increase in the value of these liabilities after the devaluation lead to a further deterioration in the banks' balance sheets. Under these circumstances, the banking system would have collapsed in the absence of a government safety net—as it did in the United States during the Great Depression—but with the assistance of the International Monetary Fund, these countries were in some cases able to protect depositors and avoid a bank panic. However, given the loss of bank capital and the need for the government to intervene to prop up the banks, the banks' ability to lend was nevertheless sharply curtailed. As we have seen, a banking crisis of this type hinders the ability of the banks to lend and also makes adverse selection and moral hazard problems worse in financial markets, because banks are less capable of playing their traditional financial intermediation role. The banking crisis, along with other factors that increased adverse selection and moral hazard problems in the credit markets of Mexico, East Asia, and Argentina, explains the collapse of lending and hence economic activity in the aftermath of the crisis.

In the aftermath of their crises, Mexico began to recover in 1996, while the crisis countries in East Asia saw the glimmer of recovery in 1999. Argentina was still in a severe depression in 2003. In all these countries, the economic hardship caused by the financial crises was tremendous. Unemployment rose sharply, poverty increased substantially, and even the social fabric of the society was stretched thin. For example, Mexico City and Buenos Aires have become crime-ridden, while Indonesia has experienced waves of ethnic violence.

Summary

1. There are eight basic puzzles about our financial structure. The first four emphasize the importance of financial intermediaries and the relative unimportance of securities markets for the financing of corporations; the fifth recognizes that financial markets are among the most heavily regulated sectors of the economy; the sixth states that only large, well-established corporations have access to securities markets; the seventh indicates that collateral is an important feature of debt contracts; and the eighth presents debt contracts as complicated legal documents that place substantial restrictions on the behavior of the borrower.
2. Transaction costs freeze many small savers and borrowers out of direct involvement with financial markets. Financial intermediaries can take advantage of economies of scale and are better able to develop expertise to lower transaction costs, thus enabling their savers and borrowers to benefit from the existence of financial markets.
3. Asymmetric information results in two problems: adverse selection, which occurs before the transaction, and moral hazard, which occurs after the transaction. Adverse selection refers to the fact that bad credit risks are the ones most likely to seek loans, and moral hazard refers to the risk of the borrower's engaging in activities that are undesirable from the lender's point of view.
4. Adverse selection interferes with the efficient functioning of financial markets. Tools to help reduce the adverse selection problem include private production and sale of information, government regulation to increase information, financial intermediation, and collateral and net worth. The free-rider problem occurs when people who do not pay for information take advantage of information that other people have paid for. This problem explains why financial intermediaries, particularly banks, play a more important role in financing the activities of businesses than securities markets do.
5. Moral hazard in equity contracts is known as the principal–agent problem, because managers (the agents) have less incentive to maximize profits than stockholders (the principals). The principal–agent problem explains why debt contracts are so much more prevalent in financial markets than equity contracts. Tools to help reduce the principal–agent problem include monitoring, government regulation to increase information, and financial intermediation.
6. Tools to reduce the moral hazard problem in debt contracts include net worth, monitoring and enforcement of restrictive covenants, and financial intermediaries.
7. Financial crises are major disruptions in financial markets. They are caused by increases in adverse selection and moral hazard problems that prevent financial markets from channeling funds to people with productive investment opportunities, leading to a sharp contraction in economic activity. The five types of factors that lead to financial crises are increases in interest rates, increases in uncertainty, asset market effects on balance sheets, problems in the banking sector, and government fiscal imbalances.

Key Terms

agency theory, p. 175, 189
bank panic, p. 191
cash flow, p. 190
collateral, p. 172
costly state verification, p. 182
creditor, p. 188
debt deflation, p. 192
financial crisis, p. 189
free-rider problem, p. 176
incentive-compatible, p. 185
insolvent, p. 192
net worth (equity capital), p. 180
pecking order hypothesis, p. 180
principal–agent problem, p. 181
restrictive covenants, p. 172
secured debt, p. 172
unsecured debt, p. 172
venture capital firm, p. 182

Questions and Problems

Questions marked with an asterisk are answered at the end of the book in an appendix, "Answers to Selected Questions and Problems."

1. How can economies of scale help explain the existence of financial intermediaries?

*2. Describe two ways in which financial intermediaries help lower transaction costs in the economy.

3. Would moral hazard and adverse selection still arise in financial markets if information were not asymmetric? Explain.

*4. How do standard accounting principles required by the government help financial markets work more efficiently?

5. Do you think the lemons problem would be more severe for stocks traded on the New York Stock Exchange or those traded over-the-counter? Explain.

*6. Which firms are most likely to use bank financing rather than to issue bonds or stocks to finance their activities? Why?

7. How can the existence of asymmetric information provide a rationale for government regulation of financial markets?

*8. Would you be more willing to lend to a friend if she put all of her life savings into her business than you would if she had not done so? Why?

9. Rich people often worry that others will seek to marry them only for their money. Is this a problem of adverse selection?

*10. The more collateral there is backing a loan, the less the lender has to worry about adverse selection. Is this statement true, false, or uncertain? Explain your answer.

11. How does the free-rider problem aggravate adverse selection and moral hazard problems in financial markets?

*12. Explain how the separation of ownership and control in American corporations might lead to poor management.

13. Is a financial crisis more likely to occur when the economy is experiencing deflation or inflation? Explain.

*14. How can a stock market crash provoke a financial crisis?

15. How can a sharp rise in interest rates provoke a financial crisis?

Web Exercises

1. In this chapter we discuss the lemons problem and its effect on the efficient functioning of a market. This theory was initially developed by George Akerlof. Go to www.nobel.se/economics/laureates/2001/public.html. This site reports that Akerlof, Spence, and Stiglitz were awarded the Nobel prize in economics in 2001 for their work. Read this report down through the section on George Akerlof. Summarize his research ideas in one page.

2. This chapter discusses how an understanding of adverse selection and moral hazard can help us better understand financial crises. The greatest financial crisis faced by the U.S. has been the Great Depression from 1929–1933. Go to www.amatecon.com/greatdepression.html. This site contains a brief discussion of the factors that led to the Depression. Write a one-page summary explaining how adverse selection and moral hazard contributed to the Depression.

Chapter 9

Banking and the Management of Financial Institutions

PREVIEW

Because banking plays such a major role in channeling funds to borrowers with productive investment opportunities, this financial activity is important in ensuring that the financial system and the economy run smoothly and efficiently. In the United States, banks (depository institutions) supply more than $5 trillion in credit annually. They provide loans to businesses, help us finance our college educations or the purchase of a new car or home, and provide us with services such as checking and savings accounts.

In this chapter, we examine how banking is conducted to earn the highest profits possible: how and why banks make loans, how they acquire funds and manage their assets and liabilities (debts), and how they earn income. Although we focus on commercial banking, because this is the most important financial intermediary activity, many of the same principles are applicable to other types of financial intermediation.

The Bank Balance Sheet

www.bankofamerica.com/investor/index.cfm?section=700
This web site shows a sample bank balance sheet.

To understand how banking works, first we need to examine the bank **balance sheet**, a list of the bank's assets and liabilities. As the name implies, this list balances; that is, it has the characteristic that:

$$\text{total assets} = \text{total liabilities} + \text{capital}$$

Furthermore, a bank's balance sheet lists *sources* of bank funds (liabilities) and *uses* to which they are put (assets). Banks obtain funds by borrowing and by issuing other liabilities such as deposits. They then use these funds to acquire assets such as securities and loans. Banks make profits by charging an interest rate on their holdings of securities and loans that is higher than the expenses on their liabilities. The balance sheet of all commercial banks as of January 2003 appears in Table 1.

Liabilities

A bank acquires funds by issuing (selling) liabilities, which are consequently also referred to as *sources of funds*. The funds obtained from issuing liabilities are used to purchase income-earning assets.

Checkable Deposits. Checkable deposits are bank accounts that allow the owner of the account to write checks to third parties. Checkable deposits include all accounts

Table 1 Balance Sheet of All Commercial Banks (items as a percentage of the total, January 2003)

Assets (Uses of Funds)*		Liabilities (Sources of Funds)	
Reserves and cash items	5	Checkable deposits	9
Securities		Nontransaction deposits	
U.S. government and agency	15	Small-denomination time deposits (< $100,000) + savings deposits	42
State and local government and other securities	10	Large-denomination time deposits	14
Loans		Borrowings	28
Commercial and industrial	14	Bank capital	7
Real estate	29		
Consumer	9		
Interbank	4		
Other	8		
Other assets (for example, physical capital)	6		
Total	100	Total	100

*In order of decreasing liquidity.

Source: www.federalreserve.gov/releases/h8/current/.

on which checks can be drawn: non-interest-bearing checking accounts (demand deposits), interest-bearing NOW (negotiable order of withdrawal) accounts, and money market deposit accounts (MMDAs). Introduced with the Depository Institutions Act in 1982, MMDAs have features similar to those of money market mutual funds and are included in the checkable deposits category. However, MMDAs differ from checkable deposits in that they are not subject to reserve requirements (discussed later in the chapter) as checkable deposits are and are not included in the M1 definition of money. Table 1 shows that the category of checkable deposits is an important source of bank funds, making up 9% of bank liabilities. Once checkable deposits were the most important source of bank funds (over 60% of bank liabilities in 1960), but with the appearance of new, more attractive financial instruments, such as money market mutual funds, the share of checkable deposits in total bank liabilities has shrunk over time.

Checkable deposits and money market deposit accounts are payable on demand; that is, if a depositor shows up at the bank and requests payment by making a withdrawal, the bank must pay the depositor immediately. Similarly, if a person who receives a check written on an account from a bank, presents that check at the bank, it must pay the funds out immediately (or credit them to that person's account).

A checkable deposit is an asset for the depositor because it is part of his or her wealth. Conversely, because the depositor can withdraw from an account funds that

the bank is obligated to pay, checkable deposits are a liability for the bank. They are usually the lowest-cost source of bank funds because depositors are willing to forgo some interest in order to have access to a liquid asset that can be used to make purchases. The bank's costs of maintaining checkable deposits include interest payments and the costs incurred in servicing these accounts—processing and storing canceled checks, preparing and sending out monthly statements, providing efficient tellers (human or otherwise), maintaining an impressive building and conveniently located branches, and advertising and marketing to entice customers to deposit their funds with a given bank. In recent years, interest paid on deposits (checkable and time) has accounted for around 25% of total bank operating expenses, while the costs involved in servicing accounts (employee salaries, building rent, and so on) have been approximately 50% of operating expenses.

Nontransaction Deposits. Nontransaction deposits are the primary source of bank funds (56% of bank liabilities in Table 1). Owners cannot write checks on nontransaction deposits, but the interest rates are usually higher than those on checkable deposits. There are two basic types of nontransaction deposits: savings accounts and time deposits (also called certificates of deposit, or CDs).

Savings accounts were once the most common type of nontransaction deposit. In these accounts, to which funds can be added or from which funds can be withdrawn at any time, transactions and interest payments are recorded in a monthly statement or in a small book (the passbook) held by the owner of the account.

Time deposits have a fixed maturity length, ranging from several months to over five years, and have substantial penalties for early withdrawal (the forfeiture of several months' interest). Small-denomination time deposits (deposits of less than $100,000) are less liquid for the depositor than passbook savings, earn higher interest rates, and are a more costly source of funds for the banks.

Large-denomination time deposits (CDs) are available in denominations of $100,000 or over and are typically bought by corporations or other banks. Large-denomination CDs are negotiable; like bonds, they can be resold in a secondary market before they mature. For this reason, negotiable CDs are held by corporations, money market mutual funds, and other financial institutions as alternative assets to Treasury bills and other short-term bonds. Since 1961, when they first appeared, negotiable CDs have become an important source of bank funds (14%).

A bank's borrowings from the Federal Reserve System; also known as advances.

Borrowings. Banks obtain funds by borrowing from the Federal Reserve System, the Federal Home Loan banks, other banks, and corporations. Borrowings from the Fed are called **discount loans** (also known as *advances*). Banks also borrow reserves overnight in the federal (fed) funds market from other U.S. banks and financial institutions. Banks borrow funds overnight in order to have enough deposits at the Federal Reserve to meet the amount required by the Fed. (The *federal funds* designation is somewhat confusing, because these loans are not made by the federal government or by the Federal Reserve, but rather by banks to other banks.) Other sources of borrowed funds are loans made to banks by their parent companies (bank holding companies), loan arrangements with corporations (such as repurchase agreements), and borrowings of Eurodollars (deposits denominated in U.S. dollars residing in foreign banks or foreign branches of U.S. banks). Borrowings have become a more important source of bank funds over time: In 1960, they made up only 2% of bank liabilities; currently, they are 28% of bank liabilities.

Bank Capital. The final category on the liabilities side of the balance sheet is bank capital, the bank's net worth, which equals the difference between total assets and liabilities (7% of total bank assets in Table 1). The funds are raised by selling new equity (stock) or from retained earnings. Bank capital is a cushion against a drop in the value of its assets, which could force the bank into insolvency (having liabilities in excess of assets, meaning that the bank can be forced into liquidation).

Assets

A bank uses the funds that it has acquired by issuing liabilities to purchase income-earning assets. Bank assets are thus naturally referred to as *uses of funds*, and the interest payments earned on them are what enable banks to make profits.

Regulation making it obligatory for depository institutions to keep a certain fraction of their deposits in accounts with the Fed.

Reserves. All banks hold some of the funds they acquire as deposits in an account at the Fed. **Reserves** are these deposits plus currency that is physically held by banks (called **vault cash** because it is stored in bank vaults overnight). Although reserves currently do not pay any interest, banks hold them for two reasons. First, some reserves, called **required reserves**, are held because of **reserve requirements**, the regulation that for every dollar of checkable deposits at a bank, a certain fraction (10 cents, for example) must be kept as reserves. This fraction (10 percent in the example) is called the **required reserve ratio**. Banks hold additional reserves, called **excess reserves**, because they are the most liquid of all bank assets and can be used by a bank to meet its obligations when funds are withdrawn, either directly by a depositor or indirectly when a check is written on an account.

Cash Items in Process of Collection. Suppose that a check written on an account at another bank is deposited in your bank and the funds for this check have not yet been received (collected) from the other bank. The check is classified as a cash item in process of collection, and it is an asset for your bank because it is a claim on another bank for funds that will be paid within a few days.

Deposits at Other Banks. Many small banks hold deposits in larger banks in exchange for a variety of services, including check collection, foreign exchange transactions, and help with securities purchases. This is an aspect of a system called *correspondent banking*.

Collectively, reserves, cash items in process of collection, and deposits at other banks are often referred to as *cash items*. In Table 1, they constitute only 5% of total assets, and their importance has been shrinking over time: In 1960, for example, they accounted for 20% of total assets.

Securities. A bank's holdings of securities are an important income-earning asset: Securities (made up entirely of debt instruments for commercial banks, because banks are not allowed to hold stock) account for 25% of bank assets in Table 1, and they provide commercial banks with about 10% of their revenue. These securities can be classified into three categories: U.S. government and agency securities, state and local government securities, and other securities. The United States government and agency securities are the most liquid because they can be easily traded and converted into cash with low transaction costs. Because of their high liquidity, short-term U.S. government securities are called **secondary reserves**.

State and local government securities are desirable for banks to hold, primarily because state and local governments are more likely to do business with banks that

hold their securities. State and local government and other securities are less marketable (hence less liquid) and are also riskier than U.S. government securities, primarily because of default risk: There is some possibility that the issuer of the securities may not be able to make its interest payments or pay back the face value of the securities when they mature.

Loans. Banks make their profits primarily by issuing loans. In Table 1, some 64% of bank assets are in the form of loans, and in recent years they have generally produced more than half of bank revenues. A loan is a liability for the individual or corporation receiving it, but an asset for a bank, because it provides income to the bank. Loans are typically less liquid than other assets, because they cannot be turned into cash until the loan matures. If the bank makes a one-year loan, for example, it cannot get its funds back until the loan comes due in one year. Loans also have a higher probability of default than other assets. Because of the lack of liquidity and higher default risk, the bank earns its highest return on loans.

As you saw in Table 1, the largest categories of loans for commercial banks are commercial and industrial loans made to businesses and real estate loans. Commercial banks also make consumer loans and lend to each other. The bulk of these interbank loans are overnight loans lent in the federal funds market. The major difference in the balance sheets of the various depository institutions is primarily in the type of loan in which they specialize. Savings and loans and mutual savings banks, for example, specialize in residential mortgages, while credit unions tend to make consumer loans.

Other Assets. The physical capital (bank buildings, computers, and other equipment) owned by the banks is included in this category.

Basic Banking

Before proceeding to a more detailed study of how a bank manages its assets and liabilities in order to make the highest profit, you should understand the basic operation of a bank.

In general terms, banks make profits by selling liabilities with one set of characteristics (a particular combination of liquidity, risk, size, and return) and using the proceeds to buy assets with a different set of characteristics. This process is often referred to as *asset transformation*. For example, a savings deposit held by one person can provide the funds that enable the bank to make a mortgage loan to another person. The bank has, in effect, transformed the savings deposit (an asset held by the depositor) into a mortgage loan (an asset held by the bank). Another way this process of asset transformation is described is to say that the bank "borrows short and lends long" because it makes long-term loans and funds them by issuing short-dated deposits.

The process of transforming assets and providing a set of services (check clearing, record keeping, credit analysis, and so forth) is like any other production process in a firm. If the bank produces desirable services at low cost and earns substantial income on its assets, it earns profits; if not, the bank suffers losses.

To make our analysis of the operation of a bank more concrete, we use a tool called a **T-account**. A T-account is a simplified balance sheet, with lines in the form of a T, that lists only the changes that occur in balance sheet items starting from some

initial balance sheet position. Let's say that Jane Brown has heard that the First National Bank provides excellent service, so she opens a checking account with a $100 bill. She now has a $100 checkable deposit at the bank, which shows up as a $100 liability on the bank's balance sheet. The bank now puts her $100 bill into its vault so that the bank's assets rise by the $100 increase in vault cash. The T-account for the bank looks like this:

FIRST NATIONAL BANK

Assets		Liabilities	
Vault cash	+$100	Checkable deposits	+$100

Since vault cash is also part of the bank's reserves, we can rewrite the T-account as follows:

Assets		Liabilities	
Reserves	+$100	Checkable deposits	+$100

Note that Jane Brown's opening of a checking account leads to *an increase in the bank's reserves equal to the increase in checkable deposits.*

If Jane had opened her account with a $100 check written on an account at another bank, say, the Second National Bank, we would get the same result. The initial effect on the T-account of the First National Bank is as follows:

Assets		Liabilities	
Cash items in process of collection	+$100	Checkable deposits	+$100

Checkable deposits increase by $100 as before, but now the First National Bank is owed $100 by the Second National Bank. This asset for the First National Bank is entered in the T-account as $100 of cash items in process of collection because the First National Bank will now try to collect the funds that it is owed. It could go directly to the Second National Bank and ask for payment of the funds, but if the two banks are in separate states, that would be a time-consuming and costly process. Instead, the First National Bank deposits the check in its account at the Fed, and the Fed collects the funds from the Second National Bank. The result is that the Fed transfers $100 of reserves from the Second National Bank to the First National Bank, and the final balance sheet positions of the two banks are as follows:

FIRST NATIONAL BANK

Assets		Liabilities	
Reserves	+$100	Checkable deposits	+$100

SECOND NATIONAL BANK

Assets		Liabilities	
Reserves	−$100	Checkable deposits	−$100

The process initiated by Jane Brown can be summarized as follows: When a check written on an account at one bank is deposited in another, the bank receiving the deposit gains reserves equal to the amount of the check, while the bank on which the check is written sees its reserves fall by the same amount. Therefore, ***when a bank receives additional deposits, it gains an equal amount of reserves; when it loses deposits, it loses an equal amount of reserves.***

Study Guide

T-accounts are used to study various topics throughout this text. Whenever you see a T-account, try to analyze what would happen if the opposite action were taken; for example, what would happen if Jane Brown decided to close her $100 account at the First National Bank by writing a $100 check and depositing it in a new checking account at the Second National Bank?

Now that you understand how banks gain and lose reserves, we can examine how a bank rearranges its balance sheet to make a profit when it experiences a change in its deposits. Let's return to the situation when the First National Bank has just received the extra $100 of checkable deposits. As you know, the bank is obliged to keep a certain fraction of its checkable deposits as required reserves. If the fraction (the required reserve ratio) is 10%, the First National Bank's required reserves have increased by $10, and we can rewrite its T-account as follows:

FIRST NATIONAL BANK

Assets		Liabilities	
Required reserves	+$10	Checkable deposits	+$100
Excess reserves	+$90		

Let's see how well the bank is doing as a result of the additional checkable deposits. Because reserves pay no interest, it has no income from the additional $100 of assets. But servicing the extra $100 of checkable deposits is costly, because the bank must keep records, pay tellers, return canceled checks, pay for check clearing, and so forth. The bank is taking a loss! The situation is even worse if the bank makes interest payments on the deposits, as with NOW accounts. If it is to make a profit, the bank must put to productive use all or part of the $90 of excess reserves it has available.

Let us assume that the bank chooses not to hold any excess reserves but to make loans instead. The T-account then looks like this:

Assets		Liabilities	
Required reserves	+$10	Checkable deposits	+$100
Loans	+$90		

The bank is now making a profit because it holds short-term liabilities such as checkable deposits and uses the proceeds to buy longer-term assets such as loans

with higher interest rates. As mentioned earlier, this process of asset transformation is frequently described by saying that banks are in the business of "borrowing short and lending long." For example, if the loans have an interest rate of 10% per year, the bank earns $9 in income from its loans over the year. If the $100 of checkable deposits is in a NOW account with a 5% interest rate and it costs another $3 per year to service the account, the cost per year of these deposits is $8. The bank's profit on the new deposits is then $1 per year (a 1% return on assets).

General Principles of Bank Management

Now that you have some idea of how a bank operates, let's look at how a bank manages its assets and liabilities in order to earn the highest possible profit. The bank manager has four primary concerns. The first is to make sure that the bank has enough ready cash to pay its depositors when there are **deposit outflows**, that is, when deposits are lost because depositors make withdrawals and demand payment. To keep enough cash on hand, the bank must engage in **liquidity management**, the acquisition of sufficiently liquid assets to meet the bank's obligations to depositors. Second, the bank manager must pursue an acceptably low level of risk by acquiring assets that have a low rate of default and by diversifying asset holdings (**asset management**). The third concern is to acquire funds at low cost (**liability management**). Finally, the manager must decide the amount of capital the bank should maintain and then acquire the needed capital (**capital adequacy management**).

To understand bank and other financial institution management fully, we must go beyond the general principles of bank asset and liability management described next and look in more detail at how a financial institution manages its assets. The two sections following this one provide an in-depth discussion of how a financial institution manages **credit risk**, the risk arising because borrowers may default, and how it manages **interest-rate risk**, the riskiness of earnings and returns on bank assets that results from interest-rate changes.

Liquidity Management and the Role of Reserves

Let us see how a typical bank, the First National Bank, can deal with deposit outflows that occur when its depositors withdraw cash from checking or savings accounts or write checks that are deposited in other banks. In the example that follows, we assume that the bank has ample excess reserves and that all deposits have the same required reserve ratio of 10% (the bank is required to keep 10% of its time and checkable deposits as reserves). Suppose that the First National Bank's initial balance sheet is as follows:

Assets		Liabilities	
Reserves	$20 million	Deposits	$100 million
Loans	$80 million	Bank capital	$ 10 million
Securities	$10 million		

The bank's required reserves are 10% of $100 million, or $10 million. Since it holds $20 million of reserves, the First National Bank has excess reserves of $10 million. If a deposit outflow of $10 million occurs, the bank's balance sheet becomes:

Assets		Liabilities	
Reserves	$10 million	Deposits	$90 million
Loans	$80 million	Bank capital	$10 million
Securities	$10 million		

The bank loses $10 million of deposits *and* $10 million of reserves, but since its required reserves are now 10% of only $90 million ($9 million), its reserves still exceed this amount by $1 million. In short, ***if a bank has ample reserves, a deposit outflow does not necessitate changes in other parts of its balance sheet.***

The situation is quite different when a bank holds insufficient excess reserves. Let's assume that instead of initially holding $10 million in excess reserves, the First National Bank makes loans of $10 million, so that it holds no excess reserves. Its initial balance sheet would be:

Assets		Liabilities	
Reserves	$10 million	Deposits	$100 million
Loans	$90 million	Bank capital	$ 10 million
Securities	$10 million		

When it suffers the $10 million deposit outflow, its balance sheet becomes:

Assets		Liabilities	
Reserves	$ 0	Deposits	$90 million
Loans	$90 million	Bank capital	$10 million
Securities	$10 million		

After $10 million has been withdrawn from deposits and hence reserves, the bank has a problem: It has a reserve requirement of 10% of $90 million, or $9 million, but it has no reserves! To eliminate this shortfall, the bank has four basic options. One is to acquire reserves to meet a deposit outflow by borrowing them from other banks in the federal funds market or by borrowing from corporations.[1] If the First National Bank acquires the $9 million shortfall in reserves by borrowing it from other banks or corporations, its balance sheet becomes:

Assets		Liabilities	
Reserves	$ 9 million	Deposits	$90 million
Loans	$90 million	Borrowings from other	
Securities	$10 million	banks or corporations	$ 9 million
		Bank capital	$10 million

[1]One way that the First National Bank can borrow from other banks and corporations is by selling negotiable certificates of deposit. This method for obtaining funds is discussed in the section on liability management.

The cost of this activity is the interest rate on these borrowings, such as the federal funds rate.

A second alternative is for the bank to sell some of its securities to help cover the deposit outflow. For example, it might sell $9 million of its securities and deposit the proceeds with the Fed, resulting in the following balance sheet:

Assets		Liabilities	
Reserves	$ 9 million	Deposits	$90 million
Loans	$90 million	Bank capital	$10 million
Securities	$ 1 million		

The bank incurs some brokerage and other transaction costs when it sells these securities. The U.S. government securities that are classified as secondary reserves are very liquid, so the transaction costs of selling them are quite modest. However, the other securities the bank holds are less liquid, and the transaction cost can be appreciably higher.

A third way that the bank can meet a deposit outflow is to acquire reserves by borrowing from the Fed. In our example, the First National Bank could leave its security and loan holdings the same and borrow $9 million in discount loans from the Fed. Its balance sheet would be:

Assets		Liabilities	
Reserves	$ 9 million	Deposits	$90 million
Loans	$90 million	Discount loans	
Securities	$10 million	from the Fed	$ 9 million
		Bank capital	$10 million

The cost associated with discount loans is the interest rate that must be paid to the Fed (called the **discount rate**).

Finally, a bank can acquire the $9 million of reserves to meet the deposit outflow by reducing its loans by this amount and depositing the $9 million it then receives with the Fed, thereby increasing its reserves by $9 million. This transaction changes the balance sheet as follows:

Assets		Liabilities	
Reserves	$ 9 million	Deposits	$90 million
Loans	$81 million	Bank capital	$10 million
Securities	$10 million		

The First National Bank is once again in good shape because its $9 million of reserves satisfies the reserve requirement.

However, this process of reducing its loans is the bank's costliest way of acquiring reserves when there is a deposit outflow. If the First National Bank has numerous short-term loans renewed at fairly short intervals, it can reduce its total amount of

loans outstanding fairly quickly by *calling in* loans—that is, by not renewing some loans when they come due. Unfortunately for the bank, this is likely to antagonize the customers whose loans are not being renewed because they have not done anything to deserve such treatment. Indeed, they are likely to take their business elsewhere in the future, a very costly consequence for the bank.

A second method for reducing its loans is for the bank to sell them off to other banks. Again, this is very costly because other banks do not personally know the customers who have taken out the loans and so may not be willing to buy the loans at their full value (This is just the lemons adverse selection problem described in Chapter 8.)

The foregoing discussion explains why banks hold excess reserves even though loans or securities earn a higher return. When a deposit outflow occurs, holding excess reserves allows the bank to escape the costs of (1) borrowing from other banks or corporations, (2) selling securities, (3) borrowing from the Fed, or (4) calling in or selling off loans. ***Excess reserves are insurance against the costs associated with deposit outflows. The higher the costs associated with deposit outflows, the more excess reserves banks will want to hold.***

Just as you and I would be willing to pay an insurance company to insure us against a casualty loss such as the theft of a car, a bank is willing to pay the cost of holding excess reserves (the opportunity cost, which is the earnings forgone by not holding income-earning assets such as loans or securities) in order to insure against losses due to deposit outflows. Because excess reserves, like insurance, have a cost, banks also take other steps to protect themselves; for example, they might shift their holdings of assets to more liquid securities (secondary reserves).

Study Guide

Bank management is easier to grasp if you put yourself in the banker's shoes and imagine what you would do in the situations described. To understand a bank's possible responses to deposit outflows, imagine how you as a banker might respond to two successive deposit outflows of $10 million.

Asset Management

Now that you understand why a bank has a need for liquidity, we can examine the basic strategy a bank pursues in managing its assets. To maximize its profits, a bank must simultaneously seek the highest returns possible on loans and securities, reduce risk, and make adequate provisions for liquidity by holding liquid assets. Banks try to accomplish these three goals in four basic ways.

First, banks try to find borrowers who will pay high interest rates and are unlikely to default on their loans. They seek out loan business by advertising their borrowing rates and by approaching corporations directly to solicit loans. It is up to the bank's loan officer to decide if potential borrowers are good credit risks who will make interest and principal payments on time (i.e., engage in screening to reduce the adverse selection problem). Typically, banks are conservative in their loan policies; the default rate is usually less than 1%. It is important, however, that banks not be so conservative that they miss out on attractive lending opportunities that earn high interest rates.

Second, banks try to purchase securities with high returns and low risk. Third, in managing their assets, banks must attempt to lower risk by diversifying. They accomplish this by purchasing many different types of assets (short- and long-term, U.S. Treasury, and municipal bonds) and approving many types of loans to a number of

customers. Banks that have not sufficiently sought the benefits of diversification often come to regret it later. For example, banks that had overspecialized in making loans to energy companies, real estate developers, or farmers suffered huge losses in the 1980s with the slump in energy, property, and farm prices. Indeed, many of these banks went broke because they had "put too many eggs in one basket."

Finally, the bank must manage the liquidity of its assets so that it can satisfy its reserve requirements without bearing huge costs. This means that it will hold liquid securities even if they earn a somewhat lower return than other assets. The bank must decide, for example, how much in excess reserves must be held to avoid costs from a deposit outflow. In addition, it will want to hold U.S. government securities as secondary reserves so that even if a deposit outflow forces some costs on the bank, these will not be terribly high. Again, it is not wise for a bank to be too conservative. If it avoids all costs associated with deposit outflows by holding only excess reserves, losses are suffered because reserves earn no interest, while the bank's liabilities are costly to maintain. The bank must balance its desire for liquidity against the increased earnings that can be obtained from less liquid assets such as loans.

Liability Management

Before the 1960s, liability management was a staid affair: For the most part, banks took their liabilities as fixed and spent their time trying to achieve an optimal mix of assets. There were two main reasons for the emphasis on asset management. First, over 60% of the sources of bank funds were obtained through checkable (demand) deposits that by law could not pay any interest. Thus banks could not actively compete with one another for these deposits by paying interest on them, and so their amount was effectively a given for an individual bank. Second, because the markets for making overnight loans between banks were not well developed, banks rarely borrowed from other banks to meet their reserve needs.

Starting in the 1960s, however, large banks (called **money center banks**) in key financial centers, such as New York, Chicago, and San Francisco, began to explore ways in which the liabilities on their balance sheets could provide them with reserves and liquidity. This led to an expansion of overnight loan markets, such as the federal funds market, and the development of new financial instruments such as negotiable CDs (first developed in 1961), which enabled money center banks to acquire funds quickly.[2]

This new flexibility in liability management meant that banks could take a different approach to bank management. They no longer needed to depend on checkable deposits as the primary source of bank funds and as a result no longer treated their sources of funds (liabilities) as given. Instead, they aggressively set target goals for their asset growth and tried to acquire funds (by issuing liabilities) as they were needed.

For example, today, when a money center bank finds an attractive loan opportunity, it can acquire funds by selling a negotiable CD. Or, if it has a reserve shortfall, funds can be borrowed from another bank in the federal funds market without incurring high transaction costs. The federal funds market can also be used to finance loans. Because of the increased importance of liability management, most banks now

[2]Because small banks are not as well known as money center banks and so might be a higher credit risk, they find it harder to raise funds in the negotiable CD market. Hence they do not engage nearly as actively in liability management.

manage both sides of the balance sheet together in a so-called asset–liability management (ALM) committee.

The emphasis on liability management explains some of the important changes over the past three decades in the composition of banks' balance sheets. While negotiable CDs and bank borrowings have greatly increased in importance as a source of bank funds in recent years (rising from 2% of bank liabilities in 1960 to 42% by the end of 2002), checkable deposits have decreased in importance (from 61% of bank liabilities in 1960 to 9% in 2002). Newfound flexibility in liability management and the search for higher profits have also stimulated banks to increase the proportion of their assets held in loans, which earn higher income (from 46% of bank assets in 1960 to 64% in 2002).

Capital Adequacy Management

Banks have to make decisions about the amount of capital they need to hold for three reasons. First, bank capital helps prevents *bank failure*, a situation in which the bank cannot satisfy its obligations to pay its depositors and other creditors and so goes out of business. Second, the amount of capital affects returns for the owners (equity holders) of the bank. And third, a minimum amount of bank capital (bank capital requirements) is required by regulatory authorities.

How Bank Capital Helps Prevent Bank Failure. Let's consider two banks with identical balance sheets, except that the High Capital Bank has a ratio of capital to assets of 10% while the Low Capital Bank has a ratio of 4%.

HIGH CAPITAL BANK

Assets		Liabilities	
Reserves	$10 million	Deposits	$90 million
Loans	$90 million	Bank capital	$10 million

LOW CAPITAL BANK

Assets		Liabilities	
Reserves	$10 million	Deposits	$96 million
Loans	$90 million	Bank capital	$ 4 million

Suppose that both banks get caught up in the euphoria of the telecom market, only to find that $5 million of their telecom loans became worthless later. When these bad loans are written off (valued at zero), the total value of assets declines by $5 million, and so bank capital, which equals total assets minus liabilities, also declines by $5 million. The balance sheets of the two banks now look like this:

HIGH CAPITAL BANK

Assets		Liabilities	
Reserves	$10 million	Deposits	$90 million
Loans	$85 million	Bank capital	$ 5 million

LOW CAPITAL BANK

Assets		Liabilities	
Reserves	$10 million	Deposits	$96 million
Loans	$85 million	Bank capital	−$ 1 million

The High Capital Bank takes the $5 million loss in stride because its initial cushion of $10 million in capital means that it still has a positive net worth (bank capital) of $5 million after the loss. The Low Capital Bank, however, is in big trouble. Now the value of its assets has fallen below its liabilities, and its net worth is now −$1 million. Because the bank has a negative net worth, it is insolvent: It does not have sufficient assets to pay off all holders of its liabilities (creditors). When a bank becomes insolvent, government regulators close the bank, its assets are sold off, and its managers are fired. Since the owners of the Low Capital Bank will find their investment wiped out, they would clearly have preferred the bank to have had a large enough cushion of bank capital to absorb the losses, as was the case for the High Capital Bank. We therefore see an important rationale for a bank to maintain a high level of capital: ***A bank maintains bank capital to lessen the chance that it will become insolvent.***

How the Amount of Bank Capital Affects Returns to Equity Holders. Because owners of a bank must know whether their bank is being managed well, they need good measures of bank profitability. A basic measure of bank profitability is the **return on assets (ROA)**, the net profit after taxes per dollar of assets:

$$\text{ROA} = \frac{\text{net profit after taxes}}{\text{assets}}$$

The return on assets provides information on how efficiently a bank is being run, because it indicates how much profits are generated on average by each dollar of assets.

However, what the bank's owners (equity holders) care about most is how much the bank is earning on their equity investment. This information is provided by the other basic measure of bank profitability, the **return on equity (ROE)**, the net profit after taxes per dollar of equity (bank) capital:

$$\text{ROE} = \frac{\text{net profit after taxes}}{\text{equity capital}}$$

There is a direct relationship between the return on assets (which measures how efficiently the bank is run) and the return on equity (which measures how well the owners are doing on their investment). This relationship is determined by the so-called **equity multiplier (EM)**, which is the amount of assets per dollar of equity capital:

$$\text{EM} = \frac{\text{assets}}{\text{equity capital}}$$

To see this, we note that:

$$\frac{\text{net profit after taxes}}{\text{equity capital}} = \frac{\text{net profit after taxes}}{\text{assets}} \times \frac{\text{assets}}{\text{equity capital}}$$

which, using our definitions, yields:

$$\text{ROE} = \text{ROA} \times \text{EM} \tag{1}$$

The formula in Equation 1 tells us what happens to the return on equity when a bank holds a smaller amount of capital (equity) for a given amount of assets. As we have seen, the High Capital Bank initially has $100 million of assets and $10 million of equity, which gives it an equity multiplier of 10 (= $100 million/$10 million). The

Low Capital Bank, by contrast, has only $4 million of equity, so its equity multiplier is higher, equaling 25 (= $100 million/$4 million). Suppose that these banks have been equally well run so that they both have the same return on assets, 1%. The return on equity for the High Capital Bank equals 1% × 10 = 10%, while the return on equity for the Low Capital Bank equals 1% × 25 = 25%. The equity holders in the Low Capital Bank are clearly a lot happier than the equity holders in the High Capital Bank because they are earning more than twice as high a return. We now see why owners of a bank may not want it to hold too much capital. ***Given the return on assets, the lower the bank capital, the higher the return for the owners of the bank.***

Trade-off Between Safety and Returns to Equity Holders. We now see that bank capital has benefits and costs. Bank capital benefits the owners of a bank in that it makes their investment safer by reducing the likelihood of bankruptcy. But bank capital is costly because the higher it is, the lower will be the return on equity for a given return on assets. In determining the amount of bank capital, managers must decide how much of the increased safety that comes with higher capital (the benefit) they are willing to trade off against the lower return on equity that comes with higher capital (the cost).

In more uncertain times, when the possibility of large losses on loans increases, bank managers might want to hold more capital to protect the equity holders. Conversely, if they have confidence that loan losses won't occur, they might want to reduce the amount of bank capital, have a high equity multiplier, and thereby increase the return on equity.

Bank Capital Requirements. Banks also hold capital because they are required to do so by regulatory authorities. Because of the high costs of holding capital for the reasons just described, bank managers often want to hold less bank capital relative to assets than is required by the regulatory authorities. In this case, the amount of bank capital is determined by the bank capital requirements. We discuss the details of bank capital requirements and why they are such an important part of bank regulation in Chapter 11.

Application

Strategies for Managing Bank Capital

Suppose that as the manager of the First National Bank, you have to make decisions about the appropriate amount of bank capital. Looking at the balance sheet of the bank, which like the High Capital Bank has a ratio of bank capital to assets of 10% ($10 million of capital and $100 million of assets), you are concerned that the large amount of bank capital is causing the return on equity to be too low. You conclude that the bank has a capital surplus and should increase the equity multiplier to increase the return on equity. What should you do?

To lower the amount of capital relative to assets and raise the equity multiplier, you can do any of three things: (1) You can reduce the amount of bank capital by buying back some of the bank's stock. (2) You can reduce the bank's capital by paying out higher dividends to its stockholders, thereby reducing the bank's retained earnings. (3) You can keep bank capital constant but increase the bank's assets by acquiring new funds—say, by issuing CDs—and then seeking out loan business or purchasing more securities with these new funds. Because you think that it would enhance your position with the

stockholders, you decide to pursue the second alternative and raise the dividend on the First National Bank stock.

Now suppose that the First National Bank is in a similar situation to the Low Capital Bank and has a ratio of bank capital to assets of 4%. You now worry that the bank is short on capital relative to assets because it does not have a sufficient cushion to prevent bank failure. To raise the amount of capital relative to assets, you now have the following three choices: (1) You can raise capital for the bank by having it issue equity (common stock). (2) You can raise capital by reducing the bank's dividends to shareholders, thereby increasing retained earnings that it can put into its capital account. (3) You can keep capital at the same level but reduce the bank's assets by making fewer loans or by selling off securities and then using the proceeds to reduce its liabilities. Suppose that raising bank capital is not easy to do at the current time because capital markets are tight or because shareholders will protest if their dividends are cut. Then you might have to choose the third alternative and decide to shrink the size of the bank.

In past years, many banks experienced capital shortfalls and had to restrict asset growth, as you might have to do if the First National Bank were short of capital. The important consequences of this for the credit markets are discussed in the application that follows.

Application

Did the Capital Crunch Cause a Credit Crunch in the Early 1990s?

During the 1990–1991 recession and the year following, there occurred a slowdown in the growth of credit that was unprecedented in the post–World War II era. Many economists and politicians have claimed that there was a "credit crunch" during this period in which credit was hard to get, and as a result the performance of the economy in 1990–1992 was very weak. Was the slowdown in credit growth a manifestation of a credit crunch, and if so, what caused it?

Our analysis of how a bank manages bank capital suggests that a credit crunch was likely to have occurred in 1990–1992 and that it was caused at least in part by the so-called capital crunch in which shortfalls of bank capital led to slower credit growth.

The period of the late 1980s saw a boom and then a major bust in the real estate market that led to huge losses for banks on their real estate loans. As our example of how bank capital helps prevent bank failures demonstrates, the loan losses caused a substantial fall in the amount of bank capital. At the same time, regulators were raising capital requirements (a subject discussed in Chapter 11). The resulting capital shortfalls meant that banks had to either raise new capital or restrict their asset growth by cutting back on lending. Because of the weak economy at the time, raising new capital was extremely difficult for banks, so they chose the latter course. Banks did restrict their lending, and borrowers found it harder to obtain loans, leading to complaints from banks' customers. Only with the stronger recovery of the economy in 1993, helped by a low-interest-rate policy at the Federal Reserve, did these complaints subside.

Managing Credit Risk

As seen in the earlier discussion of general principles of asset management, banks and also other financial institutions must make successful loans that are paid back in full (and so subject the institution to little credit risk) in order to earn high profits. The economic concepts of adverse selection and moral hazard (introduced in Chapters 2 and 8) provide a framework for understanding the principles that financial institutions have to follow to reduce credit risk and make successful loans.[3]

Adverse selection in loan markets occurs because bad credit risks (those most likely to default on their loans) are the ones who usually line up for loans; in other words, those who are most likely to produce an *adverse* outcome are the most likely to be *selected*. Borrowers with very risky investment projects have much to gain if their projects are successful, and so they are the most eager to obtain loans. Clearly, however, they are the least desirable borrowers because of the greater possibility that they will be unable to pay back their loans.

Moral hazard exists in loan markets because borrowers may have incentives to engage in activities that are undesirable from the lender's point of view. In such situations, it is more likely that the lender will be subjected to the *hazard* of default. Once borrowers have obtained a loan, they are more likely to invest in high-risk investment projects—projects that pay high returns to the borrowers if successful. The high risk, however, makes it less likely that they will be able to pay the loan back.

To be profitable, financial institutions must overcome the adverse selection and moral hazard problems that make loan defaults more likely. The attempts of financial institutions to solve these problems help explain a number of principles for managing credit risk: screening and monitoring, establishment of long-term customer relationships, loan commitments, collateral and compensating balance requirements, and credit rationing.

Screening and Monitoring

Asymmetric information is present in loan markets because lenders have less information about the investment opportunities and activities of borrowers than borrowers do. This situation leads to two information-producing activities by banks and other financial institutions—screening and monitoring. Indeed, Walter Wriston, a former head of Citicorp, the largest bank corporation in the United States, was often quoted as stating that the business of banking is the production of information.

Screening. Adverse selection in loan markets requires that lenders screen out the bad credit risks from the good ones so that loans are profitable to them. To accomplish effective screening, lenders must collect reliable information from prospective borrowers. Effective screening and information collection together form an important principle of credit risk management.

When you apply for a consumer loan (such as a car loan or a mortgage to purchase a house), the first thing you are asked to do is fill out forms that elicit a great deal of information about your personal finances. You are asked about your salary, your bank accounts and other assets (such as cars, insurance policies, and furnishings), and your outstanding loans; your record of loan, credit card, and charge

[3]Other financial intermediaries, such as insurance companies, pension funds, and finance companies, also make private loans, and the credit risk management principles we outline here apply to them as well.

account repayments; the number of years you've worked and who your employers have been. You also are asked personal questions such as your age, marital status, and number of children. The lender uses this information to evaluate how good a credit risk you are by calculating your "credit score," a statistical measure derived from your answers that predicts whether you are likely to have trouble making your loan payments. Deciding on how good a risk you are cannot be entirely scientific, so the lender must also use judgment. The loan officer, whose job is to decide whether you should be given the loan, might call your employer or talk to some of the personal references you supplied. The officer might even make a judgment based on your demeanor or your appearance. (This is why most people dress neatly and conservatively when they go to a bank to apply for a loan.)

The process of screening and collecting information is similar when a financial institution makes a business loan. It collects information about the company's profits and losses (income) and about its assets and liabilities. The lender also has to evaluate the likely future success of the business. So in addition to obtaining information on such items as sales figures, a loan officer might ask questions about the company's future plans, how the loan will be used, and the competition in the industry. The officer may even visit the company to obtain a firsthand look at its operations. The bottom line is that, whether for personal or business loans, bankers and other financial institutions need to be nosy.

Specialization in Lending. One puzzling feature of bank lending is that a bank often specializes in lending to local firms or to firms in particular industries, such as energy. In one sense, this behavior seems surprising, because it means that the bank is not diversifying its portfolio of loans and thus is exposing itself to more risk. But from another perspective, such specialization makes perfect sense. The adverse selection problem requires that the bank screen out bad credit risks. It is easier for the bank to collect information about local firms and determine their creditworthiness than to collect comparable information on firms that are far away. Similarly, by concentrating its lending on firms in specific industries, the bank becomes more knowledgeable about these industries and is therefore better able to predict which firms will be able to make timely payments on their debt.

Monitoring and Enforcement of Restrictive Covenants. Once a loan has been made, the borrower has an incentive to engage in risky activities that make it less likely that the loan will be paid off. To reduce this moral hazard, financial institutions must adhere to the principle for managing credit risk that a lender should write provisions (restrictive covenants) into loan contracts that restrict borrowers from engaging in risky activities. By monitoring borrowers' activities to see whether they are complying with the restrictive covenants and by enforcing the covenants if they are not, lenders can make sure that borrowers are not taking on risks at their expense. The need for banks and other financial institutions to engage in screening and monitoring explains why they spend so much money on auditing and information-collecting activities.

Long-Term Customer Relationships

An additional way for banks and other financial institutions to obtain information about their borrowers is through long-term customer relationships, another important principle of credit risk management.

If a prospective borrower has had a checking or savings account or other loans with a bank over a long period of time, a loan officer can look at past activity on the accounts and learn quite a bit about the borrower. The balances in the checking and

savings accounts tell the banker how liquid the potential borrower is and at what time of year the borrower has a strong need for cash. A review of the checks the borrower has written reveals the borrower's suppliers. If the borrower has borrowed previously from the bank, the bank has a record of the loan payments. Thus long-term customer relationships reduce the costs of information collection and make it easier to screen out bad credit risks.

The need for monitoring by lenders adds to the importance of long-term customer relationships. If the borrower has borrowed from the bank before, the bank has already established procedures for monitoring that customer. Therefore, the costs of monitoring long-term customers are lower than those for new customers.

Long-term relationships benefit the customers as well as the bank. A firm with a previous relationship will find it easier to obtain a loan at a low interest rate because the bank has an easier time determining if the prospective borrower is a good credit risk and incurs fewer costs in monitoring the borrower.

A long-term customer relationship has another advantage for the bank. No bank can think of every contingency when it writes a restrictive covenant into a loan contract; there will always be risky borrower activities that are not ruled out. However, what if a borrower wants to preserve a long-term relationship with a bank because it will be easier to get future loans at low interest rates? The borrower then has the incentive to avoid risky activities that would upset the bank, even if restrictions on these risky activities are not specified in the loan contract. Indeed, if a bank doesn't like what a borrower is doing even when the borrower isn't violating any restrictive covenants, it has some power to discourage the borrower from such activity: The bank can threaten not to let the borrower have new loans in the future. Long-term customer relationships therefore enable banks to deal with even unanticipated moral hazard contingencies.

Loan Commitments

Banks also create long-term relationships and gather information by issuing **loan commitments** to commercial customers. A loan commitment is a bank's commitment (for a specified future period of time) to provide a firm with loans up to a given amount at an interest rate that is tied to some market interest rate. The majority of commercial and industrial loans are made under the loan commitment arrangement. The advantage for the firm is that it has a source of credit when it needs it. The advantage for the bank is that the loan commitment promotes a long-term relationship, which in turn facilitates information collection. In addition, provisions in the loan commitment agreement require that the firm continually supply the bank with information about the firm's income, asset and liability position, business activities, and so on. A loan commitment arrangement is a powerful method for reducing the bank's costs for screening and information collection.

Collateral and Compensating Balances

Collateral requirements for loans are important credit risk management tools. Collateral, which is property promised to the lender as compensation if the borrower defaults, lessens the consequences of adverse selection because it reduces the lender's losses in the case of a loan default. If a borrower defaults on a loan, the lender can sell the collateral and use the proceeds to make up for its losses on the loan. One particular form of collateral required when a bank makes commercial loans is called **compensating balances**: A firm receiving a loan must keep a required minimum amount of funds in a checking account at the bank. For example, a business getting a $10 million loan may be required to keep compensating balances of at least $1 million in its checking account at the bank. This $1 million in compensating balances can then be taken by the bank to make up some of the losses on the loan if the borrower defaults.

A required minimum amount of funds that a firm receiving a loan must keep in a checking account at the lending bank.

Besides serving as collateral, compensating balances help increase the likelihood that a loan will be paid off. They do this by helping the bank monitor the borrower and consequently reduce moral hazard. Specifically, by requiring the borrower to use a checking account at the bank, the bank can observe the firm's check payment practices, which may yield a great deal of information about the borrower's financial condition. For example, a sustained drop in the borrower's checking account balance may signal that the borrower is having financial trouble, or account activity may suggest that the borrower is engaging in risky activities; perhaps a change in suppliers means that the borrower is pursuing new lines of business. Any significant change in the borrower's payment procedures is a signal to the bank that it should make inquiries. Compensating balances therefore make it easier for banks to monitor borrowers more effectively and are another important credit risk management tool.

Credit Rationing

Another way in which financial institutions deal with adverse selection and moral hazard is through **credit rationing**: refusing to make loans even though borrowers are willing to pay the stated interest rate or even a higher rate. Credit rationing takes two forms. The first occurs when a lender refuses to make a loan *of any amount* to a borrower, even if the borrower is willing to pay a higher interest rate. The second occurs when a lender is willing to make a loan but restricts the size of the loan to less than the borrower would like.

At first you might be puzzled by the first type of credit rationing. After all, even if the potential borrower is a credit risk, why doesn't the lender just extend the loan but at a higher interest rate? The answer is that adverse selection prevents this solution. Individuals and firms with the riskiest investment projects are exactly those that are willing to pay the highest interest rates. If a borrower took on a high-risk investment and succeeded, the borrower would become extremely rich. But a lender wouldn't want to make such a loan precisely because the investment risk is high; the likely outcome is that the borrower will *not* succeed and the lender will not be paid back. Charging a higher interest rate just makes adverse selection worse for the lender; that is, it increases the likelihood that the lender is lending to a bad credit risk. The lender would therefore rather not make any loans at a higher interest rate; instead, it would engage in the first type of credit rationing and would turn down loans.

Financial institutions engage in the second type of credit rationing to guard against moral hazard: They grant loans to borrowers, but not loans as large as the borrowers want. Such credit rationing is necessary because the larger the loan, the greater the benefits from moral hazard. If a bank gives you a $1,000 loan, for example, you are likely to take actions that enable you to pay it back because you don't want to hurt your credit rating for the future. However, if the bank lends you $10 million, you are more likely to fly down to Rio to celebrate. The larger your loan, the greater your incentives to engage in activities that make it less likely that you will repay the loan. Since more borrowers repay their loans if the loan amounts are small, financial institutions ration credit by providing borrowers with smaller loans than they seek.

Managing Interest-Rate Risk

With the increased volatility of interest rates that occurred in the 1980s, banks and other financial institutions became more concerned about their exposure to interest-rate risk, the riskiness of earnings and returns that is associated with changes in

interest rates. To see what interest-rate risk is all about, let's again take a look at the First National Bank, which has the following balance sheet:

FIRST NATIONAL BANK

Assets		Liabilities	
Rate-sensitive assets	$20 million	Rate-sensitive liabilities	$50 million
Variable-rate and short-term loans		Variable-rate CDs	
Short-term securities		Money market deposit accounts	
Fixed-rate assets	$80 million	Fixed-rate liabilities	$50 million
Reserves		Checkable deposits	
Long-term loans		Savings deposits	
Long-term securities		Long-term CDs	
		Equity capital	

A total of $20 million of its assets are rate-sensitive, with interest rates that change frequently (at least once a year), and $80 million of its assets are fixed-rate, with interest rates that remain unchanged for a long period (over a year). On the liabilities side, the First National Bank has $50 million of rate-sensitive liabilities and $50 million of fixed-rate liabilities. Suppose that interest rates rise by 5 percentage points on average, from 10% to 15%. The income on the assets rises by $1 million (= 5% × $20 million of rate-sensitive assets), while the payments on the liabilities rise by $2.5 million (= 5% × $50 million of rate-sensitive liabilities). The First National Bank's profits now decline by $1.5 million (= $1 million − $2.5 million). Conversely, if interest rates fall by 5 percentage points, similar reasoning tells us that the First National Bank's profits rise by $1.5 million. This example illustrates the following point: ***If a bank has more rate-sensitive liabilities than assets, a rise in interest rates will reduce bank profits and a decline in interest rates will raise bank profits.***

Gap and Duration Analysis

The sensitivity of bank profits to changes in interest rates can be measured more directly using **gap analysis**, in which the amount of rate-sensitive liabilities is subtracted from the amount of rate-sensitive assets. In our example, this calculation (called the "gap") is −$30 million (= $20 million − $50 million). By multiplying the gap times the change in the interest rate, we can immediately obtain the effect on bank profits. For example, when interest rates rise by 5 percentage points, the change in profits is 5% × −$30 million, which equals −$1.5 million, as we saw.

The analysis we just conducted is known as *basic gap analysis*, and it can be refined in two ways. Clearly, not all assets and liabilities in the fixed-rate category have the same maturity. One refinement, the *maturity bucket approach*, is to measure the gap for several maturity subintervals, called *maturity buckets*, so that effects of interest-rate changes over a multiyear period can be calculated. The second refinement, called *standardized gap analysis*, accounts for the differing degrees of rate sensitivity for different rate-sensitive assets and liabilities.

An alternative method for measuring interest-rate risk, called **duration analysis**, examines the sensitivity of the market value of the bank's total assets and liabilities to changes in interest rates. Duration analysis is based on what is known as Macaulay's

concept of *duration*, which measures the average lifetime of a security's stream of payments.[4] Duration is a useful concept because it provides a good approximation of the sensitivity of a security's market value to a change in its interest rate:

$$\text{percent change in market value of security} \approx -\text{percentage-point change in interest rate} \times \text{duration in years}$$

where $\approx$ denotes "approximately equals."

Duration analysis involves using the average (weighted) duration of a financial institution's assets and of its liabilities to see how its net worth responds to a change in interest rates. Going back to our example of the First National Bank, suppose that the average duration of its assets is three years (that is, the average lifetime of the stream of payments is three years), while the average duration of its liabilities is two years. In addition, the First National Bank has $100 million of assets and $90 million of liabilities, so its bank capital is 10% of assets. With a 5-percentage-point increase in interest rates, the market value of the bank's assets falls by 15% ($= -5\% \times 3$ years), a decline of $15 million on the $100 million of assets. However, the market value of the liabilities falls by 10% ($= -5\% \times 2$ years), a decline of $9 million on the $90 million of liabilities. The net result is that the net worth (the market value of the assets minus the liabilities) has declined by $6 million, or 6% of the total original asset value. Similarly, a 5-percentage-point decline in interest rates increases the net worth of the First National Bank by 6% of the total asset value.

As our example makes clear, both duration analysis and gap analysis indicate that the First National Bank will suffer if interest rates rise but will gain if they fall. Duration analysis and gap analysis are thus useful tools for telling a manager of a financial institution its degree of exposure to interest-rate risk.

Application

Strategies for Managing Interest-Rate Risk

Suppose that as the manager of the First National Bank, you have done a duration and gap analysis for the bank as discussed in the text. Now you need to decide what alternative strategies you should pursue to manage the interest-rate risk.

If you firmly believe that interest rates will fall in the future, you may be willing to take no action because you know that the bank has more rate-sensitive liabilities than rate-sensitive assets and so will benefit from the

---

[4] Algebraically, Macaulay's duration, D, is defined as:

$$D = \sum_{\tau=1}^{N} \tau \frac{CP_\tau}{(1+i)^\tau} \Big/ \sum_{\tau=1}^{N} \frac{CP_\tau}{(1+i)^\tau}$$

where τ = time until cash payment is made
CP_τ = cash payment (interest plus principal) at time τ
i = interest rate
N = time to maturity of the security

For a more detailed discussion of duration gap analysis using the concept of Macaulay's duration, you can look at an appendix to this chapter that is on this book's web site at www.aw.com/mishkin.

expected interest-rate decline. However, you also realize that the First National Bank is subject to substantial interest-rate risk because there is always a possibility that interest rates will rise rather than fall. What should you do to eliminate this interest-rate risk? One thing you could do is to shorten the duration of the bank's assets to increase their rate sensitivity. Alternatively, you could lengthen the duration of the liabilities. By this adjustment of the bank's assets and liabilities, the bank's income will be less affected by interest-rate swings.

One problem with eliminating the First National Bank's interest-rate risk by altering the balance sheet is that doing so might be very costly in the short run. The bank may be locked in to assets and liabilities of particular durations because of where its expertise lies. Fortunately, recently developed financial instruments known as financial derivatives—financial forwards and futures, options, and swaps—can help the bank reduce its interest-rate risk exposure but do not require that the bank rearrange its balance sheet. We discuss these instruments and how banks and other financial institutions can use them to manage interest-rate risk in Chapter 13.

Off-Balance-Sheet Activities

Although asset and liability management has traditionally been the major concern of banks, in the more competitive environment of recent years banks have been aggressively seeking out profits by engaging in off-balance-sheet activities.[5] **Off-balance-sheet activities** involve trading financial instruments and generating income from fees and loan sales, activities that affect bank profits but do not appear on bank balance sheets. Indeed, off-balance-sheet activities have been growing in importance for banks: The income from these activities as a percentage of assets has nearly doubled since 1980.

Loan Sales

One type of off-balance-sheet activity that has grown in importance in recent years involves income generated by loan sales. A **loan sale**, also called a *secondary loan participation*, involves a contract that sells all or part of the cash stream from a specific loan and thereby removes the loan from the bank's balance sheet. Banks earn profits by selling loans for an amount slightly greater than the amount of the original loan. Because the high interest rate on these loans makes them attractive, institutions are willing to buy them, even though the higher price means that they earn a slightly lower interest rate than the original interest rate on the loan, usually on the order of 0.15 percentage point.

Generation of Fee Income

Another type of off-balance-sheet activity involves the generation of income from fees that banks receive for providing specialized services to their customers, such as making foreign exchange trades on a customer's behalf, servicing a mortgage-backed security by collecting interest and principal payments and then paying them out, guaranteeing debt securities such as banker's acceptances (by which the bank promises

[5]Managers of financial institutions also need to know how well their banks are doing at any point in time. A second appendix to this chapter discusses how bank performance is measured; it can be found on the book's web site at www.aw.com/mishkin.

to make interest and principal payments if the party issuing the security cannot), and providing backup lines of credit. There are several types of backup lines of credit. We have already mentioned the most important, the loan commitment, under which for a fee the bank agrees to provide a loan at the customer's request, up to a given dollar amount, over a specified period of time. Credit lines are also now available to bank depositors with "overdraft privileges"—these bank customers can write checks in excess of their deposit balances and, in effect, write themselves a loan. Other lines of credit for which banks get fees include standby letters of credit to back up issues of commercial paper and other securities and credit lines (called *note issuance facilities*, NIFs, and *revolving underwriting facilities*, RUFs) for underwriting Euronotes, which are medium-term Eurobonds.

Off-balance-sheet activities involving guarantees of securities and backup credit lines increase the risk a bank faces. Even though a guaranteed security does not appear on a bank balance sheet, it still exposes the bank to default risk: If the issuer of the security defaults, the bank is left holding the bag and must pay off the security's owner. Backup credit lines also expose the bank to risk because the bank may be forced to provide loans when it does not have sufficient liquidity or when the borrower is a very poor credit risk.

Trading Activities and Risk Management Techniques

We have already mentioned that banks' attempts to manage interest-rate risk led them to trading in financial futures, options for debt instruments, and interest-rate swaps. Banks engaged in international banking also conduct transactions in the foreign exchange market. All transactions in these markets are off-balance-sheet activities because they do not have a direct effect on the bank's balance sheet. Although bank trading in these markets is often directed toward reducing risk or facilitating other bank business, banks also try to outguess the markets and engage in speculation. This speculation can be a very risky business and indeed has led to bank insolvencies, the most dramatic being the failure of Barings, a British bank, in 1995.

www.federalreserve.gov/boarddocs/SupManual/default.htm#trading

The Federal Reserve Bank Trading and Capital Market Activities Manual offers an in-depth discussion of a wide range of risk management issues encountered in trading operations.

Trading activities, although often highly profitable, are dangerous because they make it easy for financial institutions and their employees to make huge bets quickly. A particular problem for management of trading activities is that the principal-agent problem, discussed in Chapter 8, is especially severe. Given the ability to place large bets, a trader (the agent), whether she trades in bond markets, in foreign exchange markets or in financial derivatives, has an incentive to take on excessive risks: If her trading strategy leads to large profits, she is likely to receive a high salary and bonuses, but if she takes large losses, the financial institution (the principal) will have to cover them. As the Barings Bank failure in 1995 so forcefully demonstrated, a trader subject to the principal–agent problem can take an institution that is quite healthy and drive it into insolvency very fast (see Box 1).

To reduce the principal–agent problem, managers of financial institutions must set up internal controls to prevent debacles like the one at Barings. Such controls include the complete separation of the people in charge of trading activities from those in charge of the bookkeeping for trades. In addition, managers must set limits on the total amount of traders' transactions and on the institution's risk exposure. Managers must also scrutinize risk assessment procedures using the latest computer technology. One such method involves the so-called value-at-risk approach. In this approach, the institution develops a statistical model with which it can calculate the

Box 1: Global

Barings, Daiwa, Sumitomo, and Allied Irish

Rogue Traders and the Principal–Agent Problem. The demise of Barings, a venerable British bank over a century old, is a sad morality tale of how the principal–agent problem operating through a rogue trader can take a financial institution that has a healthy balance sheet one month and turn it into an insolvent tragedy the next.

In July 1992, Nick Leeson, Barings's new head clerk at its Singapore branch, began to speculate on the Nikkei, the Japanese version of the Dow Jones stock index. By late 1992, Leeson had suffered losses of $3 million, which he hid from his superiors by stashing the losses in a secret account. He even fooled his superiors into thinking he was generating large profits, thanks to a failure of internal controls at his firm, which allowed him to execute trades on the Singapore exchange *and* oversee the bookkeeping of those trades. (As anyone who runs a cash business, such as a bar, knows, there is always a lower likelihood of fraud if more than one person handles the cash. Similarly for trading operations, you never mix management of the back room with management of the front room; this principle was grossly violated by Barings management.)

Things didn't get better for Leeson, who by late 1994 had losses exceeding $250 million. In January and February 1995, he bet the bank. On January 17, 1995, the day of the Kobe earthquake, he lost $75 million, and by the end of the week had lost more than $150 million. When the stock market declined on February 23, leaving him with a further loss of $250 million, he called it quits and fled Singapore. Three days later, he turned himself in at the Frankfurt airport. By the end of his wild ride, Leeson's losses, $1.3 billion in all, ate up Barings's capital and caused the bank to fail. Leeson was subsequently convicted and sent to jail in Singapore for his activities. He was released in 1999 and apologized for his actions.

Our asymmetric information analysis of the principal–agent problem explains Leeson's behavior and the danger of Barings's management lapse. By letting Leeson control both his own trades and the back room, it increased asymmetric information, because it reduced the principal's (Barings's) knowledge about Leeson's trading activities. This lapse increased the moral hazard incentive for him to take risks at the bank's expense, as he was now less likely to be caught. Furthermore, once he had experienced large losses, he had even greater incentives to take on even higher risk because if his bets worked out, he could reverse his losses and keep in good standing with the company, whereas if his bets soured, he had little to lose since he was out of a job anyway. Indeed, the bigger his losses, the more he had to gain by bigger bets, which explains the escalation of the amount of his trades as his losses mounted. If Barings's managers had understood the principal–agent problem, they would have been more vigilant at finding out what Leeson was up to, and the bank might still be here today.

Unfortunately, Nick Leeson is no longer a rarity in the rogue traders' billionaire club, those who have lost more than $1 billion. Over 11 years, Toshihide Iguchi, an officer in the New York branch of Daiwa Bank, also had control of both the bond trading operation and the back room, and he racked up $1.1 billion in losses over the period. In July 1995, Iguchi disclosed his losses to his superiors, but the management of the bank did not disclose them to its regulators. The result was that Daiwa was slapped with a $340 million fine and the bank was thrown out of the country by U.S. bank regulators. Yasuo Hamanaka is another member of the billionaire club. In July 1996, he topped Leeson's and Iguchi's record, losing $2.6 billion for his employer, the Sumitomo Corporation, one of Japan's top trading companies. John Rusnak lost *only* $691 million for his bank, Allied Irish Banks, over the period from 1997 until he was caught in February 2002. The moral of these stories is that management of firms engaged in trading activities must reduce the principal–agent problem by closely monitoring their traders' activities, or the rogues' gallery will continue to grow.

maximum loss that its portfolio is likely to sustain over a given time interval, dubbed the value at risk, or VAR. For example, a bank might estimate that the maximum loss it would be likely to sustain over one day with a probability of 1 in 100 is $1 million; the $1 million figure is the bank's calculated value at risk. Another approach is called "stress testing." In this approach, a manager asks models what would happen if a doomsday scenario occurs; that is, she looks at the losses the institution would sustain if an unusual combination of bad events occurred. With the value-at-risk approach and stress testing, a financial institution can assess its risk exposure and take steps to reduce it.

Because of the increased risk that banks are facing from their off-balance-sheet activities, U.S. bank regulators have become concerned about increased risk from banks' off-balance-sheet activities and, as we will see in Chapter 11, are encouraging banks to pay increased attention to risk management. In addition, the Bank for International Settlements is developing additional bank capital requirements based on value-at-risk calculations for a bank's trading activities.

Summary

1. The balance sheet of commercial banks can be thought of as a list of the sources and uses of bank funds. The bank's liabilities are its sources of funds, which include checkable deposits, time deposits, discount loans from the Fed, borrowings from other banks and corporations, and bank capital. The bank's assets are its uses of funds, which include reserves, cash items in process of collection, deposits at other banks, securities, loans, and other assets (mostly physical capital).
2. Banks make profits through the process of asset transformation: They borrow short (accept deposits) and lend long (make loans). When a bank takes in additional deposits, it gains an equal amount of reserves; when it pays out deposits, it loses an equal amount of reserves.
3. Although more-liquid assets tend to earn lower returns, banks still desire to hold them. Specifically, banks hold excess and secondary reserves because they provide insurance against the costs of a deposit outflow. Banks manage their assets to maximize profits by seeking the highest returns possible on loans and securities while at the same time trying to lower risk and making adequate provisions for liquidity. Although liability management was once a staid affair, large (money center) banks now actively seek out sources of funds by issuing liabilities such as negotiable CDs or by actively borrowing from other banks and corporations. Banks manage the amount of capital they hold to prevent bank failure and to meet bank capital requirements set by the regulatory authorities. However, they do not want to hold too much capital because by so doing they will lower the returns to equity holders.
4. The concepts of adverse selection and moral hazard explain many credit risk management principles involving loan activities: screening and monitoring, establishment of long-term customer relationships and loan commitments, collateral and compensating balances, and credit rationing.
5. With the increased volatility of interest rates that occurred in the 1980s, financial institutions became more concerned about their exposure to interest-rate risk. Gap and duration analyses tell a financial institution if it has more rate-sensitive liabilities than assets (in which case a rise in interest rates will reduce profits and a fall in interest rates will raise profits). Financial institutions manage their interest-rate risk by modifying their balance sheets but can also use strategies (outlined in Chapter 13) involving financial derivatives.
6. Off-balance-sheet activities consist of trading financial instruments and generating income from fees and loan sales, all of which affect bank profits but are not visible on bank balance sheets. Because these off-balance-sheet activities expose banks to increased risk, bank management must pay particular attention to risk assessment procedures and internal controls to restrict employees from taking on too much risk.

Key Terms

asset management, p. 208
balance sheet, p. 201
capital adequacy management, p. 208
compensating balance, p. 219
credit rationing, p. 220
credit risk, p. 208
deposit outflows, p. 208
discount loans, p. 203
discount rate, p. 210
duration analysis, p. 221
equity multiplier (EM), p. 214
excess reserves, p. 204
gap analysis, p. 221
interest-rate risk, p. 208
liability management, p. 208
liquidity management, p. 208
loan commitment, p. 219
loan sale, p. 223
money center banks, p. 212
off-balance-sheet activities, p. 223
required reserve ratio, p. 204
required reserves, p. 204
reserve requirements, p. 204
reserves, p. 204
return on assets (ROA), p. 214
return on equity (ROE), p. 214
secondary reserves, p. 204
T-account, p. 205
vault cash, p. 204

Questions and Problems

Questions marked with an asterisk are answered at the end of the book in an appendix, "Answers to Selected Questions and Problems."

1. Why might a bank be willing to borrow funds from other banks at a higher rate than it can borrow from the Fed?

*2. Rank the following bank assets from most to least liquid:
   a. Commercial loans
   b. Securities
   c. Reserves
   d. Physical capital

3. Using the T-accounts of the First National Bank and the Second National Bank, describe what happens when Jane Brown writes a $50 check on her account at the First National Bank to pay her friend Joe Green, who in turn deposits the check in his account at the Second National Bank.

*4. What happens to reserves at the First National Bank if one person withdraws $1,000 of cash and another person deposits $500 of cash? Use T-accounts to explain your answer.

5. The bank you own has the following balance sheet:

Assets		Liabilities	
Reserves	$ 75 million	Deposits	$500 million
Loans	$525 million	Bank capital	$100 million

If the bank suffers a deposit outflow of $50 million with a required reserve ratio on deposits of 10%, what actions must you take to keep your bank from failing?

*6. If a deposit outflow of $50 million occurs, which balance sheet would a bank rather have initially, the balance sheet in Problem 5 or the following balance sheet? Why?

Assets		Liabilities	
Reserves	$100 million	Deposits	$500 million
Loans	$500 million	Bank capital	$100 million

7. Why has the development of overnight loan markets made it more likely that banks will hold fewer excess reserves?

*8. If the bank you own has no excess reserves and a sound customer comes in asking for a loan, should you automatically turn the customer down, explaining that you don't have any excess reserves to lend out? Why or why not? What options are available for you to provide the funds your customer needs?

9. If a bank finds that its ROE is too low because it has too much bank capital, what can it do to raise its ROE?

*10. If a bank is falling short of meeting its capital requirements by $1 million, what three things can it do to rectify the situation?

11. Why is being nosy a desirable trait for a banker?

*12. A bank almost always insists that the firms it lends to keep compensating balances at the bank. Why?

13. "Because diversification is a desirable strategy for avoiding risk, it never makes sense for a bank to specialize in making specific types of loans." Is this statement true, false, or uncertain? Explain your answer.

*14. Suppose that you are the manager of a bank whose $100 billion of assets have an average duration of four years and whose $90 billion of liabilities have an average duration of six years. Conduct a duration analysis for the bank, and show what will happen to the net worth of the bank if interest rates rise by 2 percentage points. What actions could you take to reduce the bank's interest-rate risk?

15. Suppose that you are the manager of a bank that has $15 million of fixed-rate assets, $30 million of rate-sensitive assets, $25 million of fixed-rate liabilities, and $20 million of rate-sensitive liabilities. Conduct a gap analysis for the bank, and show what will happen to bank profits if interest rates rise by 5 percentage points. What actions could you take to reduce the bank's interest-rate risk?

Web Exercises

1. Table 1 reports the balance sheet of all commercial banks based on aggregate data found in the Federal Reserve *Bulletin*. Compare this table to the balance sheet reported by Wachovia found at www.wachovia.com/investor/annualfinancials.asp. Does Wachovia have more or less of its portfolio in loans than the average bank? What type of loan is most common?

2. It is relatively easy to find up-to-date information on banks because of their extensive reporting requirements. Go to www2.fdic.gov/qbp/. This site is sponsored by the Federal Deposit Insurance Corporation. You will find summary data on financial institutions. Go to the most recent Quarterly Banking Profile. Scroll to the bottom and open Table 1-A.
   a. Have banks' return on assets been increasing or decreasing over the last few years?
   b. Has the core capital been increasing and how does it compare to the capital ratio reported in Table 1 in the text?
   c. How many institutions are currently reporting to the FDIC?

appendix 1 to chapter 9

Duration Gap Analysis

An alternative method for measuring interest-rate risk, called *duration gap analysis*, examines the sensitivity of the market value of the financial institution's net worth to changes in interest rates. Duration analysis is based on Macaulay's concept of *duration*, which measures the average lifetime of a security's stream of payments (described in the appendix to Chapter 4). Recall that duration is a useful concept, because it provides a good approximation, particularly when interest-rate changes are small, of the sensitivity of a security's market value to a change in its interest rate using the following formula:

$$\%\Delta P \approx -DUR \times \frac{\Delta i}{1 + i} \tag{1}$$

where

$\%\Delta P = (P_{t+1} - P_t)/P_t$ = percent change in market value of the security
DUR = duration
i = interest rate

After having determined the duration of all assets and liabilities on the bank's balance sheet, the bank manager could use this formula to calculate how the market value of each asset and liability changes when there is a change in interest rates and then calculate the effect on net worth. There is, however, an easier way to go about doing this, derived from the basic fact about duration we learned in the appendix to Chapter 4: Duration is additive; that is, the duration of a portfolio of securities is the weighted average of the durations of the individual securities, with the weights reflecting the proportion of the portfolio invested in each. What this means is that the bank manager can figure out the effect that interest-rate changes will have on the market value of net worth by calculating the average duration for assets and for liabilities and then using those figures to estimate the effects of interest-rate changes.

To see how a bank manager would do this, let's return to the balance sheet of the First National Bank. The bank manager has already used the procedures outlined in the appendix to Chapter 4 to calculate the duration of each asset and liability, as listed in Table 1. For each asset, the manager then calculates the weighted duration by multiplying the duration times the amount of the asset divided by total assets, which in this case is $100 million. For example, in the case of securities with maturities less than one year, the manager multiplies the 0.4 year of duration times $5 million divided by $100 million to get a weighted duration of 0.02. (Note that physical assets have no cash payments, so they have a duration of zero years.) Doing this for all the

Table 1 Duration of the First National Bank's Assets and Liabilities

	Amount ($ millions)	Duration (years)	Weighted Duration (years)
Assets			
Reserves and cash items	5	0.0	0.00
Securities			
Less than 1 year	5	0.4	0.02
1 to 2 years	5	1.6	0.08
Greater than 2 years	10	7.0	0.70
Residential mortgages			
Variable-rate	10	0.5	0.05
Fixed-rate (30-year)	10	6.0	0.60
Commercial loans			
Less than 1 year	15	0.7	0.11
1 to 2 years	10	1.4	0.14
Greater than 2 years	25	4.0	1.00
Physical capital	5	0.0	0.00
Average duration			2.70
Liabilities			
Checkable deposits	15	2.0	0.32
Money market deposit accounts	5	0.1	0.01
Savings deposits	15	1.0	0.16
CDs			
Variable-rate	10	0.5	0.05
Less than 1 year	15	0.2	0.03
1 to 2 years	5	1.2	0.06
Greater than 2 years	5	2.7	0.14
Fed funds	5	0.0	0.00
Borrowings			
Less than 1 year	10	0.3	0.03
1 to 2 years	5	1.3	0.07
Greater than 2 years	5	3.1	0.16
Average duration			1.03

assets and adding them up, the bank manager gets a figure for the average duration of the assets of 2.70 years.

The manager follows a similar procedure for the liabilities, noting that total liabilities excluding capital are $95 million. For example, the weighted duration for checkable deposits is determined by multiplying the 2.0-year duration by $15 million divided by $95 million to get 0.32. Adding up these weighted durations, the manager obtains an average duration of liabilities of 1.03 years.

EXAMPLE 1: Duration Gap Analysis

The bank manager wants to know what happens when interest rates rise from 10% to 11%. The total asset value is $100 million, and the total liability value is $95 million. Use Equation 1 to calculate the change in the market value of the assets and liabilities.

Solution

With a total asset value of $100 million, the market value of assets falls by $2.5 million ($100 million × 0.025 = $2.5 million):

$$\%\Delta P = -DUR \times \frac{\Delta i}{1 + i}$$

where

$$\begin{aligned} DUR &= \text{duration} &&= 2.70 \\ \Delta i &= \text{change in interest rate} = 0.11 - 0.10 &&= 0.01 \\ i &= \text{interest rate} &&= 0.10 \end{aligned}$$

Thus:

$$\%\Delta P \approx -2.70 \times \frac{0.01}{1 + 0.10} = -0.025 = -2.5\%$$

With total liabilities of $95 million, the market value of liabilities falls by $0.9 million ($95 million × 0.009 = −$0.9 million):

$$\%\Delta P \approx -DUR \times \frac{\Delta i}{1 + i}$$

where

$$\begin{aligned} DUR &= \text{duration} &&= 1.03 \\ \Delta i &= \text{change in interest rate} = 0.11 - 0.10 &&= 0.01 \\ i &= \text{interest rate} &&= 0.10 \end{aligned}$$

Thus:

$$\%\Delta P \approx -1.03 \times \frac{0.01}{1 + 0.10} = -0.009 = -0.9\%$$

The result is that the net worth of the bank would decline by $1.6 million (−$2.5 million − (−$0.9 million) = −$2.5 million + $0.9 million = −$1.6 million).

The bank manager could have gotten to the answer even more quickly by calculating what is called a *duration gap*, which is defined as follows:

$$DUR_{gap} = DUR_a - \left(\frac{L}{A} \times DUR_l\right) \tag{2}$$

where DUR_a = average duration of assets
DUR_l = average duration of liabilities
L = market value of liabilities
A = market value of assets

EXAMPLE 2: Duration Gap Analysis

Based on the information provided in Example 1, use Equation 2 to determine the duration gap for First National Bank.

Solution

The duration gap for First National Bank is 1.72 years:

$$DUR_{gap} = DUR_a - \left(\frac{L}{A} \times DUR_l\right)$$

where

$$\begin{aligned} DUR_a &= \text{average duration of assets} &&= 2.70 \\ L &= \text{market value of liabilities} &&= 95 \\ A &= \text{market value of assets} &&= 100 \\ DUR_l &= \text{average duration of liabilities} &&= 1.03 \end{aligned}$$

Thus:

$$DUR_{gap} = 2.70 - \left(\frac{95}{100} \times 1.03\right) = 1.72 \text{ years}$$

To estimate what will happen if interest rates change, the bank manager uses the DUR_{gap} calculation in Equation 1 to obtain the change in the market value of net worth as a percentage of total assets. In other words, the change in the market value of net worth as a percentage of assets is calculated as:

$$\frac{\Delta NW}{A} \approx -DUR_{gap} \times \frac{\Delta i}{1 + i} \tag{3}$$

EXAMPLE 3: Duration Gap Analysis

What is the change in the market value of net worth as a percentage of assets if interest rates rise from 10% to 11%? (Use Equation 3.)

Solution

A rise in interest rates from 10% to 11% would lead to a change in the market value of net worth as a percentage of assets of −1.6%:

$$\frac{\Delta NW}{A} = -DUR_{gap} \times \frac{\Delta i}{1 + i}$$

where

$$\begin{aligned} DUR_{gap} &= \text{duration gap} &&= 1.72 \\ \Delta i &= \text{change in interest rate} = 0.11 - 0.10 &&= 0.01 \\ i &= \text{interest rate} &&= 0.10 \end{aligned}$$

Thus:

$$\frac{\Delta NW}{A} = -1.72 \times \frac{0.01}{1 + 0.10} = -0.016 = -1.6\%$$

With assets totaling $100 million, Example 3 indicates a fall in the market value of net worth of $1.6 million, which is the same figure that we found in Example 1.

As our examples make clear, both income gap analysis and duration gap analysis indicate that the First National Bank will suffer from a rise in interest rates. Indeed, in this example, we have seen that a rise in interest rates from 10% to 11% will cause the market value of net worth to fall by $1.6 million, which is one-third the initial amount of bank capital. Thus the bank manager realizes that the bank faces substantial interest-rate risk because a rise in interest rates could cause it to lose a lot of its capital. Clearly, income gap analysis and duration gap analysis are useful tools for telling a financial institution manager the institution's degree of exposure to interest-rate risk.

Study Guide

To make sure that you understand income gap and duration gap analysis, you should be able to verify that if interest rates fall from 10% to 5%, the First National Bank will find its income increasing and the market value of its net worth rising.

Example of a Nonbanking Financial Institution

So far we have focused on an example involving a banking institution that has borrowed short and lent long so that when interest rates rise, both income and the net worth of the institution fall. It is important to recognize that income and duration gap analysis applies equally to other financial institutions. Furthermore, it is important for you to see that some financial institutions have income and duration gaps that are opposite in sign to those of banks, so that when interest rates rise, both income and net worth rise rather than fall. To get a more complete picture of income and duration gap analysis, let us look at a nonbank financial institution, the Friendly Finance Company, which specializes in making consumer loans.

The Friendly Finance Company has the following balance sheet:

Friendly Finance Company

Assets		Liabilities	
Cash and deposits	$3 million	Commercial paper	$40 million
Securities		Bank loans	
Less than 1 year	$5 million	Less than 1 year	$3 million
1 to 2 years	$1 million	1 to 2 years	$2 million
Greater than 2 years	$1 million	Greater than 2 years	$5 million
Consumer loans		Long-term bonds and	
Less than 1 year	$50 million	other long-term debt	$40 million
1 to 2 years	$20 million	Capital	$10 million
Greater than 2 years	$15 million		
Physical capital	$5 million		
Total	$100 million	Total	$100 million

The manager of the Friendly Finance Company calculates the rate-sensitive assets to be equal to the $5 million of securities with maturities less than one year plus the

$50 million of consumer loans with maturities of less than one year, for a total of $55 million of rate-sensitive assets. The manager then calculates the rate-sensitive liabilities to be equal to the $40 million of commercial paper, all of which has a maturity of less than one year, plus the $3 million of bank loans maturing in less than a year, for a total of $43 million. The calculation of the income gap is then:

$$GAP = RSA - RSL = \$55 \text{ million} - \$43 \text{ million} = \$12 \text{ million}$$

To calculate the effect on income if interest rates rise by 1%, the manager multiplies the *GAP* of $12 million times the change in the interest rate to get the following:

$$\Delta I = GAP \times \Delta i = \$12 \text{ million} \times 1\% = \$120{,}000$$

Thus the manager finds that the finance company's income will rise by $120,000 when interest rates rise by 1%. The reason that the company has benefited from the interest-rate rise, in contrast to the First National Bank, whose profits suffer from the rise in interest rates, is that the Friendly Finance Company has a positive income gap because it has more rate-sensitive assets than liabilities.

Like the bank manager, the manager of the Friendly Finance Company is also interested in what happens to the market value of the net worth of the company when interest rates rise by 1%. So the manager calculates the weighted duration of each item in the balance sheet, adds them up as in Table 2, and obtains a duration for the assets of 1.16 years and for the liabilities, 2.77 years. The duration gap is then calculated to be:

$$DUR_{gap} = DUR_a - \left(\frac{L}{A} \times DUR_l\right) = 1.16 - \left(\frac{90}{100} \times 2.77\right) = -1.33 \text{ years}$$

Since the Friendly Finance Company has a negative duration gap, the manager realizes that a rise in interest rates by 1 percentage point from 10% to 11% will increase the market value of net worth of the firm. The manager checks this by calculating the change in the market value of net worth as a percentage of assets:

$$\Delta NW = -DUR_{gap} \times \frac{\Delta i}{1 + i} = -(-1.33) \times \frac{0.01}{1 + 0.10} = 0.012 = 1.2\%$$

With assets of $100 million, this calculation indicates that net worth will rise in market value by $1.2 million.

Even though the income and duration gap analysis indicates that the Friendly Finance Company gains from a rise in interest rates, the manager realizes that if interest rates go in the other direction, the company will suffer a fall in income and market value of net worth. Thus the finance company manager, like the bank manager, realizes that the institution is subject to substantial interest-rate risk.

Some Problems with Income and Duration Gap Analysis

Although you might think that income and duration gap analysis is complicated enough, further complications make a financial institution manager's job even harder.

One assumption that we have been using in our discussion of income and duration gap analysis is that when the level of interest rates changes, interest rates on all maturities change by exactly the same amount. That is the same as saying that we conducted our analysis under the assumption that the slope of the yield curve remains unchanged. Indeed, the situation is even worse for duration gap analysis, because the

Table 2 Duration of the Friendly Finance Company's Assets and Liabilities

	Amount ($ millions)	Duration (years)	Weighted Duration (years)
Assets			
Cash and deposits	3	0.0	0.00
Securities			
Less than 1 year	5	0.5	0.05
1 to 2 years	1	1.7	0.02
Greater than 2 years	1	9.0	0.09
Consumer loans			
Less than 1 year	50	0.5	0.25
1 to 2 years	20	1.5	0.30
Greater than 2 years	15	3.0	0.45
Physical capital	5	0.0	0.00
Average duration			1.16
Liabilities			
Commercial paper	40	0.2	0.09
Bank loans			
Less than 1 year	3	0.3	0.01
1 to 2 years	2	1.6	0.04
Greater than 2 years	5	3.5	0.19
Long-term bonds and other long-term debt	40	5.5	2.44
Average duration			2.77

duration gap is calculated assuming that interest rates for all maturities are the same—in other words, the yield curve is assumed to be flat. As our discussion of the term structure of interest rates in Chapter 6 indicated, however, the yield curve is not flat, and the slope of the yield curve fluctuates and has a tendency to change when the level of the interest rate changes. Thus to get a truly accurate assessment of interest-rate risk, a financial institution manager has to assess what might happen to the slope of the yield curve when the level of the interest rate changes and then take this information into account when assessing interest-rate risk. In addition, duration gap analysis is based on the approximation in Equation 1 and thus only works well for small changes in interest rates.

A problem with income gap analysis is that, as we have seen, the financial institution manager must make estimates of the proportion of supposedly fixed-rate assets and liabilities that may be rate-sensitive. This involves estimates of the likelihood of prepayment of loans or customer shifts out of deposits when interest rates change. Such guesses are not easy to make, and as a result, the financial institution manager's estimates of income gaps may not be very accurate. A similar problem occurs in cal-

culating durations of assets and liabilities, because many of the cash payments are uncertain. Thus the estimate of the duration gap might not be accurate either.

Do these problems mean that managers of banks and other financial institutions should give up on gap analysis as a tool for measuring interest-rate risk? Financial institutions do use more sophisticated approaches to measuring interest-rate risk, such as scenario analysis and value-at-risk analysis, which make greater use of computers to more accurately measure changes in prices of assets when interest rates change. Income and duration gap analyses, however, still provide simple frameworks to help financial institution managers to get a first assessment of interest-rate risk, and they are thus useful tools in the financial institution manager's toolkit.

Application

Strategies for Managing Interest-Rate Risk

Once financial institution managers have done the duration and income gap analysis for their institutions, they must decide which alternative strategies to pursue. If the manager of the First National Bank firmly believes that interest rates will fall in the future, he or she may be willing to take no action knowing that the bank has more rate-sensitive liabilities than rate-sensitive assets, and so will benefit from the expected interest-rate decline. However, the bank manager also realizes that the First National Bank is subject to substantial interest-rate risk, because there is always a possibility that interest rates will rise rather than fall, and as we have seen, this outcome could bankrupt the bank. The manager might try to shorten the duration of the bank's assets to increase their rate sensitivity either by purchasing assets of shorter maturity or by converting fixed-rate loans into adjustable-rate loans. Alternatively, the bank manager could lengthen the duration of the liabilities. With these adjustments to the bank's assets and liabilities, the bank would be less affected by interest-rate swings.

For example, the bank manager might decide to eliminate the income gap by increasing the amount of rate-sensitive assets to $49.5 million to equal the $49.5 million of rate-sensitive liabilities. Or the manager could reduce rate-sensitive liabilities to $32 million so that they equal rate-sensitive assets. In either case, the income gap would now be zero, so a change in interest rates would have no effect on bank profits in the coming year.

Alternatively, the bank manager might decide to immunize the market value of the bank's net worth completely from interest-rate risk by adjusting assets and liabilities so that the duration gap is equal to zero. To do this, the manager can set DUR_{gap} equal to zero in Equation 2 and solve for DUR_a:

$$DUR_a = \frac{L}{A} \times DUR_l = \frac{95}{100} \times 1.03 = 0.98$$

These calculations reveal that the manager should reduce the average duration of the bank's assets to 0.98 year. To check that the duration gap is set equal to zero, the calculation is:

$$DUR_{gap} = 0.98 - \left(\frac{95}{100} \times 1.03\right) = 0$$

In this case, as in Equation 3, the market value of net worth would remain unchanged when interest rates change. Alternatively, the bank manager could calculate the value of the duration of the liabilities that would produce a duration gap of zero. To do this would involve setting DUR_{gap} equal to zero in Equation 2 and solving for DUR_l:

$$DUR_l = DUR_a \times \frac{A}{L} = 2.70 \times \frac{100}{95} = 2.84$$

This calculation reveals that the interest-rate risk could also be eliminated by increasing the average duration of the bank's liabilities to 2.84 years. The manager again checks that the duration gap is set equal to zero by calculating:

$$DUR_{gap} = 2.70 - \left(\frac{95}{100} \times 2.84\right) = 0$$

Study Guide

To see if you understand how a financial institution manager can protect income and net worth from interest-rate risk, first calculate how the Friendly Finance Company might change the amount of its rate-sensitive assets or its rate-sensitive liabilities to eliminate the income gap. You should find that the income gap can be eliminated either by reducing the amount of rate-sensitive assets to $43 million or by raising the amount of rate-sensitive liabilities to $55 million. Also do the calculations to determine what modifications to the duration of the assets or liabilities would immunize the market value of Friendly Finance's net worth from interest-rate risk. You should find that interest-rate risk would be eliminated if the duration of the assets were set to 2.49 years or if the duration of the liabilities were set to 1.29 years.

One problem with eliminating a financial institution's interest-rate risk by altering the balance sheet is that doing so might be very costly in the short run. The financial institution may be locked into assets and liabilities of particular durations because of its field of expertise. Fortunately, recently developed financial instruments, such as financial futures, options, and interest-rate swaps, help financial institutions manage their interest-rate risk without requiring them to rearrange their balance sheets. We discuss these instruments and how they can be used to manage interest-rate risk in Chapter 13.

appendix 2 to chapter 9

Measuring Bank Performance

To understand how well a bank is doing, we need to start by looking at a bank's income statement, the description of the sources of income and expenses that affect the bank's profitability.

Bank's Income Statement

The end-of-year 2002 income statement for all federally insured commercial banks appears in Table 1.

Operating Income. *Operating income* is the income that comes from a bank's ongoing operations. Most of a bank's operating income is generated by interest on its assets, particularly loans. As we see in Table 1, in 2002 interest income represented 67.6% of commercial banks' operating income. Interest income fluctuates with the level of interest rates, and so its percentage of operating income is highest when interest rates are at peak levels. That is exactly what happened in 1981, when interest rates rose above 15% and interest income rose to 93% of total bank operating income.

Noninterest income is generated partly by service charges on deposit accounts, but the bulk of it comes from the off-balance-sheet activities, which generate fees or trading profits for the bank. The importance of these off-balance-sheet activities to bank profits has been growing in recent years. Whereas in 1980 other noninterest income from off-balance-sheet activities represented only 5% of operating income, it reached 26.8% in 2002.

Operating Expenses. *Operating expenses* are the expenses incurred in conducting the bank's ongoing operations. An important component of a bank's operating expenses is the interest payments that it must make on its liabilities, particularly on its deposits. Just as interest income varies with the level of interest rates, so do interest expenses. Interest expenses as a percentage of total operating expenses reached a peak of 74% in 1981, when interest rates were at their highest, and fell to 30.1% in 2002 as interest rates moved lower. Noninterest expenses include the costs of running a banking business: salaries for tellers and officers, rent on bank buildings, purchases of equipment such as desks and vaults, and servicing costs of equipment such as computers.

The final item listed under operating expenses is provisions for loan losses. When a bank has a bad debt or anticipates that a loan might become a bad debt in the future, it can write up the loss as a current expense in its income statement under the "provision for loan losses" heading. Provisions for loan losses are directly related to loan loss reserves. When a bank wants to increase its loan loss reserves account by, say, $1 million, it does this by adding $1 million to its provisions for loan losses. Loan loss

Table 1 Income Statement for All Federally Insured Commercial Banks, 2002

	Amount ($ billions)		Share of Operating Income or Expenses (%)	
Operating Income				
Interest income		357.7		67.6
Interest on loans	266.3		50.3	
Interest on securities	60.1		11.4	
Other interest	31.3		5.9	
Noninterest income		171.4		32.4
Service charges on deposit accounts	29.7		5.6	
Other noninterest income	141.7		26.8	
Total operating income		529.1		100.0
Operating Expenses				
Interest expenses		120.8		30.1
Interest on deposits	82.3		20.5	
Interest on fed funds and repos	10.4		2.6	
Other	28.1		7.0	
Noninterest expenses		232.6		57.9
Salaries and employee benefits	100.4		25.0	
Premises and equipment	29.4		7.3	
Other	102.8		25.6	
Provisions for loan losses		48.0		12.0
Total operating expense		401.4		100.0
Net Operating Income		127.7		
Gains (losses) on securities		6.5		
Extraordinary items, net		0.0		
Income taxes		–44.1		
Net Income		90.1		

Source: www.fdic.gov/banks/statistical/statistics/0106/cbr

reserves rise when this is done because by increasing expenses when losses have not yet occurred, earnings are being set aside to deal with the losses in the future.

Provisions for loan losses have been a major element in fluctuating bank profits in recent years. The 1980s brought the third-world debt crisis; a sharp decline in energy prices in 1986, which caused substantial losses on loans to energy producers; and a collapse in the real estate market. As a result, provisions for loan losses were particularly high in the late 1980s, reaching a peak of 13% of operating expenses in

1987. Since then, losses on loans have begun to subside, and in 2002, provisions for loan losses dropped to 12% of operating expenses.

Income. Subtracting the $401.4 billion in operating expenses from the $529.1 billion of operating income in 2002 yields net operating income of $127.7 billion. Net operating income is closely watched by bank managers, bank shareholders, and bank regulators because it indicates how well the bank is doing on an ongoing basis.

Two items, gains (or losses) on securities sold by banks (+$6.5 billion) and net extraordinary items, which are events or transactions that are both unusual and infrequent (insignificant), are added to the $127.7 billion net operating income figure to get the $134.2 billion figure for net income before taxes. Net income before taxes is more commonly referred to as profits before taxes. Subtracting the $44.1 billion of income taxes then results in $90.1 billion of net income. Net income, more commonly referred to as profits after taxes, is the figure that tells us most directly how well the bank is doing because it is the amount that the bank has available to keep as retained earnings or to pay out to stockholders as dividends.

Measures of Bank Performance

Although net income gives us an idea of how well a bank is doing, it suffers from one major drawback: It does not adjust for the bank's size, thus making it hard to compare how well one bank is doing relative to another. A basic measure of bank profitability that corrects for the size of the bank is the return on assets (*ROA*), mentioned earlier in the chapter, which divides the net income of the bank by the amount of its assets. *ROA* is a useful measure of how well a bank manager is doing on the job because it indicates how well a bank's assets are being used to generate profits. At the beginning of 2003, the assets of all federally insured commercial banks amounted to $7,075 billion, so using the $90.1 billion net income figure from Table 1 gives us a return on assets of:

$$ROA = \frac{\text{net income}}{\text{assets}} = \frac{90.1}{7{,}075} = 0.0127 = 1.27\%$$

Although *ROA* provides useful information about bank profitability, we have already seen that it is not what the bank's owners (equity holders) care about most. They are more concerned about how much the bank is earning on their equity investment, an amount that is measured by the return on equity (*ROE*), the net income per dollar of equity capital. At the beginning of 2003, equity capital for all federally insured commercial banks was $647.9 billion, so the *ROE* was therefore:

$$ROE = \frac{\text{net income}}{\text{capital}} = \frac{90.1}{647.9} = 0.1391 = 13.91\%$$

Another commonly watched measure of bank performance is called the *net interest margin* (*NIM*), the difference between interest income and interest expenses as a percentage of total assets:

$$NIM = \frac{\text{interest income} - \text{interest expenses}}{\text{assets}}$$

As we have seen earlier in the chapter, one of a bank's primary intermediation functions is to issue liabilities and use the proceeds to purchase income-earning assets. If a bank manager has done a good job of asset and liability management such that the bank earns substantial income on its assets and has low costs on its liabilities,

profits will be high. How well a bank manages its assets and liabilities is affected by the spread between the interest earned on the bank's assets and the interest costs on its liabilities. This spread is exactly what the net interest margin measures. If the bank is able to raise funds with liabilities that have low interest costs and is able to acquire assets with high interest income, the net interest margin will be high, and the bank is likely to be highly profitable. If the interest cost of its liabilities rises relative to the interest earned on its assets, the net interest margin will fall, and bank profitability will suffer.

Recent Trends in Bank Performance Measures

Table 2 provides measures of return on assets (*ROA*), return on equity (*ROE*), and the net interest margin (*NIM*) for all federally insured commercial banks from 1980 to 2002. Because the relationship between bank equity capital and total assets for all commercial banks remained fairly stable in the 1980s, both the *ROA* and *ROE* meas-

Table 2 Measures of Bank Performance, 1980–2002

Year	Return on Assets (*ROA*) (%)	Return on Equity (*ROE*) (%)	Net Interest Margin (*NIM*)(%)
1980	0.77	13.38	3.33
1981	0.79	13.68	3.31
1982	0.73	12.55	3.39
1983	0.68	11.60	3.34
1984	0.66	11.04	3.47
1985	0.72	11.67	3.62
1986	0.64	10.30	3.48
1987	0.09	1.54	3.40
1988	0.82	13.74	3.57
1989	0.50	7.92	3.58
1990	0.49	7.81	3.50
1991	0.53	8.25	3.60
1992	0.94	13.86	3.89
1993	1.23	16.30	3.97
1994	1.20	15.00	3.95
1995	1.17	14.66	4.29
1996	1.19	14.45	4.27
1997	1.23	14.69	4.21
1998	1.18	13.30	3.47
1999	1.31	15.31	4.07
2000	1.19	14.02	3.95
2001	1.13	12.45	3.28
2002	1.27	13.91	3.34

Source: www2.fdic.gov/qbp

ures of bank performance move closely together and indicate that from the early to the late 1980s, there was a sharp decline in bank profitability. The rightmost column, net interest margin, indicates that the spread between interest income and interest expenses remained fairly stable throughout the 1980s and even improved in the late 1980s and early 1990s, which should have helped bank profits. The *NIM* measure thus tells us that the poor bank performance in the late 1980s was not the result of interest-rate movements.

The explanation of the weak performance of commercial banks in the late 1980s is that they had made many risky loans in the early 1980s that turned sour. The resulting huge increase in loan loss provisions in that period directly decreased net income and hence caused the fall in *ROA* and *ROE*. (Why bank profitability deteriorated and the consequences for the economy are discussed in Chapters 9 and 11.)

Beginning in 1992, bank performance improved substantially. The return on equity rose to nearly 14% in 1992 and remained above 12% in the 1993–2003 period. Similarly, the return on assets rose from the 0.5% level in the 1990–1991 period to around the 1.2% level in 1993–2003. The performance measures in Table 2 suggest that the banking industry has returned to health.

Chapter 10

Banking Industry: Structure and Competition

PREVIEW

The operations of individual banks (how they acquire, use, and manage funds to make a profit) are roughly similar throughout the world. In all countries, banks are financial intermediaries in the business of earning profits. When you consider the structure and operation of the banking industry as a whole, however, the United States is in a class by itself. In most countries, four or five large banks typically dominate the banking industry, but in the United States there are on the order of 8,000 commercial banks, 1,500 savings and loan associations, 400 mutual savings banks, and 10,000 credit unions.

Is more better? Does this diversity mean that the American banking system is more competitive and therefore more economically efficient and sound than banking systems in other countries? What in the American economic and political system explains this large number of banking institutions? In this chapter, we try to answer these questions by examining the historical trends in the banking industry and its overall structure.

We start by examining the historical development of the banking system and how financial innovation has increased the competitive environment for the banking industry and is causing fundamental changes in it. We then go on to look at the commercial banking industry in detail and then discuss the thrift industry, which includes savings and loan associations, mutual savings banks, and credit unions. We spend more time on commercial banks because they are by far the largest depository institutions, accounting for over two-thirds of the deposits in the banking system. In addition to looking at our domestic banking system, we also examine the forces behind the growth in international banking to see how it has affected us in the United States.

Historical Development of the Banking System

The modern commercial banking industry in the Unted States began when the Bank of North America was chartered in Philadelphia in 1782. With the success of this bank, other banks opened for business, and the American banking industry was off and running. (As a study aid, Figure 1 provides a time line of the most important dates in the history of American banking before World War II.)

A major controversy involving the industry in its early years was whether the federal government or the states should charter banks. The Federalists, particularly Alexander Hamilton, advocated greater centralized control of banking and federal

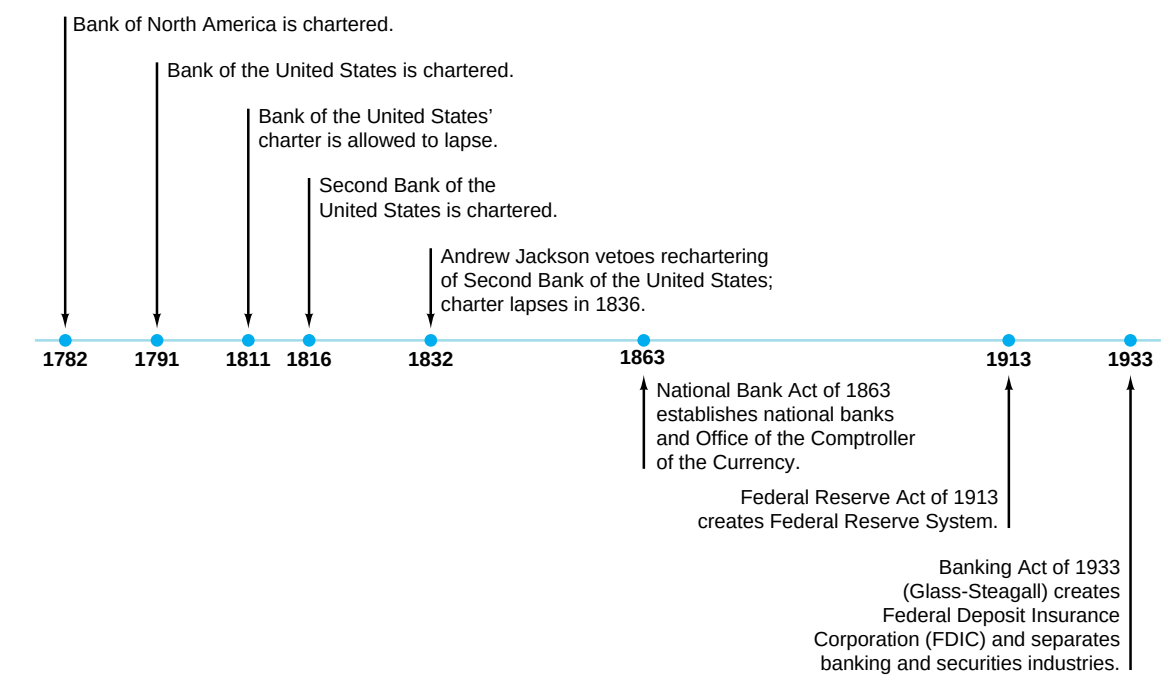


FIGURE 1 Time Line of the Early History of Commercial Banking in the United States

The government agency that oversees the banking system and is responsible for the amount of money and credit supplied in the economy; in the United States, the Federal Reserve System.

chartering of banks. Their efforts led to the creation in 1791 of the Bank of the United States, which had elements of both a private and a **central bank**, a government institution that has responsibility for the amount of money and credit supplied in the economy as a whole. Agricultural and other interests, however, were quite suspicious of centralized power and hence advocated chartering by the states. Furthermore, their distrust of moneyed interests in the big cities led to political pressures to eliminate the Bank of the United States, and in 1811 their efforts met with success, when its charter was not renewed. Because of abuses by state banks and the clear need for a central bank to help the federal government raise funds during the War of 1812, Congress was stimulated to create the Second Bank of the United States in 1816. Tensions between advocates and opponents of centralized banking power were a recurrent theme during the operation of this second attempt at central banking in the United States, and with the election of Andrew Jackson, a strong advocate of states' rights, the fate of the Second Bank was sealed. After the election in 1832, Jackson vetoed the rechartering of the Second Bank of the United States as a national bank, and its charter lapsed in 1836.

Until 1863, all commercial banks in the United States were chartered by the banking commission of the state in which each operated. No national currency existed, and banks obtained funds primarily by issuing *banknotes* (currency circulated by the banks that could be redeemed for gold). Because banking regulations were

extremely lax in many states, banks regularly failed due to fraud or lack of sufficient bank capital; their banknotes became worthless.

To eliminate the abuses of the state-chartered banks (called **state banks**), the National Bank Act of 1863 (and subsequent amendments to it) created a new banking system of federally chartered banks (called **national banks**), supervised by the Office of the Comptroller of the Currency, a department of the U.S. Treasury. This legislation was originally intended to dry up sources of funds to state banks by imposing a prohibitive tax on their banknotes while leaving the banknotes of the federally chartered banks untaxed. The state banks cleverly escaped extinction by acquiring funds through deposits. As a result, today the United States has a **dual banking system** in which banks supervised by the federal government and banks supervised by the states operate side by side.

Central banking did not reappear in this country until the Federal Reserve System (the Fed) was created in 1913 to promote an even safer banking system. All national banks were required to become members of the Federal Reserve System and became subject to a new set of regulations issued by the Fed. State banks could choose (but were not required) to become members of the system, and most did not because of the high costs of membership stemming from the Fed's regulations.

During the Great Depression years 1930–1933, some 9,000 bank failures wiped out the savings of many depositors at commercial banks. To prevent future depositor losses from such failures, banking legislation in 1933 established the Federal Deposit Insurance Corporation (FDIC), which provided federal insurance on bank deposits. Member banks of the Federal Reserve System were required to purchase FDIC insurance for their depositors, and non–Federal Reserve commercial banks could choose to buy this insurance (almost all of them did). The purchase of FDIC insurance made banks subject to another set of regulations imposed by the FDIC.

www.fdic.gov/bank/index.htm
The FDIC gathers data about individual financial institutions and the banking industry.

Because investment banking activities of the commercial banks were blamed for many bank failures, provisions in the banking legislation in 1933 (also known as the Glass-Steagall Act) prohibited commercial banks from underwriting or dealing in corporate securities (though allowing them to sell new issues of government securities) and limited banks to the purchase of debt securities approved by the bank regulatory agencies. Likewise, it prohibited investment banks from engaging in commercial banking activities. In effect, the Glass-Steagall Act separated the activities of commercial banks from those of the securities industry.

Under the conditions of the Glass-Steagall Act, which was repealed in 1999, commercial banks had to sell off their investment banking operations. The First National Bank of Boston, for example, spun off its investment banking operations into the First Boston Corporation, now part of one of the most important investment banking firms in America, Credit Suisse First Boston. Investment banking firms typically discontinued their deposit business, although J. P. Morgan discontinued its investment banking business and reorganized as a commercial bank; however, some senior officers of J. P. Morgan went on to organize Morgan Stanley, another one of the largest investment banking firms today.

Multiple Regulatory Agencies

Commercial bank regulation in the United States has developed into a crazy quilt of multiple regulatory agencies with overlapping jurisdictions. The Office of the Comptroller of the Currency has the primary supervisory responsibility for the 2,100 national banks that own more than half of the assets in the commercial banking system. The Federal Reserve and the state banking authorities have joint primary responsibility for the 1,200 state banks that are members of the Federal Reserve System. The Fed also

has regulatory responsibility over companies that own one or more banks (called **bank holding companies**) and secondary responsibility for the national banks. The FDIC and the state banking authorities jointly supervise the 5,800 state banks that have FDIC insurance but are not members of the Federal Reserve System. The state banking authorities have sole jurisdiction over the fewer than 500 state banks without FDIC insurance. (Such banks hold less than 0.2% of the deposits in the commercial banking system.)

If you find the U.S. bank regulatory system confusing, imagine how confusing it is for the banks, which have to deal with multiple regulatory agencies. Several proposals have been raised by the U.S. Treasury to rectify this situation by centralizing the regulation of all depository institutions under one independent agency. However, none of these proposals has been successful in Congress, and whether there will be regulatory consolidation in the future is highly uncertain.

Financial Innovation and the Evolution of the Banking Industry

To understand how the banking industry has evolved over time, we must first understand the process of financial innovation, which has transformed the entire financial system. Like other industries, the financial industry is in business to earn profits by selling its products. If a soap company perceives that there is a need in the marketplace for a laundry detergent with fabric softener, it develops a product to fit the need. Similarly, to maximize their profits, financial institutions develop new products to satisfy their own needs as well as those of their customers; in other words, innovation—which can be extremely beneficial to the economy—is driven by the desire to get (or stay) rich. This view of the innovation process leads to the following simple analysis: ***A change in the financial environment will stimulate a search by financial institutions for innovations that are likely to be profitable.***

Starting in the 1960s, individuals and financial institutions operating in financial markets were confronted with drastic changes in the economic environment: Inflation and interest rates climbed sharply and became harder to predict, a situation that changed demand conditions in financial markets. The rapid advance in computer technology changed supply conditions. In addition, financial regulations became more burdensome. Financial institutions found that many of the old ways of doing business were no longer profitable; the financial services and products they had been offering to the public were not selling. Many financial intermediaries found that they were no longer able to acquire funds with their traditional financial instruments, and without these funds they would soon be out of business. To survive in the new economic environment, financial institutions had to research and develop new products and services that would meet customer needs and prove profitable, a process referred to as **financial engineering**. In their case, necessity was the mother of innovation.

Our discussion of why financial innovation occurs suggests that there are three basic types of financial innovation: responses to changes in demand conditions, responses to changes in supply conditions, and avoidance of regulations. Now that we have a framework for understanding why financial institutions produce innovations, let's look at examples of how financial institutions in their search for profits have produced financial innovations of the three basic types.

Responses to Changes in Demand Conditions: Interest Rate Volatility

The most significant change in the economic environment that altered the demand for financial products in recent years has been the dramatic increase in the volatility of interest rates. In the 1950s, the interest rate on three-month Treasury bills fluctuated between 1.0% and 3.5%; in the 1970s, it fluctuated between 4.0% and 11.5%; in the 1980s, it ranged from 5% to over 15%. Large fluctuations in interest rates lead to substantial capital gains or losses and greater uncertainty about returns on investments. Recall that the risk that is related to the uncertainty about interest-rate movements and returns is called *interest-rate risk*, and high volatility of interest rates, such as we saw in the 1970s and 1980s, leads to a higher level of interest-rate risk.

We would expect the increase in interest-rate risk to increase the demand for financial products and services that could reduce that risk. This change in the economic environment would thus stimulate a search for profitable innovations by financial institutions that meet this new demand and would spur the creation of new financial instruments that help lower interest-rate risk. Two examples of financial innovations that appeared in the 1970s confirm this prediction: the development of adjustable-rate mortgages and financial derivations.

Adjustable-Rate Mortgages. Like other investors, financial institutions find that lending is more attractive if interest-rate risk is lower. They would not want to make a mortgage loan at a 10% interest rate and two months later find that they could obtain 12% in interest on the same mortgage. To reduce interest-rate risk, in 1975 savings and loans in California began to issue adjustable-rate mortgages; that is, mortgage loans on which the interest rate changes when a market interest rate (usually the Treasury bill rate) changes. Initially, an adjustable-rate mortgage might have a 5% interest rate. In six months, this interest rate might increase or decrease by the amount of the increase or decrease in, say, the six-month Treasury bill rate, and the mortgage payment would change. Because adjustable-rate mortgages allow mortgage-issuing institutions to earn higher interest rates on mortgages when rates rise, profits are kept higher during these periods.

This attractive feature of adjustable-rate mortgages has encouraged mortgage-issuing institutions to issue adjustable-rate mortgages with lower initial interest rates than on conventional fixed-rate mortgages, making them popular with many households. However, because the mortgage payment on a variable-rate mortgage can increase, many households continue to prefer fixed-rate mortgages. Hence both types of mortgages are widespread.

Financial Derivatives. Given the greater demand for the reduction of interest-rate risk, commodity exchanges such as the Chicago Board of Trade recognized that if they could develop a product that would help investors and financial institutions to protect themselves from, or **hedge**, interest-rate risk, then they could make profits by selling this new instrument. **Futures contracts**, in which the seller agrees to provide a certain standardized commodity to the buyer on a specific future date at an agreed-on price, had been around for a long time. Officials at the Chicago Board of Trade realized that if they created futures contracts in financial instruments, which are called **financial derivatives** because their payoffs are linked to previously issued securities, they could be used to hedge risk. Thus in 1975, financial derivatives were born. We will study financial derivatives later in the book, in Chapter 13.

Responses to Changes in Supply Conditions: Information Technology

The most important source of the changes in supply conditions that stimulate financial innovation has been the improvement in computer and telecommunications technology. This technology, called *information technology*, has had two effects. First, it has lowered the cost of processing financial transactions, making it profitable for financial institutions to create new financial products and services for the public. Second, it has made it easier for investors to acquire information, thereby making it easier for firms to issue securities. The rapid developments in information technology have resulted in many new financial products and services that we examine here.

Bank Credit and Debit Cards. Credit cards have been around since well before World War II. Many individual stores (Sears, Macy's, Goldwater's) institutionalized charge accounts by providing customers with credit cards that allowed them to make purchases at these stores without cash. Nationwide credit cards were not established until after World War II, when Diners Club developed one to be used in restaurants all over the country (and abroad). Similar credit card programs were started by American Express and Carte Blanche, but because of the high cost of operating these programs, cards were issued only to selected persons and businesses that could afford expensive purchases.

A firm issuing credit cards earns income from loans it makes to credit card holders and from payments made by stores on credit card purchases (a percentage of the purchase price, say 5%). A credit card program's costs arise from loan defaults, stolen cards, and the expense involved in processing credit card transactions.

Seeing the success of Diners Club, American Express, and Carte Blanche, bankers wanted to share in the profitable credit card business. Several commercial banks attempted to expand the credit card business to a wider market in the 1950s, but the cost per transaction of running these programs was so high that their early attempts failed.

In the late 1960s, improved computer technology, which lowered the transaction costs for providing credit card services, made it more likely that bank credit card programs would be profitable. The banks tried to enter this business again, and this time their efforts led to the creation of two successful bank credit card programs: BankAmericard (originally started by the Bank of America but now an independent organization called Visa) and MasterCharge (now MasterCard, run by the Interbank Card Association). These programs have become phenomenally successful; more than 200 million of their cards are in use. Indeed, bank credit cards have been so profitable that nonfinancial institutions such as Sears (which launched the Discover card), General Motors, and AT&T have also entered the credit card business. Consumers have benefited because credit cards are more widely accepted than checks to pay for purchases (particularly abroad), and they allow consumers to take out loans more easily.

The success of bank credit cards has led these institutions to come up with a new financial innovation, *debit cards*. Debit cards often look just like credit cards and can be used to make purchases in an identical fashion. However, in contrast to credit cards, which extend the purchaser a loan that does not have to be paid off immediately, a debit card purchase is immediately deducted from the card holder's bank account. Debit cards depend even more on low costs of processing transactions, since their profits are generated entirely from the fees paid by merchants on debit card purchases at their stores. Debit cards have grown increasingly popular in recent years.

Electronic Banking. The wonders of modern computer technology have also enabled banks to lower the cost of bank transactions by having the customer interact with an

electronic banking (e-banking) facility rather than with a human being. One important form of an e-banking facility is the **automated teller machine (ATM)**, an electronic machine that allows customers to get cash, make deposits, transfer funds from one account to another, and check balances. The ATM has the advantage that it does not have to be paid overtime and never sleeps, thus being available for use 24 hours a day. Not only does this result in cheaper transactions for the bank, but it also provides more convenience for the customer. Furthermore, because of their low cost, ATMs can be put at locations other than a bank or its branches, further increasing customer convenience. The low cost of ATMs has meant that they have sprung up everywhere and now number over 250,000 in the United States alone. Furthermore, it is now as easy to get foreign currency from an ATM when you are traveling in Europe as it is to get cash from your local bank. In addition, transactions with ATMs are so much cheaper for the bank than ones conducted with human tellers that some banks charge customers less if they use the ATM than if they use a human teller.

With the drop in the cost of telecommunications, banks have developed another financial innovation, *home banking*. It is now cost-effective for banks to set up an electronic banking facility in which the bank's customer is linked up with the bank's computer to carry out transactions by using either a telephone or a personal computer. Now a bank's customers can conduct many of their bank transactions without ever leaving the comfort of home. The advantage for the customer is the convenience of home banking, while banks find that the cost of transactions is substantially less than having the customer come to the bank. The success of ATMs and home banking has led to another innovation, the **automated banking machine (ABM)**, which combines in one location an ATM, an Internet connection to the bank's web site, and a telephone link to customer service.

With the decline in the price of personal computers and their increasing presence in the home, we have seen a further innovation in the home banking area, the appearance of a new type of banking institution, the **virtual bank**, a bank that has no physical location but rather exists only in cyberspace. In 1995, Security First Network Bank, based in Atlanta but now owned by Royal Bank of Canada, became the first virtual bank, planning to offer an array of banking services on the Internet—accepting checking account and savings deposits, selling certificates of deposits, issuing ATM cards, providing bill-paying facilities, and so on. The virtual bank thus takes home banking one step further, enabling the customer to have a full set of banking services at home 24 hours a day. In 1996, Bank of America and Wells Fargo entered the virtual banking market, to be followed by many others, with Bank of America now being the largest Internet bank in the United States. Will virtual banking be the predominant form of banking in the future (see Box 1)?

Junk Bonds. Before the advent of computers and advanced telecommunications, it was difficult to acquire information about the financial situation of firms that might want to sell securities. Because of the difficulty in screening out bad from good credit risks, the only firms that were able to sell bonds were very well established corporations that had high credit ratings.[1] Before the 1980s, then, only corporations that could issue bonds with ratings of Baa or above could raise funds by selling newly issued bonds. Some firms that had fallen on bad times, so-called *fallen angels*, had previously

[1]The discussion of adverse selection problems in Chapter 8 provides a more detailed analysis of why only well-established firms with high credit ratings were able to sell securities.

Box 1: E-Finance

Will "Clicks" Dominate "Bricks" in the Banking Industry?

With the advent of virtual banks ("clicks") and the convenience they provide, a key question is whether they will become the primary form in which banks do their business, eliminating the need for physical bank branches ("bricks") as the main delivery mechanism for banking services. Indeed, will stand-alone Internet banks be the wave of the future?

The answer seems to be no. Internet-only banks such as Wingspan (owned by Bank One), First-e (Dublin-based), and Egg (a British Internet-only bank owned by Prudential) have had disappointing revenue growth and profits. The result is that pure online banking has not been the success that proponents had hoped for. Why has Internet banking been a disappointment?

There have been several strikes against Internet banking. First, bank depositors want to know that their savings are secure, and so are reluctant to put their money into new institutions without a long track record. Second, customers worry about the security of their online transactions and whether their transactions will truly be kept private. Traditional banks are viewed as being more secure and trustworthy in terms of releasing private information. Third, customers may prefer services provided by physical branches. For example, banking customers seem to prefer to purchase long-term savings products face-to-face. Fourth, Internet banking has run into technical problems—server crashes, slow connections over phone lines, mistakes in conducting transactions—that will probably diminish over time as technology improves.

The wave of the future thus does not appear to be pure Internet banks. Instead it looks like "clicks and bricks" will be the predominant form of banking, in which online banking is used to complement the services provided by traditional banks. Nonetheless, the delivery of banking services is undergoing massive changes, with more and more banking services delivered over the Internet and the number of physical bank branches likely to decline in the future.

issued long-term corporate bonds that now had ratings that had fallen below Baa, bonds that were pejoratively dubbed "junk bonds."

With the improvement in information technology in the 1970s, it became easier for investors to screen out bad from good credit risks, thus making it more likely that they would buy long-term debt securities from less well known corporations with lower credit ratings. With this change in supply conditions, we would expect that some smart individual would pioneer the concept of selling new public issues of junk bonds, not for fallen angels but for companies that had not yet achieved investment-grade status. This is exactly what Michael Milken of Drexel Burnham, an investment banking firm, started to do in 1977. Junk bonds became an important factor in the corporate bond market, with the amount outstanding exceeding $200 billion by the late 1980s. Although there was a sharp slowdown in activity in the junk bond market after Milken was indicted for securities law violations in 1989, it heated up again in the 1990s.

Commercial Paper Market. *Commercial paper* is a short-term debt security issued by large banks and corporations. The commercial paper market has undergone tremendous growth since 1970, when there was $33 billion outstanding, to over $1.3 trillion outstanding at the end of 2002. Indeed, commercial paper has been one of the fastest-growing money market instruments.

Improvements in information technology also help provide an explanation for the rapid rise of the commercial paper market. We have seen that the improvement in information technology made it easier for investors to screen out bad from good credit risks, thus making it easier for corporations to issue debt securities. Not only did this make it easier for corporations to issue long-term debt securities as in the junk bond market, but it also meant that they could raise funds by issuing short-term debt securities like commercial paper more easily. Many corporations that used to do their short-term borrowing from banks now frequently raise short-term funds in the commercial paper market instead.

The development of money market mutual funds has been another factor in the rapid growth in the commercial paper market. Because money market mutual funds need to hold liquid, high-quality, short-term assets such as commercial paper, the growth of assets in these funds to around $2.1 trillion has created a ready market in commercial paper. The growth of pension and other large funds that invest in commercial paper has also stimulated the growth of this market.

Securitization. An important example of a financial innovation arising from improvements in both transaction and information technology is securitization, one of the most important financial innovations in the past two decades. **Securitization** is the process of transforming otherwise illiquid financial assets (such as residential mortgages, auto loans, and credit card receivables), which have typically been the bread and butter of banking institutions, into marketable capital market securities. As we have seen, improvements in the ability to acquire information have made it easier to sell marketable capital market securities. In addition, with low transaction costs because of improvements in computer technology, financial institutions find that they can cheaply bundle together a portfolio of loans (such as mortgages) with varying small denominations (often less than $100,000), collect the interest and principal payments on the mortgages in the bundle, and then "pass them through" (pay them out) to third parties. By dividing the portfolio of loans into standardized amounts, the financial institution can then sell the claims to these interest and principal payments to third parties as securities. The standardized amounts of these securitized loans make them liquid securities, and the fact that they are made up of a bundle of loans helps diversify risk, making them desirable. The financial institution selling the securitized loans makes a profit by servicing the loans (collecting the interest and principal payments and paying them out) and charging a fee to the third party for this service.

Avoidance of Existing Regulations

The process of financial innovation we have discussed so far is much like innovation in other areas of the economy: It occurs in response to changes in demand and supply conditions. However, because the financial industry is more heavily regulated than other industries, government regulation is a much greater spur to innovation in this industry. Government regulation leads to financial innovation by creating incentives for firms to skirt regulations that restrict their ability to earn profits. Edward Kane, an economist at Boston College, describes this process of avoiding regulations as "loophole mining." The economic analysis of innovation suggests that when the economic environment changes such that regulatory constraints are so burdensome that large profits can be made by avoiding them, loophole mining and innovation are more likely to occur.

Because banking is one of the most heavily regulated industries in America, loophole mining is especially likely to occur. The rise in inflation and interest rates from

the late 1960s to 1980 made the regulatory constraints imposed on this industry even more burdensome, leading to financial innovation.

Two sets of regulations have seriously restricted the ability of banks to make profits: reserve requirements that force banks to keep a certain fraction of their deposits as reserves (vault cash and deposits in the Federal Reserve System) and restrictions on the interest rates that can be paid on deposits. For the following reasons, these regulations have been major forces behind financial innovation.

1. Reserve requirements. The key to understanding why reserve requirements led to financial innovation is to recognize that they act, in effect, as a tax on deposits. Because the Fed does not pay interest on reserves, the opportunity cost of holding them is the interest that a bank could otherwise earn by lending the reserves out. For each dollar of deposits, reserve requirements therefore impose a cost on the bank equal to the interest rate, i, that could be earned if the reserves could be lent out times the fraction of deposits required as reserves, r. The cost of $i \times r$ imposed on the bank is just like a tax on bank deposits of $i \times r$.

It is a great tradition to avoid taxes if possible, and banks also play this game. Just as taxpayers look for loopholes to lower their tax bills, banks seek to increase their profits by mining loopholes and by producing financial innovations that allow them to escape the tax on deposits imposed by reserve requirements.

2. Restrictions on interest paid on deposits. Until 1980, legislation prohibited banks in most states from paying interest on checking account deposits, and through Regulation Q, the Fed set maximum limits on the interest rate that could be paid on time deposits. To this day, banks are not allowed to pay interest on corporate checking accounts. The desire to avoid these **deposit rate ceilings** also led to financial innovations.

If market interest rates rose above the maximum rates that banks paid on time deposits under Regulation Q, depositors withdrew funds from banks to put them into higher-yielding securities. This loss of deposits from the banking system restricted the amount of funds that banks could lend (called **disintermediation**) and thus limited bank profits. Banks had an incentive to get around deposit rate ceilings, because by so doing, they could acquire more funds to make loans and earn higher profits.

We can now look at how the desire to avoid restrictions on interest payments and the tax effect of reserve requirements led to two important financial innovations.

Money Market Mutual Funds. Money market mutual funds issue shares that are redeemable at a fixed price (usually $1) by writing checks. For example, if you buy 5,000 shares for $5,000, the money market fund uses these funds to invest in short-term money market securities (Treasury bills, certificates of deposit, commercial paper) that provide you with interest payments. In addition, you are able to write checks up to the $5,000 held as shares in the money market fund. Although money market fund shares effectively function as checking account deposits that earn interest, they are not legally deposits and so are not subject to reserve requirements or prohibitions on interest payments. For this reason, they can pay higher interest rates than deposits at banks.

The first money market mutual fund was created by two Wall Street mavericks, Bruce Bent and Henry Brown, in 1971. However, the low market interest rates from 1971 to 1977 (which were just slightly above Regulation Q ceilings of 5.25 to 5.5%) kept them from being particularly advantageous relative to bank deposits. In early 1978, the situation changed rapidly as market interest rates began to climb over 10%,

well above the 5.5% maximum interest rates payable on savings accounts and time deposits under Regulation Q. In 1977, money market mutual funds had assets under $4 billion; in 1978, their assets climbed to close to $10 billion; in 1979, to over $40 billion; and in 1982, to $230 billion. Currently, their assets are around $2 trillion. To say the least, money market mutual funds have been a successful financial innovation, which is exactly what we would have predicted to occur in the late 1970s and early 1980s when interest rates soared beyond Regulation Q ceilings.

Sweep Accounts. Another innovation that enables banks to avoid the "tax" from reserve requirements is the **sweep account**. In this arrangement, any balances above a certain amount in a corporation's checking account at the end of a business day are "swept out" of the account and invested in overnight securities that pay the corporation interest. Because the "swept out" funds are no longer classified as checkable deposits, they are not subject to reserve requirements and thus are not "taxed." They also have the advantage that they allow banks in effect to pay interest on these corporate checking accounts, which otherwise is not allowed under existing regulations. Because sweep accounts have become so popular, they have lowered the amount of required reserves to the degree that most banking institutions do not find reserve requirements binding: In other words, they voluntarily hold more reserves than they are required to.

The financial innovation of sweep accounts is particularly interesting because it was stimulated not only by the desire to avoid a costly regulation, but also by a change in supply conditions: in this case, information technology. Without low-cost computers to process inexpensively the additional transactions required by these accounts, this innovation would not have been profitable and therefore would not have been developed. Technological factors often combine with other incentives, such as the desire to get around a regulation, to produce innovation.

Financial Innovation and the Decline of Traditional Banking

The traditional financial intermediation role of banking has been to make long-term loans and to fund them by issuing short-term deposits, a process of asset transformation commonly referred to as "borrowing short and lending long." Here we examine how financial innovations have created a more competitive environment for the banking industry, causing the industry to change dramatically, with its traditional banking business going into decline.

www.financialservicefacts.org/international/INT-1.htm

Learn about the number of employees and the current profitability of commercial banks and saving institutions.

In the United States, the importance of commercial banks as a source of funds to nonfinancial borrowers has shrunk dramatically. As we can see in Figure 2, in 1974, commercial banks provided close to 40% of these funds; by 2002, their market share was down to below 30%. The decline in market share for thrift institutions has been even more precipitous: from more than 20% in the late 1970s to 6% today. Another way of viewing the declining role of banking in traditional financial intermediation is to look at the size of banks' balance sheet assets relative to those of other financial intermediaries (see Table 1 in Chapter 12, page 289). Commercial banks' share of total financial intermediary assets has fallen from about 40% in the 1960–1980 period to 30% by the end of 2002. Similarly, the share of total financial intermediary assets held by thrift institutions has declined even more from the 20% level of the 1960–1980 period to about 5% by 2002.

Clearly, the traditional financial intermediation role of banking, whereby banks make loans that are funded with deposits, is no longer as important in our financial system. However, the decline in the market share of banks in total lending and total financial intermediary assets does not necessarily indicate that the banking industry is

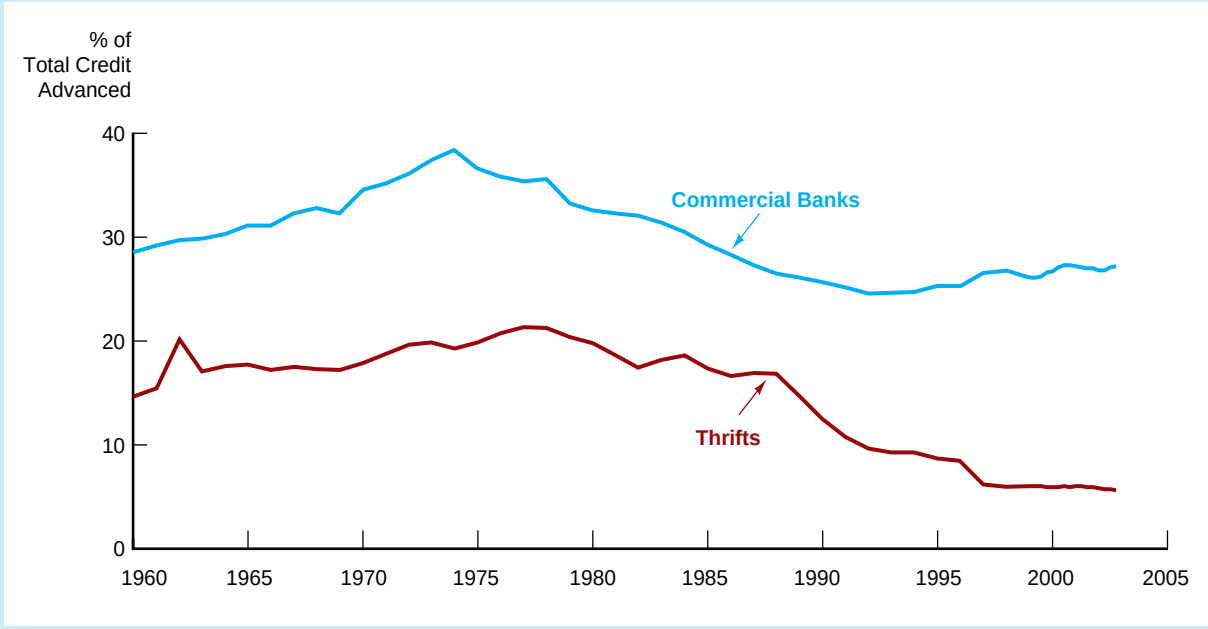


FIGURE 2 Bank Share of Total Nonfinancial Borrowing, 1960–2002

Source: Federal Reserve Flow of Funds Accounts; Federal Reserve *Bulletin*.

in decline. There is no evidence of a declining trend in bank profitability. However, overall bank profitability is not a good indicator of the profitability of traditional banking, because it includes an increasing amount of income from nontraditional off-balance-sheet activities, discussed in Chapter 9. Noninterest income derived from off-balance-sheet activities, as a share of total banking income, increased from around 7% in 1980 to more than 45% of total bank income today. Given that the overall profitability of banks has not risen, the increase in income from off-balance-sheet activities implies that the profitability of traditional banking business has declined. This decline in profitability then explains why banks have been reducing their traditional business.

To understand why traditional banking business has declined in both size and profitability, we need to look at how the financial innovations described earlier have caused banks to suffer declines in their cost advantages in acquiring funds, that is, on the liabilities side of their balance sheet, while at the same time they have lost income advantages on the assets side of their balance sheet. The simultaneous decline of cost and income advantages has resulted in reduced profitability of traditional banking and an effort by banks to leave this business and engage in new and more profitable activities.

Decline in Cost Advantages in Acquiring Funds (Liabilities). Until 1980, banks were subject to deposit rate ceilings that restricted them from paying any interest on checkable deposits and (under Regulation Q) limited them to paying a maximum interest rate of a little over 5% on time deposits. Until the 1960s, these restrictions worked to the

banks' advantage because their major source of funds (over 60%) was checkable deposits, and the zero interest cost on these deposits meant that the banks had a very low cost of funds. Unfortunately, this cost advantage for banks did not last. The rise in inflation from the late 1960s on led to higher interest rates, which made investors more sensitive to yield differentials on different assets. The result was the so-called disintermediation process, in which people began to take their money out of banks, with their low interest rates on both checkable and time deposits, and began to seek out higher-yielding investments. Also, as we have seen, at the same time, attempts to get around deposit rate ceilings and reserve requirements led to the financial innovation of money market mutual funds, which put the banks at an even further disadvantage because depositors could now obtain checking account–like services while earning high interest on their money market mutual fund accounts. One manifestation of these changes in the financial system was that the low-cost source of funds, checkable deposits, declined dramatically in importance for banks, falling from over 60% of bank liabilities to below 10% today.

The growing difficulty for banks in raising funds led to their supporting legislation in the 1980s that eliminated Regulation Q ceilings on time deposit interest rates and allowed checkable deposit accounts that paid interest. Although these changes in regulation helped make banks more competitive in their quest for funds, it also meant that their cost of acquiring funds had risen substantially, thereby reducing their earlier cost advantage over other financial institutions.

Decline in Income Advantages on Uses of Funds (Assets). The loss of cost advantages on the liabilities side of the balance sheet for American banks is one reason that they have become less competitive, but they have also been hit by a decline in income advantages on the assets side from the financial innovations we discussed earlier—junk bonds, securitization, and the rise of the commercial paper market.

We have seen that improvements in information technology have made it easier for firms to issue securities directly to the public. This has meant that instead of going to banks to finance short-term credit needs, many of the banks' best business customers now find it cheaper to go instead to the commercial paper market for funds. The loss of this competitive advantage for banks is evident in the fact that before 1970, nonfinancial commercial paper equaled less than 5% of commercial and industrial bank loans, whereas the figure has risen to 16% today. In addition, this growth in the commercial paper market has allowed finance companies, which depend primarily on commercial paper to acquire funds, to expand their operations at the expense of banks. Finance companies, which lend to many of the same businesses that borrow from banks, have increased their market share relative to banks: Before 1980, finance company loans to business equaled about 30% of commercial and industrial bank loans; currently, they are over 45%.

The rise of the junk bond market has also eaten into banks' loan business. Improvements in information technology have made it easier for corporations to sell their bonds to the public directly, thereby bypassing banks. Although Fortune 500 companies started taking this route in the 1970s, now lower-quality corporate borrowers are using banks less often because they have access to the junk bond market.

We have also seen that improvements in computer technology have led to securitization, whereby illiquid financial assets such as bank loans and mortgages are transformed into marketable securities. Computers enable other financial institutions to originate loans because they can now accurately evaluate credit risk with statistical

methods, while computers have lowered transaction costs, making it possible to bundle these loans and sell them as securities. When default risk can be easily evaluated with computers, banks no longer have an advantage in making loans. Without their former advantages, banks have lost loan business to other financial institutions even though the banks themselves are involved in the process of securitization. Securitization has been a particular problem for mortgage-issuing institutions such as S&Ls, because most residential mortgages are now securitized.

Banks' Responses. In any industry, a decline in profitability usually results in exit from the industry (often due to widespread bankruptcies) and a shrinkage of market share. This occurred in the banking industry in the United States during the 1980s via consolidations and bank failures (discussed in the next chapter).

In an attempt to survive and maintain adequate profit levels, many U.S. banks face two alternatives. First, they can attempt to maintain their traditional lending activity by expanding into new and riskier areas of lending. For example, U.S. banks increased their risk taking by placing a greater percentage of their total funds in commercial real estate loans, traditionally a riskier type of loan. In addition, they increased lending for corporate takeovers and leveraged buyouts, which are highly leveraged transaction loans. The decline in the profitability of banks' traditional business may thus have helped lead to the crisis in banking in the 1980s and early 1990s that we discuss in the next chapter.

The second way banks have sought to maintain former profit levels is to pursue new off-balance-sheet activities that are more profitable. U.S. commercial banks did this during the early 1980s, more than doubling the share of their income coming from off-balance-sheet, noninterest-income activities. This strategy, however, has generated concerns about what activities are proper for banks and whether nontraditional activities might be riskier, and thus result in excessive risk-taking by banks.

The decline of banks' traditional business has thus meant that the banking industry has been driven to seek out new lines of business. This could be beneficial because by so doing, banks can keep vibrant and healthy. Indeed, bank profitability has been high in recent years, and nontraditional, off-balance-sheet activities have been playing an important role in the resurgence of bank profits. However, there is a danger that the new directions in banking could lead to increased risk taking, and thus the decline in traditional banking requires regulators to be more vigilant. It also poses new challenges for bank regulators, who, as we will see in Chapter 11, must now be far more concerned about banks' off-balance-sheet activities.

Decline of Traditional Banking in Other Industrialized Countries. Forces similar to those in the United States have been leading to the decline of traditional banking in other industrialized countries. The loss of banks' monopoly power over depositors has occurred outside the United States as well. Financial innovation and deregulation are occurring worldwide and have created attractive alternatives for both depositors and borrowers. In Japan, for example, deregulation has opened a wide array of new financial instruments to the public, causing a disintermediation process similar to that in the United States. In European countries, innovations have steadily eroded the barriers that have traditionally protected banks from competition.

In other countries, banks have also faced increased competition from the expansion of securities markets. Both financial deregulation and fundamental economic

forces in other countries have improved the availability of information in securities markets, making it easier and less costly for firms to finance their activities by issuing securities rather than going to banks. Further, even in countries where securities markets have not grown, banks have still lost loan business because their best corporate customers have had increasing access to foreign and offshore capital markets, such as the Eurobond market. In smaller economies, like Australia, which still do not have well-developed corporate bond or commercial paper markets, banks have lost loan business to international securities markets. In addition, the same forces that drove the securitization process in the United States are at work in other countries and will undercut the profitability of traditional banking in these countries as well. The United States is not unique in seeing its banks face a more difficult competitive environment. Thus, although the decline of traditional banking has occurred earlier in the United States than in other countries, the same forces are causing a decline in traditional banking abroad.

Structure of the U.S. Commercial Banking Industry

www.fdic.gov/bank/statistical/statistics/index.html
Visit this web site to gather statistics on the banking industry.

There are approximately 8,000 commercial banks in the United States, far more than in any other country in the world. As Table 1 indicates, we have an extraordinary number of small banks. Ten percent of the banks have less than $25 million in assets. Far more typical is the size distribution in Canada or the United Kingdom, where five or fewer banks dominate the industry. In contrast, the ten largest commercial banks in the United States (listed in Table 2) together hold just 58% of the assets in their industry.

Most industries in the United States have far fewer firms than the commercial banking industry; typically, large firms tend to dominate these industries to a greater extent than in the commercial banking industry. (Consider the computer software

Table 1 Size Distribution of Insured Commercial Banks, September 30, 2002

Assets	Number of Banks	Share of Banks (%)	Share of Assets Held (%)
Less than $25 million	796	10.0	0.2
$25–$50 million	1,421	17.9	0.8
$50–$100 million	2,068	26.1	2.2
$100–$500 million	2,868	36.2	8.6
$500 million–$1 billion	381	4.8	3.7
$1–$10 billion	319	4.0	13.2
More than $10 billion	80	1.0	71.3
Total	7,933	100.0	100.0

Source: www.fdic.gov/bank/statistical/statistics/0209/allstru.html.

Table 2 Ten Largest U.S. Banks, February 2003

Bank	Assets ($ millions)	Share of All Commercial Bank Assets (%)
1. Citibank, National Association, New York	1,057,657	15.19
2. JP Morgan Chase, New York	712,508	10.23
3. Bank of America, National Association, Charlotte, N.C.	619,921	8.90
4. Wachovia National Bank, Charlotte, N.C.	319,853	4.59
5. Wells Fargo, National Association, San Francisco	311,509	4.47
6. Bank One, National Association, Chicago	262,947	3.77
7. Taunus Corporation, New York	235,867	3.39
8. Fleet National Bank, Providence, R.I.	192,032	2.76
9. ABN Amro, North America, Chicago	174,451	2.50
10. US Bancorp, Minneapolis, Minnesota	164,745	2.36
Total	4,051,490	58.16

Source: www.infoplease.com/pia/A0763206.html.

industry, which is dominated by Microsoft, or the automobile industry, which is dominated by General Motors, Ford, Daimler-Chrysler, Toyota, and Honda.) Does the large number of banks in the commercial banking industry and the absence of a few dominant firms suggest that commercial banking is more competitive than other industries?

Restrictions on Branching

The presence of so many commercial banks in the United States actually reflects past regulations that restricted the ability of these financial institutions to open **branches** (additional offices for the conduct of banking operations). Each state had its own regulations on the type and number of branches that a bank could open. Regulations on both coasts, for example, tended to allow banks to open branches throughout a state; in the middle part of the country, regulations on branching were more restrictive. The McFadden Act of 1927, which was designed to put national banks and state banks on an equal footing (and the Douglas Amendment of 1956, which closed a loophole in the McFadden Act) effectively prohibited banks from branching across state lines and forced all national banks to conform to the branching regulations in the state of their location.

The McFadden Act and state branching regulations constituted strong anticompetitive forces in the commercial banking industry, allowing many small banks to stay in existence, because larger banks were prevented from opening a branch nearby. If competition is beneficial to society, why have regulations restricting branching arisen in America? The simplest explanation is that the American public has historically been hostile to large banks. States with the most restrictive branching regulations were typically ones in which populist antibank sentiment was strongest in the nineteenth cen-

tury. (These states usually had large farming populations whose relations with banks periodically became tempestuous when banks would foreclose on farmers who couldn't pay their debts.) The legacy of nineteenth-century politics was a banking system with restrictive branching regulations and hence an inordinate number of small banks. However, as we will see later in this chapter, branching restrictions have been eliminated, and we are heading toward nationwide banking.

Response to Branching Restrictions

An important feature of the U.S. banking industry is that competition can be repressed by regulation but not completely quashed. As we saw earlier in this chapter, the existence of restrictive regulation stimulates financial innovations that get around these regulations in the banks' search for profits. Regulations restricting branching have stimulated similar economic forces and have promoted the development of two financial innovations: bank holding companies and automated teller machines.

Bank Holding Companies. A holding company is a corporation that owns several different companies. This form of corporate ownership has important advantages for banks. It has allowed them to circumvent restrictive branching regulations, because the holding company can own a controlling interest in several banks even if branching is not permitted. Furthermore, a bank holding company can engage in other activities related to banking, such as the provision of investment advice, data processing and transmission services, leasing, credit card services, and servicing of loans in other states.

The growth of the bank holding companies has been dramatic over the past three decades. Today bank holding companies own almost all large banks, and over 90% of all commercial bank deposits are held in banks owned by holding companies.

Automated Teller Machines. Another financial innovation that avoided the restrictions on branching is the automated teller machine (ATM). Banks realized that if they did not own or rent the ATM, but instead let it be owned by someone else and paid for each transaction with a fee, the ATM would probably not be considered a branch of the bank and thus would not be subject to branching regulations. This is exactly what the regulatory agencies and courts in most states concluded. Because they enable banks to widen their markets, a number of these shared facilities (such as Cirrus and NYCE) have been established nationwide. Furthermore, even when an ATM is owned by a bank, states typically have special provisions that allow wider establishment of ATMs than is permissible for traditional "brick and mortar" branches.

As we saw earlier in this chapter, avoiding regulation was not the only reason for the development of the ATM. The advent of cheaper computer and telecommunications technology enabled banks to provide ATMs at low cost, making them a profitable innovation. This example further illustrates that technological factors often combine with incentives such as the desire to avoid restrictive regulations like branching restrictions to produce financial innovation.

Bank Consolidation and Nationwide Banking

As we can see in Figure 3, after a remarkable period of stability from 1934 to the mid-1980s, the number of commercial banks began to fall dramatically. Why has this sudden decline taken place?

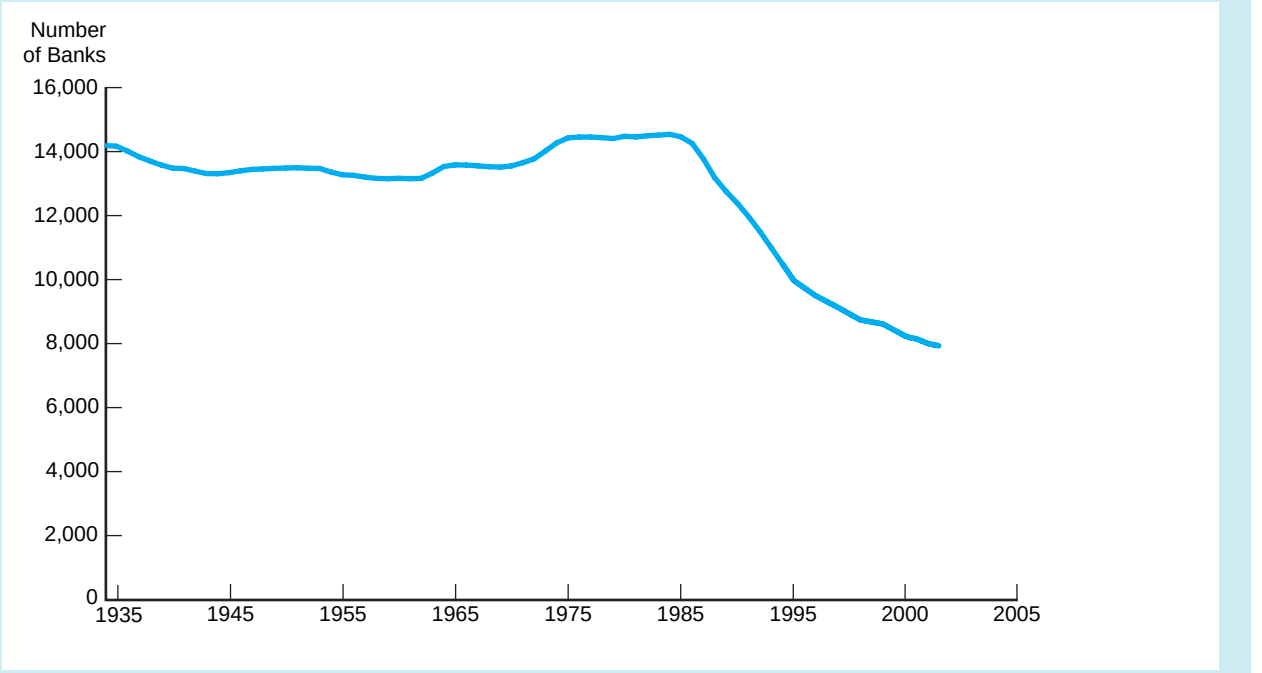


FIGURE 3 Number of Insured Commercial Banks in the United States, 1934–2002

Source: www2.fdic.gov/qbp/qbpSelect.asp?menuitem=STAT.

The banking industry hit some hard times in the 1980s and early 1990s, with bank failures running at a rate of over 100 per year from 1985 to 1992 (more on this later in the chapter and in Chapter 11). But bank failures are only part of the story. In the years 1985–1992, the number of banks declined by 3,000—more than double the number of failures. And in the period 1992–2002, when the banking industry returned to health, the number of commercial banks declined by a little over 4,100, less than 5% of which were bank failures, and most of these were of small banks. Thus we see that bank failures played an important, though not predominant, role in the decline in the number of banks in the 1985–1992 period and an almost negligible role in the decline in the number of banks since then.

So what explains the rest of the story? The answer is bank consolidation. Banks have been merging to create larger entities or have been buying up other banks. This gives rise to a new question: Why has bank consolidation been taking place in recent years?

As we have seen, loophole mining by banks has reduced the effectiveness of branching restrictions, with the result that many states have recognized that it would be in their best interest if they allowed ownership of banks across state lines. The result has been the formation of reciprocal regional compacts in which banks in one state are allowed to own banks in other states in the region. In 1975, Maine enacted the first interstate banking legislation that allowed out-of-state bank holding companies to purchase banks in that state. In 1982, Massachusetts enacted a regional compact with other New England states to allow interstate banking, and many other regional com-

pacts were adopted thereafter until by the early 1990s, almost all states allowed some form of interstate banking.

With the barriers to interstate banking breaking down in the early 1980s, banks recognized that they could gain the benefits of diversification because they would now be able to make loans in many states rather than just one. This gave them the advantage that if one state's economy was weak, another in which they operated might be strong, thus decreasing the likelihood that loans in different states would default at the same time. In addition, allowing banks to own banks in other states meant that they could take advantage of economies of scale by increasing their size through out-of-state acquisition of banks or by merging with banks in other states. Mergers and acquisitions explain the first phase of banking consolidation, which has played such an important role in the decline in the number of banks since 1985. Another result of the loosening of restrictions on interstate branching is the development of a new class of bank, the so-called **superregional banks**, bank holding companies that have begun to rival the money center banks in size but whose headquarters are not in one of the money center cities (New York, Chicago, and San Francisco). Examples of these superregional banks are Bank of America of Charlotte, North Carolina, and Banc One of Columbus, Ohio.

Not surprisingly, the advent of the Web and improved computer technology is another factor driving bank consolidation. Economies of scale have increased, because large upfront investments are required to set up many information technology platforms for financial institutions (see Box 2). To take advantage of these economies of scale, banks have needed to get bigger, and this development has led to additional

Box 2: E-Finance

Information Technology and Bank Consolidation

Achieving low costs in banking requires huge investments in information technology. In turn, such enormous investments require a business line of very large scale. This has been particularly true in the credit card business in recent years, in which huge technology investments have been made to provide customers with convenient web sites and to develop better systems to handle processing and risk analysis for both credit and fraud risk. The result has been substantial consolidation: As recently as 1995, the top five banking institutions issuing credit cards held less than 40% of total credit card debt, while today this number is above 60%.

Information technology has also spurred increasing consolidation of the bank custody business. Banks hold the actual certificate for investors when they purchase a stock or bond and provide data on the value of these securities and how much risk an investor is facing. Because this business is also computer-intensive, it also requires very large-scale investments in computer technology in order for the bank to offer these services at competitive rates. The percentage of assets at the top ten custody banks has therefore risen from 40% in 1990 to more than 90% today.

The increasing importance of e-finance, in which the computer is playing a more central role in delivering financial services, is bringing tremendous changes to the structure of the banking industry. Although banks are more than willing to offer a full range of products to their customers, they no longer find it profitable to produce all of them. Instead, they are contracting out the business, a practice that will lead to further consolidation of technology-intensive banking businesses in the future.

consolidation. Information technology has also been increasing **economies of scope**, the ability to use one resource to provide many different products and services. For example, details about the quality and creditworthiness of firms not only inform decisions about whether to make loans to them, but also can be useful in determining at what price their shares should trade. Similarly, once you have marketed one financial product to an investor, you probably know how to market another. Business people describe economies of scope by saying that there are "synergies" between different lines of business, and information technology is making these synergies more likely. The result is that consolidation is taking place not only to make financial institutions bigger, but also to increase the combination of products and services they can provide. This consolidation has had two consequences. First, different types of financial intermediaries are encroaching on each other's territory, making them more alike. Second, consolidation has led to the development of what the Federal Reserve has named **large, complex, banking organizations (LCBOs)**. This development has been facilitated by the repeal of the Glass-Steagall restrictions on combinations of banking and other financial service industries discussed in the next section.

The Riegle-Neal Interstate Banking and Branching Efficiency Act of 1994

Banking consolidation has been given further stimulus by the passage in 1994 of the Riegle-Neal Interstate Banking and Branching Efficiency Act. This legislation expands the regional compacts to the entire nation and overturns the McFadden Act and Douglas Amendment's prohibition of interstate banking. Not only does this act allow bank holding companies to acquire banks in any other state, notwithstanding any state laws to the contrary, but bank holding companies can merge the banks they own into one bank with branches in different states. States also have the option of opting out of interstate branching, a choice only Texas has made.

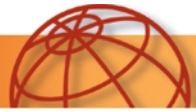
The Riegle-Neal Act finally establishes the basis for a true nationwide banking system. Although interstate banking was accomplished previously by out-of-state purchase of banks by bank holding companies, up until 1994 interstate branching was virtually nonexistent, because very few states had enacted interstate branching legislation. Allowing banks to conduct interstate banking through branching is especially important, because many bankers feel that economies of scale cannot be fully exploited through the bank holding company structure, but only through branching networks in which all of the bank's operations are fully coordinated.

Nationwide banks are now emerging. With the merger in 1998 of Bank of America and NationsBank, which created the first bank with branches on both coasts, consolidation in the banking industry is leading to banking organizations with operations in almost all of the fifty states.

What Will the Structure of the U.S. Banking Industry Look Like in the Future?

With true nationwide banking in the U.S. becoming a reality, the benefits of bank consolidation for the banking industry have increased substantially, thus driving the next phase of mergers and acquisitions and accelerating the decline in the number of commercial banks. With great changes occurring in the structure of this industry, the question naturally arises: What will the industry look like in ten years?

One view is that the industry will become more like that in many other countries (see Box 3) and we will end up with only a couple of hundred banks. A more extreme view is that the industry will look like that of Canada or the United Kingdom, with a few large banks dominating the industry. Research on this question, however, comes up with a different answer. The structure of the U.S. banking industry will still be unique, but not to the degree it once was. Most experts predict that



Box 3: Global

Comparison of Banking Structure in the United States and Abroad

The structure of the commercial banking industry in the United States is radically different from that in other industrialized nations. The United States is the only country that is just now developing a true national banking system in which banks have branches throughout the country. One result is that there are many more banks in the United States than in other industrialized countries. In contrast to the United States, which has on the order of 8,000 commercial banks, every other industrialized country has well under 1,000. Japan, for example, has fewer than 100 commercial banks—a mere fraction of the number in the United States, even though its economy and population are half the size of the United States. Another result of the past restrictions on branching in the United States is that our banks tend to be much smaller than those in other countries.

the consolidation surge will settle down as the U.S. banking industry approaches several thousand, rather than several hundred, banks.[2]

Banking consolidation will result not only in a smaller number of banks, but as the mergers between Chase Manhattan Bank and Chemical Bank and between Bank of America and NationsBank suggest, a shift in assets from smaller banks to larger banks as well. Within ten years, the share of bank assets in banks with less than $100 million in assets is expected to halve, while the amount at the so-called megabanks, those with over $100 billion in assets, is expected to more than double. Indeed, some analysts have predicted that we won't have long to wait before the first trillion-dollar bank emerges in the United States.

Are Bank Consolidation and Nationwide Banking Good Things?

Advocates of nationwide banking believe that it will produce more efficient banks and a healthier banking system less prone to bank failures. However, critics of bank consolidation fear that it will eliminate small banks, referred to as *community banks*, and that this will result in less lending to small businesses. In addition, they worry that a few banks will come to dominate the industry, making the banking business less competitive.

Most economists are skeptical of these criticisms of bank consolidation. As we have seen, research indicates that even after bank consolidation is completed, the United States will still have plenty of banks. The banking industry will thus remain highly competitive, probably even more so than now considering that banks that have been protected from competition from out-of-state banks will now have to compete with them vigorously to stay in business.

---

[2]For example, see Allen N. Berger, Anil K. Kashyap, and Joseph Scalise, "The Transformation of the U.S. Banking Industry: What a Long, Strange Trip It's Been," *Brookings Papers on Economic Activity* 2 (1995): 55–201, and Timothy Hannan and Stephen Rhoades, "Future U.S. Banking Structure, 1990–2010," *Antitrust Bulletin* 37 (1992) 737–798. For a more detailed treatment of the bank consolidation process taking place in the United States, see Frederic S. Mishkin, "Bank Consolidation: A Central Banker's Perspective," in *Mergers of Financial Institutions*, ed. Yakov Amihud and Geoffrey Wood (Boston: Kluwer Academic Publishers, 1998), pp. 3–19.

It also does not look as though community banks will disappear. When New York State liberalized branching laws in 1962, there were fears that community banks upstate would be driven from the market by the big New York City banks. Not only did this not happen, but some of the big boys found that the small banks were able to run rings around them in the local markets. Similarly, California, which has had unrestricted statewide branching for a long time, continues to have a thriving number of community banks.

Economists see some important benefits of bank consolidation and nationwide banking. The elimination of geographic restrictions on banking will increase competition and drive inefficient banks out of business, thus raising the efficiency of the banking sector. The move to larger banking organizations also means that there will be some increase in efficiency because they can take advantage of economies of scale and scope. The increased diversification of banks' loan portfolios may lower the probability of a banking crisis in the future. In the 1980s and early 1990s, bank failures were often concentrated in states with weak economies. For example, after the decline in oil prices in 1986, all the major commercial banks in Texas, which had been very profitable, now found themselves in trouble. At that time, banks in New England were doing fine. However, when the 1990–1991 recession hit New England hard, New England banks started failing. With nationwide banking, a bank could make loans in both New England and Texas and would thus be less likely to fail, because when loans go sour in one location, they would likely be doing well in the other. Thus nationwide banking is seen as a major step toward creating a banking system that is less prone to banking crises.

Two concerns remain about the effects of bank consolidation—that it may lead to a reduction in lending to small businesses and that banks rushing to expand into new geographic markets may take increased risks leading to bank failures. The jury is still out on these concerns, but most economists see the benefits of bank consolidation and nationwide banking as outweighing the costs.

Separation of the Banking and Other Financial Service Industries

Another important feature of the structure of the banking industry in the United States until recently was the separation of the banking and other financial services industries—such as securities, insurance, and real estate—mandated by the Glass-Steagall Act of 1933. As pointed out earlier in the chapter, Glass-Steagall allowed commercial banks to sell new offerings of government securities but prohibited them from underwriting corporate securities or from engaging in brokerage activities. It also prevented banks from engaging in insurance and real estate activities. In turn, it prevented investment banks and insurance companies from engaging in commercial banking activities and thus protected banks from competition.

Erosion of Glass-Steagall

Despite the Glass-Steagall prohibitions, the pursuit of profits and financial innovation stimulated both banks and other financial institutions to bypass the intent of the Glass-Steagall Act and encroach on each other's traditional territory. Brokerage firms engaged in the traditional banking business of issuing deposit instruments with the development of money market mutual funds and cash management accounts. After the Federal Reserve used a loophole in Section 20 of the Glass-Steagall Act in 1987 to

allow bank holding companies to underwrite previously prohibited classes of securities, banks began to enter this business. The loophole allowed affiliates of approved commercial banks to engage in underwriting activities as long as the revenue didn't exceed a specified amount, which started at 10% but was raised to 25% of the affiliates' total revenue. After the U.S. Supreme Court validated the Fed's action in July 1988, the Federal Reserve allowed J.P. Morgan, a commercial bank holding company, to underwrite corporate debt securities (in January 1989) and to underwrite stocks (in September 1990), with the privilege extended to other bank holding companies. The regulatory agencies later allowed banks to engage in some real estate and some insurance activities.

The Gramm-Leach-Bliley Financial Services Modernization Act of 1999: Repeal of Glass-Steagall

Because restrictions on commercial banks' securities and insurance activities put American banks at a competitive disadvantage relative to foreign banks, bills to overturn Glass-Steagall appeared in almost every session of Congress in the 1990s. With the merger in 1998 of Citicorp, the second-largest bank in the United States, and Travelers Group, an insurance company that also owned the third-largest securities firm in the country (Salomon Smith Barney), the pressure to abolish Glass-Steagall became overwhelming. Legislation to eliminate Glass-Steagall finally came to fruition in 1999. This legislation, the Gramm-Leach-Bliley Financial Services Modernization Act of 1999, allows securities firms and insurance companies to purchase banks, and allows banks to underwrite insurance and securities and engage in real estate activities. Under this legislation, states retain regulatory authority over insurance activities, while the Securities and Exchange Commission continues to have oversight of securities activities. The Office of the Comptroller of the Currency has the authority to regulate bank subsidiaries engaged in securities underwriting, but the Federal Reserve continues to have the authority to oversee bank holding companies under which all real estate and insurance activities and large securities operations will be housed.

Implications for Financial Consolidation

As we have seen, the Riegle-Neal Interstate Banking and Branching Efficiency Act of 1994 has stimulated consolidation of the banking industry. The financial consolidation process will be further hastened by the Gramm-Leach-Bliley Act of 1999, because the way is now open to consolidation in terms not only of the number of banking institutions, but also across financial service activities. Given that information technology is increasing economies of scope, mergers of banks with other financial service firms like that of Citicorp and Travelers should become increasingly common, and more mega-mergers are likely to be on the way. Banking institutions are becoming not only larger, but also increasingly complex, organizations, engaging in the full gamut of financial service activities.

Separation of Banking and Other Financial Services Industries Throughout the World

Not many other countries in the aftermath of the Great Depression followed the lead of the United States in separating the banking and other financial services industries. In fact, in the past this separation was the most prominent difference between banking regulation in the United States and in other countries. Around the world, there are three basic frameworks for the banking and securities industries.

The first framework is *universal banking*, which exists in Germany, the Netherlands, and Switzerland. It provides no separation at all between the banking and securities industries. In a universal banking system, commercial banks provide a full range of banking, securities, real estate, and insurance services, all within a single legal

entity. Banks are allowed to own sizable equity shares in commercial firms, and often they do.

The *British-style universal banking system*, the second framework, is found in the United Kingdom and countries with close ties to it, such as Canada and Australia, and now the United States. The British-style universal bank engages in securities underwriting, but it differs from the German-style universal bank in three ways: Separate legal subsidiaries are more common, bank equity holdings of commercial firms are less common, and combinations of banking and insurance firms are less common.

The third framework features some legal separation of the banking and other financial services industries, as in Japan. A major difference between the U.S. and Japanese banking systems is that Japanese banks are allowed to hold substantial equity stakes in commercial firms, whereas American banks cannot. In addition, most American banks use a bank-holding-company structure, but bank holding companies are illegal in Japan. Although the banking and securities industries are legally separated in Japan under Section 65 of the Japanese Securities Act, commercial banks are increasingly being allowed to engage in securities activities and like U.S. banks are thus becoming more like British-style universal banks.

Thrift Industry: Regulation and Structure

Not surprisingly, the regulation and structure of the thrift industry (savings and loan associations, mutual savings banks, and credit unions) closely parallels the regulation and structure of the commercial banking industry.

Savings and Loan Associations

Just as there is a dual banking system for commercial banks, savings and loan associations (S&Ls) can be chartered either by the federal government or by the states. Most S&Ls, whether state or federally chartered, are members of the Federal Home Loan Bank System (FHLBS). Established in 1932, the FHLBS was styled after the Federal Reserve System. It has 12 district Federal Home Loan banks, which are supervised by the Office of Thrift Supervision.

Federal deposit insurance (up to $100,000 per account) for S&Ls is provided by the Savings Association Insurance Fund, a subsidiary of the FDIC. The Office of Thrift Supervision regulates federally insured S&Ls by setting minimum capital requirements, requiring periodic reports, and examining the S&Ls. It is also the chartering agency for federally chartered S&Ls, and for these S&Ls it approves mergers and sets the rules for branching.

The branching regulations for S&Ls were more liberal than for commercial banks: In the past, almost all states permitted branching of S&Ls, and since 1980, federally chartered S&Ls were allowed to branch statewide in all states. Since 1981, mergers of financially troubled S&Ls were allowed across state lines, and nationwide branching of S&Ls is now a reality.

The FHLBS, like the Fed, makes loans to the members of the system (obtaining funds for this purpose by issuing bonds). However, in contrast to the Fed's discount loans, which are expected to be repaid quickly, the loans from the FHLBS often need not be repaid for long periods of time. In addition, the rates charged to S&Ls for these loans are often below the rates that the S&Ls must pay when they borrow in the open market. In this way, the FHLBS loan program provides a subsidy to the savings and

loan industry (and implicitly to the housing industry, since most of the S&L loans are for residential mortgages).

As we will see in the next chapter, the savings and loans experienced serious difficulties in the 1980s. Because savings and loans now engage in many of the same activities as commercial banks, many experts view having a separate charter and regulatory apparatus for S&Ls an anachronism that no longer makes sense.

Mutual Savings Banks

Of the 400 or so mutual savings banks, approximately half are chartered by states. Although the mutual savings banks are primarily regulated by the states in which they are located, the majority have their deposits insured by the FDIC up to the limit of $100,000 per account; these banks are also subject to many of the FDIC's regulations for state-chartered banks. As a rule, the mutual savings banks whose deposits are not insured by the FDIC have their deposits insured by state insurance funds.

The branching regulations for mutual savings banks are determined by the states in which they operate. Because these regulations are not too restrictive, there are few mutual savings banks with assets of less than $25 million.

Credit Unions

Credit unions are small cooperative lending institutions organized around a particular group of individuals with a common bond (union members or employees of a particular firm). They are the only financial institutions that are tax-exempt and can be chartered either by the states or by the federal government; over half are federally chartered. The National Credit Union Administration (NCUA) issues federal charters and regulates federally chartered credit unions by setting minimum capital requirements, requiring periodic reports, and examining the credit unions. Federal deposit insurance (up to the $100,000-per-account limit) is provided to both federally-chartered and state-chartered credit unions by a subsidiary of the NCUA, the National Credit Union Share Insurance Fund (NCUSIF). Since the majority of credit union lending is for consumer loans with fairly short terms to maturity, they did not suffer the financial difficulties of the S&Ls and mutual savings banks.

Because their members share a common bond, credit unions are typically quite small; most hold less than $10 million of assets. In addition, their ties to a particular industry or company make them more likely to fail when large numbers of workers in that industry or company are laid off and have trouble making loan payments. Recent regulatory changes allow individual credit unions to cater to a more diverse group of people by interpreting the common bond requirement less strictly, and this has encouraged an expansion in the size of credit unions that may help reduce credit union failures in the future.

Often a credit union's shareholders are dispersed over many states, and sometimes even worldwide, so branching across state lines and into other countries is permitted for federally chartered credit unions. The Navy Federal Credit Union, for example, whose shareholders are members of the U.S. Navy and Marine Corps, has branches throughout the world.

International Banking

In 1960, only eight U.S. banks operated branches in foreign countries, and their total assets were less than $4 billion. Currently, around 100 American banks have branches

abroad, with assets totaling over $500 billion. The spectacular growth in international banking can be explained by three factors.

First is the rapid growth in international trade and multinational (worldwide) corporations that has occurred since 1960. When American firms operate abroad, they need banking services in foreign countries to help finance international trade. For example, they might need a loan in a foreign currency to operate a factory abroad. And when they sell goods abroad, they need to have a bank exchange the foreign currency they have received for their goods into dollars. Although these firms could use foreign banks to provide them with these international banking services, many of them prefer to do business with the U.S. banks with which they have established long-term relationships and which understand American business customs and practices. As international trade has grown, international banking has grown with it.

Second, American banks have been able to earn substantial profits by being very active in global investment banking, in which they underwrite foreign securities. They also sell insurance abroad, and they derive substantial profits from these investment banking and insurance activities.

Third, American banks have wanted to tap into the large pool of dollar-denominated deposits in foreign countries known as Eurodollars. To understand the structure of U.S. banking overseas, let us first look at the Eurodollar market, an important source of growth for international banking.

Eurodollar Market

Eurodollars are created when deposits in accounts in the United States are transferred to a bank outside the country and are kept in the form of dollars. (For a discussion of the birth of the Eurodollar, see Box 4.) For example, if Rolls-Royce PLC deposits a $1 million check, written on an account at an American bank, in its bank in London—specifying that the deposit is payable in dollars—$1 million in Eurodollars is created.[3] Over 90% of Eurodollar deposits are time deposits, more than half of them certificates of deposit with maturities of 30 days or more. The total amount of Eurodollars outstanding is on the order of $4.4 trillion, making the Eurodollar market one of the most important financial markets in the world economy.

Why would companies like Rolls-Royce want to hold dollar deposits outside the United States? First, the dollar is the most widely used currency in international trade, so Rolls-Royce might want to hold deposits in dollars to conduct its international transactions. Second, Eurodollars are "offshore" deposits—they are held in countries that will not subject them to regulations such as reserve requirements or restrictions (called *capital controls*) on taking the deposits outside the country.[4]

The main center of the Eurodollar market is London, a major international financial center for hundreds of years. Eurodollars are also held outside Europe in locations that provide offshore status to these deposits—for example, Singapore, the Bahamas, and the Cayman Islands.

---

[3]Note that the London bank keeps the $1 million on deposit at the American bank, so the creation of Eurodollars has not caused a reduction in the amount of bank deposits in the United States.

[4]Although most offshore deposits are denominated in dollars, some are also denominated in other currencies. Collectively, these offshore deposits are referred to as Eurocurrencies. A Japanese yen-denominated deposit held in London, for example, is called a Euroyen.

Box 4: Global

Ironic Birth of the Eurodollar Market

One of capitalism's great ironies is that the Eurodollar market, one of the most important financial markets used by capitalists, was fathered by the Soviet Union. In the early 1950s, during the height of the Cold War, the Soviets had accumulated a substantial amount of dollar balances held by banks in the United States. Because the Russians feared that the U.S. government might freeze these assets in the United States, they wanted to move the deposits to Europe, where they would be safe from expropriation. (This fear was not unjustified—consider the U.S. freeze on Iranian assets in 1979 and Iraqi assets in 1990.) However, they also wanted to keep the deposits in dollars so that they could be used in their international transactions. The solution to the problem was to transfer the deposits to European banks but to keep the deposits denominated in dollars. When the Soviets did this, the Eurodollar was born.

The minimum transaction in the Eurodollar market is typically $1 million, and approximately 75% of Eurodollar deposits are held by banks. Plainly, you and I are unlikely to come into direct contact with Eurodollars. The Eurodollar market is, however, an important source of funds to U.S. banks, whose borrowing of these deposits is over $100 billion. Rather than using an intermediary and borrowing all the deposits from foreign banks, American banks decided that they could earn higher profits by opening their own branches abroad to attract these deposits. Consequently, the Eurodollar market has been an important stimulus to U.S. banking overseas.

Structure of U.S. Banking Overseas

U.S. banks have most of their foreign branches in Latin America, the Far East, the Caribbean, and London. The largest volume of assets is held by branches in London, because it is a major international financial center and the central location for the Eurodollar market. Latin America and the Far East have many branches because of the importance of U.S. trade with these regions. Parts of the Caribbean (especially the Bahamas and the Cayman Islands) have become important as tax havens, with minimal taxation and few restrictive regulations. In actuality, the bank branches in the Bahamas and the Cayman Islands are "shell operations" because they function primarily as bookkeeping centers and do not provide normal banking services.

An alternative corporate structure for U.S. banks that operate overseas is the **Edge Act corporation**, a special subsidiary engaged primarily in international banking. U.S. banks (through their holding companies) can also own a controlling interest in foreign banks and in foreign companies that provide financial services, such as finance companies. The international activities of U.S. banking organizations are governed primarily by the Federal Reserve's Regulation K.

In late 1981, the Federal Reserve approved the creation of **international banking facilities (IBFs)** within the United States that can accept time deposits from foreigners but are not subject to either reserve requirements or restrictions on interest payments. IBFs are also allowed to make loans to foreigners, but they are not allowed to make loans to domestic residents. States have encouraged the establishment of IBFs by exempting them from state and local taxes. In essence, IBFs are treated like foreign

branches of U.S. banks and are not subject to domestic regulations and taxes. The purpose of establishing IBFs is to encourage American and foreign banks to do more banking business in the United States rather than abroad. From this point of view, IBFs were a success: Their assets climbed to nearly $200 billion in the first two years, but have currently fallen to below $100 billion.

Foreign Banks in the United States

The growth in international trade has not only encouraged U.S. banks to open offices overseas, but has also encouraged foreign banks to establish offices in the United States. Foreign banks have been extremely successful in the United States. Currently, they hold more than 10% of total U.S. bank assets and do a large portion of all U.S. bank lending, with nearly a 19% market share for lending to U.S. corporations.

Foreign banks engage in banking activities in the United States by operating an agency office of the foreign bank, a subsidiary U.S. bank, or a branch of the foreign bank. An agency office can lend and transfer funds in the United States, but it cannot accept deposits from domestic residents. Agency offices have the advantage of not being subject to regulations that apply to full-service banking offices (such as requirements for FDIC insurance). A subsidiary U.S. bank is just like any other U.S. bank (it may even have an American-sounding name) and is subject to the same regulations, but it is owned by the foreign bank. A branch of a foreign bank bears the foreign bank's name and is usually a full-service office. Foreign banks may also form Edge Act corporations and IBFs.

Before 1978, foreign banks were not subject to many regulations that applied to domestic banks: They could open branches across state lines and were not expected to meet reserve requirements, for example. The passage of the International Banking Act of 1978, however, put foreign and domestic banks on a more equal footing. Now foreign banks may open new full-service branches only in the state they designate as their home state or in states that allow the entry of out-of-state banks. Limited-service branches and agency offices in any other state are permitted, however, and foreign banks are allowed to retain any full-service branches opened before ratification of the International Banking Act of 1978.

The internationalization of banking, both by U.S. banks going abroad and by foreign banks entering the United States, has meant that financial markets throughout the world have become more integrated. As a result, there is a growing trend toward international coordination of bank regulation, one example of which is the 1988 Basel agreement to standardize minimum capital requirements in industrialized countries, discussed in Chapter 11. Financial market integration has also encouraged bank consolidation abroad, culminating in the creation of the first trillion-dollar bank with the proposed merger of the Industrial Bank of Japan, Dai-Ichi Kangyo Bank, and Fuji Bank, announced in August 1999, but which took place in 2002. Another development has been the importance of foreign banks in international banking. As is shown in Table 3, in 2002, eight of the ten largest banks in the world were foreign. The implications of this financial market integration for the operation of our economy is examined further in Chapter 20 when we discuss the international financial system in more detail.

Table 3 Ten Largest Banks in the World, 2002

Bank	Assets (U.S. $ millions)
1. Mizuho Holdings, Japan	1,281,389
2. Citigroup, U.S.	1,057,657
3. Mitsubishi Tokyo Financial Group, Japan	854,749
4. Deutsche Bank, Germany	815,126
5. Allianz, Germany	805,433
6. UBS, Switzerland	753,833
7. BNP, France	734,833
8. HSBC Holdings, U.K.	694,590
9. J.P. Morgan & Chase Company, U.S.	712,508
10. Bayerische Hypo-Und Vereinsbanken, Germany	638,544

Source: American Banker, 167 (132): 17, July 12, 2002.

Summary

1. The history of banking in the United States has left us with a dual banking system, with commercial banks chartered by the states and the federal government. Multiple agencies regulate commercial banks: the Office of the Comptroller, the Federal Reserve, the FDIC, and the state banking authorities.
2. A change in the economic environment will stimulate financial institutions to search for financial innovations. Changes in demand conditions, especially the rise in interest-rate risk; changes in supply conditions, especially improvements in information technology; and the desire to avoid costly regulations have been major driving forces behind financial innovation. Financial innovation has caused banks to suffer declines in cost advantages in acquiring funds and in income advantages on their assets. The resulting squeeze has hurt profitability in banks' traditional lines of business and has led to a decline in traditional banking.
3. Restrictive state branching regulations and the McFadden Act, which prohibited branching across state lines, led to a large number of small commercial banks. The large number of commercial banks in the United States reflected the past *lack* of competition, not the presence of vigorous competition. Bank holding companies and ATMs were important responses to branching restrictions that weakened the restrictions' anticompetitive effect.
4. Since the mid-1980s, bank consolidation has been occurring at a rapid pace. The first phase of bank consolidation was the result of bank failures and the reduced effectiveness of branching restrictions. The second phase has been stimulated by information technology and the Riegle-Neal Interstate Banking and Branching Efficiency Act of 1994, which establishes the basis for a nationwide banking system. Once banking consolidation has settled down, we are likely to be left with a banking system with several thousand banks. Most economists believe that the benefits of bank consolidation and nationwide banking will outweigh the costs.
5. The Glass-Steagall Act separated commercial banking from the securities industry. Legislation in 1999,

however, repealed the Glass-Steagall Act, removing the separation of these industries.

6. The regulation and structure of the thrift industry (savings and loan associations, mutual savings banks, and credit unions) parallel closely the regulation and structure of the commercial banking industry. Savings and loans are primarily regulated by the Office of Thrift Supervision, and deposit insurance is administered by the FDIC. Mutual savings banks are regulated by the states, and federal deposit insurance is provided by the FDIC. Credit unions are regulated by the National Credit Union Administration, and deposit insurance is provided by the National Credit Union Share Insurance Fund.
7. With the rapid growth of world trade since 1960, international banking has grown dramatically. United States banks engage in international banking activities by opening branches abroad, owning controlling interests in foreign banks, forming Edge Act corporations, and operating international banking facilities (IBFs) located in the United States. Foreign banks operate in the United States by owning a subsidiary American bank or by operating branches or agency offices in the United States.

Key Terms

automated banking machine (ABM), p. 235
automated teller machine (ATM), p. 235
bank holding companies, p. 232
branches, p. 244
central bank, p. 230
deposit rate ceilings, p. 238
disintermediation, p. 238
dual banking system, p. 231
economies of scope, p. 248
Edge Act corporation, p. 255
financial derivatives, p. 233
financial engineering, p. 232
futures contracts, p. 233
hedge, p. 233
international banking facilities (IBFs), p. 255
large, complex, banking organizations (LCBOs), p. 248
national banks, p. 231
securitization, p. 237
state banks, p. 231
superregional banks, p. 247
sweep account, p. 239
virtual bank, p. 235

Questions and Problems

Questions marked with an asterisk are answered at the end of the book in an appendix, "Answers to Selected Questions and Problems."

1. Why was the United States one of the last of the major industrialized countries to have a central bank?
*2. Which regulatory agency has the primary responsibility for supervising the following categories of commercial banks?
   a. National banks
   b. Bank holding companies
   c. Non–Federal Reserve state banks
   d. Federal Reserve member state banks
3. "The commercial banking industry in Canada is less competitive than the commercial banking industry in the United States because in Canada only a few large banks dominate the industry, while in the United States there are around 8,000 commercial banks." Is this statement true, false, or uncertain? Explain your answer.
*4. Why did new technology make it harder to enforce limitations on bank branching?
5. Why has there been such a dramatic increase in bank holding companies?
*6. Why is there a higher percentage of banks with under $25 million of assets among commercial banks than among savings and loans and mutual savings banks?

7. Unlike commercial banks, savings and loans, and mutual savings banks, credit unions did not have restrictions on locating branches in other states. Why, then, are credit unions typically smaller than the other depository institutions?

*8. What incentives have regulatory agencies created to encourage international banking? Why have they done this?

9. How could the approval of international banking facilities (IBFs) by the Fed in 1981 have reduced employment in the banking industry in Europe?

*10. If the bank at which you keep your checking account is owned by Saudi Arabians, should you worry that your deposits are less safe than if the bank were owned by Americans?

11. If reserve requirements were eliminated in the future, as some economists advocate, what effects would this have on the size of money market mutual funds?

*12. Why have banks been losing cost advantages in acquiring funds in recent years?

13. "If inflation had not risen in the 1960s and 1970s, the banking industry might be healthier today." Is this statement true, false, or uncertain? Explain your answer.

*14. Why have banks been losing income advantages on their assets in recent years?

15. "The invention of the computer is the major factor behind the decline of the banking industry." Is this statement true, false, or uncertain? Explain your answer.

Web Exercises

1. Go to www.fdic.gov/bank/statistical/statistics/index.html. Select "Highlights and Trends." Choose "Number of FDIC-Insured Commercial Banks and Trust Companies." Looking at the trend in bank branches, does the public appear to have more or less access to banking facilities? How many banks were there in 1934 and how many are there now? Does the graph indicate that the trend toward consolidation is continuing?

2. Despite the regulations that protect banks from failure, some do fail. Go to www2.fdic.gov/hsob/. Select the tab labeled "Bank and Thrift Failures." How many bank failures occurred in the U.S. during the most recent complete calendar year? What were the total assets held by the banks that failed? How many banks failed in 1937?

Chapter 11

Economic Analysis of Banking Regulation

PREVIEW

As we have seen in the previous chapters, the financial system is among the most heavily regulated sectors of the economy, and banks are among the most heavily regulated of financial institutions. In this chapter, we develop an economic analysis of why regulation of banking takes the form it does.

Unfortunately, the regulatory process may not always work very well, as evidenced by recent crises in the banking systems, not only in the United States but in many countries throughout the world. Here we also use our economic analysis of banking regulation to explain the worldwide crises in banking and how the regulatory system can be reformed to prevent future disasters.

Asymmetric Information and Banking Regulation

In earlier chapters, we have seen how asymmetric information, the fact that different parties in a financial contract do not have the same information, leads to adverse selection and moral hazard problems that have an important impact on our financial system. The concepts of asymmetric information, adverse selection, and moral hazard are especially useful in understanding why government has chosen the form of banking regulation we see in the United States and in other countries. There are eight basic categories of banking regulation: the government safety net, restrictions on bank asset holdings, capital requirements, chartering and bank examination, assessment of risk management, disclosure requirements, consumer protection, and restrictions on competition.

Government Safety Net: Deposit Insurance and the FDIC

www.ny.frb.org/Pihome/regs.html
View bank regulation information.

As we saw in Chapter 8, banks are particularly well suited to solving adverse selection and moral hazard problems because they make private loans that help avoid the free-rider problem. However, this solution to the free-rider problem creates another asymmetric information problem, because depositors lack information about the quality of these private loans. This asymmetric information problem leads to two reasons why the banking system might not function well.

First, before the FDIC started operations in 1934, a **bank failure** (in which a bank is unable to meet its obligations to pay its depositors and other creditors and so must go out of business) meant that depositors would have to wait to get their deposit funds until the bank was liquidated (until its assets had been turned into cash); at that time, they would be paid only a fraction of the value of their deposits. Unable to learn

if bank managers were taking on too much risk or were outright crooks, depositors would be reluctant to put money in the bank, thus making banking institutions less viable. Second is that depositors' lack of information about the quality of bank assets can lead to bank panics, which, as we saw in Chapter 8, can have serious harmful consequences for the economy. To see this, consider the following situation. There is no deposit insurance, and an adverse shock hits the economy. As a result of the shock, 5% of the banks have such large losses on loans that they become insolvent (have a negative net worth and so are bankrupt). Because of asymmetric information, depositors are unable to tell whether their bank is a good bank or one of the 5% that are insolvent. Depositors at bad *and* good banks recognize that they may not get back 100 cents on the dollar for their deposits and will want to withdraw them. Indeed, because banks operate on a "sequential service constraint" (a first-come, first-served basis), depositors have a very strong incentive to show up at the bank first, because if they are last in line, the bank may run out of funds and they will get nothing. Uncertainty about the health of the banking system in general can lead to runs on banks both good and bad, and the failure of one bank can hasten the failure of others (referred to as the *contagion effect*). If nothing is done to restore the public's confidence, a bank panic can ensue.

Indeed, bank panics were a fact of American life in the nineteenth and early twentieth centuries, with major ones occurring every 20 years or so in 1819, 1837, 1857, 1873, 1884, 1893, 1907, and 1930–1933. Bank failures were a serious problem even during the boom years of the 1920s, when the number of bank failures averaged around 600 per year.

A government safety net for depositors can short-circuit runs on banks and bank panics, and by providing protection for the depositor, it can overcome reluctance to put funds in the banking system. One form of the safety net is deposit insurance, a guarantee such as that provided by the Federal Deposit Insurance Corporation (FDIC) in the United States in which depositors are paid off in full on the first $100,000 they have deposited in the bank no matter what happens to the bank. With fully insured deposits, depositors don't need to run to the bank to make withdrawals—even if they are worried about the bank's health—because their deposits will be worth 100 cents on the dollar no matter what. From 1930 to 1933, the years immediately preceding the creation of the FDIC, the number of bank failures averaged over 2,000 per year. After the establishment of the FDIC in 1934, bank failures averaged fewer than 15 per year until 1981.

The FDIC uses two primary methods to handle a failed bank. In the first, called the *payoff method*, the FDIC allows the bank to fail and pays off deposits up to the $100,000 insurance limit (with funds acquired from the insurance premiums paid by the banks who have bought FDIC insurance). After the bank has been liquidated, the FDIC lines up with other creditors of the bank and is paid its share of the proceeds from the liquidated assets. Typically, when the payoff method is used, account holders with deposits in excess of the $100,000 limit get back more than 90 cents on the dollar, although the process can take several years to complete.

In the second method, called the *purchase and assumption method*, the FDIC reorganizes the bank, typically by finding a willing merger partner who assumes (takes over) all of the failed bank's deposits so that no depositor loses a penny. The FDIC may help the merger partner by providing it with subsidized loans or by buying some of the failed bank's weaker loans. The net effect of the purchase and assumption method is that the FDIC has guaranteed *all* deposits, not just those under the

$100,000 limit. The purchase and assumption method was the FDIC's most common procedure for dealing with a failed bank before new banking legislation in 1991.

Deposit insurance is not the only way in which governments provide a safety net for depositors. In other countries, governments have often stood ready to provide support to domestic banks when they face runs even in the absence of explicit deposit insurance. This support is sometimes provided by lending from the central bank to troubled institutions and is often referred to as the "lender of last resort" role of the central bank. In other cases, funds are provided directly by the government to troubled institutions, or these institutions are taken over by the government and the government then guarantees that depositors will receive their money in full. However, in recent years, government deposit insurance has been growing in popularity and has spread to many countries throughtout the world. Whether this trend is desirable is discussed in Box 1.

Moral Hazard and the Government Safety Net. Although a government safety net has been successful at protecting depositors and preventing bank panics, it is a mixed blessing. The most serious drawback of the government safety net stems from moral hazard, the incentives of one party to a transaction to engage in activities detrimental to the other party. Moral hazard is an important concern in insurance arrangements in general because the existence of insurance provides increased incentives for taking

Box 1: Global

The Spread of Government Deposit Insurance Throughout the World: Is This a Good Thing?

For the first 30 years after federal deposit insurance was established in the United States, only 6 countries emulated the United States and adopted deposit insurance. However, this began to change in the late 1960s, with the trend accelerating in the 1990s, when the number of countries adopting deposit insurance doubled to over 70. Government deposit insurance has taken off throughout the world because of growing concern about the health of banking systems, particularly after the increasing number of banking crises in recent years (documented at the end of this chapter). Has this spread of deposit insurance been a good thing? Has it helped improve the performance of the financial system and prevent banking crises?

The answer seems to be no under many circumstances. Research at the World Bank has found that on average, the adoption of explicit government deposit insurance is associated with less banking sector stability and a higher incidence of banking crises.* Furthermore, on average it seems to retard financial development. However, the negative effects of deposit insurance appear only in countries with weak institutional environments: an absence of rule of law, ineffective regulation and supervision of the financial sector, and high corruption. This is exactly what might be expected because, as we will see later in this chapter, a strong institutional environment is needed to limit the incentives for banks to engage in the excessively risky behavior encouraged by deposit insurance. The problem is that developing a strong institutional environment may be very difficult to achieve in many emerging market countries. This leaves us with the following conclusion: Adoption of deposit insurance may be exactly the wrong medicine for promoting stability and efficiency of banking systems in emerging market countries.

*See World Bank, *Finance for Growth: Policy Choices in a Volatile World* (Oxford: World Bank and Oxford University Press, 2001).

risks that might result in an insurance payoff. For example, some drivers with automobile collision insurance that has a low deductible might be more likely to drive recklessly, because if they get into an accident, the insurance company pays most of the costs for damage and repairs.

Moral hazard is a prominent concern in government arrangements to provide a safety net. Because with a safety net depositors know that they will not suffer losses if a bank fails, they do not impose the discipline of the marketplace on banks by withdrawing deposits when they suspect that the bank is taking on too much risk. Consequently, banks with a government safety net have an incentive to take on greater risks than they otherwise would.

Adverse Selection and the Government Safety Net. A further problem with a government safety net like deposit insurance arises because of adverse selection, the fact that the people who are most likely to produce the adverse outcome insured against (bank failure) are those who most want to take advantage of the insurance. For example, bad drivers are more likely than good drivers to take out automobile collision insurance with a low deductible. Because depositors protected by a government safety net have little reason to impose discipline on the bank, risk-loving entrepreneurs might find the banking industry a particularly attractive one to enter—they know that they will be able to engage in highly risky activities. Even worse, because protected depositors have so little reason to monitor the bank's activities, without government intervention outright crooks might also find banking an attractive industry for their activities because it is easy for them to get away with fraud and embezzlement.

"Too Big to Fail." The moral hazard created by a government safety net and the desire to prevent bank failures have presented bank regulators with a particular quandary. Because the failure of a very large bank makes it more likely that a major financial disruption will occur, bank regulators are naturally reluctant to allow a big bank to fail and cause losses to its depositors. Indeed, consider Continental Illinois, one of the ten largest banks in the United States when it became insolvent in May 1984. Not only did the FDIC guarantee depositors up to the $100,000 insurance limit, but it also guaranteed accounts exceeding $100,000 and even prevented losses for Continental Illinois bondholders. Shortly thereafter, the Comptroller of the Currency (the regulator of national banks) testified to Congress that the FDIC's policy was to regard the 11 largest banks as "too big to fail"—in other words, the FDIC would bail them out so that no depositor or creditor would suffer a loss. The FDIC would do this by using the purchase and assumption method, giving the insolvent bank a large infusion of capital and then finding a willing merger partner to take over the bank and its deposits. The too-big-to-fail policy was extended to big banks that were not even among the 11 largest. (Note that "too big to fail" is somewhat misleading because when a bank is closed or merged into another bank, the managers are usually fired and the stockholders in the bank lose their investment.)

One problem with the too-big-to-fail policy is that it increases the moral hazard incentives for big banks. If the FDIC were willing to close a bank using the alternative payoff method, paying depositors only up to the $100,000 limit, large depositors with more than $100,000 would suffer losses if the bank failed. Thus they would have an incentive to monitor the bank by examining the bank's activities closely and pulling their money out if the bank was taking on too much risk. To prevent such a loss of deposits, the bank would be more likely to engage in less risky activities. However, once large depositors know that a bank is too big to fail, they have no incentive to

monitor the bank and pull out their deposits when it takes on too much risk: No matter what the bank does, large depositors will not suffer any losses. The result of the too-big-to-fail policy is that big banks might take on even greater risks, thereby making bank failures more likely.[1]

Financial Consolidation and the Government Safety Net. With financial innovation and the passage of the Riegle-Neal Interstate Banking and Branching and Efficiency Act of 1994 and the Gramm-Leach-Bliley Financial Services Modernization Act in 1999, financial consolidation has been proceeding at a rapid pace, leading to both larger and more complex banking organizations. Financial consolidation poses two challenges to banking regulation because of the existence of the government safety net. First, the increased size of banks as a result of financial consolidation increases the too-big-to-fail problem, because there will now be more large institutions whose failure exposes the financial system to systemic (system-wide) risk. Thus more banking institutions are likely to be treated as too big to fail, and the increased moral hazard incentives for these large institutions to take on greater risk can then increase the fragility of the financial system. Second, financial consolidation of banks with other financial services firms means that the government safety net may be extended to new activities such as securities underwriting, insurance, or real estate activities, thereby increasing incentives for greater risk taking in these activities that can also weaken the fabric of the financial system. Limiting the moral hazard incentives for the larger, more complex financial organizations that are resulting from recent changes in legislation will be one of the key issues facing banking regulators in the future.

Restrictions on Asset Holdings and Bank Capital Requirements

As we have seen, the moral hazard associated with a government safety net encourages too much risk taking on the part of banks. Bank regulations that restrict asset holdings and bank capital requirements are directed at minimizing this moral hazard, which can cost the taxpayers dearly.

Even in the absence of a government safety net, banks still have the incentive to take on too much risk. Risky assets may provide the bank with higher earnings when they pay off; but if they do not pay off and the bank fails, depositors are left holding the bag. If depositors were able to monitor the bank easily by acquiring information on its risk-taking activities, they would immediately withdraw their deposits if the bank was taking on too much risk. To prevent such a loss of deposits, the bank would be more likely to reduce its risk-taking activities. Unfortunately, acquiring information on a bank's activities to learn how much risk the bank is taking can be a difficult task. Hence most depositors are incapable of imposing discipline that might prevent banks from engaging in risky activities. A strong rationale for government regulation to reduce risk taking on the part of banks therefore existed even before the establishment of federal deposit insurance.

Bank regulations that restrict banks from holding risky assets such as common stock are a direct means of making banks avoid too much risk. Bank regulations also promote diversification, which reduces risk by limiting the amount of loans in particular categories or to individual borrowers. Requirements that banks have sufficient

[1]Evidence reveals, as our analysis predicts, that large banks took on riskier loans than smaller banks and that this led to higher loan losses for big banks; see John Boyd and Mark Gertler, "U.S. Commercial Banking: Trends, Cycles and Policy," *NBER Macroeconomics Annual*, 1993, pp. 319–368.

bank capital are another way to change the bank's incentives to take on less risk. When a bank is forced to hold a large amount of equity capital, the bank has more to lose if it fails and is thus more likely to pursue less risky activities.

Bank capital requirements take two forms. The first type is based on the so-called **leverage ratio**, the amount of capital divided by the bank's total assets. To be classified as well capitalized, a bank's leverage ratio must exceed 5%; a lower leverage ratio, especially one below 3%, triggers increased regulatory restrictions on the bank. Through most of the 1980s, minimum bank capital in the United States was set solely by specifying a minimum leverage ratio.

In the wake of the Continental Illinois and savings and loans bailouts, regulators in the United States and the rest of the world have become increasingly worried about banks' holdings of risky assets and about the increase in banks' **off-balance-sheet activities**, activities that involve trading financial instruments and generating income from fees, which do not appear on bank balance sheets but nevertheless expose banks to risk. An agreement among banking officials from industrialized nations set up the **Basel Committee on Banking Supervision** (because it meets under the auspices of the Bank for International Settlements in Basel, Switzerland), which has implemented the so-called **Basel Accord** on a second type of capital requirements, risk-based capital requirements. The Basel Accord, which required that banks hold as capital at least 8% of their risk-weighted assets, has been adopted by more than 100 countries, including the United States. Assets and off-balance-sheet activities were allocated into four categories, each with a different weight to reflect the degree of credit risk. The first category carries a zero weight and includes items that have little default risk, such as reserves and government securities in the OECD, Organization for Economic Cooperation and Development, (industrialized) countries. The second category has a 20% weight and includes claims on banks in OECD countries. The third category has a weight of 50% and includes municipal bonds and residential mortgages. The fourth category has the maximum weight of 100% and includes loans to consumers and corporations. Off-balance-sheet activities are treated in a similar manner by assigning a credit-equivalent percentage that converts them to on-balance-sheet items to which the appropriate risk weight applies. The 1996 Market Risk Amendment to the Accord set minimum capital requirements for risks in banks' trading accounts.

Over time, limitations of the Accord have become apparent, because the regulatory measure of bank risk as stipulated by the risk weights can differ substantially from the actual risk the bank faces. This has resulted in what is known as **regulatory arbitrage**, in which banks keep on their books assets that have the same risk-based capital requirement but are relatively risky, such as a loan to a company with a very low credit rating, while taking off their books low-risk assets, such as a loan to a company with a very high credit rating. The Basel Accord could thus lead to increased risk taking, the opposite of its intent. To address these limitations, the Basel Committee on Bank Supervision has released proposals for a new capital accord, often referred to as Basel 2, but it is not clear if it is workable or if it will be implemented (see Box 2).

The Basel Committee's work on bank capital requirements is never-ending. As the banking industry changes, the regulation of bank capital must change with it to ensure the safety and soundness of the banking institutions.

Bank Supervision: Chartering and Examination

Overseeing who operates banks and how they are operated, referred to as **bank supervision** or more generally as **prudential supervision**, is an important method for reducing adverse selection and moral hazard in the banking business. Because

Box 2: Global

Basel 2: Is It Spinning Out of Control?

Starting in June 1999, the Basel Committee on Banking Supervision released several proposals to reform the original 1988 Basel Accord. These efforts have culminated in what bank supervisors refer to as Basel 2, which is based on three pillars. Pillar 1 intends to link capital requirements more closely to actual risk. It does so by specifying many more categories of risk with different weights in its so-called standardized approach. Alternatively, it allows sophisticated banks to pursue instead an internal ratings–based approach that permits banks to use their own models of credit risk. Pillar 2 focuses on strengthening the supervisory process, particularly in assessing the quality of risk management in banking institutions and in evaluating whether these institutions have adequate procedures to determine how much capital they need. Pillar 3 focuses on improving market discipline through increased disclosure of details about the bank's credit exposures, its amount of reserves and capital, the officials who control the bank, and the effectiveness of its internal ratings system.

Although Basel 2 makes great strides toward limiting excessive risk taking by banking institutions, it has come at a cost of greatly increasing the complexity of the Accord. The document describing the original Basel Accord was twenty-six pages, while the second draft of Basel 2 issued in January 2001 exceeds 500 pages. The original timetable called for the completion of the final round of consultation by the end of May 2001, with the new rules taking effect by 2004. However, criticism from banks, trade associations, and national regulators has led to several postponements, with the final draft now scheduled to be published in the last quarter of 2003 and the Accord to be implemented in 2006. Will the increasing complexity of the Basel Accord lead to further postponements? Will Basel 2 eventually be put into operation? As of this writing, these questions remain unanswered.

www.federalreserve.gov/Regulations/default.htm
Access regulatory publications of the Federal Reserve Board.

banks can be used by crooks or overambitious entrepreneurs to engage in highly speculative activities, such undesirable people would be eager to run a bank. Chartering banks is one method for preventing this adverse selection problem; through chartering, proposals for new banks are screened to prevent undesirable people from controlling them.

Regular on-site bank examinations, which allow regulators to monitor whether the bank is complying with capital requirements and restrictions on asset holdings, also function to limit moral hazard. Bank examiners give banks a so-called *CAMELS rating* (the acronym is based on the six areas assessed: capital adequacy, asset quality, management, earnings, liquidity, and sensitivity to market risk). With this information about a bank's activities, regulators can enforce regulations by taking such formal actions as *cease and desist orders* to alter the bank's behavior or even close a bank if its CAMELS rating is sufficiently low. Actions taken to reduce moral hazard by restricting banks from taking on too much risk help reduce the adverse selection problem further, because with less opportunity for risk taking, risk-loving entrepreneurs will be less likely to be attracted to the banking industry. Note that the methods regulators use to cope with adverse selection and moral hazard have their counterparts in private financial markets (see Chapters 8 and 9). Chartering is similar to the screening of potential borrowers, regulations restricting risky asset holdings are similar to restrictive covenants that prevent borrowing firms from engaging in risky investment activities, bank capital requirements act like restrictive covenants that require

minimum amounts of net worth for borrowing firms, and regular bank examinations are similar to the monitoring of borrowers by lending institutions.

A commercial bank obtains a charter either from the Comptroller of the Currency (in the case of a national bank) or from a state banking authority (in the case of a state bank). To obtain a charter, the people planning to organize the bank must submit an application that shows how they plan to operate the bank. In evaluating the application, the regulatory authority looks at whether the bank is likely to be sound by examining the quality of the bank's intended management, the likely earnings of the bank, and the amount of the bank's initial capital. Before 1980, the chartering agency typically explored the issue of whether the community needed a new bank. Often a new bank charter would not be granted if existing banks in a community would be hurt by its presence. Today this anticompetitive stance (justified by the desire to prevent bank failures of existing banks) is no longer as strong in the chartering agencies.

Once a bank has been chartered, it is required to file periodic (usually quarterly) *call reports* that reveal the bank's assets and liabilities, income and dividends, ownership, foreign exchange operations, and other details. The bank is also subject to examination by the bank regulatory agencies to ascertain its financial condition at least once a year. To avoid duplication of effort, the three federal agencies work together and usually accept each other's examinations. This means that, typically, national banks are examined by the Office of the Comptroller of the Currency, the state banks that are members of the Federal Reserve System are examined by the Fed, and insured nonmember state banks are examined by the FDIC.

Bank examinations are conducted by bank examiners, who sometimes make unannounced visits to the bank (so that nothing can be "swept under the rug" in anticipation of their examination). The examiners study a bank's books to see whether it is complying with the rules and regulations that apply to its holdings of assets. If a bank is holding securities or loans that are too risky, the bank examiner can force the bank to get rid of them. If a bank examiner decides that a loan is unlikely to be repaid, the examiner can force the bank to declare the loan worthless (to write off the loan). If, after examining the bank, the examiner feels that it does not have sufficient capital or has engaged in dishonest practices, the bank can be declared a "problem bank" and will be subject to more frequent examinations.

Assessment of Risk Management

Traditionally, on-site bank examinations have focused primarily on assessment of the quality of the bank's balance sheet at a point in time and whether it complies with capital requirements and restrictions on asset holdings. Although the traditional focus is important for reducing excessive risk taking by banks, it is no longer felt to be adequate in today's world, in which financial innovation has produced new markets and instruments that make it easy for banks and their employees to make huge bets easily and quickly. In this new financial environment, a bank that is quite healthy at a particular point in time can be driven into insolvency extremely rapidly from trading losses, as forcefully demonstrated by the failure of Barings in 1995 (discussed in Chapter 9). Thus an examination that focuses only on a bank's position at a point in time may not be effective in indicating whether a bank will in fact be taking on excessive risk in the near future.

This change in the financial environment for banking institutions has resulted in a major shift in thinking about the bank supervisory process throughout the world. Bank examiners are now placing far greater emphasis on evaluating the soundness of a bank's management processes with regard to controlling risk. This shift in thinking

was reflected in a new focus on risk management in the Federal Reserve System's 1993 guidelines to examiners on trading and derivatives activities. The focus was expanded and formalized in the Trading Activities Manual issued early in 1994, which provided bank examiners with tools to evaluate risk management systems. In late 1995, the Federal Reserve and the Comptroller of the Currency announced that they would be assessing risk management processes at the banks they supervise. Now bank examiners give a separate risk management rating from 1 to 5 that feeds into the overall management rating as part of the CAMELS system. Four elements of sound risk management are assessed to come up with the risk management rating: (1) The quality of oversight provided by the board of directors and senior management, (2) the adequacy of policies and limits for all activities that present significant risks, (3) the quality of the risk measurement and monitoring systems, and (4) the adequacy of internal controls to prevent fraud or unauthorized activities on the part of employees.

This shift toward focusing on management processes is also reflected in recent guidelines adopted by the U.S. bank regulatory authorities to deal with interest-rate risk. At one point, U.S. regulators were contemplating requiring banks to use a standard model to calculate the amount of capital a bank would need to have to allow for the interest-rate risk it bears. Because coming up with a one-size-fits-all model that would work for all banks has proved difficult, the regulatory agencies have instead decided to adopt guidelines for the management of interest-rate risk, although bank examiners will continue to consider interest-rate risk in deciding on the bank's capital requirements. These guidelines require the bank's board of directors to establish interest-rate risk limits, appoint officials of the bank to manage this risk, and monitor the bank's risk exposure. The guidelines also require that senior management of a bank develop formal risk management policies and procedures to ensure that the board of director's risk limits are not violated and to implement internal controls to monitor interest-rate risk and compliance with the board's directives.

Disclosure Requirements

The free-rider problem described in Chapter 8 indicates that individual depositors and other bank creditors will not have enough incentive to produce private information about the quality of a bank's assets. To ensure that there is better information for depositors and the marketplace, regulators can require that banks adhere to certain standard accounting principles and disclose a wide range of information that helps the market assess the quality of a bank's portfolio and the amount of the bank's exposure to risk. More public information about the risks incurred by banks and the quality of their portfolio can better enable stockholders, creditors, and depositors to evaluate and monitor banks and so act as a deterrent to excessive risk taking. This view is consistent with a position paper issued by the Eurocurrency Standing Committee of the G-10 Central Banks, which recommends that estimates of financial risk generated by firms' own internal risk management systems be adapted for public disclosure purposes.[2] Such information would supplement disclosures based on tra-

[2]See Eurocurrency Standing Committee of Central Banks of Group of Ten Countries (Fisher Group), "Discussion Paper on Public Disclosure of Markets and Credit Risks by Financial Intermediaries," September 1994, and a companion piece to this report, Federal Reserve Bank of New York, "A Discussion Paper on Public Disclosure of Risks Related to Market Activity," September 1994.

ditional accounting conventions by providing information about risk exposure and risk management that is not normally included in conventional balance sheet and income statement reports.

Consumer Protection

The existence of asymmetric information also suggests that consumers may not have enough information to protect themselves fully. Consumer protection regulation has taken several forms. First is "truth in lending," mandated under the Consumer Protection Act of 1969, which requires all lenders, not just banks, to provide information to consumers about the cost of borrowing including a standardized interest rate (called the annual percentage rate, or APR) and the total finance charges on the loan. The Fair Credit Billing Act of 1974 requires creditors, especially credit card issuers, to provide information on the method of assessing finance charges and requires that billing complaints be handled quickly. Both of these acts are administered by the Federal Reserve System under Regulation Z.

Congress has also passed legislation to reduce discrimination in credit markets. The Equal Credit Opportunity Act of 1974 and its extension in 1976 forbid discrimination by lenders based on race, gender, marital status, age, or national origin. It is administered by the Federal Reserve under Regulation B. The Community Reinvestment Act (CRA) of 1977 was enacted to prevent "redlining," a lender's refusal to lend in a particular area (marked off by a hypothetical red line on a map). The Community Reinvestment Act requires that banks show that they lend in all areas in which they take deposits, and if banks are found to be in noncompliance with the act, regulators can reject their applications for mergers, branching, or other new activities.

Restrictions on Competition

Increased competition can also increase moral hazard incentives for banks to take on more risk. Declining profitability as a result of increased competition could tip the incentives of bankers toward assuming greater risk in an effort to maintain former profit levels. Thus governments in many countries have instituted regulations to protect banks from competition. These regulations have taken two forms in the United States in the past. First were restrictions on branching, such as those described in Chapter 10, which reduced competition between banks. The second form involved preventing nonbank institutions from competing with banks by engaging in banking business, as embodied in the Glass-Steagall Act, which was repealed in 1999.

Although restricting competition propped up the health of banks, restrictions on competition also had serious disadvantages: They led to higher charges to consumers and decreased the efficiency of banking institutions, which did not have to compete as hard. Thus, although the existence of asymmetric information provided a rationale for anticompetitive regulations, it did not mean that they would be beneficial. Indeed, in recent years, the impulse of governments in industrialized countries to restrict competition has been waning. Electronic banking has raised a new set of concerns for regulators to deal with. See Box 3 for a discussion of this challenge.

www.fdic.gov/regulations/laws/important/index.html

Describes the most important laws that have affected banking industry in the U.S.

Study Guide

Because so many laws regulating banking have been passed in the United States, it is hard to keep track of them all. As a study aid, Table 1 lists the major banking legislation in the twentieth century and its key provisions.

Box 3: E-Finance

Electronic Banking: New Challenges for Bank Regulation

The advent of electronic banking has raised new concerns for banking regulation, specifically about security and privacy.

Worries about the security of electronic banking and e-money are an important barrier to their increased use. With electronic banking, you might worry that criminals might access your bank account and steal your money by moving your balances to someone else's account. Indeed, a notorious case of this happened in 1995, when a Russian computer programmer got access to Citibank's computers and moved funds electronically into his and his conspirators' accounts. Private solutions to deal with this problem have arisen with the development of more secure encryption technologies to prevent this kind of fraud. However, because bank customers are not knowledgeable about computer security issues, there is a role for the government to regulate electronic banking to make sure that encryption procedures are adequate. Similar encryption issues apply to e-money, so requirements that banks make it difficult for criminals to engage in digital counterfeiting make sense. To meet these challenges, bank examiners in the United States assess how a bank deals with the special security issues raised by electronic banking and also oversee third-party providers of electronic banking platforms. Also, because consumers want to know that electronic banking transactions are executed correctly, bank examiners also assess the technical skills of banks in setting up electronic banking services and the bank's capabilities for dealing with problems. Another security issue of concern to bank customers is the validity of digital signatures. The Electronic Signatures in Global and National Commerce Act of 2000 makes electronic signatures as legally binding as written signatures in most circumstances.

Electronic banking also raises serious privacy concerns. Because electronic transactions can be stored on databases, banks are able to collect a huge amount of information about their customers—their assets, creditworthiness, what they purchase, and so on—that can be sold to other financial institutions and businesses. This potential invasion of our privacy rightfully makes us very nervous. To protect customers' privacy, the Gramm-Leach-Bliley Act of 1999 has limited the distribution of these data, but it does not go as far as the European Data Protection Directive, which prohibits the transfer of information about online transactions. How to protect consumers' privacy in our electronic age is one of the great challenges our society faces, so privacy regulations for electronic banking are likely to evolve over time.

Table 1 Major Banking Legislation in the United States in the Twentieth Century

Federal Reserve Act (1913)
Created the Federal Reserve System

McFadden Act of 1927
Effectively prohibited banks from branching across state lines
Put national and state banks on equal footing regarding branching

Banking Act of 1933 (Glass-Steagall) and 1935
Created the FDIC
Separated commercial banking from the securities industry
Prohibited interest on checkable deposits and restricted such deposits to commercial banks
Put interest-rate ceilings on other deposits

(continues)

Table 1 Major Banking Legislation in the United States in the Twentieth Century *(continued)*

Bank Holding Company Act and Douglas Amendment (1956)
Clarified the status of bank holding companies (BHCs)
Gave the Federal Reserve regulatory responsibility for BHCs

Depository Institutions Deregulation and Monetary Control Act (DIDMCA) of 1980
Gave thrift institutions wider latitude in activities
Approved NOW and sweep accounts nationwide
Phased out interest rate ceilings on deposits
Imposed uniform reserve requirements on depository institutions
Eliminated usury ceilings on loans
Increased deposit insurance to $100,000 per account

Depository Institutions Act of 1982 (Garn–St. Germain)
Gave the FDIC and the FSLIC emergency powers to merge banks and thrifts across state lines
Allowed depository institutions to offer money market deposit accounts (MMDAs)
Granted thrifts wider latitude in commercial and consumer lending

Competitive Equality in Banking Act (CEBA) of 1987
Provided $10.8 billion to the FSLIC
Made provisions for regulatory forbearance in depressed areas

Financial Institutions Reform, Recovery, and Enforcement Act (FIRREA) of 1989
Provided funds to resolve S&L failures
Eliminated the FSLIC and the Federal Home Loan Bank Board
Created the Office of Thrift Supervision to regulate thrifts
Created the Resolution Trust Corporation to resolve insolvent thrifts
Raised deposit insurance premiums
Reimposed restrictions on S&L activities

Federal Deposit Insurance Corporation Improvement Act (FDICIA) of 1991
Recapitalized the FDIC
Limited brokered deposits and the too-big-to-fail policy
Set provisions for prompt corrective action
Instructed the FDIC to establish risk-based premiums
Increased examinations, capital requirements, and reporting requirements
Included the Foreign Bank Supervision Enhancement Act (FBSEA), which strengthened the Fed's Authority to supervise foreign banks

Riegle-Neal Interstate Banking and Branching Efficiency Act of 1994
Overturned prohibition of interstate banking
Allowed branching across state lines

Gramm-Leach-Bliley Financial Services Modernization Act of 1999
Repealed Glass-Steagall and removed the separation of banking and securities industries

International Banking Regulation

Because asymmetric information problems in the banking industry are a fact of life throughout the world, bank regulation in other countries is similar to that in the United States. Banks are chartered and supervised by government regulators, just as they are in the United States. Deposit insurance is also a feature of the regulatory systems in most other developed countries, although its coverage is often smaller than in the United States and is intentionally not advertised. We have also seen that bank capital requirements are in the process of being standardized across countries with agreements like the Basel Accord.

Problems in Regulating International Banking

Particular problems in bank regulation occur when banks are engaged in international banking and thus can readily shift their business from one country to another. Bank regulators closely examine the domestic operations of banks in their country, but they often do not have the knowledge or ability to keep a close watch on bank operations in other countries, either by domestic banks' foreign affiliates or by foreign banks with domestic branches. In addition, when a bank operates in many countries, it is not always clear which national regulatory authority should have primary responsibility for keeping the bank from engaging in overly risky activities. The difficulties inherent in regulating international banking were highlighted by the collapse of the Bank of Credit and Commerce International (BCCI). BCCI, which was operating in more than 70 countries, including the United States and the United Kingdom, was supervised by Luxembourg, a tiny country unlikely to be up to the task. When massive fraud was discovered, the Bank of England closed BCCI down, but not before depositors and stockholders were exposed to huge losses. Cooperation among regulators in different countries and standardization of regulatory requirements provide potential solutions to the problems of regulating international banking. The world has been moving in this direction through agreements like the Basel Accords and oversight procedures announced by the Basel Committee in July 1992, which require a bank's worldwide operations to be under the scrutiny of a single home-country regulator with enhanced powers to acquire information on the bank's activities. Also, the Basel Committee ruled that regulators in other countries can restrict the operations of a foreign bank if they feel that it lacks effective oversight. Whether agreements of this type will solve the problem of regulating international banking in the future is an open question.

Summary

Asymmetric information analysis explains what types of banking regulations are needed to reduce moral hazard and adverse selection problems in the banking system. However, understanding the theory behind regulation does not mean that regulation and supervision of the banking system are easy in practice. Getting bank regulators and supervisors to do their job properly is difficult for several reasons. First, as we learned in the discussion of financial innovation in Chapter 10, in their search for profits, financial institutions have strong incentives to avoid existing regulations by loophole mining. Thus regulation applies to a moving target: Regulators are continually playing cat-and-mouse with financial institutions—financial institutions think up clever ways to avoid regulations, which then causes regulators to modify their regulation activities. Regulators continually face new challenges in a dynamically changing financial system, and unless they can respond rapidly to change, they may not be able to keep financial institutions from taking on excessive risk. This problem can be exacerbated if regulators and supervisors do not have the resources or expert-

ise to keep up with clever people in financial institutions who think up ways to hide what they are doing or to get around the existing regulations.

Bank regulation and supervision are difficult for two other reasons. In the regulation and supervision game, the devil is in the details. Subtle differences in the details may have unintended consequences; unless regulators get the regulation and supervision just right, they may be unable to prevent excessive risk taking. In addition, regulators and supervisors may be subject to political pressure to not do their jobs properly. For all these reasons, there is no guarantee that bank regulators and supervisors will be successful in promoting a healthy financial system. Indeed, as we will see, bank regulation and supervision have not always worked well, leading to banking crises in the United States and throughout the world.

The 1980s U.S. Banking Crisis: Why?

Before the 1980s, federal deposit insurance seemed to work exceedingly well. In contrast to the pre-1934 period, when bank failures were common and depositors frequently suffered losses, the period from 1934 to 1980 was one in which bank failures were a rarity, averaging 15 a year for commercial banks and fewer than 5 a year for savings and loans. After 1981, this rosy picture changed dramatically. Failures in both commercial banks and savings and loans climbed to levels more than ten times greater than in earlier years, as can be seen in Figure 1. Why did this happen? How did a

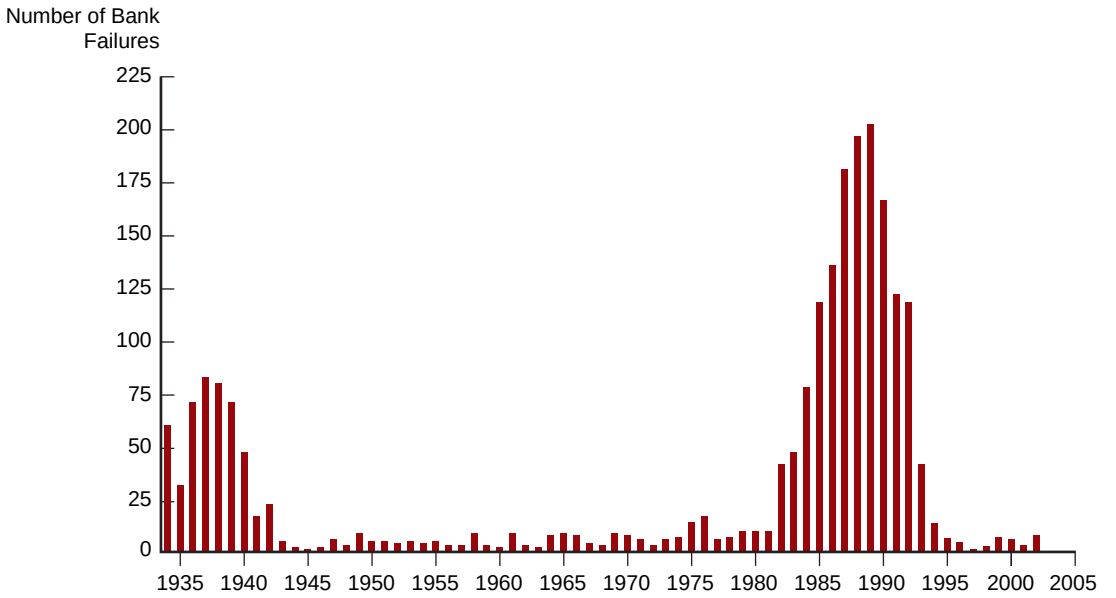


FIGURE 1 Bank Failures in the United States, 1934–2002

Source: www2.fdic.gov/qbp/2002sep/cbl.html.

deposit insurance system that seemed to be working well for half a century find itself in so much trouble?

Early Stages of the Crisis

The story starts with the burst of financial innovation in the 1960s, 1970s, and early 1980s. As we saw in the previous chapter, financial innovation decreased the profitability of certain traditional business for commercial banks. Banks now faced increased competition for their sources of funds from new financial institutions such as money market mutual funds while they were losing commercial lending business to the commercial paper market and securitization.

With the decreasing profitability of their traditional business, by the mid-1980s, commercial banks were forced to seek out new and potentially risky business to keep their profits up, by placing a greater percentage of their total loans in real estate and in credit extended to assist corporate takeovers and leveraged buyouts (called *highly leveraged transaction loans*).

The existence of deposit insurance increased moral hazard for banks because insured depositors had little incentive to keep the banks from taking on too much risk. Regardless of how much risk banks were taking, deposit insurance guaranteed that depositors would not suffer any losses.

Adding fuel to the fire, financial innovation produced new financial instruments that widened the scope for risk taking. New markets in financial futures, junk bonds, swaps, and other instruments made it easier for banks to take on extra risk—making the moral hazard problem more severe. New legislation that deregulated the banking industry in the early 1980s, the Depository Institutions Deregulation and Monetary Control Act (DIDMCA) of 1980 and the Depository Institutions (Garn–St. Germain) Act of 1982, gave expanded powers to the S&Ls and mutual savings banks to engage in new risky activities. These thrift institutions, which had been restricted almost entirely to making loans for home mortgages, now were allowed to have up to 40% of their assets in commercial real estate loans, up to 30% in consumer lending, and up to 10% in commercial loans and leases. In the wake of this legislation, S&L regulators allowed up to 10% of assets to be in junk bonds or in direct investments (common stocks, real estate, service corporations, and operating subsidiaries).

www2.fdic.gov/hsob/SelectRpt.asp?EntryTyp=30
Search for data on bank failures in any year.

In addition, DIDMCA increased the mandated amount of federal deposit insurance from $40,000 per account to $100,000 and phased out Regulation Q deposit-rate ceilings. Banks and S&Ls that wanted to pursue rapid growth and take on risky projects could now attract the necessary funds by issuing larger-denomination insured certificates of deposit with interest rates much higher than those being offered by their competitors. Without deposit insurance, high interest rates would not have induced depositors to provide the high-rolling banks with funds because of the realistic expectation that they might not get the funds back. But with deposit insurance, the government was guaranteeing that the deposits were safe, so depositors were more than happy to make deposits in banks with the highest interest rates.

Financial innovation and deregulation in the permissive atmosphere of the Reagan years led to expanded powers for the S&L industry that in turn led to several problems. First, many S&L managers did not have the required expertise to manage risk appropriately in these new lines of business. Second, the new expanded powers meant that there was a rapid growth in new lending, particularly to the real estate sector. Even if the required expertise was available initially, rapid credit growth may outstrip the available information resources of the banking institution, resulting in excessive risk taking. Third, these new powers of the S&Ls and

the lending boom meant that their activities were expanding in scope and were becoming more complicated, requiring an expansion of regulatory resources to monitor these activities appropriately. Unfortunately, regulators of the S&Ls at the Federal Savings and Loan Insurance Corporation (FSLIC) had neither the expertise nor the resources that would have enabled them to monitor these new activities sufficiently. Given the lack of expertise in both the S&L industry and the FSLIC, the weakening of the regulatory apparatus, and the moral hazard incentives provided by deposit insurance, it is no surprise that S&Ls took on excessive risks, which led to huge losses on bad loans.

In addition, the incentives of moral hazard were increased dramatically by a historical accident: the combination of sharp increases in interest rates from late 1979 until 1981 and a severe recession in 1981–1982, both of which were engineered by the Federal Reserve to bring down inflation. The sharp rises in interest rates produced rapidly rising costs of funds for the savings and loans that were not matched by higher earnings on the S&Ls' principal asset, long-term residential mortgages (whose rates had been fixed at a time when interest rates were far lower). The 1981–1982 recession and a collapse in the prices of energy and farm products hit the economies of certain parts of the country, such as Texas, very hard. As a result, there were defaults on many S&L loans. Losses for savings and loan institutions mounted to $10 billion in 1981–1982, and by some estimates over half of the S&Ls in the United States had a negative net worth and were thus insolvent by the end of 1982.

Later Stages of the Crisis: Regulatory Forbearance

At this point, a logical step might have been for the S&L regulators—the Federal Home Loan Bank Board and its deposit insurance subsidiary, the Federal Savings and Loan Insurance Fund (FSLIC), both now abolished—to close the insolvent S&Ls. Instead, these regulators adopted a stance of **regulatory forbearance**: They refrained from exercising their regulatory right to put the insolvent S&Ls out of business. To sidestep their responsibility to close ailing S&Ls, they adopted irregular regulatory accounting principles that in effect substantially lowered capital requirements. For example, they allowed S&Ls to include in their capital calculations a high value for intangible capital, called **goodwill** (an accounting entry to reflect value to the firm of its having special expertise or a particularly profitable business line).

There were three main reasons why the Federal Home Loan Bank Board and FSLIC opted for regulatory forbearance. First, the FSLIC did not have sufficient funds in its insurance fund to close the insolvent S&Ls and pay off their deposits. Second, the Federal Home Loan Bank Board was established to encourage the growth of the savings and loan industry, so the regulators were probably too close to the people they were supposed to be regulating. Third, because bureaucrats do not like to admit that their own agency is in trouble, the Federal Home Loan Bank Board and the FSLIC preferred to sweep their problems under the rug in the hope that they would go away.

Regulatory forbearance increases moral hazard dramatically because an operating but insolvent S&L (nicknamed a "zombie S&L" by economist Edward Kane because it is the "living dead") has almost nothing to lose by taking on great risk and "betting the bank": If it gets lucky and its risky investments pay off, it gets out of insolvency. Unfortunately, if, as is likely, the risky investments don't pay off, the zombie S&L's losses will mount, and the deposit insurance agency will be left holding the bag.

This strategy is similar to the "long bomb" strategy in football. When a football team is almost hopelessly behind and time is running out, it often resorts to a high-risk

play: the throwing of a long pass to try to score a touchdown. Of course, the long bomb is unlikely to be successful, but there is always a small chance that it will work. If it doesn't, the team has lost nothing, since it would have lost the game anyway.

Given the sequence of events we have discussed here, it should be no surprise that savings and loans began to take huge risks: They built shopping centers in the desert, bought manufacturing plants to convert manure to methane, and purchased billions of dollars of high-risk, high-yield junk bonds. The S&L industry was no longer the staid industry that once operated on the so-called *3-6-3 rule*: You took in money at 3%, lent it at 6%, and played golf at 3 P.M. Although many savings and loans were making money, losses at other S&Ls were colossal.

Another outcome of regulatory forbearance was that with little to lose, zombie S&Ls attracted deposits away from healthy S&Ls by offering higher interest rates. Because there were so many zombie S&Ls in Texas pursuing this strategy, above-market interest rates on deposits at Texas S&Ls were said to have a "Texas premium." Potentially healthy S&Ls now found that to compete for deposits, they had to pay higher interest rates, which made their operations less profitable and frequently pushed them into the zombie category. Similarly, zombie S&Ls in pursuit of asset growth made loans at below-market interest rates, thereby lowering loan interest rates for healthy S&Ls, and again made them less profitable. The zombie S&Ls had actually taken on attributes of vampires—their willingness to pay above-market rates for deposits and take below-market interest rates on loans was sucking the lifeblood (profits) out of healthy S&Ls.

Competitive Equality in Banking Act of 1987

Toward the end of 1986, the growing losses in the savings and loan industry were bankrupting the insurance fund of the FSLIC. The Reagan administration sought $15 billion in funds for the FSLIC, a completely inadequate sum considering that many times this amount was needed to close down insolvent S&Ls. The legislation passed by Congress, the Competitive Equality in Banking Act (CEBA) of 1987, did not even meet the administration's requests. It allowed the FSLIC to borrow only $10.8 billion through a subsidiary corporation called Financing Corporation (FICO) and, what was worse, included provisions that directed the Federal Home Loan Bank Board to continue to pursue regulatory forbearance (allow insolvent institutions to keep operating), particularly in economically depressed areas such as Texas.

The failure of Congress to deal with the savings and loan crisis was not going to make the problem go away. Consistent with our analysis, the situation deteriorated rapidly. Losses in the savings and loan industry surpassed $10 billion in 1988 and approached $20 billion in 1989. The crisis was reaching epidemic proportions. The collapse of the real estate market in the late 1980s led to additional huge loan losses that greatly exacerbated the problem.

Political Economy of the Savings and Loan Crisis

Although we now have a grasp of the regulatory and economic forces that created the S&L crisis, we still need to understand the political forces that produced the regulatory structure and activities that led to it. The key to understanding the political economy of the S&L crisis is to recognize that the relationship between voter-taxpayers and the regulators and politicians creates a particular type of moral hazard problem,

discussed in Chapter 8: the *principal–agent problem*, which occurs when representatives (agents) such as managers have incentives that differ from those of their employer (the principal) and so act in their own interest rather than in the interest of the employer.

The Principal–Agent Problem for Regulators and Politicians

Regulators and politicians are ultimately agents for voter-taxpayers (principals), because in the final analysis, taxpayers bear the cost of any losses by the deposit insurance agency. The principal–agent problem occurs because the agent (a politician or regulator) does not have the same incentives to minimize costs to the economy as the principal (the taxpayer).

To act in the taxpayers' interest and lower costs to the deposit insurance agency, regulators have several tasks, as we have seen. They must set tight restrictions on holding assets that are too risky, must impose high capital requirements, and must not adopt a stance of regulatory forbearance, which allows insolvent institutions to continue to operate. However, because of the principal–agent problem, regulators have incentives to do the opposite. Indeed, as our sad saga of the S&L debacle indicates, they have at times loosened capital requirements and restrictions on risky asset holdings and pursued regulatory forbearance. One important incentive for regulators that explains this phenomenon is their desire to escape blame for poor performance by their agency. By loosening capital requirements and pursuing regulatory forbearance, regulators can hide the problem of an insolvent bank and hope that the situation will improve. Edward Kane characterizes such behavior on the part of regulators as "bureaucratic gambling."

Another important incentive for regulators is that they want to protect their careers by acceding to pressures from the people who most influence their careers. These people are not the taxpayers but the politicians who try to keep regulators from imposing tough regulations on institutions that are major campaign contributors. Members of Congress have often lobbied regulators to ease up on a particular S&L that contributed large sums to their campaigns. Regulatory agencies that have little independence from the political process are more vulnerable to these pressures.

In addition, both Congress and the presidential administration promoted banking legislation in 1980 and 1982 that made it easier for savings and loans to engage in risk-taking activities. After the legislation passed, the need for monitoring the S&L industry increased because of the expansion of permissible activities. The S&L regulatory agencies needed more resources to carry out their monitoring activities properly, but Congress (successfully lobbied by the S&L industry) was unwilling to allocate the necessary funds. As a result, the S&L regulatory agencies became so short-staffed that they actually had to cut back on their on-site examinations just when these were needed most. In the period from January 1984 to July 1986, for example, several hundred S&Ls were not examined even once. Even worse, spurred on by the intense lobbying efforts of the S&L industry, Congress passed the Competitive Equality in Banking Act of 1987, which, as we have seen, provided inadequate funding to close down the insolvent S&Ls and also hampered the S&L regulators from doing their job properly by including provisions encouraging regulatory forbearance.

As these examples indicate, the structure of our political system has created a serious principal–agent problem; politicians have strong incentives to act in their own interests rather than in the interests of taxpayers. Because of the high cost of running

campaigns, American politicians must raise substantial contributions. This situation may provide lobbyists and other campaign contributors with the opportunity to influence politicians to act against the public interest.

Savings and Loan Bailout: The Financial Institutions Reform, Recovery, and Enforcement Act of 1989

Immediately after taking office, the first Bush administration proposed new legislation to provide adequate funding to close down the insolvent S&Ls. The resulting legislation, the Financial Institutions Reform, Recovery, and Enforcement Act (FIRREA), was signed into law on August 9, 1989. It was the most significant legislation to affect the thrift industry since the 1930s. FIRREA's major provisions were as follows: The regulatory apparatus was significantly restructured, eliminating the Federal Home Loan Bank Board and the FSLIC, both of which had failed in their regulatory tasks. The regulatory role of the Federal Home Loan Bank Board was relegated to the Office of Thrift Supervision (OTS), a bureau within the U.S. Treasury Department whose responsibilities are similar to those that the Office of the Comptroller of the Currency has over the national banks. The regulatory responsibilities of the FSLIC were given to the FDIC, and the FDIC became the sole administrator of the federal deposit insurance system with two separate insurance funds: the Bank Insurance Fund (BIF) and the Savings Association Insurance Fund (SAIF). Another new agency, the Resolution Trust Corporation (RTC), was established to manage and resolve insolvent thrifts placed in conservatorship or receivership. It was made responsible for selling more than $450 billion of real estate owned by failed institutions. After seizing the assets of about 750 insolvent S&Ls, over 25% of the industry, the RTC sold over 95% of them, with a recovery rate of over 85%. After this success, the RTC went out of business on December 31, 1995.

The cost of the bailout ended up on the order of $150 billion. The funding for the bailout came partly from capital in the Federal Home Loan Banks (owned by the S&L industry) but mostly from the sale of government debt by both the Treasury and the Resolution Funding Corporation (RefCorp).

FIRREA also imposed new restrictions on thrift activities that in essence reregulated the S&L industry to the asset choices it had before 1982. It increased the core-capital leverage requirement from 3% to 8% and imposed the same risk-based capital standards imposed on commercial banks. FIRREA also enhanced the enforcement powers of thrift regulators by making it easier for them to remove managers, issue cease and desist orders, and impose civil money penalties.

FIRREA was a serious attempt to deal with some of the problems created by the S&L crisis in that it provided substantial funds to close insolvent thrifts. However, the losses that continued to mount for the FDIC in 1990 and 1991 would have depleted its Bank Insurance Fund by 1992, requiring that this fund be recapitalized. In addition, FIRREA did not focus on the underlying adverse selection and moral hazard problems created by deposit insurance. FIRREA did, however, mandate that the U.S. Treasury produce a comprehensive study and plan for reform of the federal deposit insurance system. After this study appeared in 1991, Congress passed the Federal Deposit Insurance Corporation Improvement Act (FDICIA), which engendered major reforms in the bank regulatory system.

Federal Deposit Insurance Corporation Improvement Act of 1991

FDICIA's provisions were designed to serve two purposes: to recapitalize the Bank Insurance Fund of the FDIC and to reform the deposit insurance and regulatory system so that taxpayer losses would be minimized.

FDICIA recapitalized the Bank Insurance Fund by increasing the FDIC's ability to borrow from the Treasury and also mandated that the FDIC assess higher deposit insurance premiums until it could pay back its loans and achieve a level of reserves in its insurance funds that would equal 1.25 percent of insured deposits.

The bill reduced the scope of deposit insurance in several ways, but the most important one is that the too-big-to-fail doctrine has been substantially limited. The FDIC must now close failed banks using the least-costly method, thus making it far more likely that uninsured depositors will suffer losses. An exception to this provision, whereby a bank would be declared too big to fail so that all depositors, both insured and uninsured, would be fully protected, would be allowed only if not doing so would "have serious adverse effects on economic conditions or financial stability." Furthermore, to invoke the too-big-to-fail policy, a two-thirds majority of both the Board of Governors of the Federal Reserve System and the directors of the FDIC, as well as the approval of the Secretary of the Treasury, are required. Furthermore, FDICIA requires that the Fed share the FDIC's losses if long-term Fed lending to a bank that fails increases the FDIC's losses.

Probably the most important feature of FDICIA is its prompt corrective action provisions, which require the FDIC to intervene earlier and more vigorously when a bank gets into trouble. Banks are now classified into five groups based on bank capital. Group 1, classified as "well capitalized," are banks that significantly exceed minimum capital requirements and are allowed privileges such as the ability to do some securities underwriting. Banks in group 2, classified as "adequately capitalized," meet minimum capital requirements and are not subject to corrective actions but are not allowed the privileges of the well-capitalized banks. Banks in group 3, "undercapitalized," fail to meet capital requirements. Banks in groups 4 and 5 are "significantly undercapitalized" and "critically undercapitalized," respectively, and are not allowed to pay interest on their deposits at rates that are higher than average. In addition, for group 3 banks, the FDIC is required to take prompt corrective actions such as requiring them to submit a capital restoration plan, restrict their asset growth, and seek regulatory approval to open new branches or develop new lines of business. Banks that are so undercapitalized as to have equity capital less than 2% of assets fall into group 5, and the FDIC must take steps to close them down.

FDICIA also instructed the FDIC to come up with risk-based insurance premiums. The system that the FDIC has put in place, however, has not worked very well, because it resulted in more than 90% of the banks with over 95% of the deposits paying the same premium. Other provisions of FDICIA are listed in Table 1 on page 271.

FDICIA is an important step in the right direction, because it increases the incentives for banks to hold capital and decreases their incentives to take on excessive risk. However, concerns that it has not adequately addressed risk-based premiums or the too-big-to-fail problem have led economists to continue to search for further reforms that might help promote the safety and soundness of the banking system.[3]

[3] A further discussion of how well FDICIA has worked and other proposed reforms of the banking regulatory system appears in an appendix to this chapter that can be found on this book's web site at www.aw.com/mishkin.

Banking Crises Throughout the World

Because misery loves company, it may make you feel better to know that the United States has by no means been alone in suffering a banking crisis. Indeed, as Table 2 and Figure 2 illustrate, banking crises have struck a large number of countries throughout the world, and many of them have been substantially worse than ours. We will examine what took place in several of these other countries and see that the same forces that produced a banking crisis in the United States have been at work elsewhere too.

Scandinavia

As in the United States, an important factor in the banking crises in Norway, Sweden, and Finland was the financial liberalization that occurred in the 1980s. Before the 1980s, banks in these Scandinavian countries were highly regulated and subject to restrictions on the interest rates they could pay to depositors and on the interest rates they could earn on loans. In this noncompetitive environment, and with artificially low rates on both deposits and loans, these banks lent only to the best credit risks, and both banks and their regulators had little need to develop expertise in screening and monitoring borrowers. With the deregulated environment, a lending boom ensued, particularly in the real estate sector. Given the lack of expertise in both the banking industry and its regulatory authorities in keeping risk taking in check, banks

Table 2 The Cost of Rescuing Banks in Several Countries

Date	Country	Cost as a % of GDP
1980–1982	Argentina	55
1997–ongoing	Indonesia	50
1981–1983	Chile	41
1997–ongoing	Thailand	33
1997–ongoing	South Korea	27
1997–ongoing	Malaysia	16
1994–1997	Venezuela	22
1995	Mexico	19
1990–ongoing	Japan	20
1989–1991	Czech Republic	12
1991–1994	Finland	11
1991–1995	Hungary	10
1994–1996	Brazil	13
1987–1993	Norway	8
1998	Russia	5–7
1991–1994	Sweden	4
1984–1991	United States	3

Source: Daniela Klingebiel and Luc Laewen, eds., *Managing the Real and Fiscal Effects of Banking Crises*, World Bank Discussion Paper No. 428 (Washington: World Bank, 2002).

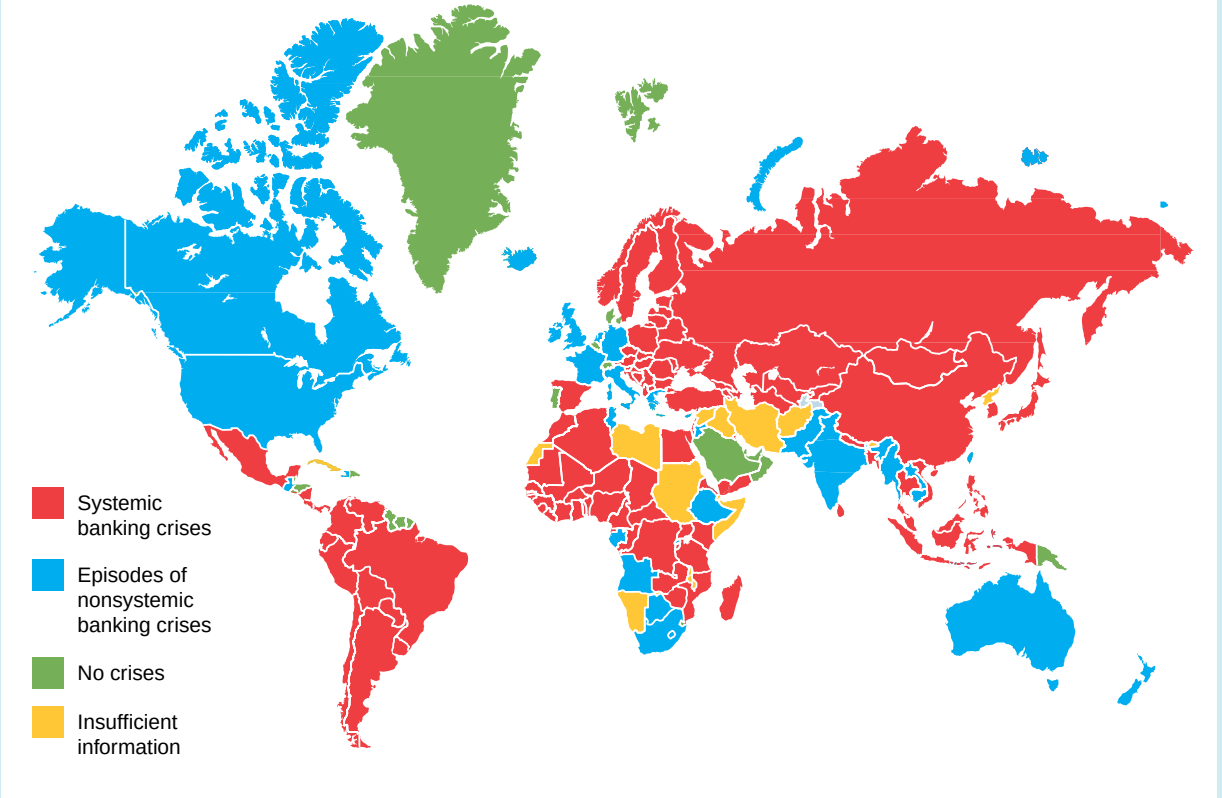


FIGURE 2 Banking Crises Throughout the World Since 1970

Source: Gerard Caprio and Daniela Klingebiel, "Episodes of Systemic and Borderline Financial Crises" mimeo., World Bank, October 1999.

engaged in risky lending. When real estate prices collapsed in the late 1980s, massive loan losses resulted. The outcome of this process was similar to what happened in the savings and loan industry in the United States. The government was forced to bail out almost the entire banking industry in these countries in the late 1980s and early 1990s on a scale that was even larger relative to GDP than in the United States (see Table 2).

Latin America

The Latin American banking crises typically show a pattern similar to those in the United States and in Scandinavia. Before the 1980s, banks in many Latin American countries were owned by the government and were subject to interest-rate restrictions as in Scandinavia. Their lending was restricted to the government and other low-risk borrowers. With the deregulation trend that was occurring world-wide, many of these countries liberalized their credit markets and privatized their banks. We then see the same pattern we saw in the United States and Scandinavia, a lending boom in the face of inadequate expertise on the part of both bankers and regulators. The result was again massive loan losses and the inevitable government bailout. The Argentine banking crisis of 2001, which is ongoing, differed from those typically seen in Latin America. Argentina's banks were well supervised and in relatively good shape before

the government coerced them into purchasing large amounts of Argentine government debt in order to help solve the government's fiscal problem. However, when market confidence in the government plummeted, spreads between Argentine government debt and U.S. Treasuries soared to more than 2,500 basis points (25 percentage points), leading to a sharp fall in the price of these securities. The losses on their holdings of government debt and rising bad loans because of the ongoing severe recession increased doubts about the solvency of the banking system.

A banking panic erupted in October and November 2001, with the Argentine public rushing to withdraw their deposits. On December 1, after losing more than $8 billion of deposits, the government imposed a $1,000 monthly limit on deposit withdrawals. Then with the collapse of the peso and the requirement that the banks must pay back their dollar deposits at a higher exchange value than they would be paid back on their dollar loans, banks' balance sheets went even further in the hole. The cost of the recent Argentine banking crisis is not yet clear, but it could very well be as large as the previous banking crisis in Argentina in the 1980–1982 period listed in Table 2 and could exceed 50% of GDP.

What is particularly striking about the Latin American experience is that the cost of the bailouts relative to GDP dwarfs that in the United States. The cost to the taxpayer of the government bailouts in Latin America has been anywhere from around 20% to more than 50% of GDP, in contrast to the 3% figure for the United States.

Russia and Eastern Europe

Before the end of the Cold War, in the communist countries of Eastern Europe and the Soviet Union, banks were owned by the state. When the downfall of communism occurred, banks in these countries had little expertise in screening and monitoring loans. Furthermore, bank regulatory and supervisory apparatus that could rein in the banks and keep them from taking on excessive risk barely existed. Given the lack of expertise on the part of regulators and banks, not surprisingly, substantial loan losses ensued, resulting in the failure or government bailout of many banks. For example, in the second half of 1993, eight banks in Hungary with 25% of the financial system's assets were insolvent, and in Bulgaria, an estimated 75% of all loans in the banking system were estimated to be substandard in 1995.

On August 24, 1995, a bank panic requiring government intervention occurred in Russia when the interbank loan market seized up and stopped functioning because of concern about the solvency of many new banks. This was not the end of troubles in the Russian banking system. On August 17, 1998, the Russian government announced that Russia would impose a moratorium on the repayment of foreign debt because of insolvencies in the banking system. In November, the Russian central bank announced that nearly half of the country's 1,500 commercial banks might go under and the cost of the bailout is expected to be on the order of $15 billion.

Japan

Japan was a latecomer to the banking crisis game. Before 1990, the vaunted Japanese economy looked unstoppable. Unfortunately, it has recently experienced many of the same pathologies that we have seen in other countries. Before the 1980s, Japan's financial markets were among the most heavily regulated in the world, with very strict restrictions on the issuing of securities and interest rates. Financial deregulation and innovation produced a more competitive environment that set off a lending boom, with banks lending aggressively in the real estate sector. As in the other countries we have examined here, financial disclosure and monitoring by regulators did not keep

pace with the new financial environment. The result was that banks could and did take on excessive risks. When property values collapsed in the early 1990s, the banks were left holding massive amounts of bad loans. For example, Japanese banks decided to get into the mortgage lending market by setting up the so-called *jusen*, home mortgage lending companies that raised funds by borrowing from banks and then lent these funds out to households. Seven of these *jusen* became insolvent, leaving banks with $60 billion or so of bad loans.

As a result the Japanese have experienced their first bank failures since World War II. In July 1995, Tokyo-based Cosmo Credit Corporation, Japan's fifth-largest credit union, failed and on August 30, the Osaka authorities announced the imminent closing of Kizu Credit Cooperative, Japan's second-largest credit union. (Kizu's story is remarkably similar to that of many U.S. savings and loans. Kizu, like many American S&Ls, began offering high rates on large time deposits and grew at a blistering pace, with deposits rising from $2.2 billion in 1988 to $12 billion by 1995 and real estate loans growing by a similar amount. When the property market collapsed, so did Kizu.) On the same day, the Ministry of Finance announced that it was liquidating Hyogo Bank, a midsize Kobe bank that was the first commercial bank to fail. Larger banks now began to follow the same path. In late 1996, the Hanwa Bank, a large regional bank, was liquidated, followed in 1997 by a government-assisted restructuring of the Nippon Credit Bank, Japan's seventeenth-largest bank. In November 1997, Hokkaido Takushoku Bank was forced to go out of business, making it the first city bank (a large commercial bank) to be closed during the crisis.

The Japanese have been going through a cycle of forbearance similar to the one that occurred in the United States in the 1980s. The Japanese regulators in the Ministry of Finance enabled banks to meet capital standards and to keep operating by allowing them to artificially inflate the value of their assets. For example, they were allowed to value their large holdings of equities at historical value, rather than market value, which was much lower. Inadequate amounts were allocated for recapitalization of the banking system, and the extent of the problem was grossly underestimated by government officials. Furthermore, until the closing of the Hokkaido Takushoku Bank, the bank regulators in the Ministry of Finance were unwilling to close down city banks and impose any losses on stockholders or any uninsured creditors.

By the middle of 1998, the Japanese government began to take some steps to attack these problems. In June, supervisory authority over financial institutions was taken away from the Ministry of Finance and transferred to the Financial Supervisory Agency (FSA), which reports directly to the prime minister. This was the first instance in half a century in which the all-powerful Ministry of Finance was stripped of some of its authority. In October, the parliament passed a bailout package of $500 billion (60 trillion yen). However, disbursement of the funds depended on the voluntary cooperation of the banks: the law did not require insolvent banks to close or to accept the funds. Indeed, acceptance of the funds required the bailed-out bank to open its books and reveal its true losses, and thus many banks remain very undercapitalized. The banking sector in Japan thus remains in very poor shape: It is burdened with bad loans and poor profitability. Indeed, many private sector analysts estimate that bad loans have reached a level of more than $1 trillion.

There has been some progress in cleaning up the banking mess: immediately after the 1998 banking law was passed, one of the ailing city banks, Long-Term Credit Bank of Japan, was taken over by the government and declared insolvent, and in December 1998, the Nippon Credit Bank was finally put out of its misery and closed

down by the government. Since then, the clean-up process has stalled and the economy has remained weak, with a growth rate from 1991–2002 averaging an anemic 1%. A new, reform-oriented prime minister, Junichiro Koizumi, who pledged to clean up the banking system, came into office in 2001; yet he has been unable to come to grips with the Japanese banking problem. This situation does not bode well for the Japanese economy.

East Asia

We discussed the banking and financial crisis in the East Asian countries (Thailand, Malaysia, Indonesia, the Philippines, and South Korea) in Chapter 8. Due to inadequate supervision of the banking system, the lending booms that arose in the aftermath of financial liberalization led to substantial loan losses, which became huge after the currency collapses that occurred in the summer of 1997. An estimated 15% to 35% of all bank loans have turned sour in Thailand, Indonesia, Malaysia, and South Korea, and the cost of the bailout for the banking system is estimated at more than 20% of GDP in these countries and over 50% of GDP in Indonesia. The Philippines is expected to fare somewhat better, with the cost below 10% of GDP.

"Déjà Vu All Over Again"

What we see in banking crises in these different countries is that history has kept on repeating itself. The parallels between the banking crisis episodes in all these countries are remarkably similar, leaving us with a feeling of déjà vu. Although financial liberalization is generally a good thing because it promotes competition and can make a financial system more efficient, as we have seen in the countries examined here, it can lead to an increase in moral hazard, with more risk taking on the part of banks if there is lax regulation and supervision; the result can then be banking crises. However, these episodes do differ in that deposit insurance has not played an important role in many of the countries experiencing banking crises. For example, the size of the Japanese equivalent of the FDIC, the Deposit Insurance Corporation, was so tiny relative to the FDIC that it did not play a prominent role in the banking system and exhausted its resources almost immediately with the first bank failures. This means that deposit insurance is not to blame for some of these banking crises. However, what is common to all the countries discussed here is the existence of a government safety net, in which the government stands ready to bail out banks whether deposit insurance is an important feature of the regulatory environment or not. It is the existence of a government safety net, and not deposit insurance per se, that increases moral hazard incentives for excessive risk taking on the part of banks.

Summary

1. The concepts of asymmetric information, adverse selection, and moral hazard help explain the eight types of banking regulation that we see in the United States and other countries: the government safety net, restrictions on bank asset holdings, capital requirements, bank supervision, assessment of risk mangement, disclosure requirements, consumer protection, and restrictions on competition.
2. Because asymmetric information problems in the banking industry are a fact of life throughout the world, bank regulation in other countries is similar to that in the United States. It is particularly problematic to regulate banks engaged in international banking, because they can readily shift their business from one country to another.

3. Because of financial innovation, deregulation, and a set of historical accidents, adverse selection and moral hazard problems increased in the 1980s and resulted in huge losses for the U.S. savings and loan industry and for taxpayers.
4. Regulators and politicians are subject to the principal–agent problem, meaning that they may not have sufficient incentives to minimize the costs of deposit insurance to taxpayers. As a result, regulators and politicians relaxed capital standards, removed restrictions on holdings of risky assets, and relied on regulatory forbearance, thereby increasing the costs of the S&L bailout.
5. The Financial Institutions Reform, Recovery, and Enforcement Act (FIRREA) of 1989 provided funds for the S&L bailout; created the Resolution Trust Corporation to manage the resolution of insolvent thrifts; eliminated the Federal Home Loan Bank Board and gave its regulatory role to the Office of Thrift Supervision; eliminated the FSLIC, whose insurance role and regulatory responsibilities were taken over by the FDIC; imposed restrictions on thrift activities similar to those in effect before 1982; increased the capital requirements to those adhered to by commercial banks; and increased the enforcement powers of thrift regulators.
6. The Federal Deposit Insurance Corporation Improvement Act (FDICIA) of 1991 recapitalized the Bank Insurance Fund of the FDIC and included reforms for the deposit insurance and regulatory system so that taxpayer losses would be minimized. This legislation limited brokered deposits and the use of the too-big-to-fail policy, mandated prompt corrective action to deal with troubled banks, and instituted risk-based deposit insurance premiums. These provisions have helped reduce the incentives of banks to take on excessive risk and so should help reduce taxpayer exposure in the future.
7. The parallels between the banking crisis episodes that have occurred in countries throughout the world are striking, indicating that similar forces are at work.

Key Terms

bank failure, p. 260
bank supervision (prudential supervision), p. 265
Basel Accord, p. 265
Basel Committee on Banking Supervision, p. 265
goodwill, p. 275
leverage ratio, p. 265
off-balance-sheet activities, p. 265
regulatory arbitrage, p. 265
regulatory forbearance, p. 275

Questions and Problems

Questions marked with an asterisk are answered at the end of the book in an appendix, "Answers to Selected Questions and Problems."

1. Give one example each of moral hazard and adverse selection in private insurance arrangements.
*2. If casualty insurance companies provided fire insurance without any restrictions, what kind of adverse selection and moral hazard problems might result?
3. What bank regulation is designed to reduce adverse selection problems for deposit insurance? Will it always work?
*4. What bank regulations are designed to reduce moral hazard problems created by deposit insurance? Will they completely eliminate the moral hazard problem?
5. What are the costs and benefits of a too-big-to-fail policy?
*6. Why did the S&L crisis not occur until the 1980s?
7. Why is regulatory forbearance a dangerous strategy for a deposit insurance agency?
*8. The FIRREA legislation of 1989 is the most comprehensive banking legislation since the 1930s. Describe its major features.

9. What steps were taken in the FDICIA legislation of 1991 to improve the functioning of federal deposit insurance?

*10. Some advocates of campaign reform believe that government funding of political campaigns and restrictions on campaign spending might reduce the principal–agent problem in our political system. Do you agree? Explain your answer.

11. How can the S&L crisis be blamed on the principal–agent problem?

*12. Do you think that eliminating or limiting the amount of deposit insurance would be a good idea? Explain your answer.

13. Do you think that removing the impediments to a nationwide banking system will be beneficial to the economy? Explain your answer.

*14. How could higher deposit insurance premiums for banks with riskier assets benefit the economy?

15. How has too-big-too-fail been limited in the FDICIA legislation? How might limiting too-big-too-fail help reduce the risk of a future banking crisis?

Web Exercises

1. Go to www.fdic.gov/regulations/laws/important/index.html. This site reports on the most significant pieces of legislation affecting banks since the 1800s. Summarize the most recently enacted bank regulation listed on this site.

2. The Office of the Comptroller of the Currency is responsible for many of the regulations affecting bank operations. Go to www.occ.treas.gov/. Click on "Regulatory Information." Now click on the 12 CFR Parts 1 to 199. What does Part 1 cover? How many parts are there in 12 CRF? Open Part 18. What topic does it cover? Summarize its purpose.

appendix to chapter 11

Evaluating FDICIA and Other Proposed Reforms of the Banking Regulatory System

FDICIA is a major step in reforming the banking regulatory system. How well will it work to solve the adverse selection and moral hazard problems of the bank regulatory system? Let's use the analysis in the chapter to evaluate the most important provisions of this legislation to answer this question.

Study Guide

Before looking at the evaluation for each set of provisions and proposals in this application, try to reason out how well they will solve the current problems with banking regulation. This exercise will help you develop a deeper understanding of the material in this chapter.

Limits on the Scope of Deposit Insurance

FDICIA's reduction of the scope of deposit insurance by limiting insurance on brokered deposits and restricting the use of the too-big-to-fail policy might have increased the incentives for uninsured depositors to monitor banks and to withdraw funds if the bank is taking on too much risk. Because banks might now fear the loss of deposits when they engage in risky activities, they might have less incentive to take on too much risk. Limitations on the use of the too-big-to-fail policy starting in 1992 have resulted in increased losses to uninsured depositors at failed banks as planned.

Although the cited elements of FDICIA strengthen the incentive of depositors to monitor banks, some critics of FDICIA would take these limitations on the scope of deposit insurance even further. Some suggest that deposit insurance should be eliminated entirely or should be reduced in amount from the current $100,000 limit to, say, $50,000 or $20,000. Another proposed reform would institute a system of *coinsurance* in which only a percentage of a deposit—say, 90%—would be covered by insurance. In this system, the insured depositor would suffer a percentage of the losses along with the deposit insurance agency. Because depositors facing a lower limit on deposit insurance or coinsurance would suffer losses if the bank goes broke, they will have an incentive to monitor the bank's activities. Other critics believe that FDICIA provides too much support for the too-big-to-fail policy. Because under FDICIA the Fed, the Treasury, and the FDIC can still agree to implement too-big-to-fail and thus bail out uninsured as well as insured depositors, big banks will not be subjected to enough discipline by uninsured depositors. These critics advocate eliminating the too-big-to-fail policy entirely, thereby decreasing the incentives of big banks to take on too much risk.

However, other experts do not believe that depositors are capable of monitoring banks and imposing discipline on them. The basic problem with reducing the scope of deposit insurance even further as proposed is that banks would be subject to runs, sudden withdrawals by nervous depositors. Such runs could by themselves lead to bank failures. In addition to protecting individual depositors, the purpose of deposit insurance is to prevent a large number of bank failures, which would lead to an unstable banking system and an unstable economy as occurred periodically before the establishment of federal deposit insurance in 1934. From this perspective, federal deposit insurance has been a resounding success. Bank panics, in which there are simultaneous failures of many banks and consequent disruption of the financial system, have not occurred since federal deposit insurance was established.

On the one hand, evidence that the largest banks benefiting from the de facto too-big-to-fail policy before 1991 were also the ones that took on the most risk suggests that limiting its application, as FDICIA does, may substantially reduce risk taking. On the other hand, eliminating the too-big-to-fail policy altogether would also cause some of the same problems that would occur if deposit insurance were eliminated or reduced: The probability of bank panics would increase. If a big bank were allowed to fail, the repercussions in the financial system might be immense. Other banks with a correspondent relationship with the failed bank (those that have deposits at the bank in exchange for a variety of services) would suffer large losses and might fail in turn, leading to a full-scale panic. In addition, the problem of liquidating the big bank's loan portfolio might create a major disruption in the financial market.

Prompt Corrective Action

The prompt corrective action provisions of FDICIA should also substantially reduce incentives for bank risk taking and reduce taxpayer losses. FDICIA uses a carrot-and-stick approach to get banks to hold more capital. If they are well capitalized, they receive valuable privileges; if their capital ratio falls, they are subject to more and more onerous regulation. Increased bank capital reduces moral hazard incentives for the bank, because the bank now has more to lose if it fails and so is less likely to take on too much risk.

In addition, encouraging banks to hold more capital reduces potential losses for the FDIC, because increased bank capital is a cushion that makes bank failure less likely. Furthermore, forcing the FDIC to close banks once their net worth is less than 2% (group 5) rather than waiting until net worth has fallen to zero makes it more likely that when a bank is closed, it will still have a positive net worth, thus limiting FDIC losses.

Prompt corrective action, which requires regulators to intervene early when bank capital begins to fall, is a serious attempt to reduce the principal–agent problem for politicians and regulators. With prompt corrective action provisions, regulators no longer have the option of regulatory forbearance, which, as we have seen, can greatly increase moral hazard incentives for banks.

Some critics of FDICIA feel that there are too many loopholes in the bill that still allow regulators too much discretion, thus leaving open the possibility of regulatory forbearance. However, an often overlooked part of the bill increases the accountability of regulators. FDICIA requires a mandatory review of any bank failure that imposes costs on the FDIC. The resulting report must be made available to any member of Congress and to the general public upon request, and the General Accounting Office must do an annual review of these reports. Opening up the actions of the regulators to public scrutiny will make regulatory forbearance less attractive to

them, thereby reducing the principal–agent problem. It will also reduce the incentives of politicians to lean on regulators to relax their regulatory supervision of banks.

Risk-Based Insurance Premiums

Under FDICIA, banks deemed to be taking on greater risk, in the form of lower capital or riskier assets, will be subjected to higher insurance premiums. Risk-based insurance premiums will consequently reduce the moral hazard incentives for banks to take on higher risk. In addition, the fact that risk-based premiums drop as the bank's capital increases encourages the bank to hold more capital, which has the benefits already mentioned.

One problem with risk-based premiums is that the scheme for determining the amount of risk the bank is taking may not be very accurate. For example, it might be hard for regulators to determine when a bank's loans are risky. Some critics have also pointed out that the classification of banks by such measures as the Basel risk-based capital standard solely reflects credit risk and does not take sufficient account of interest-rate risk. The regulatory authorities, however, are encouraged by FDICIA to modify existing risk-based standards to include interest-rate risk and, as we have seen earlier in the chapter, have proposed guidelines to encourage banks to manage interest-rate risk.

Other FDICIA Provisions

FDICIA's requirements that regulators perform bank examinations at least once a year are necessary for monitoring banks' compliance with bank capital requirements and asset restrictions. As the S&L debacle illustrates, frequent supervisory examinations of banks are necessary to keep them from taking on too much risk or committing fraud. Similarly, beefing up the ability of the Federal Reserve to monitor foreign banks might help dissuade international banks from engaging in these undesirable activities.

The stricter and more burdensome reporting requirements for banks have the advantage of providing more information to regulators to help them monitor bank activities. However, these reporting requirements have been criticized by banks, which claim that the requirements make it harder to lend to small businesses.

Other Proposed Changes in Banking Regulations

Regulatory Consolidation. The current bank regulatory system in the United States has banking institutions supervised by four federal agencies: the FDIC, the Office of the Comptroller of the Currency, the Office of Thrift Supervision, and the Federal Reserve. Critics of this system of multiple regulatory agencies with overlapping jurisdictions believe it creates a system that is too complex and too costly because it is rife with duplication. The Clinton administration proposed a consolidation in which the duties of the four regulatory agencies would be given to a new Federal Banking Commission governed by a five-member board with one member from the Treasury, one from the Federal Reserve, and three independent members appointed by the president and confirmed by the Senate. The Federal Reserve strongly opposed this proposal because it believed that it needed to have hands-on supervision of the largest banks through their bank holding companies (as is the case currently) in order to have the information that would enable the Fed to respond sufficiently quickly in a crisis. The Fed also pointed out that a monolithic regulator might be less effective than two or more regulators in providing checks and balances for regulatory supervision. The Clinton administration's proposal was not passed by Congress, but the issue of regulatory consolidation is sure to come up again.

Market-Value Accounting for Capital Requirements. We have seen that the requirement that a bank have substantial equity capital makes the bank less likely to fail. The requirement is also advantageous, because a bank with high equity capital has more to lose if it takes on risky investments and so will have less incentive to hold risky assets. Unfortunately, capital requirements, including new risk-based measures, are calculated on a historical-cost (book value) basis in which the value of an asset is set at its initial purchase price. The problem with historical-cost accounting is that changes in the value of assets and liabilities because of changes in interest rates or default risk are not reflected in the calculation of the firm's equity capital. Yet changes in the market value of assets and liabilities and hence changes in the market value of equity capital are what indicate if a firm is truly insolvent. Furthermore, it is the market value of capital that determines the incentives for a bank to hold risky assets.

Market-value accounting when calculating capital requirements is another reform that receives substantial support. All assets and liabilities could be updated to market value periodically—say, every three months—to determine if a bank's capital is sufficient to meet the minimum requirements. This market-value accounting information would let the deposit insurance agency know quickly when a bank was falling below its capital requirement. The bank could then be closed down before its net worth fell below zero, thus preventing a loss to the deposit insurance agency. The market-value-based capital requirement would also ensure that banks would not be operating with negative capital, thereby preventing the bet-the-bank strategy of taking on excessive risk.

Objections to market-value-based capital requirements center on the difficulty of making accurate and straightforward market-value estimates of capital. Historical-cost accounting has an important advantage in that accounting rules are easier to define and standardize when the value of an asset is simply set at its purchase price. Market-value accounting, by contrast, requires estimates and approximations that are harder to standardize. For example, it might be hard to assess the market value of your friend Joe's car loan, whereas it would be quite easy to value a government bond. In addition, conducting market-value accounting would prove costly to banks because estimation of market values requires the collection of more information about the characteristics of assets and liabilities. Nevertheless, proponents of market-value accounting for capital requirements point out that although market-value accounting involves some estimates and approximations, it would still provide regulators with more accurate assessment of bank equity capital than historical-cost accounting does.

Overall Evaluation

FDICIA appears to be an important step in the right direction, because it increases the incentives for banks to hold capital and decreases their incentives to take on excessive risk. However, more could be done to improve the incentives for banks to limit their risk taking. Yet eliminating deposit insurance and the too-big-to-fail policy altogether may be going too far, because these proposals might make the banking system too prone to a banking panic.

Chapter 12

Nonbank Finance

PREVIEW

Banking is not the only type of financial intermediation you are likely to experience. You might decide to purchase insurance, take out an installment loan from a finance company, or buy a share of stock. In each of these transactions you will be engaged in nonbank finance and will deal with nonbank financial institutions. In our economy, nonbank finance also plays an important role in channeling funds from lender-savers to borrower-spenders. Furthermore, the process of financial innovation we discussed in Chapter 10 has increased the importance of nonbank finance and is blurring the distinction between different financial institutons. This chapter examines in more detail how institutions engaged in nonbank finance operate, how they are regulated, and recent trends in nonbank finance.

Insurance

www.iii.org
The Insurance Information Institute publishes facts and statistics about the insurance industry.

Every day we face the possibility of the occurrence of certain catastrophic events that could lead to large financial losses. A spouse's earnings might disappear due to death or illness; a car accident might result in costly repair bills or payments to an injured party. Because financial losses from crises could be large relative to our financial resources, we protect ourselves against them by purchasing insurance coverage that will pay a sum of money if catastrophic events occur. Life insurance companies sell policies that provide income if a person dies, is incapacitated by illness, or retires. Property and casualty companies specialize in policies that pay for losses incurred as a result of accidents, fire, or theft.

Life Insurance

The first life insurance company in the United States (Presbyterian Ministers' Fund in Philadelphia) was established in 1759 and is still in existence. There are currently about 1,400 life insurance companies, which are organized in two forms: as stock companies or as mutuals. Stock companies are owned by stockholders; mutuals are technically owned by the policyholders. Although over 90% of life insurance companies are organized as stock companies, some of the largest ones are organized as mutuals.

Unlike commercial banks and other depository institutions, life insurance companies have never experienced widespread failures, so the federal government has not seen the need to regulate the industry. Instead, regulation is left to the states in which a company operates. State regulation is directed at sales practices, the provision of

adequate liquid assets to cover losses, and restrictions on the amount of risky assets (such as common stock) that the companies can hold. The regulatory authority is typically a state insurance commissioner.

Because death rates for the population as a whole are predictable with a high degree of certainty, life insurance companies can accurately predict what their payouts to policyholders will be in the future. Consequently, they hold long-term assets that are not particularly liquid—corporate bonds and commercial mortgages as well as some corporate stock.

There are two principal forms of life insurance policies: permanent life insurance (such as whole, universal, and variable life) and temporary insurance (such as term). Permanent life insurance policies have a constant premium throughout the life of the policy. In the early years of the policy, the size of this premium exceeds the amount needed to insure against death because the probability of death is low. Thus the policy builds up a cash value in its early years, but in later years the cash value declines because the constant premium falls below the amount needed to insure against death, the probability of which is now higher. The policyholder can borrow against the cash value of the permanent life policy or can claim it by canceling the policy.

www.federalreserve.gov/releases/Z1/

The Flow of Funds Accounts of the United States reports details about the current state of the insurance industry. Scroll down through the table of contents to find the location of data on insurance companies.

Term insurance, by contrast, has a premium that is matched every year to the amount needed to insure against death during the period of the term (such as one year or five years). As a result, term policies have premiums that rise over time as the probability of death rises (or level premiums with a decline in the amount of death benefits). Term policies have no cash value and thus, in contrast to permanent life policies, provide insurance only, with no savings aspect.

Weak investment returns on permanent life insurance in the 1960s and 1970s led to slow growth of demand for life insurance products. The result was a shrinkage in the size of the life insurance industry relative to other financial intermediaries, with their share of total financial intermediary assets falling from 19.6% at the end of 1960 to 11.5% at the end of 1980. (See Table 1, which shows the relative shares of financial intermediary assets for each of the financial intermediaries discussed in this chapter.)

Beginning in the mid-1970s, life insurance companies began to restructure their business to become managers of assets for pension funds. An important factor behind this restructuring was 1974 legislation that encouraged pension funds to turn fund management over to life insurance companies. Now more than half of the assets managed by life insurance companies are for pension funds and not for life insurance. Insurance companies have also begun to sell investment vehicles for retirement such as **annuities**, arrangements whereby the customer pays an annual premium in exchange for a future stream of annual payments beginning at a set age, say 65, and continuing until death. The result of this new business has been that the market share of life insurance companies as a percentage of total financial intermediary assets has held steady since 1980.

Property and Casualty Insurance

There are on the order of 3,000 property and casualty insurance companies in the United States, the two largest of which are State Farm and Allstate. Property and casualty companies are organized as both stock and mutual companies and are regulated by the states in which they operate.

Although property and casualty insurance companies had a slight increase in their share of total financial intermediary assets from 1960 to 1990 (see Table 1), in recent years they have not fared well, and insurance premiums have skyrocketed. With the high interest rates in the 1970s and 1980s, insurance companies had high

Table 1 Relative Shares of Total Financial Intermediary Assets, 1960–2002 (percent)

	1960	1970	1980	1990	2002
Insurance Companies					
Life insurance	19.6	15.3	11.5	12.5	13.6
Property and casualty	4.4	3.8	4.5	4.9	3.7
Pension Funds					
Private	6.4	8.4	12.5	14.9	14.7
Public (state and local government)	3.3	4.6	4.9	6.7	7.9
Finance Companies	4.7	4.9	5.1	5.6	3.2
Mutual Funds					
Stock and bond	2.9	3.6	1.7	5.9	10.6
Money market	0.0	0.0	1.9	4.6	8.8
Depository Institutions (Banks)					
Commercial banks	38.6	38.5	36.7	30.4	29.8
S&L and mutual savings banks	19.0	19.4	19.6	12.5	5.6
Credit unions	1.1	1.4	1.6	2.0	2.3
Total	100.0	100.0	100.0	100.0	100.0

Source: Federal Reserve Flow of Funds Accounts.

investment income that enabled them to keep insurance rates low. Since then, however, investment income has fallen with the decline in interest rates, while the growth in lawsuits involving property and casualty insurance and the explosion in amounts awarded in such cases have produced substantial losses for companies.

To return to profitability, insurance companies have raised their rates dramatically—sometimes doubling or even tripling premiums—and have refused to provide coverage for some people. They have also campaigned actively for limits on insurance payouts, particularly for medical malpractice. In the search for profits, insurance companies are also branching out into uncharted territory by insuring the payment of interest on municipal and corporate bonds and on mortgage-backed securities. One worry is that the insurance companies may be taking on excessive risk in order to boost their profits. One result of the concern about the health of the property and casualty insurance industry is that insurance regulators have proposed new rules that would impose risk-based capital requirements on these companies based on the riskiness of their assets and operations.

The investment policies of these companies are affected by two basic facts. First, because they are subject to federal income taxes, the largest share of their assets is held in tax-exempt municipal bonds. Second, because property losses are more uncertain than the death rate in a population, these insurers are less able to predict how much they will have to pay policyholders than life insurance companies are. Natural or

unnatural disasters such as the Los Angeles earthquake in 1994 and Hurricane Floyd in 1999, which devastated parts of the East Coast, and the September 11, 2001 destruction of the World Trade Center, exposed the property and casualty insurance companies to billions of dollars of losses. Therefore, property and casualty insurance companies hold more liquid assets than life insurance companies; municipal bonds and U.S. government securities amount to over half their assets, and most of the remainder is held in corporate bonds and corporate stock.

Property and casualty insurance companies will insure against losses from almost any type of event, including fire, theft, negligence, malpractice, earthquakes, and automobile accidents. If a possible loss being insured is too large for any one firm, several firms may join together to write a policy in order to share the risk. Insurance companies may also reduce their risk exposure by obtaining **reinsurance**. Reinsurance allocates a portion of the risk to another company in exchange for a portion of the premium and is particularly important for small insurance companies. You can think of reinsurance as insurance for the insurance company. The most famous risk-sharing operation is Lloyd's of London, an association in which different insurance companies can underwrite a fraction of an insurance policy. Lloyd's of London has claimed that it will insure against any contingency—for a price.

The Competitive Threat from the Banking Industry

Until recently, banks have been restricted in their ability to sell life insurance products. This has been changing rapidly, however. Over two-thirds of the states allow banks to sell life insurance in one form or another. In recent years, the bank regulatory authorities, particularly the Office of the Comptroller of the Currency (OCC), have also encouraged banks to enter the insurance field because getting into insurance would help diversify banks' business, thereby improving their economic health and making bank failures less likely. For example, in 1990, the OCC ruled that selling annuities was a form of investment that was incidental to the banking business and so was a permissible banking activity. As a result, the banks' share of the annuities market has surpassed 20%. Currently, more than 40% of banks sell insurance products, and the number is expected to grow in the future.

Insurance companies and their agents reacted to this competitive threat with both lawsuits and lobbying actions to block banks from entering the insurance business. Their efforts were set back by several Supreme Court rulings that favored the banks. Particularly important was a ruling in favor of Barnett Bank in March 1996, which held that state laws to prevent banks from selling insurance can be superseded by federal rulings from banking regulators that allow banks to sell insurance. The decision gave banks a green light to further their insurance activities, and with the passage of the Gramm-Leach-Bliley Act of 1999, banking institutions will further engage in the insurance business, thus blurring the distinction between insurance companies and banks.

Application

Insurance Management

Insurance, like banking, is in the financial intermediation business of transforming one type of asset into another for the public. Insurance providers use the premiums paid on policies to invest in assets such as bonds, stocks, mortgages, and other loans; the earnings from these assets are then used to pay out claims on the policies. In effect, insurers transform assets such as

bonds, stocks, and loans into insurance policies that provide a set of services (for example, claim adjustments, savings plans, friendly insurance agents). If the insurer's production process of asset transformation efficiently provides its customers with adequate insurance services at low cost and if it can earn high returns on its investments, it will make profits; if not, it will suffer losses.

In Chapter 9 the economic concepts of adverse selection and moral hazard allowed us to understand principles of bank management related to managing credit risk; many of these same principles also apply to the lending activities of insurers. Here again we apply the adverse selection and moral hazard concepts to explain many management practices specific to insurance.

In the case of an insurance policy, moral hazard arises when the existence of insurance encourages the insured party to take risks that increase the likelihood of an insurance payoff. For example, a person covered by burglary insurance might not take as many precautions to prevent a burglary because the insurance company will reimburse most of the losses if a theft occurs. Adverse selection holds that the people most likely to receive large insurance payoffs are the ones who will want to purchase insurance the most. For example, a person suffering from a terminal disease would want to take out the biggest life and medical insurance policies possible, thereby exposing the insurance company to potentially large losses. Both adverse selection and moral hazard can result in large losses to insurance companies, because they lead to higher payouts on insurance claims. Lowering adverse selection and moral hazard to reduce these payouts is therefore an extremely important goal for insurance companies, and this goal explains the insurance practices we will discuss here.

Screening

To reduce adverse selection, insurance providers try to screen out good insurance risks from poor ones. Effective information collection procedures are therefore an important principle of insurance management.

When you apply for auto insurance, the first thing your insurance agent does is ask you questions about your driving record (number of speeding tickets and accidents), the type of car you are insuring, and certain personal matters (age, marital status). If you are applying for life insurance, you go through a similar grilling, but you are asked even more personal questions about such things as your health, smoking habits, and drug and alcohol use. The life insurer even orders a medical evaluation (usually done by an independent company) that involves taking blood and urine samples. Just as a bank calculates a credit score to evaluate a potential borrower, the insurers use the information you provide to allocate you to a risk class—a statistical estimate of how likely you are to have an insurance claim. Based on this information, the insurer can decide whether to accept you for the insurance or to turn you down because you pose too high a risk and thus would be an unprofitable customer.

Risk-Based Premiums

Charging insurance premiums on the basis of how much risk a policyholder poses for the insurance provider is a time-honored principle of insurance management. Adverse selection explains why this principle is so important to insurance company profitability.

To understand why an insurance provider finds it necessary to have risk-based premiums, let's examine an example of risk-based insurance premiums that at first glance seems unfair. Harry and Sally, both college students with no accidents or speeding tickets, apply for auto insurance. Normally, Harry will be charged a much higher premium than Sally. Insurance providers do this because young males have a much higher accident rate than young females. Suppose, though, that one insurer did not base its premiums on a risk classification but rather just charged a premium based on the average combined risk for males and females. Then Sally would be charged too much and Harry too little. Sally could go to another insurer and get a lower rate, while Harry would sign up for the insurance. Because Harry's premium isn't high enough to cover the accidents he is likely to have, on average the insurer would lose money on Harry. Only with a premium based on a risk classification, so that Harry is charged more, can the insurance provider make a profit.[1]

Restrictive Provisions

Restrictive provisions in policies are an insurance management tool for reducing moral hazard. Such provisions discourage policyholders from engaging in risky activities that make an insurance claim more likely. For example, life insurers have provisions in their policies that eliminate death benefits if the insured person commits suicide within the first two years that the policy is in effect. Restrictive provisions may also require certain behavior on the part of the insured. A company renting motor scooters may be required to provide helmets for renters in order to be covered for any liability associated with the rental. The role of restrictive provisions is not unlike that of restrictive covenants on debt contracts described in Chapter 8: Both serve to reduce moral hazard by ruling out undesirable behavior.

Prevention of Fraud

Insurance providers also face moral hazard because an insured person has an incentive to lie to the insurer and seek a claim even if the claim is not valid. For example, a person who has not complied with the restrictive provisions of an insurance contract may still submit a claim. Even worse, a person may file claims for events that did not actually occur. Thus an important management principle for insurance providers is conducting investigations to prevent fraud so that only policyholders with valid claims receive compensation.

Cancellation of Insurance

Being prepared to cancel policies is another insurance management tool. Insurers can discourage moral hazard by threatening to cancel a policy when the insured person engages in activities that make a claim more likely. If your auto insurance company makes it clear that coverage will be canceled if a driver gets too many speeding tickets, you will be less likely to speed.

Deductibles

The **deductible** is the fixed amount by which the insured's loss is reduced when a claim is paid off. A $250 deductible on an auto policy, for example,

[1] Note that the example here is in fact the lemons problem described in Chapter 8.

means that if you suffer a loss of $1,000 because of an accident, the insurer will pay you only $750. Deductibles are an additional management tool that helps insurance providers reduce moral hazard. With a deductible, you experience a loss along with the insurer when you make a claim. Because you also stand to lose when you have an accident, you have an incentive to drive more carefully. A deductible thus makes a policyholder act more in line with what is profitable for the insurer; moral hazard has been reduced. And because moral hazard has been reduced, the insurance provider can lower the premium by more than enough to compensate the policyholder for the existence of the deductible. Another function of the deductible is to eliminate the administrative costs of handling small claims by forcing the insured to bear these losses.

Coinsurance

When a policyholder shares a percentage of the losses along with the insurer, their arrangement is called **coinsurance**. For example, some medical insurance plans provide coverage for 80% of medical bills, and the insured person pays 20% after a certain deductible has been met. Coinsurance works to reduce moral hazard in exactly the same way that a deductible does. A policyholder who suffers a loss along with the insurer has less incentive to take actions, such as going to the doctor unnecessarily, that involve higher claims. Coinsurance is thus another useful management tool for insurance providers.

Limits on the Amount of Insurance

Another important principle of insurance management is that there should be limits on the amount of insurance provided, even though a customer is willing to pay for more coverage. The higher the insurance coverage, the more the insured person can gain from risky activities that make an insurance payoff more likely and hence the greater the moral hazard. For example, if Zelda's car were insured for more than its true value, she might not take proper precautions to prevent its theft, such as making sure that the key is always removed or putting in an alarm system. If it were stolen, she comes out ahead because the excessive insurance payment would allow her to buy an even better car. By contrast, when the insurance payments are lower than the value of her car, she will suffer a loss if it is stolen and will thus take precautions to prevent this from happening. Insurance providers must always make sure that their coverage is not so high that moral hazard leads to large losses.

Summary

Effective insurance management requires several practices: information collection and screening of potential policyholders, risk-based premiums, restrictive provisions, prevention of fraud, cancellation of insurance, deductibles, coinsurance, and limits on the amount of insurance. All of these practices reduce moral hazard and adverse selection by making it harder for policyholders to benefit from engaging in activities that increase the amount and likelihood of claims. With smaller benefits available, the poor insurance risks (those who are more likely to engage in the activities in the first place) see less benefit from the insurance and are thus less likely to seek it out.

Pension Funds

In performing the financial intermediation function of asset transformation, pension funds provide the public with another kind of protection: income payments on retirement. Employers, unions, or private individuals can set up pension plans, which acquire funds through contributions paid in by the plan's participants. As we can see in Table 1, pension plans both public and private have grown in importance, with their share of total financial intermediary assets rising from 10% at the end of 1960 to 22.6% at the end of 2002. Federal tax policy has been a major factor behind the rapid growth of pension funds because employer contributions to employee pension plans are tax-deductible. Furthermore, tax policy has also encouraged employee contributions to pension funds by making them tax-deductible as well and enabling self-employed individuals to open up their own tax-sheltered pension plans, Keogh plans, and individual retirement accounts (IRAs).

Because the benefits paid out of the pension fund each year are highly predictable, pension funds invest in long-term securities, with the bulk of their asset holdings in bonds, stocks, and long-term mortgages. The key management issues for pension funds revolve around asset management: Pension fund managers try to hold assets with high expected returns and lower risk through diversification. They also use techniques we discussed in Chapter 9 to manage credit and interest-rate risk. The investment strategies of pension plans have changed radically over time. In the aftermath of World War II, most pension fund assets were held in government bonds, with less than 1% held in stock. However, the strong performance of stocks in the 1950s and 1960s afforded pension plans higher returns, causing them to shift their portfolios into stocks, currently on the order of two-thirds of their assets. As a result, pension plans now have a much stronger presence in the stock market: In the early 1950s, they held on the order of 1% of corporate stock outstanding; currently they hold on the order of 25%. Pension funds are now the dominant players in the stock market.

Although the purpose of all pension plans is the same, they can differ in a number of attributes. First is the method by which payments are made: If the benefits are determined by the contributions into the plan and their earnings, the pension is a **defined-contribution plan**; if future income payments (benefits) are set in advance, the pension is a **defined-benefit plan**. In the case of a defined-benefit plan, a further attribute is related to how the plan is funded. A defined-benefit plan is **fully funded** if the contributions into the plan and their earnings over the years are sufficient to pay out the defined benefits when they come due. If the contributions and earnings are not sufficient, the plan is **underfunded**. For example, if Jane Brown contributes $100 per year into her pension plan and the interest rate is 10%, after ten years the contributions and their earnings would be worth $1,753.[2] If the defined benefit on her

[2]The $100 contributed in year 1 would become worth $100 × (1 + 0.10)^{10} = $259.37 at the end of ten years; the $100 contributed in year 2 would become worth $100 × (1 + 0.10)^{9} = $235.79; and so on until the $100 contributed in year 10 would become worth $100 × (1 + 0.10) = $110. Adding these together, we get the total value of these contributions and their earnings at the end of ten years:

$$\$259.37 + \$235.79 + \$214.36 + \$194.87 + \$177.16 + \$161.05 + \$146.41 + \$133.10 + \$121.00 + \$110.00 = \$1,753.11$$

pension plan pays her $1,753 or less after ten years, the plan is fully funded because her contributions and earnings will fully pay for this payment. But if the defined benefit is $2,000, the plan is underfunded, because her contributions and earnings do not cover this amount.

A second characteristic of pension plans is their *vesting*, the length of time that a person must be enrolled in the pension plan (by being a member of a union or an employee of a company) before being entitled to receive benefits. Typically, firms require that an employee work five years for the company before being vested and qualifying to receive pension benefits; if the employee leaves the firm before the five years are up, either by quitting or being fired, all rights to benefits are lost.

Private Pension Plans

Private pension plans are administered by a bank, a life insurance company, or a pension fund manager. In employer-sponsored pension plans, contributions are usually shared between employer and employee. Many companies' pension plans are underfunded because they plan to meet their pension obligations out of current earnings when the benefits come due. As long as companies have sufficient earnings, underfunding creates no problems, but if not, they may not be able to meet their pension obligations. Because of potential problems caused by corporate underfunding, mismanagement, fraudulent practices, and other abuses of private pension funds (Teamsters pension funds are notorious in this regard), Congress enacted the Employee Retirement Income Security Act (ERISA) in 1974. This act established minimum standards for the reporting and disclosure of information, set rules for vesting and the degree of underfunding, placed restrictions on investment practices, and assigned the responsibility of regulatory oversight to the Department of Labor.

www.pbgc.gov/

The web site for the Pension Benefit Guarantee Corporation contains information about pensions and the insurance that it provides.

ERISA also created the Pension Benefit Guarantee Corporation (called "Penny Benny"), which performs a role similar to that of the FDIC. It insures pension benefits up to a limit (currently over $40,000 per year per person) if a company with an underfunded pension plan goes bankrupt or is unable to meet its pension obligations for other reasons. Penny Benny charges pension plans premiums to pay for this insurance, and it can also borrow funds up to $100 million from the U.S. Treasury. Unfortunately, the problem of pension plan underfunding has been growing worse in recent years. In 1993, the secretary of labor indicated that underfunding had reached levels in excess of $45 billion, with one company's pension plan alone, that of General Motors, underfunded to the tune of $11.8 billion. As a result, Penny Benny, which insures the pensions of one of every three workers, may have to foot the bill if companies with large underfunded pensions go broke.

Public Pension Plans

The most important public pension plan is Social Security (Old Age and Survivors' Insurance Fund), which covers virtually all individuals employed in the private sector. Funds are obtained from workers through Federal Insurance Contribution Act (FICA) deductions from their paychecks and from employers through payroll taxes. Social Security benefits include retirement income, Medicare payments, and aid to the disabled.

www.ssa.gov/

The web site for the Social Security Administration contains information on your benefits available from social security.

When Social Security was established in 1935, the federal government intended to operate it like a private pension fund. However, unlike a private pension plan, benefits are typically paid out from current contributions, not tied closely to a participant's past contributions. This "pay as you go" system at one point led to a massive underfunding, estimated at over $1 trillion.

The problems of the Social Security system could become worse in the future because of the growth in the number of retired people relative to the working

Box 1

Should Social Security Be Privatized?

In recent years, public confidence in the Social Security system has reached a new low. Some surveys suggest that young people have more confidence in the existence of flying saucers than they do in the government's promise to pay them their Social Security benefits. Without some overhaul of the system, Social Security will not be able to meet its future obligations. The government has set up advisory commissions and has been holding hearings to address this problem.

Currently, the assets of the Social Security system, which reside in a trust fund, are all invested in U.S. Treasury securities. Because stocks and corporate bonds have higher returns than Treasury securities, many proposals to save the Social Security system suggest investing part of the trust fund in corporate securities and thus partially privatizing the system.

Suggestions for privatization take three basic forms:

1. *Government investment of trust fund assets in corporate securities.* This plan has the advantage of possibly improving the trust fund's overall return, while minimizing transactions costs because it exploits the economies of scale of the trust fund. Critics warn that government ownership of private assets could lead to increased government intervention in the private sector.

2. *Shift of trust fund assets to individual accounts that can be invested in private assets.* This option has the advantage of possibly increasing the return on investments and does not involve the government in the ownership of private assets. However, critics warn that it might expose individuals to greater risk and to transaction costs on individual accounts that might be very high because of the small size of many of these accounts.

3. *Individual accounts in addition to those in the trust fund.* This option has advantages and disadvantages similar to those of option 2 and may provide more funds to individuals at retirement. However, some increase in taxes would be required to fund these accounts.

Whether some privatization of the Social Security system occurs is an open question. In the short term, Social Security reform is likely to involve an increase in taxes, a reduction in benefits, or both. For example, the age at which benefits begin is already scheduled to increase from 65 to 67, and might be increased further to 70. It is also likely that the cap on wages subject to the Social Security tax will be raised further, thereby increasing taxes paid into the system.

population. Congress has been grappling with the problems of the Social Security system for years, but the prospect of a huge bulge in new retirees when the 77 billion baby boomers born between 1946 and 1964 start to retire in 2011 has resulted in calls for radical surgery on Social Security (see Box 1).

State and local governments and the federal government, like private employers, have also set up pension plans for their employees. These plans are almost identical in operation to private pension plans and hold similar assets. Underfunding of the plans is also prevalent, and some investors in municipal bonds worry that it may lead to future difficulties in the ability of state and local governments to meet their debt obligations.

Finance Companies

Finance companies acquire funds by issuing commercial paper or stocks and bonds or borrowing from banks, and they use the proceeds to make loans (often for small amounts) that are particularly well suited to consumer and business needs. The finan-

www.federalreserve.gov/Releases/G20/current/default.htm
Federal reserve information about financial companies.

cial intermediation process of finance companies can be described by saying that they borrow in large amounts but often lend in small amounts—a process quite different from that of banking institutions, which collect deposits in small amounts and then often make large loans.

A key feature of finance companies is that although they lend to many of the same customers that borrow from banks, they are virtually unregulated compared to commercial banks and thrift institutions. States regulate the maximum amount they can loan to individual consumers and the terms of the debt contract, but there are no restrictions on branching, the assets they hold, or how they raise their funds. The lack of restrictions enables finance companies to tailor their loans to customer needs better than banking institutions can.

There are three types of finance companies: sales, consumer, and business.

1. *Sales finance companies* are owned by a particular retailing or manufacturing company and make loans to consumers to purchase items from that company. Sears, Roebuck Acceptance Corporation, for example, finances consumer purchases of all goods and services at Sears stores, and General Motors Acceptance Corporation finances purchases of GM cars. Sales finance companies compete directly with banks for consumer loans and are used by consumers because loans can frequently be obtained faster and more conveniently at the location where an item is purchased.
2. *Consumer finance companies* make loans to consumers to buy particular items such as furniture or home appliances, to make home improvements, or to help refinance small debts. Consumer finance companies are separate corporations (like Household Finance Corporation) or are owned by banks (Citigroup owns Person-to-Person Finance Company, which operates offices nationwide). Typically, these companies make loans to consumers who cannot obtain credit from other sources and charge higher interest rates.
3. *Business finance companies* provide specialized forms of credit to businesses by making loans and purchasing accounts receivable (bills owed to the firm) at a discount; this provision of credit is called *factoring*. For example, a dressmaking firm might have outstanding bills (accounts receivable) of $100,000 owed by the retail stores that have bought its dresses. If this firm needs cash to buy 100 new sewing machines, it can sell its accounts receivable for, say, $90,000 to a finance company, which is now entitled to collect the $100,000 owed to the firm. Besides factoring, business finance companies also specialize in leasing equipment (such as railroad cars, jet planes, and computers), which they purchase and then lease to businesses for a set number of years.

Mutual Funds

www.ici.org/facts_figures/factbook_toc.html
The Mutual Fund Fact Book published by Investment Company Institute includes information about the mutual funds industry's history, regulation, taxation, and shareholders.

Mutual funds are financial intermediaries that pool the resources of many small investors by selling them shares and using the proceeds to buy securities. Through the asset transformation process of issuing shares in small denominations and buying large blocks of securities, mutual funds can take advantage of volume discounts on brokerage commissions and purchase diversified holdings (portfolios) of securities. Mutual funds allow the small investor to obtain the benefits of lower transaction costs in purchasing securities and to take advantage of the reduction of risk by diversifying the portfolio of securities held. Many mutual funds are run by brokerage firms, but others are run by banks or independent investment advisers such as Fidelity or Vanguard.

Mutual funds have seen a large increase in their market share since 1980 (see Table 1), due primarily to the then-booming stock market. Another source of growth has been mutual funds that specialize in debt instruments, which first appeared in the 1970s. Before 1970, mutual funds invested almost solely in common stocks. Funds that purchase common stocks may specialize even further and invest solely in foreign securities or in specialized industries, such as energy or high technology. Funds that purchase debt instruments may specialize further in corporate, U.S. government, or tax-exempt municipal bonds or in long-term or short-term securities.

Mutual funds are primarily held by households (around 80%) with the rest held by other financial institutions and nonfinancial businesses. Mutual funds have become increasingly important in household savings. In 1980, only 6% of households held mutual fund shares; this number has risen to around 50% in recent years. The age group with the greatest participation in mutual fund ownership includes individuals between 50 and 70, which makes sense because they are the most interested in saving for retirement. Interestingly, Generation X (18–30) is the second most active age group in mutual fund ownership, suggesting that they have a greater tolerance for investment risk than those who are somewhat older. Generation X is also leading the way in Internet access to mutual funds (see Box 2).

The growing importance of investors in mutual funds and pension funds, so-called *institutional investors*, has resulted in their controlling over 50% of the outstanding stock in the United States. Thus, institutional investors are the predominant players in the stock markets, with over 70% of the total daily volume in the stock market due to their trading. Increased ownership of stocks has also meant that institutional investors have more clout with corporate boards, often forcing changes in leadership or in corporate policies.

Mutual funds are structured in two ways. The more common structure is an **open-end fund**, from which shares can be redeemed at any time at a price that is tied

Box 2: E-Finance

Mutual Funds and the Internet

The Investment Company Institute estimates that as of 2000, 68% of households owning mutual funds use the Internet, and nearly half of those online shareholders visit fund-related web sites. The Internet increases the attractiveness of mutual funds because it enables shareholders to review performance information and share prices and personal account information.

Of all U.S. households that conducted mutual funds transactions between April 1999 and March 2000, 18% bought or sold fund shares online. The median number of funds transactions conducted over the Internet during the 12-month period was four, while the average number was eight, indicating that a high volume of online transactions were conducted by a small number of shareholders.

Online shareholders were typically younger, had greater household income, and were better educated than those not using the Internet. The median online shareholder was 42 years old, had a household income of $100,900, and was college-educated. The median shareholder not using the Internet was 51 years old, had a household income of $41,000, and did not have a college degree.

The use of the Internet to track and trade mutual funds is rapidly increasing. The number of shareholders who visited web sites offering fund shares nearly doubled between April 1999 and March 2000.

to the asset value of the fund. Mutual funds also can be structured as a **closed-end fund**, in which a fixed number of nonredeemable shares are sold at an initial offering and are then traded like a common stock. The market price of these shares fluctuates with the value of the assets held by the fund. In contrast to the open-end fund, however, the price of the shares may be above or below the value of the assets held by the fund, depending on factors such as the liquidity of the shares or the quality of the management. The greater popularity of the open-end funds is explained by the greater liquidity of their redeemable shares relative to the nonredeemable shares of closed-end funds.

Originally, shares of most open-end mutual funds were sold by salespeople (usually brokers) who were paid a commission. Since this commission is paid at the time of purchase and is immediately subtracted from the redemption value of the shares, these funds are called **load funds**. Most mutual funds are currently **no-load funds**; they are sold directly to the public with no sales commissions. In both types of funds, the managers earn their living from management fees paid by the shareholders. These fees amount to approximately 0.5% of the asset value of the fund per year.

Mutual funds are regulated by the Securities and Exchange Commission, which was given the ability to exercise almost complete control over investment companies in the Investment Company Act of 1940. Regulations require periodic disclosure of information on these funds to the public and restrictions on the methods of soliciting business.

Money Market Mutual Funds

An important addition to the family of mutual funds resulting from the financial innovation process described in earlier chapters is the *money market mutual fund*. Recall that this type of mutual fund invests in short-term debt (money market) instruments of very high quality, such as Treasury bills, commercial paper, and bank certificates of deposit. There is some fluctuation in the market value of these securities, but because their maturity is typically less than six months, the change in the market value is small enough that these funds allow their shares to be redeemed at a fixed value. (Changes in the market value of the securities are figured into the interest paid out by the fund.) Because these shares can be redeemed at a fixed value, the funds allow shareholders to redeem shares by writing checks on the fund's account at a commercial bank. In this way, shares in money market mutual funds effectively function as checkable deposits that earn market interest rates on short-term debt securities.

In 1977, the assets in money market mutual funds were less than $4 billion; by 1980, they had climbed to over $50 billion and now stand at $2.1 trillion, with a share of financial intermediary assets that has grown to nearly 9% (see Table 1). Currently, money market mutual funds account for around one-quarter of the asset value of all mutual funds.

Hedge Funds

Hedge funds are a special type of mutual fund, with estimated assets of more than $500 billion. Hedge funds have received considerable attention recently due to the shock to the financial system resulting from the near collapse of Long-Term Capital Management, once one of the most important hedge funds (Box 3). Well-known hedge funds include Moore Capital Management and the Quantum group of funds associated with George Soros. Like mutual funds, hedge funds accumulate money from many people and invest on their behalf, but several features distinguish them from traditional mutual funds. Hedge funds have a minimum investment requirement

Box 3

The Long-Term Capital Management Debacle

Long-Term Capital Management was a hedge fund with a star cast of managers, including 25 PhDs, two Nobel Prize winners in economics (Myron Scholes and Robert Merton), a former vice-chairman of the Federal Reserve System (David Mullins), and one of Wall Street's most successful bond traders (John Meriwether). It made headlines in September 1998 because its near collapse roiled markets and required a private rescue plan organized by the Federal Reserve Bank of New York.

The experience of Long-Term Capital demonstrates that hedge funds are far from risk-free, despite their use of market-neutral strategies. Long-Term Capital got into difficulties when it thought that the spread between prices on long-term Treasury bonds and long-term corporate bonds was too high, and bet that this "anomaly" would disappear and the spread would narrow. In the wake of the collapse of the Russian financial system in August 1998, investors increased their assessment of the riskiness of corporate securities and as we saw in Chapter 6, the spread between corporates and Treasuries rose rather than narrowed as Long-Term Capital had predicted. The result was that Long-Term Capital took big losses on its positions, eating up much of its equity position.

By mid-September, Long-Term Capital was unable to raise sufficient funds to meet the demands of its creditors. With Long-Term Capital facing the potential need to liquidate its portfolio of $80 billion in securities and more than $1 trillion of notional value in derivatives (discussed in Chapter 13), the Federal Reserve Bank of New York stepped in on September 23 and organized a rescue plan with its creditors. The Fed's rationale for stepping in was that a sudden liquidation of Long-Term Capital's portfolio would create unacceptable systemic risk. Tens of billions of dollars of illiquid securities would be dumped on an already jittery market, causing potentially huge losses to numerous lenders and other institutions. The rescue plan required creditors, banks and investment banks, to supply an additional $3.6 billion of funds to Long-Term Capital in exchange for much tighter management control of funds and a 90% reduction in the managers' equity stake. In the middle of 1999, John Meriwether began to wind down the funds operations.

Even though no public funds were expended, the Fed's involvement in organizing the rescue of Long-Term Capital was highly controversial. Some critics argue that the Fed intervention increased moral hazard by weakening discipline imposed by the market on fund managers because future Fed interventions of this type would be expected. Others think that the Fed's action was necessary to prevent a major shock to the financial system that could have provoked a financial crisis. The debate on whether the Fed should have intervened is likely to go on for some time.

between $100,000 and $20 million, with the typical minimum investment being $1 million. Long-Term Capital Management required a $10 million minimum investment. Federal law limits hedge funds to have no more than 99 investors (limited partners) who must have steady annual incomes of $200,000 or more or a net worth of $1 million, excluding their homes. These restrictions are aimed at allowing hedge funds to be largely unregulated, on the theory that the rich can look out for themselves. Many of the 4,000 hedge funds are located offshore to escape regulatory restrictions.

Hedge funds also differ from traditional mutual funds in that they usually require that investors commit their money for long periods of time, often several years. The purpose of this requirement is to give managers breathing room to pursue long-run

strategies. Hedge funds also typically charge large fees to investors. The typical fund charges a 1% annual fee on the assets it manages plus 20% of profits, and some charge significantly more. Long-Term Capital, for example, charged investors a 2% asset management fee and took 25% of the profits.

The term *hedge fund* is highly misleading, because the word "hedge" typically indicates strategies to avoid risk. As the near failure of Long-Term Capital illustrates, despite their name, these funds can and do take big risks. Many hedge funds engage in what are called "market-neutral" strategies where they buy a security, such as a bond, that seems cheap and sell an equivalent amount of a similar security that appears to be overvalued. If interest rates as a whole go up or down, the fund is hedged, because the decline in value of one security is matched by the rise in value of the other. However, the fund is speculating on whether the spread between the price on the two securities moves in the direction predicted by the fund managers. If the fund bets wrong, it can lose a lot of money, particularly if it has leveraged up its positions; that is, has borrowed heavily against these positions so that its equity stake is small relative to the size of its portfolio. When Long-Term Capital was rescued, it had a leverage ratio of 50 to 1; that is, its assets were fifty times larger than its equity, and even before it got into trouble, it was leveraged 20 to 1.

In the wake of the near collapse of Long-Term Capital, many U.S. politicians have called for regulation of these funds. However, because many of these funds operate offshore in places like the Cayman Islands and are outside of U.S. jurisdiction, they would be extremely hard to regulate. What U.S. regulators can do is ensure that U.S. banks and investment banks have clear guidelines on the amount of lending they can provide to hedge funds and require that these institutions get the appropriate amount of disclosure from hedge funds as to the riskiness of their positions.

Government Financial Intermediation

The government has become involved in financial intermediation in two basic ways: first, by setting up federal credit agencies that directly engage in financial intermediation and, second, by supplying government guarantees for private loans.

Federal Credit Agencies

To promote residential housing, the government has created three government agencies that provide funds to the mortgage market by selling bonds and using the proceeds to buy mortgages: the Government National Mortgage Association (GNMA, or "Ginnie Mae"), the Federal National Mortgage Association (FNMA, or "Fannie Mae"), and the Federal Home Loan Mortgage Corporation (FHLMC, or "Freddie Mac"). Except for Ginnie Mae, which is a federal agency and is thus an entity of the U.S. government, the other agencies are federally sponsored agencies (FSEs) that function as private corporations with close ties to the government. As a result, the debt of sponsored agencies is not explicitly backed by the U.S. government, as is the case for Treasury bonds. As a practical matter, however, it is unlikely that the federal government would allow a default on the debt of these sponsored agencies.

Agriculture is another area in which financial intermediation by government agencies plays an important role. The Farm Credit System (composed of Banks for Cooperatives, Farm Credit banks, and various farm credit associations) issues securities and then uses the proceeds to make loans to farmers.

Box 4

Are Fannie Mae and Freddie Mac Getting Too Big for Their Britches?

With the growth of Fannie Mae and Freddie Mac to immense proportions, there are rising concerns that these federally sponsored agencies could threaten the health of the financial system. Fannie Mae and Freddie Mac either own or insure the risk on close to half of U.S. residential mortgages (amounting to $2 trillion). In fact, their publicly issued debt is well over half that issued by the federal government. A failure of either of these institutions would therefore pose a grave shock to the financial system. Although the federal government would be unlikely to stand by and let them fail, in such a case, the taxpayer would face substantial costs, as in the S&L crisis.

Concerns about the safety and soundness of these institutions arise because they have much smaller capital-to-asset ratios than banks. Critics also charge that Fannie Mae and Freddie Mac have become so large that they wield too much political influence. In addition, these federally sponsored agencies have conflicts of interest, because they have to serve two masters: as publicly traded corporations, they are supposed to maximize profits for the shareholders, but as government agencies, they are supposed to work in the interests of the public. These concerns have led to calls for reform of these agencies, with many advocating full privatization as was done voluntarily by the Student Loan Market Association ("Sallie Mae") in the mid-1990s.

In recent years, government financial intermediaries experienced financial difficulties. The Farm Credit System is one example. The rising tide of farm bankruptcies meant losses in the billions of dollars for the Farm Credit System, and as a result it required a bailout from the federal government in 1987. The agency was authorized to borrow up to $4 billion to be repaid over a 15-year period and received over $1 billion in assistance. There is growing concern in Washington about the health of the federal credit agencies. To head off government bailouts like that for the Farm Credit System, the Federal Credit Reform Act of 1990 set new rules that require such agencies to increase their capital to provide a greater cushion to offset any potential losses. However, there have been growing concerns about Fannie Mae and Freddie Mac (Box 4).

Securities Market Operations

The smooth functioning of securities markets, in which bonds and stocks are traded, involves several financial institutions, including securities brokers and dealers, investment banks, and organized exchanges. None of these institutions were included in our list of financial intermediaries in Chapter 2, because they do not perform the intermediation function of acquiring funds by issuing liabilities and then using the funds to acquire financial assets. Nonetheless, they are important in the process of channeling funds from savers to spenders and can be thought of as "financial facilitators."

First, however, we must recall the distinction between primary and secondary securities markets discussed in Chapter 2. In a primary market, new issues of a security are sold to buyers by the corporation or government agency borrowing the funds. A secondary market then trades the securities that have been sold in the primary mar-

ket (and so are secondhand). *Investment banks* assist in the initial sale of securities in the primary market; *securities brokers* and *dealers* assist in the trading of securities in the secondary markets, some of which are organized into exchanges.

Investment Banking

Mutual funds sold directly to the public on which no sales commissions are charged.

When a corporation wishes to borrow (raise) funds, it normally hires the services of an investment banker to help sell its securities. (Despite its name, an investment banker is not a banker in the ordinary sense; that is, it is not engaged in financial intermediation that takes in deposits and then lends them out.) Some of the well-known U.S. investment banking firms are Merrill Lynch, Salomon Smith Barney, Morgan Stanley Dean Witter, Goldman Sachs, Lehman Brothers, and Credit Suisse First Boston, which have been very successful not only in the United States but outside it as well.

Investment bankers assist in the sale of securities as follows. First, they advise the corporation on whether it should issue bonds or stock. If they suggest that the corporation issue bonds, investment bankers give advice on what the maturity and interest payments on the bonds should be. If they suggest that the corporation should sell stock, they give advice on what the price should be. This is fairly easy to do if the firm has prior issues currently selling in the market, called **seasoned issues**. However, when a firm issues stock for the first time in what is called an **initial public offering (IPO)**, it is more difficult to determine what the correct price should be. All the skills and expertise of the investment banking firm then need to be brought to bear to determine the most appropriate price. IPOs have become very important in the U.S. economy, because they are a major source of financing for Internet companies, which became all the rage on Wall Street in the late 1990s. Not only have IPOs helped these companies to acquire capital to substantially expand their operations, but they have also made the original owners of these firms very rich. Many a nerdy 20- to 30-year-old became an instant millionaire when his stake in his Internet company was given a high valuation after the initial public offering of shares in the company. However, with the bursting of the tech bubble in 2000, many of them lost much of their wealth when the value of their shares came down to earth.

www.ipo.com

The site reports initial public offering news and information and includes advanced search tools for IPO offerings, venture capital research reports, and so on.

When the corporation decides which kind of financial instrument it will issue, it offers them to **underwriters**—investment bankers that guarantee the corporation a price on the securities and then sell them to the public. If the issue is small, only one investment banking firm underwrites it (usually the original investment banking firm hired to provide advice on the issue). If the issue is large, several investment banking firms form a syndicate to underwrite the issue jointly, thus limiting the risk that any one investment bank must take. The underwriters sell the securities to the general public by contacting potential buyers, such as banks and insurance companies, directly and by placing advertisements in newspapers like the *Wall Street Journal* (see the "Following the Financial News" box).

The activities of investment bankers and the operation of primary markets are heavily regulated by the Securities and Exchange Commission (SEC), which was created by the Securities and Exchange Acts of 1933 and 1934 to ensure that adequate information reaches prospective investors. Issuers of new securities to the general public (for amounts greater than $1.5 million in a year with a maturity longer than 270 days) must file a registration statement with the SEC and must provide to potential investors a prospectus containing all relevant information on the securities. The issuer must then wait 20 days after the registration statement is filed with the SEC before it can sell any of the securities. If the SEC does not object during the 20-day waiting period, the securities can be sold.

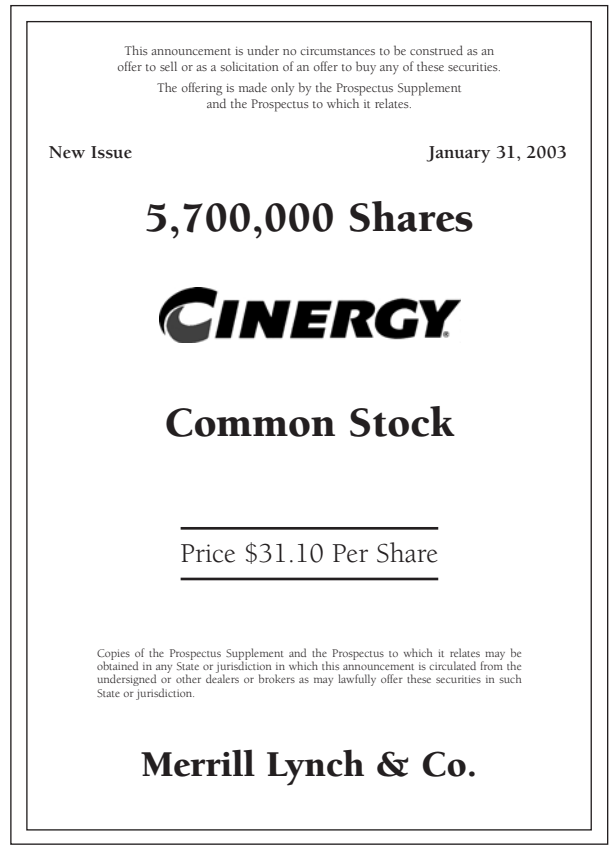
Following the Financial News

New Securities Issues

Information about new securities being issued is presented in distinctive advertisements published in the *Wall Street Journal* and other newspapers. These advertisements, called "tombstones" because of their appearance, are typically found in the "Money and Investing" section of the *Wall Street Journal*.

The tombstone indicates the number of shares of stock being issued (5.7 million shares for Cinergy) and the investment bank involved in selling them.

Source: *Wall Street Journal*, Wednesday, February 12, 2003, p. C5.



This announcement is under no circumstances to be construed as an offer to sell or as a solicitation of an offer to buy any of these securities.

The offering is made only by the Prospectus Supplement and the Prospectus to which it relates.

New Issue **January 31, 2003**

5,700,000 Shares

CINERGY

Common Stock

Price $31.10 Per Share

Copies of the Prospectus Supplement and the Prospectus to which it relates may be obtained in any State or jurisdiction in which this announcement is circulated from the undersigned or other dealers or brokers as may lawfully offer these securities in such State or jurisdiction.

Merrill Lynch & Co.

Securities Brokers and Dealers

Securities brokers and dealers conduct trading in secondary markets. Brokers act as agents for investors in the purchase or sale of securities. Their function is to match buyers with sellers, a function for which they are paid brokerage commissions. In contrast to brokers, dealers link buyers and sellers by standing ready to buy and sell securities at given prices. Therefore, dealers hold inventories of securities and make their living by selling these securities for a slightly higher price than they paid for them—that is, on the "spread" between the asked price and the bid price. This can be a high-risk business because dealers hold securities that can rise or fall in price; in recent years, several firms specializing in bonds have collapsed. Brokers, by contrast, are not as exposed to risk because they do not own the securities involved in their business dealings.

www.sec.gov

The Securities and Exchange Commission web site contains regulatory actions, concept releases, interpretive releases, and more.

Brokerage firms engage in all three securities market activities, acting as brokers, dealers, and investment bankers. The largest in the United States is Merrill Lynch; other well-known ones are PaineWebber, Morgan Stanley Dean Witter, and Salomon Smith Barney. The SEC not only regulates the investment banking operation of the firms but also restricts brokers and dealers from misrepresenting securities and from

trading on *insider information*, nonpublic information known only to the management of a corporation.

The forces of competition led to an important development: Brokerage firms started to engage in activities traditionally conducted by commercial banks. In 1977, Merrill Lynch developed the cash management account (CMA), which provides a package of financial services that includes credit cards, immediate loans, check-writing privileges, automatic investment of proceeds from the sale of securities into a money market mutual fund, and unified record keeping. CMAs were adopted by other brokerage firms and spread rapidly. The result is that the distinction between banking activities and the activities of nonbank financial institutions has become blurred (see Box 5). Another development is the growing importance of the Internet in securities markets (Box 6).

Organized Exchanges

www.nyse.com

At the New York Stock Exchange home page, you will find listed companies, member information, real-time market indices, and current stock quotes.

As discussed in Chapter 2, secondary markets can be organized either as over-the-counter markets, in which trades are conducted using dealers, or as organized exchanges, in which trades are conducted in one central location. The New York Stock Exchange (NYSE), trading thousands of securities, is the largest organized exchange in the world, and the American Stock Exchange (AMEX) is a distant second. A number of smaller regional exchanges, which trade only a small number of securities (under 100), exist in places such as Boston and Los Angeles.

Organized stock exchanges actually function as a hybrid of an auction market (in which buyers and sellers trade with each other in a central location) and a dealer

Box 5

The Return of the Financial Supermarket?

In the 1980s, companies dreamed of creating "financial supermarkets" in which there would be one-stop shopping for financial services. Consumers would be able to make deposits into their checking accounts, buy mutual funds, get a mortgage or a student loan, get a car or life insurance policy, obtain a credit card, or buy real estate. In the early 1980s, Sears, which already owned Allstate Insurance and a consumer finance subsidiary, bought Coldwell Banker Real Estate and Dean Witter, a brokerage firm. It also acquired a $6 billion California-based savings bank and introduced its Discover Card. Unfortunately, the concept of the financial supermarket never worked at Sears. (Indeed, the concept was derided as "stocks 'n' socks.") Sears's financial service firms lost money and Sears began to sell off these businesses in the late 1980s and early 1990s.

Sears is not the only firm to find it difficult to make a go of the financial supermarket concept. In the 1980s, American Express bought Shearson, Loeb Rhodes, a brokerage and securities firm, only to find it unprofitable. Similarly, Bank of America's purchase of Charles Schwab, the discount broker, also proved to be unprofitable.

Citicorp and Travelers Group, which merged in October 1998 with the view that Congress would remove all barriers to combining banking and nonbanking businesses in a financial service firm (which the Congress subsequently did in 1999), bet that the financial supermarket is an idea whose time has come. Citigroup hopes that the time is right to take advantage of economies of scope. With Citicorp's success at retail banking and the credit card business—it is the largest credit card issuer, with over 60 million outstanding—and Travelers' success in the insurance and securities business, the merged company, Citigroup, hopes to generate huge profits by providing convenient financial shopping for the consumer.

Box 6: E-Finance

The Internet Comes to Wall Street

An important development in recent years is the growing importance of the Internet in securities markets. Initial public offerings of stock are now being sold on the Internet, and many brokerage firms allow clients to conduct securities trades online or to transmit buy and sell orders via e-mail. In June of 1999, Wall Street was rocked by the announcement that its largest full-service brokerage firm, Merrill Lynch, would begin offering online trading for as little as $29.95 a trade to its five million customers. Now online trading is ubiquitous. The brokerage business will never be the same.

market (in which dealers make the market by buying and selling securities at given prices). Securities are traded on the floor of the exchange with the help of a special kind of dealer-broker called a **specialist**. A specialist matches buy and sell orders submitted at the same price and so performs a brokerage function. However, if buy and sell orders do not match up, the specialist buys stocks or sells from a personal inventory of securities, in this manner performing a dealer function. By assuming both functions, the specialist maintains orderly trading of the securities for which he or she is responsible.

Organized exchanges in which securities are traded are also regulated by the SEC. Not only does the SEC have the authority to impose regulations that govern the behavior of brokers and dealers involved with exchanges, but it also has the authority to alter the rules set by exchanges. In 1975, for example, the SEC disallowed rules that set minimum brokerage commission rates. The result was a sharp drop in brokerage commission rates, especially for institutional investors (mutual funds and pension funds), which purchase large blocks of stock. The Securities Amendments Act of 1975 confirmed the SEC's action by outlawing the setting of minimum brokerage commissions.

Furthermore, the Securities Amendments Act directed the SEC to facilitate a national market system that consolidates trading of all securities listed on the national and regional exchanges as well as those traded in the over-the-counter market using the National Association of Securities Dealers' automated quotation system (NASDAQ). Computers and advanced telecommunications, which reduce the costs of linking these markets, have encouraged the expansion of a national market system. We thus see that legislation and modern computer technology are leading the way to a more competitive securities industry.

The growing internationalization of capital markets has encouraged another trend in securities trading. Increasingly, foreign companies are being listed on U.S. stock exchanges, and the markets are moving toward trading stocks internationally, 24 hours a day.

Summary

1. Insurance providers, which are regulated by the states, acquire funds by selling policies that pay out benefits if catastrophic events occur. Property and casualty insurance companies hold more liquid assets than life insurance companies because of greater uncertainty regarding the benefits they will have to pay out. All insurers face moral hazard and adverse selection problems that explain the use of insurance management

tools, such as information collection and screening of potential policyholders, risk-based premiums, restrictive provisions, prevention of fraud, cancellation of insurance, deductibles, coinsurance, and limits on the amount of insurance.

2. Pension plans provide income payments to people when they retire after contributing to the plans for many years. Pension funds have experienced very rapid growth as a result of encouragement by federal tax policy and now play an important role in the stock market. Many pension plans are underfunded, which means that in future years they will have to pay out higher benefits than the value of their contributions and earnings. The problem of underfunding is especially acute for public pension plans such as Social Security. To prevent abuses, Congress enacted the Employee Retirement Income Security Act (ERISA), which established minimum standards for reporting, vesting, and degree of underfunding of private pension plans. This act also created the Pension Benefit Guarantee Corporation, which insures pension benefits.
3. Finance companies raise funds by issuing commercial paper and stocks and bonds and use the proceeds to make loans that are particularly suited to consumer and business needs. Virtually unregulated in comparison to commercial banks and thrift institutions, finance companies have been able to tailor their loans to customer needs very quickly and have grown rapidly.
4. Mutual funds sell shares and use the proceeds to buy securities. Open-end funds issue shares that can be redeemed at any time at a price tied to the asset value of the firm. Closed-end funds issue nonredeemable shares, which are traded like common stock. They are less popular than open-end funds because their shares are not as liquid. Money market mutual funds hold only short-term, high-quality securities, allowing shares to be redeemed at a fixed value using checks. Shares in these funds effectively function as checkable deposits that earn market interest rates. All mutual funds are regulated by the Securities and Exchange Commission (SEC).
5. Investment bankers assist in the initial sale of securities in primary markets, whereas securities brokers and dealers assist in the trading of securities in the secondary markets, some of which are organized into exchanges. The SEC regulates the financial institutions in the securities markets and ensures that adequate information reaches prospective investors.

Key Terms

annuities, p. 288
brokerage firms, p. 304
closed-end fund, p. 299
coinsurance, p. 293
deductible, p. 292
defined-benefit plan, p. 294
defined-contribution plan, p. 294
fully funded, p. 294
hedge fund, p. 299
initial public offering (IPO), p. 303
load funds, p. 299
no-load funds, p. 299
open-end fund, p. 298
reinsurance, p. 290
seasoned issue, p. 303
specialist, p. 306
underfunded, p. 294
underwriters, p. 303

Questions and Problems

Questions marked with an asterisk are answered at the end of the book in an appendix, "Answers to Selected Questions and Problems."

*1. If death rates were to become less predictable than they are, how would life insurance companies change the types of assets they hold?
2. Why do property and casualty insurance companies have large holdings of municipal bonds but life insurance companies do not?
*3. Why are all defined contribution pension plans fully funded?
4. How can favorable tax treatment of pension plans encourage saving?

*5. "In contrast to private pension plans, government pension plans are rarely underfunded." Is this statement true, false, or uncertain? Explain your answer.

6. What explains the widespread use of deductibles in insurance policies?

*7. Why might insurance companies restrict the amount of insurance a policyholder can buy?

8. Why are restrictive provisions a necessary part of insurance policies?

*9. If you needed to take out a loan, why might you first go to your local bank rather than to a finance company?

10. Explain why shares in closed-end mutual funds typically sell for less than the market value of the stocks they hold.

*11. Why might you buy a no-load mutual fund instead of a load fund?

12. Why can a money market mutual fund allow its shareholders to redeem shares at a fixed price but other mutual funds cannot?

*13. Why might government loan guarantees be a high-cost way for the government to subsidize certain activities?

14. If you like to take risks, would you rather be a dealer, a broker, or a specialist? Why?

*15. Is investment banking a good career for someone who is afraid of taking risks? Why or why not?

Web Exercises

1. Initial public offerings (IPOs) are where securities are sold to the public for the very first time. Go to http://ipo.com. This site lists various statistics regarding the IPO market.
   a. What is the largest IPO year to date ranked by amount raised?
   b. What is the next IPO to be offered to the public?
   c. How many IPOs were priced this year?
2. The Federal Reserve maintains extensive data on finance companies. Go to www.federalreserve.gov/releases and scroll down until you find G.20 Finance Companies. Click on "Releases" and find the current release.
   a. Review the terms of credit for new car loans. What is the most recent average interest rate and what is the term to maturity? How much is the average new car loan offered by finance companies?
   b. Do finance companies make more consumer loans, real estate loans, or business loans?
   c. Which type of loan has grown most rapidly over the last 5 years?

Chapter 13

Financial Derivatives

PREVIEW

Starting in the 1970s and increasingly in the 1980s and 1990s, the world became a riskier place for the financial institutions described in this part of the book. Swings in interest rates widened, and the bond and stock markets went through some episodes of increased volatility. As a result of these developments, managers of financial institutions became more concerned with reducing the risk their institutions faced. Given the greater demand for risk reduction, the process of financial innovation described in Chapter 9 came to the rescue by producing new financial instruments that help financial institution managers manage risk better. These instruments, called **financial derivatives**, have payoffs that are linked to previously issued securities and are extremely useful risk reduction tools.

In this chapter, we look at the most important financial derivatives that managers of financial institutions use to reduce risk: forward contracts, financial futures, options, and swaps. We examine not only how markets for each of these financial derivatives work but also how they can be used by financial institutions to manage risk. We also study financial derivatives because they have become an important source of profits for financial institutions, particularly larger banks, which, as we saw in Chapter 10, have found their traditional business declining.

Hedging

Financial derivatives are so effective in reducing risk because they enable financial institutions to **hedge**; that is, engage in a financial transaction that reduces or eliminates risk. When a financial institution has bought an asset, it is said to have taken a **long position**, and this exposes the institution to risk if the returns on the asset are uncertain. On the other hand, if it has sold an asset that it has agreed to deliver to another party at a future date, it is said to have taken a **short position**, and this can also expose the institution to risk. Financial derivatives can be used to reduce risk by invoking the following basic principle of hedging: ***Hedging risk involves engaging in a financial transaction that offsets a long position by taking an additional short position, or offsets a short position by taking an additional long position***. In other words, if a financial institution has *bought* a security and has therefore taken a long position, it conducts a hedge by contracting to *sell* that security (take a short position) at some future date. Alternatively, if it has taken a short position by *selling* a security that it needs to deliver at a future date, then it conducts a hedge by contracting to *buy*

that security (take a long position) at a future date. We look at how this principle can be applied using forward and futures contracts.

Interest-Rate Forward Contracts

Forward contracts are agreements by two parties to engage in a financial transaction at a future (forward) point in time. Here we focus on forward contracts that are linked to debt instruments, called **interest-rate forward contracts**; later in the chapter, we discuss forward contracts for foreign currencies.

Interest-rate forward contracts involve the future sale of a debt instrument and have several dimensions: (1) specification of the actual debt instrument that will be delivered at a future date, (2) amount of the debt instrument to be delivered, (3) price (interest rate) on the debt instrument when it is delivered, and (4) date on which delivery will take place. An example of an interest-rate forward contract might be an agreement for the First National Bank to sell to the Rock Solid Insurance Company, one year from today, $5 million face value of the 8s of 2023 Treasury bonds (that is, coupon bonds with an 8% coupon rate that mature in 2023) at a price that yields the same interest rate on these bonds as today's, say 8%. Because Rock Solid will buy the securities at a future date, it is said to have taken a long position, while the First National Bank, which will sell the securities, has taken a short position.

Application

Hedging with Interest-Rate Forward Contracts

Why would the First National Bank want to enter into this forward contract with Rock Solid Insurance Company in the first place?

To understand, suppose that you are the manager of the First National Bank and have bought $5 million of the 8s of 2023 Treasury bonds. The bonds are currently selling at par value, so their yield to maturity is 8%. Because these are long-term bonds, you recognize that you are exposed to substantial interest-rate risk: If interest rates rise in the future, the price of these bonds will fall and result in a substantial capital loss that may cost you your job. How do you hedge this risk?

Knowing the basic principle of hedging, you see that your long position in these bonds can be offset by a short position with a forward contract. That is, you need to contract to sell these bonds at a future date at the current par value price. As a result, you agree with another party—in this case, Rock Solid Insurance Company—to sell them the $5 million of the 8s of 2023 Treasury bonds at par one year from today. By entering into this forward contract, you have successfully hedged against interest-rate risk. By locking in the future price of the bonds, you have eliminated the price risk you face from interest-rate changes.

Why would Rock Solid Insurance Company want to enter into the futures contract with the First National Bank? Rock Solid expects to receive premiums of $5 million in one year's time that it will want to invest in the 8s of 2023, but worries that interest rates on these bonds will decline between now and next year. By using the forward contract, it is able to lock in the 8% interest rate on the Treasury bonds that will be sold to it by the First National Bank.

Pros and Cons of Forward Contracts

The advantage of forward contracts is that they can be as flexible as the parties involved want them to be. This means that an institution like the First National Bank may be able to hedge completely the interest-rate risk for the exact security it is holding in its portfolio, just as it has in our example.

However, forward contracts suffer from two problems that severely limit their usefulness. The first is that it may be very hard for an institution like the First National Bank to find another party (called a *counterparty*) to make the contract with. There are brokers to facilitate the matching up of parties like the First National Bank with the Rock Solid Insurance Company, but there may be few institutions that want to engage in a forward contract specifically for the 8s of 2023. This means that it may prove impossible to find a counterparty when a financial institution like the First National Bank wants to make a specific type of forward contract. Furthermore, even if the First National Bank finds a counterparty, it may not get as high a price as it wants because there may not be anyone else to make the deal with. A serious problem for the market in interest-rate forward contracts, then, is that it may be difficult to make the financial transaction or that it will have to be made at a disadvantageous price; in the parlance of financial economists, this market suffers from a *lack of liquidity*. (Note that this use of the term *liquidity* when it is applied to a market is somewhat broader than its use when it is applied to an asset. For an asset, liquidity refers to the ease with which the asset can be turned into cash; whereas for a market, liquidity refers to the ease of carrying out financial transactions.)

The second problem with forward contracts is that they are subject to default risk. Suppose that in one year's time, interest rates rise so that the price of the 8s of 2023 falls. The Rock Solid Insurance Company might then decide that it would like to default on the forward contract with the First National Bank, because it can now buy the bonds at a price lower than the agreed price in the forward contract. Or perhaps Rock Solid may not have been rock solid after all, and may have gone bust during the year, and no longer be available to complete the terms of the forward contract. Because there is no outside organization guaranteeing the contract, the only recourse is for the First National Bank to go to the courts to sue Rock Solid, but this process will be costly. Furthermore, if Rock Solid is already bankrupt, the First National Bank will suffer a loss; the bank can no longer sell the 8s of 2023 at the price it had agreed on with Rock Solid, but instead will have to sell at a price well below that, because the price of these bonds has fallen.

The presence of default risk in forward contracts means that parties to these contracts must check each other out to be sure that the counterparty is both financially sound and likely to be honest and live up to its contractual obligations. Because this type of investigation is costly and because all the adverse selection and moral hazard problems discussed in earlier chapters apply, default risk is a major barrier to the use of interest-rate forward contracts. When the default risk problem is combined with a lack of liquidity, we see that these contracts may be of limited usefulness to financial institutions. Although there is a market for interest-rate forward contracts, particularly in Treasury and mortgage-backed securities, it is not nearly as large as the financial futures market, to which we turn next.

Financial Futures Contracts and Markets

Given the default risk and liquidity problems in the interest-rate forward market, another solution to hedging interest-rate risk was needed. This solution was provided by the development of financial futures contracts by the Chicago Board of Trade starting in 1975.

http://home.teleport.com/~rpotts/fincontr.html
Find information about financial futures contract specifications.

A **financial futures contract** is similar to an interest-rate forward contract, in that it specifies that a financial instrument must be delivered by one party to another on a stated future date. However, it differs from an interest-rate forward contract in several ways that overcome some of the liquidity and default problems of forward markets.

To understand what financial futures contracts are all about, let's look at one of the most widely traded futures contracts—that for Treasury bonds, which are traded on the Chicago Board of Trade. (An illustration of how prices on these contracts are quoted can be found in the "Following the Financial News" box.) The contract value is for $100,000 face value of bonds. Prices are quoted in points, with each point equal to $1,000, and the smallest change in price is one thirty-second of a point ($31.25). This contract specifies that the bonds to be delivered must have at least 15 years to maturity at the delivery date (and must also not be callable—that is, redeemable by the Treasury at its option—in less than 15 years). If the Treasury bonds delivered to settle the futures contract have a coupon rate different from the 8% specified in the futures contract, the amount of bonds to be delivered is adjusted to reflect the difference in value between the delivered bonds and the 8% coupon bond. In line with the terminology used for forward contracts, parties who have bought a futures contract

Following the Financial News

Financial Futures

The prices for financial futures contracts for debt instruments are published daily. In the *Wall Street Journal*, these prices are found in the "Commodities" section under the "Interest Rate" heading of the "Future Prices" columns. An excerpt is reproduced here.

INTEREST RATE

TREASURY BONDS (CBT)-$100,000; pts. 32nds of 100%

	Open	High	Low	Settle	Chg.	Lifetime High	Lifetime Low	Open Int.
Mar	111-08	111-30	110-24	111-23	21	113-28	100-05	393,546
June	109-26	110-13	109-15	110-12	21	112-15	105-00	37,713

Est vol 179,387; vol Wed 159,069; open int 431,381, +368.

Information for each contract is presented in columns, as follows. (The Chicago Board of Trade's contract for delivery of long-term Treasury bonds in March 2000 is used as an example.)

Open: Opening price; each point corresponds to $1,000 of face value—111 8/32 is $111,250 for the March contract

High: Highest traded price that day—111 30/32 is $111,938 for the March contract

Low: Lowest traded price that day—110 24/32 is $110,750 for the March contract

Settle: Settlement price, the closing price that day—111 23/32 is $111,719 for the March contract

Chg: Change in the settlement price from the previous trading day—21/32 is $656.25 for the March contract

Lifetime High: Highest price ever—113 28/32 is $113,875 for the March contract

Lifetime Low: Lowest price ever—100 5/32 is $100,156 for the March contract

Open Interest: Number of contracts outstanding—393,546 for the March contract, with a face value of $39 billion (393,546 × $100,000)

Source: Wall Street Journal, January 31, 2003, p. B6.

and thereby agreed to buy (take delivery of) the bonds are said to have taken a *long position*, and parties who have sold a futures contract and thereby agreed to sell (deliver) the bonds have taken a *short position*.

To make our understanding of this contract more concrete, let's consider what happens when you buy or sell a Treasury bond futures contract. Let's say that on February 1, you sell one $100,000 June contract at a price of 115 (that is, $115,000). By selling this contract, you agree to deliver $100,000 face value of the long-term Treasury bonds to the contract's counterparty at the end of June for $115,000. By buying the contract at a price of 115, the buyer has agreed to pay $115,000 for the $100,000 face value of bonds when you deliver them at the end of June. If interest rates on long-term bonds rise, so that when the contract matures at the end of June, the price of these bonds has fallen to 110 ($110,000 per $100,000 of face value), the buyer of the contract will have lost $5,000, because he or she paid $115,000 for the bonds but can sell them only for the market price of $110,000. But you, the seller of the contract, will have gained $5,000, because you can now sell the bonds to the buyer for $115,000 but have to pay only $110,000 for them in the market.

It is even easier to describe what happens to the parties who have purchased futures contracts and those who have sold futures contracts if we recognize the following fact: ***At the expiration date of a futures contract, the price of the contract is the same as the price of the underlying asset to be delivered.*** To see why this is the case, consider what happens on the expiration date of the June contract at the end of June when the price of the underlying $100,000 face value Treasury bond is 110 ($110,000). If the futures contract is selling below 110—say at 109—a trader can buy the contract for $109,000, take delivery of the bond, and immediately sell it for $110,000, thereby earning a quick profit of $1,000. Because earning this profit involves no risk, it is a great deal that everyone would like to get in on. That means that everyone will try to buy the contract, and as a result, its price will rise. Only when the price rises to 110 will the profit opportunity cease to exist and the buying pressure disappear. Conversely, if the price of the futures contract is above 110—say at 111—everyone will want to sell the contract. Now the sellers get $111,000 from selling the futures contract but have to pay only $110,000 for the Treasury bonds that they must deliver to the buyer of the contract, and the $1,000 difference is their profit. Because this profit involves no risk, traders will continue to sell the futures contract until its price falls back down to 110, at which price there are no longer any profits to be made. The elimination of riskless profit opportunities in the futures market is referred to as **arbitrage**, and it guarantees that the price of a futures contract at expiration equals the price of the underlying asset to be delivered.[1]

Armed with the fact that a futures contract at expiration equals the price of the underlying asset makes it even easier to see who profits and who loses from such a contract when interest rates change. When interest rates have risen so that the price of the Treasury bond is 110 on the expiration day at the end of June, the June Treasury bond futures contract will also have a price of 110. Thus if you bought the contract for 115 in February, you have a loss of 5 points, or $5,000 (5% of $100,000). But if you sold the futures contract at 115 in February, the decline in price to 110 means that you have a profit of 5 points, or $5,000.

[1] In actuality, futures contracts sometimes set conditions for delivery of the underlying assets that cause the price of the contract at expiration to differ slightly from the price of the underlying assets. Because the difference in price is extremely small, we ignore it in this chapter.

Application

Hedging with Financial Futures

First National Bank can also use financial futures contracts to hedge the interest rate risk on its holdings of $5 million of the 8s of 2023. To see how, suppose that in March 2004, the 8s of 2023 are the long-term bonds that would be delivered in the Chicago Board of Trade's T-bond futures contract expiring one year in the future, in March 2005. Also suppose that the interest rate on these bonds is expected to remain at 8% over the next year, so that both the 8s of 2023 and the futures contract are selling at par (i.e., the $5 million of bonds is selling for $5 million and the $100,000 futures contract is selling for $100,000). The basic principle of hedging indicates that you need to offset the long position in these bonds with a short position, and so you have to sell the futures contract. But how many contracts should you sell? The number of contracts required to hedge the interest-rate risk is found by dividing the amount of the asset to be hedged by the dollar value of each contract, as is shown in Equation (1):

$$NC = VA/VC \tag{1}$$

where NC = number of contracts for the hedge
VA = value of the asset
VC = value of each contract

Given that the 8s of 2023 are the long-term bonds that would be delivered in the CBT T-bond futures contract expiring one year in the future and that the interest rate on these bonds is expected to remain at 8% over the next year, so that both the 8s of 2023 and the futures contract are selling at par, how many contracts must First National sell to remove its interest-rate exposure from its $5 million holdings of the 8s of 2023?[2] Since VA = $5 million and VC = $100,000:

$$NC = \$5 \text{ million}/\$100{,}000 = 50$$

You therefore hedge the interest-rate risk by selling 50 of the Treasury Bond futures contracts.

Now suppose that over the next year, interest rates increase to 10% due to an increased threat of inflation. The value of the 8s of 2023 that the First National Bank is holding will then fall to $4,163,508 in March 2005.[3] Thus the loss from the long position in these bonds is $836,492:

Value on March 2005 @ 10% interest rate	$ 4,163,508
Value on March 2004 @ 8% interest rate	− $ 5,000,000
Loss	− $ 836,492

[2] In the real world, designing a hedge is somewhat more complicated than the example here, because the bond that is most likely to be delivered might not be an 8s of 2023.

[3] The value of the bonds can be calculated using a financial calculator as follows: FV = $5,000,000, PMT = $500,000, I = 10%, N = 19, PV = $4,163,508.

www.rmahq.org
The web site of the Risk Management Association reports useful information such as annual statement studies, online publications, and so on.

However, the short position in the 50 futures contracts that obligate you to deliver $5 million of the 8s of 2023 on March 2004 have a value equal to $4,163,568, the value of the $5 million of bonds after the interest rate has risen to 10%, as we have seen before. Yet when you sold the futures contract, the buyer was obligated to pay you $5 million on the maturity date. Thus the gain from the short position on these contracts is also $836,492:

Amount paid to you on March 2005, agreed upon in March 2004	$5,000,000
Value of bonds delivered on March 2005 @ 10% interest rate	– $4,163,508
Gain	$ 836,492

Therefore the net gain for the First National Bank is zero, indicating that the hedge has been conducted successfully.

The hedge just described is called a **micro hedge** because the financial institution is hedging the interest-rate risk for a specific asset it is holding. A second type of hedge that financial institutions engage in is called a **macro hedge**, in which the hedge is for the institution's entire portfolio. For example, if a bank has more rate-sensitive liabilities than assets, we have seen in Chapter 9 that a rise in interest rates will cause the value of the bank to decline. By selling interest-rate future contracts that will yield a profit when interest rates rise, the bank can offset the losses on its overall portfolio from an interest-rate rise and thereby hedge its interest-rate risk.

Organization of Trading in Financial Futures Markets

Financial futures contracts are traded in the United States on organized exchanges such as the Chicago Board of Trade, the Chicago Mercantile Exchange, the New York Futures Exchange, the MidAmerica Commodity Exchange, and the Kansas City Board of Trade. These exchanges are highly competitive with one another, and each organization tries to design contracts and set rules that will increase the amount of futures trading on its exchange.

The futures exchanges and all trades in financial futures in the United States are regulated by the Commodity Futures Trading Commission (CFTC), which was created in 1974 to take over the regulatory responsibilities for futures markets from the Department of Agriculture. The CFTC oversees futures trading and the futures exchanges to ensure that prices in the market are not being manipulated, and it also registers and audits the brokers, traders, and exchanges to prevent fraud and to ensure the financial soundness of the exchanges. In addition, the CFTC approves proposed futures contracts to make sure that they serve the public interest. The most widely traded financial futures contracts listed in the *Wall Street Journal* and the exchanges where they are traded (along with the number of contracts outstanding, called **open interest**, on January 30, 2003) are listed in Table 1.

www.usafutures.com/stockindexfutures.htm
Detailed information about stock index futures.

Given the globalization of other financial markets in recent years, it is not surprising that increased competition from abroad has been occurring in financial futures markets as well.

Table 1 Widely Traded Financial Futures Contracts

Type of Contract	Contract Size	Exchange*	Open Interest January 30, 2003 Reflects March 2003 Futures
Treasury Rate Contracts			
Treasury bonds	$100,000	CBT	393,546
Treasury notes	$100,000	CBT	746,015
Five-year Treasury notes	$100,000	CBT	683,499
Two-year Treasury notes	$200,000	CBT	106,184
Thirty-day Fed levels	$5 million	CBT	49,069
Treasury bills	$1 million	CME	292
One-month LIBOR	$3 million	CME	6,389
Municipal Bond Index	$1,000	CBT	2,683
Eurodollar	$4 million	CME	747,691
Euroyen	$100 million	CME	10,765
Sterling	£500,000	LIFFE	159,800
Long Gilt	£100,000	LIFFE	90,093
Three-month Euribor	€ 1 million	LIFFE	497,688
Euroswiss franc	SF 1 million	LIFFE	85,366
Ten-year Euronational bonds	€ 100,000	EUREX	812,029
Canadian banker's acceptance	C$1 million	ME	64,333
Stock Index Contracts			
Standard & Poor's 500 Index	$250 × index	CME	577,661
Standard & Poor's MIDCAP 400	$500 × index	CME	13,652
NASDAQ 100	$100 × index	CME	71,233
Nikkei 225 Stock Average	$5 × index	CME	16,193
Financial Times–Stock Exchange 100-Share Index	£10 per index point	LIFFE	460,997
Currency Contracts			
Yen	12,500,000 yen	CME	90,508
Euro	125,000 euros	CME	102,536
Canadian dollar	100,000 Canadian $	CME	89,651
British pound	100,000 pounds	CME	102,536
Swiss franc	125,000 francs	CME	55,402
Mexican peso	500,000 new pesos	CME	29,774

*Exchange abbreviations: CBT, Chicago Board of Trade; CME, Chicago Mercantile Exchange; LIFFE, London International Financial Futures Exchange; EUREX, European Exchange; ME, Montreal Exchange.

Source: Wall Street Journal, January 31, 2003, p. C21.

The Globalization of Financial Futures Markets

Because American futures exchanges were the first to develop financial futures, they dominated the trading of financial futures in the early 1980s. For example, in 1985, all of the top ten futures contracts were traded on exchanges in the United States. With the rapid growth of financial futures markets and the resulting high profits made by the American exchanges, foreign exchanges saw a profit opportunity and began to enter this business. By the 1990s, Eurodollar contracts traded on the London International Financial Futures Exchange, Japanese government bond contracts and Euroyen contracts traded on the Tokyo Stock Exchange, French government bond contracts traded on the Marché à Terme International de France, and Nikkei 225 contracts traded on the Osaka Securities Exchange all became among the most widely traded futures contracts in the world.

Foreign competition has also spurred knockoffs of the most popular financial futures contracts initially developed in the United States. These contracts traded on foreign exchanges are virtually identical to those traded in the United States and have the advantage that they can be traded when the American exchanges are closed. The movement to 24-hour-a-day trading in financial futures has been further stimulated by the development of the Globex electronic trading system, which allows traders throughout the world to trade futures even when the exchanges are not officially open. Financial futures trading has thus become completely internationalized, and competition between U.S. and foreign exchanges will be intense in the future.

Explaining the Success of Futures Markets

The tremendous success of the financial futures market in Treasury bonds is evident from the fact that the total open interest of Treasury bond contracts was over 393,000 on January 30, 2003, for a total value of over $39 billion (393,000 × $100,000). There are several differences between financial futures and forward contracts and in the organization of their markets that help explain why financial futures markets like those for Treasury bonds have been so successful.

Several features of futures contracts were designed to overcome the liquidity problem inherent in forward contracts. The first feature is that, in contrast to forward contracts, the quantities delivered and the delivery dates of futures contracts are standardized, making it more likely that different parties can be matched up in the futures market, thereby increasing the liquidity of the market. In the case of the Treasury bond contract, the quantity delivered is $100,000 face value of bonds, and the delivery dates are set to be the last business day of March, June, September, and December. The second feature is that after the futures contract has been bought or sold, it can be traded (bought or sold) again at any time until the delivery date. In contrast, once a forward contract is agreed on, it typically cannot be traded. The third feature is that in a futures contract, not just one specific type of Treasury bond is deliverable on the delivery date, as in a forward contract. Instead, any Treasury bond that matures in more than 15 years and is not callable for 15 years is eligible for delivery. Allowing continuous trading also increases the liquidity of the futures market, as does the ability to deliver a range of Treasury bonds rather than one specific bond.

Another reason why futures contracts specify that more than one bond is eligible for delivery is to limit the possibility that someone might corner the market and "squeeze" traders who have sold contracts. To corner the market, someone buys up all the deliverable securities so that investors with a short position cannot obtain from anyone else the securities that they contractually must deliver on the delivery date. As

a result, the person who has cornered the market can set exorbitant prices for the securities that investors with a short position must buy to fulfill their obligations under the futures contract. The person who has cornered the market makes a fortune, but investors with a short position take a terrific loss. Clearly, the possibility that corners might occur in the market will discourage people from taking a short position and might therefore decrease the size of the market. By allowing many different securities to be delivered, the futures contract makes it harder for anyone to corner the market, because a much larger amount of securities would have to be purchased to establish the corner. Corners are a concern to both regulators and the organized exchanges that design futures contracts.

Trading in the futures market has been organized differently from trading in forward markets to overcome the default risk problems arising in forward contracts. In both types, for every contract, there must be a buyer who is taking a long position and a seller who is taking a short position. However, the buyer and seller of a futures contract make their contract not with each other but with the clearinghouse associated with the futures exchange. This setup means that the buyer of the futures contract does not need to worry about the financial health or trustworthiness of the seller, or vice versa, as in the forward market. As long as the clearinghouse is financially solid, buyers and sellers of futures contracts do not have to worry about default risk.

To make sure that the clearinghouse is financially sound and does not run into financial difficulties that might jeopardize its contracts, buyers or sellers of futures contracts must put an initial deposit, called a **margin requirement**, of perhaps $2,000 per Treasury bond contract into a margin account kept at their brokerage firm. Futures contracts are then **marked to market** every day. What this means is that at the end of every trading day, the change in the value of the futures contract is added to or subtracted from the margin account. Suppose that after you buy the Treasury bond contract at a price of 115 on Wednesday morning, its closing price at the end of the day, the *settlement price*, falls to 114. You now have a loss of 1 point, or $1,000, on the contract, and the seller who sold you the contract has a gain of 1 point, or $1,000. The $1,000 gain is added to the seller's margin account, making a total of $3,000 in that account, and the $1,000 loss is subtracted from your account, so you now only have $1,000 in your account. If the amount in this margin account falls below the maintenance margin requirement (which can be the same as the initial requirement but is usually a little less), the trader is required to add money to the account. For example, if the maintenance margin requirement is also $2,000, you would have to add $1,000 to your account to bring it up to $2,000. Margin requirements and marking to market make it far less likely that a trader will default on a contract, thus protecting the futures exchange from losses.

A final advantage that futures markets have over forward markets is that most futures contracts do not result in delivery of the underlying asset on the expiration date, whereas forward contracts do. A trader who sold a futures contract is allowed to avoid delivery on the expiration date by making an offsetting purchase of a futures contract. Because the simultaneous holding of the long and short positions means that the trader would in effect be delivering the bonds to itself, under the exchange rules the trader is allowed to cancel both contracts. Allowing traders to cancel their contracts in this way lowers the cost of conducting trades in the futures market relative to the forward market in that a futures trader can avoid the costs of physical delivery, which is not so easy with forward contracts.

Application

Hedging Foreign Exchange Risk

As we discussed in Chapter 1, foreign exchange rates have been highly volatile in recent years. The large fluctuations in exchange rates subject financial institutions and other businesses to significant foreign exchange risk because they generate substantial gains and losses. Luckily for financial institution managers, the financial derivatives discussed in this chapter—forward and financial futures contracts—can be used to hedge foreign exchange risk.

To understand how financial institution managers manage foreign exchange risk, let's suppose that in January, the First National Bank's customer Frivolous Luxuries, Inc. is due a payment of 10 million euros in two months for $10 million worth of goods it has just sold in Germany. Frivolous Luxuries is concerned that if the value of the euro falls substantially from its current value of $1, the company might suffer a large loss because the 10 million euro payment will no longer be worth $10 million. So Sam, the CEO of Frivolous Luxuries, calls up his friend Mona, the manager of the First National Bank, and asks her to hedge this foreign exchange risk for his company. Let's see how the bank manager does this using forward and financial futures contracts.

Hedging Foreign Exchange Risk with Forward Contracts

Forward markets in foreign exchange have been highly developed by commercial banks and investment banking operations that engage in extensive foreign exchange trading and so are widely used to hedge foreign exchange risk. Mona knows that she can use this market to hedge the foreign exchange risk for Frivolous Luxuries. Such a hedge is quite straightforward for her to execute. Because the payment of euros in two months means that at that time Sam would hold a long position in euros, Mona knows that the basic principle of hedging indicates that she should offset this long position by a short position. Thus, she just enters a forward contract that obligates her to sell 10 million euros two months from now in exchange for dollars at the current forward rate of $1 per euro.[4]

In two months, when her customer receives the 10 million euros, the forward contract ensures that it is exchanged for dollars at an exchange rate of $1 per euro, thus yielding $10 million. No matter what happens to future exchange rates, Frivolous Luxuries will be guaranteed $10 million for the goods it sold in Germany. Mona calls up her friend Sam to let him know that his company is now protected from any foreign exchange movements, and he thanks her for her help.

[4]The forward exchange rate will probably differ slightly from the current spot rate of $1 per euro because the interest rates in Germany and the United States may not be equal. In that case, as we will see in Equation 2 in Chapter 19, the future expected exchange rate will not equal the current spot rate and neither will the forward rate. However, since interest differentials have typically been less than 6% at an annual rate (1% bimonthly), the expected appreciation or depreciation of the euro over a two-month period has always been less than 1%. Thus the forward rate is always close to the current spot rate, and so our assumption in the example that the forward rate and the spot rate are the same is a reasonable one.

Hedging Foreign Exchange Risk with Futures Contracts

As an alternative, Mona could have used the currency futures market to hedge the foreign exchange risk. In this case, she would see that the Chicago Mercantile Exchange has a euro contract with a contract amount of 125,000 euros and a price of $1 per euro. To do the hedge, Mona must sell euros as with the forward contract, to the tune of 10 million euros of the March futures. How many of the Chicago Mercantile Exchange March euro contracts must Mona sell in order to hedge the 10 million euro payment due in March?

Using Equation 1 with VA = 10 million euros and VC = 125,000 euros:

$$NC = 10 \text{ million}/125{,}000 = 40$$

Thus Mona does the hedge by selling 40 of the CME euro contracts. Given the $1-per-euro price, the sale of the contract yields $40 \times 125,000$ euros = $10 million. The futures hedge thus again enables her to lock in the exchange rate for Frivolous Luxuries so that it gets its payment of $10 million.

One advantage of using the futures market is that the contract size of 125,000 euros, worth $125,000, is quite a bit smaller than the minimum size of a forward contract, which is usually $1 million or more. However, in this case, the bank manager is making a large enough transaction that she can use either the forward or the futures market. Her choice depends on whether the transaction costs are lower in one market than in the other. If the First National Bank is active in the forward market, that market would probably have the lower transaction costs, but if First National rarely deals in foreign exchange forward contracts, the bank manager may do better by sticking with the futures market.

Options

Another vehicle for hedging interest-rate and stock market risk involves the use of options on financial instruments. **Options** are contracts that give the purchaser the option, or *right*, to buy or sell the underlying financial instrument at a specified price, called the **exercise price** or **strike price**, within a specific period of time (the *term to expiration*). The seller (sometimes called the *writer*) of the option is *obligated* to buy or sell the financial instrument to the purchaser if the owner of the option exercises the right to sell or buy. These option contract features are important enough to be emphasized: The *owner* or buyer of an option does not have to exercise the option; he or she can let the option expire without using it. Hence the owner of an option is not obligated to take any action, but rather has the *right* to exercise the contract if he or she so chooses. The seller of an option, by contrast, has no choice in the matter; he or she *must* buy or sell the financial instrument if the owner exercises the option.

Because the right to buy or sell a financial instrument at a specified price has value, the owner of an option is willing to pay an amount for it called a **premium**. There are two types of option contracts: **American options** can be exercised *at any time up to* the expiration date of the contract, and **European options** can be exercised only *on* the expiration date.

Option contracts are written on a number of financial instruments (an example of which is shown in the "Following the Financial News" box). Options on individ-

Following the Financial News

Futures Options

The prices for financial futures options are published daily. In the *Wall Street Journal*, they are found in the section "Futures Options Prices" under the "Interest Rate" heading. An excerpt from this listing is reproduced here.

INTEREST RATE

T-BONDS (CBT)
$100,000; points and 64ths of 100%

Strike	Calls-Settle			Puts-Settle		
Price	Mar	Apr	May	Mar	Apr	May
111	2-34	2-21	2-62	0-08	1-15	1-57
112	1-45	1-50	2-28	0-19	1-43	2-22
113	1-01	1-20	1-61	0-39	2-14	...
114	0-34	0-59	1-36	1-08	2-52	...
115	0-16	0-41	1-14	1-54	...	...
116	0-07	0-28	0-60	2-45	...	...

Est vol 78,455;
Wd vol 51,578 calls 17,896 puts
Op int Wed 252,705 calls 282,711 puts

Information for each contract is reported in columns, as follows. (The Chicago Board of Trade's option on its Treasury bonds futures contract is used as an example.)

Source: Wall Street Journal, February 14, 2003, p. C10.

Strike Price: Strike (exercise) price of each contract, which runs from 111 to 116

Calls-Settle: Premium (price) at settlement for call options on the Treasury bond futures expiring in the month listed, with each full point representing $1,000 and 64ths of a point listed to the right of the hyphen; at a strike price of 112, the March call option's premium is 1 45/64, or $1,703.10 per contract

Puts-Settle: Premium (price) at settlement for put options on the Treasury bond futures expiring in the month listed, with each full point representing $1,000 and 64ths of a point listed to the right of the hyphen; at a strike price of 112, the March put option's premium is 19/64, or $296.90 per contract

ual stocks are called **stock options**, and such options have existed for a long time. Option contracts on financial futures called **financial futures options** or, more commonly, **futures options**, were developed in 1982 and have become the most widely traded option contracts.

You might wonder why option contracts are more likely to be written on financial futures than on underlying debt instruments such as bonds or certificates of deposit. As you saw earlier in the chapter, at the expiration date, the price of the futures contract and of the deliverable debt instrument will be the same because of arbitrage. So it would seem that investors should be indifferent about having the option written on the debt instrument or on the futures contract. However, financial futures contracts have been so well designed that their markets are often more liquid than the markets in the underlying debt instruments. So investors would rather have the option contract written on the more liquid instrument, in this case the futures contract. That explains why the most popular futures options are written on many of the same futures contracts listed in Table 1.

The regulation of option markets is split between the Securities and Exchange Commission (SEC), which regulates stock options, and the Commodity Futures Trading Commission (CFTC), which regulates futures options. Regulation focuses on ensuring that writers of options have enough capital to make good on their contractual obligations and on overseeing traders and exchanges to prevent fraud and ensure that the market is not being manipulated.

Option Contracts

A **call option** is a contract that gives the owner the right to buy a financial instrument at the exercise price within a specific period of time. A **put option** is a contract that gives the owner the right to sell a financial instrument at the exercise price within a specific period of time.

Study Guide

Remembering which is a call option and which is a put option is not always easy. To keep them straight, just remember that having a *call* option to buy a financial instrument is the same as having the option to *call* in the instrument for delivery at a specified price. Having a *put* option to sell a financial instrument is the same as having the option to *put* up an instrument for the other party to buy.

Profits and Losses on Option and Futures Contracts

To understand option contracts more fully, let's first examine the option on the same June Treasury bond futures contract that we looked at earlier in the chapter. Recall that if you buy this futures contract at a price of 115 (that is, $115,000), you have agreed to pay $115,000 for $100,000 face value of long-term Treasury bonds when they are delivered to you at the end of June. If you sold this futures contract at a price of 115, you agreed, in exchange for $115,000, to deliver $100,000 face value of the long-term Treasury bonds at the end of June. An option contract on the Treasury bond futures contract has several key features: (1) It has the same expiration date as the underlying futures contract, (2) it is an American option and so can be exercised at any time before the expiration date, and (3) the premium (price) of the option is quoted in points that are the same as in the futures contract, so each point corresponds to $1,000. If, for a premium of $2,000, you buy one call option contract on the June Treasury bond contract with an exercise price of 115, you have purchased the right to buy (call in) the June Treasury bond futures contract for a price of 115 ($115,000 per contract) at any time through the expiration date of this contract at the end of June. Similarly, when for $2,000 you buy a put option on the June Treasury bond contract with an exercise price of 115, you have the right to sell (put up) the June Treasury bond futures contract for a price of 115 ($115,000 per contract) at any time until the end of June.

Futures option contracts are somewhat complicated, so to explore how they work and how they can be used to hedge risk, let's first examine how profits and losses on the call option on the June Treasury bond futures contract occur. In February, our old friend Irving the Investor buys, for a $2,000 premium, a call option on the $100,000 June Treasury bond futures contract with a strike price of 115. (We assume that if Irving exercises the option, it is on the expiration date at the end of June and not before.) On the expiration date at the end of June, suppose that the underlying Treasury bond for the futures contract has a price of 110. Recall that on the expiration date, arbitrage forces the price of the futures contract to be the same as the price of the underlying bond, so it too has a price of 110 on the expiration date at the end of June. If Irving exercises the call option and buys the futures contract at an exercise price of 115, he will lose money by buying at 115 and selling at the lower market price of 110. Because Irving is smart, he will not exercise the option, but he will be out the $2,000 premium he paid. In such a situation, in which the price of the underlying financial instrument is below the exercise price, a call option is said to be "out of the money." At the price of 110 (less than the exercise price), Irving thus suffers a

loss on the option contract of the $2,000 premium he paid. This loss is plotted as point A in panel (a) of Figure 1.

On the expiration date, if the price of the futures contract is 115, the call option is "at the money," and Irving is indifferent whether he exercises his option to buy the futures contract or not, since exercising the option at 115 when the market price is also at 115 produces no gain or loss. Because he has paid the $2,000 premium, at the price of 115 his contract again has a net loss of $2,000, plotted as point B.

If the futures contract instead has a price of 120 on the expiration day, the option is "in the money," and Irving benefits from exercising the option: He would buy the futures contract at the exercise price of 115 and then sell it for 120, thereby earning a 5-point gain ($5,000 profit) on the $100,000 Treasury bond contract. Because Irving paid a $2,000 premium for the option contract, however, his net profit is

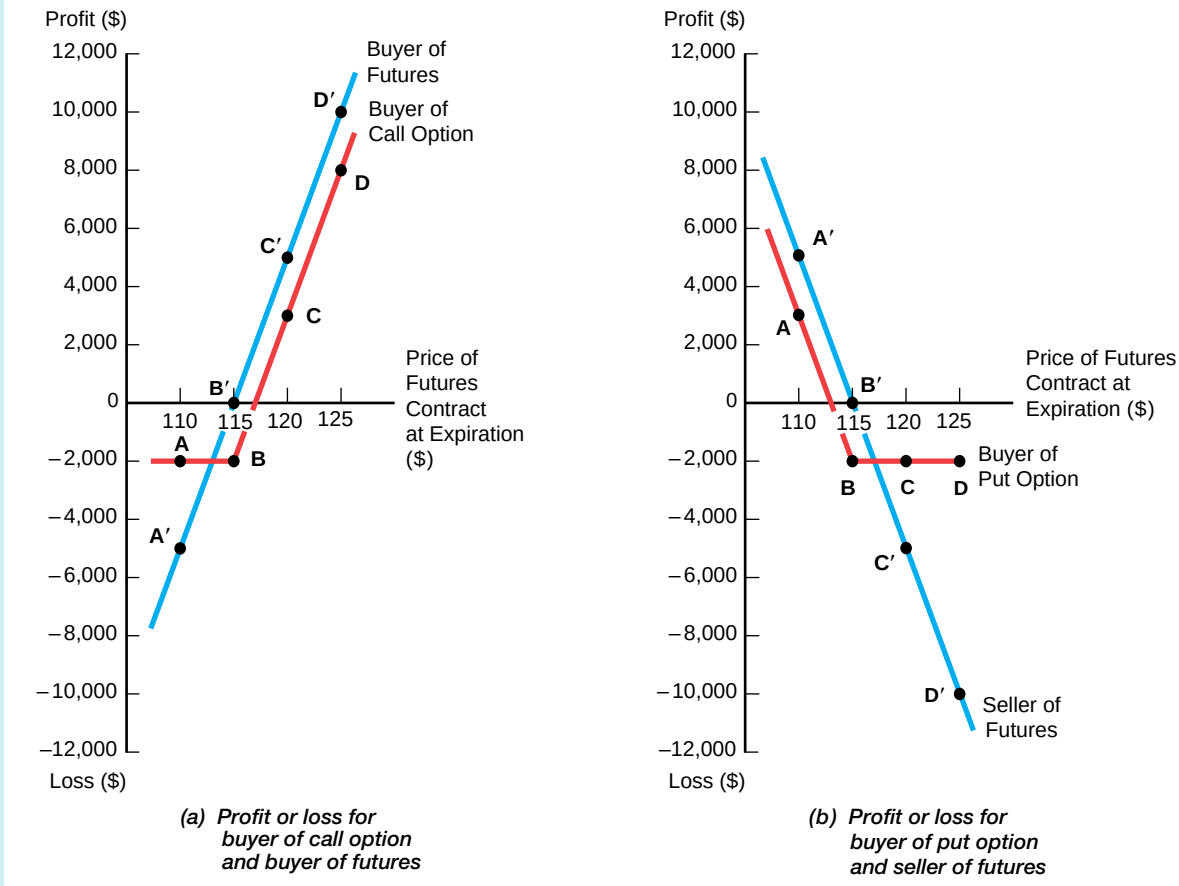


FIGURE 1 Profits and Losses on Options Versus Futures Contracts

The futures contract is the $100,000 June Treasury bond contract, and the option contracts are written on this futures contract with an exercise price of 115. Panel (a) shows the profits and losses for the buyer of the call option and the buyer of the futures contract, and panel (b) shows the profits and losses for the buyer of the put option and the seller of the futures contract.

$3,000 ($5,000 − $2,000). The $3,000 profit at a price of 120 is plotted as point C. Similarly, if the price of the futures contract rose to 125, the option contract would yield a net profit of $8,000 ($10,000 from exercising the option minus the $2,000 premium), plotted as point D. Plotting these points, we get the kinked profit curve for the call option that we see in panel (a).

Suppose that instead of purchasing the futures *option* contract in February, Irving decides instead to buy the $100,000 June Treasury bond *futures* contract at the price of 115. If the price of the bond on the expiration day at the end of June declines to 110, meaning that the price of the futures contract also falls to 110, Irving suffers a loss of 5 points, or $5,000. The loss of $5,000 on the futures contract at a price of 110 is plotted as point A′ in panel (a). At a price of 115 on the expiration date, Irving would have a zero profit on the futures contract, plotted as point B′. At a price of 120, Irving would have a profit on the contract of 5 points, or $5,000 (point C′), and at a price of 125, the profit would be 10 percentage points, or $10,000 (point D′). Plotting these points, we get the linear (straight-line) profit curve for the futures contract that appears in panel (a).

Now we can see the major difference between a futures contract and an option contract. As the profit curve for the futures contract in panel (a) indicates, the futures contract has a linear profit function: Profits grow by an equal dollar amount for every point increase in the price of the underlying financial instrument. By contrast, the kinked profit curve for the option contract is nonlinear, meaning that profits do not always grow by the same amount for a given change in the price of the underlying financial instrument. The reason for this nonlinearity is that the call option protects Irving from having losses that are greater than the amount of the $2,000 premium. In contrast, Irving's loss on the futures contract is $5,000 if the price on the expiration day falls to 110, and if the price falls even further, Irving's loss will be even greater. This insurance-like feature of option contracts explains why their purchase price is referred to as a premium. Once the underlying financial instrument's price rises above the exercise price, however, Irving's profits grow linearly. Irving has given up something by buying an option rather than a futures contract. As we see in panel (a), when the price of the underlying financial instrument rises above the exercise price, Irving's profits are always less than that on the futures contract by exactly the $2,000 premium he paid.

Panel (b) plots the results of the same profit calculations if Irving buys not a call but a put option (an option to sell) with an exercise price of 115 for a premium of $2,000 and if he sells the futures contract rather than buying one. In this case, if on the expiration date the Treasury bond futures have a price above the 115 exercise price, the put option is "out of the money." Irving would not want to exercise the put option and then have to sell the futures contract he owns as a result at a price below the market price and lose money. He would not exercise his option, and he would be out only the $2,000 premium he paid. Once the price of the futures contract falls below the 115 exercise price, Irving benefits from exercising the put option because he can sell the futures contract at a price of 115 but can buy it at a price below this. In such a situation, in which the price of the underlying instrument is below the exercise price, the put option is "in the money," and profits rise linearly as the price of the futures contract falls. The profit function for the put option illustrated in panel (b) of Figure 1 is kinked, indicating that Irving is protected from losses greater than the amount of the premium he paid. The profit curve for the sale of the futures contract is just the negative of the profit for the futures contract in panel (a) and is therefore linear.

Panel (b) of Figure 1 confirms the conclusion from panel (a) that profits on option contracts are nonlinear but profits on futures contracts are linear.

Study Guide

To make sure you understand how profits and losses on option and futures contracts are generated, calculate the net profits on the put option and the short position in the futures contract at prices on the expiration day of 110, 115, 120, and 125. Then verify that your calculations correspond to the points plotted in panel (b) of Figure 1.

Two other differences between futures and option contracts must be mentioned. The first is that the initial investment on the contracts differs. As we saw earlier in the chapter, when a futures contract is purchased, the investor must put up a fixed amount, the margin requirement, in a margin account. But when an option contract is purchased, the initial investment is the premium that must be paid for the contract. The second important difference between the contracts is that the futures contract requires money to change hands daily when the contract is marked to market, whereas the option contract requires money to change hands only when it is exercised.

Application

Hedging with Futures Options

Earlier in the chapter, we saw how the First National Bank could hedge the interest-rate risk on its $5 million holdings of 8s of 2023 by selling $5 million of T-bond futures. A rise in interest rates and the resulting fall in bond prices and bond futures contracts would lead to profits on the bank's sale of the futures contracts that would exactly offset the losses on the 8s of 2023 the bank is holding.

As panel (b) of Figure 1 suggests, an alternative way for the manager to protect against a rise in interest rates and hence a decline in bond prices is to buy $5 million of put options written on the same Treasury bond futures. As long as the exercise price is not too far from the current price as in panel (b), the rise in interest rates and decline in bond prices will lead to profits on the futures and the futures put options, profits that will offset any losses on the $5 million of Treasury bonds.

The one problem with using options rather than futures is that the First National Bank will have to pay premiums on the options contracts, thereby lowering the bank's profits in order to hedge the interest-rate risk. Why might the bank manager be willing to use options rather than futures to conduct the hedge? The answer is that the option contract, unlike the futures contract, allows the First National Bank to gain if interest rates decline and bond prices rise. With the hedge using futures contracts, the First National Bank does not gain from increases in bond prices because the profits on the bonds it is holding are offset by the losses from the futures contracts it has sold. However, as panel (b) of Figure 1 indicates, the situation when the hedge is conducted with put options is quite different: Once bond prices rise above the exercise price, the bank does not suffer additional losses on the option contracts. At the same time, the value of the Treasury bonds the bank is holding will increase, thereby leading to a profit for the bank. Thus using options rather than futures to conduct the micro hedge allows the bank to protect itself from rises in interest rates but still allows the bank to benefit from interest-rate declines (although the profit is reduced by the amount of the premium).

Similar reasoning indicates that the bank manager might prefer to use options to conduct the macro hedge to immunize the entire bank portfolio from interest-rate risk. Again, the strategy of using options rather than futures has the disadvantage that the First National Bank has to pay the premiums on these contracts up front. By contrast, using options allows the bank to keep the gains from a decline in interest rates (which will raise the value of the bank's assets relative to its liabilities), because these gains will not be offset by large losses on the option contracts.

In the case of a macro hedge, there is another reason why the bank might prefer option contracts to futures contracts. Profits and losses on futures contracts can cause accounting problems for banks because such profits and losses are not allowed to be offset by unrealized changes in the value of the rest of the bank's portfolio. Consider what happens when interest rates fall. If First National sells futures contracts to conduct the macro hedge, then when interest rates fall and the prices of the Treasury bond futures contracts rise, it will have large losses on these contracts. Of course, these losses are offset by unrealized profits in the rest of the bank's portfolio, but the bank is not allowed to offset these losses in its accounting statements. So even though the macro hedge is serving its intended purpose of immunizing the bank's portfolio from interest-rate risk, the bank would experience large accounting losses when interest rates fall. Indeed, bank managers have lost their jobs when perfectly sound hedges with interest-rate futures have led to large accounting losses. Not surprisingly, bank managers might shrink from using financial futures to conduct macro hedges for this reason.

Futures options, however, can come to the rescue of the managers of banks and other financial institutions. Suppose that First National conducted the macro hedge by buying put options instead of selling Treasury bond futures. Now if interest rates fall and bond prices rise well above the exercise price, the bank will not have large losses on the option contracts, because it will just decide not to exercise its options. The bank will not suffer the accounting problems produced by hedging with financial futures. Because of the accounting advantages of using futures options to conduct macro hedges, option contracts have become important to financial institution managers as tools for hedging interest-rate risk.

Factors Affecting the Prices of Option Premiums

If we again look closely at the *Wall Street Journal* entry for Treasury bond futures options in the "Following the Financial News" box on page 321, we learn several interesting facts about how the premiums on option contracts are priced. The first thing you may have noticed is that when the strike (exercise) price for a contract is set at a higher level, the premium for the call option is lower and the premium for the put option is higher. For example, in going from a contract with a strike price of 112 to one with 115, the premium for the March call option falls from 1 45/64 to 16/64, and the premium for the March put option rises from 19/64 to 1 54/64.

Our understanding of the profit function for option contracts illustrated in Figure 1 helps explain this fact. As we saw in panel (a), a higher price for the underlying financial instrument (in this case a Treasury bond futures contract) relative to the option's exercise price results in higher profits on the call (buy) option. Thus the lower the strike price, the higher the profits on the call option contract and the greater the

premium that investors like Irving are willing to pay. Similarly, we saw in panel (b) that a higher price for the underlying financial instrument relative to the exercise price lowers profits on the put (sell) option, so that a higher strike price increases profits and thus causes the premium to increase.

The second thing you may have noticed in the *Wall Street Journal* entry is that as the period of time over which the option can be exercised (the term to expiration) gets longer, the premiums for both call and put options rise. For example, at a strike price of 112, the premium on the call option increases from 1 45/64 in March to 1 50/64 in April and to 2 28/64 in May. Similarly, the premium on the put option increases from 19/64 in March to 1 43/64 in April and to 2 22/64 in May. The fact that premiums increase with the term to expiration is also explained by the nonlinear profit function for option contracts. As the term to expiration lengthens, there is a greater chance that the price of the underlying financial instrument will be very high or very low by the expiration date. If the price becomes very high and goes well above the exercise price, the call (buy) option will yield a high profit, but if the price becomes very low and goes well below the exercise price, the losses will be small because the owner of the call option will simply decide not to exercise the option. The possibility of greater variability of the underlying financial instrument as the term to expiration lengthens raises profits on average for the call option.

Similar reasoning tells us that the put (sell) option will become more valuable as the term to expiration increases, because the possibility of greater price variability of the underlying financial instrument increases as the term to expiration increases. The greater chance of a low price increases the chance that profits on the put option will be very high. But the greater chance of a high price does not produce substantial losses for the put option, because the owner will again just decide not to exercise the option.

Another way of thinking about this reasoning is to recognize that option contracts have an element of "heads, I win; tails, I don't lose too badly." The greater variability of where the prices might be by the expiration date increases the value of both kinds of options. Since a longer term to the expiration date leads to greater variability of where the prices might be by the expiration date, a longer term to expiration raises the value of the option contract.

The reasoning that we have just developed also explains another important fact about option premiums. When the volatility of the price of the underlying instrument is great, the premiums for both call and put options will be higher. Higher volatility of prices means that for a given expiration date, there will again be greater variability of where the prices might be by the expiration date. The "heads, I win; tails, I don't lose too badly" property of options then means that the greater variability of possible prices by the expiration date increases average profits for the option and thus increases the premium that investors are willing to pay.

Summary

Our analysis of how profits on options are affected by price movements for the underlying financial instrument leads to the following conclusions about the factors that determine the premium on an option contract:

1. The higher the strike price, everything else being equal, the lower the premium on call (buy) options and the higher the premium on put (sell) options.
2. The greater the term to expiration, everything else being equal, the higher the premiums for both call and put options.
3. The greater the volatility of prices of the underlying financial instrument, everything else being equal, the higher the premiums for both call and put options.

The results we have derived here appear in more formal models, such as the Black-Scholes model, which analyze how the premiums on options are priced. You might study such models in finance courses.

Interest-Rate Swaps

A financial contract that allows one party to exchange (swap) a set of interest payments for another set of interest payments owned by another party.

In addition to forwards, futures, and options, financial institutions use one other important financial derivative to manage risk. **Swaps** are financial contracts that obligate each party to the contract to exchange (swap) a set of payments (not assets) it owns for another set of payments owned by another party. There are two basic kinds of swaps: **Currency swaps** involve the exchange of a set of payments in one currency for a set of payments in another currency. **Interest-rate swaps** involve the exchange of one set of interest payments for another set of interest payments, all denominated in the same currency.

Interest-Rate Swap Contracts

Interest-rate swaps are an important tool for managing interest-rate risk, and they first appeared in the United States in 1982, when, as we have seen, there was an increase in the demand for financial instruments that could be used to reduce interest-rate risk. The most common type of interest-rate swap (called the *plain vanilla swap*) specifies (1) the interest rate on the payments that are being exchanged; (2) the type of interest payments (variable or fixed-rate); (3) the amount of **notional principal**, which is the amount on which the interest is being paid; and (4) the time period over which the exchanges continue to be made. There are many other more complicated versions of swaps, including forward swaps and swap options (called *swaptions*), but here we will look only at the plain vanilla swap. Figure 2 illustrates an interest-rate swap between the Midwest Savings Bank and the Friendly Finance Company. Midwest Savings agrees to pay Friendly Finance a fixed rate of 7% on $1 million of notional principal for the next ten years, and Friendly Finance agrees to pay Midwest Savings the one-year Treasury bill rate plus 1% on $1 million of notional principal for the same period. Thus, as shown in Figure 2, every year the Midwest Savings Bank would be paying the Friendly Finance Company 7% on $1 million, while Friendly Finance would be paying Midwest Savings the one-year T-bill rate plus 1% on $1 million.

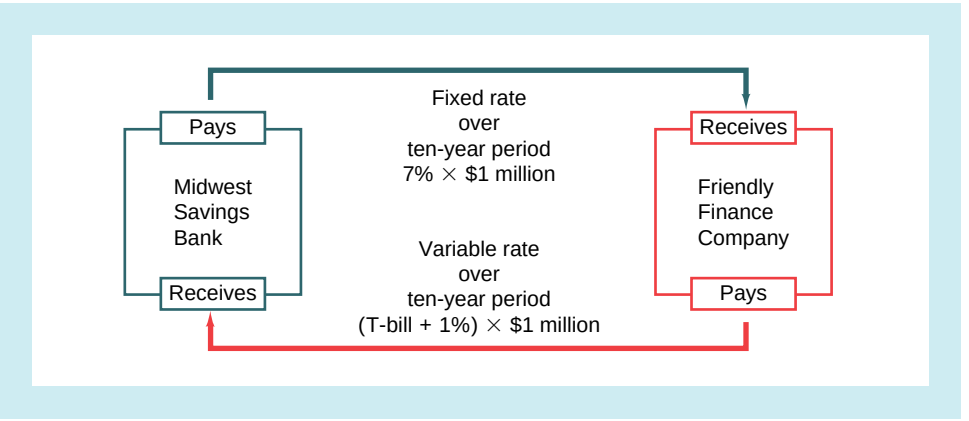


FIGURE 2 Interest-Rate Swap Payments

In this swap arrangement with a notional principal of $1 million and a term of ten years, the Midwest Savings Bank pays a fixed rate of 7% × $1 million to the Friendly Finance Company, which in turn agrees to pay the one-year Treasury bill rate plus 1% × $1 million to the Midwest Savings Bank.

Application

Hedging with Interest-Rate Swaps

You might wonder why these two parties find it advantageous to enter into this swap agreement. The answer is that it may help both of them hedge interest-rate risk.

Suppose that the Midwest Savings Bank, which tends to borrow short-term and then lend long-term in the mortgage market, has $1 million less of rate-sensitive assets than it has of rate-sensitive liabilities. As we learned in Chapter 9, this situation means that as interest rates rise, the rise in the cost of funds (liabilities) is greater than the rise in interest payments it receives on its assets, many of which are fixed-rate. The result of rising interest rates is thus a shrinking of Midwest Savings' net interest margin and a decline in its profitability. As we saw in Chapter 9, to avoid this interest-rate risk, Midwest Savings would like to convert $1 million of its fixed-rate assets into $1 million of rate-sensitive assets, in effect making rate-sensitive assets equal rate-sensitive liabilities, thereby eliminating the gap. This is exactly what happens when it engages in the interest-rate swap. By taking $1 million of its fixed-rate income and exchanging it for $1 million of rate-sensitive Treasury bill income, it has converted income on $1 million of fixed-rate assets into income on $1 million of rate-sensitive assets. Now when interest rates increase, the rise in rate-sensitive income on its assets exactly matches the rise in the rate-sensitive cost of funds on its liabilities, leaving the net interest margin and bank profitability unchanged.

The Friendly Finance Company, which issues long-term bonds to raise funds and uses them to make short-term loans, finds that it is in exactly the opposite situation to Midwest Savings: It has $1 million more of rate-sensitive assets than of rate-sensitive liabilities. It is therefore concerned that a fall in interest rates, which will result in a larger drop in income from its assets than the decline in the cost of funds on its liabilities, will cause a decline in profits. By doing the interest-rate swap, it eliminates this interest-rate risk because it has converted $1 million of rate-sensitive income into $1 million of fixed-rate income. Now the Friendly Finance Company finds that when interest rates fall, the decline in rate-sensitive income is smaller and so is matched by the decline in the rate-sensitive cost of funds on its liabilities, leaving its profitability unchanged.

Advantages of Interest-Rate Swaps

To eliminate interest-rate risk, both the Midwest Savings Bank and the Friendly Finance Company could have rearranged their balance sheets by converting fixed-rate assets into rate-sensitive assets, and vice versa, instead of engaging in an interest-rate swap. However, this strategy would have been costly for both financial institutions for several reasons. The first is that financial institutions incur substantial transaction costs when they rearrange their balance sheets. Second, different financial institutions have informational advantages in making loans to certain customers who may prefer certain maturities. Thus, adjusting the balance sheet to eliminate interest-rate risk might result in a loss of these informational advantages, which the financial institution is unwilling to give up. Interest-rate swaps solve these problems for financial institutions, because in effect, they allow the institutions to convert fixed-rate assets

into rate-sensitive assets without affecting the balance sheet. Large transaction costs are avoided, and the financial institutions can continue to make loans where they have an informational advantage.

We have seen that financial institutions can also hedge interest-rate risk with other financial derivatives such as futures contracts and futures options. Interest-rate swaps have one big advantage over hedging with these other derivatives: They can be written for very long horizons, sometimes as long as 20 years, whereas financial futures and futures options typically have much shorter horizons, not much more than a year. If a financial institution needs to hedge interest-rate risk for a long horizon, financial futures and option markets may not do it much good. Instead it can turn to the swap market.

Disadvantages of Interest-Rate Swaps

Although interest-rate swaps have important advantages that make them very popular with financial institutions, they also have disadvantages that limit their usefulness. Swap markets, like forward markets, can suffer from a lack of liquidity. Let's return to looking at the swap between the Midwest Savings Bank and the Friendly Finance Company. As with a forward contract, it might be difficult for the Midwest Savings Bank to link up with the Friendly Finance Company to arrange the swap. In addition, even if the Midwest Savings Bank could find a counterparty like the Friendly Finance Company, it might not be able to negotiate a good deal because it couldn't find any other institution with which to negotiate.

Swap contracts also are subject to the same default risk that we encountered for forward contracts. If interest rates rise, the Friendly Finance Company would love to get out of the swap contract, because the fixed-rate interest payments it receives are less than it could get in the open market. It might then default on the contract, exposing Midwest Savings to a loss. Alternatively, the Friendly Finance Company could go bust, meaning that the terms of the swap contract would not be fulfilled.

Financial Intermediaries in Interest-Rate Swaps

As we have just seen, financial institutions do have to be aware of the possibility of losses from a default on swaps. As with a forward contract, each party to a swap must have a lot of information about the other party to make sure that the contract is likely to be fulfilled. The need for information about counterparties and the liquidity problems in swap markets could limit the usefulness of these markets. However, as we saw in Chapter 8, when informational and liquidity problems crop up in a market, financial intermediaries come to the rescue. That is exactly what happens in swap markets. Intermediaries such as investment banks and especially large commercial banks have the ability to acquire information cheaply about the creditworthiness and reliability of parties to swap contracts and are also able to match up parties to a swap. Hence large commercial banks and investment banks have set up swap markets in which they act as intermediaries.

Summary

1. Interest-rate forward contracts, which are agreements to sell a debt instrument at a future (forward) point in time, can be used to hedge interest-rate risk. The advantage of forward contracts is that they are flexible, but the disadvantages are that they are subject to default risk and their market is illiquid.
2. A financial futures contract is similar to an interest-rate forward contract, in that it specifies that a debt instrument must be delivered by one party to another on a stated future date. However, it has advantages over a forward contract in that it is not subject to default risk and is more liquid. Forward and futures contracts can

be used by financial institutions to hedge (protect) against interest-rate risk.

3. An option contract gives the purchaser the right to buy (call option) or sell (put option) a security at the exercise (strike) price within a specific period of time. The profit function for options is nonlinear—profits do not always grow by the same amount for a given change in the price of the underlying financial instrument. The nonlinear profit function for options explains why their value (as reflected by the premium paid for them) is negatively related to the exercise price for call options, positively related to the exercise price for put options, positively related to the term to expiration for both call and put options, and positively related to the volatility of the prices of the underlying financial instrument for both call and put options. Financial institutions use futures options to hedge interest-rate risk in a similar fashion to the way they use financial futures and forward contracts. Futures options may be preferred for macro hedges because they suffer from fewer accounting problems than financial futures.
4. Interest-rate swaps involve the exchange of one set of interest payments for another set of interest payments and have default risk and liquidity problems similar to those of forward contracts. As a result, interest-rate swaps often involve intermediaries such as large commercial banks and investment banks that make a market in swaps. Financial institutions find that interest-rate swaps are useful ways to hedge interest-rate risk. Interest-rate swaps have one big advantage over financial futures and options: They can be written for very long horizons.

Key Terms

American option, p. 320
arbitrage, p. 313
call option, p. 322
currency swap, p. 328
European option, p. 320
exercise price (strike price), p. 320
financial derivatives, p. 309
financial futures contract, p. 312
financial futures option (futures option), p. 321
forward contract, p. 310
hedge, p. 309
interest-rate forward contract, p. 310
interest-rate swap, p. 328
long position, p. 309
macro hedge, p. 315
margin requirement, p. 318
marked to market, p. 318
micro hedge, p. 315
notional principal, p. 328
open interest, p. 315
option, p. 320
premium, p. 320
put option, p. 322
short position, p. 309
stock option, p. 321
swap, p. 328

Questions and Problems

Questions marked with an asterisk are answered at the end of the book in an appendix, "Answers to Selected Questions and Problems."

1. If the pension fund you manage expects to have an inflow of $120 million six months from now, what forward contract would you seek to enter into to lock in current interest rates?

*2. If the portfolio you manage is holding $25 million of 8s of 2023 Treasury bonds with a price of 110, what forward contract would you enter into to hedge the interest-rate risk on these bonds over the coming year?

3. If at the expiration date, the deliverable Treasury bond is selling for 101 but the Treasury bond futures contract is selling for 102, what will happen to the futures price? Explain your answer.

*4. If you buy a $100,000 June Treasury bond contract for 108 and the price of the deliverable Treasury bond at the expiration date is 102, what is your profit or loss on the contract?

5. Suppose that the pension you are managing is expecting an inflow of funds of $100 million next year and you want to make sure that you will earn the current

interest rate of 8% when you invest the incoming funds in long-term bonds. How would you use the futures market to do this?

*6. How would you use the options market to accomplish the same thing as in Problem 5? What are the advantages and disadvantages of using an options contract rather than a futures contract?

7. If you buy a put option on a $100,000 Treasury bond futures contract with an exercise price of 95 and the price of the Treasury bond is 120 at expiration, is the contract in the money, out of the money, or at the money? What is your profit or loss on the contract if the premium was $4,000?

*8. Suppose that you buy a call option on a $100,000 Treasury bond futures contract with an exercise price of 110 for a premium of $1,500. If on expiration the futures contract has a price of 111, what is your profit or loss on the contract?

9. Explain why greater volatility or a longer term to maturity leads to a higher premium on both call and put options.

*10. Why does a lower strike price imply that a call option will have a higher premium and a put option a lower premium?

11. If the finance company you manage has a gap of +$5 million (rate-sensitive assets greater than rate-sensitive liabilities by $5 million), describe an interest-rate swap that would eliminate the company's income gap.

*12. If the savings and loan you manage has a gap of −$42 million, describe an interest-rate swap that would eliminate the S&L's income risk from changes in interest rates.

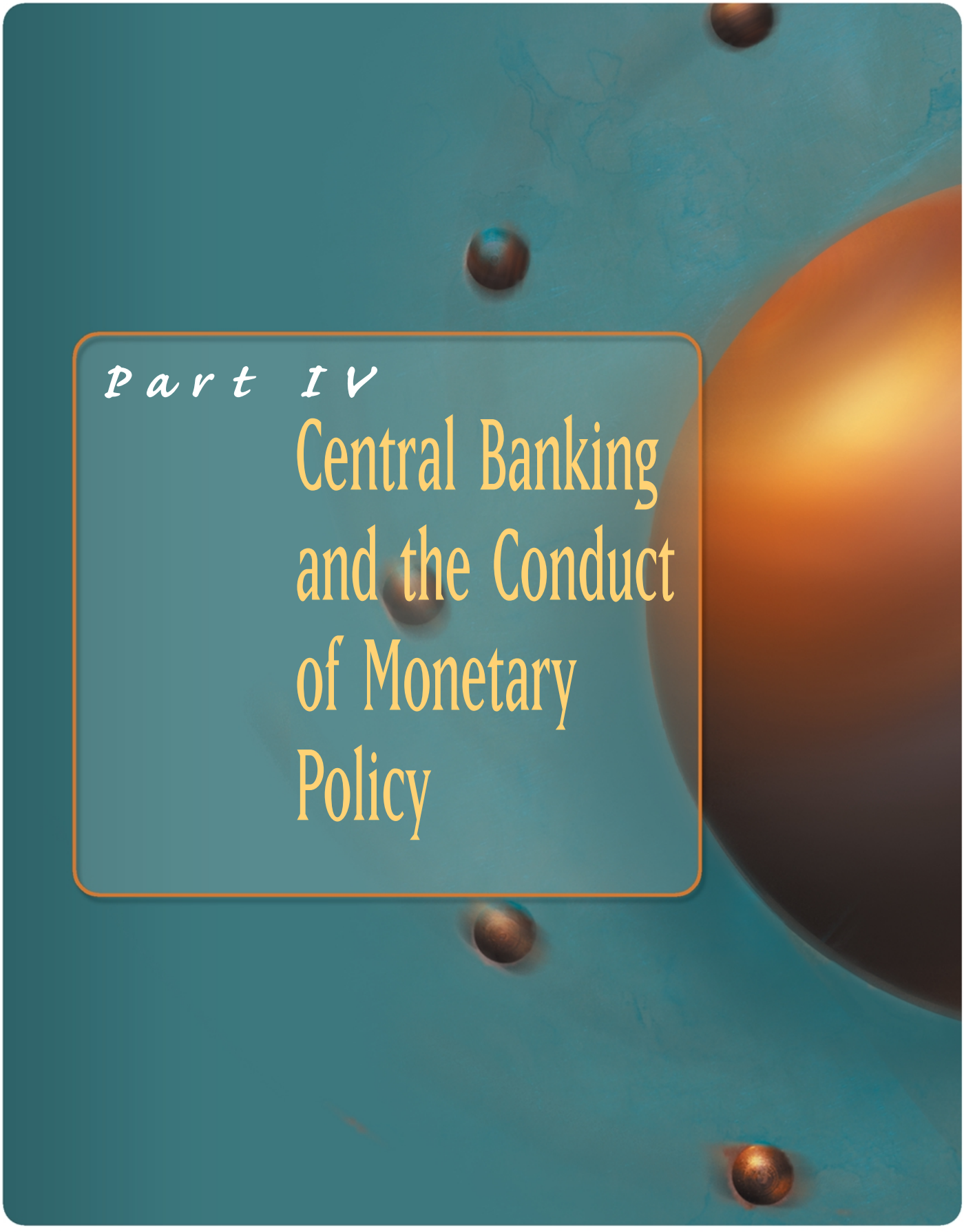
13. If your company has a payment of 200 million euros due one year from now, how would you hedge the foreign exchange risk in this payment with 125,000 euro futures contracts?

*14. If your company has to make a 10 million euro payment to a German company in June, three months from now, how would you hedge the foreign exchange risk in this payment with a 125,000 euro futures contract?

15. Suppose that your company will be receiving 30 million euros six months from now and the euro is currently selling for 1 euro per dollar. If you want to hedge the foreign exchange risk in this payment, what kind of forward contract would you want to enter into?

Web Exercises

1. We have discussed the various stock markets in detail throughout this text. Another market that is less well known is the New York Mercantile Exchange. Here contracts on a wide variety of commodities are traded on a daily basis. Go to www.nymex.com/welcome/info_01.htm and read the discussion explaining the origin and purpose of the mercantile exchange. Write a one-page summary discussing this material.
2. The following site can be used to demonstrate how the features of an option affect the option's prices. Go to www.intrepid.com/~robertl/option-pricer4.html. What happens to the price of an option under each of the following situations?
   a. The strike price increases
   b. Interest rates increase
   c. Volatility increases
   d. The time until the option matures increases



Part IV

Central Banking and the Conduct of Monetary Policy

Chapter 14

Structure of Central Banks and the Federal Reserve System

PREVIEW

Among the most important players in financial markets throughout the world are central banks, the government authorities in charge of monetary policy. Central banks' actions affect interest rates, the amount of credit, and the money supply, all of which have direct impacts not only on financial markets, but also on aggregate output and inflation. To understand the role that central banks play in financial markets and the overall economy, we need to understand how these organizations work. Who controls central banks and determines their actions? What motivates their behavior? Who holds the reins of power?

In this chapter, we look at the institutional structure of major central banks, and focus particularly on the Federal Reserve System, the most important central bank in the world. We start by focusing on the formal institutional structure of the Fed and then examine the more relevant informal structure that determines where the true power within the Federal Reserve System lies. By understanding who makes the decisions, we will have a better idea of how they are made. We then look at several other major central banks and see how they are organized. With this information, we will be more able to comprehend the actual conduct of monetary policy described in the following chapters.

Origins of the Federal Reserve System

Of all the central banks in the world, the Federal Reserve System probably has the most unusual structure. To understand why this structure arose, we must go back to before 1913, when the Federal Reserve System was created.

Before the twentieth century, a major characteristic of American politics was the fear of centralized power, as seen in the checks and balances of the Constitution and the preservation of states' rights. This fear of centralized power was one source of the American resistance to the establishment of a central bank (see Chapter 10). Another source was the traditional American distrust of moneyed interests, the most prominent symbol of which was a central bank. The open hostility of the American public to the existence of a central bank resulted in the demise of the first two experiments in central banking, whose function was to police the banking system: The First Bank of the United States was disbanded in 1811, and the national charter of the Second

Box 1: Inside the Fed

The Political Genius of the Founders of the Federal Reserve System

The history of the United States has been one of public hostility to banks and especially to a central bank. How were the politicians who founded the Federal Reserve able to design a system that has become one of the most prestigious institutions in the United States?

The answer is that the founders recognized that if power was too concentrated in either Washington or New York, cities that Americans often love to hate, an American central bank might not have enough public support to operate effectively. They thus decided to set up a decentralized system with 12 Federal Reserve banks spread throughout the country to make sure that all regions of the country were represented in monetary policy deliberations. In addition, they made the Federal Reserve banks quasi-private institutions overseen by directors from the private sector living in that district who represent views from that region and are in close contact with the president of the Federal Reserve bank. The unusual structure of the Federal Reserve System has promoted a concern in the Fed with regional issues as is evident in Federal Reserve bank publications. Without this unusual structure, the Federal Reserve System might have been far less popular with the public, making the institution far less effective.

Bank of the United States expired in 1836 after its renewal was vetoed in 1832 by President Andrew Jackson.

The termination of the Second Bank's national charter in 1836 created a severe problem for American financial markets, because there was no lender of last resort who could provide reserves to the banking system to avert a bank panic. Hence in the nineteenth and early twentieth centuries, nationwide bank panics became a regular event, occurring every twenty years or so, culminating in the panic of 1907. The 1907 panic resulted in such widespread bank failures and such substantial losses to depositors that the public was finally convinced that a central bank was needed to prevent future panics.

The hostility of the American public to banks and centralized authority created great opposition to the establishment of a single central bank like the Bank of England. Fear was rampant that the moneyed interests on Wall Street (including the largest corporations and banks) would be able to manipulate such an institution to gain control over the economy and that federal operation of the central bank might result in too much government intervention in the affairs of private banks. Serious disagreements existed over whether the central bank should be a private bank or a government institution. Because of the heated debates on these issues, a compromise was struck. In the great American tradition, Congress wrote an elaborate system of checks and balances into the Federal Reserve Act of 1913, which created the Federal Reserve System with its 12 regional Federal Reserve banks (see Box 1).

Formal Structure of the Federal Reserve System

The formal structure of the Federal Reserve System was intended by writers of the Federal Reserve Act to diffuse power along regional lines, between the private sector and the government, and among bankers, businesspeople, and the public. This initial diffusion of power has resulted in the evolution of the Federal Reserve System to

www.federalreserve.gov/pubs/frseries/frseri.htm
Information on the structure of the Federal Reserve System.

include the following entities: the **Federal Reserve banks**, the **Board of Governors of the Federal Reserve System**, the **Federal Open Market Committee (FOMC)**, the Federal Advisory Council, and around 4,800 member commercial banks. Figure 1 outlines the relationships of these entities to one another and to the three policy tools of the Fed (open market operations, the discount rate, and reserve requirements) discussed in Chapters 15 to 17.

Federal Reserve Banks

Each of the 12 Federal Reserve districts has one main Federal Reserve bank, which may have branches in other cities in the district. The locations of these districts, the Federal Reserve banks, and their branches are shown in Figure 2. The three largest

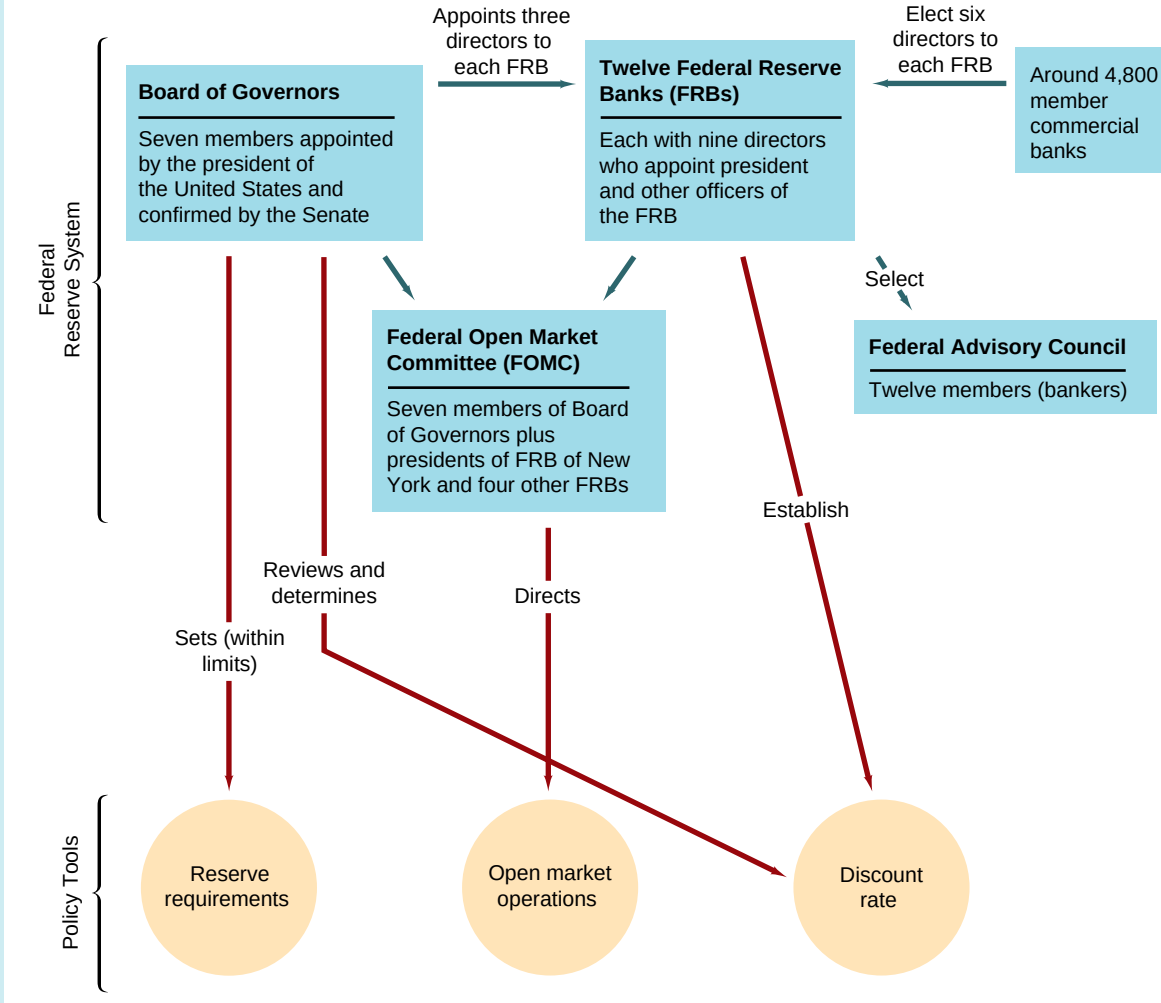


FIGURE 1 Formal Structure and Allocation of Policy Tools in the Federal Reserve

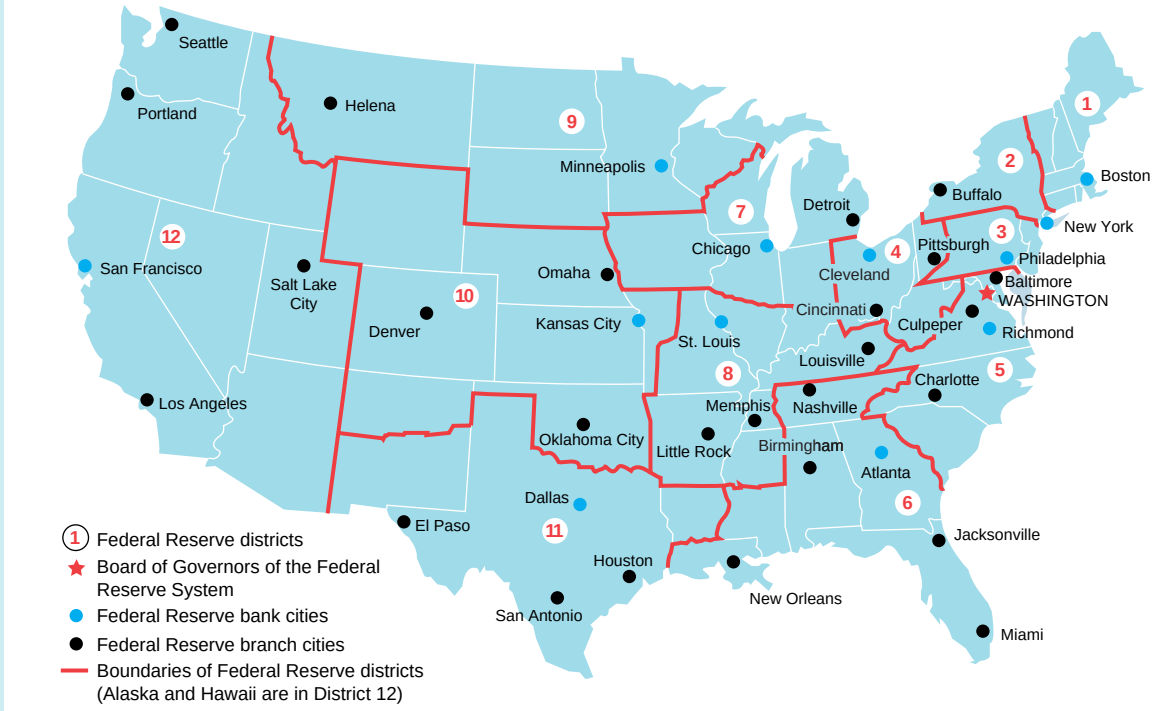


FIGURE 2 Federal Reserve System

Source: Federal Reserve *Bulletin.*

www.federalreserve.gov/otherfrb.htm

Addresses and phone numbers of Federal Reserve banks, branches, and RCPCs and links to the main pages of the 12 reserve banks and Board of Governors.

Federal Reserve banks in terms of assets are those of New York, Chicago, and San Francisco—combined they hold over 50% of the assets (discount loans, securities, and other holdings) of the Federal Reserve System. The New York bank, with around one-quarter of the assets, is the most important of the Federal Reserve banks (see Box 2).

Each of the Federal Reserve banks is a quasi-public (part private, part government) institution owned by the private commercial banks in the district that are members of the Federal Reserve System. These member banks have purchased stock in their district Federal Reserve bank (a requirement of membership), and the dividends paid by that stock are limited by law to 6% annually. The member banks elect six directors for each district bank; three more are appointed by the Board of Governors. Together, these nine directors appoint the president of the bank (subject to the approval of the Board of Governors).

The directors of a district bank are classified into three categories, A, B, and C: The three A directors (elected by the member banks) are professional bankers, and the three B directors (also elected by the member banks) are prominent leaders from industry, labor, agriculture, or the consumer sector. The three C directors, who are appointed by the Board of Governors to represent the public interest, are not allowed to be officers, employees, or stockholders of banks. This design for choosing directors was intended by the framers of the Federal Reserve Act to ensure that the directors of each Federal Reserve bank would reflect all constituencies of the American public.

Box 2: Inside the Fed

The Special Role of the Federal Reserve Bank of New York

The Federal Reserve Bank of New York plays a special role in the Federal Reserve System for several reasons. First, its district contains many of the largest commercial banks in the United States, the safety and soundness of which are paramount to the health of the U.S. financial system. The Federal Reserve Bank of New York conducts examinations of bank holding companies and state-chartered banks in its district, making it the supervisor of some of the most important financial institutions in our financial system. Not surprisingly, given this responsibility, the bank supervision group is one of the largest units of the New York Fed and is by far the largest bank supervision group in the Federal Reserve System.

The second reason for the New York Fed's special role is its active involvement in the bond and foreign exchange markets. The New York Fed houses the open market desk, which conducts open market operations—the purchase and sale of bonds—that determine the amount of reserves in the banking system. Because of this involvement in the Treasury securities market, as well as its walking-distance location near the New York and American Stock Exchanges, the officials at the Federal Reserve Bank of New York are in constant contact with the major domestic financial markets in the United States. In addition, the Federal Reserve Bank of New York also houses the foreign exchange desk, which conducts foreign exchange interventions on behalf of the Federal Reserve System and the U.S. Treasury. Its involvement in these financial markets means that the New York Fed is an important source of information on what is happening in domestic and foreign financial markets, particularly during crisis periods, as well as a liaison between officials in the Federal Reserve System and private participants in the markets.

The third reason for the Federal Reserve Bank of New York's prominence is that it is the only Federal Reserve bank to be a member of the Bank for International Settlements (BIS). Thus the president of the New York Fed, along with the chairman of the Board of Governors, represents the Federal Reserve System in its regular monthly meetings with other major central bankers at the BIS. This close contact with foreign central bankers and interaction with foreign exchange markets means that the New York Fed has a special role in international relations, both with other central bankers and with private market participants. Adding to its prominence in international circles is that the New York Fed is the repository for over $100 billion of the world's gold, an amount greater than the gold at Fort Knox.

Finally, the president of the Federal Reserve Bank of New York is the only permanent member of the FOMC among the Federal Reserve bank presidents, serving as the vice-chairman of the committee. Thus he and the chairman and vice-chairman of the Board of Governors are the three most important officials in the Federal Reserve System.

The 12 Federal Reserve banks perform the following functions:

- Clear checks
- Issue new currency
- Withdraw damaged currency from circulation
- Administer and make discount loans to banks in their districts
- Evaluate proposed mergers and applications for banks to expand their activities
- Act as liaisons between the business community and the Federal Reserve System
- Examine bank holding companies and state-chartered member banks
- Collect data on local business conditions
- Use their staffs of professional economists to research topics related to the conduct of monetary policy

The 12 Federal Reserve banks are involved in monetary policy in several ways:

1. Their directors "establish" the discount rate (although the discount rate in each district is reviewed and determined by the Board of Governors).
2. They decide which banks, member and nonmember alike, can obtain discount loans from the Federal Reserve bank.
3. Their directors select one commercial banker from each bank's district to serve on the Federal Advisory Council, which consults with the Board of Governors and provides information that helps in the conduct of monetary policy.
4. Five of the 12 bank presidents each have a vote in the Federal Open Market Committee, which directs **open market operations** (the purchase and sale of government securities that affect both interest rates and the amount of reserves in the banking system). As explained in Box 2, the president of the New York Fed always has a vote in the FOMC, making it the most important of the banks; the other four votes allocated to the district banks rotate annually among the remaining 11 presidents.

Member Banks

All *national banks* (commercial banks chartered by the Office of the Comptroller of the Currency) are required to be members of the Federal Reserve System. Commercial banks chartered by the states are not required to be members, but they can choose to join. Currently, around one-third of the commercial banks in the United States are members of the Federal Reserve System, having declined from a peak figure of 49% in 1947.

Before 1980, only member banks were required to keep reserves as deposits at the Federal Reserve banks. Nonmember banks were subject to reserve requirements determined by their states, which typically allowed them to hold much of their reserves in interest-bearing securities. Because no interest is paid on reserves deposited at the Federal Reserve banks, it was costly to be a member of the system, and as interest rates rose, the relative cost of membership rose, and more and more banks left the system.

This decline in Fed membership was a major concern of the Board of Governors (one reason was that it lessened the Fed's control over the money supply, making it more difficult for the Fed to conduct monetary policy). The chairman of the Board of Governors repeatedly called for new legislation requiring all commercial banks to be members of the Federal Reserve System. One result of the Fed's pressure on Congress was a provision in the Depository Institutions Deregulation and Monetary Control Act of 1980: All depository institutions became subject (by 1987) to the same requirements to keep deposits at the Fed, so member and nonmember banks would be on an equal footing in terms of reserve requirements. In addition, all depository institutions were given access to the Federal Reserve facilities, such as the discount window (discussed in Chapter 17) and Fed check clearing, on an equal basis. These provisions ended the decline in Fed membership and reduced the distinction between member and nonmember banks.

Board of Governors of the Federal Reserve System

At the head of the Federal Reserve System is the seven-member Board of Governors, headquartered in Washington, D.C. Each governor is appointed by the president of the United States and confirmed by the Senate. To limit the president's control over the Fed and insulate the Fed from other political pressures, the governors serve one

nonrenewable 14-year term, with one governor's term expiring every other January.[1] The governors (many are professional economists) are required to come from different Federal Reserve districts to prevent the interests of one region of the country from being overrepresented. The chairman of the Board of Governors is chosen from among the seven governors and serves a four-year term. It is expected that once a new chairman is chosen, the old chairman resigns from the Board of Governors, even if there are many years left to his or her term as a governor.

www.federalreserve.gov/bios/1199member.pdf

Lists all the members of the Board of Governors of the Federal Reserve since its inception.

The Board of Governors is actively involved in decisions concerning the conduct of monetary policy. All seven governors are members of the FOMC and vote on the conduct of open market operations. Because there are only 12 voting members on this committee (seven governors and five presidents of the district banks), the Board has the majority of the votes. The Board also sets reserve requirements (within limits imposed by legislation) and effectively controls the discount rate by the "review and determination" process, whereby it approves or disapproves the discount rate "established" by the Federal Reserve banks. The chairman of the Board advises the president of the United States on economic policy, testifies in Congress, and speaks for the Federal Reserve System to the media. The chairman and other governors may also represent the United States in negotiations with foreign governments on economic matters. The Board has a staff of professional economists (larger than those of individual Federal Reserve banks), which provides economic analysis that the board uses in making its decisions. (Box 3 discusses the role of the research staff.)

Through legislation, the Board of Governors has often been given duties not directly related to the conduct of monetary policy. In the past, for example, the Board set the maximum interest rates payable on certain types of deposits under Regulation Q. (After 1986, ceilings on time deposits were eliminated, but there is still a restriction on paying any interest on business demand deposits.) Under the Credit Control Act of 1969 (which expired in 1982), the Board had the ability to regulate and control credit once the president of the United States approved. The Board of Governors also sets margin requirements, the fraction of the purchase price of securities that has to be paid for with cash rather than borrowed funds. It also sets the salary of the president and all officers of each Federal Reserve bank and reviews each bank's budget. Finally, the Board has substantial bank regulatory functions: It approves bank mergers and applications for new activities, specifies the permissible activities of bank holding companies, and supervises the activities of foreign banks in the United States.

Federal Open Market Committee (FOMC)

The FOMC usually meets eight times a year (about every six weeks) and makes decisions regarding the conduct of open market operations, which influence the monetary base. Indeed, the FOMC is often referred to as the "Fed" in the press: for example, when the media say that the Fed is meeting, they actually mean that the FOMC is meeting. The committee consists of the seven members of the Board of Governors, the president of the Federal Reserve Bank of New York, and the presidents of four other Federal Reserve banks. The chairman of the Board of Governors also presides as the chairman of the FOMC. Even though only the presidents of five of the Federal Reserve

[1] Although technically the governor's term is nonrenewable, a governor can resign just before the term expires and then be reappointed by the president. This explains how one governor, William McChesney Martin Jr., served for 28 years. Since Martin, the chairman from 1951 to 1970, retired from the board in 1970, the practice of extending a governor's term beyond 14 years has become a rarity.

Box 3: Inside the Fed

The Role of the Research Staff

The Federal Reserve System is the largest employer of economists not just in the United States, but in the world. The system's research staff has around 1,000 people, about half of whom are economists. Of these 500 economists, 250 are at the Board of Governors, 100 are at the Federal Reserve Bank of New York, and the remainder are at the other Federal Reserve banks. What do all these economists do?

The most important task of the Fed's economists is to follow the incoming data from government agencies and private sector organizations on the economy and provide guidance to the policymakers on where the economy may be heading and what the impact of monetary policy actions on the economy might be. Before each FOMC meeting, the research staff at each Federal Reserve bank briefs its president and the senior management of the bank on its forecast for the U.S. economy and the issues that are likely to be discussed at the meeting. The research staff also provides briefing materials or a formal briefing on the economic outlook for the bank's region, something that each president discusses at the FOMC meeting. Meanwhile, at the Board of Governors, economists maintain a large econometric model (a model whose equations are estimated with statistical procedures) that helps them produce their forecasts of the national economy, and they too brief the governors on the national economic outlook.

The research staffers at the banks and the board also provide support for the bank supervisory staff, tracking developments in the banking sector and other financial markets and institutions and providing bank examiners with technical advice that they might need in the course of their examinations. Because the Board of Governors has to decide on whether to approve bank mergers, the research staff at both the board and the bank in whose district the merger is to take place prepare information on what effect the proposed merger might have on the competitive environment. To assure compliance with the Community Reinvestment Act, economists also analyze a bank's performance in its lending activities in different communities.

Because of the increased influence of developments in foreign countries on the U.S. economy, the members of the research staff, particularly at the New York Fed and the Board, produce reports on the major foreign economies. They also conduct research on developments in the foreign exchange market because of its growing importance in the monetary policy process and to support the activities of the foreign exchange desk. Economists also help support the operation of the open market desk by projecting reserve growth and the growth of the monetary aggregates.

Staff economists also engage in basic research on the effects of monetary policy on output and inflation, developments in the labor markets, international trade, international capital markets, banking and other financial institutions, financial markets, and the regional economy, among other topics. This research is published widely in academic journals and in Reserve bank publications. (Federal Reserve bank reviews are a good source of supplemental material for money and banking students.)

Another important activity of the research staff primarily at the Reserve banks is in the public education area. Staff economists are called on frequently to make presentations to the board of directors at their banks or to make speeches to the public in their district.

www.federalreserve.gov/fomc

Find general information on the FOMC, its schedule of meetings, statements, minutes, and transcripts; information on its members, and the "beige book."

banks are voting members of the FOMC, the other seven presidents of the district banks attend FOMC meetings and participate in discussions. Hence they have some input into the committee's decisions.

Because open market operations are the most important policy tool that the Fed has for controlling the money supply, the FOMC is necessarily the focal point for policymaking in the Federal Reserve System. Although reserve requirements and the discount rate are not actually set by the FOMC, decisions in regard to these policy tools

are effectively made there. The FOMC does not actually carry out securities purchases or sales. Rather it issues directives to the trading desk at the Federal Reserve Bank of New York, where the manager for domestic open market operations supervises a roomful of people who execute the purchases and sales of the government or agency securities. The manager communicates daily with the FOMC members and their staffs concerning the activities of the trading desk.

The FOMC Meeting

The FOMC meeting takes place in the boardroom on the second floor of the main building of the Board of Governors in Washington. The seven governors and the 12 Reserve Bank presidents, along with the secretary of the FOMC, the Board's director of the Research and Statistics Division and his deputy, and the directors of the Monetary Affairs and International Finance Divisions, sit around a massive conference table. Although only five of the Reserve Bank presidents have voting rights on the FOMC at any given time, all actively participate in the deliberations. Seated around the sides of the room are the directors of research at each of the Reserve banks and other senior board and Reserve Bank officials, who, by tradition, do not speak at the meeting.

Except for the meetings prior to the February and July testimony by the chairman of the Board of Governors before Congress, the meeting starts on Tuesday at 9:00 A.M. sharp with a quick approval of the minutes of the previous meeting of the FOMC. The first substantive agenda item is the report by the manager of system open market operations on foreign currency and domestic open market operations and other issues related to these topics. After the governors and Reserve Bank presidents finish asking questions and discussing these reports, a vote is taken to ratify them.

The next stage in the meeting is a presentation of the Board staff's national economic forecast, referred to as the "green book" forecast (see Box 4), by the director of the Research and Statistics Division at the board. After the governors and Reserve Bank presidents have queried the division director about the forecast, the so-called *go-round* occurs: Each bank president presents an overview of economic conditions in his or her district and the bank's assessment of the national outlook, and each governor, except for the chairman, gives a view of the national outlook. By tradition, remarks avoid the topic of monetary policy at this time.

After a coffee break, everyone returns to the boardroom and the agenda turns to current monetary policy and the domestic policy directive. The Board's director of the Monetary Affairs Division then leads off the discussion by outlining the different scenarios for monetary policy actions outlined in the blue book (see Box 4) and may describe an issue relating to how monetary policy should be conducted. After a question-and-answer period, the chairman (currently Alan Greenspan) sets the stage for the following discussion by presenting his views on the state of the economy and then typically makes a recommendation for what monetary policy action should be taken. Then each of the FOMC members as well as the nonvoting bank presidents expresses his or her views on monetary policy, and the chairman summarizes the discussion and proposes specific wording for the directive on the federal funds rate target transmitted to the open market desk. The secretary of the FOMC formally reads the proposed statement, and the members of the FOMC vote.[2]

[2]The decisions expressed in the directive may not be unanimous, and the dissenting views are made public. However, except in rare cases, the chairman's vote is always on the winning side.

Box 4: Inside the Fed

Green, Blue, and Beige

What Do These Colors Mean at the Fed? Three research documents play an important role in the monetary policy process and at Federal Open Market Committee meetings. The national forecast for the next two years, generated by the Federal Reserve Board of Governors' Research and Statistics Division, is placed between green covers and is thus known as the "green book." It is provided to all who attend the FOMC meeting. The "blue book," in blue covers, also provided to all participants at the FOMC meeting, contains the projections for the monetary aggregates prepared by the Monetary Affairs Division at the Board of Governors and typically also presents three alternative scenarios for the stance of monetary policy (labeled A, B, and C). The "beige book," with beige covers, is produced by the Reserve banks and details evidence gleaned either from surveys or from talks with key businesses and financial institutions on the state of the economy in each of the Federal Reserve districts. This is the only one of the three books that is distributed publicly, and it often receives a lot of attention in the press.

Then there is an informal buffet lunch, and while eating, the participants hear a presentation on the latest developments in Congress on banking legislation and other legislation relevant to the Federal Reserve. Around 2:15 P.M., the meeting breaks up and a public announcement is made about the outcome of the meeting: whether the target federal funds rate and discount rate have been raised, lowered, or left unchanged, and an assessment of the "balance of risks" in the future, whether toward higher inflation or toward a weaker economy.[3] The postmeeting announcement is an innovation initiated in 1994. Before then, no such announcement was made, and the markets had to guess what policy action was taken. The decision to announce this information was a step in the direction of greater openness by the Fed.

Informal Structure of the Federal Reserve System

The Federal Reserve Act and other legislation give us some idea of the formal structure of the Federal Reserve System and who makes decisions at the Fed. What is written in black and white, however, does not necessarily reflect the reality of the power and decision-making structure.

As envisioned in 1913, the Federal Reserve System was to be a highly decentralized system designed to function as 12 separate, cooperating central banks. In the original plan, the Fed was not responsible for the health of the economy through its control of the money supply and its ability to affect interest rates. Over time, it has

[3]The meetings before the February and July chairman's testimony before Congress, in which the *Monetary Report to Congress* is presented, have a somewhat different format. Rather than start Tuesday morning at 9:00 A.M. like the other meetings, they start in the afternoon on Tuesday and go over to Wednesday, with the usual announcement around 2:15 P.M. These longer meetings consider the longer-term economic outlook as well as the current conduct of open market operations.

acquired the responsibility for promoting a stable economy, and this responsibility has caused the Federal Reserve System to evolve slowly into a more unified central bank.

The framers of the Federal Reserve Act of 1913 intended the Fed to have only one basic tool of monetary policy: the control of discount loans to member banks. The use of open market operations as a tool for monetary control was not yet well understood, and reserve requirements were fixed by the Federal Reserve Act. The discount tool was to be controlled by the joint decision of the Federal Reserve banks and the Federal Reserve Board (which later became the Board of Governors), so that both would share equally in the determination of monetary policy. However, the Board's ability to "review and determine" the discount rate effectively allowed it to dominate the district banks in setting this policy.

Banking legislation during the Great Depression years centralized power within the newly created Board of Governors by giving it effective control over the remaining two tools of monetary policy, open market operations and changes in reserve requirements. The Banking Act of 1933 granted the FOMC authority to determine open market operations, and the Banking Act of 1935 gave the Board the majority of votes in the FOMC. The Banking Act of 1935 also gave the Board authority to change reserve requirements.

Since the 1930s, then, the Board of Governors has acquired the reins of control over the tools for conducting monetary policy. In recent years, the power of the Board has become even greater. Although the directors of a Federal Reserve bank choose its president with the approval of the Board, the Board sometimes suggests a choice (often a professional economist) for president of a Federal Reserve bank to the directors of the bank, who then often follow the Board's suggestions. Since the Board sets the salary of the bank's president and reviews the budget of each Federal Reserve bank, it has further influence over the district banks' activities.

If the Board of Governors has so much power, what power do the Federal Advisory Council and the "owners" of the Federal Reserve banks—the member banks—actually have within the Federal Reserve System? The answer is almost none. Although member banks own stock in the Federal Reserve banks, they have none of the usual benefits of ownership. First, they have no claim on the earnings of the Fed and get paid only a 6% annual dividend, regardless of how much the Fed earns. Second, they have no say over how their property is used by the Federal Reserve System, in contrast to stockholders of private corporations. Third, usually only a single candidate for each of the six A and B directorships is "elected" by the member banks, and this candidate is frequently suggested by the president of the Federal Reserve bank (who, in turn, is approved by the Board of Governors). The net result is that member banks are essentially frozen out of the political process at the Fed and have little effective power. Fourth, as its name implies, the Federal Advisory Council has only an advisory capacity and has no authority over Federal Reserve policymaking. Although the member bank "owners" do not have the usual power associated with being a stockholder, they do play an important, but subtle, role in the Federal Reserve System (see Box 5).

A fair characterization of the Federal Reserve System as it has evolved is that it functions as a central bank, headquartered in Washington, D.C., with branches in 12 cities. Because all aspects of the Federal Reserve System are essentially controlled by the Board of Governors, who controls the Board? Although the chairman of the Board of Governors does not have legal authority to exercise control over this body, he effectively does so through his ability to act as spokesperson for the Fed and negotiate with

Box 5: Inside the Fed

The Role of Member Banks in the Federal Reserve System

Although the member bank stockholders in each Federal Reserve bank have little direct power in the Federal Reserve System, they do play an important role. Their six representatives on the board of directors of each bank have a major oversight function. Along with the three public interest directors, they oversee the audit process for the Federal Reserve bank, making sure it is being run properly, and also share their management expertise with the senior management of the bank. Because they vote on recommendations by each bank to raise, lower, or maintain the discount rate at its current level, they engage in discussions about monetary policy and transmit their private sector views to the president and senior management of the bank. They also get to understand the inner workings of the Federal Reserve banks and the system so that they can help explain the position of the Federal Reserve to their contacts in the private and political sectors. Advisory councils like the Federal Advisory Council and others that are often set up by the district banks—for example, the Small Business and Agriculture Advisory Council and the Thrift Advisory Council at the New York Fed—are a conduit for the private sector to express views on both the economy and the state of the banking system.

So even though the owners of the Reserve banks do not have the usual voting rights, they are important to the Federal Reserve System, because they make sure it does not get out of touch with the needs and opinions of the private sector.

Congress and the president of the United States. He also exercises control by setting the agenda of Board and FOMC meetings. For example, the fact that the agenda at the FOMC has the chairman speak first about monetary policy enables him to have greater influence over what the policy action will be. The chairman also influences the Board through the force of stature and personality. Chairmen of the Board of Governors (including Marriner S. Eccles, William McChesney Martin Jr., Arthur Burns, Paul A. Volcker, and Alan Greenspan) have typically had strong personalities and have wielded great power.

The chairman also exercises power by supervising the Board's staff of professional economists and advisers. Because the staff gathers information for the Board and conducts the analyses that the Board uses in its decisions, it also has some influence over monetary policy. In addition, in the past, several appointments to the Board itself have come from within the ranks of its professional staff, making the chairman's influence even farther-reaching and longer-lasting than a four-year term.

The informal power structure of the Fed, in which power is centralized in the chairman of the Board of Governors, is summarized in Figure 3.

How Independent Is the Fed?

When we look, in the next four chapters, at how the Federal Reserve conducts monetary policy, we will want to know why it decides to take certain policy actions but not others. To understand its actions, we must understand the incentives that motivate the Fed's behavior. How free is the Fed from presidential and congressional pressures? Do economic, bureaucratic, or political considerations guide it? Is the Fed truly independent of outside pressures?

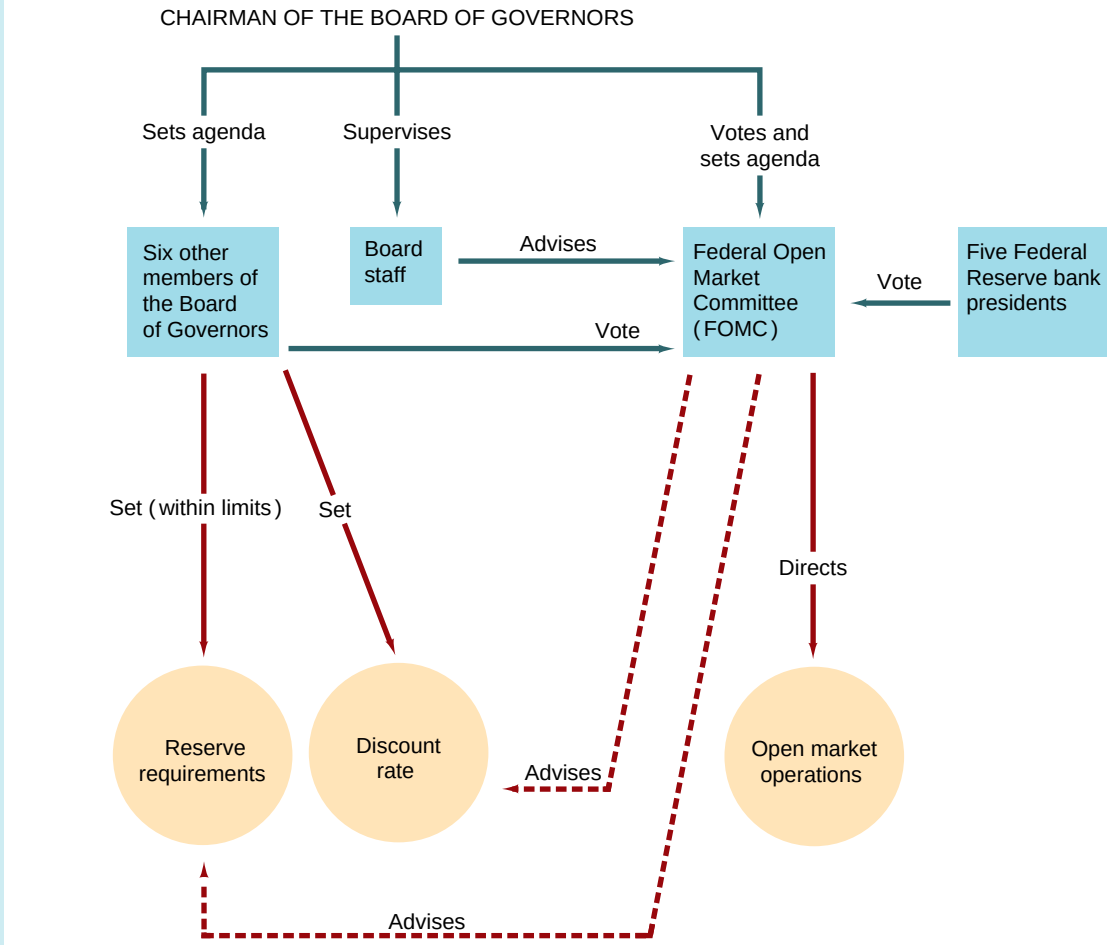


FIGURE 3 Informal Power Structure of the Federal Reserve System

Stanley Fischer, who was a professor at MIT and then the Deputy Managing Director of the International Monetary Fund, has defined two different types of independence of central banks: **instrument independence**, the ability of the central bank to set monetary policy instruments, and **goal independence**, the ability of the central bank to set the goals of monetary policy. The Federal Reserve has both types of independence and is remarkably free of the political pressures that influence other government agencies. Not only are the members of the Board of Governors appointed for a 14-year term (and so cannot be ousted from office), but also the term is technically not renewable, eliminating some of the incentive for the governors to curry favor with the president and Congress.

Probably even more important to its independence from the whims of Congress is the Fed's independent and substantial source of revenue from its holdings of securities

and, to a lesser extent, from its loans to banks. In recent years, for example, the Fed has had net earnings after expenses of around $28 billion per year—not a bad living if you can find it! Because it returns the bulk of these earnings to the Treasury, it does not get rich from its activities, but this income gives the Fed an important advantage over other government agencies: It is not subject to the appropriations process usually controlled by Congress. Indeed, the General Accounting Office, the auditing agency of the federal government, cannot audit the monetary policy or foreign exchange market functions of the Federal Reserve. Because the power to control the purse strings is usually synonymous with the power of overall control, this feature of the Federal Reserve System contributes to its independence more than any other factor.

Yet the Federal Reserve is still subject to the influence of Congress, because the legislation that structures it is written by Congress and is subject to change at any time. When legislators are upset with the Fed's conduct of monetary policy, they frequently threaten to take control of the Fed's finances and force it to submit a budget request like other government agencies. A recent example was the call by Senators Dorgan and Reid in 1996 for Congress to have budgetary authority over the nonmonetary activities of the Federal Reserve. This is a powerful club to wield, and it certainly has some effect in keeping the Fed from straying too far from congressional wishes.

Congress has also passed legislation to make the Federal Reserve more accountable for its actions. In 1975, Congress passed House Concurrent Resolution 133, which requires the Fed to announce its objectives for the growth rates of the monetary aggregates. In the Full Employment and Balanced Growth Act of 1978 (the Humphrey-Hawkins Act), the Fed is required to explain how these objectives are consistent with the economic plans of the president of the United States.

The president can also influence the Federal Reserve. Because congressional legislation can affect the Fed directly or affect its ability to conduct monetary policy, the president can be a powerful ally through his influence on Congress. Second, although ostensibly a president might be able to appoint only one or two members to the Board of Governors during each presidential term, in actual practice the president appoints members far more often. One reason is that most governors do not serve out a full 14-year term. (Governors' salaries are substantially below what they can earn in the private sector, thus providing an incentive for them to take private sector jobs before their term expires.) In addition, the president is able to appoint a new chairman of the Board of Governors every four years, and a chairman who is not reappointed is expected to resign from the board so that a new member can be appointed.

The power that the president enjoys through his appointments to the Board of Governors is limited, however. Because the term of the chairman is not necessarily concurrent with that of the president, a president may have to deal with a chairman of the Board of Governors appointed by a previous administration. Alan Greenspan, for example, was appointed chairman in 1987 by President Ronald Reagan and was reappointed to another term by another Republican president, George Bush. When Bill Clinton, a Democrat, became president in 1993, Greenspan had several years left to his term. Clinton was put under tremendous pressure to reappoint Greenspan when his term expired and did so in 1996 and again in 2000, even though Greenspan is a Republican.[4]

[4]Similarly, William McChesney Martin, Jr., the chairman from 1951 to 1970, was appointed by President Truman (Dem.) but was reappointed by Presidents Eisenhower (Rep.), Kennedy (Dem.), Johnson (Dem.), and Nixon (Rep.). Also Paul Volcker, the chairman from 1979 to 1987, was appointed by President Carter (Dem.) but was reappointed by President Reagan (Rep.).

You can see that the Federal Reserve has extraordinary independence for a government agency and is one of the most independent central banks in the world. Nonetheless, the Fed is not free from political pressures. Indeed, to understand the Fed's behavior, we must recognize that public support for the actions of the Federal Reserve plays a very important role.[5]

Structure and Independence of Foreign Central Banks

In contrast to the Federal Reserve System, which is decentralized into 12 privately owned district banks, central banks in other industrialized countries consist of one centralized unit that is owned by the government. Here we examine the structure and degree of independence of four of the most important foreign central banks: the Bank of Canada, the Bank of England, the Bank of Japan, and the European Central Bank.

Bank of Canada

www.bank-banque-canada.ca/
The website for the Bank of Canada.

Canada was late in establishing a central bank: The Bank of Canada was founded in 1934. Its directors are appointed by the government to three-year terms, and they appoint the governor, who has a seven-year term. A governing council, consisting of the four deputy governors and the governor, is the policymaking body comparable to the FOMC that makes decisions about monetary policy.

The Bank Act was amended in 1967 to give the ultimate responsibility for monetary policy to the government. So on paper, the Bank of Canada is not as instrument-independent as the Federal Reserve. In practice, however, the Bank of Canada does essentially control monetary policy. In the event of a disagreement between the bank and the government, the minister of finance can issue a directive that the bank must follow. However, because the directive must be in writing and specific and applicable for a specified period, it is unlikely that such a directive would be issued, and none has been to date. The goal for monetary policy, a target for inflation, is set jointly by the Bank of Canada and the government, so the Bank of Canada has less goal independence than the Fed.

Bank of England

www.bankofengland.co.uk Links/setframe.html
The website for the Bank of England.

Founded in 1694, the Bank of England is one of the oldest central banks. The Bank Act of 1946 gave the government statutory authority over the Bank of England. The Court (equivalent to a board of directors) of the Bank of England is made up of the governor and two deputy governors, who are appointed for five-year terms, and 16 non-executive directors, who are appointed for three-year terms.

Until 1997, the Bank of England was the least independent of the central banks examined in this chapter because the decision to raise or lower interest rates resided not within the Bank of England but with the chancellor of the Exchequer (the equivalent of the U.S. secretary of the Treasury). All of this changed when the new Labour government came to power in May 1997. At this time, the new chancellor of the Exchequer, Gordon Brown, made a surprise announcement that the Bank of England would henceforth have the power to set interest rates. However, the Bank was not granted total instrument independence: The government can overrule the Bank and

[5]An inside view of how the Fed interacts with the public and the politicians can be found in Bob Woodward, *Maestro: Greenspan's Fed and the American Boom* (New York: Simon and Schuster, 2000).

set rates "in extreme economic circumstances" and "for a limited period." Nonetheless, as in Canada, because overruling the Bank would be so public and is supposed to occur only in highly unusual circumstances and for a limited time, it likely to be a rare occurrence.

The decision to set interest rates resides in the Monetary Policy Committee, made up of the governor, two deputy governors, two members appointed by the governor after consultation with the chancellor (normally central bank officials), plus four outside economic experts appointed by the chancellor. (Surprisingly, two of the four outside experts initially appointed to this committee were not British citizens—one was Dutch and the other American, although both were residents of the United Kingdom.) The inflation target for the Bank of England is set by the Chancellor of the Exchequer, so the Bank of England is also less goal-independent than the Fed.

Bank of Japan

www.boj.or.jp/en/index.htm
The website for the Bank of Japan.

The Bank of Japan (Nippon Ginko) was founded in 1882 during the Meiji Restoration. Monetary policy is determined by the Policy Board, which is composed of the governor; two vice governors; and six outside members appointed by the cabinet and approved by the parliament, all of whom serve for five-year terms.

Until recently, the Bank of Japan was not formally independent of the government, with the ultimate power residing with the Ministry of Finance. However, the new Bank of Japan Law, which took effect in April 1998 and was the first major change in the powers of the Bank of Japan in 55 years, has changed this. In addition to stipulating that the objective of monetary policy is to attain price stability, the law granted greater instrument and goal independence to the Bank of Japan. Before this, the government had two voting members on the Policy Board, one from the Ministry of Finance and the other from the Economic Planning Agency. Now the government may send two representatives from these agencies to board meetings, but they no longer have voting rights, although they do have the ability to request delays in monetary policy decisions. In addition, the Ministry of Finance lost its authority to oversee many of the operations of the Bank of Japan, particularly the right to dismiss senior officials. However, the Ministry of Finance continues to have control over the part of the Bank's budget that is unrelated to monetary policy, which might limit its independence to some extent.

European Central Bank

www.ecb.int
The website for the European Central Bank

The Maastricht Treaty established the European Central Bank (ECB) and the European System of Central Banks (ESCB), which began operation in January 1999. The structure of the central bank is patterned after the U.S. Federal Reserve System in that central banks for each country have a role similar to that of the Federal Reserve banks. The executive board of the ECB is made up of the president, a vice president, and four other members, who are appointed for eight-year terms. The monetary policymaking body of the bank includes the six members of the executive board and the central-bank governors from each of the euro countries, all of whom must have five-year terms at a minimum.

The European Central Bank will be the most independent in the world—even more independent than the German central bank, the Bundesbank, which, before the establishment of the ECB, was considered the world's most independent central bank, along with the Swiss National Bank. The ECB is both instrument- and goal-independent of both the European Union and the national governments and has complete control over monetary policy. In addition, the ECB's mandated mission is the

pursuit of price stability. The ECB is far more independent than any other central bank in the world because its charter cannot be changed by legislation: It can be changed only by revision of the Maastricht Treaty, a difficult process, because all signatories to the treaty would have to agree.

The Trend Toward Greater Independence

As our survey of the structure and independence of the major central banks indicates, in recent years we have been seeing a remarkable trend toward increasing independence. It used to be that the Federal Reserve was substantially more independent than almost all other central banks, with the exception of those in Germany and Switzerland. Now the newly established European Central Bank is far more independent than the Fed, and greater independence has been granted to central banks like the Bank of England and the Bank of Japan, putting them more on a par with the Fed, as well as to central banks in such diverse countries as New Zealand, Sweden, and the euro nations. Both theory and experience suggest that more independent central banks produce better monetary policy, thus providing an impetus for this trend.

Explaining Central Bank Behavior

One view of government bureaucratic behavior is that bureaucracies serve the public interest (this is the *public interest view*) . Yet some economists have developed a theory of bureaucratic behavior that suggests other factors that influence how bureaucracies operate. The *theory of bureaucratic behavior* suggests that the objective of a bureaucracy is to maximize its own welfare, just as a consumer's behavior is motivated by the maximization of personal welfare and a firm's behavior is motivated by the maximization of profits. The welfare of a bureaucracy is related to its power and prestige. Thus this theory suggests that an important factor affecting a central bank's behavior is its attempt to increase its power and prestige.

What predictions does this view of a central bank like the Fed suggest? One is that the Federal Reserve will fight vigorously to preserve its autonomy, a prediction verified time and time again as the Fed has continually counterattacked congressional attempts to control its budget. In fact, it is extraordinary how effectively the Fed has been able to mobilize a lobby of bankers and businesspeople to preserve its independence when threatened.

Another prediction is that the Federal Reserve will try to avoid conflict with powerful groups that might threaten to curtail its power and reduce its autonomy. The Fed's behavior may take several forms. One possible factor explaining why the Fed is sometimes slow to increase interest rates and so smooths out their fluctuations is that it wishes to avoid a conflict with the president and Congress over increases in interest rates. The desire to avoid conflict with Congress and the president may also explain why in the past the Fed was not at all transparent about its actions and is still not fully transparent (see Box 6).

The desire of the Fed to hold as much power as possible also explains why it vigorously pursued a campaign to gain control over more banks. The campaign culminated in legislation that expanded jurisdiction of the Fed's reserve requirements to *all* banks (not just the member commercial banks) by 1987.

The theory of bureaucratic behavior seems applicable to the Federal Reserve's actions, but we must recognize that this view of the Fed as being solely concerned

Box 6: Inside the Fed

Federal Reserve Transparency

As the theory of bureaucratic behavior predicts, the Fed has incentives to hide its actions from the public and from politicians to avoid conflicts with them. In the past, this motivation led to a penchant for secrecy in the Fed, about which one former Fed official remarked that "a lot of staffers would concede that [secrecy] is designed to shield the Fed from political oversight."* For example, the Fed pursued an active defense of delaying its release of FOMC directives to Congress and the public. However, as we have seen, in 1994, it began to reveal the FOMC directive immediately after each FOMC meeting. In 1999, it also began to immediately announce the "bias" toward which direction monetary policy was likely to go, later expressed as the balance of risks in the economy. In 2002, the Fed started to report the roll call vote on the federal funds rate target taken at the FOMC meeting. Thus the Fed has increased its transparency in recent years. Yet even today, the Fed is not fully transparent: it still does not release the minutes of the FOMC meetings until six weeks after a meeting has taken place, and it does not publish its forecasts of the economy as some other central banks do.

*Quoted in "Monetary Zeal: How the Federal Reserve Under Volcker Finally Slowed Down Inflation," *Wall Street Journal*, December 7, 1984, p. 23.

with its own self-interest is too extreme. Maximizing one's welfare does not rule out altruism. (You might give generously to a charity because it makes you feel good about yourself, but in the process you are helping a worthy cause.) The Fed is surely concerned that it conduct monetary policy in the public interest. However, much uncertainty and disagreement exist over what monetary policy should be.[6] When it is unclear what is in the public interest, other motives may influence the Fed's behavior. In these situations, the theory of bureaucratic behavior may be a useful guide to predicting what motivates the Fed.

Should the Fed Be Independent?

As we have seen, the Federal Reserve is probably the most independent government agency in the United States. Every few years, the question arises in Congress whether the independence of the Fed should be curtailed. Politicians who strongly oppose a Fed policy often want to bring it under their supervision in order to impose a policy more to their liking. Should the Fed be independent, or would we be better off with a central bank under the control of the president or Congress?

The Case for Independence

The strongest argument for an independent Federal Reserve rests on the view that subjecting the Fed to more political pressures would impart an inflationary bias to monetary policy. In the view of many observers, politicians in a democratic society are

[6]One example of the uncertainty over how best to conduct monetary policy was discussed in Chapter 3: Economists are not sure how to measure money. So even if economists agreed that controlling the quantity of money is the appropriate way to conduct monetary policy (a controversial position, as we will see in later chapters), the Fed cannot be sure which monetary aggregate it should control.

shortsighted because they are driven by the need to win their next election. With this as the primary goal, they are unlikely to focus on long-run objectives, such as promoting a stable price level. Instead, they will seek short-run solutions to problems, like high unemployment and high interest rates, even if the short-run solutions have undesirable long-run consequences. For example, we saw in Chapter 5 that high money growth might lead initially to a drop in interest rates but might cause an increase later as inflation heats up. Would a Federal Reserve under the control of Congress or the president be more likely to pursue a policy of excessive money growth when interest rates are high, even though it would eventually lead to inflation and even higher interest rates in the future? The advocates of an independent Federal Reserve say yes. They believe that a politically insulated Fed is more likely to be concerned with long-run objectives and thus be a defender of a sound dollar and a stable price level.

A variation on the preceding argument is that the political process in America leads to the so-called **political business cycle**, in which just before an election, expansionary policies are pursued to lower unemployment and interest rates. After the election, the bad effects of these policies—high inflation and high interest rates—come home to roost, requiring contractionary policies that politicians hope the public will forget before the next election. There is some evidence that such a political business cycle exists in the United States, and a Federal Reserve under the control of Congress or the president might make the cycle even more pronounced.

Putting the Fed under the control of the president (making it more subject to influence by the Treasury) is also considered dangerous because the Fed can be used to facilitate Treasury financing of large budget deficits by its purchases of Treasury bonds.[7] Treasury pressure on the Fed to "help out" might lead to a more inflationary bias in the economy. An independent Fed is better able to resist this pressure from the Treasury.

Another argument for Fed independence is that control of monetary policy is too important to leave to politicians, a group that has repeatedly demonstrated a lack of expertise at making hard decisions on issues of great economic importance, such as reducing the budget deficit or reforming the banking system. Another way to state this argument is in terms of the principal–agent problem discussed in Chapters 8 and 11. Both the Federal Reserve and politicians are agents of the public (the principals), and as we have seen, both politicians and the Fed have incentives to act in their own interest rather than in the interest of the public. The argument supporting Federal Reserve independence is that the principal–agent problem is worse for politicians than for the Fed because politicians have fewer incentives to act in the public interest.

Indeed, some politicians may prefer to have an independent Fed, which can be used as a public "whipping boy" to take some of the heat off their backs. It is possible that a politician who in private opposes an inflationary monetary policy will be forced to support such a policy in public for fear of not being reelected. An independent Fed can pursue policies that are politically unpopular yet in the public interest.

[7]The Federal Reserve Act prohibited the Fed from buying Treasury bonds directly from the Treasury (except to roll over maturing securities); instead the Fed buys Treasury bonds on the open market. One possible reason for this prohibition is consistent with the foregoing argument: The Fed would find it harder to facilitate Treasury financing of large budget deficits.

The Case Against Independence

Proponents of a Fed under the control of the president or Congress argue that it is undemocratic to have monetary policy (which affects almost everyone in the economy) controlled by an elite group responsible to no one. The current lack of accountability of the Federal Reserve has serious consequences: If the Fed performs badly, there is no provision for replacing members (as there is with politicians). True, the Fed needs to pursue long-run objectives, but elected officials of Congress vote on long-run issues also (foreign policy, for example). If we push the argument further that policy is always performed better by elite groups like the Fed, we end up with such conclusions as the Joint Chiefs of Staff should determine military budgets or the IRS should set tax policies with no oversight from the president or Congress. Would you advocate this degree of independence for the Joint Chiefs or the IRS?

The public holds the president and Congress responsible for the economic well-being of the country, yet they lack control over the government agency that may well be the most important factor in determining the health of the economy. In addition, to achieve a cohesive program that will promote economic stability, monetary policy must be coordinated with fiscal policy (management of government spending and taxation). Only by placing monetary policy under the control of the politicians who also control fiscal policy can these two policies be prevented from working at cross-purposes.

Another argument against Federal Reserve independence is that an independent Fed has not always used its freedom successfully. The Fed failed miserably in its stated role as lender of last resort during the Great Depression, and its independence certainly didn't prevent it from pursuing an overly expansionary monetary policy in the 1960s and 1970s that contributed to rapid inflation in this period.

Our earlier discussion also suggests that the Federal Reserve is not immune from political pressures.[8] Its independence may encourage it to pursue a course of narrow self-interest rather than the public interest.

There is yet no consensus on whether Federal Reserve independence is a good thing, although public support for independence of the central bank seems to have been growing in both the United States and abroad. As you might expect, people who like the Fed's policies are more likely to support its independence, while those who dislike its policies advocate a less independent Fed.

Central Bank Independence and Macroeconomic Performance Throughout The World

We have seen that advocates of an independent central bank believe that macroeconomic performance will be improved by making the central bank more independent. Recent research seems to support this conjecture: When central banks are ranked from least independent to most independent, inflation performance is found to be the best for countries with the most independent central banks.[9] Although a more independent central bank appears to lead to a lower inflation rate, this is not achieved at the expense of poorer real economic performance. Countries with independent central banks are no more likely to have high unemployment or greater output fluctuations than countries with less independent central banks.

---

[8]For evidence on this issue, see Robert E. Weintraub, "Congressional Supervision of Monetary Policy," *Journal of Monetary Economics* 4 (1978): 341–362. Some economists suggest that lessening the independence of the Fed might even reduce the incentive for politically motivated monetary policy; see Milton Friedman, "Monetary Policy: Theory and Practice," *Journal of Money, Credit and Banking* 14 (1982): 98–118.

[9]Alberto Alesina and Lawrence H. Summers, "Central Bank Independence and Macroeconomic Performance: Some Comparative Evidence," *Journal of Money, Credit and Banking* 25 (1993): 151–162. However, Adam Posen, "Central Bank Independence and Disinflationary Credibility: A Missing Link," Federal Reserve Bank of New York Staff Report No. 1, May 1995, has cast some doubt on whether the causality runs from central bank independence to improved inflation performance.

Summary

1. The Federal Reserve System was created in 1913 to lessen the frequency of bank panics. Because of public hostility to central banks and the centralization of power, the Federal Reserve System was created with many checks and balances to diffuse power.
2. The formal structure of the Federal Reserve System consists of 12 regional Federal Reserve banks, around 4,800 member commercial banks, the Board of Governors of the Federal Reserve System, the Federal Open Market Committee, and the Federal Advisory Council.
3. Although on paper the Federal Reserve System appears to be decentralized, in practice it has come to function as a unified central bank controlled by the Board of Governors, especially the board's chairman.
4. The Federal Reserve is more independent than most agencies of the U.S. government, but it is still subject to political pressures because the legislation that structures the Fed is written by Congress and can be changed at any time. The theory of bureaucratic behavior suggests that one factor driving the Fed's behavior might be its attempt to increase its power and prestige. This view explains many of the Fed's actions, although the agency may also try to act in the public interest.
5. The case for an independent Federal Reserve rests on the view that curtailing the Fed's independence and subjecting it to more political pressures would impart an inflationary bias to monetary policy. An independent Fed can afford to take the long view and not respond to short-run problems that will result in expansionary monetary policy and a political business cycle. The case against an independent Fed holds that it is undemocratic to have monetary policy (so important to the public) controlled by an elite that is not accountable to the public. An independent Fed also makes the coordination of monetary and fiscal policy difficult.

Key Terms

Board of Governors of the Federal Reserve System, p. 337
Federal Open Market Committee (FOMC), p. 337
Federal Reserve banks, p. 337
goal independence, p. 347
instrument independence, p. 347
open market operations, p. 340
political business cycle, p. 353

Questions and Problems

Questions marked with an asterisk are answered at the end of the book in an appendix, "Answers to Selected Questions and Problems."

*1. Why was the Federal Reserve System set up with 12 regional Federal Reserve banks rather than one central bank, as in other countries?
2. What political realities might explain why the Federal Reserve Act of 1913 placed two Federal Reserve banks in Missouri?
*3. "The Federal Reserve System resembles the U.S. Constitution in that it was designed with many checks and balances." Discuss.
4. In what ways can the regional Federal Reserve banks influence the conduct of monetary policy?
*5. Which entities in the Federal Reserve System control the discount rate? Reserve requirements? Open market operations?
6. Do you think that the 14-year nonrenewable terms for governors effectively insulate the Board of Governors from political pressure?
*7. Over time, which entities have gained power in the Federal Reserve System and which have lost power? Why do you think this has happened?

8. The Fed is the most independent of all U.S. government agencies. What is the main difference between it and other government agencies that explains its greater independence?

*9. What is the primary tool that Congress uses to exercise some control over the Fed?

10. In the 1960s and 1970s, the Federal Reserve System lost member banks at a rapid rate. How can the theory of bureaucratic behavior explain the Fed's campaign for legislation to require all commercial banks to become members? Was the Fed successful in this campaign?

*11. "The theory of bureaucratic behavior indicates that the Fed never operates in the public interest." Is this statement true, false, or uncertain? Explain your answer.

12. Why might eliminating the Fed's independence lead to a more pronounced political business cycle?

*13. "The independence of the Fed leaves it completely unaccountable for its actions." Is this statement true, false, or uncertain? Explain your answer.

14. "The independence of the Fed has meant that it takes the long view and not the short view." Is this statement true, false, or uncertain? Explain your answer.

*15. The Fed promotes secrecy by not releasing the minutes of the FOMC meetings to Congress or the public immediately. Discuss the pros and cons of this policy.

Web Exercises

1. Go to www.federalreserve.gov/general.htm and click on the link to general information. Choose "Structure of the Federal Reserve." According to the Federal Reserve, what is the most important responsibility of the Board of Governors?

2. Go to the above site and click on "Monetary Policy" to find the beige book. According to the summary of the most recently published book, is the economy weakening or recovering?

Chapter 15

Multiple Deposit Creation and the Money Supply Process

PREVIEW

As we saw in Chapter 5 and will see in later chapters on monetary theory, movements in the money supply affect interest rates and the overall health of the economy and thus affect us all. Because of its far-reaching effects on economic activity, it is important to understand how the money supply is determined. Who controls it? What causes it to change? How might control of it be improved? In this and subsequent chapters, we answer these questions by providing a detailed description of the *money supply process*, the mechanism that determines the level of the money supply.

Because deposits at banks are by far the largest component of the money supply, understanding how these deposits are created is the first step in understanding the money supply process. This chapter provides an overview of how the banking system creates deposits, and describes the basic principles of the money supply, needed to understand later chapters.

Four Players in the Money Supply Process

The "cast of characters" in the money supply story is as follows:

1. The *central bank*—the government agency that oversees the banking system and is responsible for the conduct of monetary policy; in the United States, it is called the Federal Reserve System
2. *Banks* (depository institutions)—the financial intermediaries that accept deposits from individuals and institutions and make loans: commercial banks, savings and loan associations, mutual savings banks, and credit unions
3. *Depositors*—individuals and institutions that hold deposits in banks
4. *Borrowers* from banks—individuals and institutions that borrow from the depository institutions and institutions that issue bonds that are purchased by the depository institutions

Of the four players, the central bank—the Federal Reserve System—is the most important. The Fed's conduct of monetary policy involves actions that affect its balance sheet (holdings of assets and liabilities), to which we turn now.

The Fed's Balance Sheet

www.federalreserve.gov/boarddocs/rptcongress/annual01/default.htm
See the most recent Federal Reserve financial statement.

The operation of the Fed and its monetary policy involve actions that affect its balance sheet, its holdings of assets and liabilities. Here we discuss a simplified balance sheet that includes just four items that are essential to our understanding of the money supply process.[1]

FEDERAL RESERVE SYSTEM

Assets	Liabilities
Government securities	Currency in circulation
Discount loans	Reserves

Liabilities

The two liabilities on the balance sheet, currency in circulation and reserves, are often referred to as the *monetary liabilities* of the Fed. They are an important part of the money supply story, because increases in either or both will lead to an increase in the money supply (everything else being constant). The sum of the Fed's monetary liabilities (currency in circulation and reserves) and the U.S. Treasury's monetary liabilities (Treasury currency in circulation, primarily coins) is called the **monetary base**. When discussing the monetary base, we will focus only on the monetary liabilities of the Fed because the monetary liabilities of the Treasury account for less then 10% of the base.[2]

www.rich.frb.org/research/econed/museum/
A virtual tour of the Federal Reserve's money museum.

1. Currency in circulation. The Fed issues currency (those green-and-gray pieces of paper in your wallet that say "Federal Reserve Note" at the top). Currency in circulation is the amount of currency in the hands of the public. (Currency held by depository institutions is also a liability of the Fed, but is counted as part of the reserves.)

Federal Reserve notes are IOUs from the Fed to the bearer and are also liabilities, but unlike most, they promise to pay back the bearer solely with Federal Reserve notes; that is, they pay off IOUs with other IOUs. Accordingly, if you bring a $100 bill to the Federal Reserve and demand payment, you will receive two $50s, five $20s, ten $10s, or one hundred $1 bills.

People are more willing to accept IOUs from the Fed than from you or me because Federal Reserve notes are a recognized medium of exchange; that is, they are accepted as a means of payment and so function as money. Unfortunately, neither you nor I can convince people that our IOUs are worth anything more than the paper they are written on.[3]

---

[1] A detailed discussion of the Fed's balance sheet and the factors that affect the monetary base can be found in the appendix to this chapter, which you can find on this book's web site at www.aw.com/mishkin.

[2] It is also safe to ignore the Treasury's monetary liabilities when discussing the monetary base because the Treasury cannot actively supply its monetary liabilities to the economy due to legal restrictions.

[3] The currency item on our balance sheet refers only to currency *in circulation*; that is, the amount in the hands of the public. Currency that has been printed by the U.S. Bureau of Engraving and Printing is not automatically a liability of the Fed. For example, consider the importance of having $1 million of your own IOUs printed up. You give out $100 worth to other people and keep the other $999,900 in your pocket. The $999,900 of IOUs does not make you richer or poorer and does not affect your indebtedness. You care only about the $100 of liabilities from the $100 of circulated IOUs. The same reasoning applies for the Fed in regard to its Federal Reserve notes.

For similar reasons, the currency component of the money supply, no matter how it is defined, includes only currency in circulation. It does not include any additional currency that is not yet in the hands of the public. The fact that currency has been printed but is not circulating means that it is not anyone's asset or liability and thus cannot affect anyone's behavior. Therefore, it makes sense not to include it in the money supply.

2. *Reserves*. All banks have an account at the Fed in which they hold deposits. **Reserves** consist of deposits at the Fed plus currency that is physically held by banks (called *vault cash* because it is stored in bank vaults). Reserves are assets for the banks but liabilities for the Fed, because the banks can demand payment on them at any time and the Fed is required to satisfy its obligation by paying Federal Reserve notes. As you will see, an increase in reserves leads to an increase in the level of deposits and hence in the money supply.

Total reserves can be divided into two categories: reserves that the Fed requires banks to hold (**required reserves**) and any additional reserves the banks choose to hold (**excess reserves**). For example, the Fed might require that for every dollar of deposits at a depository institution, a certain fraction (say, 10 cents) must be held as reserves. This fraction (10%) is called the **required reserve ratio**. Currently, the Fed pays no interest on reserves.

Assets

The two assets on the Fed's balance sheet are important for two reasons. First, changes in the asset items lead to changes in reserves and consequently to changes in the money supply. Second, because these assets (government securities and discount loans) earn interest while the liabilities (currency in circulation and reserves) do not, the Fed makes billions of dollars every year—its assets earn income, and its liabilities cost nothing. Although it returns most of its earnings to the federal government, the Fed does spend some of it on "worthy causes," such as supporting economic research.

1. *Government securities*. This category of assets covers the Fed's holdings of securities issued by the U.S. Treasury. As you will see, the Fed provides reserves to the banking system by purchasing securities, thereby increasing its holdings of these assets. An increase in government securities held by the Fed leads to an increase in the money supply.

2. *Discount loans*. The Fed can provide reserves to the banking system by making discount loans to banks. An increase in discount loans can also be the source of an increase in the money supply. The interest rate charged banks for these loans is called the **discount rate**.

Control of the Monetary Base

The *monetary base* (also called **high-powered money**) equals currency in circulation C plus the total reserves in banking system R.[4] The monetary base MB can be expressed as

$$MB = C + R$$

The Federal Reserve exercises control over the monetary base through its purchases or sale of government securities in the open market, called **open market operations**, and through its extension of discount loans to banks.

Federal Reserve Open Market Operations

The primary way in which the Fed causes changes in the monetary base is through its open market operations. A purchase of bonds by the Fed is called an **open market purchase**, and a sale of bonds by the Fed is called an **open market sale**.

[4]Here currency in circulation includes both Federal Reserve currency (Federal Reserve notes) and Treasury currency (primarily coins).

Open Market Purchase from a Bank. Suppose that the Fed purchases $100 of bonds from a bank and pays for them with a $100 check. The bank will either deposit the check in its account with the Fed or cash it in for currency, which will be counted as vault cash. To understand what occurs as a result of this transaction, we look at *T-accounts*, which list only the changes that occur in balance sheet items starting from the initial balance sheet position. Either action means that the bank will find itself with $100 more reserves and a reduction in its holdings of securities of $100. The T-account for the banking system, then, is:

BANKING SYSTEM

Assets		Liabilities
Securities	−$100	
Reserves	+$100	

The Fed meanwhile finds that its liabilities have increased by the additional $100 of reserves, while its assets have increased by the $100 of additional securities that it now holds. Its T-account is:

FEDERAL RESERVE SYSTEM

Assets		Liabilities	
Securities	+$100	Reserves	+$100

The net result of this open market purchase is that reserves have increased by $100, the amount of the open market purchase. Because there has been no change of currency in circulation, the monetary base has also risen by $100.

Open Market Purchase from the Nonbank Public. To understand what happens when there is an open market purchase from the nonbank public, we must look at two cases. First, let's assume that the person or corporation that sells the $100 of bonds to the Fed deposits the Fed's check in the local bank. The nonbank public's T-account after this transaction is:

NONBANK PUBLIC

Assets		Liabilities
Securities	−$100	
Checkable deposits	+$100	

When the bank receives the check, it credits the depositor's account with the $100 and then deposits the check in its account with the Fed, thereby adding to its reserves. The banking system's T-account becomes:

BANKING SYSTEM

Assets		Liabilities	
Reserves	+$100	Checkable deposits	+$100

The effect on the Fed's balance sheet is that it has gained $100 of securities in its assets column, while it has an increase of $100 of reserves in its liabilities column:

FEDERAL RESERVE SYSTEM

Assets		Liabilities	
Securities	+$100	Reserves	+$100

As you can see in the above T-account, when the Fed's check is deposited in a bank, the net result of the Fed's open market purchase from the nonbank public is identical to the effect of its open market purchase from a bank: Reserves increase by the amount of the open market purchase, and the monetary base increases by the same amount.

If, however, the person or corporation selling the bonds to the Fed cashes the Fed's check either at a local bank or at a Federal Reserve bank for currency, the effect on reserves is different.[5] This seller will receive currency of $100 while reducing holdings of securities by $100. The bond seller's T-account will be:

NONBANK PUBLIC

Assets		Liabilities
Securities	−$100	
Currency	+$100	

The Fed now finds that it has exchanged $100 of currency for $100 of securities, so its T-account is:

FEDERAL RESERVE SYSTEM

Assets		Liabilities	
Securities	+$100	Currency in circulation	+$100

[5]If the bond seller cashes the check at the local bank, its balance sheet will be unaffected, because the $100 of vault cash that it pays out will be exactly matched by the deposit of the $100 check at the Fed. Thus its reserves will remain the same, and there will be no effect on its T-account. That is why a T-account for the banking system does not appear here.

The net effect of the open market purchase in this case is that reserves are unchanged, while currency in circulation increases by the $100 of the open market purchase. Thus the monetary base increases by the $100 amount of the open market purchase, while reserves do not. This contrasts with the case in which the seller of the bonds deposits the Fed's check in a bank; in that case, reserves increase by $100, and so does the monetary base.

The analysis reveals that ***the effect of an open market purchase on reserves depends on whether the seller of the bonds keeps the proceeds from the sale in currency or in deposits***. If the proceeds are kept in currency, the open market purchase has no effect on reserves; if the proceeds are kept as deposits, reserves increase by the amount of the open market purchase.

The effect of an open market purchase on the monetary base, however, is always the same (the monetary base increases by the amount of the purchase) whether the seller of the bonds keeps the proceeds in deposits or in currency. The impact of an open market purchase on reserves is much more uncertain than its impact on the monetary base.

Open Market Sale. If the Fed sells $100 of bonds to a bank or the nonbank public, the monetary base will decline by $100. For example, if the Fed sells the bonds to an individual who pays for them with currency, the buyer exchanges $100 of currency for $100 of bonds, and the resulting T-account is:

NONBANK PUBLIC

Assets		Liabilities
Securities	+$100	
Currency	−$100	

The Fed, for its part, has reduced its holdings of securities by $100 and has also lowered its monetary liability by accepting the currency as payment for its bonds, thereby reducing the amount of currency in circulation by $100:

FEDERAL RESERVE SYSTEM

Assets		Liabilities	
Securities	−$100	Currency in circulation	−$100

The effect of the open market sale of $100 of bonds is to reduce the monetary base by an equal amount, although reserves remain unchanged. Manipulations of T-accounts in cases in which the buyer of the bonds is a bank or the buyer pays for the bonds with a check written on a checkable deposit account at a local bank lead to the same $100 reduction in the monetary base, although the reduction occurs because the level of reserves has fallen by $100.

Study Guide

The best way to learn how open market operations affect the monetary base is to use T-accounts. Using T-accounts, try to verify that an open market sale of $100 of bonds to a bank or to a person who pays with a check written on a bank account leads to a $100 reduction in the monetary base.

The following conclusion can now be drawn from our analysis of open market purchases and sales: ***The effect of open market operations on the monetary base is much more certain than the effect on reserves.*** Therefore, the Fed can control the monetary base with open market operations more effectively than it can control reserves.

Open market operations can also be done in other assets besides government bonds and have the same effects on the monetary base we have described here. One example of this is a foreign exchange intervention by the Fed (see Box 1).

Shifts from Deposits into Currency

Even if the Fed does not conduct open market operations, a shift from deposits to currency will affect the reserves in the banking system. However, such a shift will have no effect on the monetary base, another reason why the Fed has more control over the monetary base than over reserves.

Let's suppose that Jane Brown (who opened a $100 checking account at the First National Bank in Chapter 9) decides that tellers are so abusive in all banks that she closes her account by withdrawing the $100 balance in cash and vows never to deposit it in a bank again. The effect on the T-account of the nonbank public is:

NONBANK PUBLIC

Assets		Liabilities
Checkable deposits	−$100	
Currency	+$100	

Box 1: Global

Foreign Exchange Rate Intervention and the Monetary Base

It is common to read in the newspaper about a Federal Reserve intervention to buy or sell dollars in the foreign exchange market. (Note that this intervention occurs at the request of the U.S. Treasury.) Can this intervention also be a factor that affects the monetary base? The answer is yes, because a Federal Reserve intervention in the foreign exchange market involves a purchase or sale of assets denominated in a foreign currency.

Suppose that the Fed purchases $100 of deposits denominated in euros in exchange for $100 of deposits at the Fed (a sale of dollars for euros). A Federal Reserve purchase of any asset, whether it is a U.S. government bond or a deposit denominated in a foreign currency, is still just an open market purchase and so leads to an equal rise in the monetary base. Similarly, a sale of foreign currency deposits is just an open market sale and leads to a decline in the monetary base. Federal Reserve interventions in the foreign exchange market are thus an important influence on the monetary base, a topic that we discuss further in Chapter 20.

The banking system loses $100 of deposits and hence $100 of reserves:

BANKING SYSTEM

Assets		Liabilities	
Reserves	−$100	Checkable deposits	−$100

For the Fed, Jane Brown's action means that there is $100 of additional currency circulating in the hands of the public, while reserves in the banking system have fallen by $100. The Fed's T-account is:

FEDERAL RESERVE SYSTEM

Assets		Liabilities	
		Currency in circulation	+$100
		Reserves	−$100

The net effect on the monetary liabilities of the Fed is a wash; the monetary base is unaffected by Jane Brown's disgust at the banking system. But reserves are affected. Random fluctuations of reserves can occur as a result of random shifts into currency and out of deposits, and vice versa. The same is not true for the monetary base, making it a more stable variable.

Discount Loans

In this chapter so far we have seen changes in the monetary base solely as a result of open market operations. However, the monetary base is also affected when the Fed makes a discount loan to a bank. When the Fed makes a $100 discount loan to the First National Bank, the bank is credited with $100 of reserves from the proceeds of the loan. The effects on the balance sheets of the banking system and the Fed are illustrated by the following T-accounts:

BANKING SYSTEM

Assets		Liabilities	
Reserves	+$100	Discount loans	+$100

FEDERAL RESERVE SYSTEM

Assets		Liabilities	
Discount loans	+$100	Reserves	+$100

The monetary liabilities of the Fed have now increased by $100, and the monetary base, too, has increased by this amount. However, if a bank pays off a loan from the Fed, thereby reducing its borrowings from the Fed by $100, the T-accounts of the banking system and the Fed are as follows:

BANKING SYSTEM			
Assets		Liabilities	
Reserves	−$100	Discount loans	−$100

FEDERAL RESERVE SYSTEM			
Assets		Liabilities	
Discount loans	−$100	Reserves	−$100

The net effect on the monetary liabilities of the Fed, and hence on the monetary base, is then a reduction of $100. We see that the monetary base changes one-for-one with the change in the borrowings from the Fed.

Other Factors That Affect the Monetary Base

So far in this chapter, it seems as though the Fed has complete control of the monetary base through its open market operations and discount loans. However, the world is a little bit more complicated for the Fed. Two important items that are not controlled by the Fed but affect the monetary base are *float* and *Treasury deposits at the Fed*. When the Fed clears checks for banks, it often credits the amount of the check to a bank that has deposited it (increases the bank's reserves) but only later debits (decreases the reserves of) the bank on which the check is drawn. The resulting temporary net increase in the total amount of reserves in the banking system (and hence in the monetary base) occurring from the Fed's check-clearing process is called **float**. When the U.S. Treasury moves deposits from commercial banks to its account at the Fed, leading to a rise in *Treasury deposits at the Fed*, it causes a deposit outflow at these banks like that shown in Chapter 9 and thus causes reserves in the banking system and the monetary base to fall. Thus *float* (affected by random events such as the weather, which affects how quickly checks are presented for payment) and *Treasury deposits at the Fed* (determined by the U.S. Treasury's actions) both affect the monetary base but are not fully controlled by the Fed. Decisions by the U.S. Treasury to have the Fed intervene in the foreign exchange market also affect the monetary base, as can be seen in Box 1.

Overview of the Fed's Ability to Control the Monetary Base

The factor that most affects the monetary base is the Fed's holdings of securities, which are completely controlled by the Fed through its open market operations. Factors not controlled by the Fed (for example, float and Treasury deposits with the Fed) undergo substantial short-run variations and can be important sources of fluctuations in the monetary base over time periods as short as a week. However, these fluctuations are usually quite predictable and so can be offset through open market operations. ***Although float and Treasury deposits with the Fed undergo substantial short-run fluctuations, which complicate control of the monetary base, they do not prevent the Fed from accurately controlling it.***

Multiple Deposit Creation: A Simple Model

With our understanding of how the Federal Reserve controls the monetary base and how banks operate (Chapter 9), we now have the tools necessary to explain how deposits are created. When the Fed supplies the banking system with $1 of additional

reserves, deposits increase by a multiple of this amount—a process called **multiple deposit creation**.

Deposit Creation: The Single Bank

Suppose that the $100 open market purchase described earlier was conducted with the First National Bank. After the Fed has bought the $100 bond from the First National Bank, the bank finds that it has an increase in reserves of $100. To analyze what the bank will do with these additional reserves, assume that the bank does not want to hold excess reserves because it earns no interest on them. We begin the analysis with the following T-account:

FIRST NATIONAL BANK

Assets		Liabilities	
Securities	−$100		
Reserves	+$100		

Because the bank has no increase in its checkable deposits, required reserves remain the same, and the bank finds that its additional $100 of reserves means that its excess reserves have increased by $100. Let's say that the bank decides to make a loan equal in amount to the $100 increase in excess reserves. When the bank makes the loan, it sets up a checking account for the borrower and puts the proceeds of the loan into this account. In this way, the bank alters its balance sheet by increasing its liabilities with $100 of checkable deposits and at the same time increasing its assets with the $100 loan. The resulting T-account looks like this:

FIRST NATIONAL BANK

Assets		Liabilities	
Securities	−$100	Checkable deposits	+$100
Reserves	+$100		
Loans	+$100		

The bank has created checkable deposits by its act of lending. Because checkable deposits are part of the money supply, the bank's act of lending has in fact created money.

In its current balance sheet position, the First National Bank still has excess reserves and so might want to make additional loans. However, these reserves will not stay at the bank for very long. The borrower took out a loan not to leave $100 idle at the First National Bank but to purchase goods and services from other individuals and corporations. When the borrower makes these purchases by writing checks, they will be deposited at other banks, and the $100 of reserves will leave the First National Bank. ***A bank cannot safely make loans for an amount greater than the excess reserves it has before it makes the loan.***

The final T-account of the First National Bank is:

FIRST NATIONAL BANK

Assets		Liabilities	
Securities	−$100		
Loans	+$100		

The increase in reserves of $100 has been converted into additional loans of $100 at the First National Bank, plus an additional $100 of deposits that have made their way to other banks. (All the checks written on accounts at the First National Bank are deposited in banks rather than converted into cash, because we are assuming that the public does not want to hold any additional currency.) Now let's see what happens to these deposits at the other banks.

Deposit Creation: The Banking System

To simplify the analysis, let us assume that the $100 of deposits created by First National Bank's loan is deposited at Bank A and that this bank and all other banks hold no excess reserves. Bank A's T-account becomes:

BANK A

Assets		Liabilities	
Reserves	+$100	Checkable deposits	+$100

If the required reserve ratio is 10%, this bank will now find itself with a $10 increase in required reserves, leaving it $90 of excess reserves. Because Bank A (like the First National Bank) does not want to hold on to excess reserves, it will make loans for the entire amount. Its loans and checkable deposits will then increase by $90, but when the borrower spends the $90 of checkable deposits, they and the reserves at Bank A will fall back down by this same amount. The net result is that Bank A's T-account will look like this:

BANK A

Assets		Liabilities	
Reserves	+$10	Checkable deposits	+$100
Loans	+$90		

If the money spent by the borrower to whom Bank A lent the $90 is deposited in another bank, such as Bank B, the T-account for Bank B will be:

BANK B

Assets		Liabilities	
Reserves	+$90	Checkable deposits	+$90

The checkable deposits in the banking system have increased by another $90, for a total increase of $190 ($100 at Bank A plus $90 at Bank B). In fact, the distinction between Bank A and Bank B is not necessary to obtain the same result on the overall expansion of deposits. If the borrower from Bank A writes checks to someone who deposits them at Bank A, the same change in deposits would occur. The T-accounts for Bank B would just apply to Bank A, and its checkable deposits would increase by the total amount of $190.

Bank B will want to modify its balance sheet further. It must keep 10% of $90 ($9) as required reserves and has 90% of $90 ($81) in excess reserves and so can make loans of this amount. Bank B will make an $81 loan to a borrower, who spends the proceeds from the loan. Bank B's T-account will be:

BANK B

Assets		Liabilities	
Reserves	+$ 9	Checkable deposits	+$90
Loans	+$81		

The $81 spent by the borrower from Bank B will be deposited in another bank (Bank C). Consequently, from the initial $100 increase of reserves in the banking system, the total increase of checkable deposits in the system so far is $271 (= $100 + $90 + $81).

Following the same reasoning, if all banks make loans for the full amount of their excess reserves, further increments in checkable deposits will continue (at Banks C, D, E, and so on), as depicted in Table 1. Therefore, the total increase in deposits from the initial $100 increase in reserves will be $1,000: The increase is tenfold, the reciprocal of the 0.10 reserve requirement.

If the banks choose to invest their excess reserves in securities, the result is the same. If Bank A had taken its excess reserves and purchased securities instead of making loans, its T-account would have looked like this:

BANK A

Assets		Liabilities	
Reserves	+$10	Checkable deposits	+$100
Securities	+$90		

Table 1 Creation of Deposits (assuming 10% reserve requirement and a $100 increase in reserves)

Bank	Increase in Deposits ($)	Increase in Loans ($)	Increase in Reserves ($)
First National	0.00	100.00	0.00
A	100.00	90.00	10.00
B	90.00	81.00	9.00
C	81.00	72.90	8.10
D	72.90	65.61	7.29
E	65.61	59.05	6.56
F	59.05	53.14	5.91
.	.	.	.
.	.	.	.
.	.	.	.
Total for all banks	1,000.00	1,000.00	100.00

When the bank buys $90 of securities, it writes a $90 check to the seller of the securities, who in turn deposits the $90 at a bank such as Bank B. Bank B's checkable deposits rise by $90, and the deposit expansion process is the same as before. ***Whether a bank chooses to use its excess reserves to make loans or to purchase securities, the effect on deposit expansion is the same.***

You can now see the difference in deposit creation for the single bank versus the banking system as a whole. Because a single bank can create deposits equal only to the amount of its excess reserves, it cannot by itself generate multiple deposit expansion. A single bank cannot make loans greater in amount than its excess reserves, because the bank will lose these reserves as the deposits created by the loan find their way to other banks. However, the banking system as a whole can generate a multiple expansion of deposits, because when a bank loses its excess reserves, these reserves do not leave the banking system even though they are lost to the individual bank. So as each bank makes a loan and creates deposits, the reserves find their way to another bank, which uses them to make additional loans and create additional deposits. As you have seen, this process continues until the initial increase in reserves results in a multiple increase in deposits.

The multiple increase in deposits generated from an increase in the banking system's reserves is called the **simple deposit multiplier**.[6] In our example with a 10% required reserve ratio, the simple deposit multiplier is 10. More generally, the simple deposit multiplier equals the reciprocal of the required reserve ratio, expressed as a fraction (10 = 1/0.10), so the formula for the multiple expansion of deposits can be written as:

$$\Delta D = \frac{1}{r} \times \Delta R \tag{1}$$

[6]This multiplier should not be confused with the Keynesian multiplier, which is derived through a similar step-by-step analysis. That multiplier relates an increase in income to an increase in investment, whereas the simple deposit multiplier relates an increase in deposits to an increase in reserves.

where ΔD = change in total checkable deposits in the banking system
r = required reserve ratio (0.10 in the example)
ΔR = change in reserves for the banking system ($100 in the example)[7]

Deriving the Formula for Multiple Deposit Creation

The formula for the multiple creation of deposits can also be derived directly using algebra. We obtain the same answer for the relationship between a change in deposits and a change in reserves, but more quickly.

Our assumption that banks do not hold on to any excess reserves means that the total amount of required reserves for the banking system RR will equal the total reserves in the banking system R:

$$RR = R$$

The total amount of required reserves equals the required reserve ratio r times the total amount of checkable deposits D:

$$RR = r \times D$$

Substituting $r \times D$ for RR in the first equation:

$$r \times D = R$$

and dividing both sides of the preceding equation by r gives us:

$$D = \frac{1}{r} \times R$$

Taking the change in both sides of this equation and using delta to indicate a change:

$$\Delta D = \frac{1}{r} \times \Delta R$$

which is the same formula for deposit creation found in Equation 1.

This derivation provides us with another way of looking at the multiple creation of deposits, because it forces us to look directly at the banking system as a whole rather than one bank at a time. For the banking system as a whole, deposit creation (or contraction) will stop only when all excess reserves in the banking system are gone; that is, the banking system will be in equilibrium when the total amount of required reserves equals the total amount of reserves, as seen in the equation $RR = R$. When $r \times D$ is substituted for RR, the resulting equation $R = r \times D$ tells us how high checkable deposits will have to be in order for required reserves to equal total reserves. Accordingly, a given level of reserves in the banking system determines the level of checkable deposits when the banking system is in equilibrium (when $ER = 0$); put another way, the given level of reserves supports a given level of checkable deposits.

---

[7] A formal derivation of this formula follows. Using the reasoning in the text, the change in checkable deposits is $100 ($= \Delta R \times 1$) plus $90 [$= \Delta R \times (1 - r)$] plus $81 [$= \Delta R \times (1 - r)^2$ and so on, which can be rewritten as:

$$\Delta D = \Delta R \times [1 + (1 - r) + (1 - r)^2 + (1 - r)^3 + \cdots]$$

Using the formula for the sum of an infinite series found in footnote 5 in Chapter 4, this can be rewritten as:

$$\Delta D = \Delta R \times \frac{1}{1 - (1 - r)} = \frac{1}{r} \times \Delta R$$

In our example, the required reserve ratio is 10%. If reserves increase by $100, checkable deposits must rise to $1,000 in order for total required reserves also to increase by $100. If the increase in checkable deposits is less than this, say $900, then the increase in required reserves of $90 remains below the $100 increase in reserves, so there are still excess reserves somewhere in the banking system. The banks with the excess reserves will now make additional loans, creating new deposits, and this process will continue until all reserves in the system are used up. This occurs when checkable deposits have risen to $1,000.

We can also see this by looking at the T-account of the banking system as a whole (including the First National Bank) that results from this process:

BANKING SYSTEM

Assets		Liabilities	
Securities	−$ 100	Checkable deposits	+$1,000
Reserves	+$ 100		
Loans	+$1,000		

The procedure of eliminating excess reserves by loaning them out means that the banking system (First National Bank and Banks A, B, C, D, and so on) continues to make loans up to the $1,000 amount until deposits have reached the $1,000 level. In this way, $100 of reserves supports $1,000 (ten times the quantity) of deposits.

Critique of the Simple Model

Our model of multiple deposit creation seems to indicate that the Federal Reserve is able to exercise complete control over the level of checkable deposits by setting the required reserve ratio and the level of reserves. The actual creation of deposits is much less mechanical than the simple model indicates. If proceeds from Bank A's $90 loan are not deposited but are kept in cash, nothing is deposited in Bank B, and the deposit creation process stops dead in its tracks. The total increase in checkable deposits is only $100—considerably less than the $1,000 we calculated. So if some proceeds from loans are used to raise the holdings of currency, checkable deposits will not increase by as much as our streamlined model of multiple deposit creation tells us.

Another situation ignored in our model is one in which banks do not make loans or buy securities in the full amount of their excess reserves. If Bank A decides to hold on to all $90 of its excess reserves, no deposits would be made in Bank B, and this would also stop the deposit creation process. The total increase in deposits would again be only $100 and not the $1,000 increase in our example. Hence if banks choose to hold all or some of their excess reserves, the full expansion of deposits predicted by the simple model of multiple deposit creation does not occur.

Our examples rightly indicate that the Fed is not the only player whose behavior influences the level of deposits and therefore the money supply. Banks' decisions regarding the amount of excess reserves they wish to hold, depositors' decisions regarding how much currency to hold, and borrowers' decisions on how much to borrow from banks can cause the money supply to change. In the next chapter, we stress the behavior and interactions of the four players in constructing a more realistic model of the money supply process.

Summary

1. There are four players in the money supply process: the central bank, banks (depository institutions), depositors, and borrowers from banks.
2. Four items in the Fed's balance sheet are essential to our understanding of the money supply process: the two liability items, currency in circulation and reserves, which together make up the monetary base, and the two asset items, government securities and discount loans.
3. The Federal Reserve controls the monetary base through open market operations and extension of discount loans to banks and has better control over the monetary base than over reserves. Although float and Treasury deposits with the Fed undergo substantial short-run fluctuations, which complicate control of the monetary base, they do not prevent the Fed from accurately controlling it.
4. A single bank can make loans up to the amount of its excess reserves, thereby creating an equal amount of deposits. The banking system can create a multiple expansion of deposits, because as each bank makes a loan and creates deposits, the reserves find their way to another bank, which uses them to make loans and create additional deposits. In the simple model of multiple deposit creation in which banks do not hold on to excess reserves and the public holds no currency, the multiple increase in checkable deposits (simple deposit multiplier) equals the reciprocal of the required reserve ratio.
5. The simple model of multiple deposit creation has serious deficiencies. Decisions by depositors to increase their holdings of currency or of banks to hold excess reserves will result in a smaller expansion of deposits than the simple model predicts. All four players—the Fed, banks, depositors, and borrowers from banks—are important in the determination of the money supply.

Key Terms

discount rate, p. 359
excess reserves, p. 359
float, p. 365
high-powered money, p. 359
monetary base, p. 358
multiple deposit creation, p. 366
open market operations, p. 359
open market purchase, p. 359
open market sale, p. 359
required reserve ratio, p. 359
required reserves, p. 359
reserves, p. 359
simple deposit multiplier, p. 369

Questions and Problems

Questions marked with an asterisk are answered at the end of the book in an appendix, "Answers to Selected Questions and Problems."

1. If the Fed sells $2 million of bonds to the First National Bank, what happens to reserves and the monetary base? Use T-accounts to explain your answer.

*2. If the Fed sells $2 million of bonds to Irving the Investor, who pays for the bonds with a briefcase filled with currency, what happens to reserves and the monetary base? Use T-accounts to explain your answer.

*3. If the Fed lends five banks an additional total of $100 million but depositors withdraw $50 million and hold it as currency, what happens to reserves and the monetary base? Use T-accounts to explain your answer.

4. The First National Bank receives an extra $100 of reserves but decides not to lend any of these reserves

out. How much deposit creation takes place for the entire banking system?

Unless otherwise noted, the following assumptions are made in all the remaining problems: The required reserve ratio on checkable deposits is 10%, banks do not hold any excess reserves, and the public's holdings of currency do not change.

*5. Using T-accounts, show what happens to checkable deposits in the banking system when the Fed lends an additional $1 million to the First National Bank.

6. Using T-accounts, show what happens to checkable deposits in the banking system when the Fed sells $2 million of bonds to the First National Bank.

*7. Suppose that the Fed buys $1 million of bonds from the First National Bank. If the First National Bank and all other banks use the resulting increase in reserves to purchase securities only and not to make loans, what will happen to checkable deposits?

8. If the Fed buys $1 million of bonds from the First National Bank, but an additional 10% of any deposit is held as excess reserves, what is the total increase in checkable deposits? (*Hint:* Use T-accounts to show what happens at each step of the multiple expansion process.)

*9. If a bank depositor withdraws $1,000 of currency from an account, what happens to reserves and checkable deposits?

10. If reserves in the banking system increase by $1 billion as a result of discount loans of $1 billion and checkable deposits increase by $9 billion, why isn't the banking system in equilibrium? What will continue to happen in the banking system until equilibrium is reached? Show the T-account for the banking system in equilibrium.

*11. If the Fed reduces reserves by selling $5 million worth of bonds to the banks, what will the T-account of the banking system look like when the banking system is in equilibrium? What will have happened to the level of checkable deposits?

12. If the required reserve ratio on checkable deposits increases to 20%, how much multiple deposit creation will take place when reserves are increased by $100?

*13. If a bank decides that it wants to hold $1 million of excess reserves, what effect will this have on checkable deposits in the banking system?

14. If a bank sells $10 million of bonds back to the Fed in order to pay back $10 million on the discount loan it owes, what will be the effect on the level of checkable deposits?

*15. If you decide to hold $100 less cash than usual and therefore deposit $100 in cash in the bank, what effect will this have on checkable deposits in the banking system if the rest of the public keeps its holdings of currency constant?

Web Exercises

1. Go to www.federalreserve.gov/boarddocs/rptcongress/ and find the most recent annual report of the Federal Reserve. Read the first section of the annual report that summarizes Monetary Policy and the Economic Outlook. Write a one-page summary of this section of the report.

2. Go to www.federalreserve.gov/releases/h6/hist/ and find the historical report of M1, M2, and M3. Compute the growth rate in each aggregate over each of the last 3 years (it will be easier to do if you move the data into Excel as demonstrated in Chapter 1). Does it appear that the Fed has been increasing or decreasing the rate of growth of the money supply? Is this consistent with what you understand the economy needs? Why?

appendix to chapter 15

The Fed's Balance Sheet and the Monetary Base

Just as any other bank has a balance sheet that lists its assets and liabilities, so does the Fed. We examine each of its categories of assets and liabilities because changes in them are an important way the Fed manipulates the money supply.

Assets

1. Securities. These are the Fed's holdings of securities, which consist primarily of Treasury securities but in the past have also included banker's acceptances. The total amount of securities is controlled by open market operations (the Fed's purchase and sale of these securities). As shown in Table 1, "Securities" is by far the largest category of assets in the Fed's balance sheet.

2. Discount loans. These are loans the Fed makes to banks. The amount is affected by the Fed's setting the discount rate, the interest rate that the Fed charges banks for these loans.

These first two Fed assets are important because they earn interest. Because the liabilities of the Fed do not pay interest, the Fed makes billions of dollars every year—its assets earn income, and its liabilities cost nothing. Although it returns most of its earnings to the federal government, the Fed does spend some of it on "worthy causes," such as supporting economic research.

Table 1 Consolidated Balance Sheet of the Federal Reserve System ($ billions, end of 2002)

Assets		Liabilities	
Securities: U.S. government and agency securities and banker's acceptances	668.9	Federal Reserve notes outstanding	654.3
Discount loans	10.3	Bank deposits (Reserves)	22.5
Gold and SDR certificate accounts	13.2	U.S. Treasury deposits	4.4
Coin	0.9	Foreign and other deposits	0.1
Cash items in process of collection	10.2	Deferred-availability cash items	9.4
Other Federal Reserve assets	28.5	Other Federal Reserve liabilities and capital accounts	41.3
Total	732.0	Total	732.0

Source: Federal Reserve *Bulletin*.

3. Gold and SDR certificate accounts. Special drawing rights (SDRs) are issued to governments by the International Monetary Fund (IMF) to settle international debts and have replaced gold in international financial transactions. When the Treasury acquires gold or SDRs, it issues certificates to the Fed that are claims on the gold or SDRs and is in turn credited with deposit balances at the Fed. The gold and SDR accounts are made up of these certificates issued by the Treasury.

4. Coin. This is the smallest item in the balance sheet, consisting of Treasury currency (mostly coins) held by the Fed.

5. Cash items in process of collection. These arise from the Fed's check-clearing process. When a check is given to the Fed for clearing, the Fed will present it to the bank on which it is written and will collect funds by deducting the amount of the check from the bank's deposits (reserves) with the Fed. Before these funds are collected, the check is a cash item in process of collection and is a Fed asset.

6. Other Federal Reserve assets. These include deposits and bonds denominated in foreign currencies as well as physical goods such as computers, office equipment, and buildings owned by the Federal Reserve.

Liabilities

1. Federal Reserve notes (currency) outstanding. The Fed issues currency (those green-and-gray pieces of paper in your wallet that say "Federal Reserve note" at the top). The Federal Reserve notes outstanding are the amount of this currency that is in the hands of the public. (Currency held by depository institutions is also a liability of the Fed but is counted as part of the reserves liability.)

Federal Reserve notes are IOUs from the Fed to the bearer and are also liabilities, but unlike most liabilities, they promise to pay back the bearer solely with Federal Reserve notes; that is, they pay off IOUs with other IOUs. Accordingly, if you bring a $100 bill to the Federal Reserve and demand payment, you will receive two $50s, five $20s, ten $10s, or one hundred $1 bills.

People are more willing to accept IOUs from the Fed than from you or me because Federal Reserve notes are a recognized medium of exchange; that is, they are accepted as a means of payment and so function as *money*. Unfortunately, neither you nor I can convince people that our IOUs are worth anything more than the paper they are written on.[1]

2. Reserves. All banks have an account at the Fed in which they hold deposits. Reserves consist of deposits at the Fed plus currency that is physically held by banks (called *vault cash* because it is stored in bank vaults). *Reserves* are assets for the banks but liabilities for the Fed, because the banks can demand payment on them at any time and the Fed is required to satisfy its obligation by paying Federal Reserve notes. As shown in the chapter, an increase in reserves leads to an increase in the level of deposits and hence in the money supply.

---

[1]The "Federal Reserve notes outstanding" item on the Fed's balance sheet refers only to currency in circulation, the amount in the hands of the public. Currency that has been printed by the U.S. Bureau of Engraving and Printing is not automatically a liability of the Fed. For example, consider the importance of having $1 million of your own IOUs printed up. You give out $100 worth to other people and keep the other $999,900 in your pocket. The $999,900 of IOUs does not make you richer or poorer and does not affect your indebtedness. You care only about the $100 of liabilities from the $100 of circulated IOUs. The same reasoning applies for the Fed in regard to its Federal Reserve notes.

For similar reasons, the currency component of the money supply, no matter how it is defined, includes only currency in circulation. It does not include any additional currency that is not yet in the hands of the public. The fact that currency has been printed but is not circulating means that it is not anyone's asset or liability and thus cannot affect anyone's behavior. Therefore, it makes sense not to include it in the money supply.

Total reserves can be divided into two categories: reserves that the Fed requires banks to hold (*required reserves*) and any additional reserves the banks choose to hold (*excess reserves*). For example, the Fed might require that for every dollar of deposits at a depository institution, a certain fraction (say, 10 cents) must be held as reserves. This fraction (10%) is called the *required reserve ratio*. Currently, the Fed pays no interest on reserves.

3. *U.S. Treasury deposits.* The Treasury keeps deposits at the Fed, against which it writes all its checks.

4. *Foreign and other deposits.* These include the deposits with the Fed owned by foreign governments, foreign central banks, international agencies (such as the World Bank and the United Nations), and U.S. government agencies (such as the FDIC and Federal Home Loan banks).

5. *Deferred-availability cash items.* Like cash items in process of collection, these also arise from the Fed's check-clearing process. When a check is submitted for clearing, the Fed does not immediately credit the bank that submitted the check. Instead, it promises to credit the bank within a certain prearranged time limit, which never exceeds two days. These promises are the deferred-availability items and are a liability of the Fed.

6. *Other Federal Reserve liabilities and capital accounts.* This item includes all the remaining Federal Reserve liabilities not included elsewhere on the balance sheet. For example, stock in the Federal Reserve System purchased by member banks is included here.

Monetary Base

The first two liabilities on the balance sheet, Federal Reserve notes (currency) outstanding and reserves, are often referred to as the *monetary liabilities* of the Fed. When we add to these liabilities the U.S. Treasury's monetary liabilities (Treasury currency in circulation, primarily coins), we get a construct called the *monetary base*. The monetary base is an important part of the money supply, because increases in it will lead to a multiple increase in the money supply (everything else being constant). This is why the monetary base is also called *high-powered money*. Recognizing that Treasury currency and Federal Reserve currency can be lumped together into the category *currency in circulation,* denoted by C, the monetary base equals the sum of currency in circulation plus reserves R. The monetary base MB is expressed as[2]:

$$\begin{aligned} MB &= (\text{Federal Reserve notes} + \text{Treasury currency} - \text{coin}) + \text{reserves} \\ &= C + R \end{aligned}$$

The items on the right-hand side of this equation indicate how the base is used and are called the *uses of the base*. Unfortunately, this equation does not tell us the factors that determine the base (the *sources of the base*), but the Federal Reserve balance sheet in Table 1 comes to the rescue, because like all balance sheets, it has the property that the total assets on the left-hand side must equal the total liabilities on the right-hand side. Because the "Federal Reserve notes" and "reserves" items in the uses of the base are Federal Reserve liabilities, the "assets equals liabilities" property of the Fed balance sheet enables us to solve for these items in terms of the Fed balance sheet

[2]In the member bank reserves data that the Fed publishes every week, Treasury currency outstanding is defined to include Treasury currency that is held at the Treasury (called "Treasury cash holdings"). What we have defined as "Treasury currency" is actually equal to "Treasury currency outstanding" minus "Treasury cash holdings."

items that are included in the sources of the base: Specifically, Federal Reserve notes and reserves equal the sum of all the Fed assets minus all the other Fed liabilities:

Federal Reserve notes + reserves = Securities + discount loans + gold and SDRs + coin + cash items in process of collection + other Federal Reserve assets − Treasury deposits − foreign and other deposits − deferred-availability cash items − other Federal Reserve liabilities and capital

The two balance sheet items related to check clearing can be collected into one term called *float*, defined as "Cash items in process of collection" minus "Deferred-availability cash items." Substituting all the right-hand-side items in the equation for "Federal Reserve notes + reserves" in the uses-of-the-base equation, we obtain the following expression describing the sources of the monetary base:

$$MB = \text{Securities} + \text{discount loans} + \text{gold and SDRs} + \text{float} + \text{other Federal Reserve assets} + \text{Treasury currency} - \text{Treasury deposits} - \text{foreign and other deposits} - \text{other Federal Reserve liabilities and capital} \quad (1)$$

Accounting logic has led us to a useful equation that clearly identifies the nine factors affecting the monetary base listed in Table 2. As Equation 1 and Table 2 depict, increases in the first six factors increase the monetary base, and increases in the last three reduce the monetary base.

SUMMARY **Table 2 Factors Affecting the Monetary Base**

Factor	Value ($ billions, end of 2002)	Change in Factor	Change in Monetary Base
Factors That Increase the Monetary Base			
1. Securities: U.S. government and agency securities and banker's acceptances	668.9	↑	↑
2. Discount loans	10.3	↑	↑
3. Gold and SDR certificate accounts	13.2	↑	↑
4. Float	10.2	↑	↑
5. Other Federal Reserve assets	28.5	↑	↑
6. Treasury currency	25.5	↑	↑
Subtotal 1	756.6		
Factors That Decrease the Monetary Base			
7. Treasury deposits with the Fed	4.4	↑	↓
8. Foreign and other deposits with the Fed	22.6	↑	↓
9. Other Federal Reserve liabilities and capital accounts	41.3	↑	↓
Subtotal 2	68.3		
Monetary Base			
Subtotal 1 − Subtotal 2	688.3		

Source: Federal Reserve *Bulletin*.

Chapter 16

Determinants of the Money Supply

PREVIEW

In Chapter 15, we developed a simple model of multiple deposit creation that showed how the Fed can control the level of checkable deposits by setting the required reserve ratio and the level of reserves. Unfortunately for the Fed, life isn't that simple; control of the money supply is far more complicated. Our critique of this model indicated that decisions by depositors about their holdings of currency and by banks about their holdings of excess reserves also affect the money supply. To deal with this critique, in this chapter we develop a money supply model in which depositors and banks assume their important roles. The resulting framework provides an in-depth description of the money supply process to help you understand the complexity of the Fed's role.

To simplify the analysis, we separate the development of our model into several steps. First, because the Fed can exert more precise control over the monetary base (currency in circulation plus total reserves in the banking system) than it can over total reserves alone, our model links changes in the money supply to changes in the monetary base. This link is achieved by deriving a **money multiplier** (a ratio that relates the change in the money supply to a given change in the monetary base). Finally, we examine the determinants of the money multiplier.

Study Guide

One reason for breaking the money supply model into its component parts is to help you answer questions using intuitive step-by-step logic rather than memorizing how changes in the behavior of the Fed, depositors, or banks will affect the money supply.

In deriving a model of the money supply process, we focus here on a simple definition of *money* (currency plus checkable deposits), which corresponds to M1. Although broader definitions of money—particularly, M2—are frequently used in policymaking, we conduct the analysis with an M1 definition because it is less complicated and yet provides a basic understanding of the money supply process. Furthermore, all analyses and results using the M1 definition apply equally well to the M2 definition. A somewhat more complicated money supply model for the M2 definition is developed in an appendix to this chapter, which can be viewed online at www.aw.com/mishkin.

The Money Supply Model and the Money Multiplier

Because, as we saw in Chapter 15, the Fed can control the monetary base better than it can control reserves, it makes sense to link the money supply M to the monetary base MB through a relationship such as the following:

$$M = m \times MB \tag{1}$$

The variable m is the money multiplier, which tells us how much the money supply changes for a given change in the monetary base MB. This multiplier tells us what multiple of the monetary base is transformed into the money supply. Because the money multiplier is larger than 1, the alternative name for the monetary base, *high-powered money*, is logical; a $1 change in the monetary base leads to more than a $1 change in the money supply.

The money multiplier reflects the effect on the money supply of other factors besides the monetary base, and the following model will explain the factors that determine the size of the money multiplier. Depositors' decisions about their holdings of currency and checkable deposits are one set of factors affecting the money multiplier. Another involves the reserve requirements imposed by the Fed on the banking system. Banks' decisions about excess reserves also affect the money multiplier.

Deriving the Money Multiplier

In our model of multiple deposit creation in Chapter 15, we ignored the effects on deposit creation of changes in the public's holdings of currency and banks' holdings of excess reserves. Now we incorporate these changes into our model of the money supply process by assuming that the desired level of currency C and excess reserves ER grows proportionally with checkable deposits D; in other words, we assume that the ratios of these items to checkable deposits are constants in equilibrium, as the braces in the following expressions indicate:

$$c = \{C/D\} = \text{currency ratio}$$
$$e = \{ER/D\} = \text{excess reserves ratio}$$

We will now derive a formula that describes how the currency ratio desired by depositors, the excess reserves ratio desired by banks, and the required reserve ratio set by the Fed affect the multiplier m. We begin the derivation of the model of the money supply with the equation:

$$R = RR + ER$$

which states that the total amount of reserves in the banking system R equals the sum of required reserves RR and excess reserves ER. (Note that this equation corresponds to the equilibrium condition $RR = R$ in Chapter 15, where excess reserves were assumed to be zero.)

The total amount of required reserves equals the required reserve ratio r times the amount of checkable deposits D:

$$RR = r \times D$$

Substituting $r \times D$ for RR in the first equation yields an equation that links reserves in the banking system to the amount of checkable deposits and excess reserves they can support:

$$R = (r \times D) + ER$$

A key point here is that the Fed sets the required reserve ratio r to less than 1. Thus \$1 of reserves can support more than \$1 of deposits, and the multiple expansion of deposits can occur.

Let's see how this works in practice. If excess reserves are held at zero ($ER = 0$), the required reserve ratio is set at $r = 0.10$, and the level of checkable deposits in the banking system is \$800 billion, the amount of reserves needed to support these deposits is \$80 billion ($= 0.10 \times$ \$800 billion). The \$80 billion of reserves can support ten times this amount in checkable deposits, just as in Chapter 15, because multiple deposit creation will occur.

Because the monetary base MB equals currency C plus reserves R, we can generate an equation that links the amount of monetary base to the levels of checkable deposits and currency by adding currency to both sides of the equation:

$$MB = R + C = (r \times D) + ER + C$$

Another way of thinking about this equation is to recognize that it reveals the amount of the monetary base needed to support the existing amounts of checkable deposits, currency, and excess reserves.

An important feature of this equation is that an additional dollar of MB that arises from an additional dollar of currency does not support any additional deposits. This occurs because such an increase leads to an identical increase in the right-hand side of the equation with no change occurring in D. The currency component of MB does not lead to multiple deposit creation as the reserves component does. Put another way, ***an increase in the monetary base that goes into currency is not multiplied, whereas an increase that goes into supporting deposits is multiplied.***

Another important feature of this equation is that an additional dollar of MB that goes into excess reserves ER does not support any additional deposits or currency. The reason for this is that when a bank decides to hold excess reserves, it does not make additional loans, so these excess reserves do not lead to the creation of deposits. Therefore, if the Fed injects reserves into the banking system and they are held as excess reserves, there will be no effect on deposits or currency and hence no effect on the money supply. In other words, you can think of excess reserves as an idle component of reserves that are not being used to support any deposits (although they are important for bank liquidity management, as we saw in Chapter 9). This means that for a given level of reserves, a higher amount of excess reserves implies that the banking system in effect has fewer reserves to support deposits.

To derive the money multiplier formula in terms of the currency ratio $c = \{C/D\}$ and the excess reserves ratio $e = \{ER/D\}$, we rewrite the last equation, specifying C as $c \times D$ and ER as $e \times D$:

$$MB = (r \times D) + (e \times D) + (c \times D) = (r + e + c) \times D$$

We next divide both sides of the equation by the term inside the parentheses to get an expression linking checkable deposits D to the monetary base MB:

$$D = \frac{1}{r + e + c} \times MB \qquad (2)$$

Using the definition of the money supply as currency plus checkable deposits ($M = D + C$) and again specifying C as $c \times D$,

$$M = D + (c \times D) = (1 + c) \times D$$

Substituting in this equation the expression for D from Equation 2, we have:

$$M = \frac{1 + c}{r + e + c} \times MB \tag{3}$$

Finally, we have achieved our objective of deriving an expression in the form of our earlier Equation 1. As you can see, the ratio that multiplies MB is the money multiplier that tells how much the money supply changes in response to a given change in the monetary base (high-powered money). The money multiplier m is thus:

$$m = \frac{1 + c}{r + e + c} \tag{4}$$

and it is a function of the currency ratio set by depositors c, the excess reserves ratio set by banks e, and the required reserve ratio set by the Fed r.

Although the algebraic derivation we have just completed shows you how the money multiplier is constructed, you need to understand the basic intuition behind it to understand and apply the money multiplier concept without having to memorize it.

Intuition Behind the Money Multiplier

In order to get a feel for what the money multiplier means, let us again construct a numerical example with realistic numbers for the following variables:

r = required reserve ratio = 0.10
C = currency in circulation = \$400 billion
D = checkable deposits = \$800 billion
ER = excess reserves = \$0.8 billion
M = money supply (M1) = $C + D$ = \$1,200 billion

From these numbers we can calculate the values for the currency ratio c and the excess reserves ratio e:

$$c = \frac{\$400 \text{ billion}}{\$800 \text{ billion}} = 0.5$$

$$e = \frac{\$0.8 \text{ billion}}{\$800 \text{ billion}} = 0.001$$

The resulting value of the money multiplier is:

$$m = \frac{1 + 0.5}{0.1 + 0.001 + 0.5} = \frac{1.5}{0.601} = 2.5$$

The money multiplier of 2.5 tells us that, given the required reserve ratio of 10% on checkable deposits and the behavior of depositors as represented by $c = 0.5$ and banks as represented by $e = 0.001$, a \$1 increase in the monetary base leads to a \$2.50 increase in the money supply (M1).

An important characteristic of the money multiplier is that it is less than the simple deposit multiplier of 10 found in Chapter 15. The key to understanding this result

of our money supply model is to realize that ***although there is multiple expansion of deposits, there is no such expansion for currency***. Thus if some portion of the increase in high-powered money finds its way into currency, this portion does not undergo multiple deposit expansion. In our analysis in Chapter 15, we did not allow for this possibility, and so the increase in reserves led to the maximum amount of multiple deposit creation. However, in our current model of the money multiplier, the level of currency does increase when the monetary base *MB* and checkable deposits *D* increase because *c* is greater than zero. As previously stated, any increase in *MB* that goes into an increase in currency is not multiplied, so only part of the increase in *MB* is available to support checkable deposits that undergo multiple expansion. The overall level of multiple deposit expansion must be lower, meaning that the increase in *M*, given an increase in *MB*, is smaller than the simple model in Chapter 15 indicated.[1]

Factors That Determine the Money Multiplier

To develop our intuition of the money multiplier even further, let us look at how this multiplier changes in response to changes in the variables in our model: *c*, *e*, and *r*. The "game" we are playing is a familiar one in economics: We ask what happens when one of these variables changes, leaving all other variables the same (*ceteris paribus*).

Changes in the Required Reserve Ratio *r*

If the required reserve ratio on checkable deposits increases while all the other variables stay the same, the same level of reserves cannot support as large an amount of checkable deposits; more reserves are needed because required reserves for these checkable deposits have risen. The resulting deficiency in reserves then means that banks must contract their loans, causing a decline in deposits and hence in the money supply. The reduced money supply relative to the level of *MB*, which has remained unchanged, indicates that the money multiplier has declined as well. Another way to see this is to realize that when *r* is higher, less multiple expansion of checkable deposits occurs. With less multiple deposit expansion, the money multiplier must fall.[2]

We can verify that the foregoing analysis is correct by seeing what happens to the value of the money multiplier in our numerical example when *r* increases from 10% to 15% (leaving all the other variables unchanged). The money multiplier becomes:

$$m = \frac{1 + 0.5}{0.15 + 0.001 + 0.5} = \frac{1.5}{0.651} = 2.3$$

which, as we would expect, is less than 2.5.

[1] Another reason the money multiplier is smaller is that *e* is a constant fraction greater than zero, indicating that an increase in *MB* and *D* leads to higher excess reserves. The resulting higher amount of excess reserves means that the amount of reserves used to support checkable deposits will not increase as much as it otherwise would. Hence the increase in checkable deposits and the money supply will be lower, and the money multiplier will be smaller. However, because *e* is currently so tiny—around 0.001—the impact of this ratio on the money multiplier is now quite small. But there have been periods when *e* has been much larger and so has had a more important role in lowering the money multiplier.

[2] This result can be demonstrated from the Equation 4 formula as follows: When *r* rises, the denominator of the money multiplier rises, and therefore the money multiplier must fall.

The analysis just conducted can also be applied to the case in which the required reserve ratio falls. In this case, there will be more multiple expansion for checkable deposits because the same level of reserves can now support more checkable deposits, and the money multiplier will rise. For example, if r falls from 10% to 5%, plugging this value into our money multiplier formula (leaving all the other variables unchanged) yields a money multiplier of:

$$m = \frac{1 + 0.5}{0.05 + 0.001 + 0.5} = \frac{1.5}{0.551} = 2.72$$

which is above the initial value of 2.5.

We can now state the following result: ***The money multiplier and the money supply are negatively related to the required reserve ratio r.***

Changes in the Currency Ratio *c*

Next, what happens to the money multiplier when depositor behavior causes c to increase with all other variables unchanged? An increase in c means that depositors are converting some of their checkable deposits into currency. As shown before, checkable deposits undergo multiple expansion while currency does not. Hence when checkable deposits are being converted into currency, there is a switch from a component of the money supply that undergoes multiple expansion to one that does not. The overall level of multiple expansion declines, and so must the multiplier.[3]

This reasoning is confirmed by our numerical example, where c rises from 0.50 to 0.75. The money multiplier then falls from 2.5 to:

$$m = \frac{1 + 0.75}{0.1 + 0.001 + 0.75} = \frac{1.75}{0.851} = 2.06$$

We have now demonstrated another result: ***The money multiplier and the money supply are negatively related to the currency ratio c.***

Changes in the Excess Reserves Ratio *e*

When banks increase their holdings of excess reserves relative to checkable deposits, the banking system in effect has fewer reserves to support checkable deposits. This means that given the same level of *MB*, banks will contract their loans, causing a decline in the level of checkable deposits and a decline in the money supply, and the money multiplier will fall.[4]

This reasoning is supported in our numerical example when e rises from 0.001 to 0.005. The money multiplier declines from 2.5 to:

$$m = \frac{1 + 0.5}{0.1 + 0.005 + 0.5} = \frac{1.5}{0.605} = 2.48$$

Note that although the excess reserves ratio has risen fivefold, there has been only a small decline in the money multiplier. This decline is small, because in recent years e

[3] As long as $r + e$ is less than 1 (as is the case using the realistic numbers we have used), an increase in c raises the denominator of the money multiplier proportionally by more than it raises the numerator. The increase in c causes the multiplier to fall. If you would like to know more about what explains movements in the currency ratio c, take a look at an appendix to this chapter on this topic, which can be found on this book's web site at www.aw.com/mishkin. Another appendix to this chapter, also found on the web site, discusses how the money multiplier for *M2* is determined.

[4] This result can be demonstrated from the Equation 4 formula as follows: When e rises, the denominator of the money multiplier rises, and so the money multiplier must fall.

has been extremely small, so changes in it have only a small impact on the money multiplier. However, there have been times, particularly during the Great Depression, when this ratio was far higher, and its movements had a substantial effect on the money supply and the money multiplier. Thus our final result is still an important one: ***The money multiplier and the money supply are negatively related to the excess reserves ratio e.***

To understand the factors that determine the level of e in the banking system, we must look at the costs and benefits to banks of holding excess reserves. When the costs of holding excess reserves rise, we would expect the level of excess reserves and hence e to fall; when the benefits of holding excess reserves rise, we would expect the level of excess reserves and e to rise. Two primary factors affect these costs and benefits and hence affect the excess reserves ratio: market interest rates and expected deposit outflows.

Market Interest Rates. As you may recall from our analysis of bank management in Chapter 9, the cost to a bank of holding excess reserves is its opportunity cost, the interest that could have been earned on loans or securities if they had been held instead of excess reserves. For the sake of simplicity, we assume that loans and securities earn the same interest rate i, which we call the market interest rate. If i increases, the opportunity cost of holding excess reserves rises, and the desired ratio of excess reserves to deposits falls. A decrease in i, conversely, will reduce the opportunity cost of excess reserves, and e will rise. ***The banking system's excess reserves ratio e is negatively related to the market interest rate i.***

Another way of understanding the negative effect of market interest rates on e is to return to the theory of asset demand, which states that if the expected returns on alternative assets rise relative to the expected returns on a given asset, the demand for that asset will decrease. As the market interest rate increases, the expected return on loans and securities rises relative to the zero return on excess reserves, and the excess reserves ratio falls.

Figure 1 shows us (as the theory of asset demand predicts) that there is a negative relationship between the excess reserves ratio and a representative market interest rate, the federal funds rate. The period 1960–1981 saw an upward trend in the federal funds rate and a declining trend in e, whereas in the period 1981–2002, a decline in the federal funds rate is associated with a rise in e. The empirical evidence thus supports our analysis that the excess reserves ratio is negatively related to market interest rates.

Expected Deposit Outflows. Our analysis of bank management in Chapter 9 also indicated that the primary benefit to a bank of holding excess reserves is that they provide insurance against losses due to deposit outflows; that is, they enable the bank experiencing deposit outflows to escape the costs of calling in loans, selling securities, borrowing from the Fed or other corporations, or bank failure. If banks fear that deposit outflows are likely to increase (that is, if expected deposit outflows increase), they will want more insurance against this possibility and will increase the excess reserves ratio. Another way to put it is this: If expected deposit outflows rise, the expected benefits, and hence the expected returns for holding excess reserves, increase. As the theory of asset demand predicts, excess reserves will then rise. Conversely, a decline in expected deposit outflows will reduce the insurance benefit

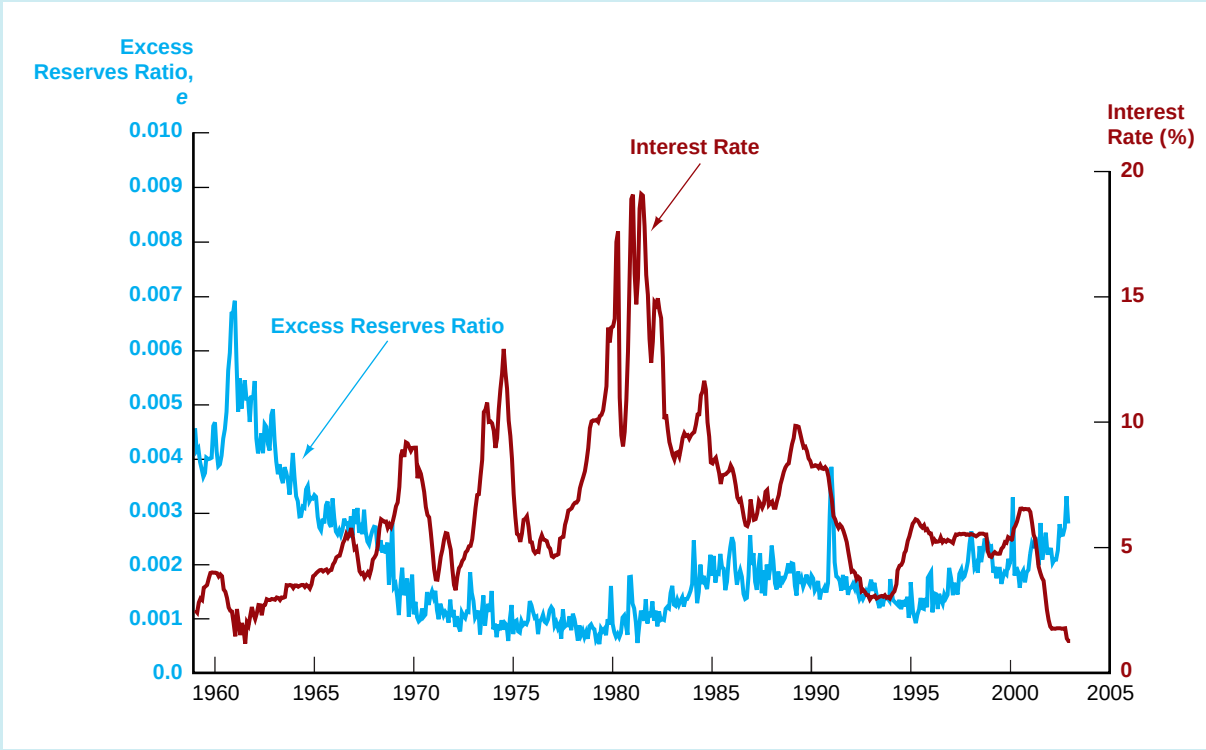


FIGURE 1 The Excess Reserves Ratio *e* and the Interest Rate (Federal Funds Rate)

Source: Federal Reserve: www.federalreserve.gov/releases/h3/hist/h3hist2.txt.

of excess reserves, and their level should fall. We have the following result: ***The excess reserves ratio e is positively related to expected deposit outflows.***

Additional Factors That Determine the Money Supply

So far we have been assuming that the Fed has accurate control over the monetary base. However, whereas the amount of open market purchases or sales is completely controlled by the Fed's placing orders with dealers in bond markets, the central bank cannot unilaterally determine, and therefore cannot perfectly predict, the amount of borrowing by banks from the Fed. The Federal Reserve sets the discount rate (interest rate on discount loans), and then banks make decisions about whether to borrow. The amount of discount loans, though influenced by the Fed's setting of the discount rate, is not completely controlled by the Fed; banks' decisions play a role, too.

Therefore, we might want to split the monetary base into two components: one that the Fed can control completely and another that is less tightly controlled. The less tightly controlled component is the amount of the base that is created by discount loans from the Fed. The remainder of the base (called the **nonborrowed monetary**

base) is under the Fed's control, because it results primarily from open market operations.[5] The nonborrowed monetary base is formally defined as the monetary base minus discount loans from the Fed:

$$MB_n = MB - DL$$

where MB_n = nonborrowed monetary base
MB = monetary base
DL = discount loans from the Fed

The reason for distinguishing the nonborrowed monetary base MB_n from the monetary base MB is that the nonborrowed monetary base, which is tied to open market operations, is directly under the control of the Fed, whereas the monetary base, which is also influenced by discount loans from the Fed, is not.

To complete the money supply model, we use $MB = MB_n + DL$ and rewrite the money supply model as:

$$M = m \times (MB_n + DL) \tag{5}$$

where the money multiplier m is defined as in Equation 4. Thus in addition to the effects on the money supply of the required reserve ratio, currency ratio, and excess reserves ratio, the expanded model stipulates that the money supply is also affected by changes in MB_n and DL. Because the money multiplier is positive, Equation 5 immediately tells us that the money supply is positively related to both the nonborrowed monetary base and discount loans. However, it is still worth developing the intuition for these results.

Changes in the Nonborrowed Monetary Base MB_n

As shown in Chapter 15, the Fed's open market purchases increase the nonborrowed monetary base, and its open market sales decrease it. Holding all other variables constant, an increase in MB_n arising from an open market purchase increases the amount of the monetary base that is available to support currency and deposits, so the money supply will increase. Similarly, an open market sale that decreases MB_n shrinks the amount of the monetary base available to support currency and deposits, thereby causing the money supply to decrease.

We have the following result: ***The money supply is positively related to the nonborrowed monetary base MB_n.***

Changes in Discount Loans *DL* from the Fed

With the nonborrowed monetary base MB_n unchanged, more discount loans from the Fed provide additional reserves (and hence higher MB) to the banking system, and these are used to support more currency and deposits. As a result, the increase in DL will lead to a rise in the money supply. If banks reduce the level of their discount loans, with all other variables held constant, the amount of MB available to support currency and deposits will decline, causing the money supply to decline.

[5] Actually, there are other items on the Fed's balance sheet (discussed in the appendix on the web site) that affect the magnitude of the nonborrowed monetary base. Since their effects on the nonborrowed base relative to open market operations are both small and predictable, these other items do not present the Fed with difficulties in controlling the nonborrowed base.

The result is this: ***The money supply is positively related to the level of discount loans DL from the Fed.*** However, because the Federal Reserve now (since January 2003) keeps the interest rate on discount loans (the discount rate) above market interest rates at which banks can borrow from each other, banks usually have little incentive to take out discount loans. Discount lending, *DL*, is thus very small except under exceptional circumstances that will be discussed in the next chapter.

Overview of the Money Supply Process

We now have a model of the money supply process in which all four of the players—the Federal Reserve System, depositors, banks, and borrowers from banks—directly influence the money supply. As a study aid, Table 1 charts the money supply (M1) response to the six variables discussed and gives a brief synopsis of the reasoning behind each result.

Study Guide

To improve your understanding of the money supply process, slowly work through the logic behind the results in Table 1 rather than just memorizing the results. Then see if you can construct your own table in which all the variables decrease rather than increase.

SUMMARY **Table 1 Money Supply (M1) Response**

Player	Variable	Change in Variable	Money Supply Response	Reason
Federal Reserve System	r	↑	↓	Less multiple deposit expansion
	MB_n	↑	↑	More *MB* to support *D* and *C*
	DL	↑	↑	More *MB* to support *D* and *C*
Depositors	c	↑	↓	Less multiple deposit expansion
Depositors and banks	Expected deposit outflows	↑	↓	e ↑ so fewer reserves to support *D*
Borrowers from banks and the other three players	i	↑	↑	e ↓ so more reserves to support *D*

Note: Only increases (↑) in the variables are shown. The effects of decreases on the money supply would be the opposite of those indicated in the "Money Supply Response" column.

www.federalreserve.gov/Releases/h3/

The Federal Reserve web site reports data about aggregate reserves and the monetary base. This site also reports on the volume of discount window lending.

www.federalreserve.gov/Releases/h6/

This site reports current and historical levels of M1, M2, and M3, and other data on the money supply.

The variables are grouped by the player or players who either influence the variable or are most influenced by it. The Federal Reserve, for example, influences the money supply by controlling the first three variables—r, MB_n, and DL, also known as the tools of the Fed. (How these tools are used is discussed in subsequent chapters.) Depositors influence the money supply through their decisions about the currency ratio c, while banks influence the money supply by their decisions about e, which are affected by their expectations about deposit outflows. Because depositors' behavior also influences bankers' expectations about deposit outflows, this variable also reflects the role of both depositors and bankers in the money supply process. Market interest rates, as represented by i, affect the money supply through the excess reserves ratio e. As shown in Chapter 5, the demand for loans by borrowers influences market interest rates, as does the supply of money. Therefore, all four players are important in the determination of i.

Application

Explaining Movements in the Money Supply, 1980–2002

To make the theoretical analysis of this chapter more concrete, we need to see whether the model of the money supply process developed here helps us understand recent movements of the money supply. We look at money supply movements from 1980 to 2002—a particularly interesting period, because the growth rate of the money supply displayed unusually high variability.

Figure 2 shows the movements of the money supply (M1) from 1980 to 2002, with the percentage next to each bracket representing the annual growth rate for the bracketed period: From January 1980 to October 1984, for example, the money supply grew at a 7.2% annual rate. The variability of money growth in the 1980–2002 period is quite apparent, swinging from 7.2% to 13.1%, down to 3.3%, then up to 11.1% and finally back down to 2.3%. What explains these sharp swings in the growth rate of the money supply?

Our money supply model, as represented by Equation 5, suggests that the movements in the money supply that we see in Figure 2 are explained by either changes in $MB_n + DL$ (the nonborrowed monetary base plus discount loans) or by changes in m (the money multiplier). Figure 3 plots these variables and shows their growth rates for the same bracketed periods as in Figure 2.

Over the whole period, the average growth rate of the money supply (5.3%) is reasonably well explained by the average growth rate of the nonborrowed monetary base MB_n (7.4%). In addition, we see that DL is rarely an important source of fluctuations in the money supply since $MB_n + DL$ is closely tied to MB_n except for the unusual period in 1984 and September 2001 when discount loans increased dramatically. (Both of these episodes involved emergency lending by the Fed and are discussed in the following chapter.)

The conclusion drawn from our analysis is this: ***Over long periods, the primary determinant of movements in the money supply is the nonborrowed monetary base MB_n, which is controlled by Federal Reserve open market operations.***

For shorter time periods, the link between the growth rates of the nonborrowed monetary base and the money supply is not always close, primarily because the money multiplier m experiences substantial short-run swings

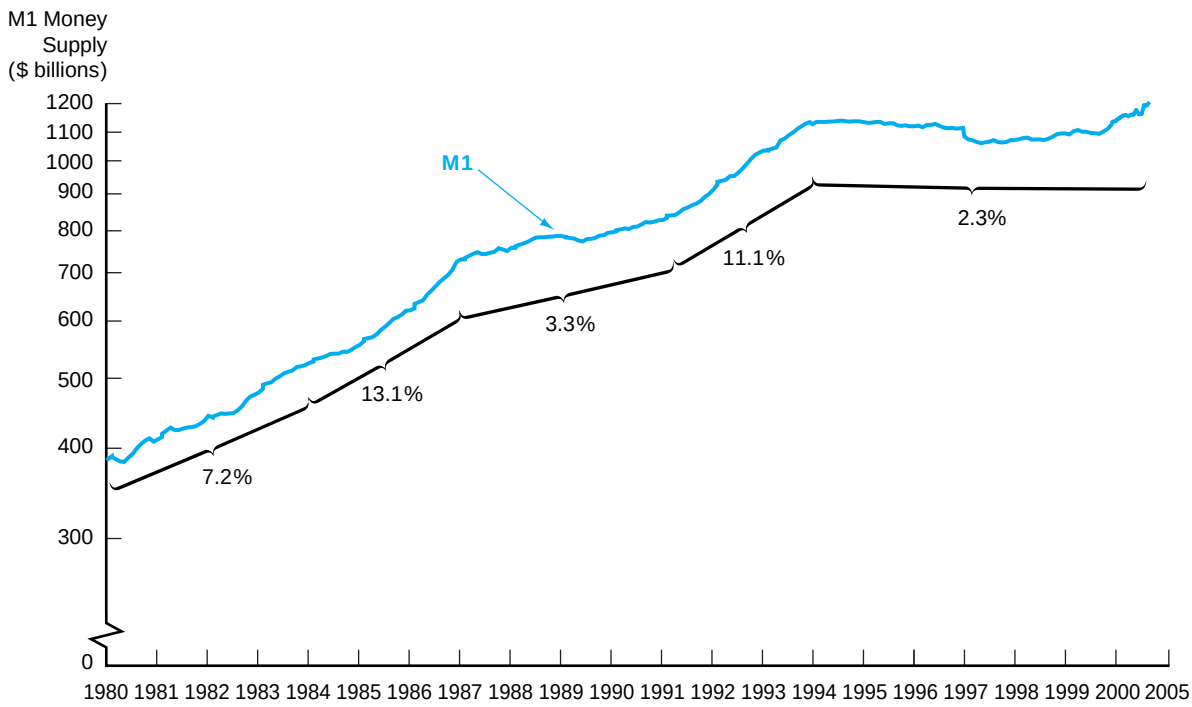


FIGURE 2 Money Supply (M1), 1980–2002

Percentage for each bracket indicates the annual growth rate of the money supply over the bracketed period.

Source: Federal Reserve: www.federalreserve.gov/releases.

that have a major impact on the growth rate of the money supply. The currency ratio c, which is also plotted in Figure 3, explains most of these movements in the money multiplier.

From January 1980 until October 1984, c is relatively constant. Unsurprisingly, there is almost no trend in the money multiplier m, so the growth rates of the money supply and the nonborrowed monetary base have similar magnitudes. The upward movement in the money multiplier from October 1984 to January 1987 is explained by the downward trend in the currency ratio. The decline in c meant that there was a shift from one component of the money supply with less multiple expansion (currency) to one with more (checkable deposits), so the money multiplier rose. In the period from January 1987 to April 1991, c underwent a substantial rise. As our money supply model predicts, the rise in c led to a fall in the money multiplier, because there was a shift from checkable deposits, with more multiple expansion, to currency, which had less. From April 1991 to December 1993, c fell somewhat. The decline in c led to a rise in the money multiplier, because there was again a shift from the currency component of the money supply with less multiple expansion to the checkable deposits component

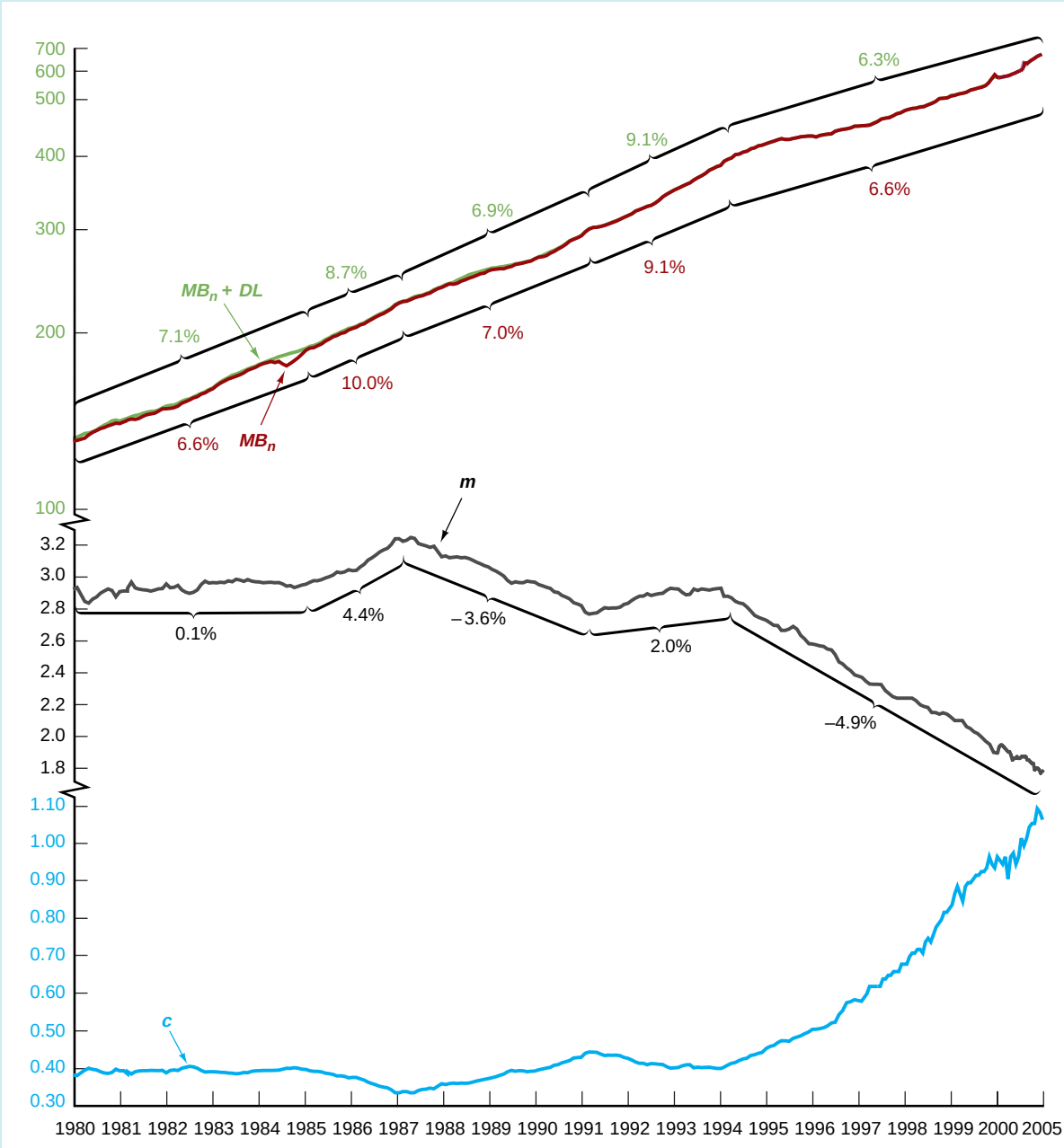


FIGURE 3 Determinants of the Money Supply, 1980–2002

Percentage for each bracket indicates the annual growth rate of the series over the bracketed period.

Source: Federal Reserve: www.federalreserve.gov/releases.

with more. Finally, the sharp rise in c from December 1993 to December 2002 should have led to a decline in the money multiplier, because the shift into currency produces less multiple deposit expansion. As our money supply model predicts, the money multiplier did indeed fall sharply in this period, and there was a dramatic deceleration of money growth.

Although our examination of the 1980–2002 period indicates that factors such as changes in c can have a major impact on the money supply over short periods, we must not forget that over the entire period, the growth rate of the money supply is closely linked to the growth rate of the nonborrowed monetary base MB_n. Indeed, empirical evidence suggests that more than three-fourths of the fluctuations in the money supply can be attributed to Federal Reserve open market operations, which determine MB_n.

Application

The Great Depression Bank Panics, 1930–1933

We can also use our money supply model to help us understand major movements in the money supply that have occurred in the past. In this application, we use the model to explain the monetary contraction that occurred during the Great Depression, the worst economic downturn in U.S. history. In Chapter 8, we discussed bank panics and saw that they could harm the economy by making asymmetric information problems more severe in credit markets, as they did during the Great Depression. Here we can see that another consequence of bank panics is that they can cause a substantial reduction in the money supply. As we will see in the chapters on monetary theory later in the book, such reductions can also cause severe damage to the economy.

Figure 4 traces the bank crisis during the Great Depression by showing the volume of deposits at failed commercial banks from 1929 to 1933. In their classic book *A Monetary History of the United States, 1867–1960*, Milton Friedman and Anna Schwartz describe the onset of the first banking crisis in late 1930 as follows:

> Before October 1930, deposits of suspended [failed] commercial banks had been somewhat higher than during most of 1929 but not out of line with experience during the preceding decade. In November 1930, they were more than double the highest value recorded since the start of monthly data in 1921. A crop of bank failures, particularly in Missouri, Indiana, Illinois, Iowa, Arkansas, and North Carolina, led to widespread attempts to convert checkable and time deposits into currency, and also, to a much lesser extent, into postal savings deposits. A contagion of fear spread among depositors, starting from the agricultural areas, which had experienced the heaviest impact of bank failures in the twenties. But failure of 256 banks with $180 million of deposits in November 1930 was followed by the failure of 532 with over $370 million of deposits in December (all figures seasonally unadjusted), the most dramatic being the failure on December 11 of the Bank of United States with over

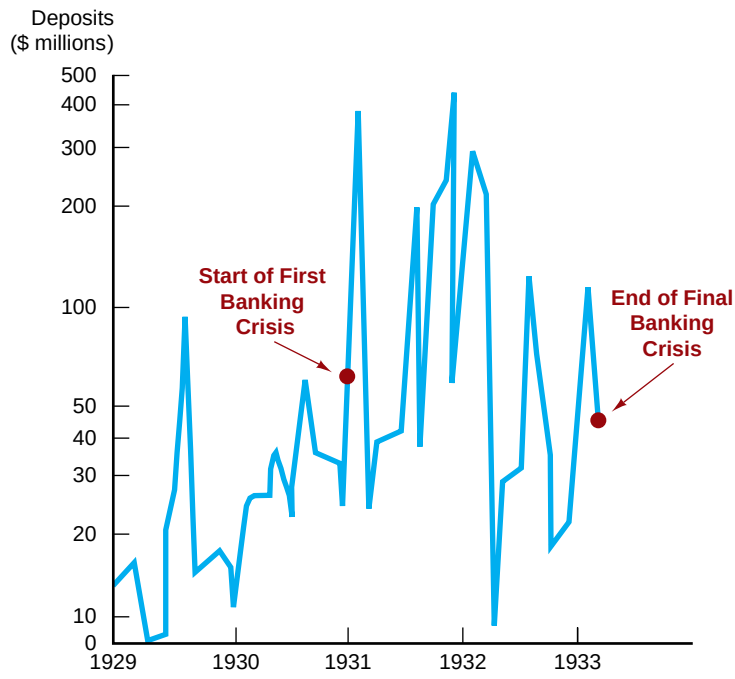


FIGURE 4 Deposits of Failed Commercial Banks, 1929–1933

Source: Milton Friedman and Anna Jacobson Schwartz, *A Monetary History of the United States, 1867–1960* (Princeton, N.J.: Princeton University Press, 1963), p. 309.

$200 million of deposits. That failure was especially important. The Bank of United States was the largest commercial bank, as measured by volume of deposits, ever to have failed up to that time in U.S. history. Moreover, though it was just an ordinary commercial bank, the Bank of United States's name had led many at home and abroad to regard it somehow as an official bank, hence its failure constituted more of a blow to confidence than would have been administered by the fall of a bank with a less distinctive name.[6]

The first bank panic, from October 1930 to January 1931, is clearly visible in Figure 4 at the end of 1930, when there is a rise in the amount of deposits at failed banks. Because there was no deposit insurance at the time (the FDIC wasn't established until 1934), when a bank failed, depositors would receive only partial repayment of their deposits. Therefore, when banks were failing during a bank panic, depositors knew that they would be likely to suffer substantial losses on deposits and thus the expected return on deposits would be negative. The theory of asset demand predicts that with the onset of the first bank crisis, depositors would shift their holdings from checkable deposits to currency by withdrawing currency from their bank

[6]Milton Friedman and Anna Jacobson Schwartz, *A Monetary History of the United States, 1867–1960* (Princeton, N.J.: Princeton University Press, 1963), pp. 308–311.

accounts, and *c* would rise. Our earlier analysis of the excess reserves ratio suggests that the resulting surge in deposit outflows would cause the banks to protect themselves by substantially increasing their excess reserves ratio *e*. Both of these predictions are borne out by the data in Figure 5. During the first bank panic (October 1930–January 1931) *c* began to climb. Even more striking is the behavior of *e*, which more than doubled from November 1930 to January 1931.

The money supply model predicts that when *e* and *c* increase, the money supply will fall. The rise in *c* results in a decline in the overall level of multiple deposit expansion, leading to a smaller money multiplier and a decline in the money supply, while the rise in *e* reduces the amount of reserves available to support deposits and also causes the money supply to fall. Thus our model predicts that the rise in *e* and *c* after the onset of the first bank crisis would result in a decline in the money supply—a prediction borne out by the evidence in Figure 6. The money supply declined sharply in December 1930 and January 1931 during the first bank panic.

Banking crises continued to occur from 1931 to 1933, and the pattern predicted by our model persisted: *c* continued to rise, and so did *e*. By the

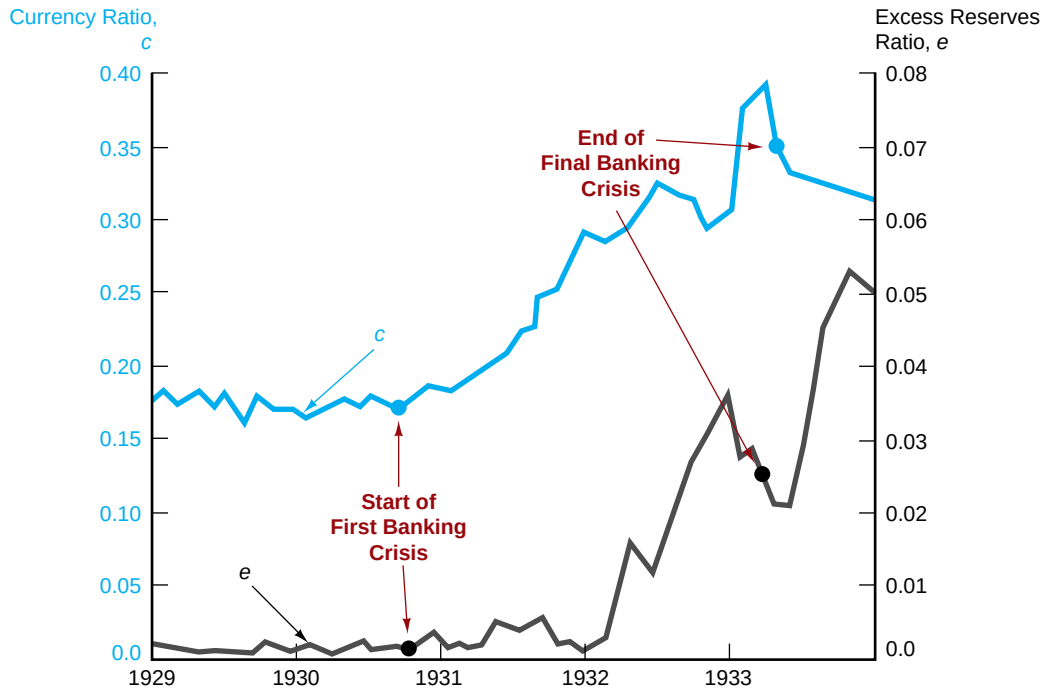


FIGURE 5 Excess Reserves Ratio and Currency Ratio, 1929–1933

Sources: Federal Reserve *Bulletin*; Milton Friedman and Anna Jacobson Schwartz, *A Monetary History of the United States, 1867–1960* (Princeton, N.J.: Princeton University Press, 1963), p. 333.

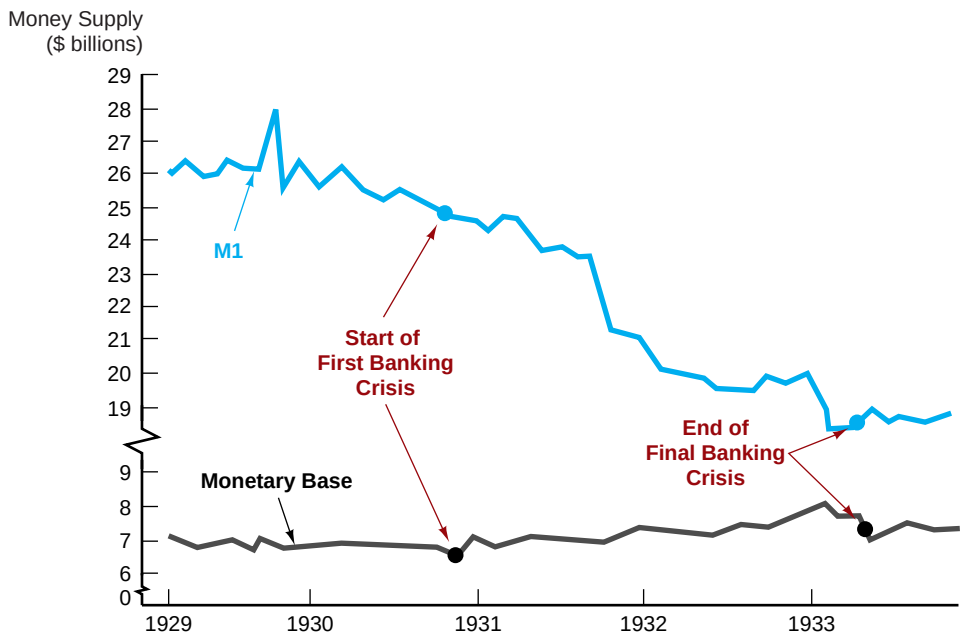


FIGURE 6 M1 and the Monetary Base, 1929–1933

Source: Milton Friedman and Anna Jacobson Schwartz, *A Monetary History of the United States, 1867–1960* (Princeton, N.J.: Princeton University Press, 1963), p. 333.

end of the crises in March 1933, the money supply (M1) had declined by over 25%—by far the largest decline in all of American history—and it coincided with the nation's worst economic contraction (see Chapter 8). Even more remarkable is that this decline occurred despite a 20% rise in the level of the monetary base—which illustrates how important the changes in c and e during bank panics can be in the determination of the money supply. It also illustrates that the Fed's job of conducting monetary policy can be complicated by depositor and bank behavior.

Summary

1. We developed a model to describe how the money supply is determined. First, we linked the monetary base to the money supply using the concept of the money multiplier, which tells us how much the money supply changes when there is a change in the monetary base.
2. The money supply is negatively related to the required reserve ratio r, the currency ratio c, and the excess reserves ratio e. It is positively related to the level of discount loans DL from the Fed and the nonborrowed base MB_n, which is determined by Fed open market

operations. The money supply model therefore allows for the behavior of all four players in the money supply process: the Fed through its setting of the required reserve ratio, the discount rate, and open market operations; depositors through their decisions about the currency ratio; the banks through their decisions about the excess reserves ratio and discount loans from the Fed; and borrowers from banks indirectly through their effect on market interest rates, which affect bank decisions regarding the excess reserves ratio and borrowings from the Fed.

Key Terms

money multiplier, p. 374

nonborrowed monetary base, p. 381

Questions and Problems

Questions marked with an asterisk are answered at the end of the book in an appendix, "Answers to Selected Questions and Problems."

*1. "The money multiplier is necessarily greater than 1." Is this statement true, false, or uncertain? Explain your answer.

2. "If reserve requirements on checkable deposits were set at zero, the amount of multiple deposit expansion would go on indefinitely." Is this statement true, false, or uncertain? Explain.

*3. During the Great Depression years 1930–1933, the currency ratio c rose dramatically. What do you think happened to the money supply? Why?

4. During the Great Depression, the excess reserves ratio e rose dramatically. What do you think happened to the money supply? Why?

*5. Traveler's checks have no reserve requirements and are included in the M1 measure of the money supply. When people travel during the summer and convert some of their checking account deposits into traveler's checks, what happens to the money supply? Why?

6. If Jane Brown closes her account at the First National Bank and uses the money instead to open a money market mutual fund account, what happens to M1? Why?

*7. Some experts have suggested that reserve requirements on checkable deposits and time deposits should be set equal because this would improve control of M2. Does this argument make sense? (*Hint:* Look at the second appendix to this chapter and think about what happens when checkable deposits are converted into time deposits or vice versa.)

8. Why might the procyclical behavior of interest rates (rising during business cycle expansions and falling during recessions) lead to procyclical movements in the money supply?

Using Economic Analysis to Predict the Future

*9. The Fed buys $100 million of bonds from the public and also lowers r. What will happen to the money supply?

10. The Fed has been discussing the possibility of paying interest on excess reserves. If this occurred, what would happen to the level of e?

*11. If the Fed sells $1 million of bonds and banks reduce their discount loans by $1 million, predict what will happen to the money supply.

12. Predict what will happen to the money supply if there is a sharp rise in the currency ratio.

*13. What do you predict would happen to the money supply if expected inflation suddenly increased?

14. If the economy starts to boom and loan demand picks up, what do you predict will happen to the money supply?

*15. Milton Friedman once suggested that Federal Reserve discount lending should be abolished. Predict what would happen to the money supply if Friedman's suggestion were put into practice.

Web Exercises

1. An important aspect of the supply of money is reserve balances. Go to www.federalreserve.gov/Releases/h41/ and locate the most recent release. This site reports changes in factors that affect depository reserve balances.
   a. What is the current reserve balance?
   b. What is the change in reserve balances since a year ago?
   c. Based on Questions a and b, does it appear that the money supply should be increasing or decreasing?

2. Refer to Figure 3: Determinants of the Money Supply, 1980–2002. Go to www.federalreserve.gov/Releases/h3/Current/ where the monetary base (*MB*) and borrowings (*DL*) are reported. Compute the growth rate in *MB* + *DL* since the end of 2002. How does this compare to previous periods reported on the graph?

appendix 1 to chapter 16

The M2 Money Multiplier

The derivation of a money multiplier for the M2 definition of money requires only slight modifications to the analysis in the chapter. The definition of M2 is:

$$M2 = C + D + T + MMF$$

where
C = currency in circulation
D = checkable deposits
T = time and savings deposits
MMF = primarily money market mutual fund shares and money market deposit accounts, plus overnight repurchase agreements and overnight Eurodollars

We again assume that all desired quantities of these variables rise proportionally with checkable deposits so that the equilibrium ratios c, $t = \{T/D\}$, and $mm = \{MMF/D\}$ set by depositors are treated as constants. Replacing C by $c \times D$, T by $t \times D$, and MMF by $mm \times D$ in the definition of M2 just given, we get:

$$\begin{aligned} M2 &= D + (c \times D) + (t \times D) + (mm \times D) \\ &= (1 + c + t + mm) \times D \end{aligned}$$

Substituting in the expression for D from Equation 2 in the chapter,[1] we have

$$M2 = \frac{1 + c + t + mm}{r + e + c} \times MB \qquad (1)$$

To see what this formula implies about the M2 money multiplier, we continue with the same numerical example in the chapter, with the additional information that T = \$2,400 billion and MMF = \$400 billion so that $t = 3$ and $mm = 0.5$. The resulting value of the multiplier for M2 is:

$$m_2 = \frac{1 + 0.5 + 3 + 0.5}{0.10 + 0.001 + 0.5} = \frac{5.0}{0.601} = 8.32$$

[1] From the derivation here it is clear that the quantity of checkable deposits D is unaffected by the depositor ratios t and mm even though time deposits and money market mutual fund shares are included in M2. This is just a consequence of the absence of reserve requirements on time deposits and money market mutual fund shares, so T and MMF do not appear in any of the equations in the derivation of D in the chapter.

An important feature of the M2 multiplier is that it is substantially above the M1 multiplier of 2.5 that we found in the chapter. The crucial concept in understanding this difference is that a lower required reserve ratio for time deposits or money market mutual fund shares means that they undergo more multiple expansion because fewer reserves are needed to support the same amount of them. Time deposits and *MMF*s have a lower required reserve ratio than checkable deposits—zero—and they will therefore have more multiple expansion than checkable deposits will. Thus the overall multiple expansion for the sum of these deposits will be greater than for checkable deposits alone, and so the M2 money multiplier will be greater than the M1 money multiplier.

Factors That Determine the M2 Money Multiplier

Changes in *r*, *c*, and *e*

The economic reasoning analyzing the effect of changes in the required reserve ratio and the currency ratio on the M2 money multiplier is identical to that used for the M1 multiplier in the chapter. An increase in the required reserve ratio r will decrease the amount of multiple deposit expansion, thus lowering the M2 money multiplier. An increase in c means that depositors have shifted out of checkable deposits into currency, and since currency has no multiple deposit expansion, the overall level of multiple deposit expansion for M2 must also fall, lowering the M2 multiplier. An increase in the excess reserves ratio e means that banks use fewer reserves to support deposits, so deposits and the M2 money multiplier fall.

We thus have the same results we found for the M1 multiplier: ***The M2 money multiplier and M2 money supply are negatively related to the required reserve ratio r, the currency ratio c, and the excess reserves ratio e.***

Response to Changes in *t* and *mm*

An increase in either t or mm leads to an increase in the M2 multiplier, because the required reserve ratios on time deposits and money market mutual fund shares are zero and hence are lower than the required reserve ratio on checkable deposits.

Both time deposits and money market mutual fund shares undergo more multiple expansion than checkable deposits. Thus a shift out of checkable deposits into time deposits or money market mutual funds, increasing t or mm, implies that the overall level of multiple expansion will increase, raising the M2 money multiplier.

A decline in t or mm will result in less overall multiple expansion, and the M2 money multiplier will decrease, leading to the following conclusion: ***The M2 money multiplier and M2 money supply are positively related to both the time deposit ratio t and the money market fund ratio mm.***

The response of the M2 money supply to all the depositor and required reserve ratios is summarized in Table 1.

SUMMARY **Table 1 Response of the M2 Money Supply to Changes in MB_n, *DL*, *r*, *e*, *c*, *t*, and *mm***

Variable	Change in Variable	M2 Money Supply Response	Reason
MB_n	↑	↑	More *MB* to support *C* and *D*
DL	↑	↑	More *MB* to support *C* and *D*
r	↑	↓	Less multiple deposit expansion
e	↑	↓	Fewer reserves to support *C* and *D*
c	↑	↓	Less overall deposit expansion
t	↑	↑	More multiple deposit expansion
mm	↑	↑	More multiple deposit expansion

Note: Only increases (↑) in the variables are shown; the effects of decreases in the variables on the money multiplier would be the opposite of those indicated in the "Response" column.

appendix 2 to chapter 16

Expanding Behavior of the Currency Ratio

The general outline of the movements of the currency ratio c since 1892 is shown in Figure 1. As you can see, several episodes stand out:

1. The declining trend in the ratio from 1892 until 1917, when the United States entered World War I
2. The sharp increase in the ratio during World War I and the decline thereafter
3. The steepest increase in the ratio that we see in the figure, which occurs during the Great Depression years from 1930 to 1933
4. The increase in the ratio during World War II
5. The reversal in the early 1960s of the downward trend in the ratio and the rise thereafter
6. The halt in the upward trend from 1980 to 1993
7. The upward trend from 1994 to 2002

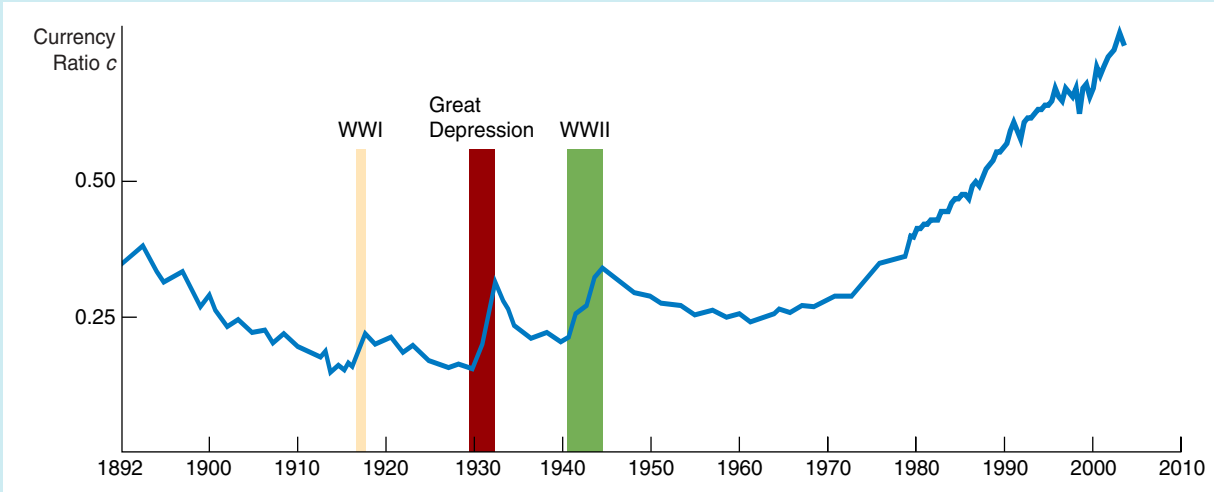


FIGURE 1 Currency-Checkable Deposits Ratio: 1892–2002

Sources: Federal Reserve *Bulletin* and *Banking and Monetary Statistics*. www.federalreserve.gov/releases/h6/hist/h6hist1.txt

To be worthwhile, our analysis of c must be able to explain these movements. These movements, however, will help us develop the analysis because they provide clues to the factors that influence c.

A natural way to approach the analysis of the relative amount of assets (currency and checkable deposits) people want to hold, hence the currency-checkable deposits ratio, is to use the theory of asset demand developed in Chapter 5. Recall the theory states that four categories of factors influence the demand for an asset such as currency or checkable deposits: (1) the total resources available to individuals, that is, wealth; (2) the expected return on one asset relative to the expected return on alternative assets; (3) the degree of uncertainty or risk associated with the return from this asset relative to the alternative assets; and (4) the liquidity of one asset relative to alternative assets. Because risk and liquidity factors have not changed independently of wealth and expected returns and lead to similar conclusions on the historical movements of c, we will focus only on how factors affecting wealth and expected returns influence c.

Effect of Changes in Wealth

What is the relative response of currency to checkable deposits when an individual's resources change? Currency is a necessity because it is used extensively by people with low incomes and little wealth, which means that the demand for currency grows proportionately less with accumulation of wealth. In contrast, checkable deposits are held by people with greater wealth, so checkable deposits are less of a necessity. Put another way, as wealth grows, the holdings of checkable deposits relative to the holdings of currency increase, and the amount of currency relative to checkable deposits falls, causing the currency ratio c to decline. A decrease in income will lead to an increase in the amount of currency relative to checkable deposits, causing c to increase. ***The currency ratio is negatively related to income or wealth.***

Effect of Changes in Expected Returns

The second factor that influences the decision to hold currency versus checkable deposits involves the expected returns on the checkable deposits relative to currency and other assets. Four primary factors influence expected returns (and hence the currency ratio): interest rates on checkable deposits, the cost of acquiring currency, bank panics, and illegal activity.[1]

Interest Rates on Checkable Deposits. By its very nature, currency cannot pay interest. Yet banks can and do pay interest on checkable deposits. One measure of the expected return on checkable deposits relative to currency is the interest rate on checkable deposits. As this interest rate increases, the theory of asset demand tells us that people will want to hold less currency relative to checkable deposits, and c will fall. Conversely, a decline in this interest rate will cause c to rise. ***The currency ratio is negatively related to the interest rate paid on checkable deposits.***

Between 1933 and 1980, regulations prevented banks from paying interest on most checkable deposits,[2] and before 1933, these interest rates were low and did not

[1]Changes in interest rates on other alternative assets (such as U.S. Treasury bills) could have a differential effect on the demand for currency versus the checkable deposits, resulting in some effect on c. However, the evidence for this effect is weak.

[2]Although banks could not pay interest on checkable deposits, they provided services to their checking account customers that can be thought of as implicit interest payments. Because these services changed only slowly over time, these implicit interest payments were not a major factor causing the demand for checkable deposits to fluctuate.

undergo substantial fluctuations. However, since 1980, banks have been allowed to pay any interest rate they choose on checkable deposits, suggesting that fluctuations in these rates can now be an important factor influencing *c* movements.

Cost of Acquiring Currency. If currency is made easier to acquire, thereby lowering the cost of using it, then in effect its expected return rises relative to deposits and the currency ratio *c* should rise. ***Lowering the cost of acquiring currency leads to a rise in the currency ratio.*** The explosion of ATMs in recent years has indeed made it easier for depositors at banks to get their hands on currency and should thus have increased its use, raising *c*.

Bank Panics. Our discussion of interest-rate effects suggests that they did not have a substantial impact on *c* before 1980. You might conclude that expected returns have had little importance in determining this ratio for most of its history. Figure 1 provides us with a clue that we are overlooking an important factor when measuring expected returns solely by the interest rates on assets. The steepest rise in *c* occurred during the Great Depression years 1930–1933, when the banking system nearly collapsed. Legend has it that during this period, people stuffed their mattresses with cash rather than keep it in banks, because they had lost confidence in them as a safe haven for their hard-earned savings. Can the theory of asset demand explain this phenomenon?

A bank failure occurs when the bank is no longer able to pay back its depositors. Before creation of the FDIC in 1933, if you had an account at a bank that failed, you would suffer a substantial loss—you could not withdraw your savings and might receive only a small fraction of the value of your deposits sometime in the future. The simultaneous failure of many banks is called a *bank panic*, and the Great Depression years 1930–1933 witnessed the worst set of bank panics in U.S. history. From the end of 1930 to the bank holiday in March 1933, more than one-third of the banks in the United States failed.

Bank panics can have a devastating effect on the expected returns from holding deposits. When a bank is likely to fail during a bank panic, depositors know that if they have deposits in this bank, they are likely to suffer substantial losses, and the expected return on deposits can be negative. The theory of asset demand predicts that depositors will shift their holdings from checkable deposits to currency by withdrawing currency from their bank accounts, and *c* will rise. This is exactly what we see in Figure 1 during the bank panics of the Great Depression period 1930–1933 and to a lesser extent 1893 and 1907, when smaller-scale bank panics occurred. *The conclusion is that* ***bank panics lead to a sharp increase in the currency ratio.*** Bank panics have been an important source of fluctuations in this ratio in the past and could be important in the future.

Illegal Activity. Expected returns on checkable deposits relative to currency can also be affected by the amount of illegal activity conducted in an economy. U.S. law allows government prosecutors access to bank records when conducting a criminal investigation. So if you were engaged in some illegal activity, you would not conduct your transactions with checks, because they are traceable and therefore a potentially powerful piece of evidence against you. Currency, however, is much harder to trace. The expected return on currency relative to checkable deposits is higher when you are engaged in illegal transactions. Hence when illegal activity in a society increases, there

is an increase in the use of currency relative to checkable deposits, and c rises. ***There is a positive association between illegal activity and the currency ratio.***[3]

Looking at Figure 1, what types of increases in illegal activity would lead to an increase in c? Beginning in the 1960s, c began to climb—just when the illegal drug trade began to experience phenomenal growth. Because illegal drug transactions are always carried out with currency, it is likely that the rise in drug trade is related to the rise in c. Supporting evidence is the current huge flow of currency into southern Florida, the major center for illegal drug importing in the United States.[4] Other illegal activities—prostitution, black markets, gambling, loan sharking, fencing of stolen goods, the employment of illegal aliens—could also be sources of a higher currency ratio.

Another interesting set of movements in c are the two increases during both world wars, which are associated with large increases in income taxes. Income taxes were raised substantially in 1917 to help finance America's entry into World War I. Although income tax rates were reduced after the war, they were again raised substantially during World War II to finance that conflict—never to return to prewar levels.

Increases in c when income tax rates rise can be explained in the following manner: Higher tax rates promote the evasion of taxes. When income tax rates rise, the incentive is high to evade taxes by conducting transactions in cash. If you receive an unreported cash payment for some service (for example, as a cab driver, waiter, or doctor), it is less likely that the Internal Revenue Service can prove that you are understating your income. If you are paid with a check or credit card, you would be wise to declare the income. The conclusion is clear: ***Higher tax rates will lead to a rise in c.***

Not only does income tax evasion explain the rise in c during the two world wars, but it also helps explain the rise in the 1960s and 1970s. This may seem surprising because the income tax rate schedule was not raised during this period. However, the burden of income taxes was increasing because the American income tax system is progressive (as income increases, the tax rate rises). A rising price level in the 1960s and 1970s raised the nominal income and pushed more individuals into higher tax brackets (a phenomenon called *bracket creep*). This meant that the effective tax rate increased even though the tax schedule was unchanged. As a result, incentives increased to evade paying taxes by not declaring income, and people would avoid the use of checkable deposits. In other words, the expected return on checkable deposits fell, so c rose.

Increased tax evasion and other illegal activities not only reflect an increase in the currency ratio but also imply that more income will go unreported to the government. The result is an understatement of statistics on economic activity such as gross domestic product (GDP), which measures the total production of goods and services in the economy.

---

[3] One exception to this is an increase in street crime. Checkable deposits have the advantage over currency that if you are mugged, the loss from carrying checks is likely to be far less than the loss from carrying currency. So if muggings are on the rise, the expected return on currency will fall relative to the expected return on checkable deposits, and you would hold less currency relative to checkable deposits. The resulting negative association of the illegal activity of street crime and c is ignored in the text because it is not an important source of fluctuations in c.

[4] The Drug Enforcement Agency has estimated that the retail value of the illegal drug trade exceeds $100 billion, making it one of the largest businesses in the United States. Evidence that the drug trade has affected c is found in Ralph C. Kimball, "Trends in the Use of Currency," *New England Economic Review*, September–October 1981, pp. 43–53.

This unreported economic activity has been labeled the *underground economy*. Evidence of its scope is the fact that the amount of currency for every man, woman, and child in the United States (as measured by currency in circulation in 2002 divided by the population) is around $1,000. Very few people hold this amount of currency; the likelihood is that much is used to conduct transactions in the underground economy. Calculations of the size of the underground economy indicate that it may exceed 10% of total economic activity. If this is true, and unreported income could be taxed, America would solve its budget deficit problems overnight!

Application

Explaining the Historical Record of *c*

The interaction of historical data with the theory of asset demand has helped us identify the factors that influence the currency ratio. We have seen that the theory of asset demand developed in Chapter 5 can help us understand how these different factors influence *c*.

To put our analysis in perspective, let us proceed to explain the major movements of *c* in Figure 1 by time periods.

Study Guide

An excellent way to test your understanding of the factors influencing *c* is to explain the movements in Figure 1 before reading this section of the text. This exercise will give you practice in using the ideas developed in the preceding discussion and should help make the abstract analysis clearer.

1892–1917. The general decline in *c* reflected in this period is explained by the increase in wealth. Because checkable deposits have a higher wealth elasticity than currency, the general trend of rising wealth over this span implies that the holdings of currency will grow more slowly than the holdings of checkable deposits, thus lowering *c*.

The upward blips in the ratio seen in 1893 and 1907 were due to bank panics, which temporarily reduced the expected return on checkable deposits and increased the risk—these factors led to a temporary increase in the holdings of currency relative to checkable deposits, temporarily increasing *c*.

1917–1919. The upward surge in *c* when America entered World War I is explained by the use of the income tax to help finance the war. The resulting attempts at tax evasion encouraged people to avoid the use of checks, which would make their income visible to the IRS; put another way, the increased desire to avoid taxes lowered the expected return on checkable deposits, resulting in a lower demand for them. The resulting increase in the use of currency relative to checkable deposits raised *c*.

1919–1921. When income taxes were reduced after the war, the demand for currency relative to checkable deposits began to fall back toward its old level, and the rise in *c* that occurred during the war was reversed. However, a severe recession in 1920–1921 led to a decline in wealth along with an increased number of bank failures, both of which might have caused a rise in *c* at that time. The decline in wealth led to a decline in the demand for both currency and checkable deposits, but the higher wealth elasticity of checkable deposits

meant that they declined more than currency, raising the currency ratio. The increased number of bank failures also made checkable deposits less desirable because it lowered their expected return, also leading to a rise in *c*.

1921–1929. During the prosperous period of the Roaring Twenties, we would expect to see the downward trend in *c* reasserting itself. The rise in wealth would lead to a fall in *c* because the holdings of currency would grow more slowly than the holdings of checkable deposits.

1929–1933. The decline in income during the Great Depression was one factor in the rise in *c*, but far more important were the bank panics that began in late 1930 and ended in March 1933. The consequent sharp rise in *c* from 1930 to 1933 was a major factor in the financial and economic collapse. These panics (the most severe in all of U.S. history) lowered the expected return on deposits, thereby raising the demand for currency relative to checkable deposits.

1933–1941. With the end of the bank panics and some restoration of the confidence in banks (helped by establishment of the FDIC), *c* fell. This decline was strengthened by a rise in wealth. However, *c* did not return to pre-Depression levels, primarily because a loss of confidence in the U.S. banking system lingered in the public mind. As a result, expected returns on deposits did not return to their pre-Depression levels, leaving a high level of *c*.

1941–1945. When income tax rates were raised to unprecedented levels to finance combat in World War II, *c* underwent a substantial rise. The incentive to evade taxes was especially strong; hence the expected return on checkable deposits fell. Price controls imposed during the war may also have contributed to the rise in *c* because they stimulated black market activity, whose transactions could be hidden using currency.

1945–Early 1960s. After the war, income tax rates were reduced slightly, but not to anywhere near their prewar levels. Income taxes remained at permanently higher levels because of the revenue needed to support an expanded role for the U.S. military as the "world's police force" and enlarged social programs such as welfare, unemployment insurance, housing and urban development, and Social Security. Although some decline in *c* occurred after the war due to a reduction in tax rates, permanently higher income tax rates left strong incentives for tax evasion, and *c* remained high. The steady rise in wealth after the war promoted the return to a declining trend in *c*, but its effect was not sufficiently strong to reduce the ratio below prewar levels.

Early 1960s–1980. The declining trend beginning at the end of World War II began to reverse in the early 1960s for a number of reasons. Most important was the growth of the underground economy, both because of the spectacular rise in illegal drug trade and because of the increased desirability of evading taxes due to bracket creep, which raised the effective tax rates. The increase in illegal activity lowered the expected return on checkable deposits, leading to an increased use of currency in relation to checkable deposits, thereby raising *c*.

1980–1993. A halt in the upward trend in *c* can be attributed to the deregulation of the banking system that allowed banks to pay interest on check-

able deposits. This raised the expected return on checkable deposits relative to currency, and the resulting reduced demand for currency helped lower c.

1994–2002. The upward trend in c can be explained by the explosion in the number of ATMs starting in the 1990s, which has been discussed in Chapter 10. The increase in the number of ATMs has dramatically lowered the cost of acquiring currency and this has, in effect, raised the expected return on currency relative to deposits, thereby raising c.

Application

Predicting the Future of *c*

A good economic model not only explains the past but also helps predict the response of economic variables to new events. The analysis of factors that influence the currency ratio outlined here has this capability. Let us consider two possible changes in the economic environment of the future and ask what our analysis would predict will happen to the currency ratio as a result. These predictions could be of interest to policymakers, who would want to know how the money supply might be affected in each of these cases.

Study Guide

Try to provide the reasoning for the predictions here without having to refer to the text. This will give you excellent practice with the economic analysis of c that we have developed in this chapter. You can get additional practice by answering problems at the end of the chapter, which also ask you to predict future movements in c.

Rise in Income Tax Rates to Balance the Budget

Much talk is circulating about balancing the budget by increasing taxes. What would happen to the currency ratio if income taxes were raised?

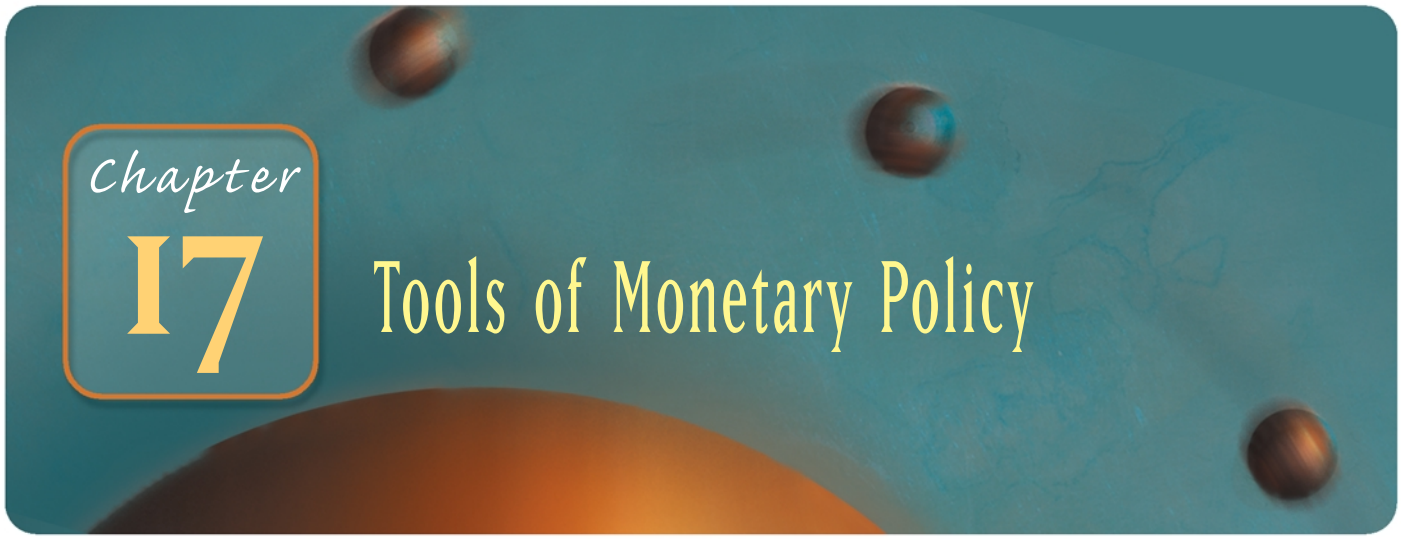
Higher tax rates would increase the incentives to evade taxes. The expected return on checkable deposits would then effectively decline. The use of currency would increase relative to checkable deposits (if other factors are held constant), and we would predict a rise in c.

Abolishment of Interest Payments on Checking Accounts

There have always been swings back and forth from deregulation to increased regulation. What if the present tide of deregulation is reversed and regulations were imposed that returned us to the situation when banks were not allowed to pay interest on checkable deposits? What would happen to the currency ratio in this case?

This policy would mean that the expected return on checkable deposits would fall below its current level, and the expected return on checkable deposits relative to currency would also fall. The resulting decreased attractiveness of checkable deposits relative to currency would mean that holdings of currency relative to checkable deposits would increase, raising c.

The usefulness of the foregoing analysis is not restricted to the predictions of the response of c to the events discussed here. With this framework, many other possible changes in our economic environment that would have an impact on c can be analyzed (a few are discussed in the problems at the end of the chapter).



Chapter 17

Tools of Monetary Policy

PREVIEW

In the chapters describing the structure of the Federal Reserve System and the money supply process, we mentioned three policy tools that the Fed can use to manipulate the money supply and interest rates: open market operations, which affect the quantity of reserves and the monetary base; changes in discount lending, which affect the monetary base; and changes in reserve requirements, which affect the money multiplier. Because the Fed's use of these policy tools has such an important impact on interest rates and economic activity, it is important to understand how the Fed wields them in practice and how relatively useful each tool is.

In recent years, the Federal Reserve has increased its focus on the **federal funds rate** (the interest rate on overnight loans of reserves from one bank to another) as the primary indicator of the stance of monetary policy. Since February 1994, the Fed announces a federal funds rate target at each FOMC meeting, an announcement that is watched closely by market participants because it affects interest rates throughout the economy. Thus, to fully understand how the Fed's tools are used in the conduct of monetary policy, we must understand not only their effect on the money supply, but their direct effects on the federal funds rate as well. The chapter thus begins with a supply-and-demand analysis of the market for reserves to explain how the Fed's settings for the three tools of monetary policy determine the federal funds rate. We then go on to look in more detail at each of the three tools—open market operations, discount rate policy, and reserve requirements—to see how they are used in practice and to ask whether the use of these tools could be modified to improve the conduct of monetary policy.

The Market for Reserves and the Federal Funds Rate

In Chapter 15, we saw how open market operations and discount lending affect the balance sheet of the Fed and the amount of reserves. The market for reserves is where the federal funds rate is determined, and this is why we turn to a supply-and-demand analysis of this market to analyze how all three tools of monetary policy affect the federal funds rate.

Supply and Demand in the Market for Reserves

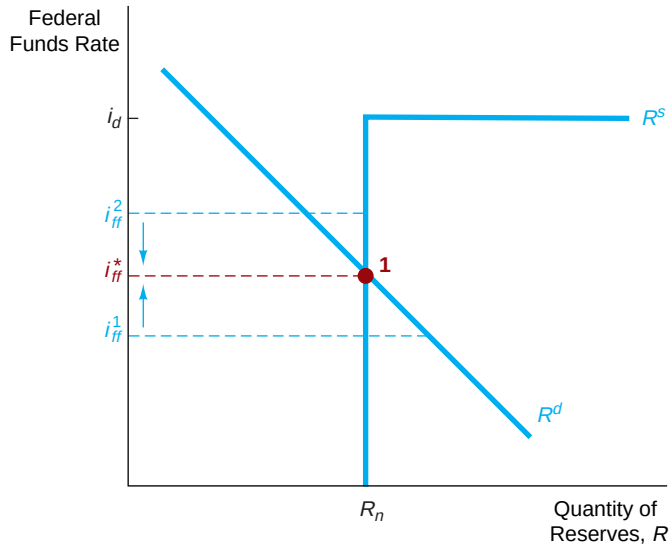
The analysis of the market for reserves proceeds in a similar fashion to the analysis of the bond market we conducted in Chapter 5. We derive a demand and supply curve for reserves. Then the market equilibrium in which the quantity of reserves demanded equals the quantity of reserves supplied determines the federal funds rate, the interest rate charged on the loans of these reserves.

Demand Curve. To derive the demand curve for reserves, we need to ask what happens to the quantity of reserves demanded, holding everything else constant, as the federal funds rate changes. Recall from Chapter 16 that the amount of reserves can be split up into two components: (1) required reserves, which equal the required reserve ratio times the amount of deposits on which reserves are required, and (2) excess reserves, the additional reserves banks choose to hold. Therefore, the quantity of reserves demanded equals required reserves plus the quantity of excess reserves demanded. Excess reserves are insurance against deposit outflows, and the cost of holding these excess reserves is their opportunity cost, the interest rate that could have been earned on lending these reserves out, which is equivalent to the federal funds rate. Thus as the federal funds rate decreases, the opportunity cost of holding excess reserves falls and, holding everything else constant, including the quantity of required reserves, the quantity of reserves demanded rises. Consequently, the demand curve for reserves, R^d, slopes downward in Figure 1.

Supply Curve. The supply of reserves, R^s, can be broken up into two components: the amount of reserves that are supplied by the Fed's open market operations, called nonborrowed reserves (R_n), and the amount of reserves borrowed from the Fed, called discount loans (DL). The primary cost of borrowing discount loans from the Fed is

FIGURE 1 Equilibrium in the Market for Reserves

Equilibrium occurs at the intersection of the supply curve R^s and the demand curve R^d at point 1 and an interest rate of i^*_{ff}.



the interest rate the Fed charges on these loans, the discount rate (i_d). Because borrowing federal funds is a substitute for taking out discount loans from the Fed, if the federal funds rate i_{ff} is below the discount rate i_d, then banks will not borrow from the Fed and discount loans will be zero because borrowing in the federal funds market is cheaper. Thus, as long as i_{ff} remains below i_d, the supply of reserves will just equal the amount of nonborrowed reserves supplied by the Fed, R_n, and so the supply curve will be vertical as shown in Figure 1. However, as the federal funds rate begins to rise above the discount rate, banks would want to keep borrowing more and more at i_d and then lending out the proceeds in the federal funds market at the higher rate, i_{ff}. The result is that the supply curve becomes flat (infinitely elastic) at i_d, as shown in Figure 1.

Market Equilibrium. Market equilibrium occurs where the quantity of reserves demanded equals the quantity supplied, $R^s = R^d$. Equilibrium therefore occurs at the intersection of the demand curve R^d and the supply curve R^s at point 1, with an equilibrium federal funds rate of i^*_{ff}. When the federal funds rate is above the equilibrium rate at i^2_{ff}, there are more reserves supplied than demanded (excess supply) and so the federal funds rate falls to i^*_{ff} as shown by the downward arrow. On the other hand, when the federal funds rate is below the equilibrium rate at i^1_{ff}, there are more reserves demanded than supplied (excess demand) and so the federal funds rate rises as shown by the upward arrow. (Note that Figure 1 is drawn so that i_d is above i^*_{ff} because the Federal Reserve now keeps the discount rate substantially above the target for the federal funds rate.)

www.federalreserve.gov/fomc/fundsrate.htm
This site lists historical federal funds rates and also discusses Federal Reserve targets.

How Changes in the Tools of Monetary Policy Affect the Federal Funds Rate

Now that we understand how the federal funds rate is determined, we can examine how changes in the three tools of monetary policy—open market operations, discount lending, and reserve requirements—affect the market for reserves and the equilibrium federal funds rate.

Open Market Operations. We have already seen that an open market purchase leads to a greater quantity of reserves supplied; this is true at any given federal funds rate because of the higher amount of nonborrowed reserves, which rises from R^1_n to R^2_n. An open market purchase therefore shifts the supply curve to the right from R^s_1 to R^s_2 and moves the equilibrium from point 1 to point 2, lowering the federal funds rate from i^1_{ff} to i^2_{ff} (see Figure 2).[1] The same reasoning implies that an open market sale decreases the quantity of reserves supplied, shifts the supply curve to the left and causes the federal funds rate to rise.

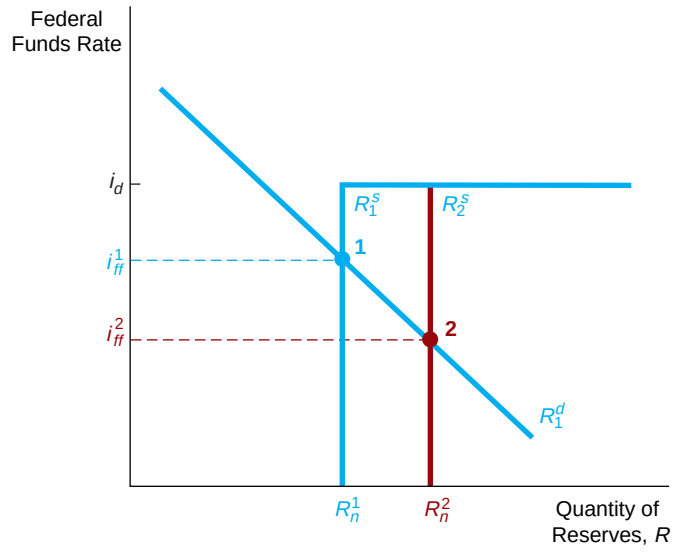
The result is that ***an open market purchase causes the federal funds rate to fall, whereas an open market sale causes the federal funds rate to rise***.

Discount Lending. The effect of a discount rate change depends on whether the demand curve intersects the supply curve in its vertical section versus its flat section. Panel a of Figure 3 shows what happens if the intersection occurs at the vertical section of the supply curve so there is no discount lending. In this case, when the discount rate

[1] We come to the same conclusion using the money supply framework in Chapter 16, along with the liquidity preference framework in Chapter 5. An open market purchase raises reserves and the money supply, and then the liquidity preference framework shows that interest rates fall.

FIGURE 2 Response to an Open Market Operation

An open market purchase increases nonborrowed reserves and hence the reserves supplied, and shifts the supply curve from R_1^s to R_2^s. The equilibrium moves from point 1 to point 2, lowering the federal funds rate from i_{ff}^1 to i_{ff}^2.



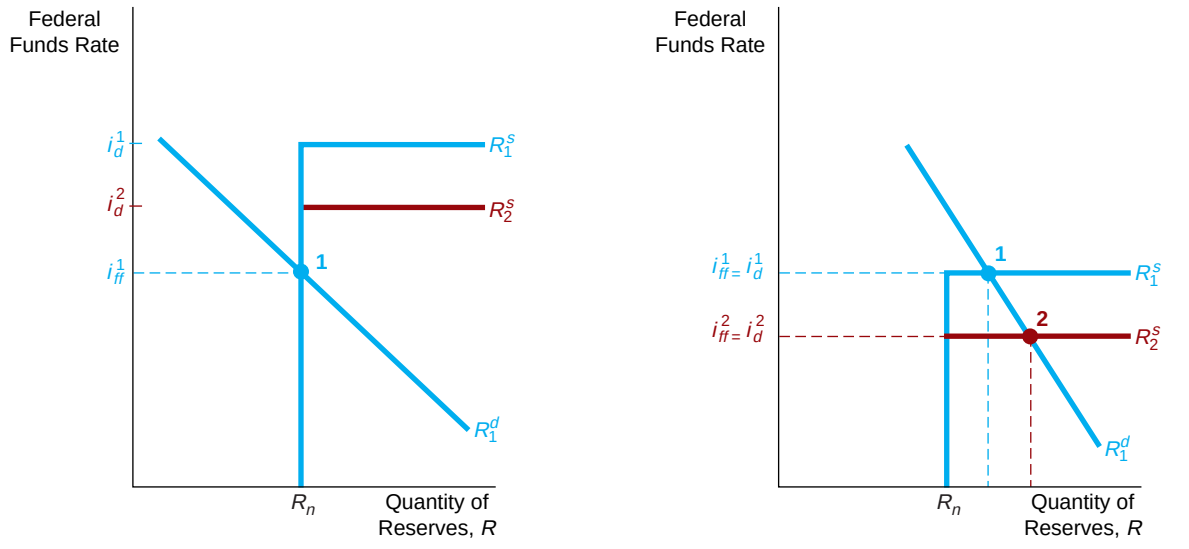


FIGURE 3 Response to a Change in the Discount Rate

In panel a when the discount rate is lowered by the Fed from i_d^1 to i_d^2, the vertical section of the supply curve just shortens, as in R_2^s, so that the equilibrium federal funds rate remains unchanged at i_{ff}^1. In panel b when the discount rate is lowered by the Fed from i_d^1 to i_d^2, the horizontal section of the supply curve R_2^s falls, and the equilibrium federal funds rate falls from i_{ff}^1 to i_{ff}^2.

www.frbdiscountwindow.org/
Information on the operation of the discount window and data on current and historical interest rates.

is lowered by the Fed from i_d^1 to i_d^2, the vertical section of the supply curve where there is no discount lending just shortens, as in R_2^s, while the intersection of the supply and demand curve remains at the same point. Thus, in this case there is no change in the equilibrium federal funds rate, which remains at i_{ff}^1. Because this is the typical situation—since the Fed now usually keeps the discount rate above its target for the federal funds rate—the conclusion is that ***most changes in the discount rate have no effect on the federal funds rate***.

However, if the demand curve intersects the supply curve on its flat section, so there is some discount lending, as in panel b of Figure 3, changes in the discount rate do affect the federal funds rate. In this case, initially discount lending is positive and the equilibrium federal funds rate equals the discount rate, $i_{ff}^1 = i_d^1$. When the discount rate is lowered by the Fed from i_d^1 to i_d^2, the horizontal section of the supply curve R_2^s falls, moving the equilibrium from point 1 to point 2, and the equilibrium federal funds rate falls from i_{ff}^1 to i_{ff}^2 $(= i_d^2)$ in panel b.

www.federalreserve.gov/monetarypolicy/reservereq.htm
Historical data and discussion about reserve requirements.

Reserve Requirements. When the required reserve ratio increases, required reserves increase and hence the quantity of reserves demanded increases for any given interest rate. Thus a rise in the required reserve ratio shifts the demand curve to the right from R_1^d to R_2^d in Figure 4, moves the equilibrium from point 1 to point 2, and in turn raises the federal funds rate from i_{ff}^1 to i_{ff}^2.

The result is that ***when the Fed raises reserve requirements, the federal funds rate rises***.[2]

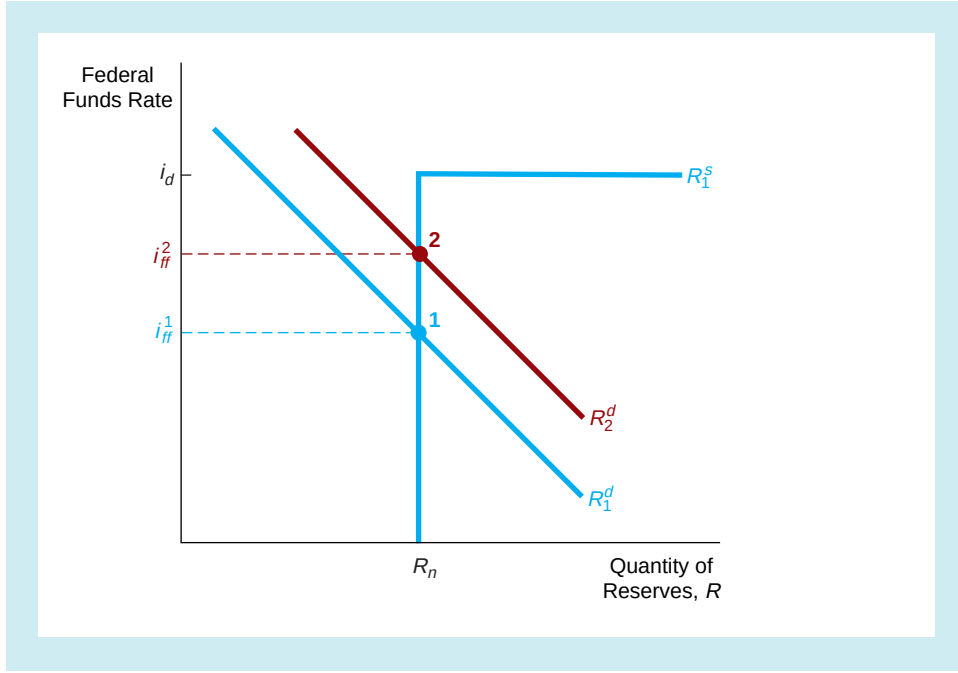


FIGURE 4 Response to a Change in Required Reserves
When the Fed raises reserve requirements, required reserves increase, which increases the demand for reserves. The demand curve shifts from R_1^d to R_2^d, the equilibrium moves from point 1 to point 2, and the federal fund rate rises from i_{ff}^1 to i_{ff}^2.

[2]Because an increase in the required reserve ratio means that the same amount of reserves is able to support a smaller amount of deposits, a rise in the required reserve ratio leads to a decline in the money supply. Using the liquidity preference framework, the fall in the money supply results in a rise in interest rates, yielding the same conclusion in the text that raising reserve requirements leads to higher interest rates.

Similarly, a decline in the required reserve ratio lowers the quantity of reserves demanded, shifts the demand curve to the left, and causes the federal funds rate to fall. ***When the Fed decreases reserve requirements, it leads to a fall in the federal funds rate.***

Now that we understand how the three tools of monetary policy—open market operations, discount lending, and reserve requirements—can be used by the Fed to manipulate the money supply and interest rates, we will look at each of them in turn to see how the Fed wields them in practice and how relatively useful each tool is.

Open Market Operations

Open market operations are the most important monetary policy tool, because they are the primary determinants of changes in interest rates and the monetary base, the main source of fluctuations in the money supply. Open market purchases expand reserves and the monetary base, thereby raising the money supply and lowering short-term interest rates. Open market sales shrink reserves and the monetary base, lowering the money supply and raising short-term interest rates. Now that we understand from Chapter 15 the factors that influence the reserves and monetary base, we can examine how the Federal Reserve conducts open market operations with the object of controlling short-term interest rates and the money supply.

There are two types of open market operations: **Dynamic open market operations** are intended to change the level of reserves and the monetary base, and **defensive open market operations** are intended to offset movements in other factors that affect reserves and the monetary base, such as changes in Treasury deposits with the Fed or float. The Fed conducts open market operations in U.S. Treasury and government agency securities, especially U.S. Treasury bills.[3] The Fed conducts most of its open market operations in Treasury securities because the market for these securities is the most liquid and has the largest trading volume. It has the capacity to absorb the Fed's substantial volume of transactions without experiencing excessive price fluctuations that would disrupt the market.

www.federalreserve.gov/fomc
A discussion about the federal open market committee, list of current members, meeting dates, and other current information.

As we saw in Chapter 14, the decision-making authority for open market operations is the Federal Open Market Committee (FOMC), which sets a target for the federal funds rate. The actual execution of these operations, however, is conducted by the trading desk at the Federal Reserve Bank of New York. The best way to see how these transactions are executed is to look at a typical day at the trading desk, located in a newly built trading room on the ninth floor of the Federal Reserve Bank of New York.

A Day at the Trading Desk

The manager of domestic open market operations supervises the analysts and traders who execute the purchases and sales of securities in order to hit the federal funds rate target. To get a grip on what might happen in the federal funds market that day, her workday and that of her staff begins with a review of developments in the federal funds market the previous day and with an update on the actual amount of reserves

[3] To avoid conflicts of interest, the Fed does not conduct open market operations in privately issued securities. (For example, think of the conflict if the Federal Reserve purchased bonds issued by a company owned by the chairman's brother-in-law.)

in the banking system the day before. Later in the morning, her staff issues updated reports that contain detailed forecasts of what will be happening to some of the short-term factors affecting the supply and demand of reserves (discussed in Chapter 15). For example, if float is predicted to decrease because good weather throughout the country is speeding up check delivery, the manager of domestic open market operations knows that she will have to conduct a defensive open market operation (in this case, a *purchase* of securities) to offset the expected decline in reserves and the monetary base from the decreased float. However, if Treasury deposits with the Fed are predicted to fall, a defensive open market *sale* would be needed to offset the expected increase in reserves. The report also predicts the change in the public's holding of currency. If currency holdings are expected to rise, then, as we have seen in Chapters 15 and 16, reserves fall, and an open market purchase is needed to raise reserves back up again.

This information will help the manager of domestic open market operations and her staff decide how large a change in reserves is needed to obtain the federal funds rate target. If the amount of reserves in the banking system is too large, many banks will have excess reserves to lend that other banks may have little desire to hold, and the federal funds rate will fall. If the level of reserves is too low, banks seeking to borrow reserves from the few banks that have excess reserves to lend may push the funds rate higher than the desired level. Also during the morning, the staff will monitor the behavior of the federal funds rate and contact some of the major participants in the funds market, which may provide independent information about whether a change in reserves is needed to achieve the desired level of the federal funds rate. Early in the morning, members of the manager's staff contact several representatives of the so-called **primary dealers**, government securities dealers (who operate out of private firms or commercial banks) that the open market desk trades with. Her staff finds out how the dealers view market conditions to get a feel for what may happen to the prices of the securities they trade in over the course of the day. They also call the Treasury to get updated information on the expected level of Treasury balances at the Fed in order to refine their estimates of the supply of reserves.

Afterward, members of the Monetary Affairs Division at the Board of Governors are contacted, and the New York Fed's forecasts of reserve supply and demand are compared with the Board's. On the basis of these projections and the observed behavior of the federal funds market, the desk will formulate and propose a course of action to be taken that day, which may involve plans to add reserves to or drain reserves from the banking system through open market operations. If an operation is contemplated, the type, size, and maturity will be discussed.

The whole process is currently completed by midmorning, at which time a daily conference call is arranged linking the desk with the Office of the Director of Monetary Affairs at the Board and with one of the four voting Reserve Bank presidents outside of New York. During the call, a member of the open market operations unit will outline the desk's proposed reserve management strategy for the day. After the plan is approved, the desk is instructed to execute immediately any temporary open market operations that were planned for that day. (Outright operations, to be described shortly, may be conducted at other times of the day.)

The desk is linked electronically with its domestic open market trading counterparties by a computer system called TRAPS (Trading Room Automated Processing System), and all open market operations are now performed over this system. A message will be electronically transmitted simultaneously to all the primary dealers over TRAPS indicating the type and maturity of the operation being arranged. The dealers

are given several minutes to respond via TRAPS with their propositions to buy or sell government securities. The propositions are then assembled and displayed on a computer screen for evaluation. The desk will select all propositions, beginning with the most attractively priced, up to the point where the desired amount is purchased or sold, and it will then notify each dealer via TRAPS which of its propositions have been chosen. The entire selection process is typically completed in a matter of minutes.

These temporary transactions are of two basic types. In a **repurchase agreement** (often called a **repo**), the Fed purchases securities with an agreement that the seller will repurchase them in a short period of time, anywhere from 1 to 15 days from the original date of purchase. Because the effects on reserves of a repo are reversed on the day the agreement matures, a repo is actually a temporary open market purchase and is an especially desirable way of conducting a defensive open market purchase that will be reversed shortly. When the Fed wants to conduct a temporary open market sale, it engages in a **matched sale–purchase transaction** (sometimes called a **reverse repo**) in which the Fed sells securities and the buyer agrees to sell them back to the Fed in the near future.

At times, the desk may see the need to address a persistent reserve shortage or surplus and wish to arrange an operation that will have a permanent impact on the supply of reserves. Outright transactions, which involve a purchase or sale of securities that is not self-reversing, are also conducted over TRAPS. These operations are traditionally executed at times of day when temporary operations are not being conducted.

Advantages of Open Market Operations

Open market operations have several advantages over the other tools of monetary policy.

1. Open market operations occur at the initiative of the Fed, which has complete control over their volume. This control is not found, for example, in discount operations, in which the Fed can encourage or discourage banks to take out discount loans by altering the discount rate but cannot directly control the volume of discount loans.

2. Open market operations are flexible and precise; they can be used to any extent. No matter how small a change in reserves or the monetary base is desired, open market operations can achieve it with a small purchase or sale of securities. Conversely, if the desired change in reserves or the base is very large, the open market operations tool is strong enough to do the job through a very large purchase or sale of securities.

3. Open market operations are easily reversed. If a mistake is made in conducting an open market operation, the Fed can immediately reverse it. If the Fed decides that the federal funds rate is too low because it has made too many open market purchases, it can immediately make a correction by conducting open market sales.

4. Open market operations can be implemented quickly; they involve no administrative delays. When the Fed decides that it wants to change the monetary base or reserves, it just places orders with securities dealers, and the trades are executed immediately.

Discount Policy

The Federal Reserve facility at which discount loans are made to banks is called the **discount window**. The easiest way to understand how the Fed affects the volume of discount loans is by looking at how the discount window operates.

Operation of the Discount Window

The Fed's discount loans to banks are of three types: primary credit, secondary credit, and seasonal credit.[4] *Primary credit* is the discount lending that plays the most important role in monetary policy. Healthy banks are allowed to borrow all they want from the primary credit facility, and it is therefore referred to as a *standing lending facility*. The interest rate on these loans is the discount rate, and as we mentioned before, it is set higher than the federal funds rate target, usually by 100 basis points (one percentage point), and thus in most circumstances the amount of discount lending under the primary credit facility is very small. Then why does the Fed have this facility?

The answer is that the facility is intended to be a backup source of liquidity for sound banks so that the federal funds rate never rises too far above the federal funds target. To see how the primary credit facility works, let's see what happens if there is a large increase in the demand for reserves, say because deposits have surged unexpectedly and have led to an increase in required reserves. This situation is analyzed in Figure 5. Suppose that initially, the demand and supply curve for reserves intersect at point 1 so that the federal funds rate is at its target level, i_{ff}^{T}. Now the increase in required reserves shifts the demand curve to R_2^d and the equilibrium moves to point 2. The result is that discount lending increases from zero to DL_2 and the federal funds rate rises to i_d and can rise no further. The primary credit facility has thus put a ceiling on the federal funds rate of i_d.

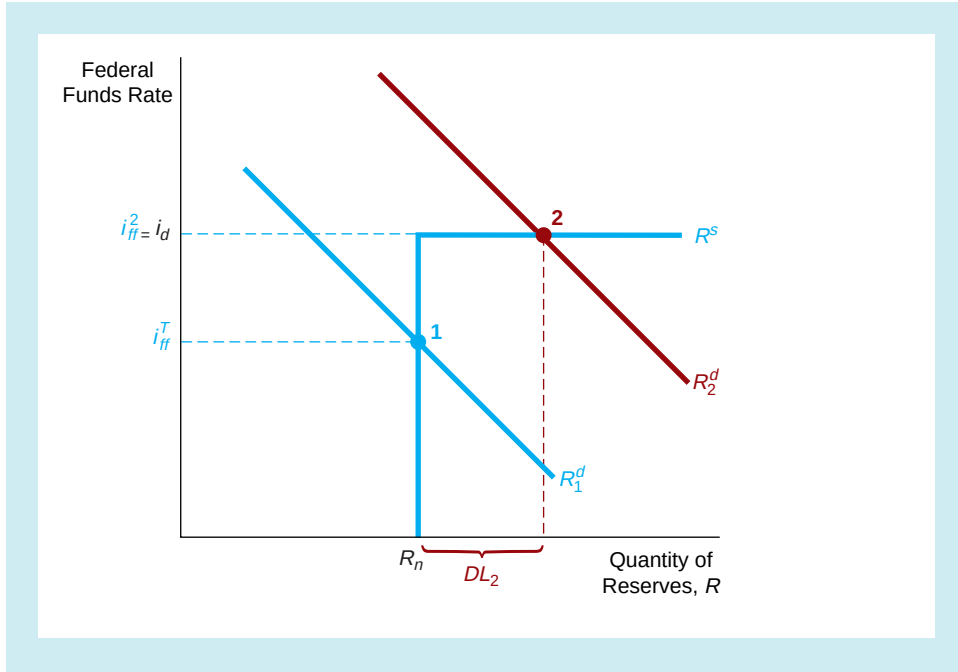


FIGURE 5 How the Primary Credit Facility Puts a Ceiling on the Federal Funds Rate

The rightward shift of the demand curve to reserves from R_1^d to R_2^d moves the equilibrium from point 1 to point 2 where $i_{ff}^{2} = i_d$ and discount lending rises from zero to DL_2.

[4] The procedures for administering the discount window were changed in January 2003. The primary credit facility replaced an adjustment credit facility whose discount rate was typically set below market interest rates, so banks were restricted in their access to this credit. In contrast, now healthy banks can borrow all they want from the primary credit facility. The secondary credit facility replaced the extended credit facility which focused somewhat more on longer-term credit extensions. The seasonal credit facility remains basically unchanged.

Secondary credit is given to banks that are in financial trouble and are experiencing severe liquidity problems. The interest rate on secondary credit is set at 50 basis points (0.5 percentage points) above the discount rate. This interest rate on these loans is set at a higher, penalty rate to reflect the less-sound condition of these borrowers. *Seasonal credit* is given to meet the needs of a limited number of small banks in vacation and agricultural areas that have a seasonal pattern of deposits. The interest rate charged on seasonal credit is tied to the average of the federal funds rate and certificate of deposit rates. The Federal Reserve has questioned the need for the seasonal credit facility because of improvements in credit markets and is thus contemplating eliminating it in the future.

Lender of Last Resort

In addition to its use as a tool to influence reserves, the monetary base, and the money supply, discounting is important in preventing financial panics. When the Federal Reserve System was created, its most important role was intended to be as the **lender of last resort**; to prevent bank failures from spinning out of control, it was to provide reserves to banks when no one else would, thereby preventing bank and financial panics. Discounting is a particularly effective way to provide reserves to the banking system during a banking crisis because reserves are immediately channeled to the banks that need them most.

Using the discount tool to avoid financial panics by performing the role of lender of last resort is an extremely important requirement of successful monetary policymaking. As we demonstrated with our money supply analysis in Chapter 16, the bank panics in the 1930–1933 period were the cause of the sharpest decline in the money supply in U.S. history, which many economists see as the driving force behind the collapse of the economy during the Great Depression. Financial panics can also severely damage the economy because they interfere with the ability of financial intermediaries and markets to move funds to people with productive investment opportunities (see Chapter 8).

Unfortunately, the discount tool has not always been used by the Fed to prevent financial panics, as the massive failures during the Great Depression attest. The Fed learned from its mistakes of that period and has performed admirably in its role of lender of last resort in the post–World War II period. The Fed has used its discount lending weapon several times to avoid bank panics by extending loans to troubled banking institutions, thereby preventing further bank failures. The largest of these occurred in 1984, when the Fed lent Continental Illinois, at that time one of the ten largest banks in the United States, more than $5 billion.

At first glance, it might seem that the presence of the FDIC, which insures depositors up to a limit of $100,000 per account from losses due to a bank's failure, would make the lender-of-last-resort function of the Fed superfluous. (The FDIC is described in detail in Chapter 11.) There are two reasons why this is not the case. First, it is important to recognize that the FDIC's insurance fund amounts to around 1% of the amount of these deposits outstanding. If a large number of bank failures occurred, the FDIC would not be able to cover all the depositors' losses. Indeed, the large number of bank failures in the 1980s and early 1990s, described in Chapter 11, led to large losses and a shrinkage in the FDIC's insurance fund, which reduced the FDIC's ability to cover depositors' losses. This fact has not weakened the confidence of small depositors in the banking system because the Fed has been ready to stand behind the banks to provide whatever reserves are needed to prevent bank panics. Second, the nearly $1 trillion of large-denomination deposits in the banking system are not

guaranteed by the FDIC, because they exceed the $100,000 limit. A loss of confidence in the banking system could still lead to runs on banks from the large-denomination depositors, and bank panics could still occur despite the existence of the FDIC. The importance of the Federal Reserve's role as lender of last resort is, if anything, more important today because of the high number of bank failures experienced in the 1980s and early 1990s.

Not only can the Fed be a lender of last resort to banks, but it can also play the same role for the financial system as a whole. The existence of the Fed's discount window can help prevent financial panics that are not triggered by bank failures, as was the case during the Black Monday stock market crash of 1987 and the terrorist destruction of the World Trade Center in September 2001 (see Box 1).

Although the Fed's role as the lender of last resort has the benefit of preventing bank and financial panics, it does have a cost. If a bank expects that the Fed will provide it with discount loans when it gets into trouble, as occurred with Continental Illinois, it will be willing to take on more risk knowing that the Fed will come to the rescue. The Fed's lender-of-last-resort role has thus created a moral hazard problem similar to the one created by deposit insurance (discussed in Chapter 11): Banks take on more risk, thus exposing the deposit insurance agency, and hence taxpayers, to greater losses. The moral hazard problem is most severe for large banks, which may believe that the Fed and the FDIC view them as "too big to fail"; that is, they will always receive Fed loans when they are in trouble because their failure would be likely to precipitate a bank panic.

Similarly, Federal Reserve actions to prevent financial panic, as occurred after the October 1987 stock market crash and the September 11, 2001 terrorist attacks, may encourage financial institutions other than banks to take on greater risk. They, too, expect the Fed to ensure that they could get loans if a financial panic seemed imminent. When the Fed considers using the discount weapon to prevent panics, it therefore needs to consider the trade-off between the moral hazard cost of its role as lender of last resort and the benefit of preventing financial panics. This trade-off explains why the Fed must be careful not to perform its role as lender of last resort too frequently.

Advantages and Disadvantages of Discount Policy

The most important advantage of discount policy is that the Fed can use it to perform its role of lender of last resort. Experiences with Continental Illinois, the Black Monday crash, and September 11, 2001 indicate that this role has become more important in the past couple of decades. In the past, discount policy was used as a tool of monetary policy, with the discount rate changed in order to affect interest rates and the monetary market. However, because the decisions to take out discount loans are made by banks and are therefore not completely controlled by the Fed, while open market operations are completely controlled by the Fed, the use of discount policy to conduct monetary policy has little to recommend it. This is why the Fed moved in January 2003 to the current system in which the discount facility is not used to set the federal funds rate, but is only a backup facility to prevent the federal funds rate from rising too far above its target.

Reserve Requirements

As we saw in Chapter 16, changes in reserve requirements affect the money supply by causing the money supply multiplier to change. A rise in reserve requirements reduces

Box 1: Inside the Fed

Discounting to Prevent a Financial Panic

The Black Monday Stock Market Crash of 1987 and the Terrorist Destruction of the World Trade Center in September 2001. Although October 19, 1987, dubbed "Black Monday," will go down in the history books as the largest one-day percentage decline in stock prices to date (the Dow Jones Industrial Average declined by more than 20%), it was on Tuesday, October 20, 1987, that financial markets almost stopped functioning. Felix Rohatyn, one of the most prominent men on Wall Street, stated flatly: "Tuesday was the most dangerous day we had in 50 years."* Much of the credit for prevention of a market meltdown after Black Monday must be given to the Federal Reserve System and the chairman of the Board of Governors, Alan Greenspan.

The stress of keeping markets functioning during the sharp decline in stock prices on Monday, October 19, meant that many brokerage houses and specialists (dealer-brokers who maintain orderly trading on the stock exchanges) were severely in need of additional funds to finance their activities. However, understandably enough, New York banks, as well as foreign and regional U.S. banks, growing very nervous about the financial health of securities firms, began to cut back credit to the securities industry at the very time when it was most needed. Panic was in the air. One chairman of a large specialist firm commented that on Monday, "from 2 P.M. on, there was total despair. The entire investment community fled the market. We were left alone on the field." It was time for the Fed, like the cavalry, to come to the rescue.

Upon learning of the plight of the securities industry, Alan Greenspan and E. Gerald Corrigan, then president of the Federal Reserve Bank of New York and the Fed official most closely in touch with Wall Street, became fearful of a spreading collapse of securities firms. To prevent this from occurring, Greenspan announced before the market opened on Tuesday, October 20, the Federal Reserve System's "readiness to serve as a source of liquidity to support the economic and financial system." In addition to this extraordinary announcement, the Fed made it clear that it would provide discount loans to any bank that would make loans to the securities industry, although this did not prove to be necessary. As one New York banker said, the Fed's message was, "We're here. Whatever you need, we'll give you."

The outcome of the Fed's timely action was that a financial panic was averted. The markets kept functioning on Tuesday, and a market rally ensued that day, with the Dow Jones Industrial Average climbing over 100 points.

A similar lender-of-last-resort operation was carried out in the aftermath of the destruction of the World Trade Center in New York City on Tuesday, September 11, 2001—the worst terrorist incident in U.S. history. Because of the disruption to the most important financial center in the world, the liquidity needs of the financial system skyrocketed. To satisfy these needs and to keep the financial system from seizing up, within a few hours of the incident, the Fed made an announcement similar to that made after the crash of 1987: "The Federal Reserve System is open and operating. The discount window is available to meet liquidity needs."** The Fed thenproceeded to provide $45 billion to banks through the discount window, a 200-fold increase over the previous week. As a result of this action, along with the injection of as much as $80 billion of reserves into the banking system through open market operations, the financial system kept functioning. When the stock market reopened on Monday, September 17, trading was orderly, although the Dow Jones average did decline 7%.

The terrorists were able to bring down the twin towers of the World Trade Center, with nearly 3,000 dead. However, they were unable to bring down the U.S. financial system because of the timely actions of the Federal Reserve.

*"Terrible Tuesday: How the Stock Market Almost Disintegrated a Day After the Crash," *Wall Street Journal*, November 20, 1987, p. 1. This article provides a fascinating and more detailed view of the events described here and is the source of all the quotations cited.

**"Economic Front: How Policy Makers Regrouped to Defend the Financial System," *Wall Street Journal*, Tuesday, September 18, 2001, p. A1, provides more detail on this episode.

the amount of deposits that can be supported by a given level of the monetary base and will lead to a contraction of the money supply. A rise in reserve requirements also increases the demand for reserves and raises the federal funds rate. Conversely, a decline in reserve requirements leads to an expansion of the money supply and a fall in the federal funds rate. The Fed has had the authority to vary reserve requirements since the 1930s; this is a powerful way of affecting the money supply and interest rates. Indeed, changes in reserve requirements have such large effects on the money supply and interest rates that the Fed rarely resorts to using this tool to control them.

The Depository Institutions Deregulation and Monetary Control Act of 1980 provided a simpler scheme for setting reserve requirements. All depository institutions, including commercial banks, savings and loan associations, mutual savings banks, and credit unions, are subject to the same reserve requirements, as follows: Required reserves on all checkable deposits—including non-interest-bearing checking accounts, NOW accounts, super-NOW accounts, and ATS (automatic transfer savings) accounts—are equal to 3% of the bank's first $42.1 million of checkable deposits[5] and 10% of the checkable deposits over $42.1 million, and the percentage set initially at 10% can be varied between 8 and 14%, at the Fed's discretion. In extraordinary circumstances, the percentage can be raised as high as 18%.

Advantages and Disadvantages of Reserve Requirement Changes

The main advantage of using reserve requirements to control the money supply and interest rates is that they affect all banks equally and have a powerful effect on the money supply. The fact that changing reserve requirements is a powerful tool, however, is probably more of a curse than a blessing, because small changes in the money supply and interest rates are hard to engineer by varying reserve requirements. With checkable deposits currently around the $600 billion level, a $\frac{1}{2}$-percentage-point increase in the reserve requirement on these deposits would reduce excess reserves by $30 billion. Because this decline in excess reserves would result in multiple deposit contraction, the decline in the money supply would be even greater. It is true that small changes in the money supply could be obtained by extremely small changes in reserve requirements (say, by 0.001 percentage point), but because it is so expensive to administer changes in reserve requirements, such a strategy is not practical. Using reserve requirements to fine-tune the money supply is like trying to use a jackhammer to cut a diamond.

Another disadvantage of using reserve requirements to control the money supply and interest rates is that raising the requirements can cause immediate liquidity problems for banks with low excess reserves. When the Fed has raised these requirements in the past, it has usually softened the blow by conducting open market purchases or by making the discount window more available, thus providing reserves to banks that needed them. Continually fluctuating reserve requirements would also create more uncertainty for banks and make their liquidity management more difficult.

The policy tool of changing reserve requirements does not have much to recommend it, and it is rarely used.

[5]The $42.1 million figure is as of the end of 2002. Each year, the figure is adjusted upward by 80% of the percentage increase in checkable deposits in the United States.

Application

Why Have Reserve Requirements Been Declining Worldwide?

In recent years, central banks in many countries in the world have been reducing or eliminating their reserve requirements. In the United States, the Federal Reserve eliminated reserve requirements on time deposits in December 1990 and lowered reserve requirements on checkable deposits from 12% to 10% in April 1992. As a result, the majority of U.S. depository institutions—but not the largest ones with the bulk of deposits—find that reserve requirements are not binding: In order to service their depositors, many depository institutions need to keep sufficient vault cash on hand (which counts toward meeting reserve requirements) that they more than meet reserve requirements voluntarily. Canada has gone a step further: Financial market legislation taking effect in June 1992 eliminated all reserve requirements over a two-year period. The central banks of Switzerland, New Zealand, and Australia have also eliminated reserve requirements entirely. What explains the downward trend for reserve requirements in most countries?

You may recall from Chapter 9 that reserve requirements act as a tax on banks. Because central banks typically do not pay interest on reserves, the bank earns nothing on them and loses the interest that could have been earned if the bank held loans instead. The cost imposed on banks from reserve requirements means that banks, in effect, have a higher cost of funds than intermediaries not subject to reserve requirements, making them less competitive. We have already seen in Chapter 10 that additional market forces have been making banks less competitive, weakening the health of banking systems throughout the world. Central banks have thus been reducing reserve requirements to make banks more competitive and stronger.[6] The Federal Reserve was explicit about this rationale for its April 1992 reduction when it announced it on February 18, 1992, stating in its press release that the reduction "will reduce funding costs for depositories and strengthen their balance sheets. Over time, it is expected that most of these cost savings will be passed on to depositors and borrowers."

Application

The Channel/Corridor System for Setting Interest Rates in Other Countries

The fall in reserve requirements has elicited the concern that if the demand for reserves falls to zero, then a central bank may not be able to exercise control over interest rates.[7] However, the so-called channel or corridor system for conducting monetary policy—which has been adopted by Canada,

[6]Many economists believe that the Fed should pay market interest rates on reserves, another suggestion for dealing with this problem.

[7]See Benjamin Friedman, "The Future of Monetary Policy: The Central Bank as an Army with Only a Signal Corps?" *International Finance* 2 (1999), pp. 321–338, and the rest of the symposium on this topic in the same journal.

Australia, and New Zealand, all of which have eliminated reserve requirements—shows that central banks can continue to effectively set overnight, interbank interest rates like the federal funds rate. How the channel/corridor system works is illustrated by Figure 6, which describes the market for reserves along the lines discussed at the beginning of this chapter.

In the channel/corridor system, the central bank sets up a standing lending facility, like the one currently in place in the United States and in most industrialized countries, in which the central bank stands ready to lend overnight any amount banks ask for at a fixed interest rate, i^l. This standing lending facility is commonly called a *lombard facility* and the interest rate charged on these loans is often called a *lombard rate*. (This name comes from Lombardy, a region in northern Italy that was an important center of banking in the middle ages.) As we saw at the beginning of the chapter, with a standing lending facility, the central bank does not limit the amount of borrowing by banks, but always stands ready to supply any amount the banks want at the lending rate i^l. Thus the quantity of reserves supplied is flat (infinitely elastic) at i^l as shown in Figure 6, because if the overnight interest rate, denoted by i_{ff}, begins to rise above i^l, banks would just keep borrowing discount loans indefinitely.

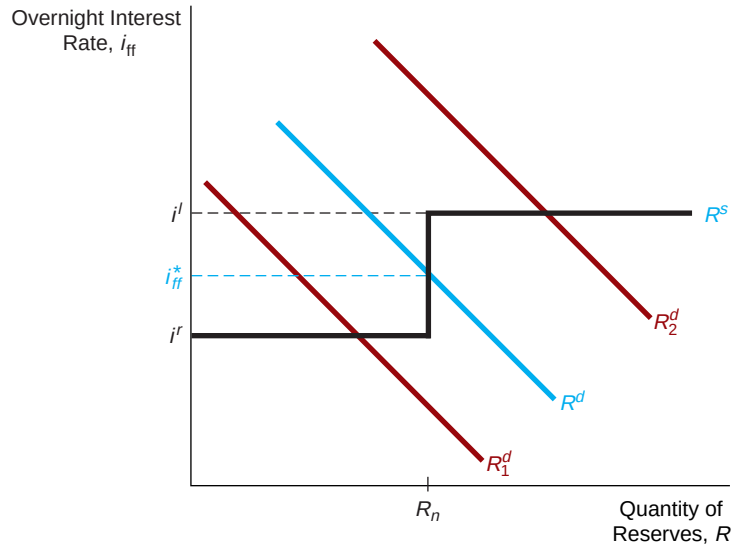
In the channel/corridor system the central bank sets up another standing facility that pays banks a fixed interest rate i^r on any reserves (deposits) they would like to keep at the central bank. The quantity of reserves supplied is also flat at i^r, because if the overnight rate begins to fall below this rate, banks would not lend in the overnight market. Instead they would keep increasing the amount of their deposits in the central bank (effectively lending to the central bank), and would thereby keep lowering the quantity of reserves the central bank is supplying. In between i^r and i^l, the quantity of reserves supplied equals nonborrowed reserves R_n, which are determined by open market operations. Nonborrowed reserves are set to zero if the demand for reserves is also expected to be zero. The supply curve for reserves R^s is thus the step function depicted in Figure 6.

The demand curve for reserves R^d has the usual downward slope. As we can see in Figure 6, when the demand curve shifts to the left to R_1^d the overnight interest rate never falls below i^r, while if the demand curve shifts to the right to R_2^d, the overnight rate never rises above i^l. Thus the channel/corridor system enables the central bank to keep the overnight interest rate in between the narrow channel/corridor with an upper limit of i^l and lower limit of i^r. In Canada, Australia, and New Zealand the lending rate i^l is set 25 basis points (0.25 percentage points) above the announced target rate, while the interest rate paid on reserves kept at the central bank is set at 25 basis points below the target. More in-depth analysis shows that banks will set the demand for reserves so that the demand curve is expected to intersect the supply curve at the announced target overnight rate of i_{ff}^*, with the result that deviations from the announced target are fairly small.[8]

[8]See Michael Woodford, "Monetary Policy in the Information Economy," in *Symposium on Economic Policy for the Information Economy* (Federal Reserve Bank of Kansas City: 2001), pp. 297–370.

FIGURE 6 The Channel/Corridor System for Setting Interest Rates

In the channel/corridor system standing facilities result in a step function supply curve, R^s. Then if the demand curve shifts between R_1^d and R_2^d, the overnight interest rate i_{ff} always remains between i^r and i^l.



The important point of this analysis is that the channel/corridor approach enables the central bank to set the overnight policy rate, whatever the demand for reserves, including zero demand. Thus in the future, continuing declines in the demand for reserves may eventually lead central banks to follow in the footsteps of the central banks of Canada, Australia, and New Zealand, and to adopt the channel/corridor system for conducting monetary policy.

Summary

1. A supply and demand analysis of the market for reserves yields the following results. When the Fed makes an open market purchase or lowers reserve requirements, the federal funds rate declines. When the Fed makes an open market sale or raises reserve requirements, the federal funds rate rises. Changes in the discount rate may also affect the federal funds rate.
2. The amount of an open market operation conducted on any given day by the trading desk of the Federal Reserve Bank of New York is determined by the amount of the dynamic open market operation intended to change reserves and the monetary base and by the amount of the defensive open market operation used to offset other factors that affect reserves and the monetary base. Open market operations are the primary tool used by the Fed to control the money supply because they occur at the initiative of the Fed, are flexible, are easily reversed, and can be implemented quickly.
3. The volume of discount loans is affected by the discount rate. Besides its effect on the monetary base and the money supply, discounting allows the Fed to perform its role as the lender of last resort. However, because the decisions by banks to take out discount loans are not controlled by the Fed, the use of discount policy to conduct monetary policy has little to recommend it.
4. Changing reserve requirements is too blunt a tool to use for controlling the money supply, and hence it is rarely used.

Key Terms

defensive open market operations, p. 398
discount window, p. 400
dynamic open market operations, p. 398
federal funds rate, p. 393
lender of last resort, p. 402
matched sale–purchase transaction (reverse repo), p. 400
primary dealers, p. 399
repurchase agreement (repo), p. 400

Questions and Problems

Questions marked with an asterisk are answered at the end of the book in an appendix, "Answers to Selected Questions and Problems."

*1. If the manager of the open market desk hears that a snowstorm is about to strike New York City, making it difficult to present checks for payment there and so raising the float, what defensive open market operations will the manager undertake?

2. During Christmastime, when the public's holdings of currency increase, what defensive open market operations typically occur? Why?

*3. If the Treasury has just paid for a supercomputer and as a result its deposits with the Fed fall, what defensive open market operations will the manager of the open market desk undertake?

4. If float decreases below its normal level, why might the manager of domestic operations consider it more desirable to use repurchase agreements to affect the monetary base rather than an outright purchase of bonds?

*5. Most open market operations are currently repurchase agreements. What does this tell us about the likely volume of defensive open market operations relative to dynamic open market operations?

6. "The only way that the Fed can affect the level of discount loans is by adjusting the discount rate." Is this statement true, false, or uncertain? Explain your answer.

*7. Using the supply and demand analysis of the market for reserves, show what happens to the federal funds rate, holding everything else constant, if the economy is surprisingly strong, leading to a rise in the amount of checkable deposits.

8. If there is a switch from deposits into currency, what happens to the federal funds rate? Use the supply and demand analysis of the market for reserves to explain your answer.

*9. "Discounting is no longer needed because the presence of the FDIC eliminates the possibility of bank panics." Discuss.

10. The benefits of using Fed discount operations to prevent bank panics are straightforward. What are the costs?

*11. You often read in the newspaper that the Fed has just lowered the discount rate. Does this signal that the Fed is moving to a more expansionary monetary policy? Why or why not?

12. How can the procyclical movement of interest rates (rising during business cycle expansions and falling during business cycle contractions) lead to a procyclical movement in the money supply as a result of Fed discounting? Why might this movement of the money supply be undesirable?

*13. "If reserve requirements were eliminated, it would be harder to control interest rates." True, false, or uncertain?

14. "Considering that raising reserve requirements to 100% makes complete control of the money supply possible, Congress should authorize the Fed to raise reserve requirements to this level." Discuss.

*15. Compare the use of open market operations, discounting, and changes in reserve requirements to control the money supply on the following criteria: flexibility, reversibility, effectiveness, and speed of implementation.

Web Exercises

1. Go to www.federalreserve.gov/fomc/. This site reports activity by the open market committee. Scroll down to Calendar and click on the statement released after the last meeting. Summarize this statement in one paragraph. Be sure to note whether the committee has decided to increase or decrease the rate of growth of reserves. Now review the statements of the last two meetings. Has the stance of the committee changed?

2. Go to www.federalreserve.gov/releases/h15/update/. What is the current Federal Funds Rate (define this rate as well)? What is the current Federal Reserve Discount rate (define this rate as well)? Is the difference between these rates similar to what is usually observed, based on Figure 4? Have short-term rates increased or declined since the end of 2002?

Chapter 18

Conduct of Monetary Policy: Goals and Targets

PREVIEW

Now that we understand the tools that central banks like the Federal Reserve use to conduct monetary policy, we can proceed to how monetary policy is actually conducted. Understanding the conduct of monetary policy is important, because it not only affects the money supply and interest rates but also has a major influence on the level of economic activity and hence on our well-being.

To explore this subject, we look at the goals that the Fed establishes for monetary policy and its strategies for attaining them. After examining the goals and strategies, we can evaluate the Fed's conduct of monetary policy in the past, with the hope that it will give us some clues to where monetary policy may head in the future.

Goals of Monetary Policy

www.federalreserve.gov/pf/pf.htm

Review what the Federal Reserve reports as its primary purposes and functions.

Six basic goals are continually mentioned by personnel at the Federal Reserve and other central banks when they discuss the objectives of monetary policy: (1) high employment, (2) economic growth, (3) price stability, (4) interest-rate stability, (5) stability of financial markets, and (6) stability in foreign exchange markets.

High Employment

The Employment Act of 1946 and the Full Employment and Balanced Growth Act of 1978 (more commonly called the Humphrey-Hawkins Act) commit the U.S. government to promoting high employment consistent with a stable price level. High employment is a worthy goal for two main reasons: (1) the alternative situation—high unemployment—causes much human misery, with families suffering financial distress, loss of personal self-respect, and increase in crime (though this last conclusion is highly controversial), and (2) when unemployment is high, the economy has not only idle workers but also idle resources (closed factories and unused equipment), resulting in a loss of output (lower GDP).

Although it is clear that high employment is desirable, how high should it be? At what point can we say that the economy is at full employment? At first, it might seem that full employment is the point at which no worker is out of a job; that is, when unemployment is zero. But this definition ignores the fact that some unemployment, called *frictional unemployment*, which involves searches by workers and firms to find suitable matchups, is beneficial to the economy. For example, a worker who decides

to look for a better job might be unemployed for a while during the job search. Workers often decide to leave work temporarily to pursue other activities (raising a family, travel, returning to school), and when they decide to reenter the job market, it may take some time for them to find the right job. The benefit of having some unemployment is similar to the benefit of having a nonzero vacancy rate in the market for rental apartments. As many of you who have looked for an apartment have discovered, when the vacancy rate in the rental market is too low, you will have a difficult time finding the right apartment.

Another reason that unemployment is not zero when the economy is at full employment is due to what is called *structural unemployment*, a mismatch between job requirements and the skills or availability of local workers. Clearly, this kind of unemployment is undesirable. Nonetheless, it is something that monetary policy can do little about.

The goal for high employment should therefore not seek an unemployment level of zero but rather a level above zero consistent with full employment at which the demand for labor equals the supply of labor. This level is called the **natural rate of unemployment**.

Although this definition sounds neat and authoritative, it leaves a troublesome question unanswered: What unemployment rate is consistent with full employment? On the one hand, in some cases, it is obvious that the unemployment rate is too high: The unemployment rate in excess of 20% during the Great Depression, for example, was clearly far too high. In the early 1960s, on the other hand, policymakers thought that a reasonable goal was 4%, a level that was probably too low, because it led to accelerating inflation. Current estimates of the natural rate of unemployment place it between $4\frac{1}{2}$ and 6%, but even this estimate is subject to a great deal of uncertainty and disagreement. In addition, it is possible that appropriate government policy, such as the provision of better information about job vacancies or job training programs, might decrease the natural rate of unemployment.

Economic Growth

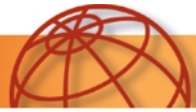
www.economagic.com/
A comprehensive listing of sites that offer a wide variety of economic summary data and graphs.

The goal of steady economic growth is closely related to the high-employment goal because businesses are more likely to invest in capital equipment to increase productivity and economic growth when unemployment is low. Conversely, if unemployment is high and factories are idle, it does not pay for a firm to invest in additional plants and equipment. Although the two goals are closely related, policies can be specifically aimed at promoting economic growth by directly encouraging firms to invest or by encouraging people to save, which provides more funds for firms to invest. In fact, this is the stated purpose of so-called supply-side economics policies, which are intended to spur economic growth by providing tax incentives for businesses to invest in facilities and equipment and for taxpayers to save more. There is also an active debate over what role monetary policy can play in boosting growth.

Price Stability

www.bls.gov/cpi/
View current data on the consumer price index.

Over the past few decades, policymakers in the United States have become increasingly aware of the social and economic costs of inflation and more concerned with a stable price level as a goal of economic policy. Indeed, price stability is increasingly viewed as the most important goal for monetary policy. (This view is also evident in Europe—see Box 1.) Price stability is desirable because a rising price level (inflation) creates uncertainty in the economy, and that uncertainty might hamper economic growth. For example, when the overall level of prices is changing, the information conveyed by the prices of goods and services is harder to interpret, which complicates



Box 1: Global

The Growing European Commitment to Price Stability

Not surprisingly, given Germany's experience with hyperinflation in the 1920s, Germans have had the strongest commitment to price stability as the primary goal for monetary policy. Other Europeans have been coming around to the view that the primary objective for a central bank should be price stability. The increased importance of this goal was reflected in the December 1991 Treaty of European Union, known as the Maastricht Treaty. This treaty created the European System of Central Banks, which functions very much like the Federal Reserve System. The statute of the European System of Central Banks sets price stability as the primary objective of this system and indicates that the general economic policies of the European Union are to be supported only if they are not in conflict with price stability.

decision making for consumers, businesses, and government. Not only do public opinion surveys indicate that the public is very hostile to inflation, but a growing body of evidence suggests that inflation leads to lower economic growth.[1] The most extreme example of unstable prices is *hyperinflation*, such as Argentina, Brazil, and Russia have experienced in the recent past. Many economists attribute the slower growth that these countries have experienced to their problems with hyperinflation.

Inflation also makes it hard to plan for the future. For example, it is more difficult to decide how much funds should be put aside to provide for a child's college education in an inflationary environment. Further, inflation can strain a country's social fabric: Conflict might result, because each group in the society may compete with other groups to make sure that its income keeps up with the rising level of prices.

Interest-Rate Stability

Interest-rate stability is desirable because fluctuations in interest rates can create uncertainty in the economy and make it harder to plan for the future. Fluctuations in interest rates that affect consumers' willingness to buy houses, for example, make it more difficult for consumers to decide when to purchase a house and for construction firms to plan how many houses to build. A central bank may also want to reduce upward movements in interest rates for the reasons we discussed in Chapter 14: Upward movements in interest rates generate hostility toward central banks like the Fed and lead to demands that their power be curtailed.

Stability of Financial Markets

As our analysis in Chapter 8 showed, financial crises can interfere with the ability of financial markets to channel funds to people with productive investment opportunities, thereby leading to a sharp contraction in economic activity. The promotion of a more stable financial system in which financial crises are avoided is thus an important goal for a central bank. Indeed, as discussed in Chapter 14, the Federal Reserve System was created in response to the bank panic of 1907 to promote financial stability.

[1]For example, see Stanley Fischer, "The Role of Macroeconomic Factors in Growth," *Journal of Monetary Economics* 32 (1993): 485–512.

The stability of financial markets is also fostered by interest-rate stability, because fluctuations in interest rates create great uncertainty for financial institutions. An increase in interest rates produces large capital losses on long-term bonds and mortgages, losses that can cause the failure of the financial institutions holding them. In recent years, more pronounced interest-rate fluctuations have been a particularly severe problem for savings and loan associations and mutual savings banks, many of which got into serious financial trouble in the 1980s and early 1990s (as we have seen in Chapter 11).

Stability in Foreign Exchange Markets

With the increasing importance of international trade to the U.S. economy, the value of the dollar relative to other currencies has become a major consideration for the Fed. A rise in the value of the dollar makes American industries less competitive with those abroad, and declines in the value of the dollar stimulate inflation in the United States. In addition, preventing large changes in the value of the dollar makes it easier for firms and individuals purchasing or selling goods abroad to plan ahead. Stabilizing extreme movements in the value of the dollar in foreign exchange markets is thus viewed as a worthy goal of monetary policy. In other countries, which are even more dependent on foreign trade, stability in foreign exchange markets takes on even greater importance.

Conflict Among Goals

Although many of the goals mentioned are consistent with each other—high employment with economic growth, interest-rate stability with financial market stability—this is not always the case. The goal of price stability often conflicts with the goals of interest-rate stability and high employment in the short run (but probably not in the long run). For example, when the economy is expanding and unemployment is falling, both inflation and interest rates may start to rise. If the central bank tries to prevent a rise in interest rates, this might cause the economy to overheat and stimulate inflation. But if a central bank raises interest rates to prevent inflation, in the short run unemployment could rise. The conflict among goals may thus present central banks like the Federal Reserve with some hard choices. We return to the issue of how central banks should choose conflicting goals in later chapters when we examine how monetary policy affects the economy.

Central Bank Strategy: Use of Targets

The central bank's problem is that it wishes to achieve certain goals, such as price stability with high employment, but it does not directly influence the goals. It has a set of tools to employ (open market operations, changes in the discount rate, and changes in reserve requirements) that can affect the goals indirectly after a period of time (typically more than a year). If the central bank waits to see what the price level and employment will be one year later, it will be too late to make any corrections to its policy—mistakes will be irreversible.

All central banks consequently pursue a different strategy for conducting monetary policy by aiming at variables that lie between its tools and the achievement of its goals. The strategy is as follows: After deciding on its goals for employment and the price level, the central bank chooses a set of variables to aim for, called **intermediate targets**, such as the monetary aggregates (M1, M2, or M3) or interest rates (short- or

long-term), which have a direct effect on employment and the price level. However, even these intermediate targets are not directly affected by the central bank's policy tools. Therefore, it chooses another set of variables to aim for, called **operating targets**, or alternatively **instrument targets**, such as reserve aggregates (reserves, nonborrowed reserves, monetary base, or nonborrowed base) or interest rates (federal funds rate or Treasury bill rate), which are more responsive to its policy tools. (Recall that nonborrowed reserves are total reserves minus borrowed reserves, which are the amount of discount loans; the nonborrowed base is the monetary base minus borrowed reserves; and the federal funds rate is the interest rate on funds loaned overnight between banks.)[2]

The central bank pursues this strategy because it is easier to hit a goal by aiming at targets than by aiming at the goal directly. Specifically, by using intermediate and operating targets, it can more quickly judge whether its policies are on the right track, rather than waiting until it sees the final outcome of its policies on employment and the price level.[3] By analogy, NASA employs the strategy of using targets when it is trying to send a spaceship to the moon. It will check to see whether the spaceship is positioned correctly as it leaves the atmosphere (we can think of this as NASA's "operating target"). If the spaceship is off course at this stage, NASA engineers will adjust its thrust (a policy tool) to get it back on target. NASA may check the position of the spaceship again when it is halfway to the moon (NASA's "intermediate target") and can make further midcourse corrections if necessary.

The central bank's strategy works in a similar way. Suppose that the central bank's employment and price-level goals are consistent with a nominal GDP growth rate of 5%. If the central bank feels that the 5% nominal GDP growth rate will be achieved by a 4% growth rate for M2 (its intermediate target), which will in turn be achieved by a growth rate of $3\frac{1}{2}$% for the monetary base (its operating target), it will carry out open market operations (its tool) to achieve the $3\frac{1}{2}$% growth in the monetary base. After implementing this policy, the central bank may find that the monetary base is growing too slowly, say at a 2% rate; then it can correct this too slow growth by increasing the amount of its open market purchases. Somewhat later, the central bank will begin to see how its policy is affecting the growth rate of the money supply. If M2 is growing too fast, say at a 7% rate, the central bank may decide to reduce its open market purchases or make open market sales to reduce the M2 growth rate.

One way of thinking about this strategy (illustrated in Figure 1) is that the central bank is using its operating and intermediate targets to direct monetary policy (the space shuttle) toward the achievement of its goals. After the initial setting of the policy tools (the liftoff), an operating target such as the monetary base, which the central bank can control fairly directly, is used to reset the tools so that monetary policy is channeled toward achieving the intermediate target of a certain rate of money supply growth. Midcourse corrections in the policy tools can be made again when the central bank

---

[2]There is some ambiguity as to whether to call a particular variable an operating target or an intermediate target. The monetary base and the Treasury bill rate are often viewed as possible intermediate targets, even though they may function as operating targets as well. In addition, if the Fed wants to pursue a goal of interest-rate stability, an interest rate can be both a goal and a target.

[3]This reasoning for the use of monetary targets has come under attack, because information on employment and the price level can be useful in evaluating policy. See Benjamin M. Friedman, "The Inefficiency of Short-Run Monetary Targets for Monetary Policy," *Brookings Papers on Economic Activity* 2 (1977): 292–346.

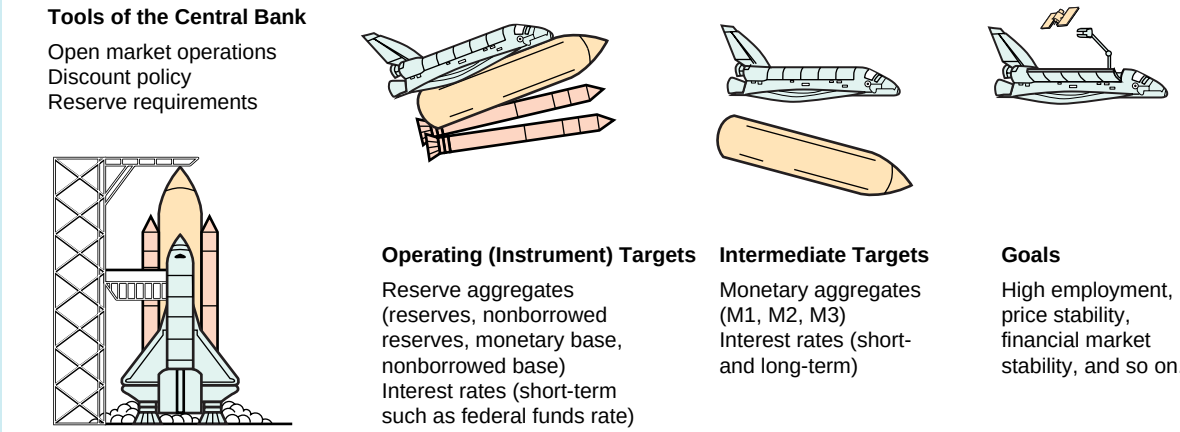


FIGURE 1 Central Bank Strategy

sees what is happening to its intermediate target, thus directing monetary policy so that it will achieve its goals of high employment and price stability (the space shuttle launches the satellite in the appropriate orbit).

Choosing the Targets

As we see in Figure 1, there are two different types of target variables: interest rates and aggregates (monetary aggregates and reserve aggregates). In our example, the central bank chose a 4% growth rate for M2 to achieve a 5% rate of growth for nominal GDP. It could have chosen to lower the interest rate on the three-month Treasury bills to, say, 3% to achieve the same goal. Can the central bank choose to pursue both of these targets at the same time? The answer is no. The application of the supply and demand analysis of the money market that we covered in Chapter 5 explains why a central bank must choose one or the other.

Let's first see why a monetary aggregate target involves losing control of the interest rate. Figure 2 contains a supply and demand diagram for the money market. Although the central bank expects the demand curve for money to be at M^{d*}, it fluctuates between $M^{d'}$ and $M^{d''}$ because of unexpected increases or decreases in output or changes in the price level. The money demand curve might also shift unexpectedly because the public's preferences about holding bonds versus money could change. If the central bank's monetary aggregate target of a 4% growth rate in M2 results in a money supply of M^*, it expects that the interest rate will be i^*. However, as the figure indicates, the fluctuations in the money demand curve between $M^{d'}$ and $M^{d''}$ will result in an interest rate fluctuating between i' and i''. Pursuing a monetary aggregate target implies that interest rates will fluctuate.

The supply and demand diagram in Figure 3 shows the consequences of an interest-rate target set at i^*. Again, the central bank expects the money demand curve to be at M^{d*}, but it fluctuates between $M^{d'}$ and $M^{d''}$ due to unexpected changes in output, the

FIGURE 2 Result of Targeting on the Money Supply
Targeting on the money supply at M^* will lead to fluctuations in the interest rate between i' and i'' because of fluctuations in the money demand curve between $M^{d'}$ and $M^{d''}$.

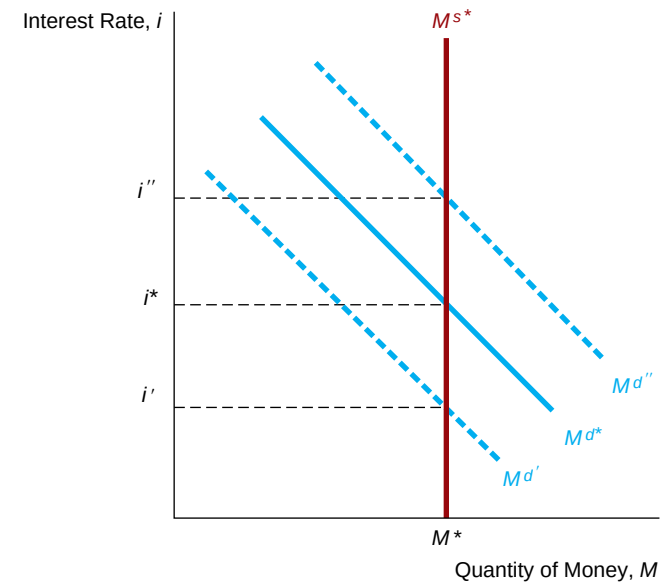
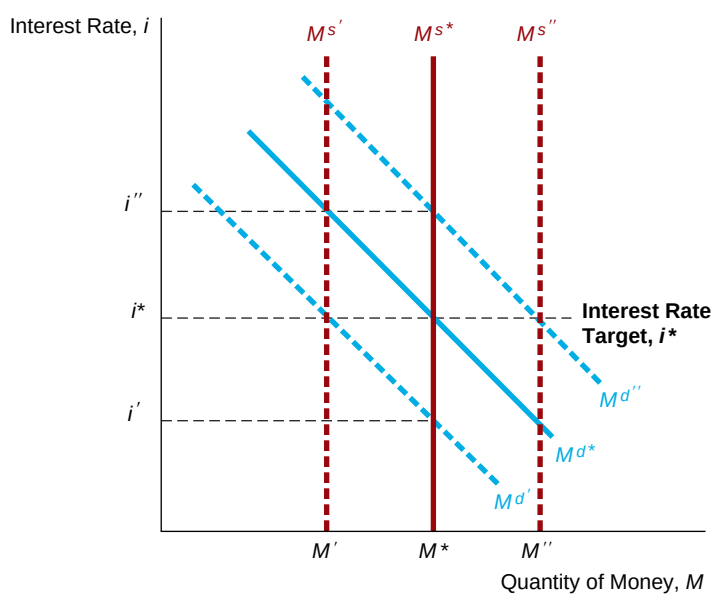


FIGURE 3 Result of Targeting on the Interest Rate
Targeting the interest rate at M^* will lead to fluctuations of the money supply between M' and M'' because of fluctuations in the money demand curve between $M^{d'}$ and $M^{d''}$.



price level, or the public's preferences toward holding money. If the demand curve falls to $M^{d'}$, the interest rate will begin to fall below i^*, and the price of bonds will rise. With an interest-rate target, the central bank will prevent the interest rate from falling by selling bonds to drive their price back down and the interest rate back up

to its former level. The central bank will make open market sales until the money supply declines to $M^{s'}$, at which point the equilibrium interest rate is again i^*. Conversely, if the demand curve rises to $M^{d''}$ and drives up the interest rate, the central bank would keep interest rates from rising by buying bonds to keep their prices from falling. The central bank will make open market purchases until the money supply rises to $M^{s''}$ and the equilibrium interest rate is i^*. The central bank's adherence to the interest-rate target thus leads to a fluctuating money supply as well as fluctuations in reserve aggregates such as the monetary base.

The conclusion from the supply and demand analysis is that interest-rate and monetary aggregate targets are incompatible: A central bank can hit one or the other but not both. Because a choice between them has to be made, we need to examine what criteria should be used to decide on the target variable.

Criteria for Choosing Intermediate Targets

The rationale behind a central bank's strategy of using targets suggests three criteria for choosing an intermediate target: It must be measurable, it must be controllable by the central bank, and it must have a predictable effect on the goal.

Measurability. Quick and accurate measurement of an intermediate-target variable is necessary, because the intermediate target will be useful only if it signals rapidly when policy is off track. What good does it do for the central bank to plan to hit a 4% growth rate for M2 if it has no way of quickly and accurately measuring M2? Data on the monetary aggregates are obtained after a two-week delay, and interest-rate data are available almost immediately. Data on a variable like GDP that serves as a goal, by contrast, are compiled quarterly and are made available with a month's delay. In addition, the GDP data are less accurate than data on the monetary aggregates or interest rates. On these grounds alone, focusing on interest rates and monetary aggregates as intermediate targets rather than on a goal like GDP can provide clearer signals about the status of the central bank's policy.

At first glance, interest rates seem to be more measurable than monetary aggregates and hence more useful as intermediate targets. Not only are the data on interest rates available more quickly than on monetary aggregates, but they are also measured more precisely and are rarely revised, in contrast to the monetary aggregates, which are subject to a fair amount of revision (as we saw in Chapter 3). However, as we learned in Chapter 4, the interest rate that is quickly and accurately measured, the nominal interest rate, is typically a poor measure of the real cost of borrowing, which indicates with more certainty what will happen to GDP. This real cost of borrowing is more accurately measured by the real interest rate—the interest rate adjusted for expected inflation ($i_r = i - \pi^e$). Unfortunately, the real interest rate is extremely hard to measure, because we have no direct way to measure expected inflation. Since both interest rate and monetary aggregates have measurability problems, it is not clear whether one should be preferred to the other as an intermediate target.

Controllability. A central bank must be able to exercise effective control over a variable if it is to function as a useful target. If the central bank cannot control an intermediate target, knowing that it is off track does little good, because the central bank has no way of getting the target back on track. Some economists have suggested that nominal GDP should be used as an intermediate target, but since the central bank has little direct control over nominal GDP, it will not provide much guidance on how the Fed should set its policy tools. A central bank does, however, have a good deal of control over the monetary aggregates and interest rates.

Our discussion of the money supply process and the central bank's policy tools indicates that a central bank does have the ability to exercise a powerful effect on the money supply, although its control is not perfect. We have also seen that open market operations can be used to set interest rates by directly affecting the price of bonds. Because a central bank can set interest rates directly, whereas it cannot completely control the money supply, it might appear that interest rates dominate the monetary aggregates on the controllability criterion. However, a central bank cannot set real interest rates, because it does not have control over expectations of inflation. So again, a clear-cut case cannot be made that interest rates are preferable to monetary aggregates as an intermediate target or vice versa.

Predictable Effect on Goals. The most important characteristic a variable must have to be useful as an intermediate target is that it must have a predictable impact on a goal. If a central bank can accurately and quickly measure the price of tea in China and can completely control its price, what good will it do? The central bank cannot use the price of tea in China to affect unemployment or the price level in its country. Because the ability to affect goals is so critical to the usefulness of an intermediate-target variable, the linkage of the money supply and interest rates with the goals—output, employment, and the price level—is a matter of much debate. The evidence on whether these goals have a closer (more predictable) link with the money supply than with interest rates is discussed in Chapter 26.

Criteria for Choosing Operating Targets

The choice of an operating target can be based on the same criteria used to evaluate intermediate targets. Both the federal funds rate and reserve aggregates are measured accurately and are available daily with almost no delay; both are easily controllable using the policy tools that we discussed in Chapter 17. When we look at the third criterion, however, we can think of the intermediate target as the goal for the operating target. An operating target that has a more predictable impact on the most desirable intermediate target is preferred. If the desired intermediate target is an interest rate, the preferred operating target will be an interest-rate variable like the federal funds rate because interest rates are closely tied to each other (as we saw in Chapter 6). However, if the desired intermediate target is a monetary aggregate, our money supply model in Chapters 15 and 16 shows that a reserve aggregate operating target such as the monetary base will be preferred. Because there does not seem to be much reason to choose an interest rate over a reserve aggregate on the basis of measurability or controllability, the choice of which operating target is better rests on the choice of the intermediate target (the goal of the operating target).

Fed Policy Procedures: Historical Perspective

The well-known adage "The road to hell is paved with good intentions" applies as much to the Federal Reserve as it does to human beings. Understanding a central bank's goals and the strategies it can use to pursue them cannot tell us how monetary policy is actually conducted. To understand the practical results of the theoretical underpinnings, we have to look at how central banks have actually conducted policy in the past. First we will look at the Federal Reserve's past policy procedures: its choice of goals, policy tools, operating targets, and intermediate targets. This historical perspective will not only show us how our central bank carries out its duties but

will also help us interpret the Fed's activities and see where U.S. monetary policy may be heading in the future. Once we are done studying the Fed, we will then examine central banks' experiences in other countries.

Study Guide

The following discussion of the Fed's policy procedures and their effect on the money supply provides a review of the money supply process and how the Fed's policy tools work. If you have trouble understanding how the particular policies described affect the money supply, it might be helpful to review the material in Chapters 15 and 16.

The Early Years: Discount Policy as the Primary Tool

When the Fed was created, changing the discount rate was the primary tool of monetary policy—the Fed had not yet discovered that open market operations were a more powerful tool for influencing the money supply, and the Federal Reserve Act made no provisions for changes in reserve requirements. The guiding principle for the conduct of monetary policy was that as long as loans were being made for "productive" purposes—that is, to support the production of goods and services—providing reserves to the banking system to make these loans would not be inflationary.[4] This theory, now thoroughly discredited, became known as the **real bills doctrine**. In practice, it meant that the Fed would make loans to member commercial banks when they showed up at the discount window with *eligible paper*, loans to facilitate the production and sale of goods and services. (Note that since the 1920s, the Fed has not conducted discount operations in this way.) The Fed's act of making loans to member banks was initially called *rediscounting*, because the original bank loans to businesses were made by discounting (loaning less than) the face value of the loan, and the Fed would be discounting them again. (Over time, when the Fed's emphasis on eligible paper diminished, the Fed's loans to banks became known as *discounts*, and the interest rate on these loans the *discount rate*, which is the terminology we use today.)

By the end of World War I, the Fed's policy of rediscounting eligible paper and keeping interest rates low to help the Treasury finance the war had led to a raging inflation; in 1919 and 1920, the inflation rate averaged 14%. The Fed decided that it could no longer follow the passive policy prescribed by the real bills doctrine because it was inconsistent with the goal of price stability, and for the first time the Fed accepted the responsibility of playing an active role in influencing the economy. In January 1920, the Fed raised the discount rate from $4\frac{3}{4}$% to 6%, the largest jump in its history, and eventually raised it further, to 7% in June 1920, where it remained for nearly a year. The result of this policy was a sharp decline in the money supply and an especially sharp recession in 1920–1921. Although the blame for this severe recession can clearly be laid at the Fed's doorstep, in one sense the Fed's policy was very successful: After an initial decline in the price level, the inflation rate went to zero, paving the way for the prosperous Roaring Twenties.

Discovery of Open Market Operations

In the early 1920s, a particularly important event occurred: The Fed accidentally discovered open market operations. When the Fed was created, its revenue came exclusively from the interest it received on the discount loans it made to member banks. After the 1920–1921 recession, the volume of discount loans shrank dramatically, and

[4] Another guiding principle was the maintenance of the gold standard, which we will discuss in Chapter 20.

the Fed was pressed for income. It solved this problem by purchasing income-earning securities. In doing so, the Fed noticed that reserves in the banking system grew and there was a multiple expansion of bank loans and deposits. This result is obvious to us now (we studied the multiple deposit creation process in Chapter 15), but to the Fed at that time it was a revelation. A new monetary policy tool was born, and by the end of the 1920s, it was the most important weapon in the Fed's arsenal.

The Great Depression

The stock market boom in 1928 and 1929 created a dilemma for the Fed. It wanted to temper the boom by raising the discount rate, but it was reluctant to do so, because that would mean raising interest rates to businesses and individuals who had legitimate needs for credit. Finally, in August 1929, the Fed raised the discount rate, but by then it was too late; the speculative excesses of the market boom had already occurred, and the Fed's action only hastened the stock market crash and pushed the economy into recession.

The weakness of the economy, particularly in the agricultural sector, led to what Milton Friedman and Anna Schwartz labeled a "contagion of fear" that triggered substantial withdrawals from banks, building to a full-fledged panic in November and December 1930. For the next two years, the Fed sat idly by while one bank panic after another occurred, culminating in the final panic in March 1933, at which point the new president, Franklin Delano Roosevelt, declared a bank holiday. (Why the Fed failed to engage in its lender-of-last-resort role during this period is discussed in Box 2.)

Box 2: Inside the Fed

Bank Panics of 1930–1933: Why Did the Fed Let Them Happen?

The Federal Reserve System was totally passive during the bank panics of the Great Depression period and did not perform its intended role of lender of last resort to prevent them. In retrospect, the Fed's behavior seems quite extraordinary, but hindsight is always clearer than foresight.

The primary reason for the Fed's inaction was that Federal Reserve officials did not understand the negative impact that bank failures could have on the money supply and economic activity. Friedman and Schwartz report that the Federal Reserve officials "tended to regard bank failures as regrettable consequences of bank management or bad banking practices, or as inevitable reactions to prior speculative excesses, or as a consequence but hardly a cause of the financial and economic collapse in process." In addition, bank failures in the early stages of the bank panics "were concentrated among smaller banks and, since the most influential figures in the system were big-city bankers who deplored the existence of smaller banks, their disappearance may have been viewed with complacency."*

Friedman and Schwartz also point out that political infighting may have played an important role in the passivity of the Fed during this period. The Federal Reserve Bank of New York, which until 1928 was the dominant force in the Federal Reserve System, strongly advocated an active program of open market purchases to provide reserves to the banking system during the bank panics. However, other powerful figures in the Federal Reserve System opposed the New York bank's position, and the bank was outvoted. (Friedman and Schwartz's discussion of the politics of the Federal Reserve System during this period makes for fascinating reading, and you might enjoy their highly readable book.)

*Milton Friedman and Anna Jacobson Schwartz, *A Monetary History of the United States, 1867–1960* (Princeton, N.J.: Princeton University Press, 1963), p. 358.

The spate of bank panics from 1930 to 1933 were the most severe in U.S. history, and Roosevelt aptly summed up the problem in his statement "The only thing we have to fear is fear itself." By the time the panics were over in March 1933, more than one-third of the commercial banks in the United States had failed.

In Chapter 16, we examined how the bank panics of this period led to a decline in the money supply by over 25%. The resulting unprecedented decline in the money supply during this period is thought by many economists, particularly monetarists, to have been the major contributing factor to the severity of the depression, never equaled before or since.

Reserve Requirements as a Policy Tool

The Thomas Amendment to the Agricultural Adjustment Act of 1933 provided the Federal Reserve's Board of Governors with emergency power to alter reserve requirements with the approval of the president of the United States. In the Banking Act of 1935, this emergency power was expanded to allow the Fed to alter reserve requirements without the president's approval.

The first use of reserve requirements as a tool of monetary control proved that the Federal Reserve was capable of adding to the blunders that it had made during the bank panics of the early 1930s. By the end of 1935, banks had increased their holdings of excess reserves to unprecedented levels, a sensible strategy, considering their discovery during the 1930–1933 period that the Fed would not always perform its intended role as lender of last resort. Bankers now understood that they would have to protect themselves against a bank run by holding substantial amounts of excess reserves. The Fed viewed these excess reserves as a nuisance that made it harder to exercise monetary control. Specifically, the Fed worried that these excess reserves might be lent out and would produce "an uncontrollable expansion of credit in the future."[5]

To improve monetary control, the Fed raised reserve requirements in three steps: August 1936, January 1937, and May 1937. The result of this action was, as we would expect from our money supply model, a slowdown of money growth toward the end of 1936 and an actual decline in 1937. The recession of 1937–1938, which commenced in May 1937, was a severe one and was especially upsetting to the American public because even at its outset unemployment was intolerably high. So not only does it appear that the Fed was at fault for the severity of the Great Depression contraction in 1929–1933, but to add insult to injury, it appears that it was also responsible for aborting the subsequent recovery. The Fed's disastrous experience with varying its reserve requirements made it far more cautious in the use of this policy tool in the future.

War Finance and the Pegging of Interest Rates: 1942–1951

With the entrance of the United States into World War II in late 1941, government spending skyrocketed, and to finance it, the Treasury issued huge amounts of bonds. The Fed agreed to help the Treasury finance the war cheaply by pegging interest rates at the low levels that had prevailed before the war: $\frac{3}{8}$% on Treasury bills and $2\frac{1}{2}$% on long-term Treasury bonds. Whenever interest rates rose above these levels and the price of bonds began to fall, the Fed would make open market purchases, thereby bidding up bond prices and driving interest rates down again. The result was a rapid

[5]Milton Friedman and Anna Jacobson Schwartz, *A Monetary History of the United States, 1867–1960* (Princeton, N.J.: Princeton University Press, 1963), p. 524.

growth in the monetary base and the money supply. The Fed had thus in effect relinquished its control of monetary policy to meet the financing needs of the government.

When the war ended, the Fed continued to peg interest rates, and because there was little pressure on them to rise, this policy did not result in an explosive growth in the money supply. When the Korean War broke out in 1950, however, interest rates began to climb, and the Fed found that it was again forced to expand the monetary base at a rapid rate. Because inflation began to heat up (the consumer price index rose 8% between 1950 and 1951), the Fed decided that it was time to reassert its control over monetary policy by abandoning the interest-rate peg. An often bitter debate ensued between the Fed and the Treasury, which wanted to keep its interest costs down and so favored a continued pegging of interest rates at low levels. In March 1951, the Fed and the Treasury came to an agreement known as the Accord, in which pegging was abandoned but the Fed promised that it would not allow interest rates to rise precipitously. After Eisenhower's election as president in 1952, the Fed was given complete freedom to pursue its monetary policy objectives.

Targeting Money Market Conditions: The 1950s and 1960s

With its freedom restored, the Federal Reserve, then under the chairmanship of William McChesney Martin Jr., took the view that monetary policy should be grounded in intuitive judgment based on a feel for the money market. The policy procedure that resulted can be described as one in which the Fed targeted on money market conditions, and particularly on interest rates.

An important characteristic of this policy procedure was that it led to more rapid growth in the money supply when the economy was expanding and a slowing of money growth when the economy was in recession. The so-called *procyclical monetary policy* (a positive association of money supply growth with the business cycle) is explained by the following step-by-step reasoning. As we learned in Chapter 5, a rise in national income ($Y\uparrow$) leads to a rise in market interest rates ($i\uparrow$). With the rise in interest rates, the Fed would purchase bonds to bid their price up and lower interest rates to their target level. The resulting increase in the monetary base caused the money supply to rise and the business cycle expansion to be accompanied by a faster rate of money growth. In summary:

$$Y\uparrow \Rightarrow i\uparrow \Rightarrow MB\uparrow \Rightarrow M\uparrow$$

In a recession, the opposite sequence of events would occur, and the decline in income would be accompanied by a slower rate of growth in the money supply ($Y\downarrow \Rightarrow M\downarrow$).

A further problem with using interest rates as the primary operating target is that they may encourage an inflationary spiral to get out of control. As we saw in Chapter 5, when inflation and hence expected inflation rises, nominal interest rates rise via the Fisher effect. If the Fed attempted to prevent this increase by purchasing bonds, this would also lead to a rise in the monetary base and the money supply:

$$\pi\uparrow \Rightarrow \pi^e\uparrow \Rightarrow i\uparrow \Rightarrow MB\uparrow \Rightarrow M\uparrow$$

Higher inflation could thus lead to an increase in the money supply, which would increase inflationary pressures further.

By the late 1960s, the rising chorus of criticism of procyclical monetary policy by such prominent monetarist economists such as Milton Friedman, Karl Brunner, and

Allan Meltzer and concerns about inflation finally led the Fed to abandon its focus on money market conditions.

Targeting Monetary Aggregates: The 1970s

In 1970, Arthur Burns was appointed chairman of the Board of Governors, and soon thereafter the Fed stated that it was committing itself to the use of monetary aggregates as intermediate targets. Did monetary policy cease to be procyclical? A glance at Figure 4 in Chapter 1 indicates that monetary policy was as procyclical in the 1970s as in the 1950s and 1960s. What went wrong? Why did the conduct of monetary policy not improve? The answers to these questions lie in the Fed's operating procedures during the period, which suggest that its commitment to targeting monetary aggregates was not very strong.

Every six weeks, the Federal Open Market Committee would set target ranges for the growth rates of various monetary aggregates and would determine what federal funds rate (the interest rate on funds loaned overnight between banks) it thought consistent with these aims. The target ranges for the growth in monetary aggregates were fairly broad—a typical range for M1 growth might be 3% to 6%; for M2, 4% to 7%—while the range for the federal funds rate was a narrow band, say from $7\frac{1}{2}\%$ to $8\frac{1}{4}\%$. The trading desk at the Federal Reserve Bank of New York was then instructed to meet both sets of targets, but as we saw earlier, interest-rate targets and monetary aggregate targets might not be compatible. If the two targets were incompatible—say, the federal funds rate began to climb higher than the top of its target band when M1 was growing too rapidly—the trading desk was instructed to give precedence to the federal funds rate target. In the situation just described, this would mean that although M1 growth was too high, the trading desk would make open market purchases to keep the federal funds rate within its target range.

The Fed was actually using the federal funds rate as its operating target. During the six-week period between FOMC meetings, an unexpected rise in output (which would cause the federal funds rate to hit the top of its target band) would then induce open market purchases and a too rapid growth of the money supply. When the FOMC met again, it would try to bring money supply growth back on track by raising the target range on the federal funds rate. However, if income continued to rise unexpectedly, money growth would overshoot again. This is exactly what occurred from June 1972 to June 1973, when the economy boomed unexpectedly: M1 growth greatly exceeded its target, increasing at approximately an 8% rate, while the federal funds rate climbed from $4\frac{1}{2}\%$ to $8\frac{1}{2}\%$. The economy soon became overheated, and inflationary pressures began to mount.

The opposite chain of events occurred at the end of 1974, when the economic contraction was far more severe than anyone had predicted. The federal funds rate fell dramatically, from over 12% to 5%, and persistently bumped against the bottom of its target range. The trading desk conducted open market sales to keep the federal funds rate from falling, and money growth dropped precipitously, actually turning negative by the beginning of 1975. Clearly, this sharp drop in money growth when the United States was experiencing one of the worst economic contractions of the postwar era was a serious mistake.

Using the federal funds rate as an operating target promoted a procyclical monetary policy despite the Fed's lip service to monetary aggregate targets. If the Federal Reserve really intended to pursue monetary aggregate targets, it seems peculiar that it would have chosen an interest rate for an operating target rather than a reserve aggregate. The explanation for the Fed's choice of an interest rate as an operating target is

that it was still very concerned with achieving interest-rate stability and was reluctant to relinquish control over interest-rate movements. The incompatibility of the Fed's policy procedure with its stated intent of targeting on the monetary aggregates had become very clear by October 1979, when the Fed's policy procedures underwent drastic revision.

New Fed Operating Procedures: October 1979–October 1982

In October 1979, two months after Paul Volcker became chairman of the Board of Governors, the Fed finally deemphasized the federal funds rate as an operating target by widening its target range more than fivefold: A typical range might be from 10% to 15%. The primary operating target became nonborrowed reserves, which the Fed would set after estimating the volume of discount loans the banks would borrow. Not surprisingly, the federal funds rate underwent much greater fluctuations after it was deemphasized as an operating target. What is surprising, however, is that the deemphasis of the federal funds target did not result in improved monetary control: After October 1979, the fluctuations in the rate of money supply growth *increased* rather than decreased as would have been expected. In addition, the Fed missed its M1 growth target ranges in all three years of the 1979–1982 period.[6] What went wrong?

There are several possible answers to this question. The first is that the economy was exposed to several shocks during this period that made monetary control more difficult: the acceleration of financial innovation and deregulation, which added new categories of deposits such as NOW accounts to the measures of monetary aggregates; the imposition by the Fed of credit controls from March to July 1980, which restricted the growth of consumer and business loans; and the back-to-back recessions of 1980 and 1981–1982.[7]

A more persuasive explanation for poor monetary control, however, is that controlling the money supply was never really the intent of Volcker's policy shift. Despite Volcker's statements about the need to target monetary aggregates, he was not committed to these targets. Rather, he was far more concerned with using interest-rate movements to wring inflation out of the economy. Volcker's primary reason for changing the Fed's operating procedure was to free his hand to manipulate interest rates in order to fight inflation. It was necessary to abandon interest-rate targets if Volcker were to be able to raise interest rates sharply when a slowdown in the economy was required to dampen inflation. This view of Volcker's strategy suggests that the Fed's announced attachment to monetary aggregate targets may have been a smokescreen to keep the Fed from being blamed for the high interest rates that would result from the new policy.

---

[6]The M1 target ranges and actual growth rates for 1980–1982 were as follows:

Year	*Target Range (%)*	*Actual (%)*
1980	4.5–7.0	7.5
1981	6.0–8.5	5.1
1982	2.5–5.5	8.8

Source: Board of Governors of the Federal Reserve System, *Monetary Policy Objectives, 1981–1983.*

[7]Another explanation focuses on the technical difficulties of monetary control when using a nonborrowed reserves operating target under a system of lagged reserve requirements, in which required reserves for a given week are calculated on the basis of the level of deposits two weeks earlier. See David Lindsey, "Nonborrowed Reserve Targeting and Monetary Control," in *Improving Money Stock Control*, ed. Laurence Meyer (Boston: Kluwer-Nijhoff, 1983), pp. 3–41.

Interest-rate movements during this period support this interpretation of Fed strategy. After the October 1979 announcement, short-term interest rates were driven up by nearly 5%, until in March 1980 they exceeded 15%. With the imposition of credit controls in March 1980 and the rapid decline in real GDP in the second quarter of 1980, the Fed eased up on its policy and allowed interest rates to decline sharply. When recovery began in July 1980, inflation remained persistent, still exceeding 10%. Because the inflation fight was not yet won, the Fed tightened the screws again, sending short-term rates above the 15% level for a second time. The 1981–1982 recession and its large decline in output and high unemployment began to bring inflation down. With inflationary psychology apparently broken, interest rates were allowed to fall.

The Fed's anti-inflation strategy during the October 1979–October 1982 period was neither intended nor likely to produce smooth growth in the monetary aggregates. Indeed, the large fluctuations in interest rates and the business cycle, along with financial innovation, helped generate volatile money growth.

De-emphasis of Monetary Aggregates: October 1982–Early 1990s

www.federalreserve.gov/releases/H3

Historic and current data on the aggregate reserves of depository institutions and the monetary base.

In October 1982, with inflation in check, the Fed returned, in effect, to a policy of smoothing interest rates. It did this by placing less emphasis on monetary aggregate targets and shifting to borrowed reserves (discount loan borrowings) as an operating target. To see how a borrowed reserves target produces interest-rate smoothing, let's consider what happens when the economy expands ($Y\uparrow$) so that interest rates are driven up. The rise in interest rates ($i\uparrow$) increases the incentives for banks to borrow more from the Fed, so borrowed reserves rise ($DL\uparrow$). To prevent the resulting rise in borrowed reserves from exceeding the target level, the Fed must lower interest rates by bidding up the price of bonds through open market purchases. The outcome of targeting on borrowed reserves, then, is that the Fed prevents a rise in interest rates. In doing so, however, the Fed's open market purchases increase the monetary base ($MB\uparrow$) and lead to a rise in the money supply ($M\uparrow$), which produces a positive association of money and national income ($Y\uparrow \Rightarrow M\uparrow$). Schematically,

$$Y\uparrow \Rightarrow i\uparrow \Rightarrow DL\uparrow \Rightarrow MB\uparrow \Rightarrow M\uparrow$$

A recession causes the opposite chain of events: The borrowed reserves target prevents interest rates from falling and results in a drop in the monetary base, leading to a fall in the money supply ($Y\downarrow \Rightarrow M\downarrow$).

The de-emphasis of monetary aggregates and the change to a borrowed reserves target led to much smaller fluctuations in the federal funds rate after October 1982 but continued to have large fluctuations in money supply growth. Finally, in February 1987, the Fed announced that it would no longer even set M1 targets. The abandonment of M1 targets was defended on two grounds. The first was that the rapid pace of financial innovation and deregulation had made the definition and measurement of money very difficult. The second is that there had been a breakdown in the stable relationship between M1 and economic activity (discussed in Chapter 22). These two arguments suggested that a monetary aggregate such as M1 might no longer be a reliable guide for monetary policy. As a result, the Fed switched its focus to the broader monetary aggregate M2, which it felt had a more stable relationship with economic activity. However, in the early 1990s, this relationship also broke down, and in July 1993, Board of Governors Chairman Alan Greenspan testified in Congress that the Fed would no longer use any monetary targets, including M2, as a guide for conducting monetary policy.

Finally, legislation in 2000 amending the Federal Reserve Act dropped the requirement that the Fed report target ranges for monetary aggregates to Congress.

Federal Funds Targeting Again: Early 1990s and Beyond

Having abandoned monetary aggregates as a guide for monetary policy, the Federal Reserve returned to using a federal funds target in the early 1990s. Indeed, from late 1992 until February 1994, a period of a year and a half, the Fed kept the federal funds rate targeted at the constant rate of 3%, a low level last seen in the 1960s. The explanation for this unusual period of keeping the federal funds rate pegged so low for such a long period of time was fear on the part of the Federal Reserve that the credit crunch mentioned in Chapter 9 was putting a drag on the economy (the "headwinds" referred to by Greenspan) that was producing a sluggish recovery from the 1990–1991 recession. Starting in February 1994, after the economy returned to rapid growth, the Fed began a preemptive strike to head off any future inflationary pressures by raising the federal funds rate in steps to 6% by early 1995. The Fed not only has engaged in preemptive strikes against a rise in inflation, but it has acted preemptively against negative shocks to demand. It lowered the federal funds rate in early 1996 to deal with a possible slowing in the economy and took the dramatic step of reducing the federal funds rate by $\frac{3}{4}$ of a percentage point when the collapse of Long Term Capital Management in the fall of 1998 (discussed in Chapter 12) led to concerns about the health of the financial system. With the strong growth of the economy in 1999 and heightened concerns about inflation, the Fed reversed course and began to raise the federal funds rate again. The Fed's timely actions kept the economy on track, helping to produce the longest business cycle expansion in U.S. history. With a weakening economy, in January 2001 (just before the start of the recession in March 2001) the Fed reversed course again and began to reduce sharply the federal funds rate from its height of 6.5% to near 1% eventually.

In February 1994, with the first change in the federal funds rate in a year and a half, the Fed adopted a new policy procedure. Instead of keeping the federal funds target secret, as it had done previously, the Fed now announced any federal funds rate target change. As mentioned in Chapter 14, around 2:15 P.M., after every FOMC meeting, the Fed now announces whether the federal funds rate target has been raised, lowered, or kept the same. This move to greater transparency of Fed policy was followed by another such move, when in February 1999 the Fed indicated that in the future it would announce the direction of bias to where the federal funds rate will head in the future. However, dissatisfaction with the confusion that the bias announcement created for market participants led the Fed to revise its policy, and starting in February 2000, the Fed switched to an announcement of a statement outlining the "balance of risks" in the future, whether toward higher inflation or toward a weaker economy. As a result of these announcements, the outcome of the FOMC meeting is now big news, and the media devote much more attention to FOMC meeting, because announced changes in the federal funds rate feeds into changes in other interest rates that affect consumers and businesses.

International Considerations

The increasing importance of international trade to the American economy has brought international considerations to the forefront of Federal Reserve policymaking in recent years. By 1985, the strength of the dollar had contributed to a deterioration in American competitiveness with foreign businesses. In public pronouncements, Chairman Volcker and other Fed officials made it clear that the dollar was at too high a value and needed to come down. Because, as we will see in Chapter 19, expansionary monetary policy is

Box 3: Global

International Policy Coordination

The Plaza Agreement and the Louvre Accord. By 1985, the decrease in the competitiveness of American corporations as a result of the strong dollar was raising strong sentiment in Congress for restricting imports. This protectionist threat to the international trading system stimulated finance ministers and the heads of central banks from the Group of Five (G-5) industrial countries—the United States, the United Kingdom, France, West Germany, and Japan—to reach an agreement at New York's Plaza Hotel in September 1985 to bring down the value of the dollar. From September 1985 until the beginning of 1987, the value of the dollar did indeed undergo a substantial decline, falling by 35 percent on average relative to foreign currencies. At this point, there was growing controversy over the decline in the dollar, and another meeting of policymakers from the G-5 countries plus Canada took place in February 1987 at the Louvre Museum in Paris. There the policymakers agreed that exchange rates should be stabilized around the levels currently prevailing. Although the value of the dollar did continue to fluctuate relative to foreign currencies after the Louvre Accord, its downward trend had been checked as intended.

Because subsequent exchange rate movements were pretty much in line with the Plaza Agreement and the Louvre Accord, these attempts at international policy coordination have been considered successful. However, other aspects of the agreements were not adhered to by all signatories. For example, West German and Japanese policymakers agreed that their countries should pursue more expansionary policies by increasing government spending and cutting taxes, and the United States agreed to try to bring down its budget deficit. At that time, the United States was not particularly successful in lowering its deficit, and the Germans were reluctant to pursue expansionary policies because of their concerns about inflation.

www.federalreserve.gov/centralbanks.htm
The Federal Reserve provides links to other central bank web pages.

one way to lower the value of the dollar, it is no surprise that the Fed engineered an acceleration in the growth rates of the monetary aggregates in 1985 and 1986 and that the value of the dollar declined. By 1987, policymakers at the Fed agreed that the dollar had fallen sufficiently, and sure enough, monetary growth in the United States slowed. These monetary policy actions by the Fed were encouraged by the process of **international policy coordination** (agreements among countries to enact policies cooperatively) that led to the Plaza Agreement in 1985 and the Louvre Accord in 1987 (see Box 3).

International considerations also played a role in the Fed's decision to lower the federal funds rate by $\frac{3}{4}$ of a percentage point in the fall of 1998. Concerns about the potential for a worldwide financial crisis in the wake of the collapse of the Russian financial system at that time and weakness in economies abroad, particularly in Asia, stimulated the Fed to take a dramatic step to calm down markets. International considerations, although not the primary focus of the Federal Reserve, are likely to be a major factor in the conduct of American monetary policy in the future.

The Taylor Rule, NAIRU, and the Phillips Curve

As we have seen, the Federal Reserve currently conducts monetary policy by setting a target for the federal funds rate. But how should this target be chosen?

John Taylor of Stanford University has come up with an answer, his so-called **Taylor rule**. The Taylor rule indicates that the federal (fed) funds rate should be set equal to the inflation rate plus an "equilibrium" real fed funds rate (the real fed funds

rate that is consistent with full employment in the long run) plus a weighted average of two gaps: (1) an inflation gap, current inflation minus a target rate, and (2) an output gap, the percentage deviation of real GDP from an estimate of its potential full employment level.[8] This rule can be written as follows:

$$\text{Federal funds rate target} = \text{inflation rate} + \text{equilibrium real fed funds rate} + 1/2\ (\text{inflation gap}) + 1/2\ (\text{output gap})$$

Taylor has assumed that the equilibrium real fed funds rate is 2% and that an appropriate target for inflation would also be 2%, with equal weights of 1/2 on the inflation and output gaps. For an example of the Taylor rule in practice suppose that the inflation rate were at 3%, leading to a positive inflation gap of 1% (= 3% − 2%), and real GDP was 1% above its potential, resulting in a positive output gap of 1%. Then the Taylor rule suggests that the federal funds rate should be set at 6% [= 3% inflation + 2% equilibrium real fed funds rate + 1/2 (1% inflation gap) + 1/2 (1% output gap)].

The presence of both an inflation gap and an output gap in the Taylor rule might indicate that the Fed should care not only about keeping inflation under control, but also about minimizing business cycle fluctuations of output around its potential. Caring about both inflation and output fluctuations is consistent with many statements by Federal Reserve officials that controlling inflation and stabilizing real output are important concerns of the Fed.

An alternative interpretation of the presence of the output gap in the Taylor rule is that the output gap is an indicator of future inflation as stipulated in **Phillips curve theory**. Phillips curve theory indicates that changes in inflation are influenced by the state of the economy relative to its productive capacity, as well as to other factors. This productive capacity can be measured by potential GDP, which is a function of the natural rate of unemployment, the rate of unemployment consistent with full employment. A related concept is the **NAIRU**, the **nonaccelerating inflation rate of unemployment**, the rate of unemployment at which there is no tendency for inflation to change.[9] Simply put, the theory states that when the unemployment rate is above NAIRU with output below potential, inflation will come down, but if it is below NAIRU with output above potential, inflation will rise. Prior to 1995, the NAIRU was thought to reside around 6%. However, with the decline in unemployment to around the 4% level in the late 1990s, with no increase in inflation and even a slight decrease, some critics have questioned the value of Phillips curve theory. Either they claim that it just doesn't work any more or alternatively believe that there is great uncertainty about the value of NAIRU, which may have fallen to below 5% for reasons that are not absolutely clear. Phillips curve theory is now highly controversial, and many economists believe that it should not be used as a guide for the conduct of monetary policy.

As Figure 4 shows, the Taylor rule does a pretty good job of describing the Fed's setting of the federal funds rate under Chairman Greenspan. It also provides a perspective on the Fed's conduct of monetary policy under Chairmen Burns and Volcker. During the Burns period, from 1970 to 1979, the federal funds rate was consistently

---

[8] John B. Taylor, "Discretion Versus Policy Rules in Practice," *Carnegie-Rochester Conference Series on Public Policy* 39 (1993): 195–214. A more intuitive discussion with a historical perspective can be found in John B. Taylor, "A Historical Analysis of Monetary Policy Rules," in *Monetary Policy Rules*, ed. John B. Taylor (Chicago: University of Chicago Press, 1999), pp. 319–341.

[9] There are however subtle differences between the two concepts as is discussed in Arturo Estrella and Frederic S. Mishkin, "The Role of NAIRU in Monetary Policy: Implications of Uncertainty and Model Selection," in *Monetary Policy Rules*, ed. John Taylor (Chicago: University of Chicago Press, 1999): 405–430.

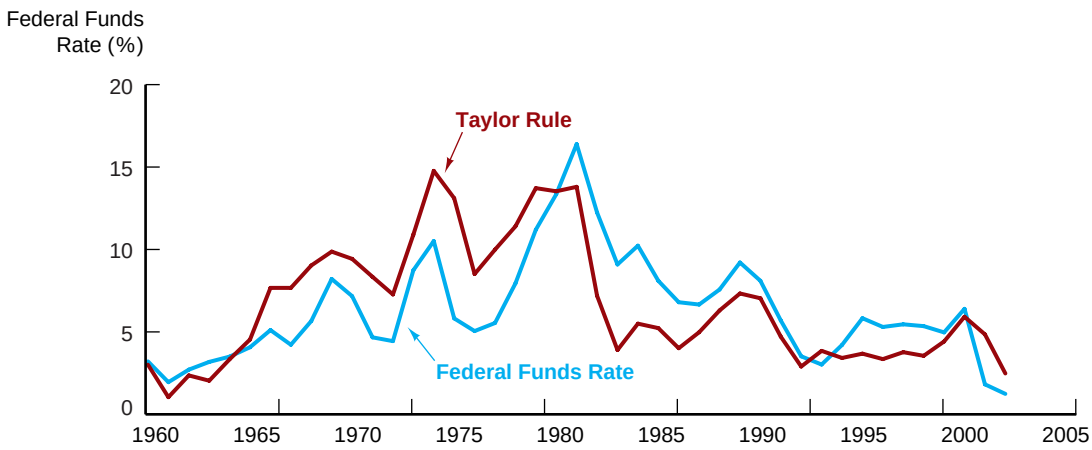


FIGURE 4 The Taylor Rule for the Federal Funds Rate, 1970–2002

Source: Federal Reserve: www.federalreserve.gov/releases and author's calculations.

lower than that indicated by the Taylor rule. This fact helps explain why inflation rose during this period. During the Volcker period, when the Fed was trying to bring inflation down quickly, the funds rate was generally higher than that recommended by the Taylor rule. The closer correspondence between the actual funds rate and the Taylor rule recommendation during the Greenspan era may help explain why the Fed's performance has been so successful in recent years.

Box 4

Fed Watching

As we have seen, the most important player in the determination of the U.S. money supply and interest rates is the Federal Reserve. When the Fed wants to inject reserves into the system, it conducts open market purchases of bonds, which cause bond prices to increase and their interest rates to fall, at least in the short term. If the Fed withdraws reserves from the system, it sells bonds, thereby depressing their price and raising their interest rates. From a longer-run perspective, if the Fed pursues an expansionary monetary policy with high money growth, inflation will rise and, as we saw in Chapter 5, interest rates will rise as well. Contractionary monetary policy is likely to lower inflation in the long run and lead to lower interest rates.

Knowing what actions the Fed might be taking can thus help investors and financial institutions to predict the future course of interest rates with greater accuracy. Because, as we have seen, changes in interest rates have a major impact on investors and financial institutions' profits, they are particularly interested in scrutinizing the Fed's behavior. To assist in this task, financial institutions hire so-called *Fed watchers*, experts on Federal Reserve behavior who may have worked in the Federal Reserve System and so have an insider's view of Federal Reserve operations. A Fed watcher who can accurately predict the course of monetary policy is a very valuable commodity, and successful Fed watchers therefore often earn very high salaries, well into the six-figure range and sometimes even higher.

Summary

1. The six basic goals of monetary policy are high employment, economic growth, price stability, interest-rate stability, stability of financial markets, and stability in foreign exchange markets.
2. By using intermediate and operating targets, a central bank like the Fed can more quickly judge whether its policies are on the right track and make midcourse corrections, rather than waiting to see the final outcome of its policies on such goals as employment and the price level. The Fed's policy tools directly affect its operating targets, which in turn affect the intermediate targets, which in turn affect the goals.
3. Because interest-rate and monetary aggregate targets are incompatible, a central bank must choose between them on the basis of three criteria: measurability, controllability, and the ability to affect goal variables predictably. Unfortunately, these criteria do not establish an overwhelming case for one set of targets over another.
4. The historical record of the Fed's conduct of monetary policy reveals that the Fed has switched its operating targets many times, returning to a federal funds rate target in recent years.
5. The Taylor rule indicates that the federal funds rate should be set equal to the inflation rate plus an "equilibrium" real funds rate plus a weighted average of two gaps: (1) an inflation gap, current inflation minus a target rate, and (2) an output gap, the percentage deviation of real GDP from an estimate of its potential full employment level. The output gap in the Taylor rule could represent an indicator of future inflation as stipulated in Phillips curve theory. However, this theory is controversial, because high output relative to potential as measured by low unemployment has not seemed to produce higher inflation in recent years.

Key Terms

instrument target, p. 415
intermediate targets, p. 414
international policy coordination, p. 428
natural rate of unemployment, p. 412
nonaccelerating inflation rate of unemployment (NAIRU), p. 429
operating target, p. 415
Phillips curve theory, p. 429
real bills doctrine, p. 420
Taylor rule, p. 428

Questions and Problems

Questions marked with an asterisk are answered at the end of the book in an appendix, "Answers to Selected Questions and Problems."

*1. "Unemployment is a bad thing, and the government should make every effort to eliminate it." Do you agree or disagree? Explain your answer.

2. Classify each of the following as either an operating target or an intermediate target, and explain why.
   a. The three-month Treasury bill rate
   b. The monetary base
   c. M2

*3. "If the demand for money did not fluctuate, the Fed could pursue both a money supply target and an interest-rate target at the same time." Is this statement true, false, or uncertain? Explain your answer.

4. If the Fed has an interest-rate target, why will an increase in money demand lead to a rise in the money supply?

*5. What procedures can the Fed use to control the three-month Treasury bill rate? Why does control of this interest rate imply that the Fed will lose control of the money supply?

6. Compare the monetary base to M2 on the grounds of controllability and measurability. Which do you prefer as an intermediate target? Why?

*7. "Interest rates can be measured more accurately and more quickly than the money supply. Hence an interest rate is preferred over the money supply as an intermediate target." Do you agree or disagree? Explain your answer.

8. Explain why the rise in the discount rate in 1920 led to a sharp decline in the money supply.

*9. How did the Fed's failure to perform its role as the lender of last resort contribute to the decline of the money supply in the 1930–1933 period?

10. Excess reserves are frequently called *idle reserves*, suggesting that they are not useful. Does the episode of the rise in reserve requirements in 1936–1937 bear out this view?

*11. "When the economy enters a recession, an interest-rate target will lead to a slower rate of growth for the money supply." Explain why this statement is true. What does it say about the use of interest rates as targets?

12. "The failure of the Fed to control the money supply in the 1970s and 1980s suggests that the Fed is not able to control the money supply." Do you agree or disagree? Explain your answer.

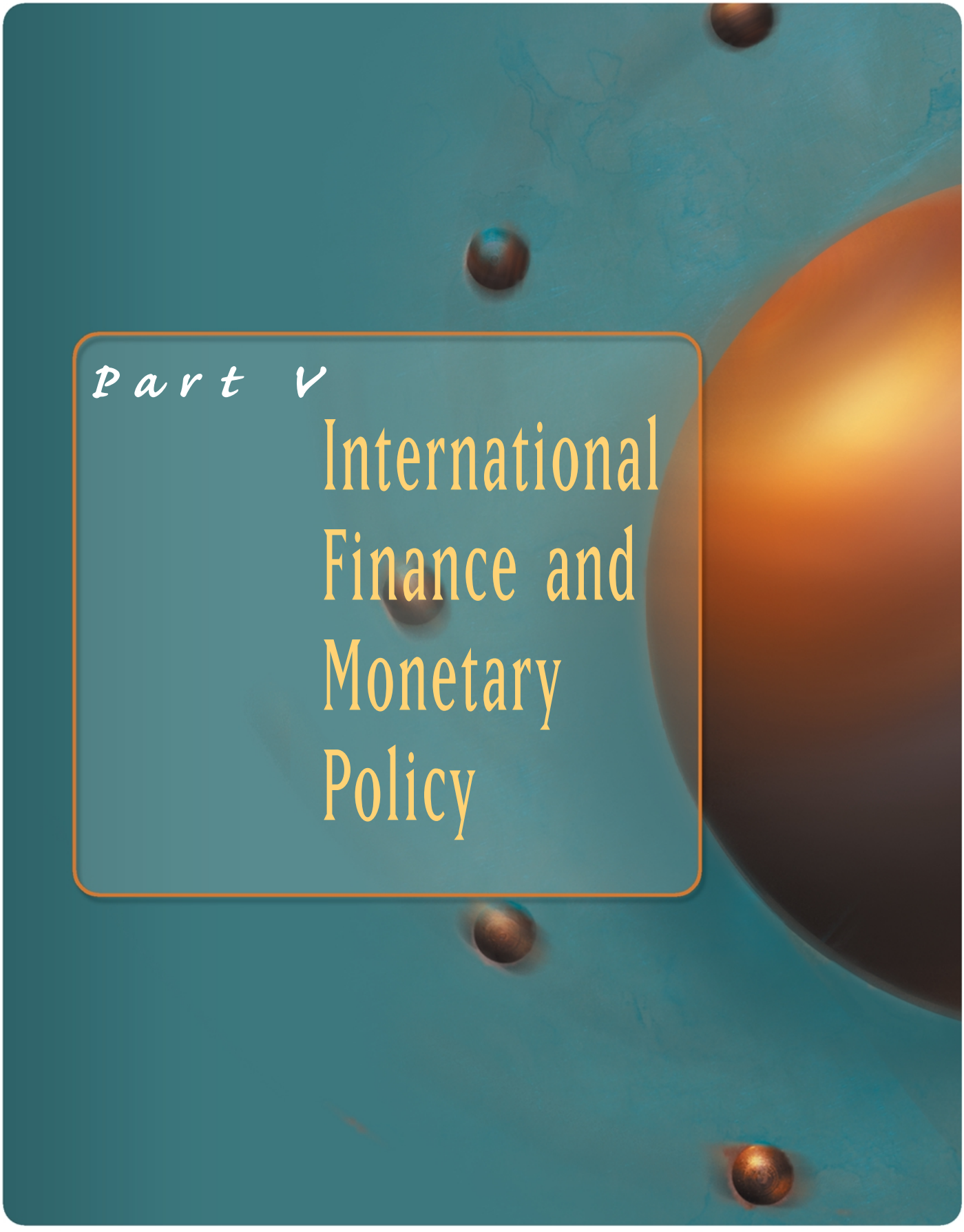
*13. Which is more likely to produce smaller fluctuations in the federal funds rate, a nonborrowed reserves target or a borrowed reserves target? Why?

14. How can bank behavior and the Fed's behavior cause money supply growth to be procyclical (rising in booms and falling in recessions)?

*15. Why might the Fed say that it wants to control the money supply but in reality not be serious about doing so?

Web Exercises

1. The Federal Open Market Committee (FOMC) meets about every six weeks to assess the state of the economy and to decide what actions the central bank should take. The minutes of this meeting are released after the next scheduled meeting; however, a brief press release is made available immediately. Find the schedule of minutes and press releases at www.federalreserve.gov/fomc/.
   a. When was the last scheduled meeting of the FOMC? When is the next meeting?
   b. Review the press release from the last meeting. What did the committee decide to do about short-term interest rates?
   c. Review the most recently published meeting minutes. What areas of the economy seemed to be of most concern to the committee members?
2. It is possible to access other central bank web sites to learn about their structure. One example is the European Central bank. Go to www.ecb.int/index.html. On the ECB home page, locate the link to the current exchange rate between the euro and the dollar. It was initially set at 1 to 1. What is it now?


Part V
International Finance and Monetary Policy

Chapter 19

The Foreign Exchange Market

PREVIEW

In the mid-1980s, American businesses became less competitive with their foreign counterparts; subsequently, in the 1990s and 2000s, their competitiveness increased. Did this swing in competitiveness occur primarily because American management fell down on the job in the 1980s and then got its act together afterwards? Not really. American business became less competitive in the 1980s because American dollars become worth more in terms of foreign currencies, making American goods more expensive relative to foreign goods. By the 1990s and 2000s, the value of the U.S. dollar had fallen appreciably from its highs in the mid-1980s, making American goods cheaper and American businesses more competitive.

The price of one currency in terms of another is called the **exchange rate**. It affects the economy and our daily lives, because when the U.S. dollar becomes more valuable relative to foreign currencies, foreign goods become cheaper for Americans and American goods become more expensive for foreigners. When the U.S. dollar falls in value, foreign goods become more expensive for Americans and American goods become cheaper for foreigners. We begin our study of international finance by examining the **foreign exchange market**, the financial market where exchange rates are determined.

As you can see in Figure 1, exchange rates are highly volatile. What factors explain the rise and fall of exchange rates? Why are exchange rates so volatile from day to day?

To answer these questions, we develop a modern view of exchange rate determination that explains recent behavior in the foreign exchange market.

Foreign Exchange Market

www.ny.frb.org/Pihome/addpub/usfxm

Get detailed information about the foreign exchange market in the United States.

Most countries of the world have their own currencies: The United States has its dollar; the European Monetary Union, the euro; Brazil, its real; and India, its rupee. Trade between countries involves the mutual exchange of different currencies (or, more usually, bank deposits denominated in different currencies). When an American firm buys foreign goods, services, or financial assets, for example, U.S. dollars (typically, bank deposits denominated in U.S. dollars) must be exchanged for foreign currency (bank deposits denominated in the foreign currency).

The trading of currency and bank deposits denominated in particular currencies takes place in the foreign exchange market. Transactions conducted in the foreign

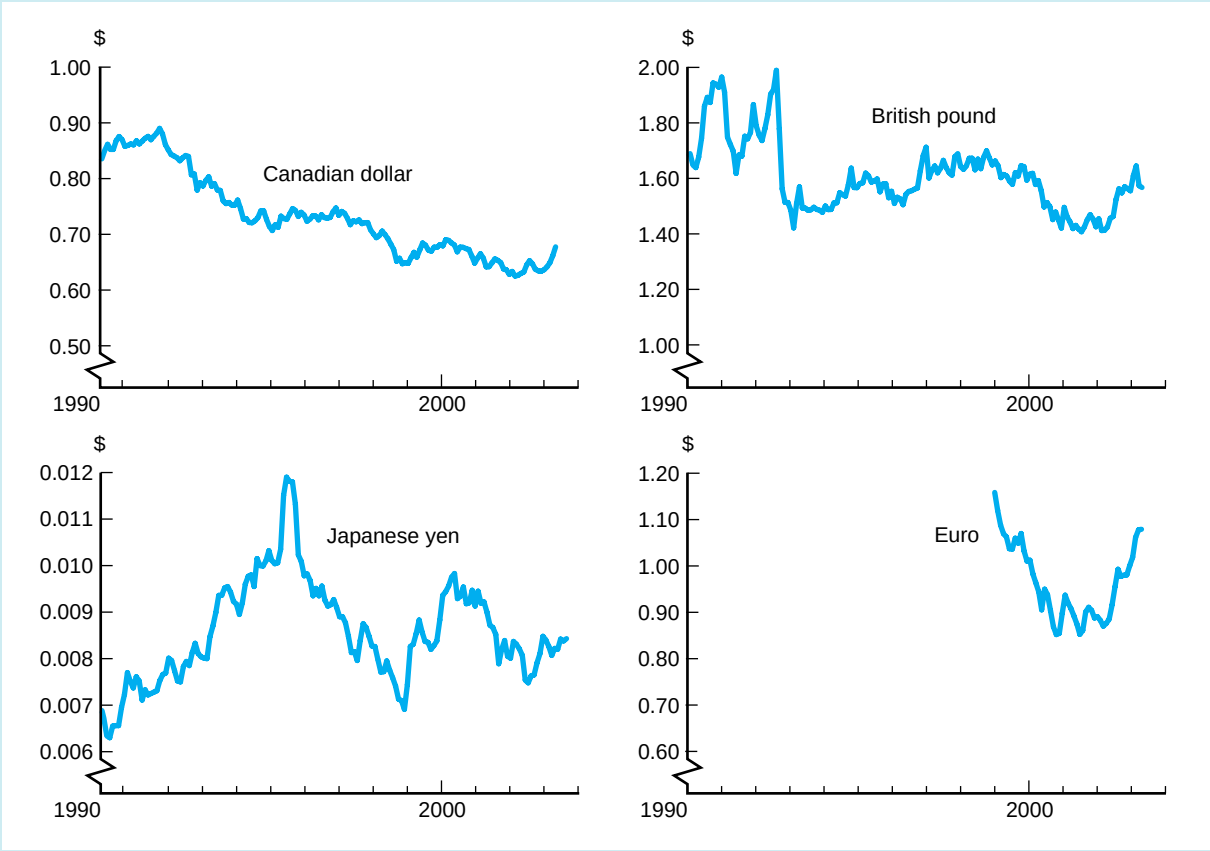


FIGURE 1 Exchange Rates, 1990–2002

Dollar prices of selected currencies. Note that a rise in these plots indicates a strengthening of the currency (weakening of the dollar).

Source: Federal Reserve: www.federalreserve.gov/releases/h10/hist.

exchange market determine the rates at which currencies are exchanged, which in turn determine the cost of purchasing foreign goods and financial assets.

What Are Foreign Exchange Rates?

http://quotes.ino.com/chart/

Go to this web site and click on "Foreign Exchange" to get market rates and time charts for the exchange rate of the U.S. dollar to major world currencies.

There are two kinds of exchange rate transactions. The predominant ones, called **spot transactions**, involve the immediate (two-day) exchange of bank deposits. **Forward transactions** involve the exchange of bank deposits at some specified future date. The **spot exchange rate** is the exchange rate for the spot transaction, and the **forward exchange rate** is the exchange rate for the forward transaction.

When a currency increases in value, it experiences **appreciation**; when it falls in value and is worth fewer U.S. dollars, it undergoes **depreciation**. At the beginning of 1999, for example, the euro was valued at 1.18 dollars, and as indicated in the "Following the Financial News" box, on February 5, 2003, it was valued at 1.08 dollars. The euro *depreciated* by 8%: (1.08 − 1.18)/1.18 = −0.08 = −8%. Equivalently, we could say that the U.S. dollar, which went from a value of 0.85 euros per dollar at the

Following the Financial News

Foreign Exchange Rates

Foreign exchange rates are published daily and appear in the "Currency Trading" column of the *Wall Street Journal*. The entries from one such column, shown here, are explained in the text.

The first entry for the euro lists the exchange rate for the spot transaction (the spot exchange rate) on February 5, 2003, and is quoted in two ways: $1.0795 per euro and 0.9264 euros per dollar. Americans generally regard the exchange rate with the euro as $1.0795 per euro, while Europeans think of it as 0.9264 euros per dollar. The three entries immediately below the spot exchange rates for some currencies give the rates for forward transactions (the forward exchange rates) that will take place one month, three months, and six months in the future.

CURRENCY TRADING

Wednesday, February 5, 2003

EXCHANGE RATES

The foreign exchange mid-range rates below apply to trading among banks in amounts of $1 million and more, as quoted at 4 p.m. Eastern time by Reuters and other sources. Retail transactions provide fewer units of foreign currency per dollar.

Country	U.S. $ Equivalent Wed	U.S. $ Equivalent Tue	Currency per U.S. $ Wed	Currency per U.S. $ Tue
Argentina (Peso)-y	.3160	.3160	3.1646	3.1646
Australia (Dollar)	.5901	.5915	1.6946	1.6906
Bahrain (Dinar)	2.6522	2.6523	.3770	.3770
Brazil (Real)	.2784	.2798	3.5920	3.5740
Canada (Dollar)	.6574	.6602	1.5211	1.5147
1-month forward	.6566	.6595	1.5230	1.5163
3-months forward	.6548	.6576	1.5272	1.5207
6-months forward	.6517	.6544	1.5344	1.5281
Chile (Peso)	.001348	.001346	741.84	742.94
China (Renminbi)	.1208	.1208	8.2781	8.2781
Colombia (Peso)	.0003372	.0003378	2965.60	2960.33
Czech. Rep. (Koruna)				
Commercial rate	.03398	.03438	29.429	29.087
Denmark (Krone)	.1453	.1463	6.8823	6.8353
Ecuador (US Dollar)	1.0000	1.0000	1.0000	1.0000
Hong Kong (Dollar)	.1282	.1282	7.8003	7.8003
Hungary (Forint)	.004406	.004454	226.96	224.52
India (Rupee)	.02099	.02094	47.642	47.756
Indonesia (Rupiah)	.0001128	.0001127	8865	8873
Israel (Shekel)	.2050	.2049	4.8780	4.8804
Japan (Yen)	.008336	.008353	119.96	119.72
1-month forward	.008344	.008362	119.85	119.59
3-months forward	.008363	.008381	119.57	119.32
6-months forward	.008391	.008408	119.18	118.93
Jordan (Dinar)	1.4094	1.4085	.7095	.7100
Kuwait (Dinar)	3.3479	3.3523	.2987	.2983
Lebanon (Pound)	.0006634	.0006634	1507.39	1507.39
Malaysia (Ringgit)-b	.2632	.2632	3.7994	3.7994
Malta (Lira)	2.5690	2.5861	.3893	.3867
Mexico (Peso)				
Floating rate	.0920	.0913	10.8648	10.9481
New Zealand (Dollar)	.5494	.5496	1.8202	1.8195
Norway (Krone)	.1434	.1448	6.9735	6.9061
Pakistan (Rupee)	.01719	.01723	58.173	58.038
Peru (new Sol)	.2866	.2863	3.4892	3.4928
Philippines (Peso)	.01852	.01853	53.996	53.967
Poland (Zloty)	.2606	.2622	3.8373	3.8139
Russia (Ruble)-a	0.3142	0.3142	31.827	31.827
Saudi Arabia	.2667	.2667	3.7495	3.7495
Singapore (Dollar)	.5742	.5755	1.7416	1.7376
Slovak Rep. (Koruna)	.02579	.02607	38.775	38.358
South Africa (Rand)	.1192	.1202	8.3893	8.3195
South Korea (Won)	.0008516	.0008529	1174.26	1172.47
Sweden (Krona)	.1169	.1177	8.5543	8.4962
Switzerland (Franc)	.7358	.7424	1.3591	1.3470
1-month forward	.7362	.7428	1.3583	1.3463
3-months forward	.7371	.7437	1.3567	1.3446
6-months forward	.7386	.7451	1.3539	1.3421
Taiwan (Dollar)	.02881	.02881	34.710	34.710
Thailand (Baht)	.02338	.02342	42.772	42.699
Turkey (Lira)	.00000061	.00000061	1639344	1639344
U.K. (Pound)	1.6423	1.6485	.6089	.6066
1-month forward	1.6391	1.6452	.6101	.6078
3-months forward	1.6322	1.6382	.6127	.6104
6-months forward	1.6221	1.6283	.6165	.6141
United Arab (Dirham)	.2723	.2723	3.6724	3.6724
Uruguay (Peso)				
Financial	.03500	.03550	28.571	28.169
Venezuela (Bolivar)	.000520	.000520	1923.08	1923.08
SDR	1.3741	1.3697	.7277	.7301
Euro	1.0795	1.0883	.9264	.9189

Special Drawing Rights (SDR) are based on exchange rates for the U.S., British, and Japanese currencies. Source: International Monetary Fund.
a-Russian Central Bank rate. b-Government rate. y-Floating rate.

Source: Wall Street Journal, Thursday, February 6, 2003, p. C12.

beginning of 1999 to a value of 0.93 euros per dollar on February 5, 2003, *appreciated* by 9%: (0.93 − 0.85)/0.85 = 0.09 = 9%.

Why Are Exchange Rates Important?

Exchange rates are important because they affect the relative price of domestic and foreign goods. The dollar price of French goods to an American is determined by the interaction of two factors: the price of French goods in euros and the euro/dollar exchange rate.

Suppose that Wanda the Winetaster, an American, decides to buy a bottle of 1961 (a very good year) Château Lafite Rothschild to complete her wine cellar. If the price of the wine in France is 1,000 euros and the exchange rate is $1.08 to the euro, the wine will cost Wanda $1,080 (= 1,000 euros × $1.08/euro). Now suppose that Wanda delays her purchase by two months, at which time the euro has appreciated to $1.20 per euro. If the domestic price of the bottle of Lafite Rothschild remains 1,000 euros, its dollar cost will have risen from $1,080 to $1,200.

The same currency appreciation, however, makes the price of foreign goods in that country less expensive. At an exchange rate of $1.08 per euro, a Compaq computer priced at $2,000 costs Pierre the Programmer 1,852 euros; if the exchange rate increases to $1.20 per euro, the computer will cost only 1,667 euros.

A depreciation of the euro lowers the cost of French goods in America but raises the cost of American goods in France. If the euro drops in value to $0.90, Wanda's bottle of Lafite Rothschild will cost her only $900 instead of $1,080, and the Compaq computer will cost Pierre 2,222 euros rather than 1,852.

Such reasoning leads to the following conclusion: ***When a country's currency appreciates (rises in value relative to other currencies), the country's goods abroad become more expensive and foreign goods in that country become cheaper (holding domestic prices constant in the two countries). Conversely, when a country's currency depreciates, its goods abroad become cheaper and foreign goods in that country become more expensive.***

Appreciation of a currency can make it harder for domestic manufacturers to sell their goods abroad and can increase competition at home from foreign goods, because they cost less. From 1980 to early 1985, the appreciating dollar hurt U.S. industries. For instance, the U.S. steel industry was hurt not just because sales abroad of the more expensive American steel declined, but also because sales of relatively cheap foreign steel in the United States increased. Although appreciation of the U.S. dollar hurt some domestic businesses, American consumers benefited because foreign goods were less expensive. Japanese videocassette recorders and cameras and the cost of vacationing in Europe fell in price as a result of the strong dollar.

How Is Foreign Exchange Traded?

You cannot go to a centralized location to watch exchange rates being determined; currencies are not traded on exchanges such as the New York Stock Exchange. Instead, the foreign exchange market is organized as an over-the-counter market in which several hundred dealers (mostly banks) stand ready to buy and sell deposits denominated in foreign currencies. Because these dealers are in constant telephone and computer contact, the market is very competitive; in effect, it functions no differently from a centralized market.

An important point to note is that while banks, companies, and governments talk about buying and selling currencies in foreign exchange markets, they do not take a fistful of dollar bills and sell them for British pound notes. Rather, most trades involve the buying and selling of bank deposits denominated in different currencies. So when

we say that a bank is buying dollars in the foreign exchange market, what we actually mean is that the bank is buying *deposits denominated in dollars*. The volume in this market is colossal, exceeding $1 trillion per day.

Trades in the foreign exchange market consist of transactions in excess of $1 million. The market that determines the exchange rates in the "Following the Financial News" box is not where one would buy foreign currency for a trip abroad. Instead, we buy foreign currency in the retail market from dealers such as American Express or from banks. Because retail prices are higher than wholesale, when we buy foreign exchange, we obtain fewer units of foreign currency per dollar than exchange rates in the box indicate.

Exchange Rates in the Long Run

Like the price of any good or asset in a free market, exchange rates are determined by the interaction of supply and demand. To simplify our analysis of exchange rates in a free market, we divide it into two parts. First, we examine how exchange rates are determined in the long run; then we use our knowledge of the long-run determinants of the exchange rate to help us understand how they are determined in the short run.

Law of One Price

The starting point for understanding how exchange rates are determined is a simple idea called the **law of one price**: If two countries produce an identical good, and transportation costs and trade barriers are very low, the price of the good should be the same throughout the world no matter which country produces it. Suppose that American steel costs $100 per ton and identical Japanese steel costs 10,000 yen per ton. For the law of one price to hold, the exchange rate between the yen and the dollar must be 100 yen per dollar ($0.01 per yen) so that one ton of American steel sells for 10,000 yen in Japan (the price of Japanese steel) and one ton of Japanese steel sells for $100 in the United States (the price of U.S. steel). If the exchange rate were 200 yen to the dollar, Japanese steel would sell for $50 per ton in the United States or half the price of American steel, and American steel would sell for 20,000 yen per ton in Japan, twice the price of Japanese steel. Because American steel would be more expensive than Japanese steel in both countries and is identical to Japanese steel, the demand for American steel would go to zero. Given a fixed dollar price for American steel, the resulting excess supply of American steel will be eliminated only if the exchange rate falls to 100 yen per dollar, making the price of American steel and Japanese steel the same in both countries.

Theory of Purchasing Power Parity

One of the most prominent theories of how exchange rates are determined is the **theory of purchasing power parity (PPP)**. It states that exchange rates between any two currencies will adjust to reflect changes in the price levels of the two countries. The theory of PPP is simply an application of the law of one price to national price levels rather than to individual prices. Suppose that the yen price of Japanese steel rises 10% (to 11,000 yen) relative to the dollar price of American steel (unchanged at $100). For the law of one price to hold, the exchange rate must rise to 110 yen to the dollar, a 10% appreciation of the dollar. Applying the law of one price to the price levels in the two countries produces the theory of purchasing power parity, which maintains that if the Japanese price level rises 10% relative to the U.S. price level, the dollar will appreciate by 10%.

www.oecd.org/EN/home/0,,EN-home-513-15-no-no-no-0,00.html

The purchasing power parities home page includes the PPP program overview, statistics, research, publications, and OECD meetings on PPP.

As our U.S./Japanese example demonstrates, the theory of PPP suggests that if one country's price level rises relative to another's, its currency should depreciate (the other country's currency should appreciate). As you can see in Figure 2, this prediction is borne out in the long run. From 1973 to the end of 2002, the British price level rose 99% relative to the U.S. price level, and as the theory of PPP predicts, the dollar appreciated against the pound; though by 73%, an amount smaller than the 99% increase predicted by PPP.

Yet, as the same figure indicates, PPP theory often has little predictive power in the short run. From early 1985 to the end of 1987, for example, the British price level rose relative to that of the United States. Instead of appreciating, as PPP theory predicts, the U.S. dollar actually depreciated by 40% against the pound. So even though PPP theory provides some guidance to the long-run movement of exchange rates, it is not perfect and in the short run is a particularly poor predictor. What explains PPP theory's failure to predict well?

Why the Theory of Purchasing Power Parity Cannot Fully Explain Exchange Rates

The PPP conclusion that exchange rates are determined solely by changes in relative price levels rests on the assumption that all goods are identical in both countries and that transportation costs and trade barriers are very low. When this assumption is true, the law of one price states that the relative prices of all these goods (that is, the relative price level between the two countries) will determine the exchange rate. The assumption that goods are identical may not be too unreasonable for American and Japanese steel, but is it a reasonable assumption for American and Japanese cars? Is a Toyota the equivalent of a Chevrolet?

Because Toyotas and Chevys are obviously not identical, their prices do not have to be equal. Toyotas can be more expensive relative to Chevys and both Americans and Japanese will still purchase Toyotas. Because the law of one price does not hold for all goods, a rise in the price of Toyotas relative to Chevys will not necessarily mean

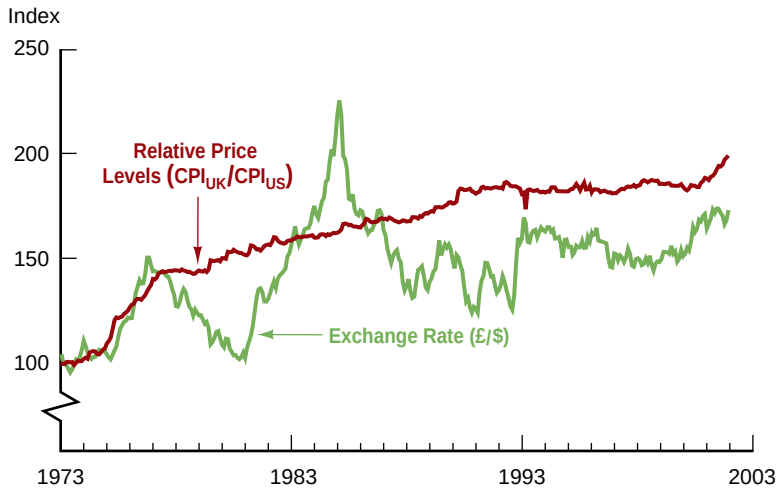


FIGURE 2 Purchasing Power Parity, United States/United Kingdom, 1973–2002
(Index: March 1973 = 100.)

Source: www.statistics.gov.uk/statbase/tsdataset2.asp.

that the yen must depreciate by the amount of the relative price increase of Toyotas over Chevys.

PPP theory furthermore does not take into account that many goods and services (whose prices are included in a measure of a country's price level) are not traded across borders. Housing, land, and services such as restaurant meals, haircuts, and golf lessons are not traded goods. So even though the prices of these items might rise and lead to a higher price level relative to another country's, there would be little direct effect on the exchange rate.

Factors That Affect Exchange Rates in the Long Run

Our analysis indicates that in the long run, four major factors affect the exchange rate: relative price levels, tariffs and quotas, preferences for domestic versus foreign goods, and productivity. We examine how each of these factors affects the exchange rate while holding the others constant.

The basic reasoning proceeds along the following lines: Anything that increases the demand for domestic goods relative to foreign goods tends to appreciate the domestic currency because domestic goods will continue to sell well even when the value of the domestic currency is higher. Similarly, anything that increases the demand for foreign goods relative to domestic goods tends to depreciate the domestic currency because domestic goods will continue to sell well only if the value of the domestic currency is lower.

Relative Price Levels. In line with PPP theory, when prices of American goods rise (holding prices of foreign goods constant), the demand for American goods falls and the dollar tends to depreciate so that American goods can still sell well. By contrast, if prices of Japanese goods rise so that the relative prices of American goods fall, the demand for American goods increases, and the dollar tends to appreciate, because American goods will continue to sell well even with a higher value of the domestic currency. ***In the long run, a rise in a country's price level (relative to the foreign price level) causes its currency to depreciate, and a fall in the country's relative price level causes its currency to appreciate.***

Trade Barriers. Barriers to free trade such as **tariffs** (taxes on imported goods) and **quotas** (restrictions on the quantity of foreign goods that can be imported) can affect the exchange rate. Suppose that the United States increases its tariff or puts a lower quota on Japanese steel. These increases in trade barriers increase the demand for American steel, and the dollar tends to appreciate because American steel will still sell well even with a higher value of the dollar. ***Increasing trade barriers cause a country's currency to appreciate in the long run.***

Preferences for Domestic Versus Foreign Goods. If the Japanese develop an appetite for American goods—say, for Florida oranges and American movies—the increased demand for American goods (exports) tends to appreciate the dollar, because the American goods will continue to sell well even at a higher value for the dollar. Likewise, if Americans decide that they prefer Japanese cars to American cars, the increased demand for Japanese goods (imports) tends to depreciate the dollar. ***Increased demand for a country's exports causes its currency to appreciate in the long run; conversely, increased demand for imports causes the domestic currency to depreciate.***

Productivity. If one country becomes more productive than other countries, businesses in that country can lower the prices of domestic goods relative to foreign goods and still earn a profit. As a result, the demand for domestic goods rises, and the domestic currency tends to appreciate. If, however, its productivity lags behind that of other countries, its goods become relatively more expensive, and the currency tends to depreciate. ***In the long run, as a country becomes more productive relative to other countries, its currency appreciates.***[1]

Study Guide

The trick to figuring out what long-run effect a factor has on the exchange rate is to remember the following: ***If a factor increases the demand for domestic goods relative to foreign goods, the domestic currency will appreciate, and if a factor decreases the relative demand for domestic goods, the domestic currency will depreciate.*** See how this works by explaining what happens to the exchange rate when any of the factors in Table 1 decreases rather than increases.

Our long-run theory of exchange rate behavior is summarized in Table 1. We use the convention that the exchange rate E is quoted so that an appreciation of the currency corresponds to a rise in the exchange rate. In the case of the United States, this means that we are quoting the exchange rate as units of foreign currency per dollar (say, yen per dollar).[2]

SUMMARY Table 1 Factors That Affect Exchange Rates in the Long Run

Factor	Change in Factor	Response of the Exchange Rate, E*
Domestic price level†	↑	↓
Trade barriers†	↑	↑
Import demand	↑	↓
Export demand	↑	↑
Productivity†	↑	↑

*Units of foreign currency per dollar: ↑ indicates domestic currency appreciation; ↓, depreciation.
†Relative to other countries.
Note: Only increases (↑) in the factors are shown; the effects of decreases in the variables on the exchange rate are the opposite of those indicated in the "Response" column.

[1] A country might be so small that a change in productivity or the preferences for domestic or foreign goods would have no effect on prices of these goods relative to foreign goods. In this case, changes in productivity or changes in preferences for domestic or foreign goods affect the country's income but will not necessarily affect the value of the currency. In our analysis, we are assuming that these factors can affect relative prices and consequently the exchange rate.

[2] Exchange rates can be quoted either as units of foreign currency per domestic currency or alternatively as units of domestic currency per foreign currency. In professional writing, many economists quote exchange rates as units of domestic currency per foreign currency so that an appreciation of the domestic currency is portrayed as a fall in the exchange rate. The opposite convention is used in the text here, because it is more intuitive to think of an appreciation of the domestic currency as a rise in the exchange rate.

Exchange Rates in the Short Run

www.federalreserve.gov/releases/

The Federal Reserve reports current and historical exchange rates for many countries.

We have developed a theory of the long-run behavior of exchange rates. However, if we are to understand why exchange rates exhibit such large changes (sometimes several percent) from day to day, we must develop a theory of how current exchange rates (spot exchange rates) are determined in the short run.

The key to understanding the short-run behavior of exchange rates is to recognize that an exchange rate is the price of domestic bank deposits (those denominated in the domestic currency) in terms of foreign bank deposits (those denominated in the foreign currency). Because the exchange rate is the price of one asset in terms of another, the natural way to investigate the short-run determination of exchange rates is through an asset market approach that relies heavily on the theory of asset demand developed in Chapter 5. As you will see, however, the long-run determinants of the exchange rate we have just outlined also play an important role in the short-run asset market approach.[3]

Earlier approaches to exchange rate determination emphasized the role of import and export demand. The more modern asset market approach used here does not emphasize the flows of purchases of exports and imports over short periods, because these transactions are quite small relative to the amount of domestic and foreign bank deposits at any given time. For example, foreign exchange transactions in the United States each year are well over 25 times greater than the amount of U.S. exports and imports. Thus over short periods such as a year, decisions to hold domestic or foreign assets play a much greater role in exchange rate determination than the demand for exports and imports does.

Comparing Expected Returns on Domestic and Foreign Deposits

In this analysis, we treat the United States as the home country, so as an example, domestic bank deposits are denominated in dollars. For simplicity, we use euros to stand for any foreign country's currency, so foreign bank deposits are denominated in euros. The theory of asset demand suggests that the most important factor affecting the demand for domestic (dollar) deposits and foreign (euro) deposits is the expected return on these assets relative to each other. When Americans or foreigners expect the return on dollar deposits to be high relative to the return on foreign deposits, there is a higher demand for dollar deposits and a correspondingly lower demand for euro deposits. To understand how the demands for dollar and foreign deposits change, we need to compare the expected returns on dollar deposits and foreign deposits.

To illustrate further, suppose that dollar deposits have an interest rate (expected return payable in dollars) of i^D, and foreign bank deposits have an interest rate (expected return payable in the foreign currency, euros) of i^F. To compare the expected returns on dollar deposits and foreign deposits, investors must convert the returns into the currency unit they use.

First let us examine how François the Foreigner compares the returns on dollar deposits and foreign deposits denominated in his currency, the euro. When he considers the expected return on dollar deposits in terms of euros, he recognizes that it does not equal i^D; instead, the expected return must be adjusted for any expected appreciation or depreciation of the dollar. If the dollar were expected to appreciate by 7%, for example, the expected return on dollar deposits in terms of euros would be

[3]For a further description of the modern asset market approach to exchange rate determination that we use here, see Paul Krugman and Maurice Obstfeld, *International Economics*, 6th ed. (Reading, Mass.: Addison Wesley Longman, 2003).

7% higher because the dollar has become worth 7% more in terms of euros. Thus if the interest rate on dollar deposits is 10%, with an expected appreciation of the dollar of 7%, the expected return on dollar deposits in terms of euros is 17%: the 10% interest rate plus the 7% expected appreciation of the dollar. Conversely, if the dollar were expected to depreciate by 7% over the year, the expected return on dollar deposits in terms of euros would be only 3%: the 10% interest rate minus the 7% expected depreciation of the dollar.

Writing the currency exchange rate (the spot exchange rate) as E_t and the expected exchange rate for the next period as E^e_{t+1}, we can write the expected rate of appreciation of the dollar as $(E^e_{t+1} - E_t)/E_t$. Our reasoning indicates that the expected return on dollar deposits R^D in terms of foreign currency can be written as the sum of the interest rate on dollar deposits plus the expected appreciation of the dollar:[4]

$$R^D \text{ in terms of euros} = i^D + \frac{E^e_{t+1} - E_t}{E_t}$$

However, François's expected return on foreign deposits R^F in terms of euros is just i^F. Thus in terms of euros, the relative expected return on dollar deposits (that is, the difference between the expected return on dollar deposits and euro deposits) is calculated by subtracting i^F from the expression just given to yield

$$\text{Relative } R^D = i^D - i^F + \frac{E^e_{t+1} - E_t}{E_t} \tag{1}$$

As the relative expected return on dollar deposits increases, foreigners will want to hold more dollar deposits and fewer foreign deposits.

Next let us look at the decision to hold dollar deposits versus euro deposits from Al the American's point of view. Following the same reasoning we used to evaluate the decision for François, we know that the expected return on foreign deposits R^F in terms of dollars is the interest rate on foreign deposits i^F plus the expected appreciation of the foreign currency, equal to minus the expected appreciation of the dollar, $-(E^e_{t+1} - E_t)/E_t$, that is:

$$R^F \text{ in terms of dollars} = i^F - \frac{E^e_{t+1} - E_t}{E_t}$$

---

[4]This expression is actually an approximation of the expected return in terms of euros, which can be more precisely calculated by thinking how a foreigner invests in the dollar deposit. Suppose that François decides to put one euro into dollar deposits. First he buys $1/E_t$ of U.S. dollar deposits (recall that E_t, the exchange rate between dollar and euro deposits, is quoted in euros per dollar), and at the end of the period he is paid $(1 + i^D)(1/E_t)$ in dollars. To convert this amount into the number of euros he expects to receive at the end of the period, he multiplies this quantity by E^e_{t+1}. François's expected return on his initial investment of one euro can thus be written as $(1 + i^D)(E^e_{t+1}/E_t)$ minus his initial investment of one euro:

$$(1 + i^D)\left(\frac{E^e_{t+1}}{E_t}\right) - 1$$

which can be rewritten as

$$i^D\left(\frac{E^e_{t+1}}{E_t}\right) + \frac{E^e_{t+1} - E_t}{E_t}$$

which is approximately equal to the expression in the text because E^e_{t+1}/E_t is typically close to 1.

If the interest rate on euro deposits is 5%, for example, and the dollar is expected to appreciate by 4%, then the expected return on euro deposits in terms of dollars is 1%. Al earns the 5% interest rate, but he expects to lose 4% because he expects the euro to be worth 4% less in terms of dollars as a result of the dollar's appreciation.

Al's expected return on the dollar deposits R^D in terms of dollars is just i^D. Hence in terms of dollars, the relative expected return on dollar deposits is calculated by subtracting the expression just given from i^D to obtain:

$$\text{Relative } R^D = i^D - \left(i^F - \frac{E^e_{t+1} - E_t}{E_t}\right) = i^D - i^F + \frac{E^e_{t+1} - E_t}{E_t}$$

This equation is the same as the one describing François's relative expected return on dollar deposits (calculated in terms of euros). The key point here is that the relative expected return on dollar deposits is the same whether it is calculated by François in terms of euros or by Al in terms of dollars. Thus as the relative expected return on dollar deposits increases, both foreigners and domestic residents respond in exactly the same way—both will want to hold more dollar deposits and fewer foreign deposits.

Interest Parity Condition

We currently live in a world in which there is **capital mobility**: Foreigners can easily purchase American assets such as dollar deposits, and Americans can easily purchase foreign assets such as euro deposits. Because foreign bank deposits and American bank deposits have similar risk and liquidity and because there are few impediments to capital mobility, it is reasonable to assume that the deposits are perfect substitutes (that is, equally desirable). When capital is mobile and when bank deposits are perfect substitutes, if the expected return on dollar deposits is above that on foreign deposits, both foreigners and Americans will want to hold only dollar deposits and will be unwilling to hold foreign deposits. Conversely, if the expected return on foreign deposits is higher than on dollar deposits, both foreigners and Americans will not want to hold any dollar deposits and will want to hold only foreign deposits. For existing supplies of both dollar deposits and foreign deposits to be held, it must therefore be true that there is no difference in their expected returns; that is, the relative expected return in Equation 1 must equal zero. This condition can be rewritten as:

$$i^D = i^F - \frac{E^e_{t+1} - E_t}{E_t} \tag{2}$$

This equation is called the **interest parity condition**, and it states that the domestic interest rate equals the foreign interest rate minus the expected appreciation of the domestic currency. Equivalently, this condition can be stated in a more intuitive way: The domestic interest rate equals the foreign interest rate plus the expected appreciation of the foreign currency. If the domestic interest rate is above the foreign interest rate, this means that there is a positive expected appreciation of the foreign currency, which compensates for the lower foreign interest rate. A domestic interest rate of 15% versus a foreign interest rate of 10% means that the expected appreciation of the foreign currency must be 5% (or, equivalently, that the expected depreciation of the dollar must be 5%).

There are several ways to look at the interest parity condition. First, we should recognize that interest parity means simply that the expected returns are the same on both dollar deposits and foreign deposits. To see this, note that the left side of the interest parity condition (Equation 2) is the expected return on dollar deposits, while

the right side is the expected return on foreign deposits, both calculated in terms of a single currency, the U.S. dollar. Given our assumption that domestic and foreign bank deposits are perfect substitutes (equally desirable), the interest parity condition is an equilibrium condition for the foreign exchange market. Only when the exchange rate is such that expected returns on domestic and foreign deposits are equal—that is, when interest parity holds—will the outstanding domestic and foreign deposits be willingly held.

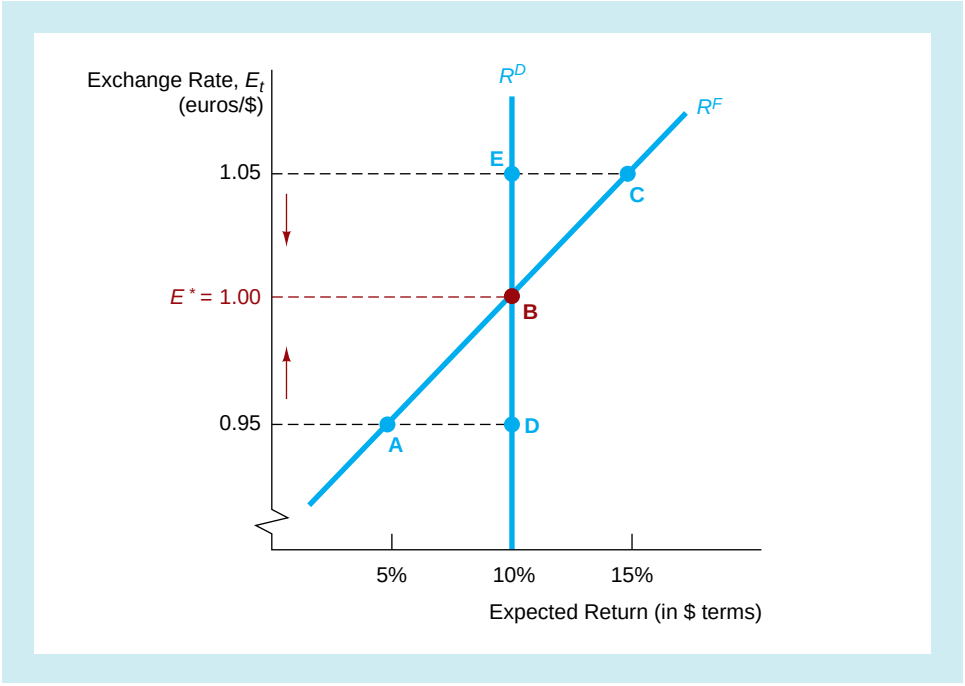
Equilibrium in the Foreign Exchange Market

To see how the interest parity equilibrium condition works in determining the exchange rate, our first step is to examine how the expected returns on euro and dollar deposits change as the current exchange rate changes.

Expected Return on Euro Deposits. As we demonstrated earlier, the expected return in terms of dollars on foreign deposits R^F is the foreign interest rate minus the expected appreciation of the domestic currency: $i^F - (E^e_{t+1} - E_t)/E_t$. Suppose that the foreign interest rate i^F is 10% and that the expected exchange rate next period E^e_{t+1} is 1 euro per dollar. When the current exchange rate E_t is 0.95 euros per dollar, the expected appreciation of the dollar is $(1.00 - 0.95)/0.95 = 0.052 = 5.2\%$, so the expected return on euro deposits R^F in terms of dollars is 4.8% (equal to the 10% foreign interest rate minus the 5.2% dollar appreciation). This expected return when $E_t = 0.95$ euros per dollar is plotted as point A in Figure 3. At a higher current exchange rate of $E_t = 1$ euro per dollar, the expected appreciation of the dollar is zero because E^e_{t+1} also equals 1 euro per dollar. Hence R^F, the expected dollar return on euro deposits, is now just $i^F = 10\%$. This expected return on euro deposits when $E_t = 1$ euro per dollar is plotted as point B. At an even higher exchange rate of $E_t = 1.05$ euros per

FIGURE 3 Equilibrium in the Foreign Exchange Market

Equilibrium in the foreign exchange market occurs at the intersection of the schedules for the expected return on euro deposits R^F and the expected return on dollar deposits R^D at point B. The equilibrium exchange rate is $E^* = 1$ euro per dollar.



dollar, the expected change in the value of the dollar is now −4.8% [= (1.00 – 1.05)/1.05 = −0.048], so the expected dollar return on foreign deposits R^F has now risen to 14.8% [= 10% −(−4.8%)]. This combination of exchange rate and expected return on euro deposits is plotted as point C.

The curve connecting these points is the schedule for the expected return on euro deposits in Figure 3, labeled R^F, and as you can see, it slopes upward; that is, as the exchange rate E_t rises, the expected return on euro deposits rises. The intuition for this upward slope is that because the expected next-period exchange rate is held constant as the current exchange rate rises, there is less expected appreciation of the dollar. Hence a higher current exchange rate means a greater expected appreciation of the foreign currency in the future, which increases the expected return on foreign deposits in terms of dollars.

Expected Return on Dollar Deposits. The expected return on dollar deposits in terms of dollars R^D is always the interest rate on dollar deposits i^D no matter what the exchange rate is. Suppose that the interest rate on dollar deposits is 10%. The expected return on dollar deposits, whether at an exchange rate of 0.95, 1.00, or 1.05 euros per dollar, is always 10% (points D, B, and E) since no foreign-exchange transaction is needed to convert the interest payments into dollars. The line connecting these points is the schedule for the expected return on dollar deposits, labeled R^D in Figure 3.

Equilibrium. The intersection of the schedules for the expected return on dollar deposits R^D and the expected return on euro deposits R^F is where equilibrium occurs in the foreign exchange market; in other words,

$$R^D = R^F$$

At the equilibrium point B where the exchange rate E^* is 1 euro per dollar, the interest parity condition is satisfied because the expected returns on dollar deposits and on euro deposits are equal.

To see that the exchange rate actually heads toward the equilibrium exchange rate E^*, let's see what happens if the exchange rate is 1.05 euros per dollar, a value above the equilibrium exchange rate. As we can see in Figure 3, the expected return on euro deposits at point C is greater than the expected return on dollar deposits at point E. Since dollar and euro deposits are perfect substitutes, people will not want to hold any dollar deposits, and holders of dollar deposits will try to sell them for euro deposits in the foreign exchange market (which is referred to as "selling dollars" and "buying euros"). However, because the expected return on these dollar deposits is below that on euro deposits, no one holding euros will be willing to exchange them for dollar deposits. The resulting excess supply of dollar deposits means that the price of the dollar deposits relative to euro deposits must fall; that is, the exchange rate (amount of euros per dollar) falls as is illustrated by the downward arrow drawn in the figure at the exchange rate of 1.05 euros per dollar. The decline in the exchange rate will continue until point B is reached at the equilibrium exchange rate of 1 euro per dollar, where the expected return on dollar and euro deposits is now equalized.

Now let us look at what happens when the exchange rate is 0.95 euros per dollar, a value below the equilibrium level. Here the expected return on dollar deposits is greater than that on euro deposits. No one will want to hold euro deposits, and everyone will try to sell them to buy dollar deposits ("sell euros" and "buy dollars"), thus driving up the exchange rate as illustrated by the upward arrow. As the exchange

rate rises, there is a higher expected depreciation of the dollar and so a higher expected appreciation of the euro, thereby increasing the expected return on euro deposits. Finally, when the exchange rate has risen to $E^* = 1$ euro per dollar, the expected return on euro deposits has risen enough so that it again equals the expected return on dollar deposits.

Explaining Changes in Exchange Rates

To explain how an exchange rate changes over time, we have to understand the factors that shift the expected-return schedules for domestic (dollar) deposits and foreign (euro) deposits.

Shifts in the Expected-Return Schedule for Foreign Deposits

As we have seen, the expected return on foreign (euro) deposits depends on the foreign interest rate i^F minus the expected appreciation of the dollar $(E^e_{t+1} - E_t)/E_t$. Because a change in the current exchange rate E_t results in a movement along the expected-return schedule for euro deposits, factors that shift this schedule must work through the foreign interest rate i^F and the expected future exchange rate E^e_{t+1}. We examine the effect of changes in these factors on the expected-return schedule for euro deposits R^F, holding everything else constant.

Study Guide

To grasp how the expected-return schedule for euro deposits shifts, just think of yourself as an investor who is considering putting funds into foreign deposits. When a variable changes (i^F, for example), decide whether at a given level of the current exchange rate, holding all other variables constant, you would earn a higher or lower expected return on euro deposits.

Changes in the Foreign Interest Rate. If the interest rate on foreign deposits i^F increases, holding everything else constant, the expected return on these deposits must also increase. Hence at a given exchange rate, the increase in i^F leads to a rightward shift in the expected-return schedule for euro deposits from R^F_1 to R^F_2 in Figure 4. As you can see in the figure, the outcome is a depreciation of the dollar from E_1 to E_2. An alternative way to see this is to recognize that the increase in the expected return on euro deposits at the original equilibrium exchange rate resulting from the rise in i^F means that people will want to buy euros and sell dollars, so the value of the dollar must fall. Our analysis thus generates the following conclusion: ***An increase in the foreign interest rate i^F shifts the R^F schedule to the right and causes the domestic currency to depreciate ($E\downarrow$).***

Conversely, if i^F falls, the expected return on euro deposits falls, the R^F schedule shifts to the left, and the exchange rate rises. This yields the following conclusion: ***A decrease in i^F shifts the R^F schedule to the left and causes the domestic currency to appreciate ($E\uparrow$).***

Changes in the Expected Future Exchange Rate. Any factor that causes the expected future exchange rate E^e_{t+1} to fall decreases the expected appreciation of the dollar and hence raises the expected appreciation of the euro. The result is a higher expected return

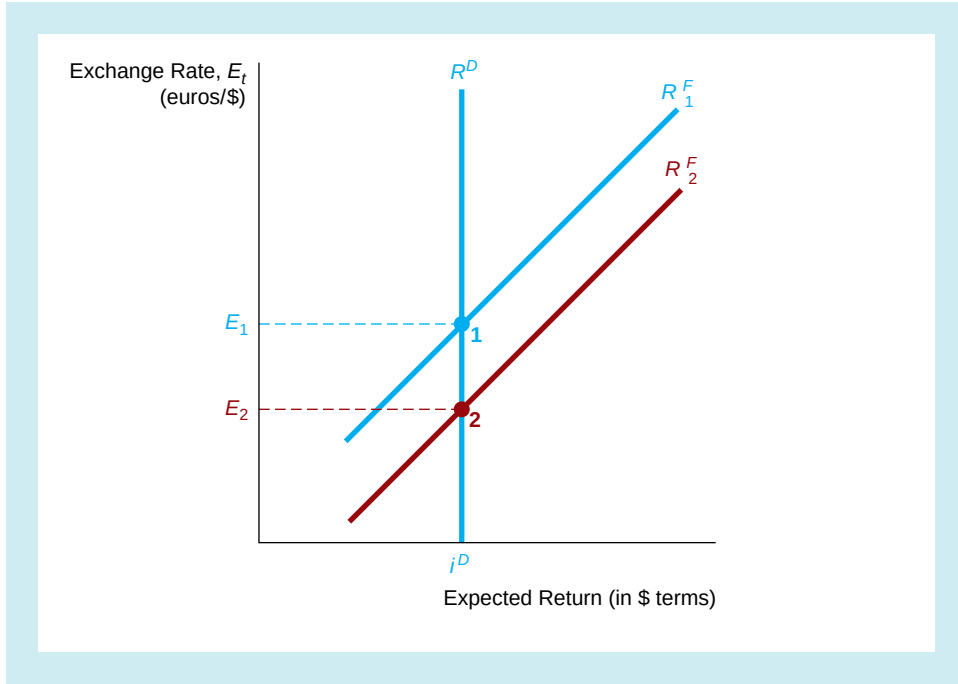


FIGURE 4 Shifts in the Schedule for the Expected Return on Foreign Deposits R^F

An increase in the expected return on foreign deposits, which occurs when either the foreign interest rate rises or the expected future exchange rate falls, shifts the schedule for the expected return on foreign deposits from R^F_1 to R^F_2, and the exchange rate falls from E_1 to E_2.

on euro deposits, which shifts the schedule for the expected return on euro deposits to the right and leads to a decline in the exchange rate as in Figure 4. Conversely, a rise in E^e_{t+1} raises the expected appreciation of the dollar, lowers the expected return on foreign deposits, shifts the R^F schedule to the left, and raises the exchange rate. To summarize, ***a rise in the expected future exchange rate shifts the R^F schedule to the left and causes an appreciation of the domestic currency; a fall in the expected future exchange rate shifts the R^F schedule to the right and causes a depreciation of the domestic currency.***

Summary. Our analysis of the long-run determinants of the exchange rate indicates the factors that influence the expected future exchange rate: the relative price level, relative tariffs and quotas, import demand, export demand, and relative productivity (refer to Table 1). The theory of purchasing power parity suggests that if a higher American price level relative to the foreign price level is expected to persist, the dollar will depreciate in the long run. A higher expected relative American price level should thus have a tendency to lower E^e_{t+1}, raise the expected return on euro deposits, shift the R^F schedule to the right, and lower the current exchange rate.

Similarly, the other long-run determinants of the exchange rate we discussed earlier can also influence the expected return on euro deposits and the current exchange rate. Briefly, the following changes will lower E^e_{t+1}, increase the expected return on euro deposits, shift the R^F schedule to the right, and cause a depreciation of the domestic currency, the dollar: (1) expectations of a rise in the American price level relative to the foreign price level, (2) expectations of lower American trade barriers relative to foreign trade barriers, (3) expectations of higher American import demand, (4) expectations of lower foreign demand for American exports, and (5) expectations of lower American productivity relative to foreign productivity.

Shifts in the Expected-Return Schedule for Domestic Deposits

Since the expected return on domestic (dollar) deposits is just the interest rate on these deposits i^D, this interest rate is the only factor that shifts the schedule for the expected return on dollar deposits.

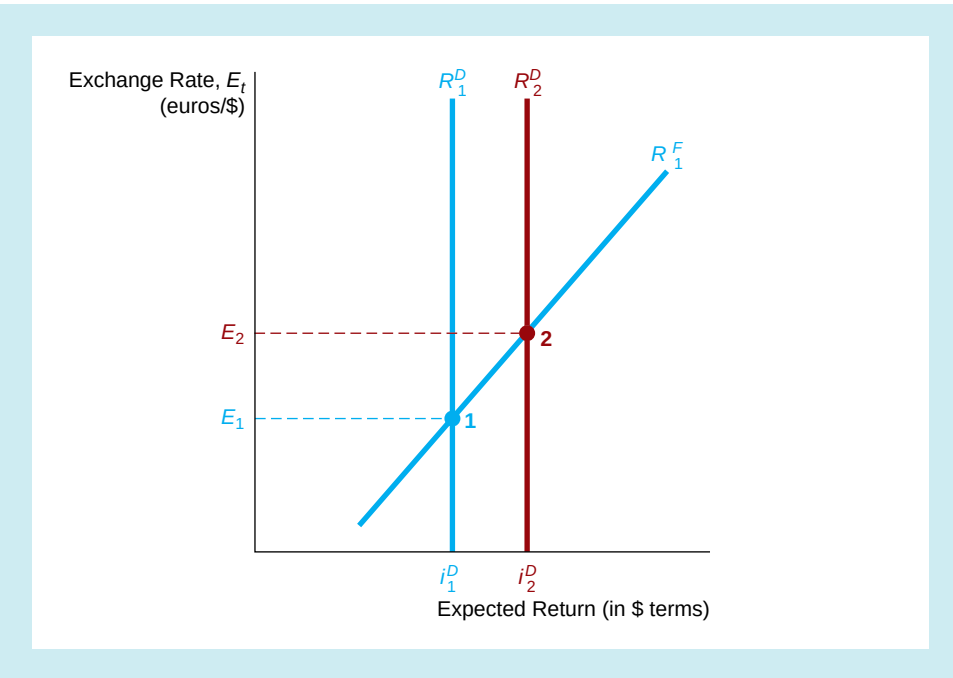
Changes in the Domestic Interest Rate. A rise in i^D raises the expected return on dollar deposits, shifts the R^D schedule to the right, and leads to a rise in the exchange rate, as is shown in Figure 5. Another way of seeing this is to recognize that a rise in i^D, which raises the expected return on dollar deposits, creates an excess demand for dollar deposits at the original equilibrium exchange rate, and the resulting purchases of dollar deposits cause an appreciation of the dollar. ***A rise in the domestic interest rate i^D shifts the R^D schedule to the right and causes an appreciation of the domestic currency; a fall in i^D shifts the R^D schedule to the left and causes a depreciation of the domestic currency.***

Study Guide

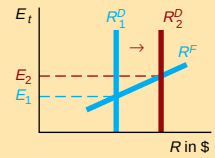
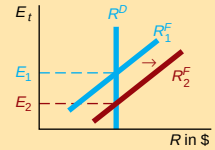
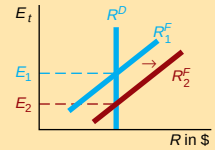
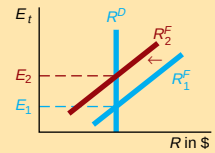
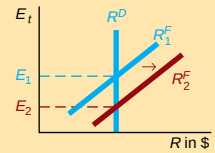
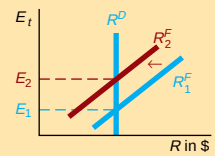
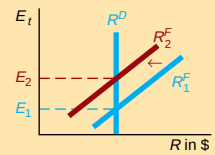
As a study aid, the factors that shift the R^F and R^D schedules and lead to changes in the current exchange rate E_t are listed in Table 2. The table shows what happens to the exchange rate when there is an increase in each of these variables, holding everything else constant. To give yourself practice, see if you can work out what happens to the R^F and R^D schedules and to the exchange rate if each of these factors falls rather than rises. Check your answers by seeing if you get the opposite change in the exchange rate to those indicated in Table 2.

FIGURE 5 Shifts in the Schedule for the Expected Return on Domestic Deposits R^D

An increase in the expected return on dollar deposits i^D shifts the expected return on domestic (dollar) deposits from R_1^D to R_2^D and the exchange rate from E_1 to E_2.



SUMMARY Table 2 Factors That Shift the R^F and R^D Schedules and Affect the Exchange Rate

Factor	Change in Factor	Response of Exchange Rate E_t	
Domestic interest rate i^D	↑	↑	
Foreign interest rate i^F	↑	↓	
Expected domestic price level*	↑	↓	
Expected trade barriers*	↑	↑	
Expected import demand	↑	↓	
Expected export demand	↑	↑	
Expected productivity*	↑	↑	

*Relative to other countries.

Note: Only increases (↑) in the factors are shown; the effects of decreases in the variables on the exchange rate are the opposite of those indicated in the "Response" column.

Application

Changes in the Equilibrium Exchange Rate: Two Examples

Our analysis has revealed the factors that affect the value of the equilibrium exchange rate. Now we use this analysis to take a close look at the response of the exchange rate to changes in interest rates and money growth.

Changes in Interest Rates

Changes in domestic interest rates i^D are often cited as a major factor affecting exchange rates. For example, we see headlines in the financial press like this one: "Dollar Recovers As Interest Rates Edge Upward." But is the view presented in this headline always correct?

Not necessarily, because to analyze the effects of interest rate changes, we must carefully distinguish the sources of the changes. The Fisher equation (Chapter 4) states that a (nominal) interest rate equals the *real* interest rate plus expected inflation: $i = i_r + \pi^e$. The Fisher equation indicates that an interest rate i can change for two reasons: Either the real interest rate i_r changes or the expected inflation rate π^e changes. The effect on the exchange rate is quite different, depending on which of these two factors is the source of the change in the nominal interest rate.

Suppose that the domestic real interest rate increases so that the nominal interest rate i^D rises while expected inflation remains unchanged. In this case, it is reasonable to assume that the expected appreciation of the dollar will be unchanged because expected inflation is unchanged, and so the expected return on foreign deposits will remain unchanged for any given exchange rate. The result is that the R^F schedule stays put and the R^D schedule shifts to the right, and we end up with the situation depicted in Figure 5, which analyzes an increase in i^D, holding everything else constant. Our model of the foreign exchange market produces the following result: ***When domestic real interest rates rise, the domestic currency appreciates.***

When the nominal interest rate rises because of an increase in expected inflation, we get a different result from the one shown in Figure 5. The rise in expected domestic inflation leads to a decline in the expected appreciation of the dollar (a higher appreciation of the euro), which is typically thought to be larger than the increase in the domestic interest rate i^D.[5] As a result, at any given exchange rate, the expected return on foreign deposits rises more than the expected return on dollar deposits. Thus, as we see in Figure 6, the R^F schedule shifts to the right more than the R^D schedule, and the exchange rate falls. Our analysis leads to this conclusion: ***When domestic interest rates rise due to an expected increase in inflation, the domestic currency depreciates.***

Because this conclusion is completely different from the one reached when the rise in the domestic interest rate is associated with a higher real

[5]This conclusion is standard in asset market models of exchange rate determination; see Rudiger Dornbusch, "Expectations and Exchange Rate Dynamics," *Journal of Political Economy* 84 (1976): 1061–1076. It is also consistent with empirical evidence that suggests that nominal interest rates do not rise one-for-one with increases in expected inflation. See Frederic S. Mishkin, "The Real Interest Rate: An Empirical Investigation," *Carnegie-Rochester Conference Series on Public Policy* 15 (1981): 151–200; and Lawrence Summers, "The Nonadjustment of Nominal Interest Rates: A Study of the Fisher Effect," in *Macroeconomics, Prices and Quantities*, ed. James Tobin (Washington, D.C.: Brookings Institution, 1983), pp. 201–240.

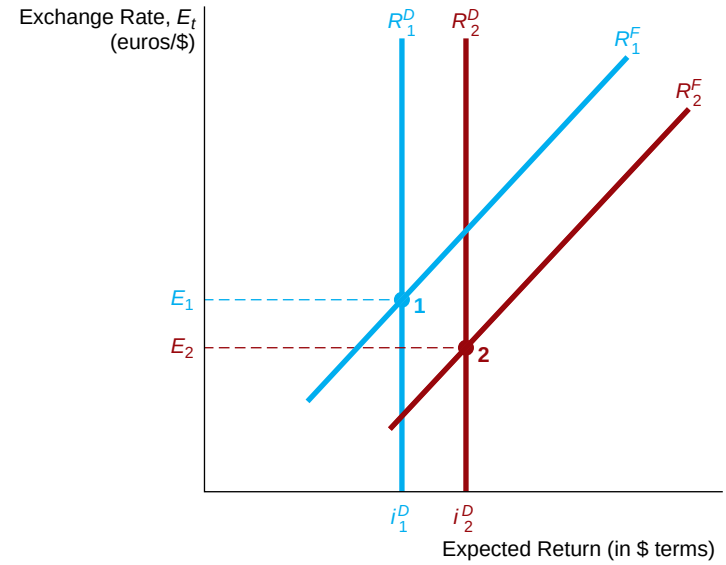


FIGURE 6 Effect of a Rise in the Domestic Nominal Interest Rate as a Result of an Increase in Expected Inflation

Because a rise in domestic expected inflation leads to a decline in expected dollar appreciation that is larger than the resulting increase in the domestic interest rate, the expected return on foreign deposits rises by more than the expected return on domestic (dollar) deposits. R^F shifts to the right more than R^D, and the equilibrium exchange rate falls from E_1 to E_2.

interest rate, we must always distinguish between *real* and *nominal* measures when analyzing the effects of interest rates on exchange rates.

Changes in the Money Supply

Suppose that the Federal Reserve decides to increase the level of the money supply in order to reduce unemployment, which it believes to be excessive. The higher money supply will lead to a higher American price level in the long run (as we will see in Chapter 25) and hence to a lower expected future exchange rate. The resulting decline in the expected appreciation of the dollar increases the expected return on foreign deposits at any given current exchange rate and so shifts the R^F schedule rightward from R_1^F to R_2^F in Figure 7. In addition, the higher money supply will lead to a higher real money supply M/P because the price level does not immediately increase in the short run. As suggested in Chapter 5, the resulting rise in the real money supply causes the domestic interest rate to fall from i_1^D to i_2^D, which lowers the expected return on domestic (dollar) deposits, shifting the R^D schedule leftward from R_1^D to R_2^D. As we can see in Figure 7, the result is a decline in the exchange rate from E_1 to E_2. The conclusion is this: ***A higher domestic money supply causes the domestic currency to depreciate.***

Exchange Rate Overshooting

Our analysis of the effect of an increase in the money supply on the exchange rate is not yet over—we still need to look at what happens to the exchange rate in the long run. A basic proposition in monetary theory, called **monetary neutrality**, states that in the long run, a one-time percentage rise in the money supply is matched by the same one-time percentage rise in the price level, leaving unchanged the real money supply and all other economic variables such as interest rates. An intuitive way to understand this proposition

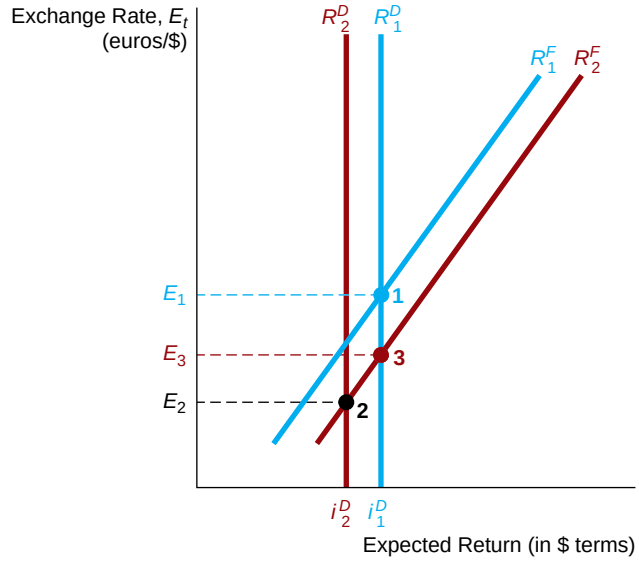
is to think of what would happen if our government announced overnight that an old dollar would now be worth 100 new dollars. The money supply in new dollars would be 100 times its old value and the price level would also be 100 times higher, but nothing in the economy would really have changed; real and nominal interest rates and the real money supply would remain the same. Monetary neutrality tells us that in the long run, the rise in the money supply would not lead to a change in the domestic interest rate and so it would return to i^D_1 in the long run, and the schedule for the expected return on domestic deposits would return to R^D_1. As we can see in Figure 7, this means that the exchange rate would rise from E_2 to E_3 in the long run.

The phenomenon we have described here in which the exchange rate falls by more in the short run than it does in the long run when the money supply increases is called **exchange rate overshooting**. It is important because, as we will see in the following application, it can help explain why exchange rates exhibit so much volatility.

Another way of thinking about why exchange rate overshooting occurs is to recognize that when the domestic interest rate falls in the short run, equilibrium in the foreign exchange market means that the expected return on foreign deposits must be lower. With the foreign interest rate given, this lower expected return on foreign deposits means that there must be an expected appreciation of the dollar (depreciation of the euro) in order for the expected return on foreign deposits to decline when the domestic interest rate falls. This can occur only if the current exchange rate falls below its long-run value.

FIGURE 7 Effect of a Rise in the Money Supply

A rise in the money supply leads to a higher domestic price level in the long run, which in turn leads to a lower expected future exchange rate. The resulting decline in the expected appreciation of the dollar raises the expected return on foreign deposits, shifting the R^F schedule rightward from R^F_1 to R^F_2. In the short run, the domestic interest rate i^D falls, shifting R^D from R^D_1 to R^D_2. The short-run outcome is that the exchange rate falls from E_1 to E_2. In the long run, however, the interest rate returns to i^D_1 and R^D returns to R^D_1. The exchange rate thus rises from E_2 to E_3 in the long run.



Application

Why Are Exchange Rates So Volatile?

The high volatility of foreign exchange rates surprises many people. Thirty or so years ago, economists generally believed that allowing exchange rates to be determined in the free market would not lead to large fluctuations in their values. Recent experience has proved them wrong. If we return to Figure 1, we see that exchange rates over the 1980–2002 period have been very volatile.

The asset market approach to exchange rate determination that we have outlined in this chapter gives a straightforward explanation of volatile exchange rates. Because expected appreciation of the domestic currency affects the expected return on foreign deposits, expectations about the price level, inflation, trade barriers, productivity, import demand, export demand, and the money supply play important roles in determining the exchange rate. When expectations about any of these variables change, our model indicates that there will be an immediate effect on the expected return on foreign deposits and therefore on the exchange rate. Since expectations on all these variables change with just about every bit of news that appears, it is not surprising that the exchange rate is volatile. In addition, we have seen that our exchange rate analysis produces exchange rate overshooting when the money supply increases. Exchange rate overshooting is an additional reason for the high volatility of exchange rates.

Because earlier models of exchange rate behavior focused on goods markets rather than asset markets, they did not emphasize changing expectations as a source of exchange rate movements, and so these earlier models could not predict substantial fluctuations in exchange rates. The failure of earlier models to explain volatility is one reason why they are no longer so popular. The more modern approach developed here emphasizes that the foreign exchange market is like any other asset market in which expectations of the future matter. The foreign exchange market, like other asset markets such as the stock market, displays substantial price volatility, and foreign exchange rates are notoriously hard to forecast.

Application

The Dollar and Interest Rates, 1973–2002

In the chapter preview, we mentioned that the dollar was weak in the late 1970s, rose substantially from 1980 to 1985, and declined thereafter. We can use our analysis of the foreign exchange market to understand exchange rate movements and help explain the dollar's rise and fall in the 1980s.

Some important information for tracing the dollar's changing value is presented in Figure 8, which plots measures of real and nominal interest rates and the value of the dollar in terms of a basket of foreign currencies (called an **effective exchange rate index**). We can see that the value of the dollar and the measure of real interest rates tend to rise and fall together. In the late 1970s, real interest rates were at low levels, and so was the value of the dollar. Beginning in 1980, however, real interest rates in the United States began

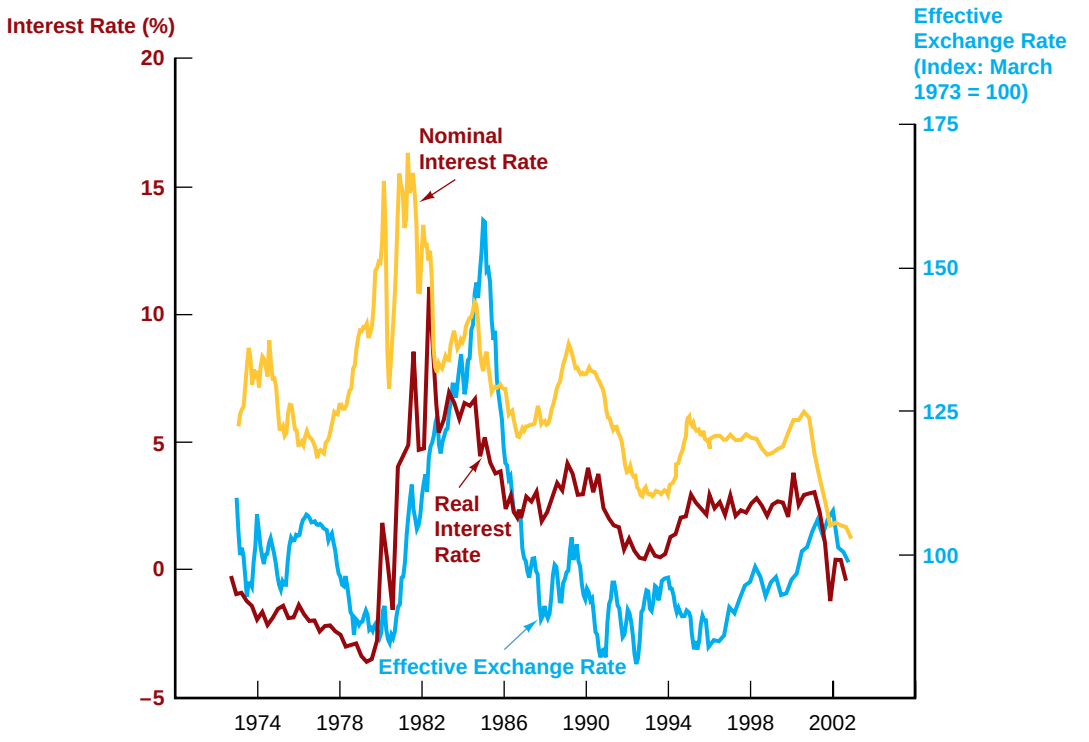


FIGURE 8 Value of the Dollar and Interest Rates, 1973–2002

Sources: Federal Reserve: www.federalreserve.gov/releases/h10/summary/indexn_m.txt; real interest rate from Figure 1 in Chapter 4.

to climb sharply, and at the same time so did the dollar. After 1984, the real interest rate declined substantially, as did the dollar.

Our model of exchange rate determination helps explain the rise and fall in the dollar in the 1980s. As Figure 5 indicates, a rise in the U.S. real interest rate raises the expected return on dollar deposits while leaving the expected return on foreign deposits unchanged. The resulting increased demand for dollar deposits then leads to purchases of dollar deposits (and sales of foreign deposits), which raise the exchange rate. This is exactly what occurred in the 1980–1984 period. The subsequent fall in U.S. real interest rates then lowered the expected return on dollar deposits relative to foreign deposits, and the resulting sales of dollar deposits (and purchases of foreign deposits) lowered the exchange rate.

The plot of *nominal* interest rates in Figure 8 also demonstrates that the correspondence between nominal interest rates and exchange rate movements is not nearly as close as that between *real* interest rates and exchange rate movements. This is also exactly what our analysis predicts. The rise in nominal interest rates in the late 1970s was not reflected in a corresponding rise in the value of the dollar; indeed, the dollar actually fell in the late 1970s.

Figure 8 explains why the rise in nominal rates in the late 1970s did not produce a rise in the dollar. As a comparison of the real and nominal interest rates in the late 1970s indicates, the rise in nominal interest rates reflected an increase in expected inflation and not an increase in real interest rates. As our analysis in Figure 6 demonstrates, the rise in nominal interest rates stemming from a rise in expected inflation should lead to a decline in the dollar, and that is exactly what happened.

If there is a moral to the story, it is that a failure to distinguish between real and nominal interest rates can lead to poor predictions of exchange rate movements: The weakness of the dollar in the late 1970s and the strength of the dollar in the early 1980s can be explained by movements in *real* interest rates but not by movements in *nominal* interest rates.

Application

The Euro's First Four Years

With much fanfare, the euro debuted on January 1, 1999, at an exchange rate of 1.18 dollars per euro. Despite initial hopes that the euro would be a strong currency, it has proved to be weak, declining 30% to a low of 83 cents per euro in October 2000, only to recover to 1.05 dollars per euro by the beginning of 2003. What explains the weakness of the euro in its first two years, and its recovery in its third and fourth year?

The previous application has shown that changes in real interest rates are an important factor determining the exchange rate. When the domestic real interest rate falls relative to the foreign real interest rate, the domestic currency declines in value. Indeed, this is exactly what has happened to the euro. While the euro was coming into existence, European economies were experiencing only slow recoveries from recession, thus causing both real and nominal interest rates to fall. In contrast, in 1999 and 2000, the United States experienced very rapid growth, substantially above their European counterparts. As in the analysis of the previous application, low real interest rates in Europe relative to those in the United States drove down the value of the euro.

With the slowing of the U.S. economy, which entered into recession in the spring of 2001, the process reversed. The U.S. growth rate fell slightly behind Europe's, so that U.S. relative real and nominal interest rates fell, setting the stage for a recovery in the euro.

Application

Reading the *Wall Street Journal*: The "Currency Trading" Column

Now that we have an understanding of how exchange rates are determined, we can use our analysis to understand discussions about developments in the foreign exchange market reported in the financial press.

Every day, the *Wall Street Journal* reports on developments in the foreign exchange market on the previous business day in its "Currency

Trading" column, an example of which is presented in the "Following the Financial News" box.

The column indicates that concerns about a possible war against Iraq and weak economic data have put downward pressure on the U.S. dollar. Our analysis of the foreign exchange market explains why these developments have led to a weak dollar.

Following the Financial News

The "Currency Trading" Column

The "Currency Trading" column appears daily in the *Wall Street Journal*; an example is presented here. It is usually found in the third section, "Money and Investing."

CURRENCY TRADING

Concerns About War Put Pressure on the Dollar

By Grainne McCarthy
Dow Jones Newswires

NEW YORK—Having fallen swiftly on the back of some surprisingly weak U.S. employment data, the dollar is set to remain under pressure this week, increasingly vulnerable to the drumbeat of war surrounding Iraq and nuclear saber-rattling in North Korea.

"People are positioned for Armageddon on the dollar. In that scenario, you can get wacky moves," says Paul Podolsky, chief strategist at Fleet Global Markets in Boston.

Investors unsure of the dollar's vulnerability to the U.S. economic data got a resounding wake-up call Friday, with the currency tumbling swiftly after the government reported a dismal December payrolls report that fueled concerns about the lingering soft spot dogging the world's largest economy. The dollar hit a fresh three-year trough against the euro, while sliding to its weakest point against the Swiss franc in four years.

In late New York trading Friday, the euro was at $1.0579, up steeply from $1.0488 late Thursday. Against the Swiss franc, the dollar was at 1.3799 francs, down sharply from 1.3912 francs, while sterling was at $1.6084, modestly up from $1.6064. The dollar was at 119.16 yen, modestly lower than 119.30 yen Thursday.

Even as Canada reported another remarkably strong month of employment growth, job losses in the U.S. soared to 101,000, far from consensus forecasts for a modest increase of 20,000.

There were clearly some seasonal explanations for the leap, but the report still underscored a view that the sluggish U.S. economic recovery isn't creating jobs. That potentially bodes ill for the dollar at a time when it is already being undermined by war concerns.

"Until Iraq goes away and the outlook for consumer confidence and business spending improves, the dollar is going to remain under pressure," said Jay Bryson, global economist at Wachovia Securities in Charlotte, N.C.

There will certainly be plenty of economic data this week for dollar investors to sink their teeth into, with the focus most likely on somewhat stronger economic activity and benign inflation. Retail sales for December, to be reported tomorrow, are expected to come in very firm, mostly because of the 18% jump in auto sales already reported. But excluding autos, economists anticipate just a 0.3% increase.

The U.S. will also get December's producer-price and consumer-price indexes on Wednesday and Thursday, respectively. The focus for Friday will be squarely on the initial University of Michigan consumer-sentiment report for January, which should provide a glimpse of how confidence is holding up amid growing war jitters.

But aside from the clear significance of much of these data, many analysts expect the dollar to look more to the Pentagon, State Department and, ultimately, the White House for signposts for near-term direction.

As the central emblem in financial markets of the world's only superpower, the dollar is beset by multiple threats to global stability that are breaking out on several fronts. As well as the situations in North Korea and Iraq, the continuing battle against terrorist network al Qaeda is high on the list of issues facing the Bush administration. The U.S., given its position of global hegemony, has almost by default become the first line of defense in tackling these challenges.

"Connect the dots, and what emerges is hardly encouraging for the dollar in particular and the financial markets in general," said Joseph Quinlan, global economist at Johns Hopkins University.

He argues that investors in U.S. assets, while certainly cognizant of a war risk, may have priced in an overly rosy scenario under which the war on terrorism has already been won, the war against Iraq has already been priced in, and a war on the Korean peninsula is too remote a possibility to take seriously.

An upset to this more optimistic picture could weigh more heavily on global capital flows, ultimately depressing the dollar, given the U.S.'s status as a creditor nation dependent on capital inflows to finance the current account.

Source: Wall Street Journal, Monday, January 13, 2003, p. C16.

The column starts by pointing out that surprisingly weak U.S. employment data has led to a falling dollar. The weakness in the U.S. economy suggests that real interest rates in the United States are likely to fall in the future. As a result, in the future we have the opposite scenario to Figure 5 occurring, the R^D curve shifts to the left, lowering the value of the dollar. The future decline in the dollar then means that the foreign currency is expected to appreciate, thus raising the expected return on foreign deposits today and shifting the R^F curve to the right. Thus as we see in Figure 4 the dollar declines today.

Concerns about a war with Iraq also have a similar impact. The possibility that the war in Iraq might not go well and could lead to increased terrorist incidents suggests that the U.S. economy might suffer negative consequences. This provides an additional reason for a possible weakening of the economy and we then get the identical analysis to that in the paragraph above, which shows that the U.S. dollar would fall in response to these war fears.

Summary

1. Foreign exchange rates (the price of one country's currency in terms of another's) are important because they affect the price of domestically produced goods sold abroad and the cost of foreign goods bought domestically.
2. The theory of purchasing power parity suggests that long-run changes in the exchange rate between two countries are determined by changes in the relative price levels in the two countries. Other factors that affect exchange rates in the long run are tariffs and quotas, import demand, export demand, and productivity.
3. Exchange rates are determined in the short run by the interest parity condition, which states that the expected return on domestic deposits is equal to the expected return on foreign deposits.
4. Any factor that changes the expected returns on domestic or foreign deposits will lead to changes in the exchange rate. Such factors include changes in the interest rates on domestic and foreign deposits as well as changes in any of the factors that affect the long-run exchange rate and hence the expected future exchange rate. Changes in the money supply lead to exchange rate overshooting, causing the exchange rate to change by more in the short run than in the long run.
5. The asset market approach to exchange rate determination can explain both the volatility of exchange rates and the rise of the dollar in the 1980–1984 period and its subsequent fall.

Key Terms

appreciation, p. 436
capital mobility, p. 445
depreciation, p. 436
effective exchange rate index, p. 455
exchange rate, p. 435
exchange rate overshooting, p. 454
foreign exchange market, p. 435
forward exchange rate, p. 436
forward transaction, p. 436
interest parity condition, p. 445
law of one price, p. 439
monetary neutrality, p. 453

quotas, p. 441
spot exchange rate, p. 436
spot transaction, p. 436
tariffs, p. 441
theory of purchasing power parity (PPP), p. 439

Questions and Problems

Questions marked with an asterisk are answered at the end of the book in an appendix, "Answers to Selected Questions and Problems."

1. When the euro appreciates, are you more likely to drink California or French wine?

*2. "A country is always worse off when its currency is weak (falls in value)." Is this statement true, false, or uncertain? Explain your answer.

3. In a newspaper, check the exchange rates for the foreign currencies listed in the "Following the Financial News" box on page 437. Which of these currencies have appreciated and which have depreciated since February 5, 2003?

*4. If the Japanese price level rises by 5% relative to the price level in the United States, what does the theory of purchasing power parity predict will happen to the value of the Japanese yen in terms of dollars?

5. If the demand for a country's exports falls at the same time that tariffs on imports are raised, will the country's currency tend to appreciate or depreciate in the long run?

*6. In the mid- to late 1970s, the yen appreciated relative to the dollar even though Japan's inflation rate was higher than America's. How can this be explained by an improvement in the productivity of Japanese industry relative to American industry?

Using Economic Analysis to Predict the Future

Answer the remaining questions by drawing the appropriate exchange market diagrams.

7. The president of the United States announces that he will reduce inflation with a new anti-inflation program. If the public believes him, predict what will happen to the U.S. exchange rate.

*8. If the British central bank prints money to reduce unemployment, what will happen to the value of the pound in the short run and the long run?

9. If the Canadian government unexpectedly announces that it will be imposing higher tariffs on foreign goods one year from now, what will happen to the value of the Canadian dollar today?

*10. If nominal interest rates in America rise but real interest rates fall, predict what will happen to the U.S. exchange rate.

11. If American auto companies make a breakthrough in automobile technology and are able to produce a car that gets 60 miles to the gallon, what will happen to the U.S. exchange rate?

*12. If Americans go on a spending spree and buy twice as much French perfume, Japanese TVs, English sweaters, Swiss watches, and Italian wine, what will happen to the value of the U.S. dollar?

13. If expected inflation drops in Europe so that interest rates fall there, predict what will happen to the U.S. exchange rate.

*14. If the European central bank decides to contract the money supply in order to fight inflation, what will happen to the value of the U.S. dollar?

15. If there is a strike in France, making it harder to buy French goods, what will happen to the value of the euro?

Web Exercises

1. The Federal Reserve maintains a web site that lists the exchange rate between the U.S. dollar and many other currencies. Go to www.federalreserve.gov/releases/H10/hist/. Go to the historical data from 1999 and later and find the Euro.
   a. What has the percentage change in the Euro-dollar exchange rate been between introduction and now?
   b. What has been the annual percentage change in the Euro-dollar exchange rate for each year since the Euro's introduction?
2. International travelers and business people frequently need to accurately convert from one currency to another. It is often easy to find the rate needed to convert the U.S. dollar into another currency. It can be more difficult to find cross-conversion rates. Go to www.oanda.com/convert/classic. This site lets you convert from any currency into any other currency. How many Lithuanian Litas can you currently buy with one Chilean Peso?

Chapter 20

The International Financial System

PREVIEW

Thanks to the growing interdependence between the U.S. economy and the economies of the rest of the world, a country's monetary policy can no longer be conducted without taking international considerations into account. In this chapter, we examine how international financial transactions and the structure of the international financial system affect monetary policy. We also examine the evolution of the international financial system during the past half century and consider where it may be heading in the future.

Intervention in the Foreign Exchange Market

In Chapter 19, we analyzed the foreign exchange market as if it were a completely free market that responds to all market pressures. Like many other markets, however, the foreign exchange market is not free of government intervention; central banks regularly engage in international financial transactions called **foreign exchange interventions** in order to influence exchange rates. In our current international financial arrangement, called a **managed float regime** (or a **dirty float**), exchange rates fluctuate from day to day, but central banks attempt to influence their countries' exchange rates by buying and selling currencies. The exchange rate analysis we developed in Chapter 19 is used here to explain the impact that central bank intervention has on the foreign exchange market.

Foreign Exchange Intervention and the Money Supply

The first step in understanding how central bank intervention in the foreign exchange market affects exchange rates is to see the impact on the monetary base from a central bank sale in the foreign exchange market of some of its holdings of assets denominated in a foreign currency (called **international reserves**). Suppose that the Fed decides to sell $1 billion of its foreign assets in exchange for $1 billion of U.S. currency. (This transaction is conducted at the foreign exchange desk at the Federal Reserve Bank of New York—see Box 1.) The Fed's purchase of dollars has two effects. First, it reduces the Fed's holding of international reserves by $1 billion. Second, because its purchase of currency removes it from the hands of the public, currency in

Box 1: Inside the Fed

A Day at the Federal Reserve Bank of New York's Foreign Exchange Desk

Although the U.S. Treasury is primarily responsible for foreign exchange policy, decisions to intervene in the foreign exchange market are made jointly by the U.S. Treasury and the Federal Reserve as represented by the FOMC (Federal Open Market Committee). The actual conduct of foreign exchange intervention is the responsibility of the foreign exchange desk at the Federal Reserve Bank of New York, which is right next to the open market desk.

The manager of foreign exchange operations at the New York Fed supervises the traders and analysts who follow developments in the foreign exchange market. Every morning at 7:30, a trader on his staff who has arrived at the New York Fed in the predawn hours speaks on the telephone with counterparts at the U.S. Treasury and provides an update on overnight activity in overseas financial and foreign exchange markets. Later in the morning, at 9:30, the manager and his staff hold a conference call with senior staff at the Board of Governors of the Federal Reserve in Washington. In the afternoon, at 2:30, they have a second conference call, which is a joint briefing of officials at the board and the Treasury. Although by statute the Treasury has the lead role in setting foreign exchange policy, it strives to reach a consensus among all three parties—the Treasury, the Board of Governors, and the Federal Reserve Bank of New York. If they decide that a foreign exchange intervention is necessary that day—an unusual occurrence, as a year may go by without a U.S. foreign exchange intervention—the manager instructs his traders to carry out the agreed-on purchase or sale of foreign currencies. Because funds for exchange rate intervention are held separately by the Treasury (in its Exchange Stabilization Fund) and the Federal Reserve, the manager and his staff are not trading the funds of the Federal Reserve Bank of New York; rather they act as an agent for the Treasury and the FOMC in conducting these transactions.

As part of their duties, before every FOMC meeting, the staff help prepare a lengthy document full of data for the FOMC members, other Reserve bank presidents, and Treasury officials. It describes developments in the domestic and foreign markets over the previous five or six weeks, a task that keeps them especially busy right before the FOMC meeting.

circulation falls by $1 billion. We can see this in the following T-account for the Federal Reserve:

FEDERAL RESERVE SYSTEM

Assets		Liabilities	
Foreign assets (international reserves)	−$1 billion	Currency in circulation	−$1 billion

Because the monetary base is made up of currency in circulation plus reserves, this decline in currency implies that the monetary base has fallen by $1 billion.

If instead of paying for the foreign assets sold by the Fed with currency, the persons buying the foreign assets pay for them by checks written on accounts at domestic banks,

then the Fed deducts the $1 billion from the deposit accounts these banks have with the Fed. The result is that deposits with the Fed (reserves) decline by $1 billion, as shown in the following T-account:

FEDERAL RESERVE SYSTEM

Assets		Liabilities	
Foreign assets (international reserves)	−$1 billion	Deposits with the Fed (reserves)	−$1 billion

In this case, the outcome of the Fed sale of foreign assets and the purchase of dollar deposits is a $1 billion decline in reserves and a $1 billion decline in the monetary base because reserves are also a component of the monetary base.

We now see that the outcome for the monetary base is exactly the same when a central bank sells foreign assets to purchase domestic bank deposits or domestic currency. This is why when we say that a central bank has purchased its domestic currency, we do not have to distinguish whether it actually purchased currency or bank deposits denominated in the domestic currency. We have thus reached an important conclusion: ***A central bank's purchase of domestic currency and corresponding sale of foreign assets in the foreign exchange market leads to an equal decline in its international reserves and the monetary base.***

We could have reached the same conclusion by a more direct route. A central bank sale of a foreign asset is no different from an open market sale of a government bond. We learned in our exploration of the money supply process that an open market sale leads to an equal decline in the monetary base; therefore, a sale of foreign assets also leads to an equal decline in the monetary base. By similar reasoning, a central bank purchase of foreign assets paid for by selling domestic currency, like an open market purchase, leads to an equal rise in the monetary base. Thus we reach the following conclusion: ***A central bank's sale of domestic currency to purchase foreign assets in the foreign exchange market results in an equal rise in its international reserves and the monetary base.***

The intervention we have just described, in which a central bank allows the purchase or sale of domestic currency to have an effect on the monetary base, is called an **unsterilized foreign exchange intervention**. But what if the central bank does not want the purchase or sale of domestic currency to affect the monetary base? All it has to do is to counter the effect of the foreign exchange intervention by conducting an offsetting open market operation in the government bond market. For example, in the case of a $1 billion purchase of dollars by the Fed and a corresponding $1 billion sale of foreign assets, which, as we have seen, would decrease the monetary base by $1 billion, the Fed can conduct an open market purchase of $1 billion of government bonds, which would increase the monetary base by $1 billion. The resulting T-account for the foreign exchange intervention and the offsetting open market operation leaves the monetary base unchanged:

FEDERAL RESERVE SYSTEM

Assets		Liabilities	
Foreign assets (international reserves)	−$1 billion	Monetary base (reserves)	0
Government bonds	+$1 billion		

A foreign exchange intervention with an offsetting open market operation that leaves the monetary base unchanged is called a **sterilized foreign exchange intervention**.

Now that we understand that there are two types of foreign exchange interventions—unsterilized and sterilized—let's look at how each affects the exchange rate.

Unsterilized Intervention

Your intuition might lead you to suspect that if a central bank wants to lower the value of the domestic currency, it should sell its currency in the foreign exchange market and purchase foreign assets. Indeed, this intuition is correct for the case of an unsterilized intervention.

Recall that in an unsterilized intervention, if the Federal Reserve decides to sell dollars in order to buy foreign assets in the foreign exchange market, this works just like an open market purchase of bonds to increase the monetary base. Hence the sale of dollars leads to an increase in the money supply, and we find ourselves analyzing exactly the situation already described in Figure 7 of Chapter 19, which is reproduced here as Figure 1. The higher money supply leads to a higher U.S. price level in the long run and so to a lower expected future exchange rate. The resulting decline in the expected appreciation of the dollar increases the expected return on foreign deposits and shifts the R^F schedule to the right. In addition, the increase in the money supply will lead to a higher real money supply in the short run, which causes the interest rate on dollar deposits to fall. The resulting lower expected return on dollar deposits translates as a leftward shift in the R^D schedule. The fall in the expected return on dollar deposits and the increase in the expected return on foreign deposits means that foreign assets have a higher expected return than dollar deposits at the old equilibrium exchange rate. Hence people will try to sell their dollar deposits, and the exchange rate will fall. Indeed, as we saw in the previous chapter, the increase in the money supply will lead to exchange rate overshooting, whereby the exchange rate falls by more in the short run than it does in the long run.

Our analysis leads us to the following conclusion about unsterilized interventions in the foreign exchange market: ***An unsterilized intervention in which domestic currency is sold to purchase foreign assets leads to a gain in international reserves, an increase in the money supply, and a depreciation of the domestic currency.***

The reverse result is found for an unsterilized intervention in which domestic currency is purchased by selling foreign assets. The purchase of domestic currency by selling foreign assets (reducing international reserves) works like an open market sale to reduce the monetary base and the money supply. The decrease in the money supply raises the interest rate on dollar deposits and shifts R^D rightward, while causing

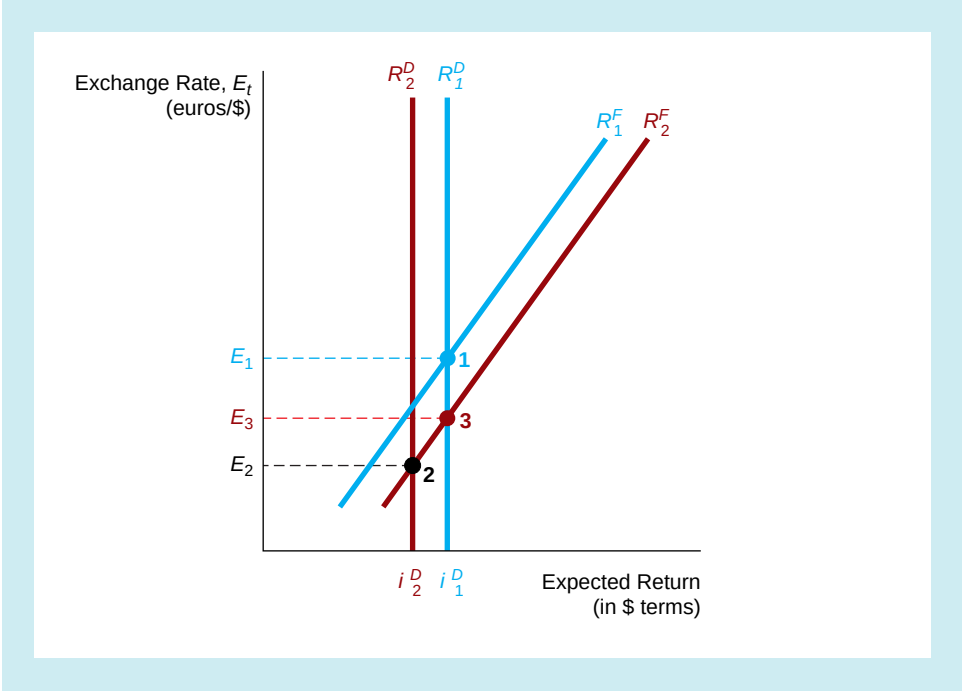


FIGURE 1 Effect of a Sale of Dollars and a Purchase of Foreign Assets

A sale of dollars and the consequent open market purchase of foreign assets increase the monetary base. The resulting rise in the money supply leads to a higher domestic price level in the long run, which leads to a lower expected future exchange rate. The resulting decline in the expected appreciation of the dollar raises the expected return on foreign deposits, shifting the R^F schedule rightward from R_1^F to R_2^F. In the short run, the domestic interest rate i^D falls, shifting R^D from R_1^D to R_2^D. The short-run outcome is that the exchange rate falls from E_1 to E_2. In the long run, however, the interest rate returns to i_1^D, and R^D returns to R_1^D. The exchange rate therefore rises from E_2 to E_3 in the long run.

R^F to shift leftward, because it leads to a lower U.S. price level in the long run and thus to a higher expected appreciation of the dollar, and hence a lower expected return on foreign deposits. The increase in the expected return on dollar deposits relative to foreign deposits will mean that people will want to buy more dollar deposits, and the exchange rate will rise. ***An unsterilized intervention in which domestic currency is purchased by selling foreign assets leads to a drop in international reserves, a decrease in the money supply, and an appreciation of the domestic currency.***

Sterilized Intervention

The key point to remember about a sterilized intervention is that the central bank engages in offsetting open market operations, so that there is no impact on the monetary base and the money supply. In the context of the model of exchange rate determination we have developed here, it is straightforward to show that a sterilized intervention has *no effect* on the exchange rate. Remember that in our model, foreign and domestic deposits are perfect substitutes, so equilibrium in the foreign exchange market occurs when the expected returns on foreign and domestic deposits are equal. A sterilized intervention leaves the money supply unchanged and so has no way of directly affecting interest rates or the expected future exchange rate.[1] Because the

[1]Note that a sterilized intervention could indicate what central banks want to happen to the future exchange rate and so might provide a signal about the course of future monetary policy. In this way, a sterilized intervention could lead to shifts in the R^F schedule, but in reality it is the future change in monetary policy, not the sterilized intervention, that is the ultimate source of exchange rate effects. For a discussion of the signaling effect, see Maurice Obstfeld, "The Effectiveness of Foreign Exchange Intervention: Recent Experience, 1985–1988," in *International Policy Coordination and Exchange Rate Fluctuations*, ed. William H. Branson, Jacob A. Frenkel, and Morris Goldstein (Chicago: University of Chicago Press, 1990), pp. 197–237.

expected returns on dollar and foreign deposits are unaffected, the expected return schedules remain at R_1^D and R_1^F in Figure 1, and the exchange rate remains unchanged at E_1.

At first it might seem puzzling that a central bank purchase or sale of domestic currency that is sterilized does not lead to a change in the exchange rate. A central bank purchase of domestic currency cannot raise the exchange rate, because with no effect on the domestic money supply or interest rates, any resulting rise in the exchange rate would mean that the expected return on foreign deposits would be greater than the expected return on domestic deposits. Given our assumption that foreign and domestic deposits are perfect substitutes (equally desirable), this would mean that no one would want to hold domestic deposits.[2] So the exchange rate would have to fall back to its previous level, where the expected returns on domestic and foreign deposits were equal.

Balance of Payments

http://research.stlouisfed.org/fred/data/exchange.html

This web site contains exchange rates, balance of payments, and trade data.

Because international financial transactions such as foreign exchange interventions have considerable effects on monetary policy, it is worth knowing how these transactions are measured. The **balance of payments** is a bookkeeping system for recording all receipts and payments that have a direct bearing on the movement of funds between a nation (private sector and government) and foreign countries.

Here we examine the key items in the balance of payments that you often hear about in the media.

The **current account** shows international transactions that involve currently produced goods and services. The difference between merchandise exports and imports, the net receipts from trade, is called the **trade balance**. When merchandise imports are greater than exports (by $427 billion), we have a trade deficit; if exports are greater than imports, we have a trade surplus.

Three additional items included in the current account are the net receipts that arise from investment income, the purchase and sale of services, and unilateral transfers (gifts, pensions, and foreign aid). In 2001, for example, net investment income was negative $19 billion for the United States because Americans received less investment income from abroad than they paid out. Americans bought less in services from foreigners than foreigners bought from Americans, so net services generated $79 billion in receipts. Since Americans made more unilateral transfers to foreign countries (especially foreign aid) than foreigners made to the United States, net unilateral transfers were negative $50 billion. The sum of the previous three items plus the trade balance is the current account balance, which in 2001 showed a deficit of $417 billion ($-\$427 - \$19 + \$79 - \$50 = -\$417$ billion).

Another important item in the balance of payments is the **capital account**, the net receipts from capital transactions. In 2001 the capital account was + $416 billion,

[2] If domestic and foreign deposits are not perfect substitutes, a sterilized intervention can affect the exchange rate. However, most studies find little evidence to support the position that sterilized intervention has a significant impact on foreign exchange rates. For a further discussion of the effects of sterilized versus unsterilized intervention, see Paul Krugman and Maurice Obstfeld, *International Economics,* 5th ed. (Reading, Mass.: Addison Wesley Longman, 2000).

indicating that $416 billion more capital came into the United States than went out. Another way of saying this is that the United States had a net capital inflow of $416 billion.[3] The sum of the current account and the capital account equals the **official reserve transactions balance**, which was negative $1 billion in 2001 (−$417 + $416 = −$1 billion). When economists refer to a surplus or deficit in the balance of payments, they actually mean a surplus or deficit in the official reserve transactions balance.

Because the balance of payments must balance, the official reserve transactions balance, which equals the current account plus the capital account, tells us the net amount of international reserves that must move between governments (as represented by their central banks) to finance international transactions: i.e.,

$$\text{Current account} + \text{capital account} = \text{net change in government international reserves}$$

This equation shows us why the current account receives so much attention from economists and the media. The current account balance tells us whether the United States (private sector and government combined) is increasing or decreasing its claims on foreign wealth. A surplus indicates that America is increasing its claims on foreign wealth, and a deficit, as in 2001, indicates that the country is reducing its claims on foreign wealth.[4]

Financial analysts follow the current account balance closely because they believe that it can provide information on the future movement of exchange rates. The current account balance provides some indication of what is happening to the demand for imports and exports, which, as we saw in the previous chapter, can affect the exchange rate. In addition, the current account balance provides information about what will be happening to U.S. claims on foreign wealth in the long run. Because a movement of foreign wealth to American residents can affect the demand for dollar assets, changes in U.S. claims on foreign wealth, reflected in the current account balance, can affect the exchange rate over time.[5]

Evolution of the International Financial System

Before examining the impact of international financial transactions on monetary policy, we need to understand the past and current structure of the international financial system.

[3]Note that the capital account balance number reported above includes a statistical discrepancy item that represents errors due to unrecorded transactions involving smuggling and other capital flows (−$39 billion in 2001). Many experts believe that the statistical discrepancy item, which keeps the balance of payments in balance, is primarily the result of large hidden capital flows, and this is why it is included in the capital account balance reported above.

[4]The current account balance can also be viewed as showing the amount by which total saving exceeds private sector and government investment in the United States. We can see this by noting that total U.S. saving equals the increase in total wealth held by the U.S. private sector and government. Total investment equals the increase in the U.S. capital stock (wealth physically in the United States). The difference between them is the increase in U.S. claims on foreign wealth.

[5]If American residents have a greater preference for dollar assets than foreigners do, a movement of foreign wealth to American residents when there is a balance-of-payments surplus will increase the demand for dollar assets over time and will cause the dollar to appreciate.

Gold Standard

Before World War I, the world economy operated under the **gold standard**, meaning that the currency of most countries was convertible directly into gold. American dollar bills, for example, could be turned in to the U.S. Treasury and exchanged for approximately $\frac{1}{20}$ ounce of gold. Likewise, the British Treasury would exchange $\frac{1}{4}$ ounce of gold for £1 sterling. Because an American could convert $20 into 1 ounce of gold, which could be used to buy £4, the exchange rate between the pound and the dollar was effectively fixed at $5 to the pound. Tying currencies to gold resulted in an international financial system with fixed exchange rates between currencies. The fixed exchange rates under the gold standard had the important advantage of encouraging world trade by eliminating the uncertainty that occurs when exchange rates fluctuate.

To see how the gold standard operated in practice, let us see what occurs if, under the gold standard, the British pound begins to appreciate above the $5 par value. If an American importer of £100 of English tweed tries to pay for the tweed with dollars, it costs more than the $500 it cost before. Nevertheless, the importer has another option involving the purchase of gold that can reduce the cost of the tweed. Instead of using dollars to pay for the tweed, the American importer can exchange the $500 for gold, ship the gold to Britain, and convert it into £100. The shipment of gold to Britain is cheaper as long as the British pound is above the $5 par value (plus a small amount to pay for the cost of shipping the gold).

Under the gold standard, the appreciation of the pound leads to a British gain of international reserves (gold) and an equal U.S. loss. Because a change in a country's holdings of international reserves (gold) leads to an equal change in its monetary base, the movement of gold from the United States to Britain causes the British monetary base to rise and the American monetary base to fall. The resulting rise in the British money supply raises the British price level, while the fall in the U.S. money supply lowers the U.S. price level. The resulting increase in the British price level relative to the United States then causes the pound to depreciate. This process will continue until the value of the pound falls back down to its $5 par value.

A depreciation of the pound below the $5 par value, on the contrary, stimulates gold shipments from Britain to the United States. These shipments raise the American money supply and lower the British money supply, causing the pound to appreciate back toward the $5 par value. We thus see that under the gold standard, a rise or fall in the exchange rate sets in motion forces that return it to the par value.

As long as countries abided by the rules under the gold standard and kept their currencies backed by and convertible into gold, exchange rates remained fixed. However, adherence to the gold standard meant that a country had no control over its monetary policy, because its money supply was determined by gold flows between countries. Furthermore, monetary policy throughout the world was greatly influenced by the production of gold and gold discoveries. When gold production was low in the 1870s and 1880s, the money supply throughout the world grew slowly and did not keep pace with the growth of the world economy. The result was deflation (falling price levels). Gold discoveries in Alaska and South Africa in the 1890s then greatly expanded gold production, causing money supplies to increase rapidly and price levels to rise (inflation) until World War I.

Bretton Woods System

World War I caused massive trade disruptions. Countries could no longer convert their currencies into gold, and the gold standard collapsed. Despite attempts to revive it in the interwar period, the worldwide depression, beginning in 1929, led to its permanent

demise. As the Allied victory in World War II was becoming certain in 1944, the Allies met in Bretton Woods, New Hampshire, to develop a new international monetary system to promote world trade and prosperity after the war. In the agreement worked out among the Allies, central banks bought and sold their own currencies to keep their exchange rates fixed at a certain level (called a **fixed exchange rate regime**). The agreement lasted from 1945 to 1971 and was known as the **Bretton Woods system**.

The Bretton Woods agreement created the **International Monetary Fund (IMF)**, headquartered in Washington, D.C., which had 30 original member countries in 1945 and currently has over 180. The IMF was given the task of promoting the growth of world trade by setting rules for the maintenance of fixed exchange rates and by making loans to countries that were experiencing balance-of-payments difficulties. As part of its role of monitoring the compliance of member countries with its rules, the IMF also took on the job of collecting and standardizing international economic data.

The Bretton Woods agreement also set up the International Bank for Reconstruction and Development, commonly referred to as the **World Bank**, also headquartered in Washington, D.C., which provides long-term loans to help developing countries build dams, roads, and other physical capital that would contribute to their economic development. The funds for these loans are obtained primarily by issuing World Bank bonds, which are sold in the capital markets of the developed countries.[6]

Because the United States emerged from World War II as the world's largest economic power, with over half of the world's manufacturing capacity and the greater part of the world's gold, the Bretton Woods system of fixed exchange rates was based on the convertibility of U.S. dollars into gold (for foreign governments and central banks only) at $35 per ounce. The fixed exchange rates were to be maintained by intervention in the foreign exchange market by central banks in countries besides the United States who bought and sold dollar assets, which they held as international reserves. The U.S. dollar, which was used by other countries to denominate the assets that they held as international reserves, was called the **reserve currency**. Thus an important feature of the Bretton Woods system was the establishment of the United States as the reserve currency country. Even after the breakup of the Bretton Woods system, the U.S. dollar has kept its position as the reserve currency in which most international financial transactions are conducted. However, with the creation of the euro in 1999, the supremacy of the U.S. dollar may be subject to a serious challenge (see Box 2).

How a Fixed Exchange Rate Regime Works. The most important feature of the Bretton Woods system was that it set up a fixed exchange rate regime. Figure 2 shows how a fixed exchange rate regime works in practice using the model of exchange rate determination we learned in the previous chapter. Panel (a) describes a situation in which the domestic currency is initially overvalued: The schedule for the expected return on foreign deposits R_1^F intersects the schedule for the expected return on domestic deposits R_1^D at exchange rate E_1, which is lower than the par (fixed) value of the exchange rate E_{par}. To keep the exchange rate at E_{par}, the central bank must intervene in the foreign exchange market to purchase domestic currency by selling foreign assets, and this action, like an open market sale, means that the monetary base and

[6]In 1960, the World Bank established an affiliate, the International Development Association (IDA), which provides particularly attractive loans to third-world countries (with 50-year maturities and zero interest rates, for example). Funds for these loans are obtained by direct contributions of member countries.

Box 2: Global

The Euro's Challenge to the Dollar

With the creation of the European Monetary System and the euro in 1999, the U.S. dollar may face a challenge to its position as the key reserve currency in international financial transactions. Adoption of the euro increases integration of Europe's financial markets, which could help them rival those in the United States. The resulting increase in the use of euros in financial markets will make it more likely that international transactions are carried out in the euro. The economic clout of the European Union rivals that of the United States: Both have a similar share of world GDP (around 20%) and world exports (around 15%). If the European Central Bank can make sure that inflation remains low so that the euro becomes a sound currency, this should bode well for the euro.

However, for the euro to eat into the dollar's position as a reserve currency, the European Union must function as a cohesive political entity that is able to exert its influence on the world stage. There are serious doubts on this score, and most analysts think it will be a long time before the euro beats out the dollar in international financial transactions.

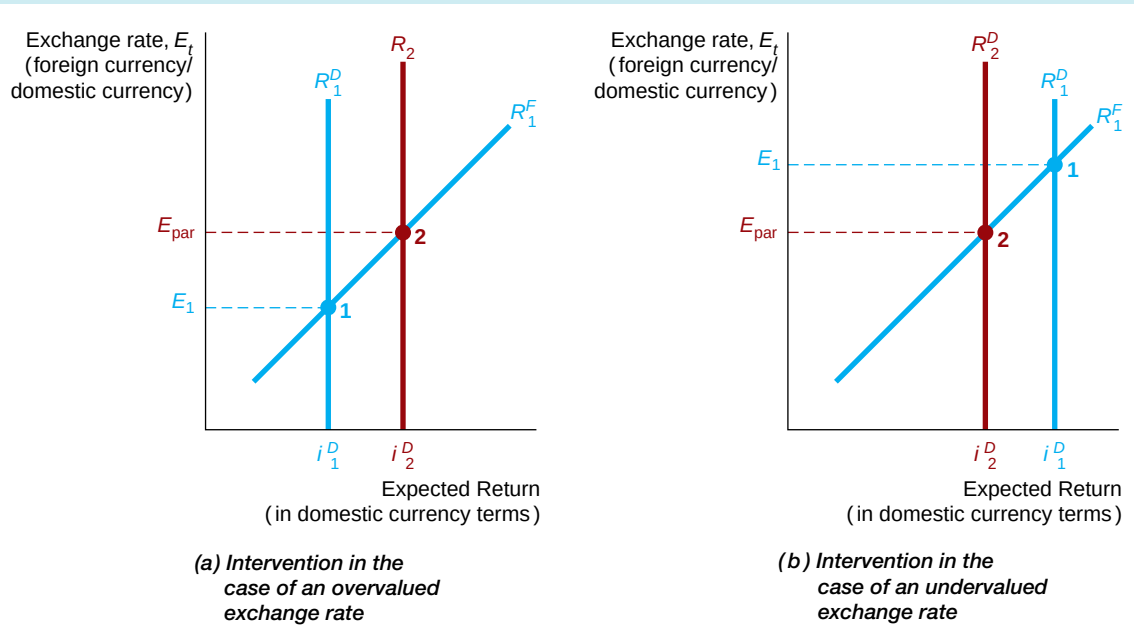


FIGURE 2 Intervention in the Foreign Exchange Market Under a Fixed Exchange Rate Regime

In panel (a), the exchange rate at E_{par} is overvalued. To keep the exchange rate at E_{par} (point 2), the central bank must purchase domestic currency to shift the schedule for the expected return on domestic deposits to R_2^D. In panel (b), the exchange rate at E_{par} is undervalued, so a central bank sale of domestic currency is needed to shift R^D to R_2^D to keep the exchange rate at E_{par} (point 2).

the money supply decline. Because the exchange rate will continue to be fixed at E_{par}, the expected future exchange rate remains unchanged, and so the schedule for the expected return on foreign deposits remains at R_1^F. However, the purchase of domestic

currency, which leads to a fall in the money supply, also causes the interest rate on domestic deposits i^D to rise. This increase in turn shifts the expected return on domestic deposits R^D to the right. The central bank will continue purchasing domestic currency and selling foreign assets until the R^D curve reaches R_2^D and the equilibrium exchange rate is at E_{par} at point 2 in panel (a).

We have thus come to the conclusion that ***when the domestic currency is overvalued, the central bank must purchase domestic currency to keep the exchange rate fixed, but as a result it loses international reserves.***

Panel (b) in Figure 2 shows how a central bank intervention keeps the exchange rate fixed at E_{par} when the exchange rate is initially undervalued; that is, when R_1^F and the initial R_1^D intersect at exchange rate E_1, which is above E_{par}. Here the central bank must sell domestic currency and purchase foreign assets, and this works like an open market purchase to raise the money supply and lower the interest rate on domestic deposits i^D. The central bank keeps selling domestic currency and lowers i^D until R^D shifts all the way to R_2^D, where the equilibrium exchange rate is at E_{par}—point 2 in panel (b). Our analysis thus leads us to the following result: ***When the domestic currency is undervalued, the central bank must sell domestic currency to keep the exchange rate fixed, but as a result, it gains international reserves.***

As we have seen, if a country's currency has an overvalued exchange rate, its central bank's attempts to keep the currency from depreciating will result in a loss of international reserves. If the country's central bank eventually runs out of international reserves, it cannot keep its currency from depreciating, and then a **devaluation** must occur, meaning that the par exchange rate is reset at a lower level.

If, by contrast, a country's currency has an undervalued exchange rate, its central bank's intervention to keep the currency from appreciating leads to a gain of international reserves. Because, as we will see shortly, the central bank might not want to acquire these international reserves, it might want to reset the par value of its exchange rate at a higher level (a **revaluation**).

Note that if domestic and foreign deposits are perfect substitutes, as is assumed in the model of exchange rate determination used here, a sterilized exchange rate intervention would not be able to keep the exchange rate at E_{par} because, as we have seen earlier in the chapter, neither R^F nor R^D will shift. For example, if the exchange rate is overvalued, a sterilized purchase of domestic currency will still leave the expected return on domestic deposits below the expected return on foreign deposits at the par exchange rate—so pressure for a depreciation of the domestic currency is not removed. If the central bank keeps on purchasing its domestic currency but continues to sterilize, it will just keep on losing international reserves until it finally runs out of them and is forced to let the value of the currency seek a lower level.

One implication of the foregoing analysis is that a country that ties its exchange rate to a larger country's currency loses control of its monetary policy. If the larger country pursues a more contractionary monetary policy and decreases its money supply, this would lead to lower expected inflation in the larger country, thus causing an appreciation of the larger country's currency and a depreciation of the smaller country's currency. The smaller country, having locked in its exchange rate, will now find its currency overvalued and will therefore have to sell the larger country's currency and buy its own to keep its currency from depreciating. The result of this foreign exchange intervention will then be a decline in the smaller country's international reserves, a contraction of the monetary base, and thus a decline in its money supply. Sterilization of this foreign exchange intervention is not an option because this would

just lead to a continuing loss of international reserves until the smaller country was forced to devalue. The smaller country no longer controls its monetary policy, because movements in its money supply are completely determined by movements in the larger country's money supply.

Another way to see that when a country fixes its exchange rate to a larger country's currency it loses control of its monetary policy is through the interest parity condition discussed in the previous chapter. There we saw that when there is capital mobility, the domestic interest rate equals the foreign interest rate minus the expected appreciation of the domestic currency. With a fixed exchange rate, expected appreciation of the domestic currency is zero, so that the domestic interest rate equals the foreign interest rate. Therefore changes in the monetary policy in the large country that affect its interest rate are directly transmitted to interest rates in the small country. Furthermore, because the monetary authorities in the small country cannot make their interest rate deviate from that of the larger country, they have no way to use monetary policy to affect their economy.

Bretton Woods System of Fixed Exchange Rates. Under the Bretton Woods system, exchange rates were supposed to change only when a country was experiencing a "fundamental disequilibrium"; that is, large persistent deficits or surpluses in its balance of payments. To maintain fixed exchange rates when countries had balance-of-payments deficits and were losing international reserves, the IMF would loan deficit countries international reserves contributed by other members. As a result of its power to dictate loan terms to borrowing countries, the IMF could encourage deficit countries to pursue contractionary monetary policies that would strengthen their currency or eliminate their balance-of-payments deficits. If the IMF loans were not sufficient to prevent depreciation of a currency, the country was allowed to devalue its currency by setting a new, lower exchange rate.

A notable weakness of the Bretton Woods system was that although deficit countries losing international reserves could be pressured into devaluing their currency or pursuing contractionary policies, the IMF had no way to force surplus countries to revise their exchange rates upward or pursue more expansionary policies. Particularly troublesome in this regard was the fact that the reserve currency country, the United States, could not devalue its currency under the Bretton Woods system even if the dollar was overvalued. When the United States attempted to reduce domestic unemployment in the 1960s by pursuing an inflationary monetary policy, a fundamental disequilibrium of an overvalued dollar developed. Because surplus countries were not willing to revise their exchange rates upward, adjustment in the Bretton Woods system did not take place, and the system collapsed in 1971. Attempts to patch up the Bretton Woods system with the Smithsonian Agreement in December 1971 proved unsuccessful, and by 1973, America and its trading partners had agreed to allow exchange rates to float.

Managed Float

Although exchange rates are currently allowed to change daily in response to market forces, central banks have not been willing to give up their option of intervening in the foreign exchange market. Preventing large changes in exchange rates makes it easier for firms and individuals purchasing or selling goods abroad to plan into the future. Furthermore, countries with surpluses in their balance of payments frequently do not want to see their currencies appreciate, because it makes their goods more expensive abroad and foreign goods cheaper in their country. Because an appreciation

might hurt sales for domestic businesses and increase unemployment, surplus countries have often sold their currency in the foreign exchange market and acquired international reserves.

Countries with balance-of-payments deficits do not want to see their currency lose value, because it makes foreign goods more expensive for domestic consumers and can stimulate inflation. To keep the value of the domestic currency high, deficit countries have often bought their own currency in the foreign exchange market and given up international reserves.

The current international financial system is a hybrid of a fixed and a flexible exchange rate system. Rates fluctuate in response to market forces but are not determined solely by them. Furthermore, many countries continue to keep the value of their currency fixed against other currencies, as was the case in the European Monetary System (to be described shortly).

www.imf.org/external/np/exr/facts/sdr.htm

Find information about special drawing rights, allocation, valuation, and SDR users' guide.

Another important feature of the current system is the continuing de-emphasis of gold in international financial transactions. Not only has the United States suspended convertibility of dollars into gold for foreign central banks, but since 1970 the IMF has been issuing a paper substitute for gold, called **special drawing rights (SDRs)**. Like gold in the Bretton Woods system, SDRs function as international reserves. Unlike gold, whose quantity is determined by gold discoveries and the rate of production, SDRs can be created by the IMF whenever it decides that there is a need for additional international reserves to promote world trade and economic growth.

The use of gold in international transactions was further de-emphasized by the IMF's elimination of the official gold price in 1975 and by the sale of gold by the U.S. Treasury and the IMF to private interests in order to demonetize it. Currently, the price of gold is determined in a free market. Investors who want to speculate in it are able to purchase and sell at will, as are jewelers and dentists who use gold in their businesses.

European Monetary System (EMS)

In March 1979, eight members of the European Economic Community (Germany, France, Italy, the Netherlands, Belgium, Luxembourg, Denmark, and Ireland) set up the European Monetary System (EMS), in which they agreed to fix their exchange rates vis-à-vis one another and to float jointly against the U.S. dollar. Spain joined the EMS in June 1989, the United Kingdom in October 1990, and Portugal in April 1992. The EMS created a new monetary unit, the *European currency unit* (ECU), whose value was tied to a basket of specified amounts of European currencies. Each member of the EMS was required to contribute 20% of its holdings of gold and dollars to the European Monetary Cooperation Fund and in return received an equivalent amount of ECUs.

The exchange rate mechanism (ERM) of the European Monetary System worked as follows. The exchange rate between every pair of currencies of the participating countries was not allowed to fluctuate outside narrow limits around a fixed exchange rate. (The limits were typically $\pm 2.25\%$ but were raised to $\pm 15\%$ in August 1993.) When the exchange rate between two countries' currencies moved outside these limits, the central banks of both countries were supposed to intervene in the foreign exchange market. If, for example, the French franc depreciated below its lower limit against the German mark, the Bank of France was required to buy francs and sell marks, thereby giving up international reserves. Similarly, the German central bank was also required to intervene to sell marks and buy francs and consequently increase its international reserves. The EMS thus required that intervention be symmetric

when a currency fell outside the limits, with the central bank with the weak currency giving up international reserves and the one with the strong currency gaining them. Central bank intervention was also very common even when the exchange rate was within the limits, but in this case, if one central bank intervened, no others were required to intervene as well.

A serious shortcoming of fixed exchange rate systems such as the Bretton Woods system or the European Monetary System is that they can lead to foreign exchange crises involving a "speculative attack" on a currency—massive sales of a weak currency or purchases of a strong currency to cause a sharp change in the exchange rate. In the following application, we use our model of exchange rate determination to understand how the September 1992 exchange rate crisis that rocked the European Monetary System came about.

Application

The Foreign Exchange Crisis of September 1992

In the aftermath of German reunification in October 1990, the German central bank, the Bundesbank, faced rising inflationary pressures, with inflation having accelerated from below 3% in 1990 to near 5% by 1992. To get monetary growth under control and to dampen inflation, the Bundesbank raised German interest rates to near double-digit levels. Figure 3 shows the consequences of these actions by the Bundesbank in the foreign exchange market for sterling. Note that in the diagram, the pound sterling is the domestic currency and R^D is the expected return on sterling deposits, whereas the foreign currency is the German mark (deutsche mark, DM), so R^F is the expected return on mark deposits.

The increase in German interest rates i^F shifted the R^F schedule rightward to R_2^F in Figure 3, so that the intersection of the R_1^D and the R_2^F schedules at point 1′ was below the lower exchange rate limit (2.778 marks per pound, denoted E_{par}) under the exchange rate mechanism. To lower the value of the mark relative to the pound and restore the pound/mark exchange rate to within the ERM limits, either the Bank of England had to pursue a contractionary monetary policy, thereby raising British interest rates to i_2^D and shifting the R_1^D schedule to the right to point 2, or the Bundesbank could pursue an expansionary monetary policy, thereby lowering German interest rates, which would shift the R^F schedule to the left to move back to point 1. (The shift in R^D to point 2 is not shown in the figure.)

The catch was that the Bundesbank, whose primary goal was fighting inflation, was unwilling to pursue an expansionary monetary policy, and the British, who were facing their worst recession in the postwar period, were unwilling to pursue a contractionary monetary policy to prop up the pound. This impasse became clear when in response to great pressure from other members of the EMS, the Bundesbank was willing to lower its lending rates by only a token amount on September 14 after a speculative attack was mounted on the currencies of the Scandinavian countries. So at some point in the near future, the value of the pound would have to decline to point 1′. Speculators now knew that the appreciation of the mark was imminent and

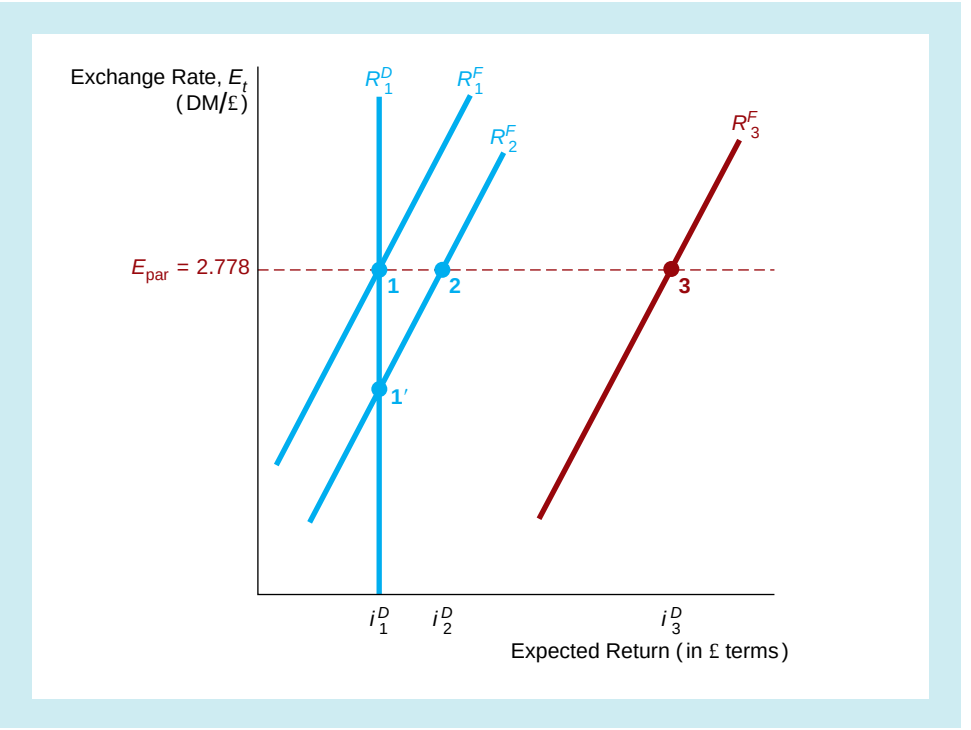


FIGURE 3 Foreign Exchange Market for British Pounds in 1992

The realization by speculators that the United Kingdom would soon devalue the pound increased the expected return on foreign (German mark, DM) deposits and shifted R_2^F rightward to R_3^F. The result was the need for a much greater purchase of pounds by the British central bank to raise the interest rate to i_3^D to keep the exchange rate at 2.778 German marks per pound.

hence that the value of foreign (mark) deposits would rise in value relative to the pound. As a result, the expected return on mark deposits increased sharply, shifting the R^F schedule to R_3^F in Figure 3.

The huge potential losses on pound deposits and potential gains on mark deposits caused a massive sell-off of pounds (and purchases of marks) by speculators. The need for the British central bank to intervene to raise the value of the pound now became much greater and required a huge rise in British interest rates, all the way to i_3^D. After a major intervention effort on the part of the Bank of England, which included a rise in its lending rate from 10% to 15%, which still wasn't enough, the British were finally forced to give up on September 16: They pulled out of the ERM indefinitely, allowing the pound to depreciate by 10% against the mark.

Speculative attacks on other currencies forced devaluation of the Spanish peseta by 5% and the Italian lira by 15%. To defend its currency, the Swedish central bank was forced to raise its daily lending rate to the astronomical level of 500%! By the time the crisis was over, the British, French, Italian, Spanish, and Swedish central banks had intervened to the tune of $100 billion; the Bundesbank alone had laid out $50 billion for foreign exchange intervention. Because foreign exchange crises lead to large changes in central banks' holdings of international reserves and thus affect the official reserve asset items in the balance of payments, these crises are also referred to as **balance-of-payments crises**.

The attempt to prop up the European Monetary System was not cheap for these central banks. It is estimated that they lost $4 to $6 billion as a result of

exchange rate intervention during the crisis. What the central banks lost, the speculators gained. A speculative fund run by George Soros ran up $1 billion of profits during the crisis, and Citibank traders are reported to have made $200 million. When an exchange rate crisis comes, life can certainly be sweet for exchange rate speculators.

Application

Recent Foreign Exchange Crises in Emerging Market Countries: Mexico 1994, East Asia 1997, Brazil 1999, and Argentina 2002

Major currency crises in emerging market countries have been a common occurrence in recent years. We can use Figure 3 to understand the sequence of events during the currency crises in Mexico in 1994, East Asia in 1997, Brazil in 1999, and Argentina in 2002. To do so, we just need to recognize that dollars are the foreign currency, so that R^F is the expected return on dollar deposits, while R^D is the expected return on deposits denominated in the domestic currency, whether it was pesos, baht, or reals. (Note that the exchange rate label on the vertical axis would be in terms of dollars/domestic currency and that the label on the horizontal axis would be expected return in the domestic currency—say, pesos.)

In Mexico in March 1994, political instability (the assassination of the ruling party's presidential candidate) sparked investors' concerns that the peso might be devalued. The result was that the expected return on dollar deposits rose, thus moving the R^F schedule from R_1^F to R_2^F in Figure 3. In the case of Thailand in May 1997, the large current account deficit and the weakness of the Thai financial system raised similar concerns about the devaluation of the domestic currency, with the same effect on the R^F schedule. In Brazil in late 1998 and Argentina in 2001, concerns about fiscal situations that could lead to the printing of money to finance the deficit, and thereby raise inflation, also meant that a devaluation was more likely to occur. The concerns thus raised the expected return on dollar deposits and shifted the R^F schedule from R_1^F to R_2^F. In all of these cases, the result was that the intersection of the R_1^D and R_2^F curves was below the pegged value of the domestic currency at E_{par}.

To keep their domestic currencies from falling below E_{par}, these countries' central banks needed to buy the domestic currency and sell dollars to raise interest rates to i_2^D and shift the R^D curve to the right, in the process losing international reserves. At first, the central banks were successful in containing this speculative attack. However, when more bad news broke, speculators became even more confident that these countries could not defend their currencies. (The bad news was everywhere: in Mexico, with an uprising in Chiappas and revelations about problems in the banking system; in Thailand, there was a major failure of a financial institution; in Brazil, a worsening fiscal situation was reported, along with a threat by a governor to default on his state's debt; and in Argentina, a full-scale bank panic and an actual default on the government debt occurred.) As a result, the expected returns on dollar deposits shot up further, and R^F moved much farther to the right, to R_3^F; and

the central banks lost even more international reserves. Given the stress on the economy from rising interest rates and the loss of reserves, eventually the monetary authorities could no longer continue to defend the currency and were forced to give up and let their currencies depreciate. This scenario happened in Mexico in December 1994, in Thailand in July 1997, in Brazil in January 1999, and in Argentina in January 2002.

Concerns about similar problems in other countries then triggered speculative attacks against them as well. This contagion occurred in the aftermath of the Mexican crisis (jauntily referred to as the "Tequila effect") with speculative attacks on other Latin American currencies, but there were no further currency collapses. In the East Asian crisis, however, fears of devaluation spread throughout the region, leading to a scenario akin to that depicted in Figure 3. Consequently, one by one, Indonesia, Malaysia, South Korea, and the Philippines were forced to devalue sharply. Even Hong Kong, Singapore, and Taiwan were subjected to speculative attacks, but because these countries had healthy financial systems, the attacks were successfully averted.

As we saw in Chapter 8, the sharp depreciations in Mexico, East Asia, and Argentina led to full-scale financial crises that severely damaged these countries' economies. The foreign exchange crisis that shocked the European Monetary System in September 1992 cost central banks a lot of money, but the public in European countries were not seriously affected. By contrast, the public in Mexico, Argentina, and the crisis countries of East Asia were not so lucky: The collapse of these currencies triggered by speculative attacks led to the financial crises described in Chapter 8, producing severe depressions that caused hardship and political unrest.

Capital Controls

Because capital flows have been an important element in the currency crises in Mexico and East Asia, politicians and some economists have advocated that capital mobility in emerging market countries should be restricted with capital controls in order to avoid financial instability. Are capital controls a good idea?

Controls on Capital Outflows

Capital outflows can promote financial instability in emerging market countries, because when domestic residents and foreigners pull their capital out of a country, the resulting capital outflow forces a country to devalue its currency. This is why some politicians in emerging market countries have recently found capital controls particularly attractive. For example, Prime Minister Mahathir of Malaysia instituted capital controls in 1998 to restrict outflows in the aftermath of the East Asian crisis.

Although these controls sound like a good idea, they suffer from several disadvantages. First, empirical evidence indicates that controls on capital outflows are seldom effective during a crisis because the private sector finds ingenious ways to evade them and has little difficulty moving funds out of the country.[7] Second, the evidence

[7] See Sebastian Edwards, "How Effective are Capital Controls?" *Journal of Economic Perspectives*, Winter 2000; vol. 13, no. 4, pp. 65–84.

suggests that capital flight may even increase after controls are put into place, because confidence in the government is weakened. Third, controls on capital outflows often lead to corruption, as government officials get paid off to look the other way when domestic residents are trying to move funds abroad. Fourth, controls on capital outflows may lull governments into thinking they do not have to take the steps to reform their financial systems to deal with the crisis, with the result that opportunities are lost to improve the functioning of the economy.

Controls on Capital Inflows

Although most economists find the arguments against controls on capital outflows persuasive, controls on capital inflows receive more support. Supporters reason that if speculative capital cannot come in, then it cannot go out suddenly and create a crisis. Our analysis of the financial crises in East Asia in Chapter 8 provides support for this view by suggesting that capital inflows can lead to a lending boom and excessive risk taking on the part of banks, which then helps trigger a financial crisis.

However, controls on capital inflows have the undesirable feature that they may block from entering a country funds that would be used for productive investment opportunities. Although such controls may limit the fuel supplied to lending booms through capital flows, over time they produce substantial distortions and misallocation of resources as households and businesses try to get around them. Indeed, just as with controls on capital outflows, controls on capital inflows can lead to corruption. There are serious doubts whether capital controls can be effective in today's environment, in which trade is open and where there are many financial instruments that make it easier to get around these controls.

On the other hand, there is a strong case for improving bank regulation and supervision so that capital inflows are less likely to produce a lending boom and encourage excessive risk taking by banking institutions. For example, restricting banks in how fast their borrowing could grow might have the impact of substantially limiting capital inflows. Supervisory controls of this type, focusing on the sources of financial fragility rather than the symptoms, can enhance the efficiency of the financial system rather than hampering it.

The Role of the IMF

The International Monetary Fund was originally set up under the Bretton Woods system to help countries deal with balance-of-payments problems and stay with the fixed exchange rate by lending to deficit countries. With the collapse of the Bretton Woods system of fixed exchange rates in 1971, the IMF has taken on new roles.

The IMF continues to function as a data collector and provide technical assistance to its member countries. Although the IMF no longer attempts to encourage fixed exchange rates, its role as an international lender has become more important recently. This role first came to the fore in the 1980s during the third-world debt crisis, in which the IMF assisted developing countries in repaying their loans. The financial crises in Mexico in 1994–1995 and in East Asia in 1997–1998 led to huge loans by the IMF to these and other affected countries to help them recover from their financial crises and to prevent the spread of these crises to other countries. This role, in which the IMF acts like an international lender of last resort to cope with financial instability, is indeed highly controversial.

Should the IMF Be an International Lender of Last Resort?

As we saw in Chapter 17, in industrialized countries when a financial crisis occurs and the financial system threatens to seize up, domestic central banks can address matters with a lender-of-last-resort operation to limit the degree of instability in the banking system. In emerging market countries, however, where the credibility of the central bank as an inflation-fighter may be in doubt and debt contracts are typically short-term and denominated in foreign currencies, a lender-of-last-resort operation becomes a two-edged sword—as likely to exacerbate the financial crisis as to alleviate it. For example, when the U.S. Federal Reserve engaged in a lender-of-last-resort operation during the 1987 stock market crash (Chapter 17), there was almost no sentiment in the markets that there would be substantially higher inflation. However, for a central bank having less inflation-fighting credibility than the Fed, central bank lending to the financial system in the wake of a financial crisis—even under the lender-of-last-resort rhetoric—may well arouse fears of inflation spiraling out of control, causing an even greater currency depreciation and still greater deterioration of balance sheets. The resulting increase in moral hazard and adverse selection problems in financial markets, along the lines discussed in Chapter 8, would only worsen the financial crisis.

Central banks in emerging market countries therefore have only a very limited ability to successfully engage in a lender-of-last-resort operation. However, liquidity provided by an international lender of last resort does not have these undesirable consequences, and in helping to stabilize the value of the domestic currency, it strengthens domestic balance sheets. Moreover, an international lender of last resort may be able to prevent contagion, the situation in which a successful speculative attack on one emerging market currency leads to attacks on other emerging market currencies, spreading financial and economic disruption as it goes. Since a lender of last resort for emerging market countries is needed at times, and since it cannot be provided domestically, there is a strong rationale for an international institution to fill this role. Indeed, since Mexico's financial crisis in 1994, the International Monetary Fund and other international agencies have stepped into the lender-of-last-resort role and provided emergency lending to countries threatened by financial instability.

However, support from an international lender of last resort brings risks of its own, especially the risk that the perception it is standing ready to bail out irresponsible financial institutions may lead to excessive risk taking of the sort that makes financial crises more likely. In the Mexican and East Asian crises, governments in the crisis countries used IMF support to protect depositors and other creditors of banking institutions from losses. This safety net creates a well-known moral hazard problem because the depositors and other creditors have less incentive to monitor these banking institutions and withdraw their deposits if the institutions are taking on too much risk. The result is that these institutions are encouraged to take on excessive risks. Indeed, critics of the IMF—most prominently, the Congressional Commission headed by Professor Alan Meltzer of Carnegie-Mellon University—contend that IMF lending in the Mexican crisis, which was used to bail out foreign lenders, set the stage for the East Asian crisis, because these lenders expected to be bailed out if things went wrong, and thus provided funds that were used to fuel excessive risk taking.[8]

An international lender of last resort must find ways to limit this moral hazard problem, or it can actually make the situation worse. The international lender of last resort can make it clear that it will extend liquidity only to governments that put the

[8] See International Financial Institution Advisory Commission, *Report* (IFIAC: Washington, D.C., 2000).

proper measures in place to prevent excessive risk taking. In addition, it can reduce the incentives for risk taking by restricting the ability of governments to bail out stockholders and large uninsured creditors of domestic financial institutions. Some critics of the IMF believe that the IMF has not put enough pressure on the governments to which it lends to contain the moral hazard problem.

One problem that arises for international organizations like the IMF engaged in lender-of-last-resort operations is that they know that if they don't come to the rescue, the emerging market country will suffer extreme hardship and possible political instability. Politicians in the crisis country may exploit these concerns and engage in a game of chicken with the international lender of last resort: They resist necessary reforms, hoping that the IMF will cave in. Elements of this game were present in the Mexican crisis of 1995 and were also a particularly important feature of the negotiations between the IMF and Indonesia during the Asian crisis.

The IMF would produce better outcomes if it made clear that it will not play this game. Just as giving in to ill-behaved children may be the easy way out in the short run, but supports a pattern of poor behavior in the long run, some critics worry that the IMF may not be tough enough when confronted by short-run humanitarian concerns. For example, these critics have been particularly critical of the IMF's lending to the Russian government, which resisted adopting appropriate reforms to stabilize its financial system.

The IMF has also been criticized for imposing on the East Asian countries so-called austerity programs that focus on tight macroeconomic policies rather than on microeconomic policies to fix the crisis-causing problems in the financial sector. Such programs are likely to increase resistance to IMF recommendations, particularly in emerging market countries. Austerity programs allow these politicians to label institutions such as the IMF as being anti-growth, rhetoric that helps the politicians mobilize the public against the IMF and avoid doing what they really need to do to reform the financial system in their country. IMF programs focused instead on microeconomic policies related to the financial sector would increase the likelihood that the IMF will be seen as a helping hand in the creation of a more efficient financial system.

An important historical feature of successful lender-of-last-resort operations is that the faster the lending is done, the lower is the amount that actually has to be lent. An excellent example occurred in the aftermath of the stock market crash on October 19, 1987 (Chapter 17). At the end of that day, in order to service their customers' accounts, securities firms needed to borrow several billion dollars to maintain orderly trading. However, given the unprecedented developments, banks were very nervous about extending further loans to these firms. Upon learning this, the Federal Reserve engaged in an immediate lender-of-last-resort operation, with the Fed making it clear that it would provide liquidity to banks making loans to the securities industry. Indeed, what is striking about this episode is that the extremely quick intervention of the Fed not only resulted in a negligible impact of the stock market crash on the economy, but also meant that the amount of liquidity that the Fed needed to supply to the economy was not very large.

The ability of the Fed to engage in a lender-of-last-resort operation within a day of a substantial shock to the financial system is in sharp contrast to the amount of time it has taken the IMF to supply liquidity during the recent crises in emerging market countries. Because IMF lending facilities were originally designed to provide funds after a country was experiencing a balance-of-payments crisis and because the conditions for the loan had to be negotiated, it took several months before the IMF made

funds available. By this time, the crises had gotten much worse—and much larger sums of funds were needed to cope with the crisis, often stretching the resources of the IMF. One reason central banks can lend so much more quickly than the IMF is that they have set up procedures in advance to provide loans, with the terms and conditions for this lending agreed upon beforehand. The need for quick provision of liquidity, to keep the loan amount manageable, argues for similar credit facilities at the international lender of last resort, so that funds can be provided quickly, as long as the borrower meets conditions such as properly supervising its banks or keeping budget deficits low. A step in this direction was made in 1999 when the IMF set up a new lending facility, the Contingent Credit Line, so that it can provide liquidity faster during a crisis.

The debate on whether the world will be better off with the IMF operating as an international lender of last resort is currently a hot one. Much attention is being focused on making the IMF more effective in performing this role, and redesign of the IMF is at the center of proposals for a new international financial architecture to help reduce international financial instability.

International Considerations and Monetary Policy

Our analysis in this chapter so far has suggested several ways in which monetary policy can be affected by international matters. Awareness of these effects can have significant implications for the way monetary policy is conducted.

Direct Effects of the Foreign Exchange Market on the Money Supply

When central banks intervene in the foreign exchange market, they acquire or sell off international reserves, and their monetary base is affected. When a central bank intervenes in the foreign exchange market, it gives up some control of its money supply. For example, in the early 1970s, the German central bank faced a dilemma. In attempting to keep the German mark from appreciating too much against the U.S. dollar, the Germans acquired huge quantities of international reserves, leading to a rate of money growth that the German central bank considered inflationary.

The Bundesbank could have tried to halt the growth of the money supply by stopping its intervention in the foreign exchange market and reasserting control over its own money supply. Such a strategy has a major drawback when the central bank is under pressure not to allow its currency to appreciate: The lower price of imports and higher price of exports as a result of an appreciation in its currency will hurt domestic producers and increase unemployment.

Because the U.S. dollar has been a reserve currency, the U.S. monetary base and money supply have been less affected by developments in the foreign exchange market. As long as foreign central banks, rather than the Fed, intervene to keep the value of the dollar from changing, American holdings of international reserves are unaffected. The ability to conduct monetary policy is typically easier when a country's currency is a reserve currency.[9]

[9]However, the central bank of a reserve currency country must worry about a shift away from the use of its currency for international reserves.

Balance-of-Payments Considerations

Under the Bretton Woods system, balance-of-payments considerations were more important than they are under the current managed float regime. When a non-reserve currency country is running balance-of-payments deficits, it necessarily gives up international reserves. To keep from running out of these reserves, under the Bretton Woods system it had to implement contractionary monetary policy to strengthen its currency—exactly what occurred in the United Kingdom before its devaluation of the pound in 1967. When policy became expansionary, the balance of payments deteriorated, and the British were forced to "slam on the brakes" by implementing a contractionary policy. Once the balance of payments improved, policy became more expansionary until the deteriorating balance of payments again forced the British to pursue a contractionary policy. Such on-again, off-again actions became known as a "stop-go" policy, and the domestic instability it created was criticized severely.

Because the United States is a major reserve currency country, it can run large balance-of-payments deficits without losing huge amounts of international reserves. This does not mean, however, that the Federal Reserve is never influenced by developments in the U.S. balance of payments. Current account deficits in the United States suggest that American businesses may be losing some of their ability to compete because the value of the dollar is too high. In addition, large U.S. balance-of-payments deficits lead to balance-of-payments surpluses in other countries, which can in turn lead to large increases in their holdings of international reserves (this was especially true under the Bretton Woods system). Because such increases put a strain on the international financial system and may stimulate world inflation, the Fed worries about U.S. balance-of-payments and current account deficits. To help shrink these deficits, the Fed might pursue a more contractionary monetary policy.

Exchange Rate Considerations

Unlike balance-of-payments considerations, which have become less important under the current managed float system, exchange rate considerations now play a greater role in the conduct of monetary policy. If a central bank does not want to see its currency fall in value, it may pursue a more contractionary monetary policy of reducing the money supply to raise the domestic interest rate, thereby strengthening its currency. Similarly, if a country experiences an appreciation in its currency, domestic industry may suffer from increased foreign competition and may pressure the central bank to pursue a higher rate of money growth in order to lower the exchange rate.

The pressure to manipulate exchange rates seems to be greater for central banks in countries other than the United States, but even the Federal Reserve is not completely immune. The growing tide of protectionism stemming from the inability of American firms to compete with foreign firms because of the strengthening dollar from 1980 to early 1985 stimulated congressional critics of the Fed to call for a more expansionary monetary policy to lower the value of the dollar. As we saw in Chapter 18, the Fed then did let money growth surge to very high levels. A policy to bring the dollar down was confirmed in the Plaza Agreement of September 1985, in which the finance ministers from the five most important industrial nations in the free world (the United States, Japan, West Germany, the United Kingdom, and France) agreed to intervene in foreign exchange markets to achieve a decline in the dollar. The dollar continued to fall rapidly after the Plaza Agreement, and the Fed played an important role in this decline by continuing to expand the money supply at a rapid rate.

Summary

1. An unsterilized central bank intervention in which the domestic currency is sold to purchase foreign assets leads to a gain in international reserves, an increase in the money supply, and a depreciation of the domestic currency. Available evidence suggests, however, that sterilized central bank interventions have little long-term effect on the exchange rate.
2. The balance of payments is a bookkeeping system for recording all payments between a country and foreign countries that have a direct bearing on the movement of funds between them. The official reserve transactions balance is the sum of the current account balance plus the items in the capital account. It indicates the amount of international reserves that must be moved between countries to finance international transactions.
3. Before World War I, the gold standard was predominant. Currencies were convertible into gold, thus fixing exchange rates between countries. After World War II, the Bretton Woods system and the IMF were established to promote a fixed exchange rate system in which the U.S. dollar was convertible into gold. The Bretton Woods system collapsed in 1971. We now have an international financial system that has elements of a managed float and a fixed exchange rate system. Some exchange rates fluctuate from day to day, although central banks intervene in the foreign exchange market, while other exchange rates are fixed.
4. Controls on capital outflows receive support because they may prevent domestic residents and foreigners from pulling capital out of a country during a crisis and make devaluation less likely. Controls on capital inflows make sense under the theory that if speculative capital cannot flow in, then it cannot go out suddenly and create a crisis. However, capital controls suffer from several disadvantages: they are seldom effective, they lead to corruption, and they may allow governments to avoid taking the steps needed to reform their financial systems to deal with the crisis.
5. The IMF has recently taken on the role of an international lender of last resort. Because central banks in emerging market countries are unlikely to be able to perform a lender-of-last-resort operation successfully, an international lender of last resort like the IMF is needed to prevent financial instability. However, the IMF's role as an international lender of last resort creates a serious moral hazard problem that can encourage excessive risk taking and make a financial crisis more likely, but avoiding the problem may be politically hard to do. In addition, it needs to be able to provide liquidity quickly during a crisis in order to keep manageable the amount of funds lent.
6. Three international considerations affect the conduct of monetary policy: direct effects of the foreign exchange market on the money supply, balance-of-payments considerations, and exchange rate considerations. Inasmuch as the United States has been a reserve currency country in the post–World War II period, U.S. monetary policy has been less affected by developments in the foreign exchange market and its balance of payments than is true for other countries. However, in recent years, exchange rate considerations have been playing a more prominent role in influencing U.S. monetary policy.

Key Terms

balance of payments, p. 467
balance-of-payments crisis, p. 476
Bretton Woods system, p. 470
capital account, p. 467
current account, p. 467
devaluation, p. 472
fixed exchange rate regime, p. 470
foreign exchange intervention, p. 462
gold standard, p. 469
International Monetary Fund (IMF), p. 470
international reserves, p. 462
managed float regime (dirty float), p. 462
official reserve transactions balance, p. 468
reserve currency, p. 470
revaluation, p. 472

special drawing rights (SDRs), p. 474
sterilized foreign exchange intervention, p. 465
trade balance, p. 467
unsterilized foreign exchange intervention, p. 464
World Bank, p. 470

Questions and Problems

Questions marked with an asterisk are answered at the end of the book in an appendix, "Answers to Selected Questions and Problems."

1. If the Federal Reserve buys dollars in the foreign exchange market but conducts an offsetting open market operation to sterilize the intervention, what will be the impact on international reserves, the money supply, and the exchange rate?

*2. If the Federal Reserve buys dollars in the foreign exchange market but does not sterilize the intervention, what will be the impact on international reserves, the money supply, and the exchange rate?

3. For each of the following, identify in which part of the balance-of-payments account it appears (current account, capital account, or method of financing) and whether it is a receipt or a payment.
   a. A British subject's purchase of a share of Johnson & Johnson stock
   b. An American's purchase of an airline ticket from Air France
   c. The Swiss government's purchase of U.S. Treasury bills
   d. A Japanese's purchase of California oranges
   e. $50 million of foreign aid to Honduras
   f. A loan by an American bank to Mexico
   g. An American bank's borrowing of Eurodollars

*4. Why does a balance-of-payments deficit for the United States have a different effect on its international reserves than a balance-of-payments deficit for the Netherlands?

5. Under the gold standard, if Britain became more productive relative to the United States, what would happen to the money supply in the two countries? Why would the changes in the money supply help preserve a fixed exchange rate between the United States and Britain?

*6. What is the exchange rate between dollars and Swiss francs if one dollar is convertible into $\frac{1}{20}$ ounce of gold and one franc is convertible into $\frac{1}{40}$ ounce of gold?

7. If a country's par exchange rate was undervalued during the Bretton Woods fixed exchange rate regime, what kind of intervention would that country's central bank be forced to undertake, and what effect would it have on its international reserves and the money supply?

*8. How can a large balance-of-payments surplus contribute to the country's inflation rate?

9. "If a country wants to keep its exchange rate from changing, it must give up some control over its money supply." Is this statement true, false, or uncertain? Explain your answer.

*10. Why can balance-of-payments deficits force some countries to implement a contractionary monetary policy?

11. "Balance-of-payments deficits always cause a country to lose international reserves." Is this statement true, false, or uncertain? Explain your answer.

*12. How can persistent U.S. balance-of-payments deficits stimulate world inflation?

13. "Inflation is not possible under the gold standard." Is this statement true, false, or uncertain? Explain your answer.

*14. Why is it that in a pure flexible exchange rate system, the foreign exchange market has no direct effects on the money supply? Does this mean that the foreign exchange market has no effect on monetary policy?

15. "The abandonment of fixed exchange rates after 1973 has meant that countries have pursued more independent monetary policies." Is this statement true, false, or uncertain? Explain your answer.

Web Exercises

1. The Federal Reserve publishes information online that explains the workings of the foreign exchange market. One such publication can be found at www.ny.frb.org/pihome/addpub/usfxm/. Review the table of contents and open Chapter 10, the evolution of the international monetary system. Read this chapter and write a one-page summary that discusses why each monetary standard was dropped in favor of the succeeding one.

2. The International Monetary Fund stands ready to help nations facing monetary crises. Go to www.imf.org. Click on the tab labeled "About IMF." What is the stated purpose of the IMF? How many nations participate and when was it established?

Chapter 21

Monetary Policy Strategy: The International Experience

PREVIEW

Getting monetary policy right is crucial to the health of the economy. Overly expansionary monetary policy leads to high inflation, which decreases the efficiency of the economy and hampers economic growth. The United States has not been exempt from inflationary episodes, but more extreme cases of inflation, in which the inflation rate climbs to over 100% per year, have been prevalent in some regions of the world such as Latin America, and have been very harmful to the economy. Monetary policy that is too tight can produce serious recessions in which output falls and unemployment rises. It can also lead to deflation, a fall in the price level, as occurred in the United States during the Great Depression and in Japan more recently. As we have seen in Chapter 8, deflation can be especially damaging to the economy, because it promotes financial instability and can even help trigger financial crises.

www.federalreserve.gov/centralbanks.htm

Features links to home pages for central banks around the world.

In Chapter 18 our discussion of the conduct of monetary policy focused primarily on the United States. However, the United States is not the source of all wisdom about how to do monetary policy well. In thinking about what strategies for the conduct of monetary policy might be best, we need to examine monetary policy experiences in other countries.

A central feature of monetary policy strategies in all countries is the use of a **nominal anchor** (a nominal variable that monetary policymakers use to tie down the price level such as the inflation rate, an exchange rate, or the money supply) as an intermediate target to achieve an ultimate goal such as price stability. We begin the chapter by examining the role a nominal anchor plays in promoting price stability. Then we examine three basic types of monetary policy strategy—exchange-rate targeting, monetary targeting, and inflation targeting—and compare them to the Federal Reserve's current monetary policy strategy, which features an implicit (not an explicit) nominal anchor. We will see that despite the recent excellent performance of monetary policy in the United States, there is much to learn from the foreign experience.

The Role of a Nominal Anchor

Adherence to a nominal anchor forces a nation's monetary authority to conduct monetary policy so that the nominal anchor variable such as the inflation rate or the money supply stays within a narrow range. A nominal anchor thus keeps the price level from growing or falling too fast and thereby preserves the value of a country's

money. Thus, a nominal anchor of some sort is a necessary element in successful monetary policy strategies.

One reason a nominal anchor is necessary for monetary policy is that it can help promote price stability, which most countries now view as the most important goal for monetary policy. A nominal anchor promotes price stability by tying inflation expectations to low levels directly through its constraint on the value of domestic money. A more subtle reason for a nominal anchor's importance is that it can limit the **time-consistency problem**, in which monetary policy conducted on a discretionary, day-by-day basis leads to poor long-run outcomes.[1]

The Time-Consistency Problem

The time-consistency problem of discretionary policy arises because economic behavior is influenced by what firms and people expect the monetary authorities to do in the future. With firms' and people's expectations assumed to remain unchanged, policymakers think they can boost economic output (or lower unemployment) by pursuing discretionary monetary policy that is more expansionary than expected, and so they have incentives to pursue this policy. This situation is described by saying that discretionary monetary policy is *time-consistent*; that is, the policy is what policymakers are likely to want to pursue at any given point in time. The problem with time-consistent, discretionary policy is that it leads to bad outcomes. Because decisions about wages and prices reflect expectations about policy, workers and firms will raise their expectations not only of inflation but also of wages and prices. On average, output will not be higher under such an expansionary strategy, but inflation will be. (We examine this more formally in Chapter 28.)

Clearly, a central bank will do better if it does not try to boost output by surprising people with an unexpectedly expansionary policy, but instead keeps inflation under control. However, even if a central bank recognizes that discretionary policy will lead to a poor outcome—high inflation with no gains on the output front—it may still fall into the time-consistency trap, because politicians are likely to apply pressure on the central bank to try to boost output with overly expansionary monetary policy.

Although the analysis sounds somewhat complicated, the time-consistency problem is actually something we encounter in everyday life. For example, at any given point in time, it seems to make sense for a parent to give in to a child to keep the child from acting up. The more the parent gives in, however, the more the demanding the child is likely to become. Thus, the discretionary time-consistent actions by the parent lead to a bad outcome—a very spoiled child—because the child's expectations *are* affected by what the parent does. How-to books on parenting suggest a solution to the time-consistency problem (although they don't call it that) by telling parents that they should set rules for their children and stick to them.

A nominal anchor is like a behavior rule. Just as rules help to prevent the time-consistency problem in parenting, a nominal anchor can help to prevent the time-

[1]The time-consistency problem is also called the time-inconsistency problem because monetary policy that leads to a good outcome by controlling inflation is not sustainable (and is thus time-inconsistent). When the central bank pursues such a policy, it has incentives to deviate from it to try to boost output by engaging in discretionary, time-consistent policy. The time-consistency problem was first outlined in Finn Kydland and Edward Prescott, "Rules Rather Than Discretion: The Inconsistency of Optimal Plans," *Journal of Political Economy* 85 (1977): 473–491; Guillermo Calvo, "On the Time Consistency of Optimal Policy in the Monetary Economy," *Econometrica* 46 (November 1978): 1411–1428; and Robert J. Barro and David Gordon, "A Positive Theory of Monetary Policy in a Natural Rate Model," *Journal of Political Economy* 91 (August 1983).

consistency problem in monetary policy by providing an expected constraint on discretionary policy. In the following sections, we examine three monetary policy strategies—exchange-rate targeting, monetary targeting, and inflation targeting—that use a nominal anchor.

Exchange-Rate Targeting

Targeting the exchange rate is a monetary policy strategy with a long history. It can take the form of fixing the value of the domestic currency to a commodity such as gold, the key feature of the gold standard described in Chapter 20. More recently, fixed exchange-rate regimes have involved fixing the value of the domestic currency to that of a large, low-inflation country like the United States or Germany (called the *anchor country*). Another alternative is to adopt a crawling target or peg, in which a currency is allowed to depreciate at a steady rate so that the inflation rate in the pegging country can be higher than that of the anchor country.

Advantages of Exchange-Rate Targeting

Exchange-rate targeting has several advantages. First, the nominal anchor of an exchange-rate target directly contributes to keeping inflation under control by tying the inflation rate for internationally traded goods to that found in the anchor country. It does this because the foreign price of internationally traded goods is set by the world market, while the domestic price of these goods is fixed by the exchange-rate target. For example, until 2002 in Argentina the exchange rate for the Argentine peso was exactly one to the dollar, so that a bushel of wheat traded internationally at five dollars had its price set at five pesos. If the exchange-rate target is credible (i.e., expected to be adhered to), the exchange-rate target has the added benefit of anchoring inflation expectations to the inflation rate in the anchor country.

Second, an exchange-rate target provides an automatic rule for the conduct of monetary policy that helps mitigate the time-consistency problem. As we saw in Chapter 20, an exchange-rate target forces a tightening of monetary policy when there is a tendency for the domestic currency to depreciate or a loosening of policy when there is a tendency for the domestic currency to appreciate, so that discretionary, time-consistent monetary policy is less of an option.

Third, an exchange-rate target has the advantage of simplicity and clarity, which makes it easily understood by the public. A "sound currency" is an easy-to-understand rallying cry for monetary policy. In the past, this aspect was important in France, where an appeal to the "franc fort" (strong franc) was often used to justify tight monetary policy.

Given its advantages, it is not surprising that exchange-rate targeting has been used successfully to control inflation in industrialized countries. Both France and the United Kingdom, for example, successfully used exchange-rate targeting to lower inflation by tying the value of their currencies to the German mark. In 1987, when France first pegged its exchange rate to the mark, its inflation rate was 3%, two percentage points above the German inflation rate. By 1992, its inflation rate had fallen to 2%, a level that can be argued is consistent with price stability, and was even below that in Germany. By 1996, the French and German inflation rates had converged, to a number slightly below 2%. Similarly, after pegging to the German mark in 1990, the United Kingdom was able to lower its inflation rate from 10% to 3% by 1992, when it was forced to abandon the exchange rate mechanism (ERM, discussed in Chapter 20).

Exchange-rate targeting has also been an effective means of reducing inflation quickly in emerging market countries. For example, before the devaluation in Mexico in 1994, its exchange-rate target enabled it to bring inflation down from levels above 100% in 1988 to below 10% in 1994.

Disadvantages of Exchange-Rate Targeting

Despite the inherent advantages of exchange-rate targeting, there are several serious criticisms of this strategy. The problem (as we saw in Chapter 20) is that with capital mobility the targeting country no longer can pursue its own independent monetary policy and so loses its ability to use monetary policy to respond to domestic shocks that are independent of those hitting the anchor country. Furthermore, an exchange-rate target means that shocks to the anchor country are directly transmitted to the targeting country, because changes in interest rates in the anchor country lead to a corresponding change in interest rates in the targeting country.

A striking example of these problems occurred when Germany reunified in 1990. In response to concerns about inflationary pressures arising from reunification and the massive fiscal expansion required to rebuild East Germany, long-term German interest rates rose until February 1991 and short-term rates rose until December 1991. This shock to the anchor country in the exchange rate mechanism (ERM) was transmitted directly to the other countries in the ERM whose currencies were pegged to the mark, and their interest rates rose in tandem with those in Germany. Continuing adherence to the exchange-rate target slowed economic growth and increased unemployment in countries such as France that remained in the ERM and adhered to the exchange-rate peg.

A second problem with exchange-rate targets is that they leave countries open to speculative attacks on their currencies. Indeed, one aftermath of German reunification was the foreign exchange crisis of September 1992. As we saw in Chapter 20, the tight monetary policy in Germany following reunification meant that the countries in the ERM were subjected to a negative demand shock that led to a decline in economic growth and a rise in unemployment. It was certainly feasible for the governments of these countries to keep their exchange rates fixed relative to the mark in these circumstances, but speculators began to question whether these countries' commitment to the exchange-rate peg would weaken. Speculators reasoned that these countries would not tolerate the rise in unemployment resulting from keeping interest rates high enough to fend off attacks on their currencies.

At this stage, speculators were, in effect, presented with a one-way bet, because the currencies of countries like France, Spain, Sweden, Italy, and the United Kingdom could go only in one direction and depreciate against the mark. Selling these currencies before the likely depreciation occurred gave speculators an attractive profit opportunity with potentially high expected returns. The result was the speculative attack in September 1992 discussed in Chapter 20. Only in France was the commitment to the fixed exchange rate strong enough so that France did not devalue. The governments in the other countries were unwilling to defend their currencies at all costs and eventually allowed their currencies to fall in value.

The different response of France and the United Kingdom after the September 1992 exchange-rate crisis illustrates the potential cost of an exchange-rate target. France, which continued to peg to the mark and was thus unable to use monetary policy to respond to domestic conditions, found that economic growth remained slow after 1992 and unemployment increased. The United Kingdom, on the other hand,

which dropped out of the ERM exchange-rate peg and adopted inflation targeting (discussed later in this chapter), had much better economic performance: economic growth was higher, the unemployment rate fell, and yet its inflation was not much worse than France's.

In contrast to industrialized countries, emerging market countries (including the so-called transition countries of Eastern Europe) may not lose much by giving up an independent monetary policy when they target exchange rates. Because many emerging market countries have not developed the political or monetary institutions that allow the successful use of discretionary monetary policy, they may have little to gain from an independent monetary policy, but a lot to lose. Thus, they would be better off by, in effect, adopting the monetary policy of a country like the United States through targeting exchange rates than by pursuing their own independent policy. This is one of the reasons that so many emerging market countries have adopted exchange-rate targeting.

Nonetheless, exchange-rate targeting is highly dangerous for these countries, because it leaves them open to speculative attacks that can have far more serious consequences for their economies than for the economies of industrialized countries. Indeed, as we saw in Chapters 8 and 20, the successful speculative attacks in Mexico in 1994, East Asia in 1997, and Argentina in 2002 plunged their economies into full-scale financial crises that devastated their economies.

An additional disadvantage of an exchange-rate target is that it can weaken the accountability of policymakers, particularly in emerging market countries. Because exchange-rate targeting fixes the exchange rate, it eliminates an important signal that can help constrain monetary policy from becoming too expansionary. In industrialized countries, particularly in the United States, the bond market provides an important signal about the stance of monetary policy. Overly expansionary monetary policy or strong political pressure to engage in overly expansionary monetary policy produces an inflation scare in which inflation expectations surge, interest rates rise because of the Fisher effect (described in Chapter 5), and there is a sharp decline in long-term bond prices. Because both central banks and the politicians want to avoid this kind of scenario, overly expansionary, time-consistent monetary policy will be less likely.

In many countries, particularly emerging market countries, the long-term bond market is essentially nonexistent. Under a flexible exchange-rate regime, however, if monetary policy is too expansionary, the exchange rate will depreciate. In these countries the daily fluctuations of the exchange rate can, like the bond market in United States, provide an early warning signal that monetary policy is too expansionary. Just as the fear of a visible inflation scare in the bond market constrains central bankers from pursuing overly expansionary monetary policy and also constrains politicians from putting pressure on the central bank to engage in overly expansionary monetary policy, fear of exchange-rate depreciations can make overly expansionary, time-consistent monetary policy less likely.

The need for signals from the foreign exchange market may be even more acute for emerging market countries, because the balance sheets and actions of the central banks are not as transparent as they are in industrialized countries. Targeting the exchange rate can make it even harder to ascertain the central bank's policy actions, as was true in Thailand before the July 1997 currency crisis. The public is less able to keep a watch on the central banks and the politicians pressuring it, which makes it easier for monetary policy to become too expansionary.

When Is Exchange-Rate Targeting Desirable for Industrialized Countries?

Given the above disadvantages with exchange-rate targeting, when might it make sense?

In industrialized countries, the biggest cost to exchange-rate targeting is the loss of an independent monetary policy to deal with domestic considerations. If an independent, domestic monetary policy can be conducted responsibly, this can be a serious cost indeed, as the comparison between the post-1992 experience of France and the United Kingdom indicates. However, not all industrialized countries have found that they are capable of conducting their own monetary policy successfully, either because of the lack of independence of the central bank or because political pressures on the central bank lead to an inflation bias in monetary policy. In these cases, giving up independent control of domestic monetary policy may not be a great loss, while the gain of having monetary policy determined by a better-performing central bank in the anchor country can be substantial.

Italy provides an example: It was not a coincidence that the Italian public was the most favorable of all those in Europe to the European Monetary Union. The past record of Italian monetary policy was not good, and the Italian public recognized that having monetary policy controlled by more responsible outsiders had benefits that far outweighed the costs of losing the ability to focus monetary policy on domestic considerations.

A second reason why industrialized countries might find targeting exchange rates useful is that it encourages integration of the domestic economy with its neighbors. Clearly this was the rationale for long-standing pegging of the exchange rate to the deutsche mark by countries such as Austria and the Netherlands, and the more recent exchange-rate pegs that preceded the European Monetary Union.

To sum up, exchange-rate targeting for industrialized countries is probably not the best monetary policy strategy to control the overall economy unless (1) domestic monetary and political institutions are not conducive to good monetary policymaking or (2) there are other important benefits of an exchange-rate target that have nothing to do with monetary policy.

When Is Exchange-Rate Targeting Desirable for Emerging Market Countries?

In countries whose political and monetary institutions are particularly weak and who therefore have been experiencing continued bouts of hyperinflation, a characterization that applies to many emerging market (including transition) countries, exchange-rate targeting may be the only way to break inflationary psychology and stabilize the economy. In this situation, exchange-rate targeting is the stabilization policy of last resort. However, if the exchange-rate targeting regimes in emerging market countries are not always transparent, they are more likely to break down, often resulting in disastrous financial crises.

Are there exchange-rate strategies that make it less likely that the exchange-rate regime will break down in emerging market countries? Two such strategies that have received increasing attention in recent years are currency boards and dollarization.

Currency Boards

http://users.erols.com/kurrency/intro.htm

A detailed discussion of the history, purpose, and function of currency boards.

One solution to the problem of lack of transparency and commitment to the exchange-rate target is the adoption of a **currency board**, in which the domestic currency is backed 100% by a foreign currency (say, dollars) and in which the note-issuing authority, whether the central bank or the government, establishes a fixed exchange rate to this foreign currency and stands ready to exchange domestic currency for the foreign currency at this rate whenever the public requests it. A currency board is just a variant of a fixed exchange-rate target in which the commitment to the fixed exchange

rate is especially strong because the conduct of monetary policy is in effect put on autopilot, taken completely out of the hands of the central bank and the government. In contrast, the typical fixed or pegged exchange-rate regime does allow the monetary authorities some discretion in their conduct of monetary policy because they can still adjust interest rates or print money.

A currency board arrangement thus has important advantages over a monetary policy strategy that just uses an exchange-rate target. First, the money supply can expand only when foreign currency is exchanged for domestic currency at the central bank. Thus the increased amount of domestic currency is matched by an equal increase in foreign exchange reserves. The central bank no longer has the ability to print money and thereby cause inflation. Second, the currency board involves a stronger commitment by the central bank to the fixed exchange rate and may therefore be effective in bringing down inflation quickly and in decreasing the likelihood of a successful speculative attack against the currency.

Although they solve the transparency and commitment problems inherent in an exchange-rate target regime, currency boards suffer from some of the same shortcomings: the loss of an independent monetary policy and increased exposure of the economy to shocks from the anchor country, and the loss of the central bank's ability to create money and act as a lender of last resort. Other means must therefore be used to cope with potential banking crises. Also, if there is a speculative attack on a currency board, the exchange of the domestic currency for foreign currency leads to a sharp contraction of the money supply, which can be highly damaging to the economy.

Currency boards have been established recently in countries such as Hong Kong (1983), Argentina (1991), Estonia (1992), Lithuania (1994), Bulgaria (1997), and Bosnia (1998). Argentina's currency board, which operated from 1991 to 2002 and required the central bank to exchange U.S. dollars for new pesos at a fixed exchange rate of 1 to 1, is one of the most interesting. Box 1 describes Argentina's experience with its currency board.

Dollarization

Another solution to the problems created by a lack of transparency and commitment to the exchange-rate target is **dollarization**, the adoption of a sound currency, like the U.S. dollar, as a country's money. Indeed, dollarization is just another variant of a fixed exchange-rate target with an even stronger commitment mechanism than a currency board provides. A currency board can be abandoned, allowing a change in the value of the currency, but a change of value is impossible with dollarization: a dollar bill is always worth one dollar whether it is held in the United States or outside of it.

Dollarization has been advocated as a monetary policy strategy for emerging market countries: It has been discussed actively by Argentine officials in the aftermath of the devaluation of the Brazilian real in January 1999 and was adopted by Ecuador in March 2000. Dollarization's key advantage is that it completely avoids the possibility of a speculative attack on the domestic currency (because there is none). (Such an attack is still a danger even under a currency board arrangement.)

Dollarization is subject to the usual disadvantages of an exchange-rate target (the loss of an independent monetary policy, increased exposure of the economy to shocks from the anchor country, and the inability of the central bank to create money and act as a lender of last resort). Dollarization has one additional disadvantage not characteristic of currency boards or other exchange-rate target regimes. Because a country adopting dollarization no longer has its own currency it loses the revenue that a government receives by issuing money, which is called **seignorage**. Because governments

Box 1: Global

Argentina's Currency Board

Argentina has had a long history of monetary instability, with inflation rates fluctuating dramatically and sometimes surging to beyond 1,000% a year. To end this cycle of inflationary surges, Argentina decided to adopt a currency board in April 1991. The Argentine currency board worked as follows. Under Argentina's convertibility law, the peso/dollar exchange rate was fixed at one to one, and a member of the public can go to the Argentine central bank and exchange a peso for a dollar, or vice versa, at any time.

The early years of Argentina's currency board looked stunningly successful. Inflation, which had been running at an 800% annual rate in 1990, fell to less than 5% by the end of 1994, and economic growth was rapid, averaging almost 8% at an annual rate from 1991 to 1994. In the aftermath of the Mexican peso crisis, however, concern about the health of the Argentine economy resulted in the public pulling money out of the banks (deposits fell by 18%) and exchanging pesos for dollars, thus causing a contraction of the Argentine money supply. The result was a sharp drop in Argentine economic activity, with real GDP shrinking by more than 5% in 1995 and the unemployment rate jumping above 15%. Only in 1996 did the economy begin to recover.

Because the central bank of Argentina had no control over monetary policy under the currency board system, it was relatively helpless to counteract the contractionary monetary policy stemming from the public's behavior. Furthermore, because the currency board did not allow the central bank to create pesos and lend them to the banks, it had very little capability to act as a lender of last resort. With help from international agencies, such as the IMF, the World Bank, and the Interamerican Development Bank, which lent Argentina over $5 billion in 1995 to help shore up its banking system, the currency board survived.

However, in 1998 Argentina entered another recession, which was both severe and very long lasting. By the end of 2001, unemployment reached nearly 20%, a level comparable to that experienced in the United States during the Great Depression of the 1930s. The result has been civil unrest and the fall of the elected government, as well as a major banking crisis and a default on nearly $150 billion of government debt. Because the Central Bank of Argentina had no control over monetary policy under the currency board system, it was unable to use monetary policy to expand the economy and get out of its recession. Furthermore, because the currency board did not allow the central bank to create pesos and lend them to banks, it had very little capability to act as a lender of last resort. In January 2002, the currency board finally collapsed and the peso depreciated by more than 70%. The result was the full-scale financial crisis described in Chapter 8, with inflation shooting up and an extremely severe depression. Clearly, the Argentine public is not as enamored of its currency board as it once was.

(or their central banks) do not have to pay interest on their currency, they earn revenue (seignorage) by using this currency to purchase income-earning assets such as bonds. In the case of the Federal Reserve in the United States, this revenue is on the order of $30 billion per year. If an emerging market country dollarizes and gives up its currency, it needs to make up this loss of revenue somewhere, which is not always easy for a poor country.

Study Guide

As a study aid, the advantages and disadvantages of exchange-rate targeting and the other monetary policy strategies are listed in Table 1.

SUMMARY **Table 1 Advantages and Disadvantages of Different Monetary Policy Strategies**

Exchange-Rate Targeting	Monetary Targeting	Inflation Targeting	Implicit Nominal Anchor
Advantages			
Directly ties down inflation of internationally traded goods			
Automatic rule for conduct of monetary policy			
Simplicity and clarity of target		Simplicity and clarity of target	
	Independent monetary policy can focus on domestic considerations	Independent monetary policy can focus on domestic considerations	Independent monetary policy can focus on domestic considerations
	Immediate signal on achievement of target		
		Does not rely on stable money–inflation relationship	Does not rely on stable money–inflation relationship
		Increased accountability of central bank	
		Reduced effects of inflationary shocks	
			Demonstrated success in U.S.
Disadvantages			
Loss of independent monetary policy			
Open to speculative attacks (less for currency board and not a problem for dollarization)			
Loss of exchange-rate signal			
	Relies on stable money–inflation relationship		
		Delayed signal about achievement of target	
		Could impose rigid rule (though not in practice)	
		Larger output fluctuations if sole focus on inflation (though not in practice)	
			Lack of transparency
			Success depends on individuals
			Low accountability

Monetary Targeting

In many countries, exchange-rate targeting is not an option, because either the country (or bloc of countries) is too large or because there is no country whose currency is an obvious choice to serve as the nominal anchor. Exchange-rate targeting is therefore clearly not an option for the United States, Japan, or the European Monetary Union. These countries must look to other strategies for the conduct of monetary policy, one of which is *monetary targeting*.

Monetary Targeting in Canada, the United Kingdom, Japan, Germany, and Switzerland

In the 1970s, monetary targeting was adopted by several countries, notably Germany, Switzerland, Canada, the United Kingdom, and Japan, as well as in the United States (already discussed in Chapter 18). This strategy involves using monetary aggregates as an intermediate target of the type described in Chapter 18 to achieve an ultimate goal such as price stability. Monetary targeting as practiced was quite different from Milton Friedman's suggestion that the chosen monetary aggregate be targeted to grow at a constant rate. Indeed, in all these countries the central banks never adhered to strict, ironclad rules for monetary growth and in some of these countries monetary targeting was not pursued very seriously.

Canada and the United Kingdom. In a move similar to that made by the United States, the Bank of Canada responded to a rise in inflation in the early 1970s by introducing a program of monetary targeting referred to as "monetary gradualism." Under this policy, which began in 1975, M1 growth would be controlled within a gradually falling target range. The British introduced monetary targeting in late 1973, also in response to mounting concerns about inflation. The Bank of England targeted M3, a broader monetary target than the Bank of Canada or the Fed used.

By 1978, only three years after monetary targeting had begun, the Bank of Canada began to distance itself from this strategy out of concern for the exchange rate. Because of the conflict with exchange-rate goals, as well as the uncertainty about M1 as a reliable guide to monetary policy, the M1 targets were abandoned in November 1982. Gerald Bouey, then governor of the Bank of Canada, described the situation by saying, "We didn't abandon monetary aggregates, they abandoned us."

In the United Kingdom, after monetary aggregates overshot their targets and inflation accelerated in the late 1970s, Prime Minister Margaret Thatcher in 1980 introduced the Medium-Term Financial Strategy, which proposed a gradual deceleration of M3 growth. Unfortunately, the M3 targets ran into problems similar to those of the M1 targets in the United States: They were not reliable indicators of the tightness of monetary policy. After 1983, arguing that financial innovation was wreaking havoc with the relationship between M3 and national income, the Bank of England began to de-emphasize M3 in favor of a narrower monetary aggregate, M0 (the monetary base). The target for M3 was temporarily suspended in October 1985 and was completely dropped in 1987.

A feature of monetary targeting in Canada and especially in the United Kingdom was that there was substantial game playing: Their central banks targeted multiple aggregates, allowed base drift (by applying target growth rates to a new base at which the target ended up every period), did not announce targets on a regular schedule, used artificial means to bring down the growth of a targeted aggregate, often overshot their targets without reversing the overshoot later, and often obscured why deviations from the monetary targets occurred.

Japan. The increase in oil prices in late 1973 was a major shock for Japan, which experienced a huge jump in the inflation rate, to greater than 20% in 1974—a surge facilitated by money growth in 1973 in excess of 20%. The Bank of Japan, like the other central banks discussed here, began to pay more attention to money growth rates. In 1978, the Bank of Japan began to announce "forecasts" at the beginning of each quarter for M2 + CDs. Although the Bank of Japan was not officially committed to monetary targeting, monetary policy appeared to be more money-focused after 1978. For example, after the second oil price shock in 1979, the Bank of Japan quickly reduced M2 + CDs growth, rather than allowing it to shoot up as occurred after the first oil shock. The Bank of Japan conducted monetary policy with operating procedures that were similar in many ways to those that the Federal Reserve has used in the United States. The Bank of Japan uses the interest rate in the Japanese interbank market (which has a function similar to that of the federal funds market in the United States) as its daily operating target, just as the Fed has done.

The Bank of Japan's monetary policy performance during the 1978–1987 period was much better than the Fed's. Money growth in Japan slowed gradually, beginning in the mid-1970s, and was much less variable than in the United States. The outcome was a more rapid braking of inflation and a lower average inflation rate. In addition, these excellent results on inflation were achieved with lower variability in real output in Japan than in the United States.

In parallel with the United States, financial innovation and deregulation in Japan began to reduce the usefulness of the M2 + CDs monetary aggregate as an indicator of monetary policy. Because of concerns about the appreciation of the yen, the Bank of Japan significantly increased the rate of money growth from 1987 to 1989. Many observers blame speculation in Japanese land and stock prices (the so-called bubble economy) on the increase in money growth. To reduce this speculation, in 1989 the Bank of Japan switched to a tighter monetary policy aimed at slower money growth. The aftermath was a substantial decline in land and stock prices and the collapse of the bubble economy.

The 1990s and afterwards has not been a happy period for the Japanese economy. The collapse of land and stock prices helped provoke a severe banking crisis, discussed in Chapter 11, that has continued to be a severe drag on the economy. The resulting weakness of the economy has even led to bouts of deflation, promoting further financial instability. The outcome has been an economy that has been stagnating for over a decade. Many critics believe that the Bank of Japan has pursued overly tight monetary policy and needs to substantially increase money growth in order to lift the economy out of its stagnation.

Germany and Switzerland. The two countries that officially engaged in monetary targeting for over 20 years starting at the end of 1974 were Germany and Switzerland, and this is why we will devote more attention to them. The success of monetary policy in these two countries in controlling inflation is the reason that monetary targeting still has strong advocates and is an element of the official policy regime for the European Central Bank (see Box 2).

The monetary aggregate chosen by the Germans was a narrow one they called *central bank money*, the sum of currency in circulation and bank deposits weighted by the 1974 required reserve ratios. In 1988, the Bundesbank switched targets from central bank money to M3. The Swiss began targeting the M1 monetary aggregate, but in 1980 switched to the narrower monetary aggregate, M0, the monetary base.

Box 2: Global

The European Central Bank's Monetary Policy Strategy

The European Central Bank (ECB) has adopted a hybrid monetary policy strategy that has much in common with the monetary targeting strategy previously used by the Bundesbank but also has some elements of inflation targeting. The ECB's strategy has two key "pillars." First is a prominent role for monetary aggregates with a "reference value" for the growth rate of a monetary aggregate (M3). Second is a broadly based assessment of the outlook for future price developments with a goal of price stability defined as a year-on-year increase in the consumer price index below 2%. After critics pointed out that a deflationary situation with negative inflation would satisfy the stated price stability criteria, the ECB provided a clarification that inflation meant positive inflation only, so that the price stability goal should be interpreted as a range for inflation of 0–2%.

The ECB's strategy is somewhat unclear and has been subjected to criticism for this reason. Although the 0–2% range for the goal of price stability sounds like an inflation target, the ECB has not been willing to live with this interpretation—it has repeatedly stated that it does not have an inflation target. On the other hand, the ECB has downgraded the importance of monetary aggregates in its strategy by using the term "reference value" rather than "target" in describing its strategy and has indicated that it will also monitor broadly based developments on the price level. The ECB seems to have decided to try to have its cake and eat it too by not committing too strongly to either a monetary or an inflation-targeting strategy. The resulting difficulty of assessing what the ECB's strategy is likely to be has the potential to reduce the accountability of this new institution.

The key fact about monetary targeting regimes in Germany and Switzerland is that the targeting regimes were very far from a Friedman-type monetary targeting rule in which a monetary aggregate is kept on a constant-growth-rate path and is the primary focus of monetary policy. As Otmar Issing, at the time the chief economist of the Bundesbank has noted, "One of the secrets of success of the German policy of money-growth targeting was that ... it often did not feel bound by monetarist orthodoxy as far as its more technical details were concerned."[2] The Bundesbank allowed growth outside of its target ranges for periods of two to three years, and overshoots of its targets were subsequently reversed. Monetary targeting in Germany and Switzerland was instead primarily a method of communicating the strategy of monetary policy focused on long-run considerations and the control of inflation.

The calculation of monetary target ranges put great stress on making policy transparent (clear, simple, and understandable) and on regular communication with the public. First and foremost, a numerical inflation goal was prominently featured in the setting of target ranges. Second, monetary targeting, far from being a rigid policy rule, was quite flexible in practice. The target ranges for money growth were missed on the order of 50% of the time in Germany, often because of the Bundesbank's concern about other objectives, including output and exchange rates. Furthermore, the Bundesbank demonstrated its flexibility by allowing its inflation goal to vary over time and to converge gradually to the long-run inflation goal.

[2]Otmar Issing, "Is Monetary Targeting in Germany Still Adequate?" In *Monetary Policy in an Integrated World Economy: Symposium 1995*, ed. Horst Siebert (Tübingen: Mohr, 1996), p. 120.

When the Bundesbank first set its monetary targets at the end of 1974, it announced a medium-term inflation goal of 4%, well above what it considered to be an appropriate long-run goal. It clarified that this medium-term inflation goal differed from the long-run goal by labeling it the "unavoidable rate of price increase." Its gradualist approach to reducing inflation led to a period of nine years before the medium-term inflation goal was considered to be consistent with price stability. When this occurred at the end of 1984, the medium-term inflation goal was renamed the "normative rate of price increase" and was set at 2%. It continued at this level until 1997, when it was changed to 1.5 to 2%. The Bundesbank also responded to negative supply shocks, restrictions in the supply of energy or raw materials that raise the price level, by raising its medium-term inflation goal: specifically, it raised the unavoidable rate of price increase from 3.5% to 4% in the aftermath of the second oil price shock in 1980.

The monetary targeting regimes in Germany and Switzerland demonstrated a strong commitment to clear communication of the strategy to the general public. The money growth targets were continually used as a framework to explain the monetary policy strategy, and both the Bundesbank and the Swiss National Bank expended tremendous effort in their publications and in frequent speeches by central bank officials to communicate to the public what the central bank was trying to achieve. Given that both central banks frequently missed their money growth targets by significant amounts, their monetary targeting frameworks are best viewed as a mechanism for transparently communicating how monetary policy is being directed to achieve inflation goals and as a means for increasing the accountability of the central bank.

The success of Germany's monetary targeting regime in producing low inflation has been envied by many other countries, explaining why it was chosen as the anchor country for the exchange rate mechanism. One clear indication of Germany's success occurred in the aftermath of German reunification in 1990. Despite a temporary surge in inflation stemming from the terms of reunification, high wage demands, and the fiscal expansion, the Bundesbank was able to keep these temporary effects from becoming embedded in the inflation process, and by 1995, inflation fell back down below the Bundesbank's normative inflation goal of 2%.

Monetary targeting in Switzerland has been more problematic than in Germany, suggesting the difficulties of targeting monetary aggregates in a small open economy that also underwent substantial changes in the institutional structure of its money markets. In the face of a 40% trade-weighted appreciation of the Swiss franc from the fall of 1977 to the fall of 1978, the Swiss National Bank decided that the country could not tolerate this high a level of the exchange rate. Thus, in the fall of 1978, the monetary targeting regime was abandoned temporarily, with a shift from a monetary target to an exchange-rate target until the spring of 1979, when monetary targeting was reintroduced (although not announced).

The period from 1989 to 1992 was also not a happy one for Swiss monetary targeting, because the Swiss National Bank failed to maintain price stability after it successfully reduced inflation. The substantial overshoot of inflation from 1989 to 1992, reaching levels above 5%, was due to two factors. The first was that the strength of the Swiss franc from 1985 to 1987 caused the Swiss National Bank to allow the monetary base to grow at a rate greater than the 2% target in 1987 and then caused it to raise the money growth target to 3% for 1988. The second arose from the introduction of a new interbank payment system, Swiss Interbank Clearing (SIC), and a wide-ranging revision of the commercial banks' liquidity requirements in 1988. The result

of the shocks to the exchange rate and the shift in the demand for monetary base arising from the above institutional changes created a serious problem for its targeted aggregate. As the 1988 year unfolded, it became clear that the Swiss National Bank had guessed wrong in predicting the effects of these shocks, so that monetary policy was too easy, even though the monetary target was undershot. The result was a subsequent rise in inflation to above the 5% level.

As a result of these problems with monetary targeting Switzerland substantially loosened its monetary targeting regime (and ultimately, adopted inflation targeting in 2000). The Swiss National Bank recognized that its money growth targets were of diminished utility as a means of signaling the direction of monetary policy. Thus, its announcement at the end of 1990 of the medium-term growth path did not specify a horizon for the target or the starting point of the growth path. At the end of 1992, the bank specified the starting point for the expansion path, and at the end of 1994, it announced a new medium-term path for money base growth for the period 1995 to 1999. By setting this path, the bank revealed retroactively that the horizon of the first path was also five years (1990–1995). Clearly, the Swiss National Bank moved to a much more flexible framework in which hitting one-year targets for money base growth has been abandoned. Nevertheless, Swiss monetary policy continued to be successful in controlling inflation, with inflation rates falling back down below the 1% level after the temporary bulge in inflation from 1989 to 1992.

There are two key lessons to be learned from our discussion of German and Swiss monetary targeting. First, a monetary targeting regime can restrain inflation in the longer run, even when the regime permits substantial target misses. Thus adherence to a rigid policy rule has not been found to be necessary to obtain good inflation outcomes. Second, the key reason why monetary targeting has been reasonably successful in these two countries, despite frequent target misses, is that the objectives of monetary policy are clearly stated and both the central banks actively engaged in communicating the strategy of monetary policy to the public, thereby enhancing the transparency of monetary policy and the accountability of the central bank.

As we will see in the next section, these key elements of a successful targeting regime—flexibility, transparency, and accountability—are also important elements in inflation-targeting regimes. German and Swiss monetary policy was actually closer in practice to inflation targeting than it was to Friedman-like monetary targeting, and thus might best be thought of as "hybrid" inflation targeting.

Advantages of Monetary Targeting

A major advantage of monetary targeting over exchange-rate targeting is that it enables a central bank to adjust its monetary policy to cope with domestic considerations. It enables the central bank to choose goals for inflation that may differ from those of other countries and allows some response to output fluctuations. Also, as with an exchange-rate target, information on whether the central bank is achieving its target is known almost immediately—figures for monetary aggregates are typically reported within a couple of weeks. Thus, monetary targets can send almost immediate signals to the public and markets about the stance of monetary policy and the intentions of the policymakers to keep inflation in check. In turn, these signals help fix inflation expectations and produce less inflation. Monetary targets also allow almost immediate accountability for monetary policy to keep inflation low, thus helping to constrain the monetary policymaker from falling into the time-consistency trap.

Disadvantages of Monetary Targeting

All of the above advantages of monetary aggregate targeting depend on a big *if*: There must be a strong and reliable relationship between the goal variable (inflation or nominal income) and the targeted aggregate. If the relationship between the monetary aggregate and the goal variable is weak, monetary aggregate targeting will not work; this seems to have been a serious problem in Canada, the United Kingdom, and Switzerland, as well as in the United States. The weak relationship implies that hitting the target will not produce the desired outcome on the goal variable and thus the monetary aggregate will no longer provide an adequate signal about the stance of monetary policy. As a result, monetary targeting will not help fix inflation expectations and be a good guide for assessing the accountability of the central bank. In addition, an unreliable relationship between monetary aggregates and goal variables makes it difficult for monetary targeting to serve as a communications device that increases the transparency of monetary policy and makes the central bank accountable to the public.

Inflation Targeting

www.ny.frb.org/rmaghome/econ_pol/897fmis.htm

Research on inflation targeting published by the Federal Reserve and coauthored by the author of this text.

Given the breakdown of the relationship between monetary aggregates and goal variables such as inflation, many countries that want to maintain an independent monetary policy have recently adopted inflation targeting as their monetary policy regime. New Zealand was the first country to formally adopt inflation targeting in 1990, followed by Canada in 1991, the United Kingdom in 1992, Sweden and Finland in 1993, and Australia and Spain in 1994. Israel, Chile, and Brazil, among others, have also adopted a form of inflation targeting.

Inflation targeting involves several elements: (1) public announcement of medium-term numerical targets for inflation; (2) an institutional commitment to price stability as the primary, long-run goal of monetary policy and a commitment to achieve the inflation goal; (3) an information-inclusive strategy in which many variables and not just monetary aggregates are used in making decisions about monetary policy; (4) increased transparency of the monetary policy strategy through communication with the public and the markets about the plans and objectives of monetary policymakers; and (5) increased accountability of the central bank for attaining its inflation objectives.

Inflation Targeting in New Zealand, Canada, and the United Kingdom

We begin our look at inflation targeting with New Zealand, because it was the first country to adopt it. We then go on to look at the experiences in Canada and the United Kingdom, which were next to adopt this strategy.[3]

New Zealand. As part of a general reform of the government's role in the economy, the New Zealand parliament passed a new Reserve Bank of New Zealand Act in 1989,

[3]For further discussion of experiences with inflation targeting, particularly in other countries, see Leonardo Leiderman and Lars E. O. Svensson, *Inflation Targeting* (London: Centre for Economic Policy Research, 1995); Frederic S. Mishkin and Adam Posen, "Inflation Targeting: Lessons from Four Countries," Federal Reserve Bank of New York, *Economic Policy Review* 3 (August 1997), pp. 9–110; and Ben S. Bernanke, Thomas Laubach, Frederic S. Mishkin, and Adam S. Posen, *Inflation Targeting: Lessons from the International Experience* (Princeton: Princeton University Press, 1999).

which became effective on February 1, 1990. Besides increasing the independence of the central bank, moving it from being one of the least independent to one of the most independent among the developed countries, the act also committed the Reserve Bank to a sole objective of price stability. The act stipulated that the minister of finance and the governor of the Reserve Bank should negotiate and make public a Policy Targets Agreement, a statement that sets out the targets by which monetary policy performance will be evaluated, specifying numerical target ranges for inflation and the dates by which they are to be reached. An unusual feature of the New Zealand legislation is that the governor of the Reserve Bank is held highly accountable for the success of monetary policy. If the goals set forth in the Policy Targets Agreement are not satisfied, the governor is subject to dismissal.

The first Policy Targets Agreement, signed by the minister of finance and the governor of the Reserve Bank on March 2, 1990, directed the Reserve Bank to achieve an annual inflation rate within a 3–5% range. Subsequent agreements lowered the range to 0–2% until the end of 1996, when the range was changed to 0–3%. As a result of tight monetary policy, the inflation rate was brought down from above 5% to below 2% by the end of 1992 (see Figure 1, panel a), but at the cost of a deep recession and a sharp rise in unemployment. Since then, inflation has typically remained within the targeted range, with the exception of a brief period in 1995 when it exceeded the range by a few tenths of a percentage point. (Under the Reserve Bank Act, the governor, Donald Brash, could have been dismissed, but after parliamentary debate he was retained in his job.) Since 1992, New Zealand's growth rate has generally been high, with some years exceeding 5%, and unemployment has come down significantly.

Canada. On February 26, 1991, a joint announcement by the minister of finance and the governor of the Bank of Canada established formal inflation targets. The target ranges were 2–4% by the end of 1992, 1.5–3.5% by June 1994, and 1–3% by December 1996. After the new government took office in late 1993, the target range was set at 1–3% from December 1995 until December 1998 and has been kept at this level. Canadian inflation has also fallen dramatically since the adoption of inflation targets, from above 5% in 1991, to a 0% rate in 1995, and to between 1 and 2% in the late 1990s (see Figure 1, panel b). As was the case in New Zealand, however, this decline was not without cost: unemployment soared to above 10% from 1991 until 1994, but then declined substantially.

United Kingdom. Once the U.K. left the European Monetary System after the speculative attack on the pound in September 1992 (discussed in Chapter 20), the British decided to turn to inflation targets instead of the exchange rate as their nominal anchor. As you may recall from Chapter 14, the central bank in the U.K., the Bank of England, did not have statutory authority over monetary policy until 1997; it could only make recommendations about monetary policy. Thus it was the chancellor of the Exchequer (the equivalent of the U.S. Treasury secretary) who announced an inflation target for the U.K. on October 8, 1992. Three weeks later he "invited" the governor of the Bank of England to begin producing an *Inflation Report*, a quarterly report on the progress being made in achieving the target—an invitation the governor accepted. The inflation target range was set at 1–4% until the next election, spring 1997 at the latest, with the intent that the inflation rate should settle down to the lower half of the range (below 2.5%). In May 1997, after the new Labour government came into power, it adopted a point target of 2.5% for inflation and gave the Bank of England the power to set interest rates henceforth, granting it a more independent role in monetary policy.

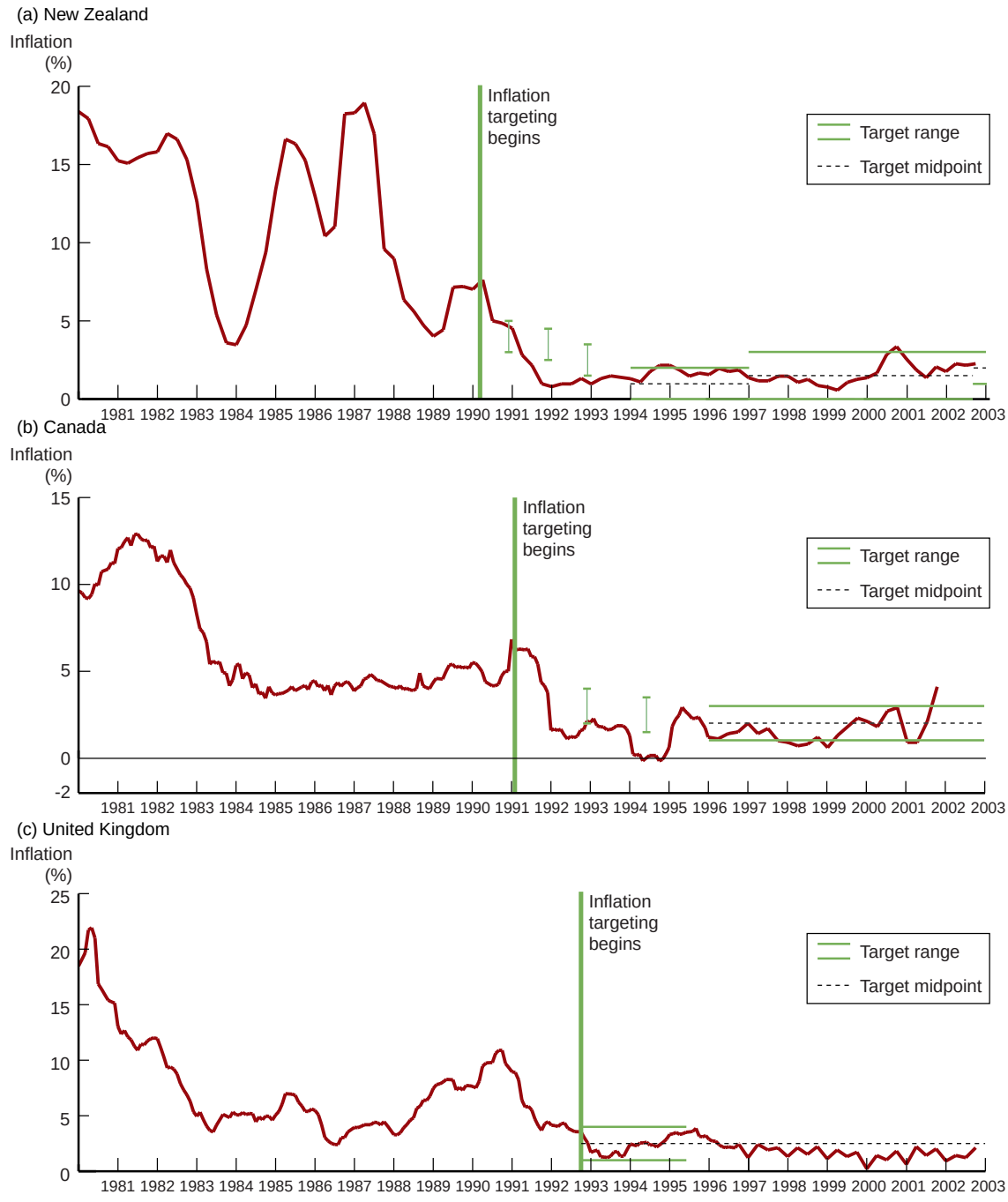


FIGURE 1 Inflation Rates and Inflation Targets for New Zealand, Canada, and the United Kingdom, 1980–2002
(a) New Zealand; (b) Canada; (c) United Kingdom

Source: Ben S. Bernanke, Thomas Laubach, Frederic S. Mishkin, and Adam S. Poson, *Inflation Targeting: Lessons from the International Experience* (Princeton: Princeton University Press, 1999), updates from the same sources, and www.rbnz.govt.nz/statistics/econind/a3/ha3.xls.

Before the adoption of inflation targets, inflation had already been falling in the U.K. from a peak of 9% at the beginning of 1991 to 4% at the time of adoption (see Figure 1, panel c). After a small upward movement in early 1993, inflation continued to fall until by the third quarter of 1994, it was at 2.2%, within the intended range articulated by the chancellor. Subsequently inflation rose, climbing slightly above the 2.5% level by 1996, but has remained around the 2.5% target since then. Meanwhile, growth of the U.K. economy has been strong, causing a substantial reduction in the unemployment rate.

Advantages of Inflation Targeting

Inflation targeting has several advantages over exchange-rate and monetary targeting as a strategy for the conduct of monetary policy. In contrast to exchange-rate targeting, but like monetary targeting, inflation targeting enables monetary policy to focus on domestic considerations and to respond to shocks to the domestic economy. Inflation targeting also has the advantage that stability in the relationship between money and inflation is not critical to its success, because it does not rely on this relationship. An inflation target allows the monetary authorities to use all available information, not just one variable, to determine the best settings for monetary policy.

Inflation targeting, like exchange-rate targeting, also has the key advantage that it is readily understood by the public and is thus highly transparent. Monetary targets, in contrast, are less likely to be easily understood by the public than inflation targets, and if the relationship between monetary aggregates and the inflation goal variable is subject to unpredictable shifts, as has occurred in many countries, monetary targets lose their transparency because they are no longer able to accurately signal the stance of monetary policy.

Because an explicit numerical inflation target increases the accountability of the central bank, inflation targeting also has the potential to reduce the likelihood that the central bank will fall into the time-consistency trap, trying to expand output and employment by pursuing overly expansionary monetary policy. A key advantage of inflation targeting is that it can help focus the political debate on what a central bank can do in the long run—that is, control inflation, rather than what it cannot do, which is permanently increase economic growth and the number of jobs through expansionary monetary policy. Thus, inflation targeting has the potential to reduce political pressures on the central bank to pursue inflationary monetary policy and thereby to reduce the likelihood of time-consistent policymaking.

Inflation-targeting regimes also put great stress on making policy transparent and on regular communication with the public. Inflation-targeting central banks have frequent communications with the government, some mandated by law and some in response to informal inquiries, and their officials take every opportunity to make public speeches on their monetary policy strategy. While these techniques are also commonly used in countries that have not adopted inflation targeting (such as Germany before EMU and the United States), inflation-targeting central banks have taken public outreach a step further: not only do they engage in extended public information campaigns, including the distribution of glossy brochures, but they publish documents like the Bank of England's *Inflation Report*. The publication of these documents is particularly noteworthy, because they depart from the usual dull-looking, formal reports of central banks and use fancy graphics, boxes, and other eye-catching design elements to engage the public's interest.

The above channels of communication are used by central banks in inflation-targeting countries to explain the following concepts to the general public, financial

market participants, and the politicians: (1) the goals and limitations of monetary policy, including the rationale for inflation targets; (2) the numerical values of the inflation targets and how they were determined, (3) how the inflation targets are to be achieved, given current economic conditions; and (4) reasons for any deviations from targets. These communications have improved private sector planning by reducing uncertainty about monetary policy, interest rates, and inflation; they have promoted public debate of monetary policy, in part by educating the public about what a central bank can and cannot achieve; and they have helped clarify the responsibilities of the central bank and of politicians in the conduct of monetary policy.

Another key feature of inflation-targeting regimes is the tendency toward increased accountability of the central bank. Indeed, transparency and communication go hand in hand with increased accountability. The strongest case of accountability of a central bank in an inflation-targeting regime is in New Zealand, where the government has the right to dismiss the Reserve Bank's governor if the inflation targets are breached, even for one quarter. In other inflation-targeting countries, the central bank's accountability is less formalized. Nevertheless, the transparency of policy associated with inflation targeting has tended to make the central bank highly accountable to the public and the government. Sustained success in the conduct of monetary policy as measured against a pre-announced and well-defined inflation target can be instrumental in building public support for a central bank's independence and for its policies. This building of public support and accountability occurs even in the absence of a rigidly defined and legalistic standard of performance evaluation and punishment.

Two remarkable examples illustrate the benefits of transparency and accountability in the inflation-targeting framework. The first occurred in Canada in 1996, when the president of the Canadian Economic Association made a speech criticizing the Bank of Canada for pursuing monetary policy that he claimed was too contractionary. His speech sparked a widespread public debate. In countries not pursuing inflation targeting, such debates often degenerate into calls for the immediate expansion of monetary policy with little reference to the long-run consequences of such a policy change. In this case, however, the very existence of inflation targeting channeled the debate into a discussion over what should be the appropriate target for inflation, with both the bank and its critics obliged to make explicit their assumptions and estimates of the costs and benefits of different levels of inflation. Indeed, the debate and the Bank of Canada's record and responsiveness increased support for the Bank of Canada, with the result that criticism of the bank and its conduct of monetary policy was not a major issue in the 1997 elections as it had been before the 1993 elections.

The second example occurred upon the granting of operational independence to the Bank of England on May 6, 1997. Prior to that date, the government, as represented by the chancellor of the Exchequer, controlled the decision to set monetary policy instruments, while the Bank of England was relegated to acting as the government's counterinflationary conscience. On May 6, the new chancellor of the Exchequer, Gordon Brown, announced that the Bank of England would henceforth have the responsibility for setting interest rates and for engaging in short-term exchange-rate interventions. Two factors were cited by Chancellor Brown that justified the government's decision: first was the bank's successful performance over time as measured against an announced clear target; second was the increased accountability that an independent central bank is exposed to under an inflation-targeting framework, making the bank more responsive to political oversight. The granting of operational independence

to the Bank of England occurred because it would operate under a monetary policy regime to ensure that monetary policy goals cannot diverge from the interests of society for extended periods of time. Nonetheless, monetary policy was to be insulated from short-run political considerations. An inflation-targeting regime makes it more palatable to have an independent central bank that focuses on long-run objectives but is consistent with a democratic society because it is accountable.

The performance of inflation-targeting regimes has been quite good. Inflation-targeting countries seem to have significantly reduced both the rate of inflation and inflation expectations beyond what would likely have occurred in the absence of inflation targets. Furthermore, once down, inflation in these countries has stayed down; following disinflations, the inflation rate in targeting countries has not bounced back up during subsequent cyclical expansions of the economy.

Inflation targeting also seems to ameliorate the effects of inflationary shocks. For example, shortly after adopting inflation targets in February 1991, the Bank of Canada was faced with a new goods and services tax (GST), an indirect tax similar to a value-added tax—an adverse supply shock that in earlier periods might have led to a ratcheting up in inflation. Instead the tax increase led to only a one-time increase in the price level; it did not generate second- and third-round increases in wages and prices that would have led to a persistent rise in the inflation rate. Another example is the experience of the United Kingdom and Sweden following their departures from the ERM exchange-rate pegs in 1992. In both cases, devaluation would normally have stimulated inflation because of the direct effects on higher export and import prices from devaluation and the subsequent effects on wage demands and price-setting behavior. Again, it seems reasonable to attribute the lack of inflationary response in these episodes to adoption of inflation targeting, which short-circuited the second- and later-round effects and helped to focus public attention on the temporary nature of the inflation shocks. Indeed, one reason why inflation targets were adopted in both countries was to achieve exactly this result.

Disadvantages of Inflation Targeting

Critics of inflation targeting cite four disadvantages/criticisms of this monetary policy strategy: delayed signaling, too much rigidity, the potential for increased output fluctuations, and low economic growth. We look at each in turn and examine the validity of these criticisms.

Delayed Signaling. In contrast to exchange rates and monetary aggregates, inflation is not easily controlled by the monetary authorities. Furthermore, because of the long lags in the effects of monetary policy, inflation outcomes are revealed only after a substantial lag. Thus, an inflation target is unable to send immediate signals to both the public and markets about the stance of monetary policy. However, we have seen that the signals provided by monetary aggregates may not be very strong and that an exchange-rate peg may obscure the ability of the foreign exchange market to signal overly expansionary policies. Hence, it is not at all clear that these other strategies are superior to inflation targeting on these grounds.

Too Much Rigidity. Some economists have criticized inflation targeting because they believe it imposes a rigid rule on monetary policymakers, limiting their discretion to respond to unforeseen circumstances. For example, policymakers in countries that adopted monetary targeting did not foresee the breakdown of the relationship between monetary aggregates and goal variables such as nominal spending or infla-

tion. With rigid adherence to a monetary rule, the breakdown in their relationship could have been disastrous. However, the traditional distinction between rules and discretion can be highly misleading. Useful policy strategies exist that are "rule-like," in that they involve forward-looking behavior that limits policymakers from systematically engaging in policies with undesirable long-run consequences. Such policies avoid the time-consistency problem and would best be described as "constrained discretion."

Indeed, inflation targeting can be described exactly in this way. Inflation targeting, as actually practiced, is far from rigid. First, inflation targeting does not prescribe simple and mechanical instructions on how the central bank should conduct monetary policy. Rather, it requires the central bank to use all available information to determine what policy actions are appropriate to achieve the inflation target. Unlike simple policy rules, inflation targeting never requires the central bank to focus solely on one key variable. Second, inflation targeting as practiced contains a substantial degree of policy discretion. Inflation targets have been modified depending on economic circumstances, as we have seen. Moreover, central banks under inflation-targeting regimes have left themselves considerable scope to respond to output growth and fluctuations through several devices.

Potential for Increased Output Fluctuations. An important criticism of inflation targeting is that a sole focus on inflation may lead to monetary policy that is too tight when inflation is above target and thus may lead to larger output fluctuations. Inflation targeting does not, however, require a sole focus on inflation—in fact, experience has shown that inflation targeters do display substantial concern about output fluctuations. All the inflation targeters have set their inflation targets above zero.[4] For example, currently New Zealand has the lowest midpoint for an inflation target, 1.5%, while Canada and Sweden set the midpoint of their inflation target at 2%; and the United Kingdom and Australia currently have their midpoints at 2.5%.

The decision by inflation targeters to choose inflation targets above zero reflects the concern of monetary policymakers that particularly low inflation can have substantial negative effects on real economic activity. Deflation (negative inflation in which the price level actually falls) is especially to be feared because of the possibility that it may promote financial instability and precipitate a severe economic contraction (Chapter 8). The deflation in Japan in recent years has been an important factor in the weakening of the Japanese financial system and economy. Targeting inflation rates of above zero makes periods of deflation less likely. This is one reason why some economists both within and outside of Japan have been calling on the Bank of Japan to adopt an inflation target at levels of 2% or higher.

Inflation targeting also does not ignore traditional stabilization goals. Central bankers in inflation-targeting countries continue to express their concern about fluctuations in output and employment, and the ability to accommodate short-run stabilization goals to some degree is built into all inflation-targeting regimes. All inflation-targeting countries have been willing to minimize output declines by gradually lowering medium-term inflation targets toward the long-run goal.

[4]CPI indices have been found to have an upward bias in the measurement of true inflation, and so it is not surprising that inflation targets would be chosen to exceed zero. However, the actual targets have been set to exceed the estimates of this measurement bias, indicating that inflation targeters have decided to have targets for inflation that exceed zero even after measurement bias is accounted for.

In addition, many inflation targeters, particularly the Bank of Canada, have emphasized that the floor of the target range should be emphasized every bit as much as the ceiling, thus helping to stabilize the real economy when there are negative shocks to demand. Inflation targets can increase the central bank's flexibility in responding to declines in aggregate spending. Declines in aggregate demand that cause the inflation rate to fall below the floor of the target range will automatically stimulate the central bank to loosen monetary policy without fearing that its action will trigger a rise in inflation expectations.

Another element of flexibility in inflation-targeting regimes is that deviations from inflation targets are routinely allowed in response to supply shocks, such as restrictions in the supply of energy or raw materials that could have substantial negative effects on output. First, the price index on which the official inflation targets are based is often defined to exclude or moderate the effects of "supply shocks"; for example, the officially targeted price index may exclude some combination of food and energy prices. Second, following (or in anticipation of) a supply shock, such as a rise in a value-added tax (similar to a sales tax), the central bank would first deviate from its planned policies as needed and then explain to the public the reasons for its action.

Low Economic Growth. Another common concern about inflation targeting is that it will lead to low growth in output and employment. Although inflation reduction has been associated with below-normal output during disinflationary phases in inflation-targeting regimes, once low inflation levels were achieved, output and employment returned to levels at least as high as they were before. A conservative conclusion is that once low inflation is achieved, inflation targeting is not harmful to the real economy. Given the strong economic growth after disinflation in many countries (such as New Zealand) that have adopted inflation targets, a case can be made that inflation targeting promotes real economic growth, in addition to controlling inflation.

Nominal GDP Targeting

The concern that a sole focus on inflation may lead to larger output fluctuations has led some economists to propose a variation on inflation targeting in which central banks would target the growth rate of nominal GDP (real GDP times the price level) rather than inflation. Relative to inflation, nominal GDP growth has the advantage that it does put some weight on output as well as prices in the policymaking process. With a nominal GDP target, a decline in projected real output growth would automatically imply an increase in the central bank's inflation target. This increase would tend to be stabilizing, because it would automatically lead to an easier monetary policy.

Nominal GDP targeting is close in spirit to inflation targeting, and although it has the advantages mentioned in the previous paragraph, it has disadvantages as well. First, a nominal GDP target forces the central bank or the government to announce a number for potential (long-term) GDP growth. Such an announcement is highly problematic, because estimates of potential GDP growth are far from precise and change over time. Announcing a specific number for potential GDP growth may thus imply a certainty that policymakers do not have and may also cause the public to mistakenly believe that this estimate is actually a fixed target for potential GDP growth. Announcing a potential GDP growth number is likely to be political dynamite, because it opens policymakers to the criticism that they are willing to settle for long-term growth rates that the public may consider too low. Indeed, a nominal GDP target may lead to an accusation that the central bank or the targeting regime is anti-growth, when the opposite is true, because a low inflation rate is a means to pro-

mote a healthy economy with high growth. In addition, if the estimate for potential GDP growth is higher than the true potential for long-term growth and becomes embedded in the public mind as a target, it can lead to a positive inflation bias.

Second, information on prices is more timely and more frequently reported than data on nominal GDP (and could be made even more so)—a practical consideration that offsets some of the theoretical appeal of nominal GDP as a target. Although collecting data on nominal GDP could be improved, measuring nominal GDP requires data on current quantities and current prices, and the need to collect two pieces of information is perhaps intrinsically more difficult to accomplish in a timely manner.

Third, the concept of inflation in consumer prices is much better understood by the public than the concept of nominal GDP, which can easily be confused with real GDP. Consequently, it seems likely that communication with the public and accountability would be better served by using an inflation rather than a nominal GDP growth target. While a significant number of central banks have adopted inflation targeting, none has adopted a nominal GDP target.

Finally, as argued earlier, inflation targeting, as it is actually practiced, allows considerable flexibility for policy in the short run, and elements of monetary policy tactics based on nominal GDP targeting could easily be built into an inflation-targeting regime. Thus it is doubtful that, in practice, nominal GDP targeting would be more effective than inflation targeting in achieving short-run stabilization.

When all is said and done, inflation targeting has almost all the benefits of nominal GDP targeting, but without the problems that arise from potential confusion about what nominal GDP is or the political complications that arise because nominal GDP requires announcement of a potential GDP growth path.

Monetary Policy with an Implicit Nominal Anchor

In recent years, the United States has achieved excellent macroeconomic performance (including low and stable inflation) without using an explicit nominal anchor such as an exchange rate, a monetary aggregate, or an inflation target. Although the Federal Reserve has not articulated an explicit strategy, a coherent strategy for the conduct of monetary policy exists nonetheless. This strategy involves an implicit but not an explicit nominal anchor in the form of an overriding concern by the Federal Reserve to control inflation in the long run. In addition, it involves forward-looking behavior in which there is careful monitoring for signs of future inflation using a wide range of information, coupled with periodic "pre-emptive strikes" by monetary policy against the threat of inflation.

As emphasized by Milton Friedman, monetary policy effects have long lags. In industrialized countries with a history of low inflation, the inflation process seems to have tremendous inertia: Estimates from large macroeconometric models of the U.S. economy, for example, suggest that monetary policy takes over a year to affect output and over two years to have a significant impact on inflation. For countries whose economies respond more quickly to exchange-rate changes or that have experienced highly variable inflation, and therefore have more flexible prices, the lags may be shorter.

The presence of long lags means that monetary policy cannot wait to respond until inflation has already reared its ugly head. If the central bank waited until overt signs of inflation appeared, it would already be too late to maintain stable prices, at least not without a severe tightening of policy: inflation expectations would already

be embedded in the wage- and price-setting process, creating an inflation momentum that would be hard to halt. Inflation becomes much harder to control once it has been allowed to gather momentum, because higher inflation expectations become ingrained in various types of long-term contracts and pricing agreements.

To prevent inflation from getting started, therefore, monetary policy needs to be forward-looking and pre-emptive: that is, depending on the lags from monetary policy to inflation, monetary policy needs to act long before inflationary pressures appear in the economy. For example, suppose it takes roughly two years for monetary policy to have a significant impact on inflation. In this case, even if inflation is currently low but policymakers believe inflation will rise over the next two years with an unchanged stance of monetary policy, they must *now* tighten monetary policy to prevent the inflationary surge.

Under Alan Greenspan, the Federal Reserve has been successful in pursuing a pre-emptive monetary policy. For example, the Fed raised interest rates from 1994 to 1995 before a rise in inflation got a toehold. As a result, inflation not only did not rise, but fell slightly. This pre-emptive monetary policy strategy is clearly also a feature of inflation-targeting regimes, because monetary policy instruments are adjusted to take account of the long lags in their effects in order to hit future inflation targets. However, the Fed's policy regime, which has no nominal anchor and so might best be described as a "just do it" policy, differs from inflation targeting in that it does not officially have a nominal anchor and is much less transparent in its monetary policy strategy.

Advantages of the Fed's Approach

The Fed's "just do it" approach, which has some of the key elements of inflation targeting, has many of the same advantages. It also enables monetary policy to focus on domestic considerations and does not rely on a stable money–inflation relationship. As with inflation targeting, the central bank uses many sources of information to determine the best settings for monetary policy. The Fed's forward-looking behavior and stress on price stability also help to discourage overly expansionary monetary policy, thereby ameliorating the time-consistency problem.

Another key argument for the "just do it" strategy is its demonstrated success. The Federal Reserve has been able to bring down inflation in the United States from double-digit levels in 1980 to around the 3% level by the end of 1991. Since then, inflation has dropped to around the 2% level, which is arguably consistent with the price stability goal. The Fed conducted a successful pre-emptive strike against inflation from February 1994 until early 1995, when in several steps it raised the federal funds rate from 3% to 6% even though inflation was not increasing during this period. The subsequent lengthy business-cycle expansion, the longest in U.S. history, brought unemployment down to around 4%, a level not seen since the 1960s, while CPI inflation fell to a level near 2%. In addition, the overall U.S. growth rate was very strong throughout the 1990s. Indeed, the performance of the U.S. economy became the envy of the industrialized world in the 1990s.

Disadvantages of the Fed's Approach

Given the success of the "just do it" strategy in the United States, why should the United States consider other monetary policy strategies? (If it ain't broke, why fix it?) The answer is that the "just do it" strategy has some disadvantages that may cause it to work less well in the future.

One disadvantage of the strategy is a lack of transparency. The Fed's close-mouthed approach about its intentions gives rise to a constant guessing game about what it is going to do. This high level of uncertainty leads to unnecessary volatility in

financial markets and creates doubt among producers and the general public about the future course of inflation and output. Furthermore, the opacity of its policymaking makes it hard to hold the Federal Reserve accountable to Congress and the general public: The Fed can't be held accountable if there are no predetermined criteria for judging its performance. Low accountability may make the central bank more susceptible to the time-consistency problem, whereby it may pursue short-term objectives at the expense of long-term ones.

Probably the most serious problem with the "just do it" approach is strong dependence on the preferences, skills, and trustworthiness of the individuals in charge of the central bank. In recent years in the United States, Federal Reserve Chairman Alan Greenspan and other Federal Reserve officials have emphasized forward-looking policies and inflation control, with great success. The Fed's prestige and credibility with the public have risen accordingly. But the Fed's leadership will eventually change, and there is no guarantee that the new team will be committed to the same approach. Nor is there any guarantee that the relatively good working relationship that has existed between the Fed and the executive branch will continue. In a different economic or political environment, the Fed might face strong pressure to engage in over-expansionary policies, raising the possibility that time consistency may become a more serious problem. In the past, after a successful period of low inflation, the Federal Reserve has reverted to inflationary monetary policy—the 1970s are one example—and without an explicit nominal anchor, this could certainly happen again.

Another disadvantage of the "just do it" approach is that it has some inconsistencies with democratic principles. As described in Chapter 14, there are good reasons—notably, insulation from short-term political pressures—for the central bank to have some degree of independence, as the Federal Reserve currently does, and the evidence does generally support central bank independence. Yet the practical economic arguments for central bank independence coexist uneasily with the presumption that government policies should be made democratically, rather than by an elite group.

In contrast, inflation targeting can make the institutional framework for the conduct of monetary policy more consistent with democratic principles and avoid some of the above problems. The inflation-targeting framework promotes the accountability of the central bank to elected officials, who are given some responsibility for setting the goals for monetary policy and then monitoring the economic outcomes. However, under inflation targeting as it has generally been practiced, the central bank has complete control over operational decisions, so that it can be held accountable for achieving its assigned objectives.

Inflation targeting thus can help to promote operational independence of the central bank. The example of the granting of independence to the Bank of England in 1997 indicates how inflation targeting can reduce the tensions between central bank independence and democratic principles and promote central bank independence. When operational independence was granted to the Bank of England in May 1997, the chancellor of the Exchequer made it clear that this action had been made possible by the adoption of an inflation-targeting regime, which had increased the transparency of policy and the accountability of the bank for achieving policy objectives set by the government.

The Fed's monetary policy strategy may move more toward inflation targeting in the future. Inflation targeting is not too far from the Fed's current policymaking philosophy, which has stressed the importance of price stability as the overriding, long-run goal of

monetary policy. Also, a move to inflation targeting is consistent with recent steps by the Fed to increase the transparency of monetary policy, such as shortening the time before the minutes of the FOMC meeting are released, the practice of announcing the FOMC's decision about whether to change the target for the federal funds rates immediately after the conclusion of the FOMC meeting, and the announcement of the "balance of risks" in the future, whether toward higher inflation or toward a weaker economy.

Summary

1. A nominal anchor is a key element in monetary policy strategies. It helps promote price stability by tying down inflation expectations and limiting the time-consistency problem, in which monetary policymakers conduct monetary policy in a discretionary way that produces poor long-run outcomes.
2. Exchange-rate targeting has the following advantages: (1) it directly keeps inflation under control by tying the inflation rate for internationally traded goods to that found in the anchor country to whom its currency is pegged; (2) it provides an automatic rule for the conduct of monetary policy that helps mitigate the time-consistency problem; and (3) it has the advantage of simplicity and clarity. Exchange-rate targeting also has serious disadvantages: (1) it results in a loss of independent monetary policy and increases the exposure of the economy to shocks from the anchor country; (2) it leaves the currency open to speculative attacks; and (3) it can weaken the accountability of policymakers because the exchange-rate signal is lost. Two strategies that make it less likely that the exchange-rate regime will break down are currency boards, in which the central bank stands ready to automatically exchange domestic for foreign currency at a fixed rate, and dollarization, in which a sound currency like the U.S. dollar is adopted as the country's money.
3. Monetary targeting has two main advantages: It enables a central bank to adjust its monetary policy to cope with domestic considerations, and information on whether the central bank is achieving its target is known almost immediately. On the other hand, monetary targeting suffers from the disadvantage that it works well only if there is a reliable relationship between the monetary aggregate and the goal variable, inflation, a relationship that has often not held in different countries.
4. Inflation targeting has several advantages: (1) it enables monetary policy to focus on domestic considerations; (2) stability in the relationship between money and inflation is not critical to its success; (3) it is readily understood by the public and is highly transparent; (4) it increases accountability of the central bank; and (5) it appears to ameliorate the effects of inflationary shocks. It does have some disadvantages, however: (1) inflation is not easily controlled by the monetary authorities, so that an inflation target is unable to send immediate signals to both the public and markets; (2) it might impose a rigid rule on policymakers, although this has not been the case in practice; and (3) a sole focus on inflation may lead to larger output fluctuations, although this has also not been the case in practice. The concern that a sole focus on inflation may lead to larger output fluctuations has led some economists to propose a variant of inflation targeting, nominal GDP targeting, in which central banks target the growth in nominal GDP rather than inflation.
5. The Federal Reserve has a strategy of having an implicit, not an explicit, nominal anchor. This strategy has the following advantages: (1) it enables monetary policy to focus on domestic considerations; (2) it does not rely on a stable money–inflation relationship; and (3) it has had a demonstrated success, producing low inflation with the longest business cycle expansion in U.S. history. However, it does have some disadvantages: (1) it has a lack of transparency; (2) it is strongly dependent on the preferences, skills, and trustworthiness of individuals in the central bank and the government; and (3) it has some inconsistencies with democratic principles, because the central bank is not highly accountable.

Key Terms

currency board, p. 492
dollarization, p. 493
nominal anchor, p. 487
seignorage, p. 493
time-consistency problem, p. 488

Questions and Problems

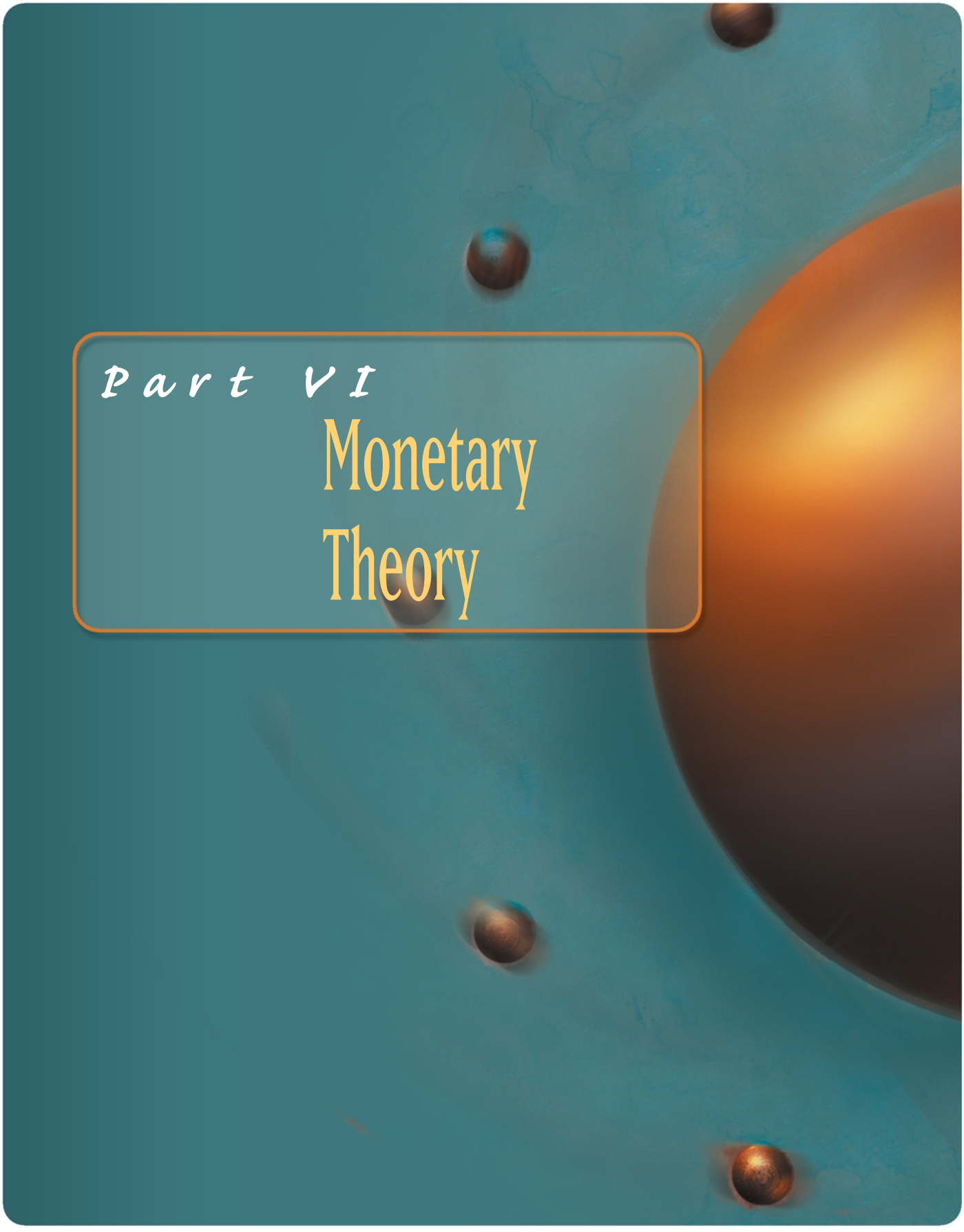
Questions marked with an asterisk are answered at the end of the book in an appendix, "Answers to Selected Questions and Problems."

1. What are the benefits of using a nominal anchor for the conduct of monetary policy?
2. Give an example of the time-consistency problem that you experience in your everyday life.
3. What incentives arise for a central bank to engage in time-consistent behavior?
*4. What are the key advantages of exchange-rate targeting as a monetary policy strategy?
5. Why did the exchange-rate peg lead to difficulties for the countries in the ERM when German reunification occurred?
*6. How can exchange-rate targets lead to a speculative attack on a currency?
7. Why may the disadvantage of exchange-rate targeting of not having an independent monetary policy be less of an issue for emerging market countries than for industrialized countries?
*8. How can the long-term bond market help reduce the time-consistency problem for monetary policy? Can the foreign exchange market also perform this role?
9. When is exchange-rate targeting likely to be a sensible strategy for industrialized countries? When is exchange-rate targeting likely to be a sensible strategy for emerging market countries?
*10. What are the advantages and disadvantages of a currency board over a monetary policy that just uses an exchange-rate target?
11. What are the key advantages and disadvantages of dollarization over other forms of exchange-rate targeting?
*12. What are the advantages of monetary targeting as a strategy for the conduct of monetary policy?
13. What is the big *if* necessary for the success of monetary targeting? Does the experience with monetary targeting suggest that the big *if* is a problem?
*14. What methods have inflation-targeting central banks used to increase communication with the public and increase the transparency of monetary policymaking?
15. Why might inflation targeting increase support for the independence of the central bank to conduct monetary policy?
*16. "Because the public can see whether a central bank hits its monetary targets almost immediately, whereas it takes time before the public can see whether an inflation target is achieved, monetary targeting makes central banks more accountable than inflation targeting does." True, false, or uncertain? Explain.
17. "Because inflation targeting focuses on achieving the inflation target, it will lead to excessive output fluctuations." True, false, or uncertain? Explain.
*18. What are the most important advantages and disadvantages of nominal GDP targeting over inflation targeting?
19. What are the key advantages and disadvantages of the monetary strategy used in the United States under Alan Greenspan in which the nominal anchor is only implicit?
*20. What is the advantage that monetary targeting, inflation targeting, and a monetary strategy with an implicit, but not an explicit, nominal anchor have in common?

Web Exercises

1. Many countries have central banks that are responsible for their nation's monetary policy. Go to www.federalreserve.gov/centralbanks.htm and select one of the central banks (for example, Norway). Review that bank's web site to determine its policies regarding application of monetary policy. How does this bank's policies compare to those of the U.S. central bank?

2. The web provides a rich source of information about international issues. The topic of dollarization has many references. Go to www.imf.org/external/pubs/ft/fandd/2000/03/berg.htm. Summarize this report sponsored by the International Monetary Fund about the value of dollarization.



Part VI

Monetary Theory

Chapter 22

The Demand for Money

PREVIEW

In earlier chapters, we spent a lot of time and effort learning what the money supply is, how it is determined, and what role the Federal Reserve System plays in it. Now we are ready to explore the role of the money supply in determining the price level and total production of goods and services (aggregate output) in the economy. The study of the effect of money on the economy is called **monetary theory**, and we examine this branch of economics in the chapters of Part VI.

When economists mention *supply*, the word *demand* is sure to follow, and the discussion of money is no exception. The supply of money is an essential building block in understanding how monetary policy affects the economy, because it suggests the factors that influence the quantity of money in the economy. Not surprisingly, another essential part of monetary theory is the demand for money.

This chapter describes how the theories of the demand for money have evolved. We begin with the classical theories refined at the start of the twentieth century by economists such as Irving Fisher, Alfred Marshall, and A. C. Pigou; then we move on to the Keynesian theories of the demand for money. We end with Milton Friedman's modern quantity theory.

A central question in monetary theory is whether or to what extent the quantity of money demanded is affected by changes in interest rates. Because this issue is crucial to how we view money's effects on aggregate economic activity, we focus on the role of interest rates in the demand for money.[1]

Quantity Theory of Money

Developed by the classical economists in the nineteenth and early twentieth centuries, the quantity theory of money is a theory of how the nominal value of aggregate income is determined. Because it also tells us how much money is held for a given amount of aggregate income, it is also a theory of the demand for money. The most important feature of this theory is that it suggests that interest rates have no effect on the demand for money.

[1] In Chapter 24, we will see that the responsiveness of the quantity of money demanded to changes in interest rates has important implications for the relative effectiveness of monetary policy and fiscal policy in influencing aggregate economic activity.

Velocity of Money and Equation of Exchange

http://cepa.newschool.edu/het/profiles/fisher.htm

A brief biography and summary of the writings of Irving Fisher.

The clearest exposition of the classical quantity theory approach is found in the work of the American economist Irving Fisher, in his influential book *The Purchasing Power of Money*, published in 1911. Fisher wanted to examine the link between the total quantity of money M (the money supply) and the total amount of spending on final goods and services produced in the economy $P \times Y$, where P is the price level and Y is aggregate output (income). (Total spending $P \times Y$ is also thought of as aggregate nominal income for the economy or as nominal GDP.) The concept that provides the link between M and $P \times Y$ is called the **velocity of money** (often reduced to *velocity*), the rate of turnover of money; that is, the average number of times per year that a dollar is spent in buying the total amount of goods and services produced in the economy. Velocity V is defined more precisely as total spending $P \times Y$ divided by the quantity of money M:

$$V = \frac{P \times Y}{M} \tag{1}$$

If, for example, nominal GDP $(P \times Y)$ in a year is \$5 trillion and the quantity of money is \$1 trillion, velocity is 5, meaning that the average dollar bill is spent five times in purchasing final goods and services in the economy.

By multiplying both sides of this definition by M, we obtain the **equation of exchange**, which relates nominal income to the quantity of money and velocity:

$$M \times V = P \times Y \tag{2}$$

The equation of exchange thus states that the quantity of money multiplied by the number of times that this money is spent in a given year must be equal to nominal income (the total nominal amount spent on goods and services in that year).[2]

As it stands, Equation 2 is nothing more than an identity—a relationship that is true by definition. It does not tell us, for instance, that when the money supply M changes, nominal income $(P \times Y)$ changes in the same direction; a rise in M, for example, could be offset by a fall in V that leaves $M \times V$ (and therefore $P \times Y$) unchanged. To convert the equation of exchange (*an identity*) into a *theory* of how nominal income is determined requires an understanding of the factors that determine velocity.

Irving Fisher reasoned that velocity is determined by the institutions in an economy that affect the way individuals conduct transactions. If people use charge accounts and credit cards to conduct their transactions and consequently use money less often when making purchases, less money is required to conduct the transactions generated by nominal income ($M\downarrow$ relative to $P \times Y$), and velocity $(P \times Y)/M$ will increase. Conversely, if it is more convenient for purchases to be paid for with cash or checks (both of which are money), more money is used to conduct the transactions generated by the same level of nominal income, and velocity will fall. Fisher took the view that

---

[2]Fisher actually first formulated the equation of exchange in terms of the nominal value of transactions in the economy PT:

$$MV_T = PT$$

where P = average price per transaction
T = number of transactions conducted in a year
$V_T = PT/M$ = transactions velocity of money

Because the nominal value of transactions T is difficult to measure, the quantity theory has been formulated in terms of aggregate output Y as follows: T is assumed to be proportional to Y so that $T = vY$, where v is a constant of proportionality. Substituting vY for T in Fisher's equation of exchange yields $MV_T = vPY$, which can be written as Equation 2 in the text, in which $V = V_T/v$.

the institutional and technological features of the economy would affect velocity only slowly over time, so velocity would normally be reasonably constant in the short run.

Quantity Theory

Fisher's view that velocity is fairly constant in the short run transforms the equation of exchange into the **quantity theory of money**, which states that nominal income is determined solely by movements in the quantity of money: When the quantity of money M doubles, $M \times V$ doubles and so must $P \times Y$, the value of nominal income. To see how this works, let's assume that velocity is 5, nominal income (GDP) is initially \$5 trillion, and the money supply is \$1 trillion. If the money supply doubles to \$2 trillion, the quantity theory of money tells us that nominal income will double to \$10 trillion ($= 5 \times$ \$2 trillion).

Because the classical economists (including Fisher) thought that wages and prices were completely flexible, they believed that the level of aggregate output Y produced in the economy during normal times would remain at the full-employment level, so Y in the equation of exchange could also be treated as reasonably constant in the short run. The quantity theory of money then implies that if M doubles, P must also double in the short run, because V and Y are constant. In our example, if aggregate output is \$5 trillion, the velocity of 5 and a money supply of \$1 trillion indicate that the price level equals 1 because 1 times \$5 trillion equals the nominal income of \$5 trillion. When the money supply doubles to \$2 trillion, the price level must also double to 2 because 2 times \$5 trillion equals the nominal income of \$10 trillion.

For the classical economists, the quantity theory of money provided an explanation of movements in the price level: ***Movements in the price level result solely from changes in the quantity of money.***

Quantity Theory of Money Demand

Because the quantity theory of money tells us how much money is held for a given amount of aggregate income, it is in fact a theory of the demand for money. We can see this by dividing both sides of the equation of exchange by V, thus rewriting it as:

$$M = \frac{1}{V} \times PY$$

where nominal income $P \times Y$ is written as PY. When the money market is in equilibrium, the quantity of money M that people hold equals the quantity of money demanded M^d, so we can replace M in the equation by M^d. Using k to represent the quantity $1/V$ (a constant, because V is a constant), we can rewrite the equation as:

$$M^d = k \times PY \tag{3}$$

Equation 3 tells us that because k is a constant, the level of transactions generated by a fixed level of nominal income PY determines the quantity of money M^d that people demand. Therefore, Fisher's quantity theory of money suggests that the demand for money is purely a function of income, and interest rates have no effect on the demand for money.[3]

Fisher came to this conclusion because he believed that people hold money only to conduct transactions and have no freedom of action in terms of the amount they want to hold. The demand for money is determined (1) by the level of transactions generated

[3]While Fisher was developing his quantity theory approach to the demand for money, a group of classical economists in Cambridge, England, came to similar conclusions, although with slightly different reasoning. They derived Equation 3 by recognizing that two properties of money motivate people to hold it: its utility as a medium of exchange and as a store of wealth.

by the level of nominal income *PY* and (2) by the institutions in the economy that affect the way people conduct transactions and thus determine velocity and hence *k*.

Is Velocity a Constant?

The classical economists' conclusion that nominal income is determined by movements in the money supply rested on their belief that velocity *PY/M* could be treated as reasonably constant.[4] Is it reasonable to assume that velocity is constant? To answer this, let's look at Figure 1, which shows the year-to-year changes in velocity from 1915 to 2002 (nominal income is represented by nominal GDP and the money supply by M1 and M2).

What we see in Figure 1 is that even in the short run, velocity fluctuates too much to be viewed as a constant. Prior to 1950, velocity exhibited large swings up and down. This may reflect the substantial instability of the economy in this period, which included two world wars and the Great Depression. (Velocity actually falls, or at least its rate of growth declines, in years when recessions are taking place.) After 1950, velocity appears to have more moderate fluctuations, yet there are large differences in

www.usagold.com/gildedopinion/puplava/20020614.html
A summary of how various factors affect the velocity of money.

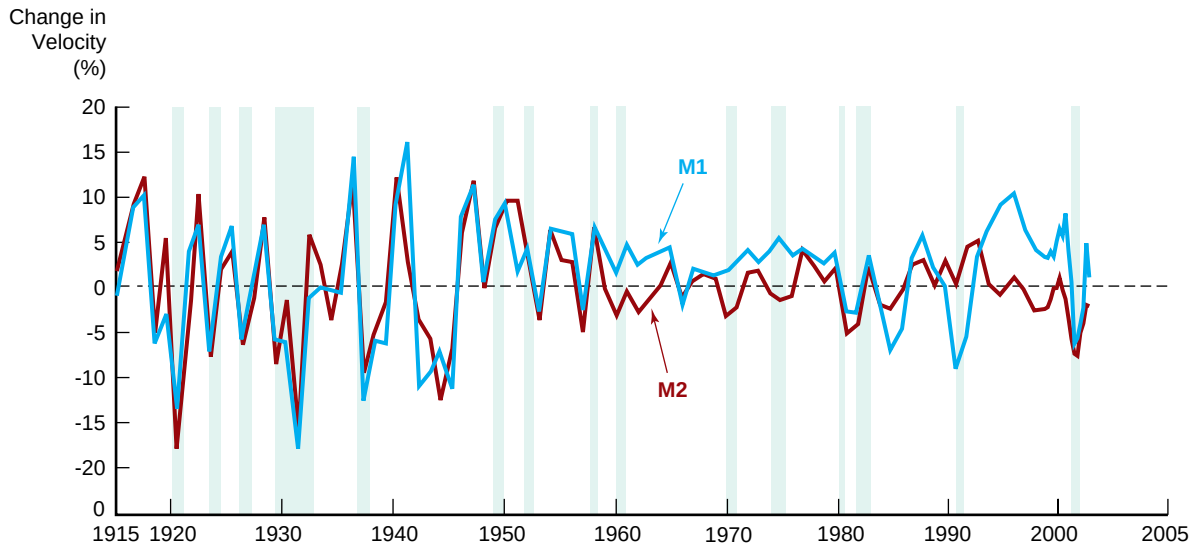


FIGURE 1 Change in the Velocity of M1 and M2 from Year to Year, 1915–2002
Shaded areas indicate recessions. Velocities are calculated using nominal GNP before 1959 and nominal GDP thereafter.
Sources: Economic Report of the President; Banking and Monetary Statistics; www.federalreserve.gov/releases/h6/.

[4]Actually, the classical conclusion still holds if velocity grows at some uniform rate over time that reflects changes in transaction technology. Hence the concept of a constant velocity should more accurately be thought of here as a lack of upward and downward fluctuations in velocity.

the growth rate of velocity from year to year. The percentage change in M1 velocity (GDP/M1) from 1981 to 1982, for example, was −2.5%, whereas from 1980 to 1981 velocity grew at a rate of 4.2%. This difference of 6.7% means that nominal GDP was 6.7% lower than it would have been if velocity had kept growing at the same rate as in 1980–1981.[5] The drop is enough to account for the severe recession that took place in 1981–1982. After 1982, M1 velocity appears to have become even more volatile, a fact that has puzzled researchers when they examine the empirical evidence on the demand for money (discussed later in this chapter). M2 velocity remained more stable than M1 velocity after 1982, with the result that the Federal Reserve dropped its M1 targets in 1987 and began to focus more on M2 targets. However, instability of M2 velocity in the early 1990s resulted in the Fed's announcement in July 1993 that it no longer felt that any of the monetary aggregates, including M2, was a reliable guide for monetary policy.

Until the Great Depression, economists did not recognize that velocity declines sharply during severe economic contractions. Why did the classical economists not recognize this fact when it is easy to see in the pre-Depression period in Figure 1? Unfortunately, accurate data on GDP and the money supply did not exist before World War II. (Only after the war did the government start to collect these data.) Economists had no way of knowing that their view of velocity as a constant was demonstrably false. The decline in velocity during the Great Depression years was so great, however, that even the crude data available to economists at that time suggested that velocity was not constant. This explains why, after the Great Depression, economists began to search for other factors influencing the demand for money that might help explain the large fluctuations in velocity.

Let us now examine the theories of money demand that arose from this search for a better explanation of the behavior of velocity.

Keynes's Liquidity Preference Theory

http://www-gap.dcs.st-and.ac.uk/~history/Mathematicians/Keynes.html
A brief history of John Maynard Keynes.

In his famous 1936 book *The General Theory of Employment, Interest, and Money*, John Maynard Keynes abandoned the classical view that velocity was a constant and developed a theory of money demand that emphasized the importance of interest rates. His theory of the demand for money, which he called the **liquidity preference theory**, asked the question: Why do individuals hold money? He postulated that there are three motives behind the demand for money: the transactions motive, the precautionary motive, and the speculative motive.

Transactions Motive

In the classical approach, individuals are assumed to hold money because it is a medium of exchange that can be used to carry out everyday transactions. Following the classical tradition, Keynes emphasized that this component of the demand for money is determined primarily by the level of people's transactions. Because he believed that these transactions were proportional to income, like the classical economists, he took the transactions component of the demand for money to be proportional to income.

[5] We reach a similar conclusion if we use M2 velocity. The percentage change in M2 velocity (GDP/M2) from 1981 to 1982 was −5.0%, whereas from 1980 to 1981 it was +2.3%. This difference of 7.3% means that nominal GDP was 7.3% lower than it would have been if M2 velocity had kept growing at the same rate as in 1980–1981.

Precautionary Motive

Keynes went beyond the classical analysis by recognizing that in addition to holding money to carry out current transactions, people hold money as a cushion against an unexpected need. Suppose that you've been thinking about buying a fancy stereo; you walk by a store that is having a 50%-off sale on the one you want. If you are holding money as a precaution for just such an occurrence, you can purchase the stereo right away; if you are not holding precautionary money balances, you cannot take advantage of the sale. Precautionary money balances also come in handy if you are hit with an unexpected bill, say for car repair or hospitalization.

Keynes believed that the amount of precautionary money balances people want to hold is determined primarily by the level of transactions that they expect to make in the future and that these transactions are proportional to income. Therefore, he postulated, the demand for precautionary money balances is proportional to income.

Speculative Motive

If Keynes had ended his theory with the transactions and precautionary motives, income would be the only important determinant of the demand for money, and he would not have added much to the classical approach. However, Keynes took the view that money is a store of wealth and called this reason for holding money the *speculative motive*. Since he believed that wealth is tied closely to income, the speculative component of money demand would be related to income. However, Keynes looked more carefully at the factors that influence the decisions regarding how much money to hold as a store of wealth, especially interest rates.

Keynes divided the assets that can be used to store wealth into two categories: money and bonds. He then asked the following question: Why would individuals decide to hold their wealth in the form of money rather than bonds?

Thinking back to the discussion of the theory of asset demand (Chapter 5), you would want to hold money if its expected return was greater than the expected return from holding bonds. Keynes assumed that the expected return on money was zero because in his time, unlike today, most checkable deposits did not earn interest. For bonds, there are two components of the expected return: the interest payment and the *expected* rate of capital gains.

You learned in Chapter 4 that when interest rates rise, the price of a bond falls. If you expect interest rates to rise, you expect the price of the bond to fall and therefore suffer a negative capital gain—that is, a capital loss. If you expect the rise in interest rates to be substantial enough, the capital loss might outweigh the interest payment, and your *expected* return on the bond would be negative. In this case, you would want to store your wealth as money because its expected return is higher; its zero return exceeds the negative return on the bond.

Keynes assumed that individuals believe that interest rates gravitate to some normal value (an assumption less plausible in today's world). If interest rates are below this normal value, individuals expect the interest rate on bonds to rise in the future and so expect to suffer capital losses on them. As a result, individuals will be more likely to hold their wealth as money rather than bonds, and the demand for money will be high.

What would you expect to happen to the demand for money when interest rates are above the normal value? In general, people will expect interest rates to fall, bond prices to rise, and capital gains to be realized. At higher interest rates, they are more likely to expect the return from holding a bond to be positive, thus exceeding the expected return from holding money. They will be more likely to hold bonds than money, and the demand for money will be quite low. From Keynes's reasoning, we can conclude that as interest rates rise, the demand for money falls, and therefore ***money demand is negatively related to the level of interest rates***.

Putting the Three Motives Together

In putting the three motives for holding money balances together into a demand for money equation, Keynes was careful to distinguish between nominal quantities and real quantities. Money is valued in terms of what it can buy. If, for example, all prices in the economy double (the price level doubles), the same nominal quantity of money will be able to buy only half as many goods. Keynes thus reasoned that people want to hold a certain amount of **real money balances** (the quantity of money in real terms)—an amount that his three motives indicated would be related to real income Y and to interest rates i. Keynes wrote down the following demand for money equation, known as the *liquidity preference function*, which says that the demand for real money balances M^d/P is a function of (related to) i and Y:[6]

$$\frac{M^d}{P} = f(\underset{-}{i}, \underset{+}{Y}) \tag{4}$$

The minus sign below i in the liquidity preference function means that the demand for real money balances is negatively related to the interest rate i, and the plus sign below Y means that the demand for real money balances and real income Y are positively related. This money demand function is the same one that was used in our analysis of money demand discussed in Chapter 5. Keynes's conclusion that the demand for money is related not only to income but also to interest rates is a major departure from Fisher's view of money demand, in which interest rates can have no effect on the demand for money.

By deriving the liquidity preference function for velocity PY/M, we can see that Keynes's theory of the demand for money implies that velocity is not constant, but instead fluctuates with movements in interest rates. The liquidity preference equation can be rewritten as:

$$\frac{P}{M^d} = \frac{1}{f(i, Y)}$$

Multiplying both sides of this equation by Y and recognizing that M^d can be replaced by M because they must be equal in money market equilibrium, we solve for velocity:

$$V = \frac{PY}{M} = \frac{Y}{f(i, Y)} \tag{5}$$

We know that the demand for money is negatively related to interest rates; when i goes up, $f(i, Y)$ declines, and therefore velocity rises. In other words, a rise in interest rates encourages people to hold lower real money balances for a given level of income; therefore, the rate of turnover of money (velocity) must be higher. This reasoning implies that because interest rates have substantial fluctuations, the liquidity preference theory of the demand for money indicates that velocity has substantial fluctuations as well.

An interesting feature of Equation 5 is that it explains some of the velocity movements in Figure 1, in which we noted that when recessions occur, velocity falls or its rate of growth declines. What fact regarding the cyclical behavior of interest rates (discussed in Chapter 5) might help us explain this phenomenon? You might recall that

---

[6]The classical economists' money demand equation can also be written in terms of real money balances by dividing both sides of Equation 3 by the price level P to obtain:

$$\frac{M^d}{P} = k \times Y$$

interest rates are procyclical, rising in expansions and falling in recessions. The liquidity preference theory indicates that a rise in interest rates will cause velocity to rise also. The procyclical movements of interest rates should induce procyclical movements in velocity, and that is exactly what we see in Figure 1.

Keynes's model of the speculative demand for money provides another reason why velocity might show substantial fluctuations. What would happen to the demand for money if the view of the normal level of interest rates changes? For example, what if people expect the future normal interest rate to be higher than the current normal interest rate? Because interest rates are then expected to be higher in the future, more people will expect the prices of bonds to fall and will anticipate capital losses. The expected returns from holding bonds will decline, and money will become more attractive relative to bonds. As a result, the demand for money will increase. This means that $f(i, Y)$ will increase and so velocity will fall. Velocity will change as expectations about future normal levels of interest rates change, and unstable expectations about future movements in normal interest rates can lead to instability of velocity. This is one more reason why Keynes rejected the view that velocity could be treated as a constant.

Study Guide

Keynes's explanation of how interest rates affect the demand for money will be easier to understand if you think of yourself as an investor who is trying to decide whether to invest in bonds or to hold money. Ask yourself what you would do if you expected the normal interest rate to be lower in the future than it is currently. Would you rather be holding bonds or money?

To sum up, Keynes's liquidity preference theory postulated three motives for holding money: the transactions motive, the precautionary motive, and the speculative motive. Although Keynes took the transactions and precautionary components of the demand for money to be proportional to income, he reasoned that the speculative motive would be negatively related to the level of interest rates.

Keynes's model of the demand for money has the important implication that velocity is not constant, but instead is positively related to interest rates, which fluctuate substantially. His theory also rejected the constancy of velocity, because changes in people's expectations about the normal level of interest rates would cause shifts in the demand for money that would cause velocity to shift as well. Thus Keynes's liquidity preference theory casts doubt on the classical quantity theory that nominal income is determined primarily by movements in the quantity of money.

Further Developments in the Keynesian Approach

After World War II, economists began to take the Keynesian approach to the demand for money even further by developing more precise theories to explain the three Keynesian motives for holding money. Because interest rates were viewed as a crucial element in monetary theory, a key focus of this research was to understand better the role of interest rates in the demand for money.

Transactions Demand

William Baumol and James Tobin independently developed similar demand for money models, which demonstrated that even money balances held for transactions

purposes are sensitive to the level of interest rates.[7] In developing their models, they considered a hypothetical individual who receives a payment once a period and spends it over the course of this period. In their model, money, which earns zero interest, is held only because it can be used to carry out transactions.

To refine this analysis, let's say that Grant Smith receives $1,000 at the beginning of the month and spends it on transactions that occur at a constant rate during the course of the month. If Grant keeps the $1,000 in cash in order to carry out his transactions, his money balances follow the sawtooth pattern displayed in panel (a) of Figure 2. At the beginning of the month he has $1,000, and by the end of the month he has no cash left because he has spent it all. Over the course of the month, his holdings of money will on average be $500 (his holdings at the beginning of the month, $1,000, plus his holdings at the end of the month, $0, divided by 2).

At the beginning of the next month, Grant receives another $1,000 payment, which he holds as cash, and the same decline in money balances begins again. This process repeats monthly, and his average money balance during the course of the year is $500. Since his yearly nominal income is $12,000 and his holdings of money average $500, the velocity of money ($V = PY/M$) is $12,000/$500 = 24.

Suppose that as a result of taking a money and banking course, Grant realizes that he can improve his situation by not always holding cash. In January, then, he decides to hold part of his $1,000 in cash and puts part of it into an income-earning security such as bonds. At the beginning of each month, Grant keeps $500 in cash and uses the other $500 to buy a Treasury bond. As you can see in panel (b), he starts out each

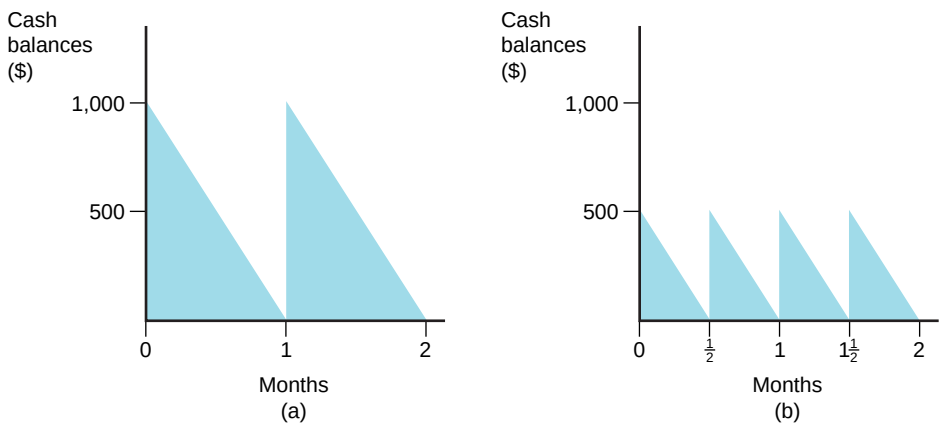


FIGURE 2 Cash Balances in the Baumol-Tobin Model
In panel (a), the $1,000 payment at the beginning of the month is held entirely in cash and is spent at a constant rate until it is exhausted by the end of the month. In panel (b), half of the monthly payment is put into cash and the other half into bonds. At the middle of the month, cash balances reach zero and bonds must be sold to bring balances up to $500. By the end of the month, cash balances again dwindle to zero.

[7] William J. Baumol, "The Transactions Demand for Cash: An Inventory Theoretic Approach," *Quarterly Journal of Economics* 66 (1952): 545–556; James Tobin, "The Interest Elasticity of the Transactions Demand for Cash," *Review of Economics and Statistics* 38 (1956): 241–247.

month with $500 of cash, and by the middle of the month, his cash balance has run down to zero. Because bonds cannot be used directly to carry out transactions, Grant must sell them and turn them into cash so that he can carry out the rest of the month's transactions. At the middle of the month, then, Grant's cash balance rises back up to $500. By the end of the month, the cash is gone. When he again receives his next $1,000 monthly payment, he again divides it into $500 of cash and $500 of bonds, and the process continues. The net result of this process is that the average cash balance held during the month is $500/2 = $250—just half of what it was before. Velocity has doubled to $12,000/$250 = 48.

What has Grant Smith gained from his new strategy? He has earned interest on $500 of bonds that he held for half the month. If the interest rate is 1% per month, he has earned an additional $2.50 (= $\frac{1}{2} \times \$500 \times 1\%$) per month.

Sounds like a pretty good deal, doesn't it? In fact, if he had kept $333.33 in cash at the beginning of the month, he would have been able to hold $666.67 in bonds for the first third of the month. Then he could have sold $333.33 of bonds and held on to $333.34 of bonds for the next third of the month. Finally, two-thirds of the way through the month, he would have had to sell the remaining bonds to raise cash. The net result of this is that Grant would have earned $3.33 per month [= ($\frac{1}{3} \times \$666.67 \times 1\%$) + ($\frac{1}{3} \times \$333.34 \times 1\%$)]. This is an even better deal. His average cash holdings in this case would be $333.33/2 = $166.67. Clearly, the lower his average cash balance, the more interest he will earn.

As you might expect, there is a catch to all this. In buying bonds, Grant incurs transaction costs of two types. First, he must pay a straight brokerage fee for the buying and selling of the bonds. These fees increase when average cash balances are lower because Grant will be buying and selling bonds more often. Second, by holding less cash, he will have to make more trips to the bank to get the cash, once he has sold some of his bonds. Because time is money, this must also be counted as part of the transaction costs.

Grant faces a trade-off. If he holds very little cash, he can earn a lot of interest on bonds, but he will incur greater transaction costs. If the interest rate is high, the benefits of holding bonds will be high relative to the transaction costs, and he will hold more bonds and less cash. Conversely, if interest rates are low, the transaction costs involved in holding a lot of bonds may outweigh the interest payments, and Grant would then be better off holding more cash and fewer bonds.

The conclusion of the Baumol-Tobin analysis may be stated as follows: As interest rates increase, the amount of cash held for transactions purposes will decline, which in turn means that velocity will increase as interest rates increase.[8] Put another way, the ***transactions component of the demand for money is negatively related to the level of interest rates***.

The basic idea in the Baumol-Tobin analysis is that there is an opportunity cost of holding money—the interest that can be earned on other assets. There is also a benefit to holding money—the avoidance of transaction costs. When interest rates increase, people will try to economize on their holdings of money for transactions purposes, because the opportunity cost of holding money has increased. By using

[8]Similar reasoning leads to the conclusion that as brokerage fees increase, the demand for transactions money balances increases as well. When these fees rise, the benefits from holding transactions money balances increase because by holding these balances, an individual will not have to sell bonds as often, thereby avoiding these higher brokerage costs. The greater benefits to holding money balances relative to the opportunity cost of interest forgone, then, lead to a higher demand for transactions balances.

simple models, Baumol and Tobin revealed something that we might not otherwise have seen: that the transactions demand for money, and not just the speculative demand, will be sensitive to interest rates. The Baumol-Tobin analysis presents a nice demonstration of the value of economic modeling.[9]

Study Guide

The idea that as interest rates increase, the opportunity cost of holding money increases so that the demand for money falls, can be stated equivalently with the terminology of expected returns used earlier. As interest rates increase, the expected return on the other asset, bonds, increases, causing the relative expected return on money to fall, thereby lowering the demand for money. These two explanations are in fact identical, because as we saw in Chapter 5, changes in the opportunity cost of an asset are just a description of what is happening to the relative expected return. The opportunity cost terminology was used by Baumol and Tobin in their work on the transactions demand for money, and that is why we used this terminology in the text. To make sure you understand the equivalence of the two terminologies, try to translate the reasoning in the precautionary demand discussion from opportunity cost terminology to expected returns terminology.

Precautionary Demand

Models that explore the precautionary motive of the demand for money have been developed along lines similar to the Baumol-Tobin framework, so we will not go into great detail about them here. We have already discussed the benefits of holding precautionary money balances, but weighed against these benefits must be the opportunity cost of the interest forgone by holding money. We therefore have a trade-off similar to the one for transactions balances. As interest rates rise, the opportunity cost of holding precautionary balances rises, and so the holdings of these money balances fall. We then have a result similar to the one found for the Baumol-Tobin analysis.[10] ***The precautionary demand for money is negatively related to interest rates.***

Speculative Demand

Keynes's analysis of the speculative demand for money was open to several serious criticisms. It indicated that an individual holds only money as a store of wealth when the expected return on bonds is less than the expected return on money and holds only bonds when the expected return on bonds is greater than the expected return on money. Only when people have expected returns on bonds and money that are exactly equal (a rare instance) would they hold both. Keynes's analysis therefore implies that practically no one holds a diversified portfolio of bonds and money simultaneously as a store of wealth. Since diversification is apparently a sensible strategy for choosing which assets to hold, the fact that it rarely occurs in Keynes's analysis is a serious shortcoming of his theory of the speculative demand for money.

Tobin developed a model of the speculative demand for money that attempted to avoid this criticism of Keynes's analysis.[11] His basic idea was that not only do people

[9]The mathematics behind the Baumol-Tobin model can be found in an appendix to this chapter on this book's web site at www.aw.com/mishkin.

[10]These models of the precautionary demand for money also reveal that as uncertainty about the level of future transactions grows, the precautionary demand for money increases. This is so because greater uncertainty means that individuals are more likely to incur transaction costs if they are not holding precautionary balances. The benefit of holding such balances then increases relative to the opportunity cost of forgone interest, and so the demand for them rises.

[11]James Tobin, "Liquidity Preference as Behavior Towards Risk," *Review of Economic Studies* 25 (1958): 65–86.

care about the expected return on one asset versus another when they decide what to hold in their portfolio, but they also care about the riskiness of the returns from each asset. Specifically, Tobin assumed that most people are risk-averse—that they would be willing to hold an asset with a lower expected return if it is less risky. An important characteristic of money is that its return is certain; Tobin assumed it to be zero. Bonds, by contrast, can have substantial fluctuations in price, and their returns can be quite risky and sometimes negative. So even if the expected returns on bonds exceed the expected return on money, people might still want to hold money as a store of wealth because it has less risk associated with its return than bonds do.

The Tobin analysis also shows that people can reduce the total amount of risk in a portfolio by diversifying; that is, by holding both bonds and money. The model suggests that individuals will hold bonds and money simultaneously as stores of wealth. Since this is probably a more realistic description of people's behavior than Keynes's, Tobin's rationale for the speculative demand for money seems to rest on more solid ground.

Tobin's attempt to improve on Keynes's rationale for the speculative demand for money was only partly successful, however. It is still not clear that the speculative demand even exists. What if there are assets that have no risk—like money—but earn a higher return? Will there be any speculative demand for money? No, because an individual will always be better off holding such an asset rather than money. The resulting portfolio will enjoy a higher expected return yet has no higher risk. Do such assets exist in the American economy? The answer is yes. U.S. Treasury bills and other assets that have no default risk provide certain returns that are greater than those available on money. Therefore, why would anyone want to hold money balances as a store of wealth (ignoring for the moment transactions and precautionary reasons)?

Although Tobin's analysis did not explain why money is held as a store of wealth, it was an important development in our understanding of how people should choose among assets. Indeed, his analysis was an important step in the development of the academic field of finance, which examines asset pricing and portfolio choice (the decision to buy one asset over another).

To sum up, further developments of the Keynesian approach have attempted to give a more precise explanation for the transactions, precautionary, and speculative demand for money. The attempt to improve Keynes's rationale for the speculative demand for money has been only partly successful; it is still not clear that this demand even exists. However, the models of the transactions and precautionary demand for money indicate that these components of money demand are negatively related to interest rates. Hence Keynes's proposition that the demand for money is sensitive to interest rates—suggesting that velocity is not constant and that nominal income might be affected by factors other than the quantity of money—is still supported.

Friedman's Modern Quantity Theory of Money

In 1956, Milton Friedman developed a theory of the demand for money in a famous article, "The Quantity Theory of Money: A Restatement."[12] Although Friedman frequently refers to Irving Fisher and the quantity theory, his analysis of the demand for money is actually closer to that of Keynes than it is to Fisher's.

[12]Milton Friedman, "The Quantity Theory of Money: A Restatement," in *Studies in the Quantity Theory of Money*, ed. Milton Friedman (Chicago: University of Chicago Press, 1956), pp. 3–21.

Like his predecessors, Friedman pursued the question of why people choose to hold money. Instead of analyzing the specific motives for holding money, as Keynes did, Friedman simply stated that the demand for money must be influenced by the same factors that influence the demand for any asset. Friedman then applied the theory of asset demand to money.

The theory of asset demand (Chapter 5) indicates that the demand for money should be a function of the resources available to individuals (their wealth) and the expected returns on other assets relative to the expected return on money. Like Keynes, Friedman recognized that people want to hold a certain amount of real money balances (the quantity of money in real terms). From this reasoning, Friedman expressed his formulation of the demand for money as follows:

$$\frac{M^d}{P} = f(\underset{+}{Y_p}, \underset{-}{r_b - r_m}, \underset{-}{r_e - r_m}, \underset{-}{\pi^e - r_m}) \tag{6}$$

where M^d/P = demand for real money balances
Y_p = Friedman's measure of wealth, known as *permanent income* (technically, the present discounted value of all expected future income, but more easily described as expected average long-run income)
r_m = expected return on money
r_b = expected return on bonds
r_e = expected return on equity (common stocks)
π^e = expected inflation rate

and the signs underneath the equation indicate whether the demand for money is positively (+) related or negatively (−) related to the terms that are immediately above them.[13]

Let us look in more detail at the variables in Friedman's money demand function and what they imply for the demand for money.

Because the demand for an asset is positively related to wealth, money demand is positively related to Friedman's wealth concept, permanent income (indicated by the plus sign beneath it). Unlike our usual concept of income, permanent income (which can be thought of as expected average long-run income) has much smaller short-run fluctuations, because many movements of income are transitory (short-lived). For example, in a business cycle expansion, income increases rapidly, but because some of this increase is temporary, average long-run income does not change very much. Hence in a boom, permanent income rises much less than income. During a recession, much of the income decline is transitory, and average long-run income (hence permanent income) falls less than income. One implication of Friedman's use of the concept of permanent income as a determinant of the demand for money is that the demand for money will not fluctuate much with business cycle movements.

[13]Friedman also added to his formulation a term h that represented the ratio of human to nonhuman wealth. He reasoned that if people had more permanent income coming from labor income and thus from their human capital, they would be less liquid than if they were receiving income from financial assets. In this case, they might want to hold more money because it is a more liquid asset than the alternatives. The term h plays no essential role in Friedman's theory and has no important implications for monetary theory. That is why we ignore it in the money demand function.

An individual can hold wealth in several forms besides money; Friedman categorized them into three types of assets: bonds, equity (common stocks), and goods. The incentives for holding these assets rather than money are represented by the expected return on each of these assets relative to the expected return on money, the last three terms in the money demand function. The minus sign beneath each indicates that as each term rises, the demand for money will fall.

The expected return on money r_m, which appears in all three terms, is influenced by two factors:

1. The services provided by banks on deposits included in the money supply, such as provision of receipts in the form of canceled checks or the automatic paying of bills. When these services are increased, the expected return from holding money rises.
2. The interest payments on money balances. NOW accounts and other deposits that are included in the money supply currently pay interest. As these interest payments rise, the expected return on money rises.

The terms $r_b - r_m$ and $r_e - r_m$ represent the expected return on bonds and equity relative to money; as they rise, the relative expected return on money falls, and the demand for money falls. The final term, $\pi^e - r_m$, represents the expected return on goods relative to money. The expected return from holding goods is the expected rate of capital gains that occurs when their prices rise and hence is equal to the expected inflation rate π^e. If the expected inflation rate is 10%, for example, then goods' prices are expected to rise at a 10% rate, and their expected return is 10%. When $\pi^e - r_m$ rises, the expected return on goods relative to money rises, and the demand for money falls.

Distinguishing Between the Friedman and Keynesian Theories

There are several differences between Friedman's theory of the demand for money and the Keynesian theories. One is that by including many assets as alternatives to money, Friedman recognized that more than one interest rate is important to the operation of the aggregate economy. Keynes, for his part, lumped financial assets other than money into one big category—bonds—because he felt that their returns generally move together. If this is so, the expected return on bonds will be a good indicator of the expected return on other financial assets, and there will be no need to include them separately in the money demand function.

Also in contrast to Keynes, Friedman viewed money and goods as substitutes; that is, people choose between them when deciding how much money to hold. That is why Friedman included the expected return on goods relative to money as a term in his money demand function. The assumption that money and goods are substitutes indicates that changes in the quantity of money may have a direct effect on aggregate spending.

In addition, Friedman stressed two issues in discussing his demand for money function that distinguish it from Keynes's liquidity preference theory. First, Friedman did not take the expected return on money to be a constant, as Keynes did. When interest rates rise in the economy, banks make more profits on their loans, and they want to attract more deposits to increase the volume of their now more profitable loans. If there are no restrictions on interest payments on deposits, banks attract deposits by paying higher interest rates on them. Because the industry is competitive, the expected return on money held as bank deposits then rises with the higher interest rates on bonds and loans. The banks compete to get deposits until there are no

excess profits, and in doing so they close the gap between interest earned on loans and interest paid on deposits. The net result of this competition in the banking industry is that $r_b - r_m$ stays relatively constant when the interest rate i rises.[14]

What if there are restrictions on the amount of interest that banks can pay on their deposits? Will the expected return on money be a constant? As interest rates rise, will $r_b - r_m$ rise as well? Friedman thought not. He argued that although banks might be restricted from making pecuniary payments on their deposits, they can still compete on the quality dimension. For example, they can provide more services to depositors by hiring more tellers, paying bills automatically, or making more cash machines available at more accessible locations. The result of these improvements in money services is that the expected return from holding deposits will rise. So despite the restrictions on pecuniary interest payments, we might still find that a rise in market interest rates will raise the expected return on money sufficiently so that $r_b - r_m$ will remain relatively constant.[15] ***Unlike Keynes's theory, which indicates that interest rates are an important determinant of the demand for money, Friedman's theory suggests that changes in interest rates should have little effect on the demand for money.***

Therefore, Friedman's money demand function is essentially one in which permanent income is the primary determinant of money demand, and his money demand equation can be approximated by:

$$\frac{M^d}{P} = f(Y_p) \tag{7}$$

In Friedman's view, the demand for money is insensitive to interest rates—not because he viewed the demand for money as insensitive to changes in the incentives for holding other assets relative to money, but rather because changes in interest rates should have little effect on these incentive terms in the money demand function. The incentive terms remain relatively constant, because any rise in the expected returns on other assets as a result of the rise in interest rates would be matched by a rise in the expected return on money.

The second issue Friedman stressed is the stability of the demand for money function. In contrast to Keynes, Friedman suggested that random fluctuations in the demand for money are small and that the demand for money can be predicted accurately by the money demand function. When combined with his view that the demand for money is insensitive to changes in interest rates, this means that velocity is highly predictable. We can see this by writing down the velocity that is implied by the money demand equation (Equation 7):

$$V = \frac{Y}{f(Y_p)} \tag{8}$$

Because the relationship between Y and Y_p is usually quite predictable, a stable money demand function (one that does not undergo pronounced shifts, so that it predicts the

[14]Friedman does suggest that there is some increase in $r_b - r_m$ when i rises because part of the money supply (especially currency) is held in forms that cannot pay interest in a pecuniary or nonpecuniary form. See, for example, Milton Friedman, "Why a Surge of Inflation Is Likely Next Year," *Wall Street Journal*, September 1, 1983, p. 24.

[15]Competing on the quality of services is characteristic of many industries that are restricted from competing on price. For example, in the 1960s and early 1970s, when airfares were set high by the Civil Aeronautics Board, airlines were not allowed to lower their fares to attract customers. Instead, they improved the quality of their service by providing free wine, fancier food, piano bars, movies, and wider seats.

demand for money accurately) implies that velocity is predictable as well. If we can predict what velocity will be in the next period, a change in the quantity of money will produce a predictable change in aggregate spending. Even though velocity is no longer assumed to be constant, the money supply continues to be the primary determinant of nominal income as in the quantity theory of money. Therefore, Friedman's theory of money demand is indeed a restatement of the quantity theory, because it leads to the same conclusion about the importance of money to aggregate spending.

You may recall that we said that the Keynesian liquidity preference function (in which interest rates are an important determinant of the demand for money) is able to explain the procyclical movements of velocity that we find in the data. Can Friedman's money demand formulation explain this procyclical velocity phenomenon as well?

The key clue to answering this question is the presence of permanent income rather than measured income in the money demand function. What happens to permanent income in a business cycle expansion? Because much of the increase in income will be transitory, permanent income rises much less than income. Friedman's money demand function then indicates that the demand for money rises only a small amount relative to the rise in measured income, and as Equation 8 indicates, velocity rises. Similarly, in a recession, the demand for money falls less than income, because the decline in permanent income is small relative to income, and velocity falls. In this way, we have the procyclical movement in velocity.

To summarize, Friedman's theory of the demand for money used a similar approach to that of Keynes but did not go into detail about the motives for holding money. Instead, Friedman made use of the theory of asset demand to indicate that the demand for money will be a function of permanent income and the expected returns on alternative assets relative to the expected return on money. There are two major differences between Friedman's theory and Keynes's. Friedman believed that changes in interest rates have little effect on the expected returns on other assets relative to money. Thus, in contrast to Keynes, he viewed the demand for money as insensitive to interest rates. In addition, he differed from Keynes in stressing that the money demand function does not undergo substantial shifts and is therefore stable. These two differences also indicate that velocity is predictable, yielding a quantity theory conclusion that money is the primary determinant of aggregate spending.

Empirical Evidence on the Demand for Money

As we have seen, the alternative theories of the demand for money can have very different implications for our view of the role of money in the economy. Which of these theories is an accurate description of the real world is an important question, and it is the reason why evidence on the demand for money has been at the center of many debates on the effects of monetary policy on aggregate economic activity. Here we examine the empirical evidence on the two primary issues that distinguish the different theories of money demand and affect their conclusions about whether the quantity of money is the primary determinant of aggregate spending: Is the demand for money sensitive to changes in interest rates, and is the demand for money function stable over time?[16]

[16]If you are interested in a more detailed discussion of the empirical research on the demand for money, you can find it in an appendix to this chapter on this book's web site at www.aw.com/mishkin.

Interest Rates and Money Demand

Earlier in the chapter, we saw that if interest rates do not affect the demand for money, velocity is more likely to be a constant—or at least predictable—so that the quantity theory view that aggregate spending is determined by the quantity of money is more likely to be true. However, the more sensitive the demand for money is to interest rates, the more unpredictable velocity will be, and the less clear the link between the money supply and aggregate spending will be. Indeed, there is an extreme case of ultrasensitivity of the demand for money to interest rates, called the *liquidity trap*, in which monetary policy has no effect on aggregate spending, because a change in the money supply has no effect on interest rates. (If the demand for money is ultrasensitive to interest rates, a tiny change in interest rates produces a very large change in the quantity of money demanded. Hence in this case, the demand for money is completely flat in the supply and demand diagrams of Chapter 5. Therefore, a change in the money supply that shifts the money supply curve to the right or left results in it intersecting the flat money demand curve at the same unchanged interest rate.)

The evidence on the interest sensitivity of the demand for money found by different researchers is remarkably consistent. Neither extreme case is supported by the data: The demand for money is sensitive to interest rates, but there is little evidence that a liquidity trap has ever existed.

Stability of Money Demand

If the money demand function, like Equation 4 or 6, is unstable and undergoes substantial unpredictable shifts, as Keynes thought, then velocity is unpredictable, and the quantity of money may not be tightly linked to aggregate spending, as it is in the modern quantity theory. The stability of the money demand function is also crucial to whether the Federal Reserve should target interest rates or the money supply (see Chapter 18 and 24). Thus it is important to look at the question of whether the money demand function is stable, because it has important implications for how monetary policy should be conducted.

By the early 1970s, evidence strongly supported the stability of the money demand function. However, after 1973, the rapid pace of financial innovation, which changed what items could be counted as money, led to substantial instability in estimated money demand functions. The recent instability of the money demand function calls into question whether our theories and empirical analyses are adequate. It also has important implications for the way monetary policy should be conducted, because it casts doubt on the usefulness of the money demand function as a tool to provide guidance to policymakers. In particular, because the money demand function has become unstable, velocity is now harder to predict, and as discussed in Chapter 21, setting rigid money supply targets in order to control aggregate spending in the economy may not be an effective way to conduct monetary policy.

Summary

1. Irving Fisher developed a transactions-based theory of the demand for money in which the demand for real balances is proportional to real income and is insensitive to interest-rate movements. An implication of his theory is that velocity, the rate of turnover of money, is constant. This generates the quantity theory of money, which implies that aggregate spending is determined solely by movements in the quantity of money.
2. The classical view that velocity can be effectively treated as a constant is not supported by the data. The nonconstancy of velocity became especially clear to the

economics profession after the sharp drop in velocity during the years of the Great Depression.

3. John Maynard Keynes suggested three motives for holding money: the transactions motive, the precautionary motive, and the speculative motive. His resulting liquidity preference theory views the transactions and precautionary components of money demand as proportional to income. However, the speculative component of money demand is viewed as sensitive to interest rates as well as to expectations about the future movements of interest rates. This theory, then, implies that velocity is unstable and cannot be treated as a constant.
4. Further developments in the Keynesian approach provided a better rationale for the three Keynesian motives for holding money. Interest rates were found to be important to the transactions and precautionary components of money demand as well as to the speculative component.
5. Milton Friedman's theory of money demand used a similar approach to that of Keynes. Treating money like any other asset, Friedman used the theory of asset demand to derive a demand for money that is a function of the expected returns on other assets relative to the expected return on money and permanent income. In contrast to Keynes, Friedman believed that the demand for money is stable and insensitive to interest-rate movements. His belief that velocity is predictable (though not constant) in turn leads to the quantity theory conclusion that money is the primary determinant of aggregate spending.
6. There are two main conclusions from the research on the demand for money: The demand for money is sensitive to interest rates, but there is little evidence that the liquidity trap has ever existed; and since 1973, money demand has been found to be unstable, with the most likely source of the instability being the rapid pace of financial innovation.

Key Terms

equation of exchange, p. 518
liquidity preference theory, p. 521
monetary theory, p. 517
quantity theory of money, p. 519
real money balances, p. 523
velocity of money, p. 518

Questions and Problems

Questions marked with an asterisk are answered at the end of the book in an appendix, "Answers to Selected Questions and Problems."

*1. The money supply *M* has been growing at 10% per year, and nominal GDP *PY* has been growing at 20% per year. The data are as follows (in billions of dollars):

	2001	2002	2003
M	100	110	121
PY	1,000	1,200	1,440

Calculate the velocity in each year. At what rate is velocity growing?

2. Calculate what happens to nominal GDP if velocity remains constant at 5 and the money supply increases from $200 billion to $300 billion.
*3. What happens to nominal GDP if the money supply grows by 20% but velocity declines by 30%?
4. If credit cards were made illegal by congressional legislation, what would happen to velocity? Explain your answer.
*5. If velocity and aggregate output are reasonably constant (as the classical economists believed), what happens to the price level when the money supply increases from $1 trillion to $4 trillion?
6. If velocity and aggregate output remain constant at 5 and 1,000, respectively, what happens to the price level if the money supply declines from $400 billion to $300 billion?
*7. Looking at Figure 1 in the chapter, when were the two largest falls in velocity? What do declines like this sug-

gest about how velocity moves with the business cycle? Given the data in Figure 1, is it reasonable to assume, as the classical economists did, that declines in aggregate spending are caused by declines in the quantity of money?

8. Using data from the *Economic Report of the President*, calculate velocity for the M2 definition of the money supply in the past five years. Does velocity appear to be constant?

*9. In Keynes's analysis of the speculative demand for money, what will happen to money demand if people suddenly decide that the normal level of the interest rate has declined? Why?

10. Why is Keynes's analysis of the speculative demand for money important to his view that velocity will undergo substantial fluctuations and thus cannot be treated as constant?

*11. If interest rates on bonds go to zero, what does the Baumol-Tobin analysis suggest Grant Smith's average holdings of money balances should be?

12. If brokerage fees go to zero, what does the Baumol-Tobin analysis suggest Grant Smith's average holdings of money should be?

*13. "In Tobin's analysis of the speculative demand for money, people will hold both money and bonds, even if bonds are expected to earn a positive return." Is this statement true, false, or uncertain? Explain your answer.

14. Both Keynes's and Friedman's theories of the demand for money suggest that as the relative expected return on money falls, demand for it will fall. Why does Friedman think that money demand is unaffected by changes in interest rates, but Keynes thought that it is affected?

*15. Why does Friedman's view of the demand for money suggest that velocity is predictable, whereas Keynes's view suggests the opposite?

Web Exercises

1. Refer to Figure 1. The formula for computing the velocity of money is GDP/M1. Go to www.research.stlouisfed.org/fred/data/gdp.html and look up the GDP. Next go to www.federalreserve.gov/Releases/h6/Current/ and find M1. Compute the most recent year's velocity of money and compare it to its level in 2002. Has it risen or fallen? Suggest reasons for its change since that time.

2. John Maynard Keynes is among the most well known economic theorists. Go to www-gap.dcsn.st-and.ac.uk/~history/Mathematicians/Keynes.html and write a one-page summary of his life and contributions.

appendix 1 to chapter 22

A Mathematical Treatment of the Baumol-Tobin and Tobin Mean-Variance Models

Baumol-Tobin Model of Transactions Demand for Money

The basic idea behind the Baumol-Tobin model was laid out in the chapter. Here we explore the mathematics that underlie the model. The assumptions of the model are as follows:

1. An individual receives income of T_0 at the beginning of every period.
2. An individual spends this income at a constant rate, so at the end of the period, all income T_0 has been spent.
3. There are only two assets—cash and bonds. Cash earns a nominal return of zero, and bonds earn an interest rate i.
4. Every time an individual buys or sells bonds to raise cash, a fixed brokerage fee of b is incurred.

Let us denote the amount of cash that the individual raises for each purchase or sale of bonds as C, and n = the number of times the individual conducts a transaction in bonds. As we saw in Figure 3 in the chapter, where $T_0 = 1{,}000$, $C = 500$, and $n = 2$:

$$n = \frac{T_0}{C}$$

Because the brokerage cost of each bond transaction is b, the total brokerage costs for a period are:

$$nb = \frac{bT_0}{C}$$

Not only are there brokerage costs, but there is also an opportunity cost to holding cash rather than bonds. This opportunity cost is the bond interest rate i times average cash balances held during the period, which, from the discussion in the chapter, we know is equal to $C/2$. The opportunity cost is then:

$$\frac{iC}{2}$$

Combining these two costs, we have the total costs for an individual equal to:

$$COSTS = \frac{bT_0}{C} + \frac{iC}{2}$$

The individual wants to minimize costs by choosing the appropriate level of C. This is accomplished by taking the derivative of costs with respect to C and setting it to zero.[1] That is:

$$\frac{d\,COSTS}{dC} = \frac{-\,bT_0}{C^2} + \frac{i}{2} = 0$$

Solving for C yields the optimal level of C:

$$C = \sqrt{\frac{2bT_0}{i}}$$

Because money demand M^d is the average desired holding of cash balances $C/2$,

$$M^d = \frac{1}{2}\sqrt{\frac{2bT_0}{i}} = \sqrt{\frac{bT_0}{2i}} \tag{1}$$

This is the famous *square root rule*.[2] It has these implications for the demand for money:

1. The transactions demand for money is negatively related to the interest rate i.
2. The transactions demand for money is positively related to income, but there are economies of scale in money holdings—that is, the demand for money rises less than proportionally with income. For example, if T_0 quadruples in Equation 1, the demand for money only doubles.
3. A lowering of the brokerage costs due to technological improvements would decrease the demand for money.
4. There is no money illusion in the demand for money. If the price level doubles, T_0 and b will double. Equation 1 then indicates that M will double as well. Thus the demand for real money balances remains unchanged, which makes sense because neither the interest rate nor real income has changed.

---

[1]To minimize costs, the second derivative must be greater than zero. We find that it is, because:

$$\frac{d^2COSTS}{dC^2} = \frac{-2}{C^3}(-bT_0) = \frac{2bT_0}{C^3} > 0$$

[2]An alternative way to get Equation 1 is to have the individual maximize profits, which equal the interest on bonds minus the brokerage costs. The average holding of bonds over a period is just:

$$\frac{T_0}{2} - \frac{C}{2}$$

Thus profits are:

$$PROFITS = -\frac{i}{2}(T_0 - C) - \frac{bT_0}{C}$$

Then:

$$\frac{d\,PROFITS}{dC} = \frac{-i}{2} + \frac{bT_0}{C^2} = 0$$

This equation yields the same square root rule as Equation 1.

Tobin Mean-Variance Model

Tobin's mean-variance analysis of money demand is just an application of the basic ideas in the theory of portfolio choice. Tobin assumes that the utility that people derive from their assets is positively related to the expected return on their portfolio of assets and is negatively related to the riskiness of this portfolio as represented by the variance (or standard deviation) of its returns. This framework implies that an individual has indifference curves that can be drawn as in Figure 1. Notice that these indifference curves slope upward because an individual is willing to accept more risk if offered a higher expected return. In addition, as we go to higher indifference curves, utility is higher, because for the same level of risk, the expected return is higher.

Tobin looks at the choice of holding money, which earns a certain zero return, or bonds, whose return can be stated as:

$$R_B = i + g$$

where i = interest rate on the bond
g = capital gain

Tobin also assumes that the expected capital gain is zero[3] and its variance is σ_g^2. That is,

$$E(g) = 0 \quad \text{and so} \quad E(R_B) = i + 0 = i$$
$$Var(g) = E[g - E(g)]^2 = E(g^2) = \sigma_g^2$$

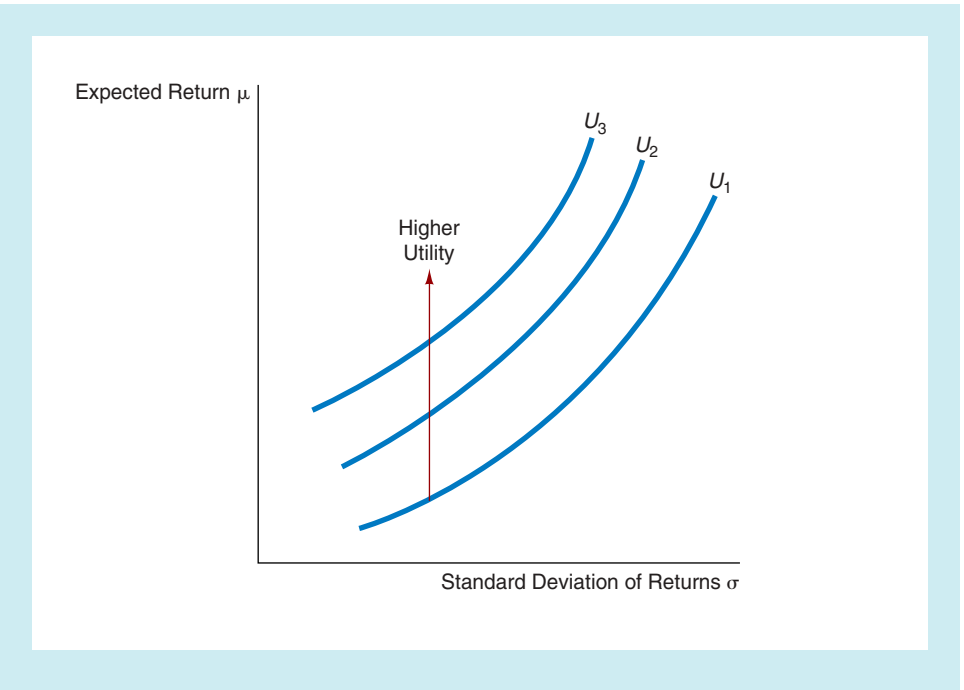


FIGURE 1 Indifference Curves in a Mean-Variace Model

The indifference curves are upward-sloping, and higher indifference curves indicate that utility is higher. In other words, $U_3 > U_2 > U_1$.

[3] This assumption is not critical to the results. If $E(g) \neq 0$, it can be added to the interest term i, and the analysis proceeds as indicated.

where E = expectation of the variable inside the parentheses
Var = variance of the variable inside the parentheses

If A is the fraction of the portfolio put into bonds ($0 \leq A \leq 1$) and $1 - A$ is the fraction of the portfolio held as money, the return R on the portfolio can be written as:

$$R = AR_B + (1 - A)(0) = AR_B = A(i + g)$$

Then the mean and variance of the return on the portfolio, denoted respectively as μ and σ^2, can be calculated as follows:

$$\mu = E(R) = E(AR_B) = AE(R_B) = Ai$$

$$\sigma^2 = E(R - \mu)^2 = E[A(i + g) - Ai]^2 = E(Ag)^2 = A^2E(g^2) = A^2\sigma_g^2$$

Taking the square root of both sides of the equation directly above and solving for A yields:

$$A = \frac{1}{\sigma_g}\sigma \tag{2}$$

Substituting for A in the equation $\mu = Ai$ using the preceding equation gives us:

$$\mu = \frac{i}{\sigma_g}\sigma \tag{3}$$

Equation 3 is known as the *opportunity locus* because it tells us the combinations of μ and σ that are feasible for the individual. This equation is written in a form in which the μ variable corresponds to the Y axis and the σ variable to the X axis. The opportunity locus is a straight line going through the origin with a slope of i/σ_g. It is drawn in the top half of Figure 2 along with the indifference curves from Figure 1.

The highest indifference curve is reached at point B, the tangency of the indifference curve and the opportunity locus. This point determines the optimal level of risk σ^* in the figure. As Equation 2 indicates, the optimal level of A, A^*, is:

$$A^* = \frac{\sigma^*}{\sigma_g}$$

This equation is solved in the bottom half of Figure 2. Equation 2 for A is a straight line through the origin with a slope of $1/\sigma_g$. Given σ^*, the value of A read off this line is the optimal value A^*. Notice that the bottom part of the figure is drawn so that as we move down, A is increasing.

Now let's ask ourselves what happens when the interest rate increases from i_1 to i_2. This situation is shown in Figure 3. Because σ_g is unchanged, the Equation 2 line in the bottom half of the figure does not change. However, the slope of the opportunity locus does increase as i increases. Thus the opportunity locus rotates up and we move to point C at the tangency of the new opportunity locus and the indifference curve. As you can see, the optimal level of risk increases from σ_1^* and σ_2^* the optimal fraction of the portfolio in bonds rises from A_1^* to A_2^*. The result is that as the interest

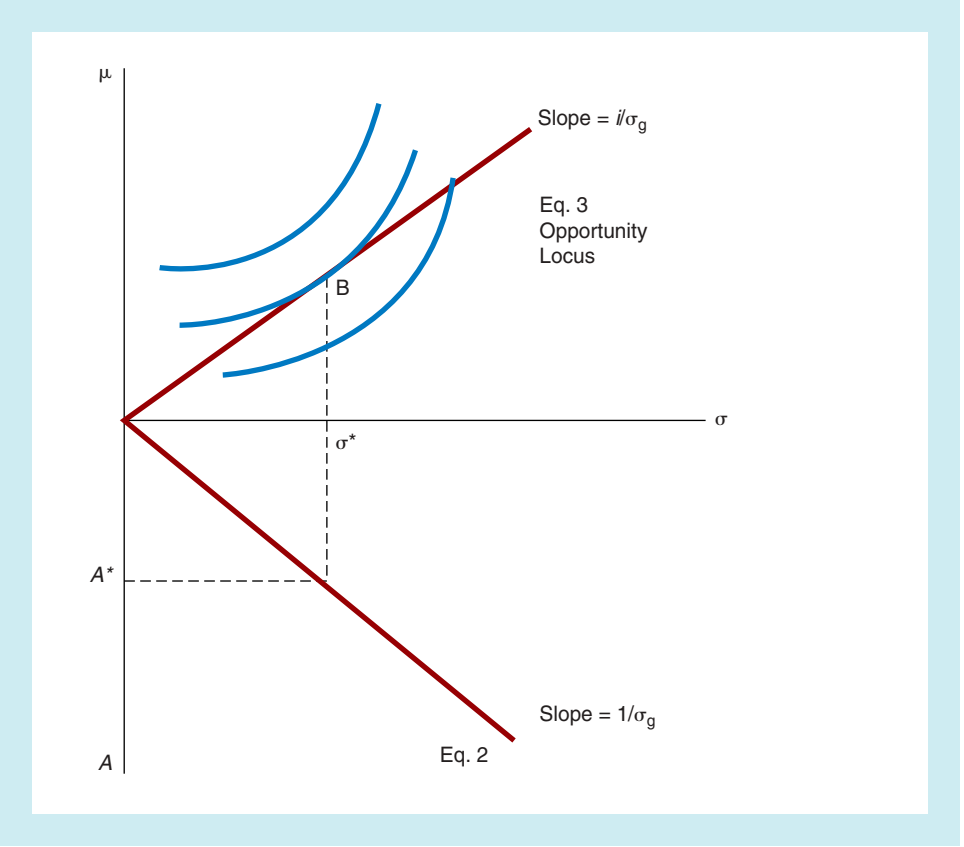


FIGURE 2 Optimal Choice of the Fraction of the Portfolio in Bonds
The highest indifference curve is reached at a point B, the tangency of the indifference curve with the opportunity locus. This point determines the optimal risk σ^*, and using Equation 2 in the bottom half of the figure, we solve for the optimal fraction of the portfolio in bonds A^*.

rate on bonds rises, the demand for money falls; that is, $1 - A$, the fraction of the portfolio held as money, declines.[4]

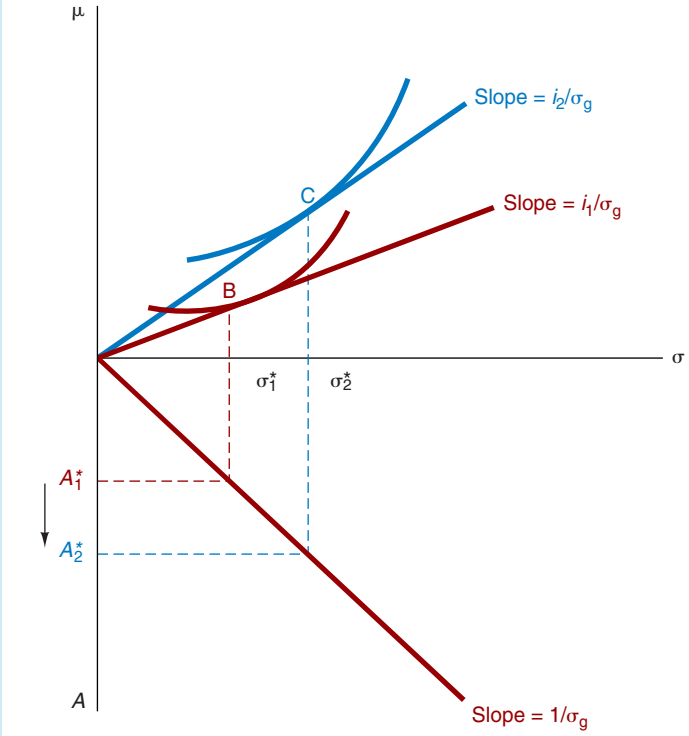
Tobin's model then yields the same result as Keynes's analysis of the speculative demand for money: It is negatively related to the level of interest rates. This model, however, makes two important points that Keynes's model does not:

1. Individuals diversify their portfolios and hold money *and* bonds at the same time.
2. Even if the expected return on bonds is greater than the expected return on money, individuals will still hold money as a store of wealth because its return is more certain.

[4]The indifference curves have been drawn so that the usual result is obtained that as i goes up, A^* goes up as well. However, there is a subtle issue of income versus substitution effects. If, as people get wealthier, they are willing to bear less risk, and if this income effect is larger than the substitution effect, then it is possible to get the opposite result that as i increases, A^* declines. This set of conditions is unlikely, which is why the figure is drawn so that the usual result is obtained. For a discussion of income versus substitution effects, see David Laidler, *The Demand for Money: Theories and Evidence*, 4th ed. (New York: HarperCollins, 1993).

FIGURE 3 Optimal Choice of the Fraction of the Portfolio in Bonds as the Interest Rate Rises

The interest rate on bonds rises from i_1 to i_2, rotating the opportunity locus upward. The highest indifference curve is now at point C, where it is tangent to the new opportunity locus. The optimal level of risk rises from σ_1^* to σ_2^*, and then Equation 2, in the bottom haf of the figure, shows that the optimal fraction of the portfolio in bonds rises from A_1^* to A_2^*.



appendix 2 to chapter 22

Empirical Evidence on the Demand for Money

Here we examine the empirical evidence on the two primary issues that distinguish the different theories of money demand and affect their conclusions about whether the quantity of money is the primary determinant of aggregate spending: Is the demand for money sensitive to changes in interest rates, and is the demand for money function stable over time?

Interest Rates and Money Demand

James Tobin conducted one of the earliest studies on the link between interest rates and money demand using U.S. data.[1] Tobin separated out transactions balances from other money balances, which he called "idle balances," assuming that transactions balances were proportional to income only, and idle balances were related to interest rates only. He then looked at whether his measure of idle balances was inversely related to interest rates in the period 1922–1941 by plotting the average level of idle balances each year against the average interest rate on commercial paper that year. When he found a clear-cut inverse relationship between interest rates and idle balances, Tobin concluded that the demand for money is sensitive to interest rates.[2]

Additional empirical evidence on the demand for money strongly confirms Tobin's finding.[3] Does this sensitivity ever become so high that we approach the case of the liquidity trap in which monetary policy is ineffective? The answer is almost certainly no. Keynes suggested in *The General Theory* that a liquidity trap might occur when interest rates are extremely low. (However, he did state that he had never yet seen an occurrence of a liquidity trap.)

Typical of the evidence demonstrating that the liquidity trap has never occurred is that of David Laidler, Karl Brunner, and Allan Meltzer, who looked at whether the interest sensitivity of money demand increased in periods when interest rates were

---

[1]James Tobin, "Liquidity Preference and Monetary Policy," *Review of Economics and Statistics* 29 (1947): 124–131.

[2]A problem with Tobin's procedure is that idle balances are not really distinguishable from transactions balances. As the Baumol-Tobin model of transactions demand for money makes clear, transactions balances will be related to both income and interest rates, just like idle balances.

[3]See David E. W. Laidler, *The Demand for Money: Theories and Evidence*, 4th ed. (New York: HarperCollins, 1993). Only one major study has found that the demand for money is insensitive to interest rates: Milton Friedman, "The Demand for Money: Some Theoretical and Empirical Results," *Journal of Political Economy* 67 (1959): 327–351. He concluded that the demand for money is not sensitive to interest-rate movements, but as later work by David Laidler (using the same data as Friedman) demonstrated, Friedman used a faulty statistical procedure that biased his results: David E. W. Laidler, "The Rate of Interest and the Demand for Money: Some Empirical Evidence," *Journal of Political Economy* 74 (1966): 545–555. When Laidler employed the correct statistical procedure, he found the usual result that the demand for money is sensitive to interest rates. In later work, Friedman has also concluded that the demand for money is sensitive to interest rates.

very low.[4] Laidler and Meltzer looked at this question by seeing whether the interest sensitivity of money demand differed across periods, especially in periods such as the 1930s when interest rates were particularly low.[5] They found that there was no tendency for interest sensitivity to increase as interest rates fell—in fact, interest sensitivity did not change from period to period. Brunner and Meltzer explored this question by recognizing that higher interest sensitivity in the 1930s as a result of a liquidity trap implies that a money demand function estimated for this period should not predict well in more normal periods. What Brunner and Meltzer found was that a money demand function, estimated mostly with data from the 1930s, accurately predicted the demand for money in the 1950s. This result provided little evidence in favor of the existence of a liquidity trap during the Great Depression period.

The evidence on the interest sensitivity of the demand for money found by different researchers is remarkably consistent. Neither extreme case is supported by the data: The demand for money is sensitive to interest rates, but there is little evidence that a liquidity trap has ever existed.

Stability of Money Demand

If the money demand function, like Equation 4 or 6 in Chapter 22, is unstable and undergoes substantial unpredictable shifts, as Keynes thought, then velocity is unpredictable, and the quantity of money may not be tightly linked to aggregate spending, as it is in the modern quantity theory. The stability of the money demand function is also crucial to whether the Federal Reserve should target interest rates or the money supply (see Chapter 24). Thus it is important to look at the question of whether the money demand function is stable, because it has important implications for how monetary policy should be conducted.

As our discussion of the Brunner and Meltzer article indicates, evidence on the stability of the demand for money function is related to the evidence on the existence of a liquidity trap. Brunner and Meltzer's finding that a money demand function estimated using data mostly from the 1930s predicted the demand for money well in the postwar period not only suggests that a liquidity trap did not exist in the 1930s, but also indicates that the money demand function has been stable over long periods of time. The evidence that the interest sensitivity of the demand for money did not change from period to period also suggests that the money demand function is stable, since a changing interest sensitivity would mean that the demand for money function estimated in one period would not be able to predict well in another period.

By the early 1970s, the evidence using quarterly data from the postwar period strongly supported the stability of the money demand function when M1 was used as the definition of the money supply. For example, a well-known study by Stephen Goldfeld published in 1973 found not only that the interest sensitivity of M1 money demand did not undergo changes in the postwar period, but also that the M1 money demand function predicted extremely well throughout the postwar period.[6] As a

---

[4]David E. W. Laidler, "Some Evidence on the Demand for Money," *Journal of Political Economy* 74 (1966): 55–68; Allan H. Meltzer, "The Demand for Money: The Evidence from the Time Series," *Journal of Political Economy* 71 (1963): 219–246; Karl Brunner and Allan H. Meltzer, "Predicting Velocity: Implications for Theory and Policy," *Journal of Finance* 18 (1963): 319–354.

[5]Interest sensitivity is measured by the interest elasticity of money demand, which is defined as the percentage change in the demand for money divided by the percentage change in the interest rate.

[6]Stephen M. Goldfeld, "The Demand for Money Revisited," *Brookings Papers on Economic Activity* 3 (1973): 577–638.

result of this evidence, the M1 money demand function became the conventional money demand function used by economists.

The Case of the Missing Money. The stability of the demand for money, then, was a well-established fact when, starting in 1974, the conventional M1 money demand function began to severely overpredict the demand for money. Stephen Goldfeld labeled this phenomenon of instability in the demand for money function "the case of the missing money."[7] It presented a serious challenge to the usefulness of the money demand function as a tool for understanding how monetary policy affects aggregate economic activity. In addition, it had important implications for how monetary policy should be conducted. As a result, the instability of the M1 money demand function stimulated an intense search for a solution to the mystery of the missing money so that a stable money demand function could be resurrected.

The search for a stable money demand function took two directions. The first direction focused on whether an incorrect definition of money could be the reason why the demand for money function had become so unstable. Inflation, high nominal interest rates, and advances in computer technology caused the payments mechanism and cash management techniques to undergo rapid changes after 1974. In addition, many new financial instruments emerged and have grown in importance. This has led some researchers to suspect that the rapid pace of financial innovation since 1974 has meant that the conventional definitions of the money supply no longer apply. They searched for a stable money demand function by actually looking directly for the missing money; that is, they looked for financial instruments that have been incorrectly left out of the definition of money used in the money demand function.

Overnight repurchase agreements (RPs) are one example. These are one-day loans with little default risk because they are structured to provide Treasury bills as collateral. (The appendix to Chapter 2 gives a more detailed discussion of the structure of this type of loan.) Corporations with demand deposit accounts at commercial banks frequently lend out substantial amounts of their account balances overnight with these RPs, lowering the measures of the money supply. However, the amounts lent out are very close substitutes for money, since the corporation can quickly make a decision to decrease these loans if it needs more money in its demand deposit account to pay its bills. Gillian Garcia and Simon Pak, for example, found that including overnight RPs in measures of the money supply substantially reduced the degree to which money demand functions overpredicted the money supply.[8] More recent evidence using later data has cast some doubt on whether including overnight RPs and other highly liquid assets in measures of the money supply produces money demand functions that are stable.[9]

The second direction of search for a stable money demand function was to look for new variables to include in the money demand function that will make it stable.

---

[7]Stephen M. Goldfeld, "The Case of the Missing Money," *Brookings Papers on Economic Activity* 3 (1976): 683–730.

[8]Gillian Garcia and Simon Pak, "Some Clues in the Case of the Missing Money," *American Economic Review* 69 (1979): 330–334.

[9]See the survey in John P. Judd and John L. Scadding, "The Search for a Stable Money Demand Function," *Journal of Economic Literature* 20 (1982): 993–1023.

Michael Hamburger, for example, found that including the average dividend–price ratio on common stocks (average dividends divided by average price) as a measure of their interest rate resulted in a money demand function that is stable.[10] Other researchers, such as Heller and Khan, added the entire term structure of interest rates to their money demand function and found that this produces a stable money demand function.[11]

These attempts to produce a stable money demand function have been criticized on the grounds that these additional variables do not accurately measure the opportunity cost of holding money, and so the theoretical justification for including them in the money demand function is weak.[12] Also, later research questions whether these alterations to the money demand function will lead to continuing stability in the future.[13]

Velocity Slowdown in the 1980s. The woes of conventional money demand functions increased in the 1980s. We have seen that they overpredicted money demand in the middle and late 1970s; that is, they underpredicted velocity (PY/M), which rose faster than expected. The tables turned beginning in 1982; as can be seen in Figure 1 in Chapter 22, economists now faced a surprising slowdown in M1 velocity, which conventional money demand functions also could not predict. Although researchers have tried to explain this velocity slowdown, they have not been entirely successful.[14]

M2 to the Rescue? As we saw in Figure 1, M2 velocity remained far more stable than M1 velocity in the 1980s. The relative stability of M2 velocity suggests that money demand functions in which the money supply is defined as M2 might perform substantially better than those in which the money supply is defined as M1. Researchers at the Federal Reserve found that M2 money demand functions performed well in the 1980s, with M2 velocity moving quite closely with the opportunity cost of holding M2 (market interest rates minus an average of the interest paid on deposits and financial instruments that make up M2).[15] However, in the early 1990s, M2 growth underwent a dramatic slowdown, which some researchers believe cannot be explained by

---

[10]Michael Hamburger, "Behavior of the Money Stock: Is There a Puzzle?" *Journal of Monetary Economics* 3 (1977): 265–288. The stability of his money demand function also depends on his assumption that the income elasticity of the demand for money is unity. This assumption has been strongly criticized by many critics, including R. W. Hafer and Scott E. Hein, "Evidence on the Temporal Stability of the Demand for Money Relationship in the United States," Federal Reserve Bank of St. Louis *Review* (1979): 3–14, who find that this assumption is strongly rejected by the data.

[11]H. Heller and Moshin S. Khan, "The Demand for Money and the Term Structure of Interest Rates," *Journal of Political Economy* 87 (1979): 109–129.

[12]Frederic S. Mishkin, "Discussion of Asset Substitutability and the Impact of Federal Deficits," in *The Economic Consequences of Government Deficits,* ed. Laurence H. Meyer (Boston: Kluwer-Nijhoff, 1983), pp. 117–120; Frederic S. Mishkin, "Discussion of Recent Velocity Behavior: The Demand for Money and Monetary Policy," in *Monetary Targeting and Velocity* (San Francisco: Federal Reserve Bank of San Francisco, 1983), pp. 129–132.

[13]This research is discussed in Judd and Scadding (note 9).

[14]See, for example, Robert H. Rasche, "M1 Velocity and Money-Demand Functions: Do Stable Relationships Exist?" *Empirical Studies of Velocity, Real Exchange Rates, Unemployment, and Productivity, Carnegie-Rochester Conference Series on Public Policy* 17 (Autumn 1987), pp. 9–88.

[15]See David H. Small and Richard D. Porter, "Understanding the Behavior of M2 and V2," *Federal Reserve Bulletin* 75 (1989): 244–254.

traditional money demand functions.[16] In the late 1990s, M2 velocity seemed to settle down, suggesting a more normal relationship between M2 demand and macroeconomic variables. However, doubts continue to arise about the stability of money demand.[17]

Conclusion. The main conclusion from the research on the money demand function seems to be that the most likely cause of its instability is the rapid pace of financial innovation occurring after 1973, which has changed what items can be counted as money. The evidence is still somewhat tentative, however, and a truly stable and satisfactory money demand function has not yet been found. And so the search for a stable money demand function goes on.

The recent instability of the money demand function calls into question whether our theories and empirical analyses are adequate.[18] It also has important implications for the way monetary policy should be conducted because it casts doubt on the usefulness of the money demand function as a tool to provide guidance to policymakers. In particular, because the money demand function has become unstable, velocity is now harder to predict, and as discussed in Chapter 21, setting rigid money supply targets in order to control aggregate spending in the economy may not be an effective way to conduct monetary policy.

---

[16]See, for example, Bryon Higgins, "Policy Implications of Recent M2 Behavior," Federal Reserve Bank of Kansas City *Economic Review* (Third Quarter 1992): 21–36. For a contrary view, see Robert L. Hetzel, "How Useful Is M2 Today," Federal Reserve Bank of Richmond *Economic Review* (September–October 1992): 12–26.

[17]For example, see Kelly Ragan and Bharat Trehan, "Is It Time to Look at M2 Again?" Federal Reserve Bank of San Francisco *Economic Letter* #98-07 (March 6, 1998).

[18]Thomas F. Cooley and Stephen F. Le Roy, "Identification and Estimation of Money Demand," *American Economic Review* 71 (1981): 825–844, is especially critical of the empirical research on the demand for money.

Chapter 23

The Keynesian Framework and the *ISLM* Model

PREVIEW

In the media, you often see forecasts of GDP and interest rates by economists and government agencies. At times, these forecasts seem to come from a crystal ball, but economists actually make their predictions using a variety of economic models. One model widely used by economic forecasters is the *ISLM* model, which was developed by Sir John Hicks in 1937 and is based on the analysis in John Maynard Keynes's influential book *The General Theory of Employment, Interest, and Money*, published in 1936.[1] The *ISLM* model explains how interest rates and total output produced in the economy (aggregate output or, equivalently, aggregate income) are determined, given a fixed price level.

The *ISLM* model is valuable not only because it can be used in economic forecasting, but also because it provides a deeper understanding of how government policy can affect aggregate economic activity. In Chapter 24 we use it to evaluate the effects of monetary and fiscal policy on the economy and to learn some lessons about how monetary policy might best be conducted.

In this chapter, we begin by developing the simplest framework for determining aggregate output, in which all economic actors (consumers, firms, and others) except the government play a role. Government fiscal policy (spending and taxes) is then added to the framework to see how it can affect the determination of aggregate output. Finally, we achieve a complete picture of the *ISLM* model by adding monetary policy variables: the money supply and the interest rate.

Determination of Aggregate Output

http://research.stlouisfed.org/fred/index.html
Information about the macroeconomic variables discussed in this chapter.

Keynes was especially interested in understanding movements of aggregate output because he wanted to explain why the Great Depression had occurred and how government policy could be used to increase employment in a similar economic situation. Keynes's analysis started with the recognition that the total quantity demanded of an economy's output was the sum of four types of spending: (1) **consumer expenditure** (C), the total demand for consumer goods and services (hamburgers, stereos, rock concerts, visits to the doctor, and so on); (2) **planned investment spending** (I),

[1] John Hicks, "Mr. Keynes and the Classics: A Suggested Interpretation," *Econometrica* (1937): 147–159.

the total planned spending by businesses on new physical capital (machines, computers, factories, raw materials, and the like) plus planned spending on new homes; (3) **government spending** (G), the spending by all levels of government on goods and services (aircraft carriers, government workers, red tape, and so forth); and (4) **net exports** (NX), the net foreign spending on domestic goods and services, equal to exports minus imports.[2] The total quantity demanded of an economy's output, called **aggregate demand** (Y^{ad}), can be written as:

$$Y^{ad} = C + I + G + NX \tag{1}$$

Using the common-sense concept from supply and demand analysis, Keynes recognized that equilibrium would occur in the economy when total quantity of output supplied (aggregate output produced) Y equals quantity of output demanded Y^{ad}, that is, when:

$$Y = Y^{ad} \tag{2}$$

When this equilibrium condition is satisfied, producers are able to sell all of their output and have no reason to change their production. Keynes's analysis explains two things: (1) why aggregate output is at a certain level (which involves understanding what factors affect each component of aggregate demand) and (2) how the sum of these components can add up to an output smaller than the economy is capable of producing, resulting in less than full employment of resources.

Keynes was especially concerned with explaining the low level of output and employment during the Great Depression. Because inflation was not a serious problem during this period, he assumed that output could change without causing a change in prices. ***Keynes's analysis assumes that the price level is fixed***; that is, dollar amounts for variables such as consumer expenditure, investment, and aggregate output do not have to be adjusted for changes in the price level to tell us how much the real quantities of these variables change. Because the price level is assumed to be fixed, when we talk in this chapter about changes in nominal quantities, we are talking about changes in real quantities as well.

Our discussion of Keynes's analysis begins with a simple framework of aggregate output determination in which the role of government, net exports, and the possible effects of money and interest rates are ignored. Because we are assuming that government spending and net exports are zero ($G = 0$ and $NX = 0$), we need only examine consumer expenditure and investment spending to explain how aggregate output is determined. This simple framework is unrealistic, because both government and monetary policy are left out of the picture, and because it makes other simplifying assumptions, such as a fixed price level. Still, the model is worth studying, because its simplified view helps us understand the key factors that explain how the economy works. It also clearly illustrates the Keynesian idea that the economy can come to rest at a level of aggregate output below the full employment level. Once you understand this simple framework, we can proceed to more complex and more realistic models.

[2] Imports are subtracted from exports in arriving at the net exports component of the total quantity demanded of an economy's output because imports are already counted in *C*, *I*, and *G* but do not add to the demand for the economy's output.

Consumer Expenditure and the Consumption Function

Ask yourself what determines how much you spend on consumer goods and services. Your likely response is that your income is the most important factor, because if your income rises, you will be willing to spend more. Keynes reasoned similarly that consumer expenditure is related to **disposable income**, the total income available for spending, equal to aggregate income (which is equivalent to aggregate output) minus taxes ($Y - T$). He called this relationship between disposable income Y_D and consumer expenditure C the **consumption function** and expressed it as:

$$C = a + (mpc \times Y_D) \tag{3}$$

The term *mpc*, the **marginal propensity to consume**, is the slope of the consumption function line ($\Delta C/\Delta Y_D$) and reflects the change in consumer expenditure that results from an additional dollar of disposable income. Keynes assumed that *mpc* was a constant between the values of 0 and 1. If, for example, a $1.00 increase in disposable income leads to an increase in consumer expenditure of $0.50, then $mpc = 0.5$.

The term a stands for **autonomous consumer expenditure**, the amount of consumer expenditure that is independent of disposable income. It tells us how much consumers will spend when disposable income is 0 (they still must have food, clothing, and shelter). If a is $200 billion when disposable income is 0, consumer expenditure will equal $200 billion.[3]

A numerical example of a consumption function using the values of $mpc = 0.5$ and $a = 200$ will clarify the preceding concept. The $200 billion of consumer expenditure at a disposable income of 0 is listed in the first row of Table 1 and is plotted as point E in Figure 1. (Remember that throughout this chapter, dollar amounts for all variables in the figures correspond to real quantities, because Keynes assumed that the price level is fixed.) Because $mpc = 0.5$, when disposable income increases by $400 billion, the change in consumer expenditure—ΔC in column 3 of Table 1—is $200 billion ($0.5 \times$ $400 billion). Thus when disposable income is $400 billion, consumer expenditure is $400 billion (initial value of $200 billion when income is 0 plus the $200 billion change in consumer expenditure). This combination of consumer expenditure and disposable income is listed in the second row of Table 1 and is plotted as point F in Figure 1. Similarly, at point G, where disposable income has increased by another $400 billion to $800 billion, consumer expenditure will rise by another $200 billion to $600 billion. By the same reasoning, at point H, at which disposable income is $1,200 billion, consumer expenditure will be $800 billion. The line connecting these points in Figure 1 graphs the consumption function.

Study Guide

The consumption function is an intuitive concept that you can readily understand if you think about how your own spending behavior changes as you receive more disposable income. One way to make yourself more comfortable with this concept is to estimate your marginal propensity to consume (for example, it might be 0.8) and your level of consumer expenditure when your disposable income is 0 (it might be $2,000) and then construct a consumption function similar to that in Table 1.

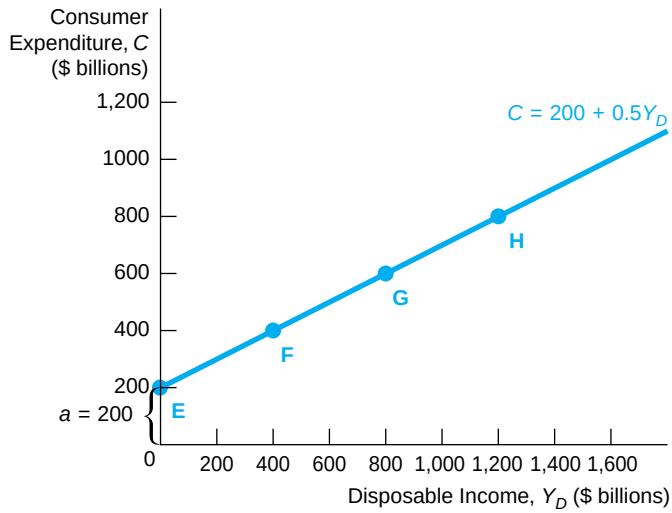
[3]Consumer expenditure can exceed income if people have accumulated savings to tide them over bad times. An alternative is to have parents who will give you money for food (or to pay for school) when you have no income. The situation in which consumer expenditure is greater than disposable income is called *dissaving*.

Table 1 Consumption Function: Schedule of Consumer Expenditure *C* When *mpc* = 0.5 and *a* = 200 ($ billions)

Point in Figure 1	Disposable income Y_D (1)	Change in Disposable Income ΔY_D (2)	Change in Consumer Expenditure ΔC $(0.5 \times \Delta Y_D)$ (3)	Consumer Expenditure C (4)
E	0	—	—	200 (= a)
F	400	400	200	400
G	800	400	200	600
H	1,200	400	200	800

FIGURE 1 Consumption Function

The consumption function plotted here is from Table 1; $mpc = 0.5$ and $a = 200$.



http://nova.umuc.edu/~black/consf1000.html

A dynamic interactive demonstration of how changing the inputs to the consumption function alters the results.

Investment Spending

It is important to understand that there are two types of investment. The first type, **fixed investment**, is the spending by firms on equipment (machines, computers, airplanes) and structures (factories, office buildings, shopping centers) and planned spending on residential housing. The second type, **inventory investment**, is spending by firms on additional holdings of raw materials, parts, and finished goods, calculated as the change in holdings of these items in a given time period—say a year. (Box 1 explains how economists' use of the word *investment* differs from everyday use of the term.)

Suppose that Compaq, a company that produces personal computers, has 100,000 computers sitting in its warehouses on December 31, 2003, ready to be shipped to

Box 1

Meaning of the Word *Investment*

Economists use the word *investment* somewhat differently than other people do. When people say that they are making an investment, they are normally referring to the purchase of common stocks or bonds, purchases that do not necessarily involve newly produced goods and services. But when economists speak of investment spending, they are referring to the purchase of *new* physical assets such as new machines or new houses—purchases that add to aggregate demand.

dealers. If each computer has a wholesale price of $1,000, Compaq has an inventory worth $100 million. If by December 31, 2004, its inventory of personal computers has risen to $150 million, its inventory investment in 2004 is $50 million, the *change* in the level of its inventory over the course of the year ($150 million minus $100 million). Now suppose that there is a drop in the level of inventories; inventory investment will then be negative.

Compaq may also have additional inventory investment if the level of raw materials and parts that it is holding to produce these computers increases over the course of the year. If on December 31, 2003, it holds $20 million of computer chips used to produce its computers and on December 31, 2004, it holds $30 million, it has an additional $10 million of inventory investment in 2001.

An important feature of inventory investment is that—in contrast to fixed investment, which is always planned—some inventory investment can be unplanned. Suppose that the reason Compaq finds itself with an additional $50 million of computers on December 31, 2004 is that $50 million less of its computers were sold in 2004 than expected. This $50 million of inventory investment in 2004 was unplanned. In this situation, Compaq is producing more computers than it can sell and will cut production.

Planned investment spending, a component of aggregate demand Y^{ad}, is equal to planned fixed investment plus the amount of inventory investment *planned* by firms. Keynes mentioned two factors that influence planned investment spending: interest rates and businesses' expectations about the future. How these factors affect investment spending is discussed later in this chapter. For now, planned investment spending will be treated as a known value. At this stage, we want to explain how aggregate output is determined for a given level of planned investment spending; we can then examine how interest rates and business expectations influence aggregate output by affecting planned investment spending.

Equilibrium and the Keynesian Cross Diagram

We have now assembled the building blocks (consumer expenditure and planned investment spending) that will enable us to see how aggregate output is determined when we ignore the government. Although unrealistic, this stripped-down analysis clarifies the basic principles of output determination. In the next section, government enters the picture and makes our model more realistic.

The diagram in Figure 2, known as the *Keynesian cross diagram*, shows how aggregate output is determined. The vertical axis measures aggregate demand, and the horizontal axis measures the level of aggregate output. The 45° line shows all the points

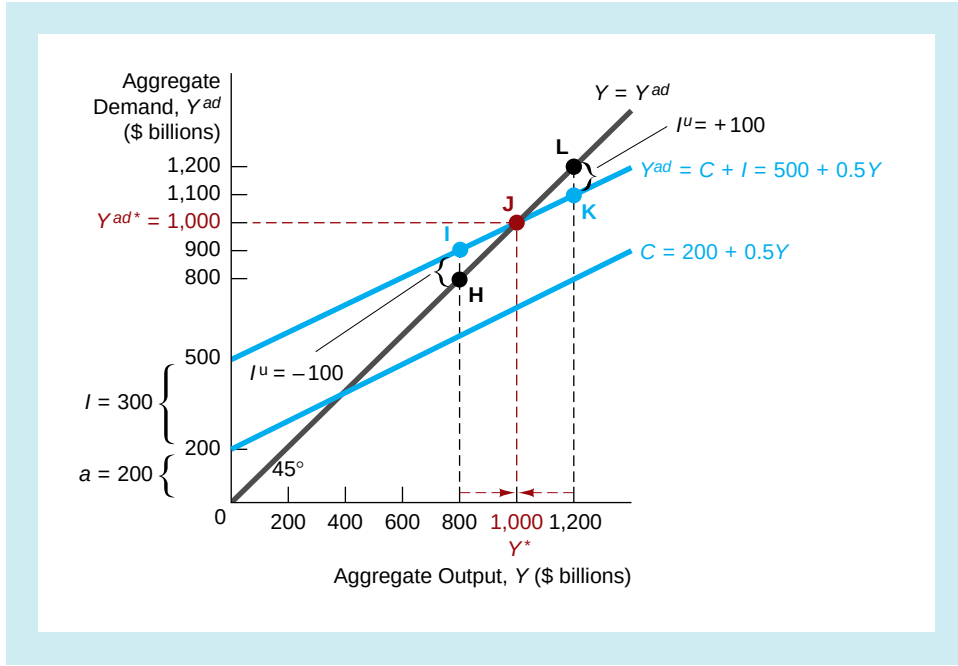


FIGURE 2 Keynesian Cross Diagram

When $I = 300$ and $C = 200 + 0.5Y$, equilibrium output occurs at $Y^* = 1{,}000$, where the aggregate demand function $Y^{ad} = C + I$ intersects with the 45° line $Y = Y^{ad}$.

at which aggregate output Y equals aggregate demand Y^{ad}; that is, it shows all the points at which the equilibrium condition $Y = Y^{ad}$ is satisfied. Since government spending and net exports are zero ($G = 0$ and $NX = 0$), aggregate demand is:

$$Y^{ad} = C + I$$

Because there is no government sector to collect taxes, there are none in our simplified economy; disposable income Y_D then equals aggregate output Y (remember that aggregate income and aggregate output are equivalent; see the appendix to Chapter 1). Thus the consumption function with $a = 200$ and $mpc = 0.5$ plotted in Figure 1 can be written as $C = 200 + 0.5Y$ and is plotted in Figure 2. Given that planned investment spending is $300 billion, aggregate demand can then be expressed as:

$$Y^{ad} = C + I = 200 + 0.5Y + 300 = 500 + 0.5Y$$

This equation, plotted in Figure 2, represents the quantity of aggregate demand at any given level of aggregate output and is called the **aggregate demand function**.

The aggregate demand function $Y^{ad} = C + I$ is the vertical sum of the consumption function line ($C = 200 + 0.5Y$) and planned investment spending ($I = 300$). The point at which the aggregate demand function crosses the 45° line $Y = Y^{ad}$ indicates the equilibrium level of aggregate demand and aggregate output. In Figure 2, equilibrium occurs at point J, with both aggregate output Y^* and aggregate demand Y^{ad*} at $1,000 billion.

As you learned in Chapter 5, the concept of equilibrium is useful only if there is a tendency for the economy to settle there. To see whether the economy heads toward the equilibrium output level of $1,000 billion, let's first look at what happens if the

amount of output produced in the economy is $1,200 billion and is therefore above the equilibrium level. At this level of output, aggregate demand is $1,100 billion (point K), $100 billion less than the $1,200 billion of output (point L on the 45° line). Since output exceeds aggregate demand by $100 billion, firms are saddled with $100 billion of unsold inventory. To keep from accumulating unsold goods, firms will cut production. As long as it is above the equilibrium level, output will exceed aggregate demand and firms will cut production, sending aggregate output toward the equilibrium level.

Another way to observe a tendency of the economy to head toward equilibrium at point J is from the viewpoint of inventory investment. When firms do not sell all output produced, they add unsold output to their holdings of inventory, and inventory investment increases. At an output level of $1,200 billion, for instance, the $100 billion of unsold goods leads to $100 billion of unplanned inventory investment, which firms do not want. Companies will decrease production to reduce inventory to the desired level, and aggregate output will fall (indicated by the arrow near the horizontal axis). This viewpoint means that unplanned inventory investment for the entire economy I^u equals the excess of output over aggregate demand. In our example, at an output level of $1,200 billion, I^u = $100 billion. If I^u is positive, firms will cut production and output will fall. Output will stop falling only when it has returned to its equilibrium level at point J, where $I^u = 0$.

What happens if aggregate output is below the equilibrium level of output? Let's say output is $800 billion. At this level of output, aggregate demand at point I is $900 billion, $100 billion higher than output (point H on the 45° line). At this level, firms are selling $100 billion more goods than they are producing, so inventory falls below the desired level. The negative unplanned inventory investment ($I^u = -\$100$ billion) will induce firms to increase their production in order to raise inventory to desired levels. As a result, output rises toward the equilibrium level, shown by the arrow in Figure 2. As long as output is below the equilibrium level, unplanned inventory investment will remain negative, firms will continue to raise production, and output will continue to rise. We again see the tendency for the economy to settle at point J, where aggregate demand Y equals output Y^{ad} and unplanned inventory investment is zero ($I^u = 0$).

Expenditure Multiplier

Now that we understand that equilibrium aggregate output is determined by the position of the aggregate demand function, we can examine how different factors shift the function and consequently change aggregate output. We will find that either a rise in planned investment spending or a rise in autonomous consumer expenditure shifts the aggregate demand function upward and leads to an increase in aggregate output.

Output Response to a Change in Planned Investment Spending. Suppose that a new electric motor is invented that makes all factory machines three times more efficient. Because firms are suddenly more optimistic about the profitability of investing in new machines that use this new motor, planned investment spending increases by $100 billion from an initial level of I_1 = $300 billion to I_2 = $400 billion. What effect does this have on output?

The effects of this increase in planned investment spending are analyzed in Figure 3 using a Keynesian cross diagram. Initially, when planned investment spending I_1 is $300 billion, the aggregate demand function is Y_1^{ad}, and equilibrium occurs at point 1, where output is $1,000 billion. The $100 billion increase in planned investment

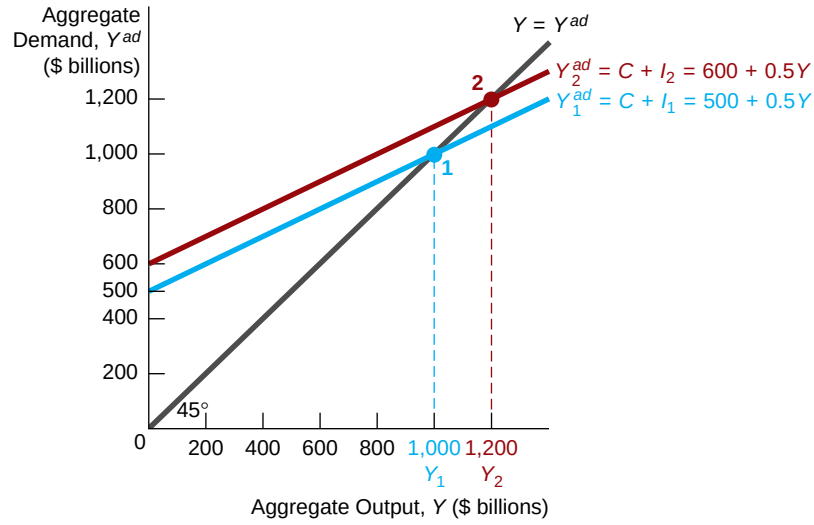


FIGURE 3 Response of Aggregate Output to a Change in Planned Investment

A $100 billion increase in planned investment spending from $I_1 = 300$ to $I_2 = 400$ shifts the aggregate demand function upward from Y_1^{ad} to Y_2^{ad}. The equilibrium moves from point 1 to point 2, and equilibrium output rises from $Y_1 = 1{,}000$ to $Y_2 = 1{,}200$.

spending adds directly to aggregate demand and shifts the aggregate demand function upward to Y_2^{ad}. Aggregate demand now equals output at the intersection of Y_2^{ad} with the 45° line $Y = Y^{ad}$ (point 2). As a result of the $100 billion increase in planned investment spending, equilibrium output rises by $200 billion to $1,200 billion ($Y_2$). For every dollar increase in planned investment spending, aggregate output has increased twofold.

The ratio of the change in aggregate output to a change in planned investment spending, $\Delta Y/\Delta I$, is called the **expenditure multiplier**. (This multiplier should not be confused with the money supply multiplier developed in Chapter 16, which measures the ratio of the change in the money supply to a change in the monetary base.) In Figure 3, the expenditure multiplier is 2.

Why does a change in planned investment spending lead to an even larger change in aggregate output so that the expenditure multiplier is greater than 1? The expenditure multiplier is greater than 1 because an increase in planned investment spending, which raises output, also leads to an additional increase in consumer expenditure ($mpc \times \Delta Y$). The increase in consumer expenditure in turn raises aggregate demand and output further, resulting in a multiple change of output from a given change in planned investment spending. This conclusion can be derived algebraically by solving for the unknown value of Y in terms of a, mpc, and I, resulting in the following equation:[4]

$$Y = (a + I) \times \frac{1}{1 - mpc} \tag{4}$$

[4]Substituting the consumption function $C = a + (mpc \times Y)$ into the aggregate demand function $Y^{ad} = C + I$ yields:

$$Y^{ad} = a + (mpc \times Y) + I$$

Because I is multiplied by the term $1/(1 - mpc)$, this equation tells us that a \$1 change in I leads to a $\$1/(1 - mpc)$ change in aggregate output; thus $1/(1 - mpc)$ is the expenditure multiplier. When $mpc = 0.5$, the change in output for a \$1 change in I is \$2 [$= 1/(1 - 0.5)$]; if $mpc = 0.8$, the change in output for a \$1 change in I is \$5. The larger the marginal propensity to consume, the higher the expenditure multiplier.

Response to Changes in Autonomous Spending. Because a is also multiplied by the term $1/(1 - mpc)$ in Equation 4, a \$1 change in autonomous consumer expenditure a also changes aggregate output by $1/(1 - mpc)$, the amount of the expenditure multiplier. Therefore, we see that the expenditure multiplier applies equally well to changes in autonomous consumer expenditure. In fact, Equation 4 can be rewritten as:

$$Y = A \times \frac{1}{1 - mpc} \tag{5}$$

in which A = autonomous spending $= a + I$.

This rewritten equation tells us that any change in autonomous spending, whether from a change in a, in I, or in both, will lead to a multiplied change in Y. If both a and I decrease by \$100 billion each, so that A decreases by \$200 billion, and $mpc = 0.5$, the expenditure multiplier is 2 [$= 1/(1 - 0.5)$], and aggregate output Y will fall by $2 \times \$200$ billion $= \$400$ billion. Conversely, a rise in I by \$100 billion that is offset by a \$100 billion decline in a will leave autonomous spending A, and hence Y, unchanged. The expenditure multiplier $1/(1 - mpc)$ can therefore be defined more generally as the ratio of the change in aggregate output to a change in autonomous spending ($\Delta Y/\Delta A$).

Another way to reach this conclusion—that any change in autonomous spending will lead to a multiplied change in aggregate output—is to recognize that the shift in the aggregate demand function in Figure 3 did not have to come from an increase in I; it could also have come from an increase in a, which directly raises consumer expenditure and therefore aggregate demand. Alternatively, it could have come from an increase in both a and I. Changes in the attitudes of consumers and firms about the future, which cause changes in their spending, will result in multiple changes in aggregate output.

Keynes believed that changes in autonomous spending are dominated by unstable fluctuations in planned investment spending, which is influenced by emotional waves of optimism and pessimism—factors he labeled "**animal spirits**." His view was colored by the collapse in investment spending during the Great Depression, which he saw as the primary reason for the economic contraction. We will examine the consequences of this fall in investment spending in the following application.

---

[4] *continued*

In equilibrium, where aggregate output equals aggregate demand,

$$Y = Y^{ad} = a + (mpc \times Y) + I$$

Subtracting the term $mpc \times Y$ from both sides of this equation in order to collect the terms involving Y on the left side, we have:

$$Y - (mpc \times Y) = Y(1 - mpc) = a + I$$

Dividing both sides by $1 - mpc$ to solve for Y leads to Equation 4 in the text.

Application

The Collapse of Investment Spending and the Great Depression

From 1929 to 1933, the U.S. economy experienced the largest percentage decline in investment spending ever recorded. One explanation for the investment collapse was the ongoing set of financial crises during this period, described in Chapter 8. In 1996 dollars, investment spending fell from $218 billion to $36 billion—a decline of over 80%. What does the Keynesian analysis developed so far suggest should have happened to aggregate output in this period?

Figure 4 demonstrates how the $182 billion drop in planned investment spending would shift the aggregate demand function downward from Y_1^{ad} to Y_2^{ad}, moving the economy from point 1 to point 2. Aggregate output would then fall sharply; real GDP actually fell by $330 billion (a multiple of the $182 billion drop in investment spending), from $1,111 billion to $781 billion (in 1996 dollars). Because the economy was at full employment in 1929, the fall in output resulted in massive unemployment, with over 25% of the labor force unemployed in 1933.

Government's Role

After witnessing the events in the Great Depression, Keynes took the view that an economy would continually suffer major output fluctuations because of the volatility of autonomous spending, particularly planned investment spending. He was especially worried about sharp declines in autonomous spending, which would inevitably

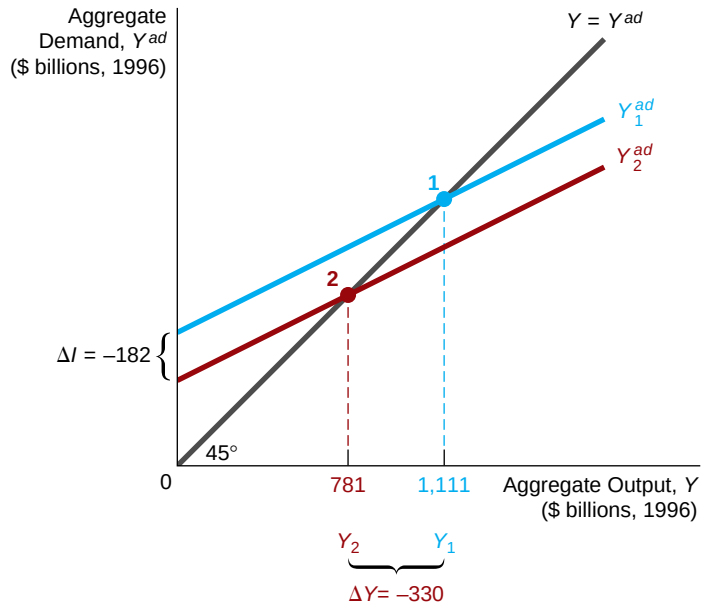


FIGURE 4 Response of Aggregate Output to the Collapse of Investment Spending, 1929–1933

The decline of $182 billion (in 1996 dollars) in planned investment spending from 1929 to 1933 shifted the aggregate demand function down from Y_1^{ad} to Y_2^{ad} and caused the economy to move from point 1 to point 2, where output fell by $330 billion.

Source: Economic Report of the President.

lead to large declines in output and an equilibrium with high unemployment. If autonomous spending fell sharply, as it did during the Great Depression, how could an economy be restored to higher levels of output and more reasonable levels of unemployment? Not by an increase in autonomous investment and consumer spending, since the business outlook was so grim. Keynes's answer to this question involved looking at the role of government in determining aggregate output.

Keynes realized that government spending and taxation could also affect the position of the aggregate demand function and hence be manipulated to restore the economy to full employment. As shown in the aggregate demand equation $Y^{ad} = C + I + G + NX$, government spending G adds directly to aggregate demand. Taxes, however, do not affect aggregate demand directly, as government spending does. Instead, taxes lower the amount of income that consumers have available for spending and affect aggregate demand by influencing consumer expenditure; that is, when there are taxes, disposable income Y_D does not equal aggregate output; it equals aggregate output Y minus taxes T: $Y_D = Y - T$. The consumption function $C = a + (mpc \times Y_D)$ can be rewritten as follows:

$$C = a + [mpc \times (Y - T)] = a + (mpc \times Y) - (mpc \times T) \tag{6}$$

This consumption function looks similar to the one used in the absence of taxes, but it has the additional term $-(mpc \times T)$ on the right side. This term indicates that if taxes increase by \$100, consumer expenditure declines by mpc multiplied by this amount; if $mpc = 0.5$, consumer expenditure declines by \$50. This occurs because consumers view \$100 of taxes as equivalent to a \$100 reduction in income and reduce their expenditure by the marginal propensity to consume times this amount.

To see how the inclusion of government spending and taxes modifies our analysis, first we will observe the effect of a positive level of government spending on aggregate output in the Keynesian cross diagram of Figure 5. Let's say that in the absence of government spending or taxes, the economy is at point 1, where the aggregate demand function $Y_1^{ad} = C + I = 500 + 0.5Y$ crosses the 45° line $Y = Y^{ad}$. Here equilibrium output is at \$1,000 billion. Suppose, however, that the economy reaches full employment at an aggregate output level of \$1,800 billion. How can government spending be used to restore the economy to full employment at \$1,800 billion of aggregate output?

If government spending is set at \$400 billion, the aggregate demand function shifts upward to $Y_2^{ad} = C + I + G = 900 + 0.5Y$. The economy moves to point 2, and aggregate output rises by \$800 billion to \$1,800 billion. Figure 5 indicates that aggregate output is positively related to government spending and that a change in government spending leads to a multiplied change in aggregate output, equal to the expenditure multiplier, $1/(1 - mpc) = 1/(1 - 0.5) = 2$. Therefore, declines in planned investment spending that produce high unemployment (as occurred during the Great Depression) can be offset by raising government spending.

What happens if the government decides that it must collect taxes of \$400 billion to balance the budget? Before taxes are raised, the economy is in equilibrium at the same point 2 found in Figure 5. Our discussion of the consumption function (which allows for taxes) indicates that taxes T reduce consumer expenditure by $mpc \times T$ because there is T less income now available for spending. In our example, $mpc = 0.5$, so consumer expenditure and the aggregate demand function shift downward by \$200 billion ($= 0.5 \times 400$); at the new equilibrium, point 3, the level of output has declined by twice this amount (the expenditure multiplier) to \$1,400 billion.

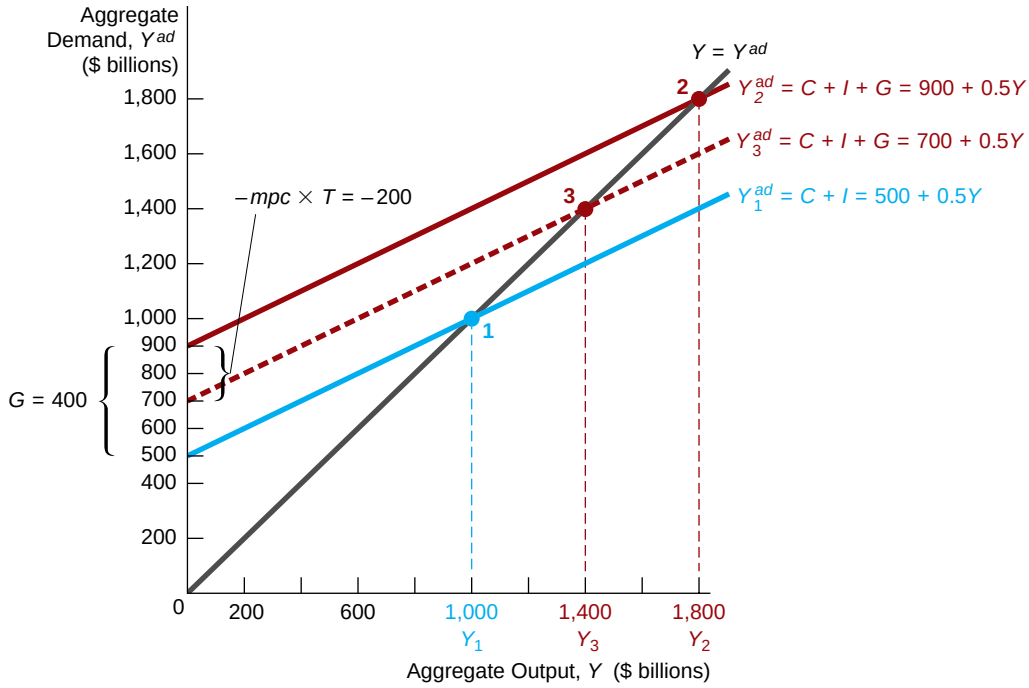


FIGURE 5 Response of Aggregate Output to Government Spending and Taxes

With no government spending or taxes, the aggregate demand function is Y_1^{ad}, and equilibrium output is $Y_1 = 1{,}000$. With government spending of \$400 billion, the aggregate demand function shifts upward to Y_2^{ad}, and aggregate output rises by \$800 billion to $Y_2 = \$1{,}800$ billion. Taxes of \$400 billion lower consumer expenditure and the aggregate demand function by \$200 billion from Y_2^{ad} to Y_3^{ad}, and aggregate output falls by \$400 billion to $Y_3 = \$1{,}400$ billion.

Although you can see that aggregate output is negatively related to the level of taxes, it is important to recognize that the change in aggregate output from the \$400 billion increase in taxes ($\Delta Y = -\$400$ billion) is smaller than the change in aggregate output from the \$400 billion increase in government spending ($\Delta Y = \$800$ billion). If both taxes and government spending are raised equally—by \$400 billion, as occurs in going from point 1 to point 3 in Figure 5, aggregate output will rise.

The Keynesian framework indicates that the government can play an important role in determining aggregate output by changing the level of government spending or taxes. If the economy enters a deep recession, in which output drops severely and unemployment climbs, the analysis we have just developed provides a prescription for restoring the economy to health. The government might raise aggregate output by increasing government spending, or it could lower taxes and reverse the process described in Figure 5 (that is, a tax cut makes more income available for spending at any level of output, shifting the aggregate demand function upward and causing the equilibrium level of output to rise).

Role of International Trade

International trade also plays a role in determining aggregate output because net exports (exports minus imports) are a component of aggregate demand. To analyze the effect of net exports in the Keynesian cross diagram of Figure 6, suppose that initially net exports are equal to zero ($NX_1 = 0$) so that the economy is at point 1, where the aggregate demand function $Y_1^{ad} = C + I + G + NX_1 = 500 + 0.5Y$ crosses the 45° line $Y = Y_1^{ad}$. Equilibrium output is again at $1,000 billion. Now foreigners suddenly get an urge to buy more American products so that net exports rise to $100 billion ($NX_2 = 100$). The $100 billion increase in net exports adds directly to aggregate demand and shifts the aggregate demand function upward to $Y_2^{ad} = C + I + G + NX_2 = 600 + 0.5Y$. The economy moves to point 2, and aggregate output rises by $200 billion to $1,200 billion ($Y_2$). Figure 6 indicates that, just as we found for planned investment spending and government spending, a rise in net exports leads to a multiplied rise in aggregate output, equal to the expenditure multiplier, $1/(1 - mpc) = 1/(1 - 0.5) = 2$. Therefore, changes in net exports can be another important factor affecting fluctuations in aggregate output.

Summary of the Determinants of Aggregate Output

Our analysis of the Keynesian framework so far has identified five autonomous factors (factors independent of income) that shift the aggregate demand function and hence the level of aggregate output:

1. Changes in autonomous consumer expenditure (a)
2. Changes in planned investment spending (I)
3. Changes in government spending (G)
4. Changes in taxes (T)
5. Changes in net exports (NX)

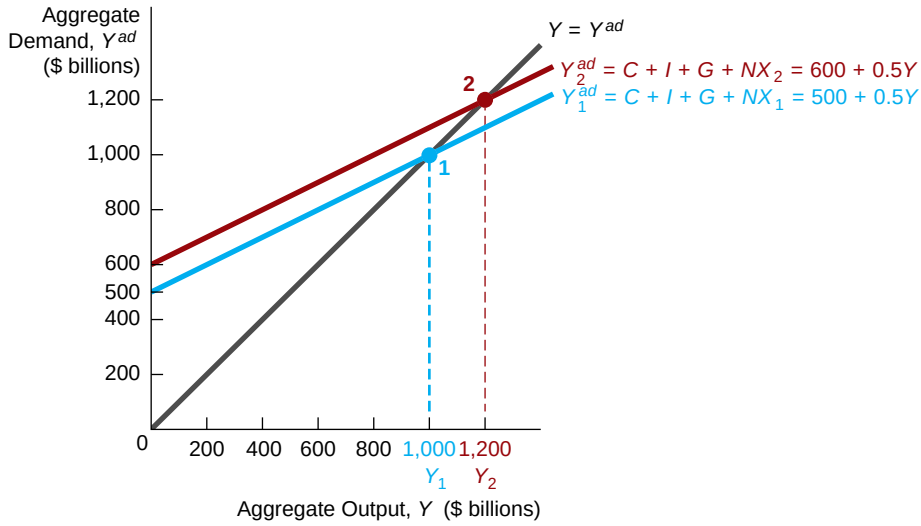


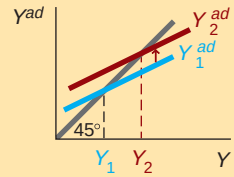
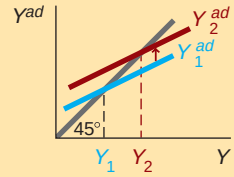
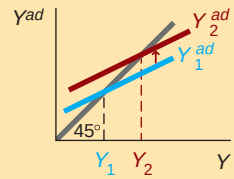
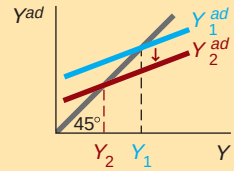
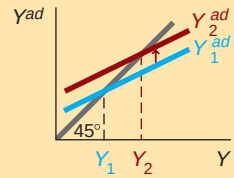
FIGURE 6 Response of Aggregate Output to a Change in Net Exports

A $100 billion increase in net exports from $NX_1 = 0$ to $NX_2 = 100$ shifts the aggregate demand function upward from Y_1^{ad} to Y_2^{ad}. The equilibrium moves from point 1 to point 2, and equilibrium output rises from $Y_1 = $1,000 billion to $Y_2 = $1,200 billion.

The effects of changes in each of these variables on aggregate output are summarized in Table 2 and discussed next in the text.

Changes in Autonomous Consumer Spending (*a*). A rise in autonomous consumer expenditure *a* (say, because consumers become more optimistic about the economy when the stock market booms) directly raises consumer expenditure and shifts the aggregate demand function upward, resulting in an increase in aggregate output. A decrease in *a* causes consumer expenditure to fall, leading ultimately to a decline in

SUMMARY Table 2 Response of Aggregate Output *Y* to Autonomous Changes in *a*, *I*, *G*, *T*, and *NX*

Variable	Change in Variable	Response of Aggregate Output, *Y*	
Autonomous consumer expenditure, *a*	↑	↑	
Investment, *I*	↑	↑	
Government spending, *G*	↑	↑	
Taxes, *T*	↑	↓	
Net exports, *NX*	↑	↑	

Note: Only increases (↑) in the variables are shown; the effects of decreases in the variables on aggregate output would be the opposite of those indicated in the "Response" column.

aggregate output. Therefore, ***aggregate output is positively related to autonomous consumer expenditure a***.

Changes in Planned Investment Spending (*I*). A rise in planned investment spending adds directly to aggregate demand, thus raising the aggregate demand function and aggregate output. A fall in planned investment spending lowers aggregate demand and causes aggregate output to fall. Therefore, ***aggregate output is positively related to planned investment spending I***.

Changes in Government Spending (*G*). A rise in government spending also adds directly to aggregate demand and raises the aggregate demand function, increasing aggregate output. A fall directly reduces aggregate demand, lowers the aggregate demand function, and causes aggregate output to fall. Therefore, ***aggregate output is positively related to government spending G***.

Changes in Taxes (*T*). A rise in taxes does not affect aggregate demand directly, but does lower the amount of income available for spending, reducing consumer expenditure. The decline in consumer expenditure then leads to a fall in the aggregate demand function, resulting in a decline in aggregate output. A lowering of taxes makes more income available for spending, raises consumer expenditure, and leads to higher aggregate output. Therefore, ***aggregate output is negatively related to the level of taxes T***.

Changes in Net Exports (*NX*). A rise in net exports adds directly to aggregate demand and raises the aggregate demand function, increasing aggregate output. A fall directly reduces aggregate demand, lowers the aggregate demand function, and causes aggregate output to fall. Therefore, ***aggregate output is positively related to net exports NX***.

Size of the Effects from the Five Factors. The aggregate demand function in the Keynesian cross diagrams shifts vertically by the full amount of the change in a, I, G, or NX, resulting in a multiple effect on aggregate output through the effects of the expenditure multiplier, $1/(1 - mpc)$. A change in taxes has a smaller effect on aggregate output, because consumer expenditure changes only by mpc times the change in taxes ($-mpc \times \Delta T$), which in the case of $mpc = 0.5$ means that aggregate demand shifts vertically by only half of the change in taxes.

If there is a change in one of these autonomous factors that is offset by a change in another (say, I rises by \$100 billion, but a, G, or NX falls by \$100 billion or T rises by \$200 billion when $mpc = 0.5$), the aggregate demand function will remain in the same position, and aggregate output will remain unchanged.[5]

---

[5]These results can be derived algebraically as follows. Substituting the consumption function allowing for taxes (Equation 6) into the aggregate demand function (Equation 1), we have:

$$Y^{ad} = a - (mpc \times T) + (mpc \times Y) + I + G + NX$$

If we assume that taxes T are unrelated to income, we can define autonomous spending in the aggregate demand function to be:

$$A = a - (mpc \times T) + I + G + NX$$

The expenditure equation can be rewritten as:

$$Y^{ad} = A + (mpc \times Y)$$

In equilibrium, aggregate demand equals aggregate output:

$$Y = A + (mpc \times Y)$$

Study Guide

To test your understanding of the Keynesian analysis of how aggregate output changes in response to changes in the factors described, see if you can use Keynesian cross diagrams to illustrate what happens to aggregate output when each variable decreases rather than increases. Also, be sure to do the problems at the end of the chapter that ask you to predict what will happen to aggregate output when certain economic variables change.

The *ISLM* Model

So far, our analysis has excluded monetary policy. We now include money and interest rates in the Keynesian framework in order to develop the more intricate *ISLM* model of how aggregate output is determined, in which monetary policy plays an important role. Why another complex model? The *ISLM* model is more versatile and allows us to understand economic phenomena that cannot be analyzed with the simpler Keynesian cross framework used earlier. The *ISLM* model will help you understand how monetary policy affects economic activity and interacts with fiscal policy (changes in government spending and taxes) to produce a certain level of aggregate output; how the level of interest rates is affected by changes in investment spending as well as by changes in monetary and fiscal policy; how best to conduct monetary policy; and how the *ISLM* model generates the aggregate demand curve, an essential building block for the aggregate supply and demand analysis used in Chapter 25 and thereafter.

Like our simplified Keynesian model, the full Keynesian *ISLM* model examines an equilibrium in which aggregate output produced equals aggregate demand, and since it assumes a fixed price level, real and nominal quantities are the same. The first step in constructing the *ISLM* model is to examine the effect of interest rates on planned investment spending and hence on aggregate demand. Next we use a Keynesian cross diagram to see how the interest rate affects the equilibrium level of aggregate output. The resulting relationship between equilibrium aggregate output and the interest rate is known as the ***IS* curve**.

Just as a demand curve alone cannot tell us the quantity of goods sold in a market, the *IS* curve by itself cannot tell us what the level of aggregate output will be because the interest rate is still unknown. We need another relationship, called the ***LM* curve**, which describes the combinations of interest rates and aggregate output for which the quantity of money demanded equals the quantity of money supplied.

---

which can be solved for Y. The resulting equation:

$$Y = A \times \frac{1}{1 - mpc}$$

is the same equation that links autonomous spending and aggregate output in the text (Equation 5), but it now allows for additional components of autonomous spending in A. We see that any increase in autonomous expenditure leads to a multiple increase in output. Thus any component of autonomous spending that enters A with a positive sign (a, I, G, and NX) will have a positive relationship with output, and any component with a negative sign ($-mpc \times T$) will have a negative relationship with output. This algebraic analysis also shows us that any rise in a component of A that is offset by a movement in another component of A, leaving A unchanged, will also leave output unchanged.

When the *IS* and *LM* curves are combined in the same diagram, the intersection of the two determines the equilibrium level of aggregate output as well as the interest rate. Finally, we will have obtained a more complete analysis of the determination of aggregate output in which monetary policy plays an important role.

Equilibrium in the Goods Market: The *IS* Curve

In Keynesian analysis, the primary way that interest rates affect the level of aggregate output is through their effects on planned investment spending and net exports. After explaining why interest rates affect planned investment spending and net exports, we will use Keynesian cross diagrams to learn how interest rates affect equilibrium aggregate output.[6]

Interest Rates and Planned Investment Spending. Businesses make investments in physical capital (machines, factories, and raw materials) as long as they expect to earn more from the physical capital than the interest cost of a loan to finance the investment. When the interest rate is high, few investments in physical capital will earn more than the cost of borrowed funds, so planned investment spending is low. When the interest rate is low, many investments in physical capital will earn more than the interest cost of borrowed funds. Therefore, when interest rates are lower, business firms are more likely to undertake an investment in physical capital, and planned investment spending will be higher.

Even if a company has surplus funds and does not need to borrow to undertake an investment in physical capital, its planned investment spending will be affected by the interest rate. Instead of investing in physical capital, it could purchase a security, such as a bond. If the interest rate on this security is high, the opportunity cost (forgone interest earnings) of an investment is high, and planned investment spending will be low, because the firm would probably prefer to purchase the security than to invest in physical capital. As the interest rate and the opportunity cost of investing fall, planned investment spending will increase because investments in physical capital are more likely than the security to earn greater income for the firm.

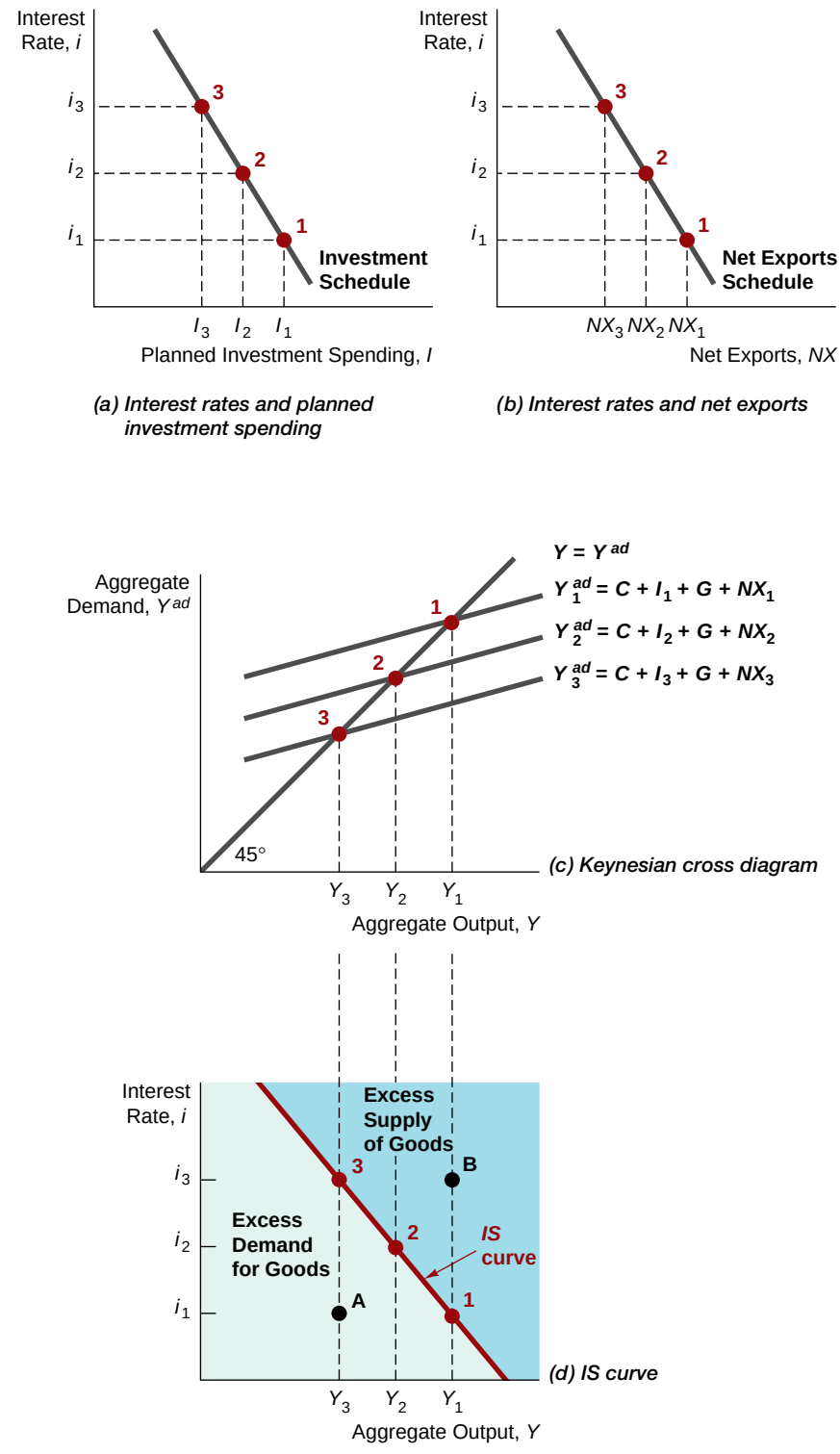
The relationship between the amount of planned investment spending and any given level of the interest rate is illustrated by the investment schedule in panel (a) of Figure 7. The downward slope of the schedule reflects the negative relationship between planned investment spending and the interest rate. At a low interest rate i_1, the level of planned investment spending I_1 is high; for a high interest rate i_3, planned investment spending I_3 is low.

Interest Rates and Net Exports. As discussed in more detail in Chapter 19, when interest rates rise in the United States (with the price level fixed), U.S. dollar bank deposits become more attractive relative to deposits denominated in foreign currencies, thereby causing a rise in the value of dollar deposits relative to other currency deposits; that is, a rise in the exchange rate. The higher value of the dollar resulting from the rise in interest rates makes domestic goods more expensive than foreign goods, thereby causing a fall in net exports. The resulting negative relationship between interest rates and net exports is shown in panel (b) of Figure 7. At a low

[6]More modern Keynesian approaches suggest that consumer expenditure, particularly for consumer durables (cars, furniture, appliances), is influenced by the interest rate. This interest sensitivity of consumer expenditure can be allowed for in the model here by defining planned investment spending more generally to include the interest-sensitive component of consumer expenditure.

FIGURE 7 Deriving the *IS* Curve

The investment schedule in panel (a) shows that as the interest rate rises from i_1 to i_2 to i_3, planned investment spending falls from I_1 to I_2 to I_3, and panel (b) shows that net exports also fall from NX_1 to NX_2 to NX_3 as the interest rate rises. Panel (c) then indicates the levels of equilibrium output Y_1, Y_2, and Y_3 that correspond to those three levels of planned investment and net exports. Finally, panel (d) plots the level of equilibrium output corresponding to each of the three interest rates; the line that connects these points is the *IS* curve.



interest rate i_1, the exchange rate is low and net exports NX_1 are high; at a high interest rate i_3, the exchange rate is high and net exports NX_3 are low.

Deriving the *IS* Curve. We can now use what we have learned about the relationship of interest rates to planned investment spending and net exports in panels (a) and (b) to examine the relationship between interest rates and the equilibrium level of aggregate output (holding government spending and autonomous consumer expenditure constant). The three levels of planned investment spending and net exports in panels (a) and (b) are represented in the three aggregate demand functions in the Keynesian cross diagram of panel (c). The lowest interest rate i_1 has the highest level of both planned investment spending I_1 and net exports NX_1, and hence the highest aggregate demand function Y_1^{ad}. Point 1 in panel (d) shows the resulting equilibrium level of output Y_1, which corresponds to interest rate i_1. As the interest rate rises to i_2, both planned investment spending and net exports fall, to I_2 and NX_2, so equilibrium output falls to Y_2. Point 2 in panel (d) shows the lower level of output Y_2, which corresponds to interest rate i_2. Finally, the highest interest rate i_3 leads to the lowest level of planned investment spending and net exports, and hence the lowest level of equilibrium output, which is plotted as point 3.

The line connecting the three points in panel (d), the *IS* curve, shows the combinations of interest rates and equilibrium aggregate output for which aggregate output produced equals aggregate demand.[7] The negative slope indicates that higher interest rates result in lower planned investment spending and net exports, and hence lower equilibrium output.

What the *IS* Curve Tells Us. The *IS* curve traces out the points at which the total quantity of goods produced equals the total quantity of goods demanded. It describes points at which the goods market is in equilibrium. For each given level of the interest rate, the *IS* curve tells us what aggregate output must be for the goods market to be in equilibrium. As the interest rate rises, planned investment spending and net exports fall, which in turn lowers aggregate demand; aggregate output must be lower in order for it to equal aggregate demand and satisfy goods market equilibrium.

The *IS* curve is a useful concept because output tends to move toward points on the curve that satisfy goods market equilibrium. If the economy is located in the area to the right of the *IS* curve, it has an excess supply of goods. At point B, for example, aggregate output Y_1 is greater than the equilibrium level of output Y_3 on the *IS* curve. This excess supply of goods results in unplanned inventory accumulation, which causes output to fall toward the *IS* curve. The decline stops only when output is again at its equilibrium level on the *IS* curve.

If the economy is located in the area to the left of the *IS* curve, it has an excess demand for goods. At point A, aggregate output Y_3 is below the equilibrium level of output Y_1 on the *IS* curve. The excess demand for goods results in an unplanned decrease in inventory, which causes output to rise toward the *IS* curve, stopping only when aggregate output is again at its equilibrium level on the *IS* curve.

Significantly, equilibrium in the goods market does not produce a unique equilibrium level of aggregate output. Although we now know where aggregate output will head for a given level of the interest rate, we cannot determine aggregate output

[7] The *IS* was so named by Sir John Hicks because in the simplest Keynesian framework with no government sector, equilibrium in the Keynesian cross diagram occurs when investment spending *I* equals savings *S*.

because we do not know what the interest rate is. To complete our analysis of aggregate output determination, we need to introduce another market that produces an additional relationship that links aggregate output and interest rates. The market for money fulfills this function with the *LM* curve. When the *LM* curve is combined with the *IS* curve, a unique equilibrium that determines both aggregate output and the interest rate is obtained.

Equilibrium in the Market for Money: The *LM* Curve

Just as the *IS* curve is derived from the equilibrium condition in the goods market (aggregate output equals aggregate demand), the *LM* curve is derived from the equilibrium condition in the market for money, which requires that the quantity of money demanded equal the quantity of money supplied. The main building block in Keynes's analysis of the market for money is the demand for money he called *liquidity preference*. Let us briefly review his theory of the demand for money (discussed at length in Chapters 5 and 22).

Keynes's liquidity preference theory states that the demand for money in real terms M^d/P depends on income Y (aggregate output) and interest rates i. The demand for money is positively related to income for two reasons. First, a rise in income raises the level of transactions in the economy, which in turn raises the demand for money because it is used to carry out these transactions. Second, a rise in income increases the demand for money because it increases the wealth of individuals who want to hold more assets, one of which is money. The opportunity cost of holding money is the interest sacrificed by not holding other assets (such as bonds) instead. As interest rates rise, the opportunity cost of holding money rises, and the demand for money falls. According to the liquidity preference theory, the demand for money is positively related to aggregate output and negatively related to interest rates.

Deriving the *LM* Curve. In Keynes's analysis, the level of interest rates is determined by equilibrium in the market for money, at which point the quantity of money demanded equals the quantity of money supplied. Figure 8 depicts what happens to equilibrium in the market for money as the level of output changes. Because the LM curve is derived holding the money supply at a fixed level, it is fixed at the level of $\overline{M}$ in panel (a).[8] Each level of aggregate output has its own money demand curve because as aggregate output changes, the level of transactions in the economy changes, which in turn changes the demand for money.

When aggregate output is Y_1, the money demand curve is $M^d(Y_1)$: It slopes downward because a lower interest rate means that the opportunity cost of holding money is lower, so the quantity of money demanded is higher. Equilibrium in the market for money occurs at point 1, at which the interest rate is i_1. When aggregate output is at the higher level Y_2, the money demand curve shifts rightward to $M^d(Y_2)$ because the higher level of output means that at any given interest rate, the quantity of money demanded is higher. Equilibrium in the market for money now occurs at point 2, at which the interest rate is at the higher level of i_2. Similarly, a still higher level of aggregate output Y_3 results in an even higher level of the equilibrium interest rate i_3.

[8] As pointed out in earlier chapters on the money supply process, the money supply is positively related to interest rates, and so the M^s curve in panel (a) should actually have a positive slope. The M^s curve is assumed to be vertical in panel (a) in order to simplify the graph, but allowing for a positive slope leads to identical results.

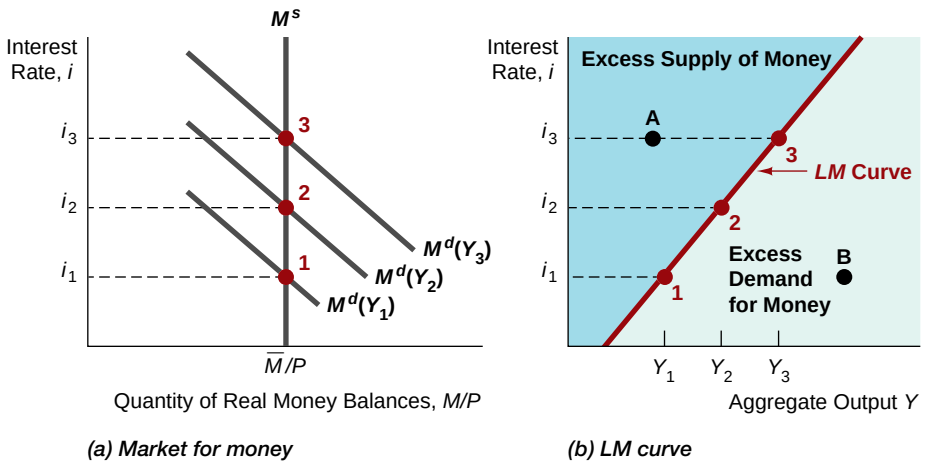


FIGURE 8 Deriving the *LM* Curve

Panel (a) shows the equilibrium levels of the interest rate in the market for money that arise when aggregate output is at Y_1, Y_2, and Y_3. Panel (b) plots the three levels of the equilibrium interest rate i_1, i_2, and i_3 corresponding to these three levels of output; the line that connects these points is the *LM* curve.

Panel (b) plots the equilibrium interest rates that correspond to the different output levels, with points 1, 2, and 3 corresponding to the equilibrium points 1, 2, and 3 in panel (a). The line connecting these points is the *LM* curve, which shows the combinations of interest rates and output for which the market for money is in equilibrium.[9] The positive slope arises because higher output raises the demand for money and thus raises the equilibrium interest rate.

What the *LM* Curve Tells Us. The *LM* curve traces out the points that satisfy the equilibrium condition that the quantity of money demanded equals the quantity of money supplied. For each given level of aggregate output, the *LM* curve tells us what the interest rate must be for there to be equilibrium in the market for money. As aggregate output rises, the demand for money increases and the interest rate rises, so that money demanded equals money supplied and the market for money is in equilibrium.

Just as the economy tends to move toward the equilibrium points represented by the *IS* curve, it also moves toward the equilibrium points on the *LM* curve. If the economy is located in the area to the left of the *LM* curve, there is an excess supply of money. At point A, for example, the interest rate is i_3 and aggregate output is Y_1. The interest rate is above the equilibrium level, and people are holding more money than they want to. To eliminate their excess money balances, they will purchase bonds, which causes the price of the bonds to rise and their interest rate to fall. (The inverse relationship between the price of a bond and its interest rate is discussed in

[9]Hicks named this the *LM* curve to indicate that it represents the combinations of interest rates and output for which money demand, which Keynes denoted as *L* to represent liquidity preference, equals money supply *M*.

Chapter 4.) As long as an excess supply of money exists, the interest rate will fall until it comes to rest on the *LM* curve.

If the economy is located in the area to the right of the *LM* curve, there is an excess demand for money. At point B, for example, the interest rate i_1 is below the equilibrium level, and people want to hold more money than they currently do. To acquire this money, they will sell bonds and drive down bond prices, and the interest rate will rise. This process will stop only when the interest rate rises to an equilibrium point on the *LM* curve.

ISLM Approach to Aggregate Output and Interest Rates

Now that we have derived the *IS* and *LM* curves, we can put them into the same diagram (Figure 9) to produce a model that enables us to determine both aggregate output and the interest rate. The only point at which the goods market and the market for money are in simultaneous equilibrium is at the intersection of the *IS* and *LM* curves, point E. At this point, aggregate output equals aggregate demand (*IS*) and the quantity of money demanded equals the quantity of money supplied (*LM*). At any other point in the diagram, at least one of these equilibrium conditions is not satisfied, and market forces move the economy toward the general equilibrium, point E.

To learn how this works, let's consider what happens if the economy is at point A, which is on the *IS* curve but not the *LM* curve. Even though at point A the goods market is in equilibrium, so that aggregate output equals aggregate demand, the interest rate is above its equilibrium level, so the demand for money is less than the supply. Because people have more money than they want to hold, they will try to get rid of it by buying bonds. The resulting rise in bond prices causes a fall in interest rates,

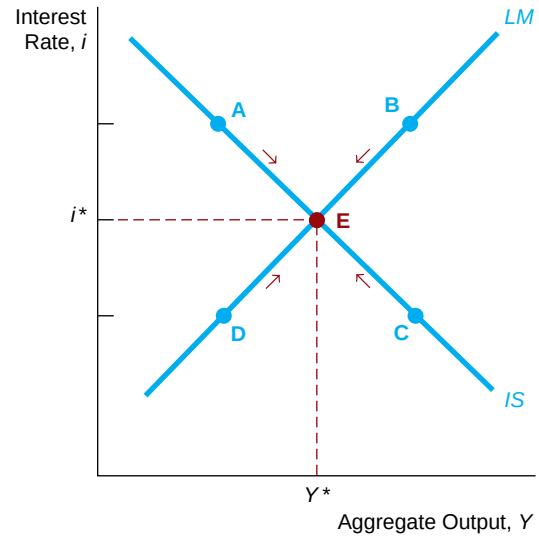


FIGURE 9 *ISLM* Diagram: Simultaneous Determination of Output and the Interest Rate

Only at point E, when the interest rate is i^* and output is Y^*, is there equilibrium simultaneously in both the goods market (as measured by the *IS* curve) and the market for money (as measured by the *LM* curve). At other points, such as A, B, C, or D, one of the two markets is not in equilibrium, and there will be a tendency to head toward the equilibrium, point E.

which in turn causes both planned investment spending and net exports to rise, and thus aggregate output rises. The economy then moves down along the *IS* curve, and the process continues until the interest rate falls to i^* and aggregate output rises to Y^*—that is, until the economy is at equilibrium point E.

If the economy is on the *LM* curve but off the *IS* curve at point B, it will also head toward the equilibrium at point E. At point B, even though money demand equals money supply, output is higher than the equilibrium level and exceeds aggregate demand. Firms are unable to sell all their output, and unplanned inventory accumulates, prompting them to cut production and lower output. The decline in output means that the demand for money will fall, lowering interest rates. The economy then moves down along the *LM* curve until it reaches equilibrium point E.

Study Guide

To test your understanding of why the economy heads toward equilibrium point E at the intersection of the *IS* and *LM* curves, see if you can provide the reasoning behind the movement to point E from points such as C and D in the figure.

We have finally developed a model, the *ISLM* model, that tells us how both interest rates and aggregate output are determined when the price level is fixed. Although we have demonstrated that the economy will head toward an aggregate output level of Y^*, there is no reason to assume that at this level of aggregate output the economy is at full employment. If the unemployment rate is too high, government policymakers might want to increase aggregate output to reduce it. The *ISLM* apparatus indicates that they can do this by manipulating monetary and fiscal policy. We will conduct an *ISLM* analysis of how monetary and fiscal policy can affect economic activity in the next chapter.

Summary

1. In the simple Keynesian framework in which the price level is fixed, output is determined by the equilibrium condition in the goods market that aggregate output equals aggregate demand. Aggregate demand equals the sum of consumer expenditure, planned investment spending, government spending, and net exports. Consumer expenditure is described by the consumption function, which indicates that consumer expenditure will rise as disposable income increases. Keynes's analysis shows that aggregate output is positively related to autonomous consumer expenditure, planned investment spending, government spending, and net exports and negatively related to the level of taxes. A change in any of these factors leads, through the expenditure multiplier, to a multiple change in aggregate output.
2. The *ISLM* model determines aggregate output and the interest rate for a fixed price level using the *IS* and *LM* curves. The *IS* curve traces out the combinations of the interest rate and aggregate output for which the goods market is in equilibrium, and the *LM* curve traces out the combinations for which the market for money is in equilibrium. The *IS* curve slopes downward, because higher interest rates lower planned investment spending and net exports and so lower equilibrium output. The *LM* curve slopes upward, because higher aggregate output raises the demand for money and so raises the equilibrium interest rate.
3. The simultaneous determination of output and interest rates occurs at the intersection of the *IS* and *LM* curves, where both the goods market and the market for money are in equilibrium. At any other level of interest rates and output, at least one of the markets will be out of equilibrium, and forces will move the economy toward the general equilibrium point at the intersection of the *IS* and *LM* curves.

Key Terms

aggregate demand, p. 537
aggregate demand function, p. 541
"animal spirits," p. 544
autonomous consumer expenditure, p. 538
consumer expenditure, p. 536
consumption function, p. 538
disposable income, p. 538
expenditure multiplier, p. 543
fixed investment, p. 539
government spending, p. 537
inventory investment, p. 539
IS curve, p. 551
LM curve, p. 551
marginal propensity to consume, p. 538
net exports, p. 537
planned investment spending, p. 536

Questions and Problems

Questions marked with an asterisk are answered at the end of the book in an appendix, "Answers to Selected Questions and Problems."

1. Calculate the value of the consumption function at each level of disposable income in Table 1 if $a = 100$ and $mpc = 0.9$.

*2. Why do companies cut production when they find that their unplanned inventory investment is greater than zero? If they didn't cut production, what effect would this have on their profits? Why?

3. Plot the consumption function $C = 100 + 0.75Y$ on graph paper.
   a. Assuming no government sector, if planned investment spending is 200, what is the equilibrium level of aggregate output? Show this equilibrium level on the graph you have drawn.
   b. If businesses become more pessimistic about the profitability of investment and planned investment spending falls by 100, what happens to the equilibrium level of output?

*4. If the consumption function is $C = 100 + 0.8Y$ and planned investment spending is 200, what is the equilibrium level of output? If planned investment falls by 100, how much does the equilibrium level of output fall?

5. Why are the multipliers in Problems 3 and 4 different? Explain intuitively why one is higher than the other.

*6. If firms suddenly become more optimistic about the profitability of investment and planned investment spending rises by $100 billion, while consumers become more pessimistic and autonomous consumer spending falls by $100 billion, what happens to aggregate output?

7. "A rise in planned investment spending by $100 billion at the same time that autonomous consumer expenditure falls by $50 billion has the same effect on aggregate output as a rise in autonomous consumer expenditure alone by $50 billion." Is this statement true, false, or uncertain? Explain your answer.

*8. If the consumption function is $C = 100 + 0.75Y$, $I = 200$, and government spending is 200, what will be the equilibrium level of output? Demonstrate your answer with a Keynesian cross diagram. What happens to aggregate output if government spending rises by 100?

9. If the marginal propensity to consume is 0.5, how much would government spending have to rise in order to raise output by $1,000 billion?

*10. Suppose that government policymakers decide that they will change taxes to raise aggregate output by $400 billion, and $mpc = 0.5$. By how much will taxes have to be changed?

11. What happens to aggregate output if both taxes and government spending are lowered by $300 billion and $mpc = 0.5$? Explain your answer.

*12. Will aggregate output rise or fall if an increase in autonomous consumer expenditure is matched by an equal increase in taxes?

13. If a change in the interest rate has no effect on planned investment spending, trace out what happens to the equilibrium level of aggregate output as interest rates fall. What does this imply about the slope of the *IS* curve?

*14. Using a supply and demand diagram for the market for money, show what happens to the equilibrium level of the interest rate as aggregate output falls. What does this imply about the slope of the *LM* curve?

15. "If the point describing the combination of the interest rate and aggregate output is not on either the *IS* or the *LM* curve, the economy will have no tendency to head toward the intersection of the two curves." Is this statement true, false, or uncertain? Explain your answer.

Web Exercises

1. The study tip on page 538 suggests that you construct a consumption function based on your own propensity to consume. This process can be automated by using the tools available at http://nova.umuc.edu/~black/consf1000.html. Assume your level of consumer expenditure as $2,000 (input as 20) and that your marginal propensity to spend is 0.8. Review the resulting graphs. At an income level of $5,000 (50 on the graph), what is your expenditure?
2. Refer to question 1. Again go to http://nova.umuc.edu/~black/consf1000.html. Input any level of consumer expenditure and marginal propensity to consume. According to the resulting graph, where do the 45° line and the consumption function cross? What is the significance of this point? Will you be a saver or dissaver at income levels below this point?

Chapter 24

Monetary and Fiscal Policy in the *ISLM* Model

PREVIEW

Since World War II, government policymakers have tried to promote high employment without causing inflation. If the economy experiences a recession such as the one that began in March 2001, policymakers have two principal sets of tools that they can use to affect aggregate economic activity: *monetary policy*, the control of interest rates or the money supply, and *fiscal policy*, the control of government spending and taxes.

The *ISLM* model can help policymakers predict what will happen to aggregate output and interest rates if they decide to increase the money supply or increase government spending. In this way, *ISLM* analysis enables us to answer some important questions about the usefulness and effectiveness of monetary and fiscal policy in influencing economic activity.

But which is better? When is monetary policy more effective than fiscal policy at controlling the level of aggregate output, and when is it less effective? Will fiscal policy be more effective if it is conducted by changing government spending rather than changing taxes? Should the monetary authorities conduct monetary policy by manipulating the money supply or interest rates?

In this chapter, we use the *ISLM* model to help answer these questions and to learn how the model generates the aggregate demand curve featured prominently in the aggregate demand and supply framework (examined in Chapter 25), which is used to understand changes not only in aggregate output but in the price level as well. Our analysis will show why economists focus so much attention on topics such as the stability of the demand for money function and whether the demand for money is strongly influenced by interest rates.

First, however, let's examine the *ISLM* model in more detail to see how the *IS* and *LM* curves developed in Chapter 23 shift and the implications of these shifts. (We continue to assume that the price level is fixed so that real and nominal quantities are the same.)

Factors That Cause the *IS* Curve to Shift

You have already learned that the *IS* curve describes equilibrium points in the goods market—the combinations of aggregate output and interest rate for which aggregate output produced equals aggregate demand. The *IS* curve shifts whenever a change in autonomous factors (independent of aggregate output) occurs that is unrelated to the interest rate. (A change in the interest rate that affects equilibrium aggregate output

http://cepa.newschool.edu/het/essays/keynes/hickshansen.htm
A detailed discussion of *ISLM* analysis.

causes a movement only along the *IS* curve.) We have already identified five candidates as autonomous factors that can shift aggregate demand and hence affect the level of equilibrium output. We can now ask how changes in each of these factors affect the *IS* curve.

1. Changes in Autonomous Consumer Expenditure. A rise in autonomous consumer expenditure shifts aggregate demand upward and shifts the *IS* curve to the right (see Figure 1). To see how this shift occurs, suppose that the *IS* curve is initially at IS_1 in panel (a) and a huge oil field is discovered in Wyoming, perhaps containing more oil than in Saudi Arabia. Consumers now become more optimistic about the future health of the economy, and autonomous consumer expenditure rises. What happens to the

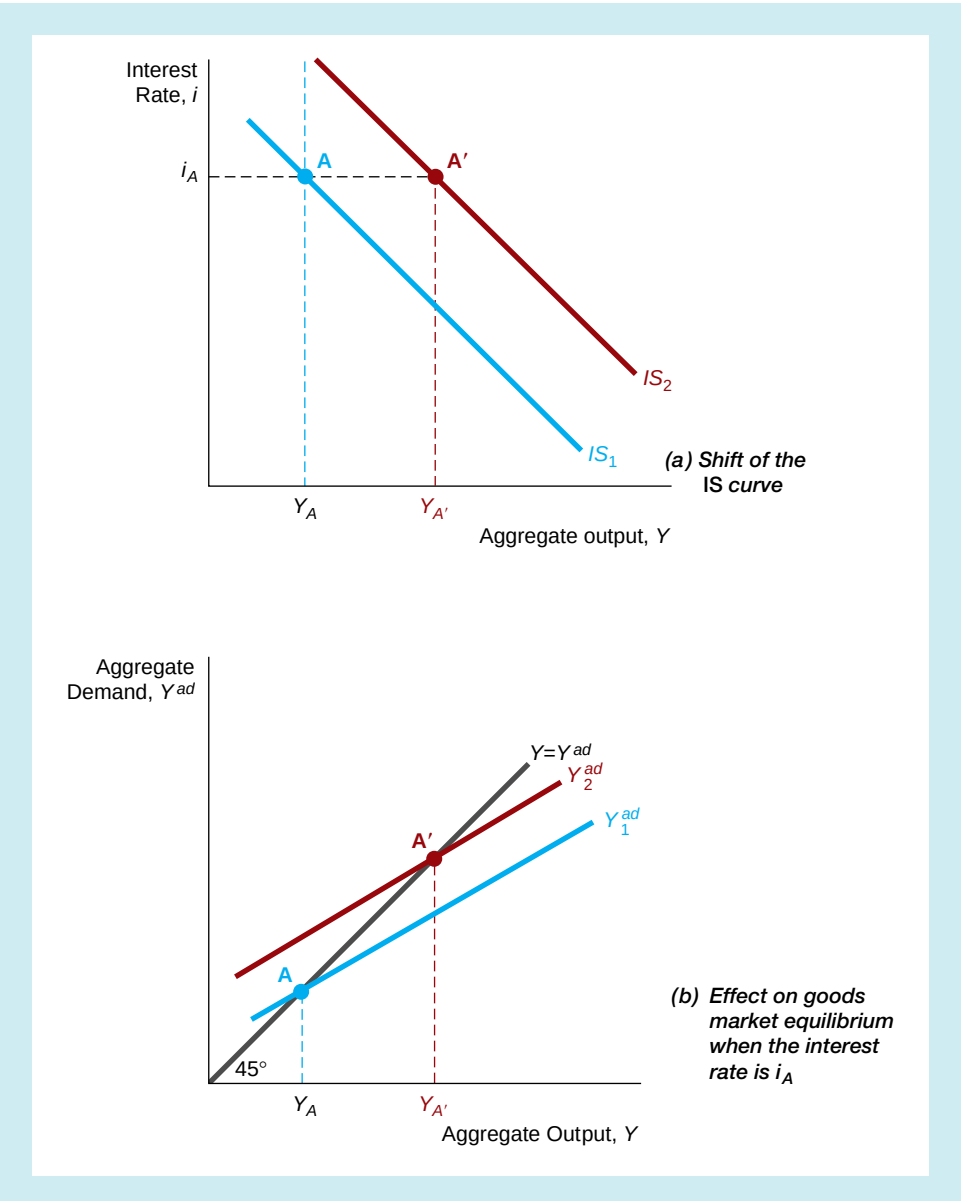


FIGURE 1 Shift in the *IS* Curve
The *IS* curve will shift from IS_1 to IS_2 as a result of (1) an increase in autonomous consumer spending, (2) an increase in planned investment spending due to business optimism, (3) an increase in government spending, (4) a decrease in taxes, or (5) an increase in net exports that is unrelated to interest rates. Panel (b) shows how changes in these factors lead to the rightward shift in the *IS* curve using a Keynesian cross diagram. For any given interest rate (here i_A), these changes shift the aggregate demand function upward and raise equilibrium output from Y_A to $Y_{A'}$.

equilibrium level of aggregate output as a result of this rise in autonomous consumer expenditure when the interest rate is held constant at i_A?

The IS_1 curve tells us that equilibrium aggregate output is at Y_A when the interest rate is at i_A (point A). Panel (b) shows that this point is an equilibrium in the goods market because the aggregate demand function Y_1^{ad} at an interest rate i_A crosses the 45° line $Y = Y^{ad}$ at an aggregate output level of Y_A. When autonomous consumer expenditure rises because of the oil discovery, the aggregate demand function shifts upward to Y_2^{ad} and equilibrium output rises to $Y_{A'}$. This rise in equilibrium output from Y_A to $Y_{A'}$ when the interest rate is i_A is plotted in panel (a) as a movement from point A to point A′. The same analysis can be applied to every point on the initial IS_1 curve; therefore, the rise in autonomous consumer expenditure shifts the *IS* curve to the right from IS_1 to IS_2 in panel (a).

A decline in autonomous consumer expenditure reverses the direction of the analysis. For any given interest rate, the aggregate demand function shifts downward, the equilibrium level of aggregate output falls, and the *IS* curve shifts to the left.

2. Changes in Investment Spending Unrelated to the Interest Rate. In Chapter 23, we learned that changes in the interest rate affect planned investment spending and hence the equilibrium level of output, but this change in investment spending merely causes a movement along the *IS* curve and not a shift. A rise in planned investment spending unrelated to the interest rate (say, because companies become more confident about investment profitability after the Wyoming oil discovery) shifts the aggregate demand function upward, as in panel (b) of Figure 1. For any given interest rate, the equilibrium level of aggregate output rises, and the *IS* curve will shift to the right, as in panel (a).

A decrease in investment spending because companies become more pessimistic about investment profitability shifts the aggregate demand function downward for any given interest rate; the equilibrium level of aggregate output falls, shifting the *IS* curve to the left.

3. Changes in Government Spending. An increase in government spending will also cause the aggregate demand function at any given interest rate to shift upward, as in panel (b). The equilibrium level of aggregate output rises at any given interest rate, and the *IS* curve shifts to the right. Conversely, a decline in government spending shifts the aggregate demand function downward, and the equilibrium level of output falls, shifting the *IS* curve to the left.

4. Changes in Taxes. Unlike changes in other factors that directly affect the aggregate demand function, a decline in taxes shifts the aggregate demand function by raising consumer expenditure and shifting the aggregate demand function upward at any given interest rate. A decline in taxes raises the equilibrium level of aggregate output at any given interest rate and shifts the *IS* curve to the right (as in Figure 1). Recall, however, that a change in taxes has a smaller effect on aggregate demand than an equivalent change in government spending. So for a given change in taxes, the *IS* curve will shift less than for an equal change in government spending.

A rise in taxes lowers the aggregate demand function and reduces the equilibrium level of aggregate output at each interest rate. Therefore, a rise in taxes shifts the *IS* curve to the left.

5. Changes in Net Exports Unrelated to the Interest Rate. As with planned investment spending, changes in net exports arising from a change in interest rates merely cause a movement along the *IS* curve and not a shift. An autonomous rise in net exports unrelated to the interest rate—say, because American-made jeans become more chic than French-made jeans—shifts the aggregate demand function upward and causes the *IS* curve to shift to the right, as in Figure 1. Conversely, an autonomous

fall in net exports shifts the aggregate demand function downward, and the equilibrium level of output falls, shifting the *IS* curve to the left.

Factors That Cause the *LM* Curve to Shift

http://web.mit.edu/rigobon/www/Pdfs/islm.pdf
Visit this web site for an additional discussion of factors that cause shifts in the *LM* curve.

The *LM* curve describes the equilibrium points in the market for money—the combinations of aggregate output and interest rate for which the quantity of money demanded equals the quantity of money supplied. Whereas five factors can cause the *IS* curve to shift (changes in autonomous consumer expenditure, planned investment spending unrelated to the interest rate, government spending, taxes, and net exports unrelated to the interest rate), only two factors can cause the *LM* curve to shift: autonomous changes in money demand and changes in the money supply. How do changes in these two factors affect the *LM* curve?

1. *Changes in the Money Supply.* A rise in the money supply shifts the *LM* curve to the right, as shown in Figure 2. To see how this shift occurs, suppose that the *LM* curve is initially at LM_1 in panel (a) and the Federal Reserve conducts open market purchases that increase the money supply. If we consider point A, which is on the initial LM_1 curve, we can examine what happens to the equilibrium level of the interest rate, holding output constant at Y_A.

Panel (b), which contains a supply and demand diagram for the market for money, depicts the equilibrium interest rate initially as i_A at the intersection of the supply curve for money M_1^s and the demand curve for money M^d. The rise in the quantity of money supplied shifts the supply curve to M_2^s, and, holding output constant at Y_A, the equilibrium interest rate falls to $i_{A'}$. In panel (a), this decline in the equilibrium interest rate from i_A to $i_{A'}$ is shown as a movement from point A to point A′. The same analysis can be applied to every point on the initial LM_1 curve, leading to the conclusion that at any given level of aggregate output, the equilibrium interest rate falls when the money supply increases. Thus LM_2 is below and to the right of LM_1.

Reversing this reasoning, a decline in the money supply shifts the *LM* curve to the left. A decline in the money supply results in a shortage of money at points on the initial *LM* curve. This condition of excess demand for money can be eliminated by a rise in the interest rate, which reduces the quantity of money demanded until it again equals the quantity of money supplied.

2. *Autonomous Changes in Money Demand.* The theory of asset demand outlined in Chapter 5 indicates that there can be an autonomous rise in money demand (not caused by a change in the price level, aggregate output, or the interest rate). For example, an increase in the volatility of bond returns would make bonds riskier relative to money and would increase the quantity of money demanded at any given interest rate, price level, or amount of aggregate output. The resulting autonomous increase in the demand for money shifts the *LM* curve to the left, as shown in Figure 3. Consider point A on the initial LM_1 curve. Suppose that a massive financial panic occurs, sending many companies into bankruptcy. Because bonds have become a riskier asset, people want to shift from holding bonds to holding money; they will hold more money at all interest rates and output levels. The resulting increase in money demand at an output level of Y_A is shown by the shift of the money demand curve from M_1^d to M_2^d in panel (b). The new equilibrium in the market for money now indicates that if aggregate output is constant at Y_A, the equilibrium interest rate will rise to $i_{A'}$, and the point of equilibrium moves from A to A′.

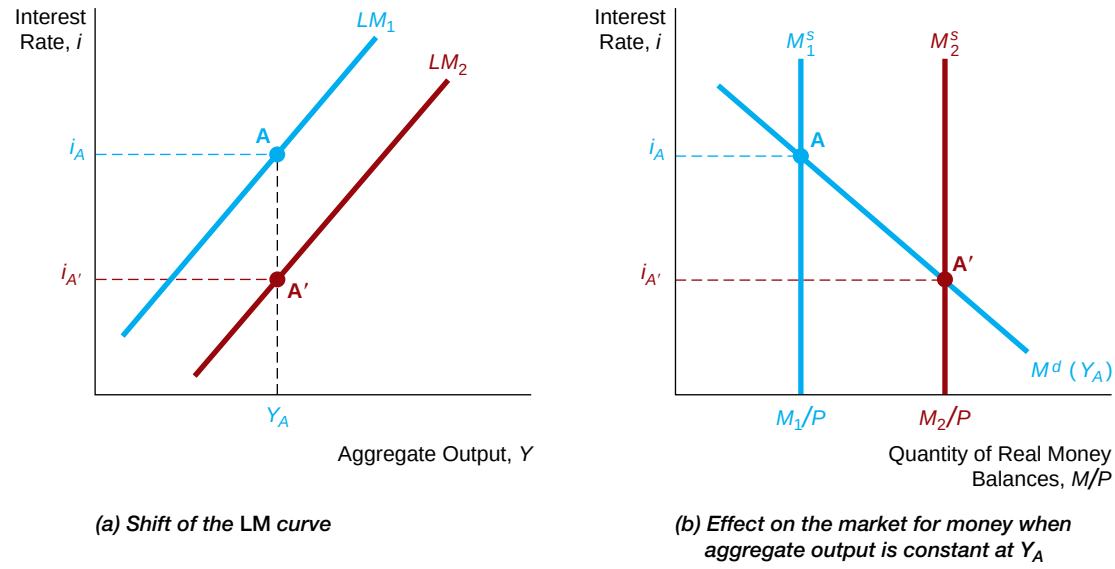


FIGURE 2 Shift in the *LM* Curve from an Increase in the Money Supply

The *LM* curve shifts to the right from LM_1 to LM_2 when the money supply increases because, as indicated in panel (b), at any given level of aggregate output (say, Y_A), the equilibrium interest rate falls (point A to A′).

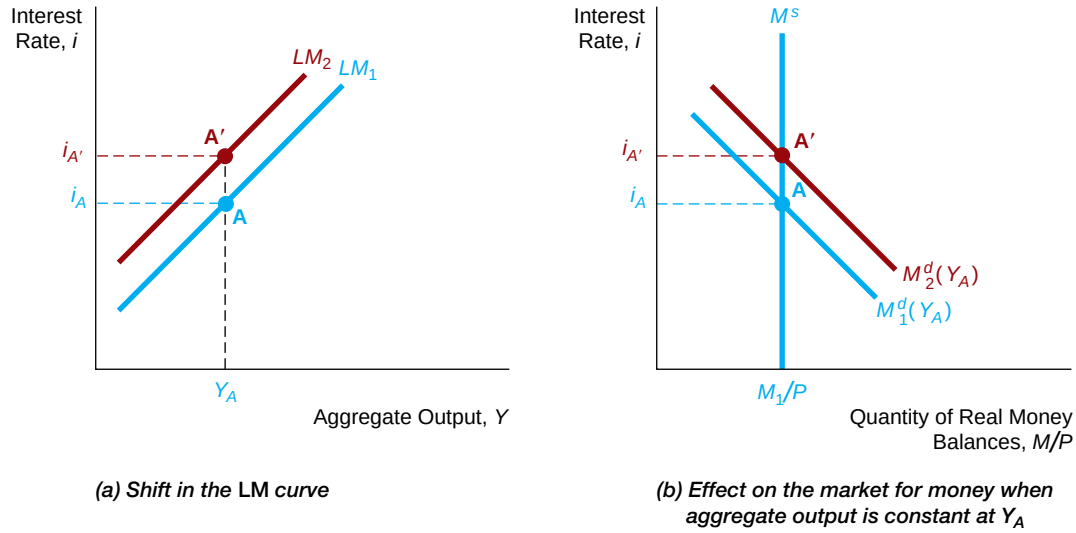


FIGURE 3 Shift in the *LM* Curve When Money Demand Increases

The *LM* curve shifts to the left from LM_1 to LM_2 when money demand increases because, as indicated in panel (b), at any given level of aggregate output (say, Y_A), the equilibrium interest rate rises (point A to A′).

Conversely, an autonomous decline in money demand would lead to a rightward shift in the *LM* curve. The fall in money demand would create an excess supply of money, which is eliminated by a rise in the quantity of money demanded from a decline in the interest rate.

Changes in Equilibrium Level of the Interest Rate and Aggregate Output

You can now use your knowledge of factors that cause the *IS* and *LM* curves to shift for the purpose of analyzing how the equilibrium levels of the interest rate and aggregate output change in response to changes in monetary and fiscal policies.

Response to a Change in Monetary Policy

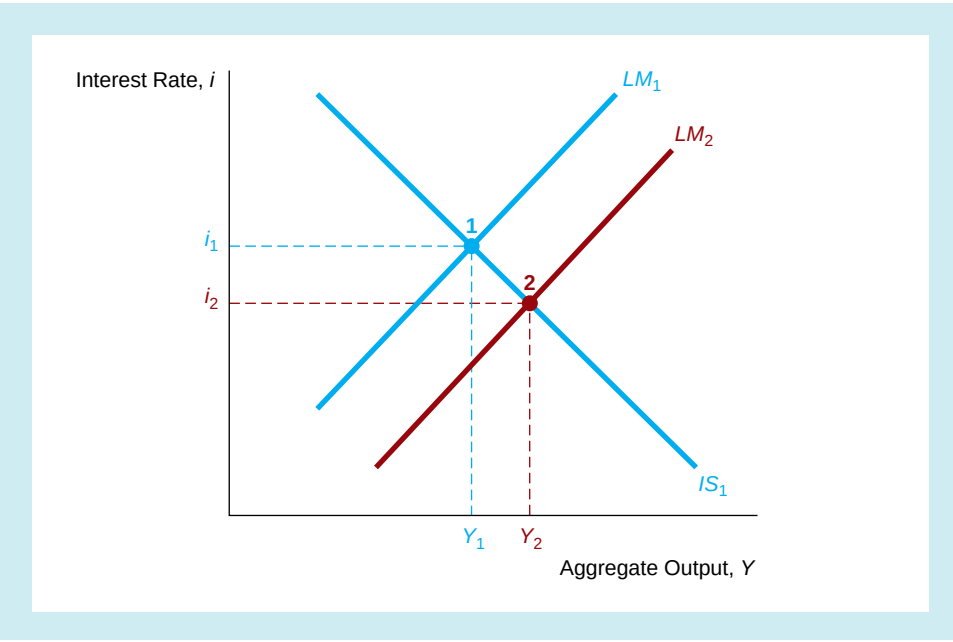
Figure 4 illustrates the response of output and interest rate to an increase in the money supply. Initially, the economy is in equilibrium for both the goods market and the market for money at point 1, the intersection of IS_1 and LM_1. Suppose that at the resulting level of aggregate output Y_1, the economy is suffering from an unemployment rate of 10%, and the Federal Reserve decides it should try to raise output and reduce unemployment by raising the money supply. Will the Fed's change in monetary policy have the intended effect?

The rise in the money supply causes the *LM* curve to shift rightward to LM_2, and the equilibrium point for both the goods market and the market for money moves to point 2 (intersection of IS_1 and LM_2). As a result of an increase in the money supply, the interest rate declines to i_2, as we found in Figure 2, and aggregate output rises to Y_2; the Fed's policy has been successful in improving the health of the economy.

For a clear understanding of why aggregate output rises and the interest rate declines, think about exactly what has happened in moving from point 1 to point 2.

FIGURE 4 Response of Aggregate Output and the Interest Rate to an Increase in the Money Supply

The increase in the money supply shifts the *LM* curve to the right from LM_1 to LM_2; the economy moves to point 2, where output has increased to Y_2 and the interest rate has declined to i_2.



When the economy is at point 1, the increase in the money supply (rightward shift of the *LM* curve) creates an excess supply of money, resulting in a decline in the interest rate. The decline causes investment spending and net exports to rise, which in turn raises aggregate demand and causes aggregate output to rise. The excess supply of money is eliminated when the economy reaches point 2 because both the rise in output and the fall in the interest rate have raised the quantity of money demanded until it equals the new higher level of the money supply.

A decline in the money supply reverses the process; it shifts the *LM* curve to the left, causing the interest rate to rise and output to fall. Accordingly, ***aggregate output is positively related to the money supply***; aggregate output expands when the money supply increases and falls when it decreases.

Response to a Change in Fiscal Policy

Suppose that the Federal Reserve is not willing to increase the money supply when the economy is suffering from a 10% unemployment rate at point 1. Can the federal government come to the rescue and manipulate government spending and taxes to raise aggregate output and reduce the massive unemployment?

The *ISLM* model demonstrates that it can. Figure 5 depicts the response of output and the interest rate to an expansionary fiscal policy (increase in government spending or decrease in taxes). An increase in government spending or a decrease in taxes causes the *IS* curve to shift to IS_2, and the equilibrium point for both the goods market and the market for money moves to point 2 (intersection of IS_2 with LM_1). The result of the change in fiscal policy is a rise in aggregate output to Y_2 and a rise in the interest rate to i_2. Note the difference in the effect on the interest rate between an expansionary fiscal policy and an expansionary monetary policy. In the case of an expansionary fiscal policy, the interest rate rises, whereas in the case of an expansionary monetary policy, the interest rate falls.

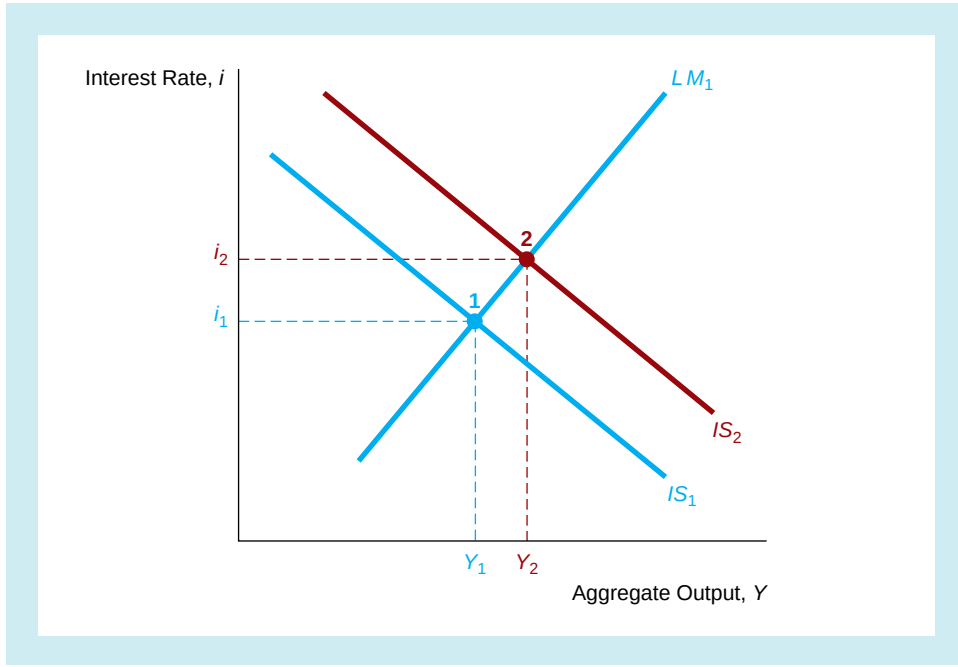


FIGURE 5 Response of Aggregate Output and the Interest Rate to an Expansionary Fiscal Policy
Expansionary fiscal policy (a rise in government spending or a decrease in taxes) shifts the *IS* curve to the right from IS_1 to IS_2; the economy moves to point 2, aggregate output increases to Y_2, and the interest rate rises to i_2.

Why does an increase in government spending or a decrease in taxes move the economy from point 1 to point 2, causing a rise in both aggregate output and the interest rate? An increase in government spending raises aggregate demand directly; a decrease in taxes makes more income available for spending and raises aggregate demand by raising consumer expenditure. The resulting increase in aggregate demand causes aggregate output to rise. The higher level of aggregate output raises the quantity of money demanded, creating an excess demand for money, which in turn causes the interest rate to rise. At point 2, the excess demand for money created by a rise in aggregate output has been eliminated by a rise in the interest rate, which lowers the quantity of money demanded.

A contractionary fiscal policy (decrease in government spending or increase in taxes) reverses the process described in Figure 5; it causes aggregate demand to fall, which shifts the *IS* curve to the left and causes both aggregate output and the interest rate to fall. ***Aggregate output and the interest rate are positively related to government spending and negatively related to taxes***.

Study Guide

As a study aid, Table 1 indicates the effect on aggregate output and interest rates of a change in the seven factors that shift the *IS* and *LM* curves. In addition, the table provides schematics describing the reason for the output and interest-rate response. *ISLM* analysis is best learned by practicing applications. To get this practice, you might try to develop the reasoning for your own Table 1 in which all the factors decrease rather than increase or answer Problems 5–7 and 13–15 at the end of this chapter.

Effectiveness of Monetary Versus Fiscal Policy

http://ingrimayne.saintjoe.edu/econ/optional/ISLM/Limitations.html

A paper discussing limitations of *ISLM* analysis, posted by the Federal Reserve.

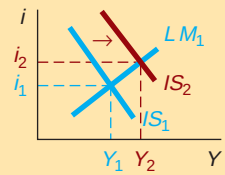
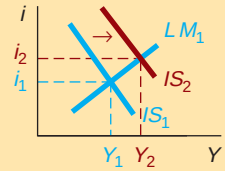
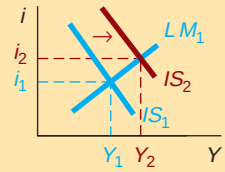
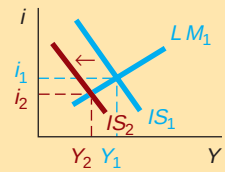
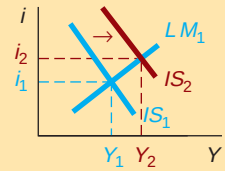
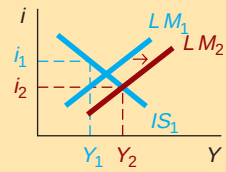
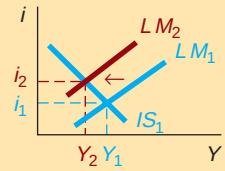
Our discussion of the effects of fiscal and monetary policy suggests that a government can easily lift an economy out of a recession by implementing any of a number of policies (changing the money supply, government spending, or taxes). But how can policymakers decide which of these policies to use if faced with too much unemployment? Should they decrease taxes, increase government spending, raise the money supply, or do all three? And if they decide to increase the money supply, by how much? Economists do not pretend to have all the answers, and although the *ISLM* model will not clear the path to aggregate economic bliss, it can help policymakers decide which policies may be most effective under certain circumstances.

Monetary Policy Versus Fiscal Policy: The Case of Complete Crowding Out

The *ISLM* model developed so far in this chapter shows that both monetary and fiscal policy affect the level of aggregate output. To understand when monetary policy is more effective than fiscal policy, we will examine a special case of the *ISLM* model in which money demand is unaffected by the interest rate (money demand is said to be interest-inelastic) so that monetary policy affects output but fiscal policy does not.

Consider the slope of the *LM* curve if the demand for money is unaffected by changes in the interest rate. If point 1 in panel (a) of Figure 6 is such that the quantity of money demanded equals the quantity of money supplied, then it is on the *LM* curve. If the interest rate rises to, say, i_2, the quantity of money demanded is unaffected, and it will continue to equal the *unchanged* quantity of money supplied only if

SUMMARY Table 1 Effects from Factors That Shift the *IS* and *LM* Curves

Factor	Autonomous Change in Factor	Response	Reason	
Consumer expenditure C	↑	$Y\uparrow, i\uparrow$	$C\uparrow \Rightarrow Y^{ad}\uparrow \Rightarrow$ *IS* shifts right	
Investment I	↑	$Y\uparrow, i\uparrow$	$I\uparrow \Rightarrow Y^{ad}\uparrow \Rightarrow$ *IS* shifts right	
Government spending G	↑	$Y\uparrow, i\uparrow$	$G\uparrow \Rightarrow Y^{ad}\uparrow \Rightarrow$ *IS* shifts right	
Taxes T	↑	$Y\downarrow, i\downarrow$	$T\uparrow \Rightarrow C\downarrow \Rightarrow Y^{ad}\downarrow \Rightarrow$ *IS* shifts left	
Net exports NX	↑	$Y\uparrow, i\uparrow$	$NX\uparrow \Rightarrow Y^{ad}\uparrow \Rightarrow$ *IS* shifts right	
Money supply M^s	↑	$Y\uparrow, i\downarrow$	$M^s\uparrow \Rightarrow i\downarrow \Rightarrow$ *LM* shifts right	
Money demand M^d	↑	$Y\downarrow, i\uparrow$	$M^d\uparrow \Rightarrow i\uparrow \Rightarrow$ *LM* shifts left	

Note: Only increases (↑) in the factors are shown. The effect of decreases in the factors would be the opposite of those indicated in the "Response" column.

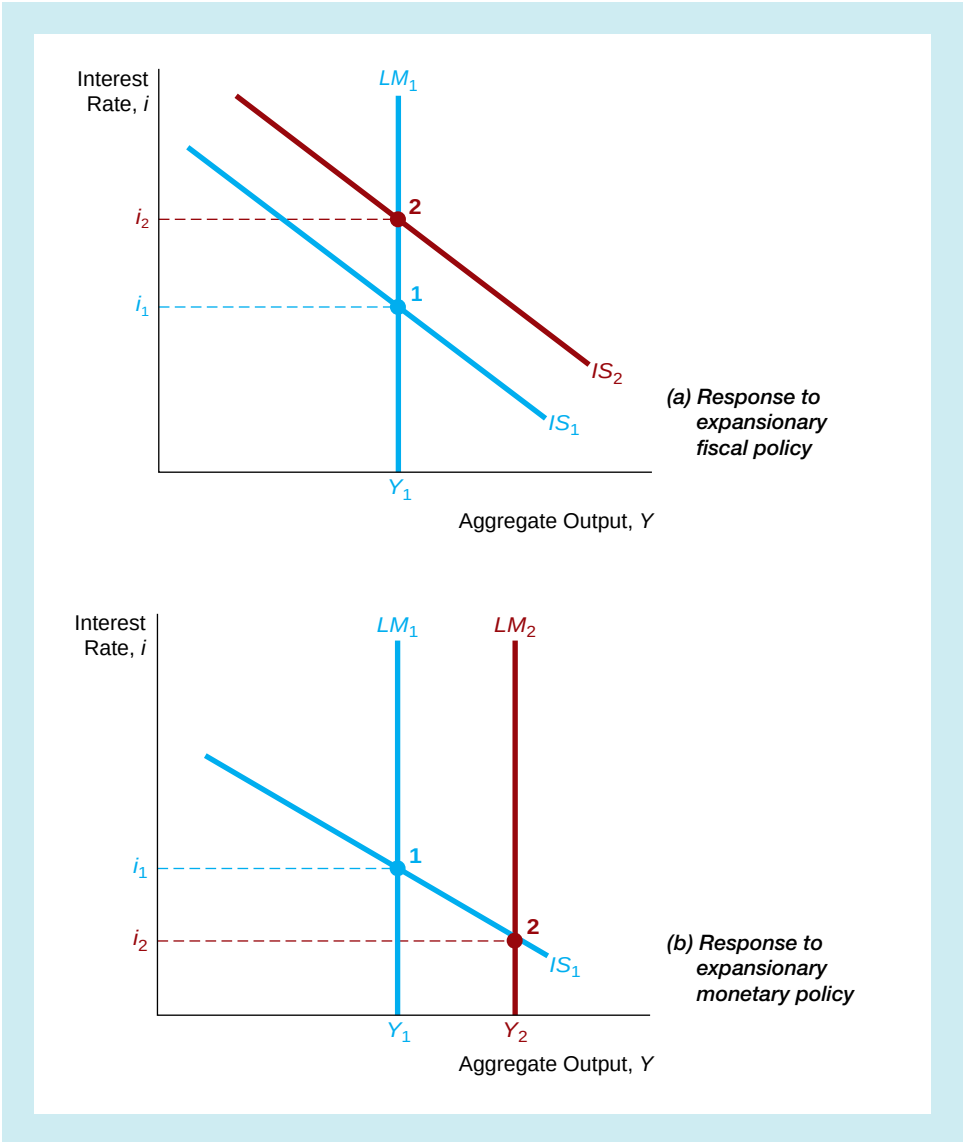


FIGURE 6 Effectiveness of Monetary and Fiscal Policy When Money Demand Is Unaffected by the Interest Rate

When the demand for money is unaffected by the interest rate, the *LM* curve is vertical. In panel (a), an expansionary fiscal policy (increase in government spending or a cut in taxes) shifts the *IS* curve from IS_1 to IS_2 and leaves aggregate output unchanged at Y_1. In panel (b), an increase in the money supply shifts the *LM* curve from LM_1 to LM_2 and raises aggregate output from Y_1 to Y_2. Therefore, monetary policy is effective, but fiscal policy is not.

www.bothell.washington.edu/faculty/danby/islm/animation.html

An animated explanation of *ISLM*.

aggregate output remains *unchanged* at Y_1 (point 2). Equilibrium in the market for money will occur at the same level of aggregate output regardless of the interest rate, and the *LM* curve will be vertical, as shown in both panels of Figure 6.

Suppose that the economy is suffering from a high rate of unemployment, which policymakers try to eliminate with either expansionary fiscal or monetary policy. Panel (a) depicts what happens when an expansionary fiscal policy (increase in government spending or cut in taxes) is implemented, shifting the *IS* curve to the right from IS_1 to IS_2. As you can see in panel (a), the fiscal expansion has no effect on output; aggregate output remains at Y_1 when the economy moves from point 1 to point 2.

In our earlier analysis, expansionary fiscal policy always increased aggregate demand and raised the level of output. Why doesn't that happen in panel (a)? The

answer is that because the *LM* curve is vertical, the rightward shift of the *IS* curve raises the interest rate to i_2, which causes investment spending and net exports to fall enough to offset completely the increased spending of the expansionary fiscal policy. Put another way, increased spending that results from expansionary fiscal policy has *crowded out* investment spending and net exports, which decrease because of the rise in the interest rate. This situation in which expansionary fiscal policy does not lead to a rise in output is frequently referred to as a case of **complete crowding out**.[1]

Panel (b) shows what happens when the Federal Reserve tries to eliminate high unemployment through an expansionary monetary policy (increase in the money supply). Here the *LM* curve shifts to the right from LM_1 to LM_2, because at each interest rate, output must rise so that the quantity of money demanded rises to match the increase in the money supply. Aggregate output rises from Y_1 to Y_2 (the economy moves from point 1 to point 2), and expansionary monetary policy does affect aggregate output in this case.

We conclude from the analysis in Figure 6 that if the demand for money is unaffected by changes in the interest rate (money demand is interest-inelastic), monetary policy is effective but fiscal policy is not. An even more general conclusion can be reached: ***The less interest-sensitive money demand is, the more effective monetary policy is relative to fiscal policy***.[2]

Because the interest sensitivity of money demand is important to policymakers' decisions regarding the use of monetary or fiscal policy to influence economic activity, the subject has been studied extensively by economists and has been the focus of many debates. Findings on the interest sensitivity of money demand are discussed in Chapter 22.

Application

Targeting Money Supply Versus Interest Rates

In the 1970s and early 1980s, central banks in many countries pursued a strategy of *monetary targeting*—that is, they used their policy tools to hit a money supply target (tried to make the money supply equal to a target value). However, as we saw in Chapter 18, many of these central banks abandoned monetary targeting in the 1980s to pursue interest-rate targeting instead, because of the breakdown of the stable relationship between the money supply and economic activity. The *ISLM* model has important implications for which variable a central bank should target and we can apply it

[1]When the demand for money is affected by the interest rate, the usual case in which the *LM* curve slopes upward but is not vertical, some crowding out occurs. The rightward shift of the *IS* curve also raises the interest rate, which causes investment spending and net exports to fall somewhat. However, as Figure 5 indicates, the rise in the interest rate is not sufficient to reduce investment spending and net exports to the point where aggregate output does not increase. Thus expansionary fiscal policy increases aggregate output, and only partial crowding out occurs.

[2]This result and many others in this and the previous chapter can be obtained more directly by using algebra. An algebraic treatment of the *ISLM* model can be found in an appendix to this chapter, which is on this book's web site at www.aw.com/mishkin.

to explain why central banks have abandoned monetary targeting for interest-rate targeting.[3]

As we saw in Chapter 18, when the Federal Reserve attempts to hit a money supply target, it cannot at the same time pursue an interest-rate target; it can hit one target or the other but not both. Consequently, it needs to know which of these two targets will produce more accurate control of aggregate output.

In contrast to the textbook world you have been inhabiting, in which the *IS* and *LM* curves are assumed to be fixed, the real world is one of great uncertainty in which *IS* and *LM* curves shift because of unanticipated changes in autonomous spending and money demand. To understand whether the Fed should use a money supply target or an interest-rate target, we need to look at two cases: first, one in which uncertainty about the *IS* curve is far greater than uncertainty about the *LM* curve and another in which uncertainty about the *LM* curve is far greater than uncertainty about the *IS* curve.

The *ISLM* diagram in Figure 7 illustrates the outcome of the two targeting strategies for the case in which the *IS* curve is unstable and uncertain, and so it fluctuates around its expected value of IS^* from IS' to IS'', while the *LM* curve is stable and certain, so it stays at LM^*. Since the central bank knows that the expected position of the *IS* curve is at IS^* and desires aggregate output of Y^*, it will set its interest-rate target at i^* so that the expected level of output is Y^*. This policy of targeting the interest rate at i^* is labeled "Interest-Rate Target."

How would the central bank keep the interest rate at its target level of i^*? Recall from Chapter 18 that the Fed can hit its interest-rate target by buying and selling bonds when the interest rate differs from i^*. When the *IS* curve shifts out to IS'', the interest rate would rise above i^* with the money supply unchanged. To counter this rise in interest rates, however, the central bank would need to buy bonds just until their price is driven back up so that the interest rate comes back down to i^*. (The result of these open market purchases, as we have seen in Chapters 15 and 16, is that the monetary base and the money supply rise until the *LM* curve shifts to the right to intersect the IS'' curve at i^*—not shown in the diagram for simplicity.) When the interest rate is below i^*, the central bank needs to sell bonds to lower their price and raise the interest rate back up to i^*. (These open market sales reduce the monetary base and the money supply until the *LM* curve shifts to the left to intersect the *IS* curve at IS'—again not shown in the diagram.) The result of pursuing the interest-rate target is that aggregate output fluctuates between Y'_1 and Y''_1 in Figure 7.

If, instead, the Fed pursues a money supply target, it will set the money supply so that the resulting *LM* curve LM^* intersects the IS^* curve at the desired output level of Y^*. This policy of targeting the money supply is

[3] The classic paper on this topic is William Poole, "The Optimal Choice of Monetary Policy Instruments in a Simple Macro Model," *Quarterly Journal of Economics* 84 (1970): 192–216. A less mathematical version of his analysis, far more accessible to students, is contained in William Poole, "Rules of Thumb for Guiding Monetary Policy," in *Open Market Policies and Operating Procedures: Staff Studies* (Washington, D.C.: Board of Governors of the Federal Reserve System, 1971).

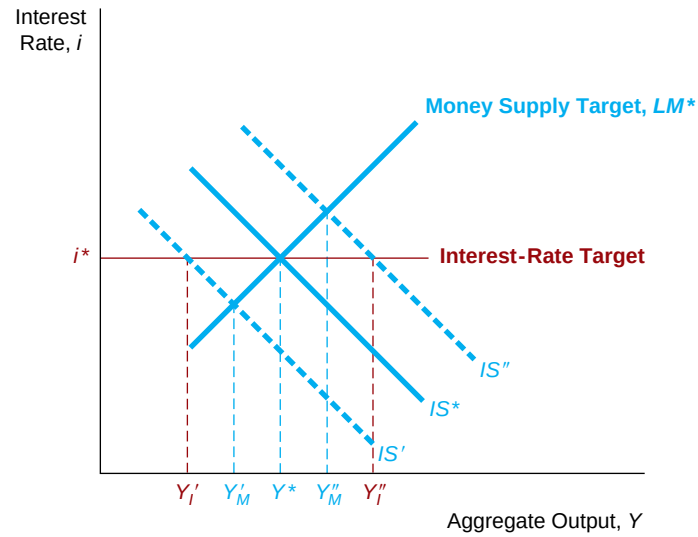


FIGURE 7 Money Supply and Interest-Rate Targets When the *IS* Curve Is Unstable and the *LM* Curve Is Stable

The unstable *IS* curve fluctuates between *IS*′ and *IS*″. The money supply target produces smaller fluctuations in output (Y_M' to Y_M'') than the interest rate targets (Y_I' to Y_I''). Therefore, the money supply target is preferred.

labeled "Money Supply Target." Because it is not changing the money supply and so keeps the *LM* curve at *LM**, aggregate output will fluctuate between Y_M' and Y_M'' for the money supply target policy.

As you can see in the figure, the money supply target leads to smaller output fluctuations around the desired level than the interest-rate target. A rightward shift of the *IS* curve to *IS*″, for example, causes the interest rate to rise, given a money supply target, and this rise in the interest rate leads to a lower level of investment spending and net exports and hence to a smaller increase in aggregate output than occurs under an interest-rate target. Because smaller output fluctuations are desirable, the conclusion is that ***if the* IS *curve is more unstable than the* LM *curve, a money supply target is preferred***.

The outcome of the two targeting strategies for the case of a stable *IS* curve and an unstable *LM* curve caused by unanticipated changes in money demand is illustrated in Figure 8. Again, the interest-rate and money supply targets are set so that the expected level of aggregate output equals the desired level *Y**. Because the *LM* curve is now unstable, it fluctuates between *LM*′ and *LM*″ even when the money supply is fixed, causing aggregate output to fluctuate between Y_M' and Y_M''.

The interest-rate target, by contrast, is not affected by uncertainty about the *LM* curve, because it is set by the Fed's adjusting the money supply whenever the interest rate tries to depart from *i**. When the interest rate begins to rise above *i** because of an increase in money demand, the central bank again just buys bonds, driving up their price and bringing the interest rate back down to *i**. The result of these open market purchases is a rise in the monetary base and the money supply. Similarly, if the interest rate falls below *i**, the central bank sells bonds to lower their price and raise the interest rate back to *i**, thereby causing a decline in the monetary base and the

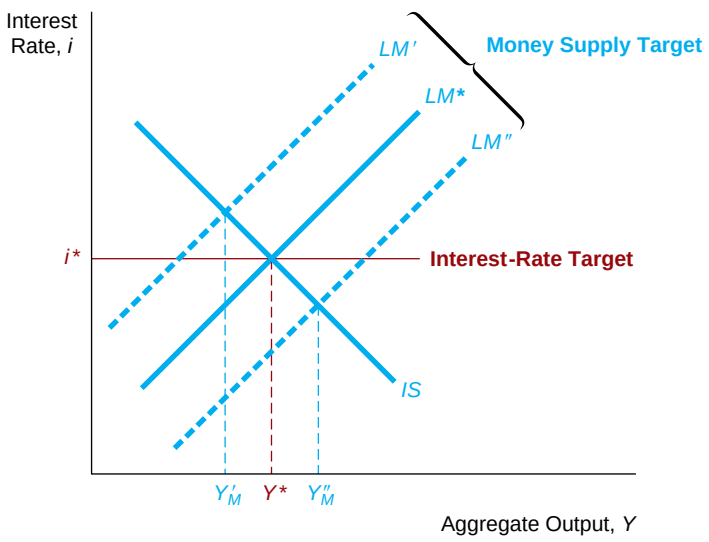


FIGURE 8 Money Supply and Interest-Rate Targets When the *LM* Curve Is Unstable and the *IS* Curve Is Stable

The unstable *LM* curve fluctuates between LM' and LM''. The money supply target then produces bigger fluctuations in output (Y'_M to Y''_M) than the interest-rate target (which leaves output fixed at Y^*). Therefore, the interest-rate target is preferred.

money supply. The only effect of the fluctuating *LM* curve, then, is that the money supply fluctuates more as a result of the interest-rate target policy. The outcome of the interest-rate target is that output will be exactly at the desired level with no fluctuations.

Since smaller output fluctuations are desirable, the conclusion from Figure 8 is that ***if the* LM *curve is more unstable than the* IS *curve, an interest-rate target is preferred***.

We can now see why many central banks decided to abandon monetary targeting for interest-rate targeting in the 1980s. With the rapid proliferation of new financial instruments whose presence can affect the demand for money (see Chapter 22), money demand (which is embodied in the *LM* curve) became highly unstable in many countries. Thus central banks in these countries recognized that they were more likely to be in the situation in Figure 8 and decided that they would be better off with an interest-rate target than a money supply target.[4]

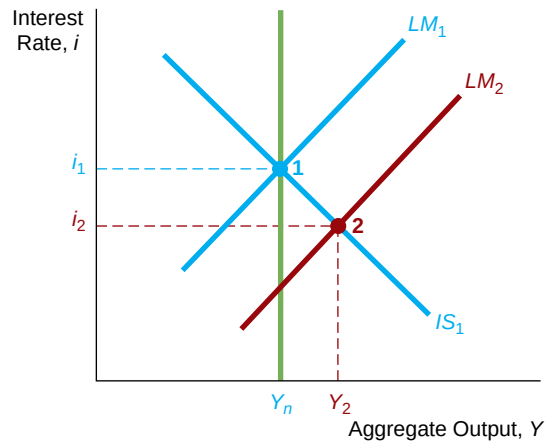
[4]It is important to recognize, however, that the crucial factor in deciding which target is preferred is the *relative* instability of the *IS* and *LM* curves. Although the *LM* curve has been unstable recently, the evidence supporting a stable *IS* curve is also weak. Instability in the money demand function does not automatically mean that money supply targets should be abandoned for an interest-rate target. Furthermore, the analysis so far has been conducted assuming that the price level is fixed. More realistically, when the price level can change, so that there is uncertainty about expected inflation, the case for an interest-rate target is less strong. As we learned in Chapters 4 and 5, the interest rate that is more relevant to investment decisions is not the nominal interest rate but the real interest rate (the nominal interest rate minus expected inflation). Hence when expected inflation rises, at each given nominal interest rate, the real interest rate falls and investment and net exports rise, shifting the *IS* curve to the right. Similarly, a fall in expected inflation raises the real interest rate at each given nominal interest rate, lowers investment and net exports, and shifts the *IS* curve to the left. Since in the real world, expected inflation undergoes large fluctuations, the *IS* curve in Figure 8 will also have substantial fluctuations, making it less likely that the interest-rate target is preferable to the money supply target.

ISLM Model in the Long Run

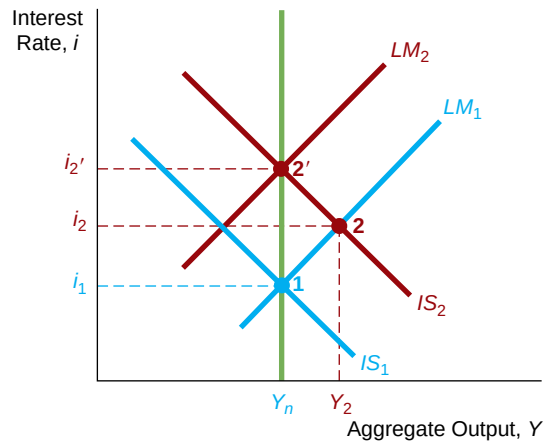
So far in our *ISLM* analysis, we have been assuming that the price level is fixed so that nominal values and real values are the same. This is a reasonable assumption for the short run, but in the long run the price level does change. To see what happens in the *ISLM* model in the long run, we make use of the concept of the **natural rate level of output** (denoted by Y_n), which is the rate of output at which the price level has no tendency to rise or fall. When output is above the natural rate level, the booming economy will cause prices to rise; when output is below the natural rate level, the slack in the economy will cause prices to fall.

Because we now want to examine what happens when the price level changes, we can no longer assume that real and nominal values are the same. The spending variables that affect the *IS* curve (consumer expenditure, investment spending, government spending, and net exports) describe the demand for goods and services and are *in real terms*; they describe the physical quantities of goods that people want to buy. Because these quantities do not change when the price level changes, a change in the price level has no effect on the *IS* curve, which describes the combinations of the interest rate and aggregate output *in real terms* that satisfy goods market equilibrium.

Figure 9 shows what happens in the *ISLM* model when output rises above the natural rate level, which is marked by a vertical line at Y_n. Suppose that initially the *IS* and *LM* curves intersect at point 1, where output $Y = Y_n$. Panel (a) examines what



(a) Response to a rise in the money supply M



(b) Response to a rise in government spending G

FIGURE 9 The *ISLM* Model in the Long Run

In panel (a), a rise in the money supply causes the *LM* curve to shift rightward to LM_2, and the equilibrium moves to point 2, where the interest rate falls to i_2 and output rises to Y_2. Because output at Y_2 is above the natural rate level Y_n, the price level rises, the real money supply falls, and the *LM* curve shifts back to LM_1; the economy has returned to the original equilibrium at point 1. In panel (b), an increase in government spending shifts the *IS* curve to the right to IS_2, and the economy moves to point 2, at which the interest rate has risen to i_2 and output has risen to Y_2. Because output at Y_2 is above the natural rate level Y_n, the price level begins to rise, real money balances M/P begin to fall, and the *LM* curve shifts to the left to LM_2. The long-run equilibrium at point 2′ has an even higher interest rate at i_2', and output has returned to Y_n.

happens to output and interest rates when there is a rise in the money supply. As we saw in Figure 2, the rise in the money supply causes the *LM* curve to shift to LM_2, and the equilibrium moves to point 2 (the intersection of IS_1 and LM_2), where the interest rate falls to i_2 and output rises to Y_2. However, as we can see in panel (a), the level of output at Y_2 is greater than the natural rate level Y_n, and so the price level begins to rise.

In contrast to the *IS* curve, which is unaffected by a rise in the price level, the *LM* curve is affected by the price level rise because the liquidity preference theory states that the demand for money *in real terms* depends on real income and interest rates. This makes sense because money is valued in terms of what it can buy. However, the money supply that you read about in newspapers is not the money supply in real terms; it is a nominal quantity. As the price level rises, the quantity of money *in real terms* falls, and the effect on the *LM* curve is identical to a fall in the nominal money supply with the price level fixed. The lower value of the real money supply creates an excess demand for money, causing the interest rate to rise at any given level of aggregate output, and the *LM* curve shifts back to the left. As long as the level of output exceeds the natural rate level, the price level will continue to rise, shifting the *LM* curve to the left, until finally output is back at the natural rate level Y_n. This occurs when the *LM* curve has returned to LM_1, where real money balances *M/P* have returned to the original level and the economy has returned to the original equilibrium at point 1. The result of the expansion in the money supply in the long run is that the economy has the same level of output and interest rates.

The fact that the increase in the money supply has left output and interest rates unchanged in the long run is referred to as **long-run monetary neutrality**. The only result of the increase in the money supply is a higher price level, which has increased proportionally to the increase in the money supply so that real money balances *M/P* are unchanged.

Panel (b) looks at what happens to output and interest rates when there is expansionary fiscal policy such as an increase in government spending. As we saw earlier, the increase in government spending shifts the *IS* curve to the right to IS_2, and in the short run the economy moves to point 2 (the intersection of IS_2 and LM_1), where the interest rate has risen to i_2 and output has risen to Y_2. Because output at Y_2 is above the natural rate level Y_n, the price level begins to rise, real money balances *M/P* begin to fall, and the *LM* curve shifts to the left. Only when the *LM* curve has shifted to LM_2 and the equilibrium is at point 2′, where output is again at the natural rate level Y_n, does the price level stop rising and the *LM* curve come to rest. The resulting long-run equilibrium at point 2′ has an even higher interest rate at $i_{2'}$ and output has not risen from Y_n. Indeed, what has occurred in the long run is complete crowding out: The rise in the price level, which has shifted the *LM* curve to LM_2, has caused the interest rate to rise to $i_{2'}$, causing investment and net exports to fall enough to offset the increased government spending completely. What we have discovered is that even though complete crowding out does not occur in the short run in the *ISLM* model (when the *LM* curve is not vertical), it does occur in the long run.

Our conclusion from examining what happens in the *ISLM* model from an expansionary monetary or fiscal policy is that ***although monetary and fiscal policy can affect output in the short run, neither affects output in the long run***. Clearly, an important issue in deciding on the effectiveness of monetary and fiscal policy to raise output is how soon the long run occurs. This is a topic that we explore in the next chapter.

ISLM Model and the Aggregate Demand Curve

We now examine further what happens in the *ISLM* model when the price level changes. When we conduct the *ISLM* analysis with a changing price level, we find that as the price level falls, the level of aggregate output rises. Thus we obtain a relationship between the price level and quantity of aggregate output for which the goods market and the market for money are in equilibrium, called the **aggregate demand curve**. This aggregate demand curve is a central element in the aggregate supply and demand analysis of Chapter 25, which allows us to explain changes not only in aggregate output but also in the price level.

Deriving the Aggregate Demand Curve

Now that you understand how a change in the price level affects the *LM* curve, we can analyze what happens in the *ISLM* diagram when the price level changes. This exercise is carried out in Figure 10. Panel (a) contains an *ISLM* diagram for a given value of the nominal money supply. Let us first consider a price level of P_1. The *LM* curve at this price level is $LM(P_1)$, and its intersection with the *IS* curve is at point 1, where output is Y_1. The equilibrium output level Y_1 that occurs when the price level is P_1 is also plotted in panel (b) as point 1. If the price level rises to P_2, then *in real terms* the money supply has fallen. The effect on the *LM* curve is identical to a decline in the nominal money supply when the price level is fixed: The *LM* curve will shift leftward to $LM(P_2)$. The new equilibrium level of output has fallen to Y_2, because planned investment and net exports fall when the interest rate rises. Point 2

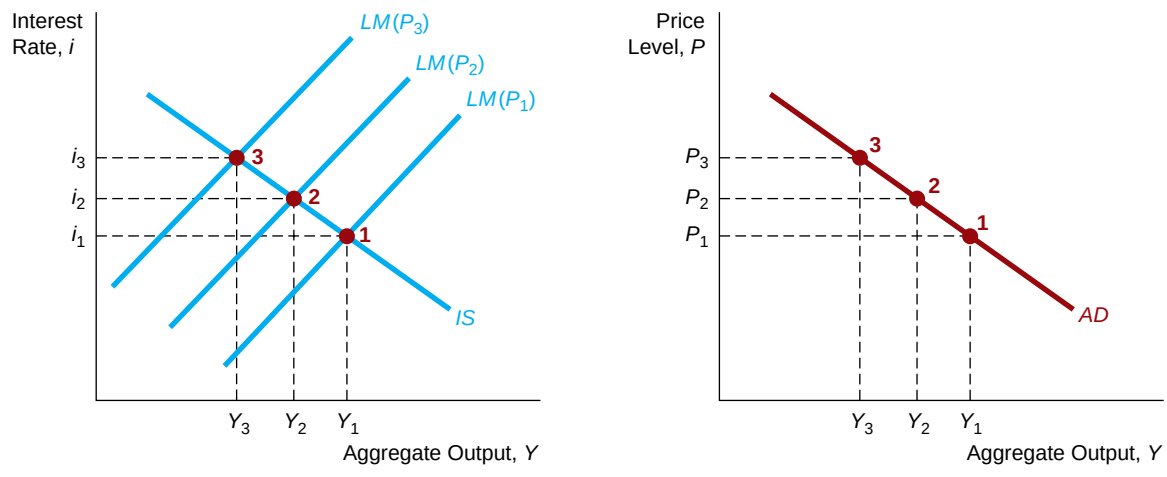


FIGURE 10 Deriving the Aggregate Demand Curve

The *ISLM* diagram in panel (a) shows that with a given nominal money supply as the price level rises from P_1 to P_2 to P_3, the *LM* curve shifts to the left, and equilibrium output falls. The combinations of the price level and equilibrium output from panel (a) are then plotted in panel (b), and the line connecting them is the aggregate demand curve *AD*.

in panel (b) plots this level of output for price level P_2. A further increase in the price level to P_3 causes a further decline in the real money supply, leading to a further increase in the interest rate and a further decline in planned investment and net exports, and output declines to Y_3. Point 3 in panel (b) plots this level of output for price level P_3.

The line that connects the three points in panel (b) is the aggregate demand curve *AD*, and it indicates the level of aggregate output consistent with equilibrium in the goods market and the market for money at any given price level. This aggregate demand curve has the usual downward slope, because a higher price level reduces the money supply in real terms, raises interest rates, and lowers the equilibrium level of aggregate output.

Factors That Cause the Aggregate Demand Curve to Shift

ISLM analysis demonstrates how the equilibrium level of aggregate output changes for a given price level. A change in any factor (except a change in the price level) that causes the *IS* or *LM* curve to shift causes the aggregate demand curve to shift. To see how this works, let's first look at what happens to the aggregate demand curve when the *IS* curve shifts.

www.worldbank.org.ru/wbimo/islmcl/islmcl.html

The World Bank has designed an animated *ISLM* model that lets you set various parameters and observe the results.

Shifts in the *IS* Curve. Five factors cause the *IS* curve to shift: changes in autonomous consumer spending, changes in investment spending related to business confidence, changes in government spending, changes in taxes, and autonomous changes in net exports. How changes in these factors lead to a shift in the aggregate demand curve is examined in Figure 11.

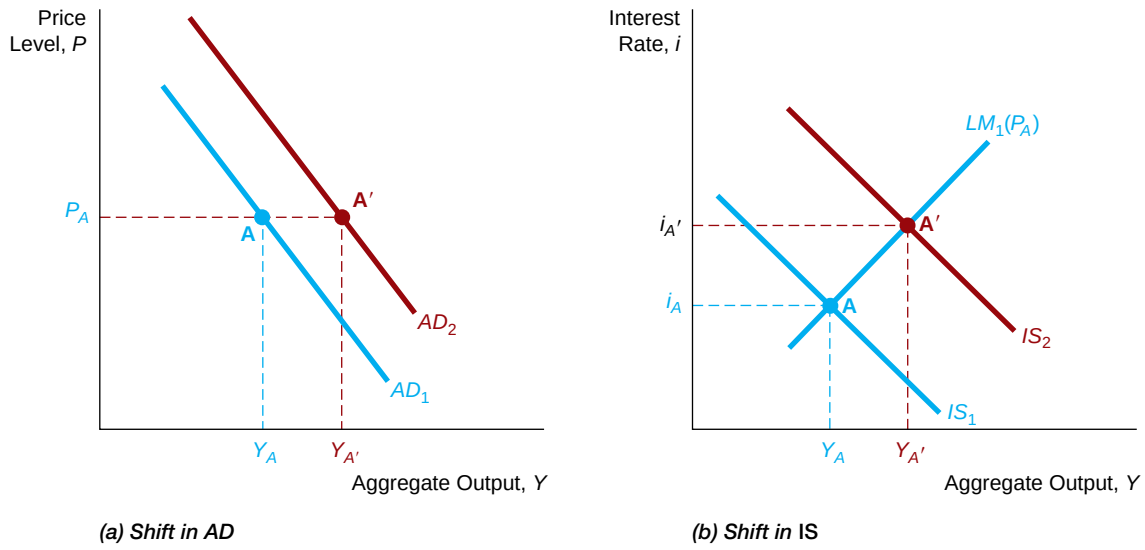


FIGURE 11 Shift in the Aggregate Demand Curve Caused by a Shift in the *IS* Curve

Expansionary fiscal policy, a rise in net exports, or more optimistic consumers and firms shift the *IS* curve to the right in panel (b), and at a price level of P_A, equilibrium output rises from Y_A to $Y_{A'}$. This change in equilibrium output is shown as a movement from point A to point A′ in panel (a); hence the aggregate demand curve shifts to the right, from AD_1 to AD_2.

Suppose that initially the aggregate demand curve is at AD_1 and there is a rise in, for example, government spending. The *ISLM* diagram in panel (b) shows what then happens to equilibrium output, holding the price level constant at P_A. Initially, equilibrium output is at Y_A at the intersection of IS_1 and LM_1. The rise in government spending (holding the price level constant at P_A) shifts the *IS* curve to the right and raises equilibrium output to $Y_{A'}$. In panel (a), this rise in equilibrium output is shown as a movement from point A to point A′, and the aggregate demand curve shifts to the right (to AD_2).

The conclusion from Figure 11 is that ***any factor that shifts the* IS *curve shifts the aggregate demand curve in the same direction***. Therefore, "animal spirits" that encourage a rise in autonomous consumer spending or planned investment spending, a rise in government spending, a fall in taxes, or an autonomous rise in net exports—all of which shift the *IS* curve to the right—will also shift the aggregate demand curve to the right. Conversely, a fall in autonomous consumer spending, a fall in planned investment spending, a fall in government spending, a rise in taxes, or a fall in net exports will cause the aggregate demand curve to shift to the left.

Shifts in the *LM* Curve. Shifts in the *LM* curve are caused by either an autonomous change in money demand (not caused by a change in *P*, *Y*, or *i*) or a change in the money supply. Figure 12 shows how either of these changes leads to a shift in the aggregate demand curve. Again, we are initially at the AD_1 aggregate demand curve, and we look at what happens to the level of equilibrium output when the price level is held

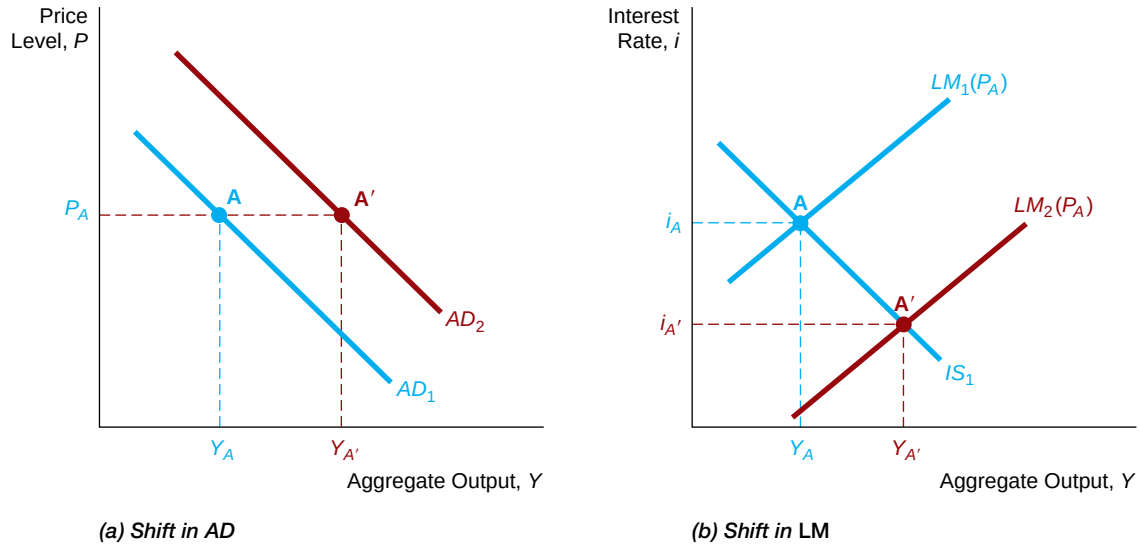


FIGURE 12 Shift in the Aggregate Demand Curve Caused by a Shift in the *LM* Curve

A rise in the money supply or a fall in money demand shifts the *LM* curve to the right in panel (b), and at a price level of P_A, equilibrium output rises from Y_A to $Y_{A'}$. This change in equilibrium output is shown as a movement from point A to point A′ in panel (a); hence the aggregate demand curve shifts to the right, from AD_1 to AD_2.

constant at P_A. A rise in the money supply shifts the *LM* curve to the right and raises equilibrium output to $Y_{A'}$. This rise in equilibrium output is shown as a movement from point A to point A′ in panel (a), and the aggregate demand curve shifts to the right.

Our conclusion from Figure 12 is similar to that of Figure 11: ***Holding the price level constant, any factor that shifts the* LM *curve shifts the aggregate demand curve in the same direction***. Therefore, a decline in money demand as well as an increase in the money supply, both of which shift the *LM* curve to the right, also shift the aggregate demand curve to the right. The aggregate demand curve will shift to the left, however, if the money supply declines or money demand rises.

You have now derived and analyzed the aggregate demand curve—an essential element in the aggregate demand and supply framework that we examine in Chapter 25. The aggregate demand and supply framework is particularly useful, because it demonstrates how the price level is determined and enables us to examine factors that affect aggregate output when the price level varies.

Summary

1. The *IS* curve is shifted to the right by a rise in autonomous consumer spending, a rise in planned investment spending related to business confidence, a rise in government spending, a fall in taxes, or an autonomous rise in net exports. A movement in the opposite direction of these five factors will shift the *IS* curve to the left.
2. The *LM* curve is shifted to the right by a rise in the money supply or an autonomous fall in money demand; it is shifted to the left by a fall in the money supply or an autonomous rise in money demand.
3. A rise in the money supply raises equilibrium output, but lowers the equilibrium interest rate. Expansionary fiscal policy (a rise in government spending or a fall in taxes) raises equilibrium output, but, in contrast to expansionary monetary policy, also raises the interest rate.
4. The less interest-sensitive money demand is, the more effective monetary policy is relative to fiscal policy.
5. The *ISLM* model provides the following conclusion about the conduct of monetary policy: When the *IS* curve is more unstable than the *LM* curve, pursuing a money supply target provides smaller output fluctuations than pursuing an interest-rate target and is preferred; when the *LM* curve is more unstable than the *IS* curve, pursuing an interest-rate target leads to smaller output fluctuations and is preferred.
6. The conclusion from examining what happens in the *ISLM* model from an expansionary monetary or fiscal policy is that although monetary and fiscal policy can affect output in the short run, neither affects output in the long run.
7. The aggregate demand curve tells us the level of aggregate output consistent with equilibrium in the goods market and the market for money for any given price level. It slopes downward because a lower price level creates a higher level of the real money supply, lowers the interest rate, and raises equilibrium output. The aggregate demand curve shifts in the same direction as a shift in the *IS* or *LM* curve; hence it shifts to the right when government spending increases, taxes decrease, "animal spirits" encourage consumer and business spending, autonomous net exports increase, the money supply increases, or money demand decreases.

Key Terms

aggregate demand curve, p. 577
complete crowding out, p. 571
long-run monetary neutrality, p. 576
natural rate level of output, p. 575

Questions and Problems

Questions marked with an asterisk are answered at the end of the book in an appendix, "Answers to Selected Questions and Problems."

1. If taxes and government spending rise by equal amounts, what will happen to the position of the *IS* curve? Explain this with a Keynesian cross diagram.

*2. What happened to the *IS* curve during the Great Depression when investment spending collapsed? Why?

3. What happens to the position of the *LM* curve if the Fed decides that it will decrease the money supply to fight inflation and if, at the same time, the demand for money falls?

*4. "An excess demand for money resulting from a rise in the demand for money can be eliminated only by a rise in the interest rate." Is this statement true, false, or uncertain? Explain your answer.

In Problems 5–15, demonstrate your answers with an *ISLM* diagram.

5. In late 1969, the Federal Reserve reduced the money supply while the government raised taxes. What do you think should have happened to interest rates and aggregate output?

*6. "The high level of interest rates and the rapidly growing economy during Ronald Reagan's third and fourth years as president can be explained by a tight monetary policy combined with an expansionary fiscal policy." Do you agree? Why or why not?

7. Suppose that the Federal Reserve wants to keep interest rates from rising when the government sharply increases military spending. How can the Fed do this?

*8. Evidence indicates that lately the demand for money has become quite unstable. Why is this finding important to Federal Reserve policymakers?

9. "As the price level rises, the equilibrium level of output determined in the *ISLM* model also rises." Is this statement true, false, or uncertain? Explain your answer.

*10. What will happen to the position of the aggregate demand curve if the money supply is reduced when government spending increases?

11. An equal rise in government spending and taxes will have what effect on the position of the aggregate demand curve?

*12. If money demand is unaffected by changes in the interest rate, what effect will a rise in government spending have on the position of the aggregate demand curve?

Using Economic Analysis to Predict the Future

13. Predict what will happen to interest rates and output if a stock market crash causes autonomous consumer expenditure to fall.

*14. Predict what will happen to interest rates and aggregate output when there is an autonomous export boom.

15. If a series of defaults in the bond market make bonds riskier and as a result the demand for money rises, predict what will happen to interest rates and aggregate output.

Web Exercises

1. We can continue our study of the ISLM framework by reviewing a dynamic interactive site. Go to http://nova.umuc.edu/~black/econ0.html. Assume that the change in government spending is $25, the tax rate is 30%, the velocity of money is 12, and the money supply is increased by $2. What is the resulting change in interest rates? (Be sure to check the button above ISLM.)

2. An excellent way to learn about how changes in various factors affect the *IS* and *LM* curves is to visit www.worldbank.org.ru/wbimo/islmcl/islmcl.html. This site, sponsored by the World Bank, allows you to make changes and to observe immediately their impact on the *ISLM* model.
   a. Increase G from 1,200 to 1,400. What happens to the interest rate?
   b. Reduce T_0 to .08. What happens to aggregate output *Y*?
   c. Increase the *M* to 1,100. What happens to the interest rate and aggregate output?

appendix to chapter 24

Algebra of the *ISLM* Model

The use of algebra to analyze the *ISLM* model allows us to extend the multiplier analysis in Chapter 23 and to obtain many of the results of Chapters 23 and 24 very quickly.

Basic Closed-Economy *ISLM* Model

The goods market can be described by the following equations:

Consumption function:	$C = \overline{C} + mpc\ (Y - T)$	(1)
Investment function:	$I = \overline{I} - di$	(2)
Taxes:	$T = \overline{T}$	(3)
Government spending:	$G = \overline{G}$	(4)
Goods market equilibrium condition:	$Y = Y^{ad} = C + I + G$	(5)

The money market is described by these equations:

Money demand function:	$M^d = \overline{M}^d + eY - fi$	(6)
Money supply:	$M^s = \overline{M}$	(7)
Money market equilibrium condition:	$M^d = M^s$	(8)

The uppercase terms are the variables of the model; $\overline{G}$, $\overline{T}$, and $\overline{M}$, are the values of the policy variables that are set exogenously (outside the model); and $\overline{C}$, $\overline{I}$, and $\overline{M}^d$ are autonomous components of consumer expenditure, investment spending, and money demand that are also determined exogenously (outside the model). Except for the interest rate i, the lowercase terms are the parameters, the givens of the model, and all are assumed to be positive. The definitions of these variables and parameters are as follows:

C = consumer spending
I = investment spending
$G = \overline{G}$ = government spending
Y = output
$T = \overline{T}$ = taxes
M^d = money demand
$M^s = \overline{M}$ = money supply
i = interest rate
$\overline{C}$ = autonomous consumer spending

d = interest sensitivity of investment spending
$\bar{I}$ = autonomous investment spending related to business confidence
$\bar{M}^d$ = autonomous money demand
e = income sensitivity of money demand
f = interest sensitivity of money demand
mpc = marginal propensity to consume

IS and *LM* Curves

Substituting for C, I, and G in the goods market equilibrium condition and then solving for Y, we obtain the *IS* curve:

$$Y = \frac{1}{1 - mpc}(\bar{C} + \bar{I} - mpc\,\bar{T} + \bar{G} - di) \tag{9}$$

Solving for i from Equations 6, 7, and 8, we obtain the *LM* curve:

$$i = \frac{\bar{M}^d - \bar{M} + eY}{f} \tag{10}$$

Solution of the Model

The solution to the model occurs at the intersection of the *IS* and *LM* curves, which involves solving for Y and i simultaneously, using Equations 9 and 10, as follows:

$$Y = \frac{1}{1 - mpc + de/f}\left(\bar{C} + \bar{I} - mpc\,\bar{T} + \bar{G} - \frac{d\bar{M}^d}{f} + \frac{d\bar{M}}{f}\right) \tag{11}$$

$$i = \frac{1}{f(1 - mpc) + d}[e(\bar{C} + \bar{I} - mpc\,\bar{T} + \bar{G}) + \bar{M}^d(1 - mpc) - \bar{M}(1 - mpc)] \tag{12}$$

Implications

The conclusions reached with these algebraic solutions are the same as those reached in Chapters 23 and 24; for example:

1. Because all the coefficients are positive, Equation 11 indicates that a rise in $\bar{C}$, $\bar{I}$, $\bar{G}$, and $\bar{M}$ leads to a rise in Y and that a rise in $\bar{T}$ or $\bar{M}^d$ leads to a fall in Y.
2. Equation 12 indicates that a rise in $\bar{C}$, $\bar{I}$, $\bar{G}$, and $\bar{M}^d$ leads to a rise in i and that a rise in $\bar{M}$ or $\bar{T}$ leads to a fall in i.
3. As f, the interest sensitivity of money demand, increases, the multiplier term:

$$\frac{1}{1 - mpc + de/f}$$

increases, and so fiscal policy ($\bar{G}$, $\bar{T}$) has more effect on output; conversely, the term multiplying $\bar{M}$,

$$\frac{d}{f}\left(\frac{1}{1 - mpc + de/f}\right) = \frac{d}{f(1 - mpc) + de}$$

declines, so monetary policy has less effect on output.
4. By similar reasoning, as d, the interest sensitivity of investment spending, increases, monetary policy has more effect on output and fiscal policy has less effect on output.

Open-Economy *ISLM* Model

To make the basic *ISLM* model into an open-economy model, we need to include net exports in the goods market equilibrium condition so that Equation 5 becomes Equation 5':

$$Y = Y^{ad} = C + I + G + NX \tag{5'}$$

As the discussion in Chapter 24 suggests, the net exports and exchange rate relations can be written:

$$NX = \overline{NX} - hE \tag{13}$$

$$E = \overline{E} + ji \tag{14}$$

where NX = net exports
$\overline{NX}$ = autonomous net exports
h = exchange rate sensitivity of net exports
E = exchange rate (value of domestic currency)
$\overline{E}$ = autonomous exchange rate
j = interest sensitivity of exchange rate

Substituting for net exports in the goods market equilibrium condition (Equation 5') using the net exports and exchange rate relations and then solving for Y as in the basic model, we obtain the open-economy *IS* curve:

$$Y = \frac{1}{1 - mpc}[\overline{C} + \overline{I} - mpc\,\overline{T} + \overline{G} + \overline{NX} - h\overline{E} - (d + hj)i] \tag{15}$$

The *LM* curve is the same as in the basic model, and so the solutions for Y and i are as follows:

$$Y = \frac{1}{1 - mpc + (d + hj)e/f}$$

$$\times \left(\overline{C} + \overline{I} - mpc\,\overline{T} + \overline{G} - \frac{d + hj}{f}\overline{M}^d + \frac{d + hj}{f}\overline{M} + \overline{NX} - h\overline{E}\right) \tag{16}$$

$$i = \frac{1}{f(1 - mpc) + (d + hj)e}$$

$$\times [e(\overline{C} + \overline{I} - mpc\,\overline{T} + \overline{G} + \overline{NX} - h\overline{E}) + \overline{M}^d(1 - mpc) - \overline{M}(1 - mpc)] \tag{17}$$

Implications

1. As the *IS* curve in Equation 15 indicates, including net exports in aggregate demand provides an additional reason for the negative relationship between Y and i (the downward slope of the *IS* curve). This additional reason for the negative relationship of Y and i is represented by hj in the term $-(d + hj)i$.

2. Equations 16 and 17 indicate that all the results we found for the basic model still hold.
3. Equation 16 indicates that a rise in $\overline{NX}$ leads to a rise in Y, and an autonomous rise in the value of the domestic currency $\overline{E}$ leads to a decline in Y.
4. Equation 17 indicates that a rise in $\overline{NX}$ leads to a rise in i, and a rise in $\overline{E}$ leads to a decline in i.

Chapter 25

Aggregate Demand and Supply Analysis

PREVIEW

In earlier chapters, we focused considerable attention on monetary policy, because it touches our everyday lives by affecting the prices of the goods we buy and the quantity of available jobs. In this chapter, we develop a basic tool, aggregate demand and supply analysis, that will enable us to study the effects of money on output and prices. **Aggregate demand** is the total quantity of an economy's final goods and services demanded at different price levels. **Aggregate supply** is the total quantity of final goods and services that firms in the economy want to sell at different price levels. As with other supply and demand analyses, the actual quantity of output and the price level are determined by equating aggregate demand and aggregate supply.

Aggregate demand and supply analysis will enable us to explore how aggregate output and the price level are determined. (The "Following the Financial News" box indicates when data on aggregate output and the price level are published.) Not only will the analysis help us interpret recent episodes in the business cycle, but it will also enable us to understand the debates on how economic policy should be conducted.

Aggregate Demand

The first building block of aggregate supply and demand analysis is the **aggregate demand curve**, which describes the relationship between the quantity of aggregate output demanded and the price level when all other variables are held constant. **Monetarists** (led by Milton Friedman) view the aggregate demand curve as downward-sloping with one primary factor that causes it to shift—changes in the quantity of money. **Keynesians** (followers of Keynes) also view the aggregate demand curve as downward-sloping, but they believe that changes in government spending and taxes or in consumer and business willingness to spend can also cause it to shift.

Monetarist View of Aggregate Demand

The monetarist view of aggregate demand links the quantity of money M with total nominal spending on goods and services $P \times Y$ (P = price level and Y = aggregate real output or, equivalently, aggregate real income). To do this it uses the concept of the **velocity of money**: the average number of times per year that a dollar is spent on

Following the Financial News

Aggregate Output, Unemployment, and the Price Level

Newspapers and Internet sites periodically report data that provide information on the level of aggregate output, unemployment, and the price level. Here is a list of the relevant data series, their frequency, and when they are published.

Aggregate Output and Unemployment

Real GDP: Quarterly (January–March, April–June, July–September, October–December); published three to four weeks after the end of a quarter.

Industrial production: Monthly. Industrial production is not as comprehensive a measure of aggregate output as real GDP, because it measures only manufacturing output; the estimate for the previous month is reported in the middle of the following month.

Unemployment rate: Monthly; previous month's figure is usually published on the Friday of the first week of the following month.

Price Level

GDP deflator: Quarterly. This comprehensive measure of the price level (described in the appendix to Chapter 1) is published at the same time as the real GDP data.

Consumer price index (CPI): Monthly. The CPI is a measure of the price level for consumers (also described in the appendix to Chapter 1); the value for the previous month is published in the third or fourth week of the following month.

Producer price index (PPI): Monthly. The PPI is a measure of the average level of wholesale prices charged by producers and is published at the same time as industrial production data.

www.bls.gov/data/home.htm
The home page of the Bureau of Labor Statistics lists information on unemployment and price levels.

final goods and services. More formally, velocity V is calculated by dividing nominal spending $P \times Y$ by the money supply M:

$$V = \frac{P \times Y}{M}$$

Suppose that the total nominal spending in a year was \$2 trillion and the money supply was \$1 trillion; velocity would then be \$2 trillion/\$1 trillion = 2. On average, the money supply supports a level of transactions associated with 2 times its value in final goods and services in the course of a year. By multiplying both sides by M, we obtain the **equation of exchange**, which relates the money supply to aggregate spending:

$$M \times V = P \times Y \tag{1}$$

At this point, the equation of exchange is nothing more than an identity; that is, it is true by definition. It does not tell us that when M rises, aggregate spending will rise as well. For example, the rise in M could be offset by a fall in V, with the result that $M \times V$ does not rise. However, Friedman's analysis of the demand for money (discussed in detail in Chapter 22) suggests that velocity varies over time in a predictable manner unrelated to changes in the money supply. With this analysis, the equation of

exchange is transformed into a theory of how aggregate spending is determined and is called the **modern quantity theory of money**.

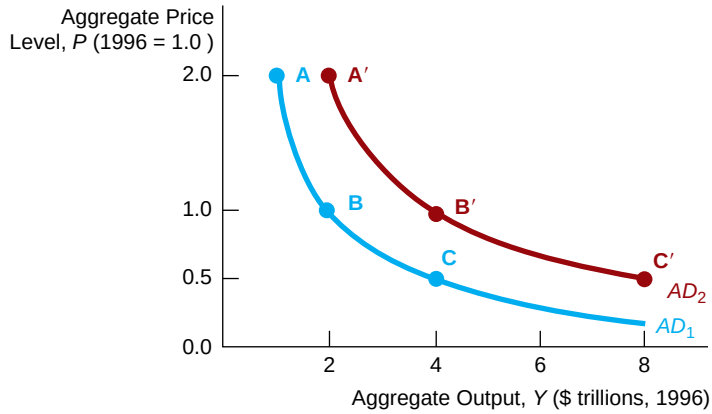
To see how the theory works, let's look at an example. If velocity is predicted to be 2 and the money supply is $1 trillion, the equation of exchange tells us that aggregate spending will be $2 trillion (2 × $1 trillion). If the money supply doubles to $2 trillion, Friedman's analysis suggests that velocity will continue to be 2 and aggregate spending will double to $4 trillion (2 × $2 trillion). Thus Friedman's modern quantity theory of money concludes that ***changes in aggregate spending are determined primarily by changes in the money supply***.

Deriving the Aggregate Demand Curve. To learn how the modern quantity theory of money generates the aggregate demand curve, let's look at an example in which we measure aggregate output in trillions of 1996 dollars, with the price level in 1996 having a value of 1.0. As just shown, with a predicted velocity of 2 and a money supply of $1 trillion, aggregate spending will be $2 trillion. If the price level is given at 2.0, the quantity of aggregate output demanded is $1 trillion because aggregate spending $P \times Y$ then continues to equal 2.0 × $1 trillion = $2 trillion, the value of $M \times V$. This combination of a price level of 2.0 and aggregate output of 1 is marked as point A in Figure 1. If the price level is given as 1.0 instead, aggregate output demanded is $2 trillion (point B), so aggregate spending continues to equal $2 trillion (= 1.0 × 2 trillion). Similarly, at an even lower price level of 0.5, the quantity of output demanded rises to $4 trillion, shown by point C. The curve connecting these points, marked AD_1, is the aggregate demand curve, given a money supply of $1 trillion. As you can see, it has the usual downward slope of a demand curve, indicating that as the price level falls (everything else held constant), the quantity of output demanded rises.

Shifts in the Aggregate Demand Curve. In Friedman's modern quantity theory, changes in the money supply are the primary source of the changes in aggregate spending and shifts in the aggregate demand curve. To see how a change in the money supply shifts the aggregate demand curve in Figure 1, let's look at what happens when the money supply increases to $2 trillion. Now aggregate spending rises to 2 × $2 trillion = $4 trillion,

FIGURE 1 Aggregate Demand Curve

An aggregate demand curve is drawn for a *fixed* level of the money supply. A rise in the money supply from $1 trillion to $2 trillion leads to a shift in the aggregate demand curve from AD_1 to AD_2.



and at a price level of 2.0, the quantity of aggregate output demanded will rise to \$2 trillion so that 2.0 × 2 trillion = \$4 trillion. Therefore, at a price level of 2.0, the aggregate demand curve moves from point A to A′. At a price level of 1.0, the quantity of output demanded rises from \$2 to \$4 trillion (from point B to B′), and at a price level of 0.5, output demanded rises from \$4 to \$8 trillion (from point C to C′). The result is that the rise in the money supply to \$2 trillion shifts the aggregate demand curve outward to AD_2.

Similar reasoning indicates that a decline in the money supply lowers aggregate spending proportionally and reduces the quantity of aggregate output demanded at each price level. Thus a decline in the money supply shifts the aggregate demand curve to the left.

Keynesian View of Aggregate Demand

Rather than determining aggregate demand from the equation of exchange, Keynesians analyze aggregate demand in terms of its four component parts: **consumer expenditure**, the total demand for consumer goods and services; **planned investment spending**,[1] the total planned spending by business firms on new machines, factories, and other inputs to production, plus planned spending on new homes; **government spending**, spending by all levels of government (federal, state, and local) on goods and services (paper clips, computers, computer programming, missiles, government workers, and so on); and **net exports**, the net foreign spending on domestic goods and services, equal to exports minus imports. Using the symbols C for consumer expenditure, I for planned investment spending, G for government spending, and NX for net exports, we can write the following expression for aggregate demand Y^{ad}:

$$Y^{ad} = C + I + G + NX \tag{2}$$

Deriving the Aggregate Demand Curve. Keynesian analysis, like monetarist analysis, suggests that the aggregate demand curve is downward-sloping because a lower price level ($P\downarrow$), holding the nominal quantity of money (M) constant, leads to a larger quantity of money in real terms (in terms of the goods and services that it can buy, $M/P\uparrow$). The larger quantity of money in real terms ($M/P\uparrow$) that results from the lower price level causes interest rates to fall ($i\downarrow$), as suggested in Chapter 5 and 24. The resulting lower cost of financing purchases of new physical capital makes investment more profitable and stimulates planned investment spending ($I\uparrow$). Because, as shown in Equation 2, the increase in planned investment spending adds directly to aggregate demand ($Y^{ad}\uparrow$), the lower price level leads to a higher level of aggregate demand ($P\downarrow \Rightarrow Y^{ad}\uparrow$). Schematically, we can write the mechanism just described as follows:

$$P\downarrow \Rightarrow M/P\uparrow \Rightarrow i\downarrow \Rightarrow I\uparrow \Rightarrow Y^{ad}\uparrow$$

Another mechanism that generates a downward-sloping aggregate demand curve operates through international trade. Because a lower price level ($P\downarrow$) leads to a larger quantity of money in real terms ($M/P\uparrow$) and lower interest rates ($i\downarrow$), U.S. dollar bank deposits become less attractive relative to deposits denominated in foreign currencies, thereby causing a fall in the value of dollar deposits relative to other currency deposits

[1] Recall that economists restrict use of the word *investment* to the purchase of new physical capital, such as a new machine or a new house, that adds to expenditure.

(a decline in the exchange rate, denoted by $E\downarrow$). The lower value of the dollar, which makes domestic goods cheaper relative to foreign goods, then causes net exports to rise, which in turn increases aggregate demand:

$$P\downarrow \Rightarrow M/P\uparrow \Rightarrow i\downarrow \Rightarrow E\downarrow \Rightarrow NX\uparrow \Rightarrow Y^{ad}\uparrow$$

Shifts in the Aggregate Demand Curve. The mechanisms described also indicate why Keynesian analysis suggests that changes in the money supply shift the aggregate demand curve. For a given price level, a rise in the money supply causes the real money supply to increase ($M/P\uparrow$), which leads to an increase in aggregate demand, as shown. Thus an increase in the money supply shifts the aggregate demand curve to the right (as in Figure 1), because it lowers interest rates and stimulates planned investment spending and net exports. Similarly, a decline in the money supply shifts the aggregate demand curve to the left.[2]

In contrast to monetarists, Keynesians believe that other factors (manipulation of government spending and taxes, changes in net exports, and changes in consumer and business spending) are also important causes of shifts in the aggregate demand curve. For instance, if the government spends more ($G\uparrow$) or net exports increase ($NX\uparrow$), aggregate demand rises, and the aggregate demand curve shifts to the right. A decrease in government taxes ($T\downarrow$) leaves consumers with more income to spend, so consumer expenditure rises ($C\uparrow$). Aggregate demand also rises, and the aggregate demand curve shifts to the right. Finally, if consumer and business optimism increases, consumer expenditure and planned investment spending rise ($C\uparrow$, $I\uparrow$), again shifting the aggregate demand curve to the right. Keynes described these waves of optimism and pessimism as "**animal spirits**" and considered them a major factor affecting the aggregate demand curve and an important source of business cycle fluctuations.

The Crowding-Out Debate

You have seen that both monetarists and Keynesians agree that the aggregate demand curve is downward-sloping and shifts in response to changes in the money supply. However, monetarists see only one important source of movements in the aggregate demand curve—changes in the money supply—while Keynesians suggest that other factors—fiscal policy, net exports, and "animal spirits"—are equally important sources of shifts in the aggregate demand curve.

Because aggregate demand can be written as the sum of $C + I + G + NX$, it might appear that any factor affecting one of its components must cause aggregate demand to change. Then it would seem that a fiscal policy change such as a rise in government spending (holding the money supply constant) would necessarily shift the aggregate demand curve. Because monetarists view changes in the money supply as the only important source of shifts in the aggregate demand curve, they must be able to explain why the foregoing reasoning is invalid.

Monetarists agree that an increase in government spending will raise aggregate demand if the other components of aggregate demand—C, I, and NX—remained unchanged after the government spending rise. They contend, however, that the increase in government spending will *crowd out* private spending (C, I, and NX), which will fall by exactly the amount of the government spending increase. For example, an increase of $50 billion in government spending might be offset by a decline of $30 billion in consumer expenditure, $10 billion in investment spending, and $10

[2] A complete demonstration of the Keynesian analysis of the aggregate demand curve is given in Chapters 23 and 24.

billion in net exports. This phenomenon of an exactly offsetting movement of private spending to an expansionary fiscal policy, such as a rise in government spending, is called **complete crowding out**.

How might complete crowding out occur? When government spending increases ($G\uparrow$), the government has to finance this spending by competing with private borrowers for funds in the credit market. Interest rates will rise ($i\uparrow$), increasing the cost of financing purchases of both physical capital and consumer goods and lowering net exports. The result is that private spending will fall ($C\downarrow$, $I\downarrow$, $NX\downarrow$), and so aggregate demand may remain unchanged. This chain of reasoning can be summarized as follows:

$$G\uparrow \Rightarrow i\uparrow \Rightarrow C\downarrow, I\downarrow, NX\downarrow$$

Therefore, $C + I + G + NX = Y^{ad}$ is unchanged.

Keynesians do not deny the validity of the first set of steps. They agree that an increase in government spending raises interest rates, which in turn lowers private spending; indeed, this is a feature of the Keynesian analysis of aggregate demand (see Chapters 23 and 24). However, they contend that in the short run only **partial crowding out** occurs—some decline in private spending that does not completely offset the rise in government spending.

The Keynesian crowding-out picture suggests that when government spending rises, aggregate demand does increase, and the aggregate demand curve shifts to the right. The extent to which crowding out occurs is the issue that separates monetarist and Keynesian views of the aggregate demand curve. We will discuss the evidence on this issue in Chapter 26.

Aggregate Supply

The key feature of aggregate supply is that as the price level increases, the quantity of output supplied increases *in the short run*. Figure 2 illustrates the positive relationship between quantity of output supplied and price level. Suppose that initially the quantity

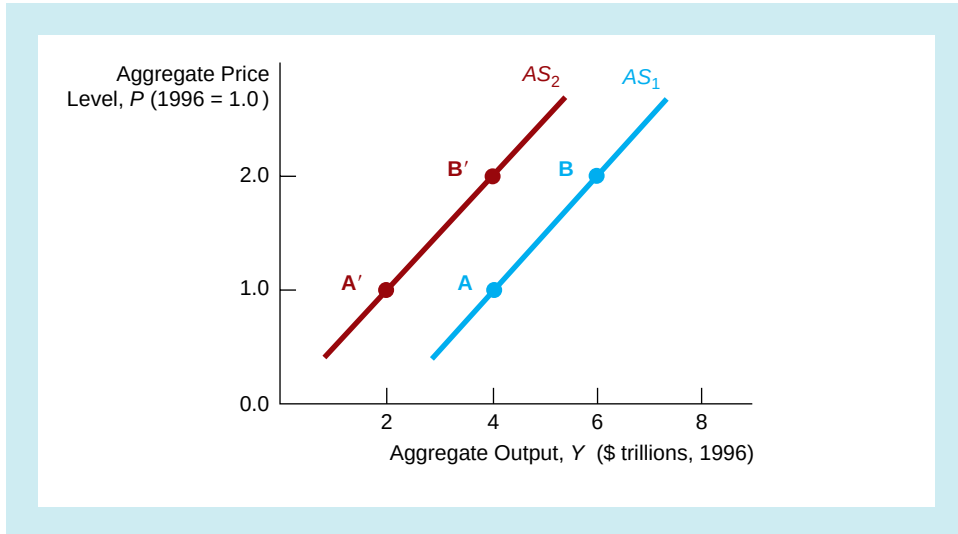


FIGURE 2 Aggregate Supply Curve in the Short Run

A rise in the costs of production shifts the supply curve leftward from AS_1 to AS_2.

of output supplied at a price level of 1.0 is $4 trillion, represented by point A. A rise in the price level to 2.0 leads, in the short run, to an increase to $6 trillion in the quantity of output supplied (point B). The line AS_1 connecting points A and B describes the relationship between the quantity of output supplied in the short run and the price level and is called the **aggregate supply curve**; as you can see, it is upward-sloping.

To understand why the aggregate supply curve slopes upward, we have to look at the factors that cause the quantity of output supplied to change. Because the goal of business is to maximize profits, the quantity of output supplied is determined by the profit made on each unit of output. If profit rises, more output will be produced, and the quantity of output supplied will increase; if it falls, less output will be produced, and the quantity of output supplied will fall.

Profit on a unit of output equals the price for the unit minus the costs of producing it. In the short run, costs of many factors that go into producing goods and services are fixed; wages, for example, are often fixed for periods of time by labor contracts (sometimes as long as three years), and raw materials are often bought by firms under long-term contracts that fix the price. Because these costs of production are fixed in the short run, when the overall price level rises, the price for a unit of output will be rising relative to the costs of producing it, and the profit per unit will rise. Because the higher price level results in higher profits in the short run, firms increase production, and the quantity of aggregate output supplied rises, resulting in an upward-sloping aggregate supply curve.

Frequent mention of the *short run* in the preceding paragraph hints that the aggregate supply curve (AS_1 in Figure 2) may not remain fixed as time passes. To see what happens over time, we need to understand what makes the aggregate supply curve shift.[3]

Shifts in the Aggregate Supply Curve

We have seen that the profit on a unit of output determines the quantity of output supplied. If the cost of producing a unit of output rises, profit on a unit of output falls, and the quantity of output supplied falls. To learn what this implies for the position of the aggregate supply curve, let's consider what happens at a price level of 1.0 when the costs of production increase. Now that firms are earning a lower profit per unit of output, they reduce production, and the quantity of aggregate output supplied falls from $4 (point A) to $2 trillion (point A′). Applying the same reasoning at point B indicates that aggregate output supplied falls to point B′. What we see is that ***the aggregate supply curve shifts to the left when costs of production increase and to the right when costs decrease***.

Equilibrium in Aggregate Supply and Demand Analysis

http://hadm.sph.sc.edu/Courses/Econ/SD/SD.html
An interactive lecture on aggregate supply and demand.

The equilibrium level of aggregate output and the price level will occur at the point where the quantity of aggregate output demanded equals the quantity of aggregate output supplied. However, in the context of aggregate supply and demand analysis, there are two types of equilibrium: short-run and long-run.

[3]The aggregate supply curve is closely linked to the Phillips curve discussed in Chapter 18. More information on the Phillips and aggregate supply curve can be found in an appendix to this chapter, which is on this book's web site at www.aw.com/mishkin.

Equilibrium in the Short Run

Figure 3 illustrates an equilibrium in the short run in which the quantity of aggregate output demanded equals the quantity of output supplied; that is, where the aggregate demand curve *AD* and the aggregate supply curve *AS* intersect at point E. The equilibrium level of aggregate output equals Y^*, and the equilibrium price level equals P^*.

As in our earlier supply and demand analyses, equilibrium is a useful concept only if there is a tendency for the economy to head toward it. We can see that the economy heads toward the equilibrium at point E by first looking at what happens when we are at a price level above the equilibrium price level P^*. If the price level is at P'', the quantity of aggregate output supplied at point D is greater than the quantity of aggregate output demanded at point A. Because people want to sell more goods and services than others want to buy (a condition of *excess supply*), the prices of goods and services will fall, and the aggregate price level will drop, as shown by the downward arrow. This decline in the price level will continue until it has reached its equilibrium level of P^* at point E.

When the price level is below the equilibrium price level, say at P', the quantity of output demanded is greater than the quantity of output supplied. Now the price level will rise, as shown by the upward arrow, because people want to buy more goods than others want to sell (a condition of *excess demand*). This rise in the price level will continue until it has again reached its equilibrium level of P^* at point E.

Equilibrium in the Long Run

Usually in supply and demand analysis, once we find the equilibrium at which the quantity demanded equals the quantity supplied, there is no need for additional discussion. In *aggregate* supply and demand analysis, however, that is not the case. Even when the quantity of aggregate output demanded equals the quantity supplied, forces operate that can cause the equilibrium to move over time. To understand why, we must remember that if costs of production change, the aggregate supply curve will shift.

The most important component of production costs is wages (approximately 70% of production costs), which are determined in the labor market. If the economy is booming, employers will find that they have difficulty hiring qualified workers and may even have a hard time keeping their present employees. In this case, the labor

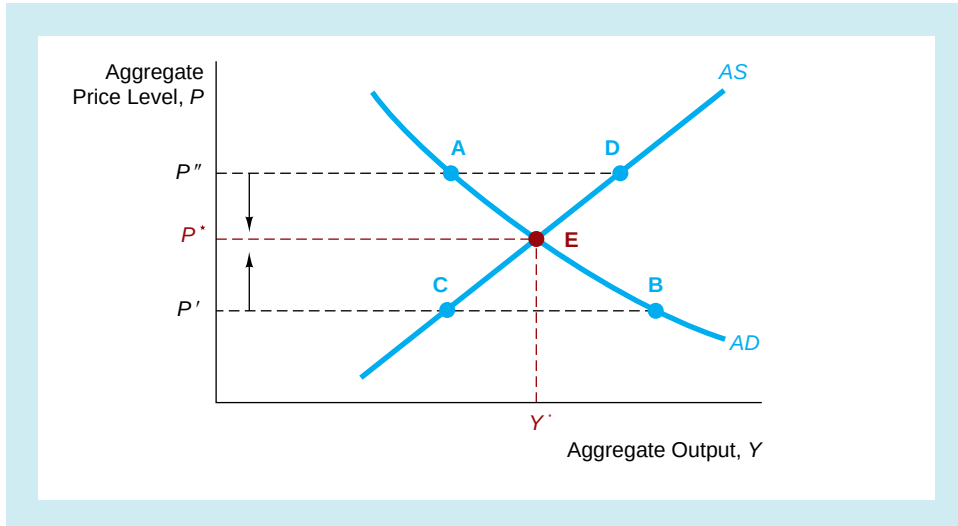


FIGURE 3 Equilibrium in the Short Run

Equilibrium occurs at point E at the intersection of the aggregate demand curve *AD* and the aggregate supply curve *AS*.

market is tight, because the demand for labor exceeds the supply; employers will raise wages to attract needed workers, and this raises the costs of production. The higher costs of production lower the profits per unit of output at each price level, and the aggregate supply curve shifts to the left (see Figure 2).

By contrast, if the economy enters a recession and the labor market is slack, because demand for labor is less than supply, workers who cannot find jobs will be willing to work for lower wages. In addition, employed workers may be willing to make wage concessions to keep from losing their jobs. Therefore, in a slack labor market in which the quantity of labor demanded is less than the quantity supplied, wages and hence costs of production will fall, profits per unit of output will rise, and the aggregate supply curve will shift to the right.

Our analysis suggests that the aggregate supply curve will shift depending on whether the labor market is tight or slack. How do we decide which it is? One helpful concept is the **natural rate of unemployment**, the rate of unemployment to which the economy gravitates in the long run at which demand for labor equals supply. (A related concept is the **NAIRU**, the **nonaccelerating inflation rate of unemployment**, the rate of unemployment at which there is no tendency for inflation to change.) Many economists believe that the rate is currently around 5%. When unemployment is at, say, 4%, below the natural rate of unemployment of 5%, the labor market is tight; wages will rise, and the aggregate supply curve will shift leftward. When unemployment is at, say, 8%, above the natural rate of unemployment, the labor market is slack; wages will fall, and the aggregate supply curve will shift rightward. Only when unemployment is at the natural rate will no pressure exist from the labor market for wages to rise or fall, so the aggregate supply need not shift.

The level of aggregate output produced at the natural rate of unemployment is called the **natural rate level of output**. Because, as we have seen, the aggregate supply curve will not remain stationary when unemployment and aggregate output differ from their natural rate levels, we need to look at how the short-run equilibrium changes over time in response to two situations: when equilibrium is initially below the natural rate level and when it is initially above the natural rate level.

In panel (a) of Figure 4, the initial equilibrium occurs at point 1, the intersection of the aggregate demand curve AD and the initial aggregate supply curve AS_1. Because the level of equilibrium output Y_1 is greater than the natural rate level Y_n, unemployment is less than the natural rate, and excessive tightness exists in the labor market. This tightness drives wages up, raises production costs, and shifts the aggregate supply curve to AS_2. The equilibrium is now at point 2, and output falls to Y_2. Because aggregate output Y_2 is still above the natural rate level, Y_n, wages continue to be driven up, eventually shifting the aggregate supply curve to AS_3. The equilibrium reached at point 3 is on the vertical line at Y_n and is a long-run equilibrium. Because output is at the natural rate level, there is no further pressure on wages to rise and thus no further tendency for the aggregate supply curve to shift.

The movements in panel (a) indicate that the economy will not remain at a level of output higher than the natural rate level because the aggregate supply curve will shift to the left, raise the price level, and cause the economy to slide upward along the aggregate demand curve until it comes to rest at a point on the vertical line through the natural rate level of output Y_n. Because the vertical line through Y_n is the only place at which the aggregate supply curve comes to rest, this vertical line indicates the quantity of output supplied in the long run for any given price level. We can characterize this as the **long-run aggregate supply curve**.

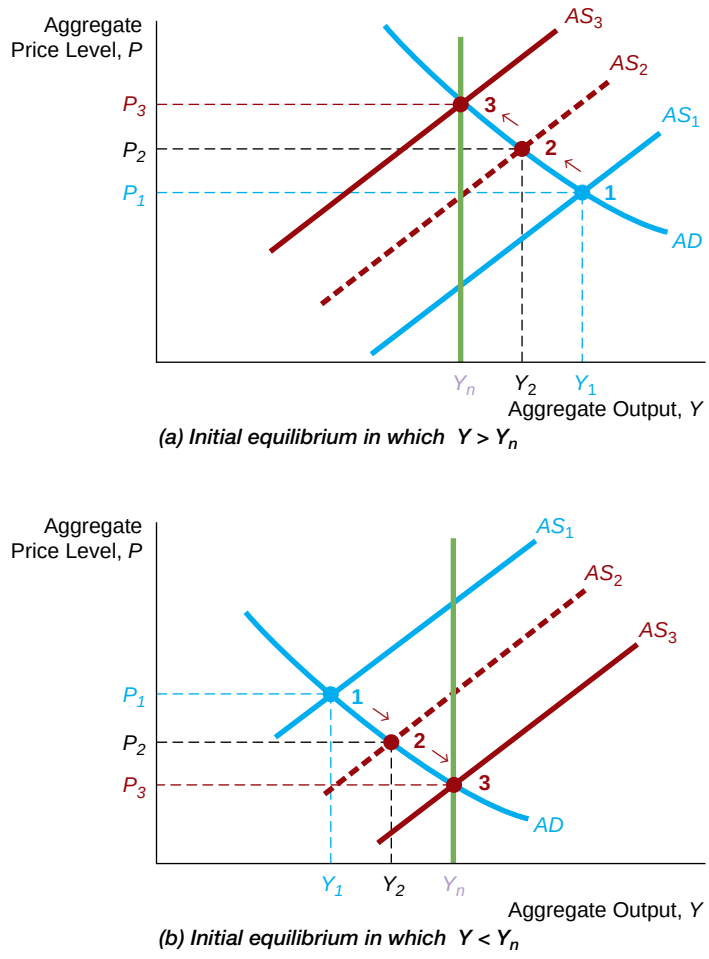


FIGURE 4 Adjustment to Long-Run Equilibrium in Aggregate Supply and Demand Analysis

In both panels, the initial equilibrium is at point 1 at the intersection of AD and AS_1. In panel (a), $Y_1 > Y_n$, so the aggregate supply curve keeps shifting to the left until it reaches AS_3, where output has returned to Y_n. In panel (b), $Y_1 < Y_n$, so the aggregate supply curve keeps shifting to the right until output is again returned to Y_n. Hence in both cases, the economy displays a self-correcting mechanism that returns it to the natural rate level of output.

A characteristic of the economy that causes output to return eventually to the natural rate level regardless of where it is initially.

In panel (b), the initial equilibrium at point 1 is one at which output Y_1 is below the natural rate level. Because unemployment is higher than the natural rate, wages begin to fall, shifting the aggregate supply curve rightward until it comes to rest at AS_3. The economy slides downward along the aggregate demand curve until it reaches the long-run equilibrium point 3, the intersection of the aggregate demand curve AD and the long-run aggregate supply curve at Y_n. Here, as in panel (a), the economy comes to rest when output has again returned to the natural rate level.

A striking feature of both panels of Figure 4 is that regardless of where output is initially, it returns eventually to the natural rate level. This feature is described by saying that the economy has a **self-correcting mechanism**.

An important issue for policymakers is how rapidly this self-correcting mechanism works. Many economists, particularly Keynesians, believe that the self-correcting mechanism takes a long time, so the approach to long-run equilibrium is slow. This

view is reflected in Keynes's often quoted remark, "In the long run, we are all dead." These economists view the self-correcting mechanism as slow, because wages are inflexible, particularly in the downward direction when unemployment is high. The resulting slow wage and price adjustments mean that the aggregate supply curve does not move quickly to restore the economy to the natural rate of unemployment. Hence when unemployment is high, these economists (called **activists**) are more likely to see the need for active government policy to restore the economy to full employment.

Other economists, particularly monetarists, believe that wages are sufficiently flexible that the wage and price adjustment process is reasonably rapid. As a result of this flexibility, adjustment of the aggregate supply curve to its long-run position and the economy's return to the natural rate levels of output and unemployment will occur quickly. Thus these economists (called **nonactivists**) see much less need for active government policy to restore the economy to the natural rate levels of output and unemployment when unemployment is high. Indeed, monetarists advocate the use of a rule whereby the money supply or the monetary base grows at a constant rate so as to minimize fluctuations in aggregate demand that might lead to output fluctuations. We will return in Chapter 27 to the debate about whether active government policy to keep the economy near full employment is beneficial.

Shifts in Aggregate Demand

http://ecedweb.unomaha.edu/Dem_Sup/demand.htm

An interactive tutorial on demand and how various factors cause changes in the demand curve.

You are now ready to analyze what happens when the aggregate demand curve shifts. Our discussion of the Keynesian and monetarist views of aggregate demand indicates that six factors can affect the aggregate demand curve: the money supply, government spending, net exports, taxes, consumer optimism, and business optimism—the last two ("animal spirits") affecting willingness to spend. The possible effect on the aggregate demand curve of these six factors is summarized in Table 1.

Figure 5 depicts the effect of a rightward shift in the aggregate demand curve caused by an increase in the money supply ($M\uparrow$), an increase in government spending ($G\uparrow$), an increase in net exports ($NX\uparrow$), a decrease in taxes ($T\downarrow$), or an increase in the willingness of consumers and businesses to spend because they become more

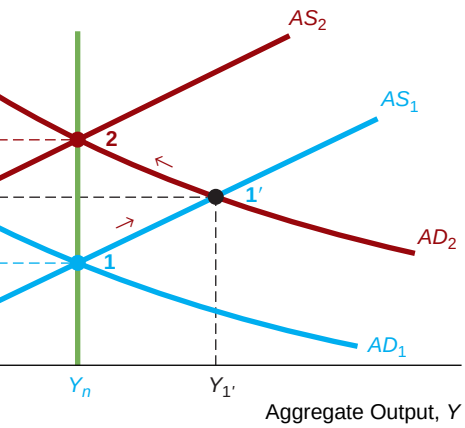
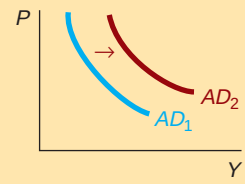
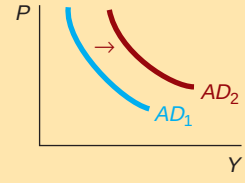
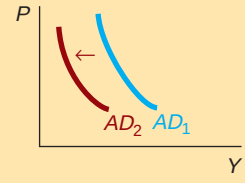
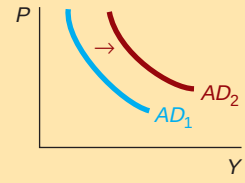
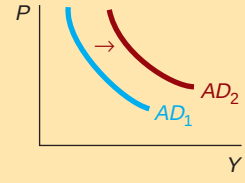
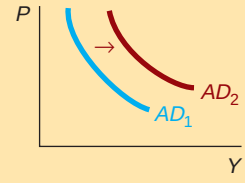


FIGURE 5 Response of Output and the Price Level to a Shift in the Aggregate Demand Curve

A shift in the aggregate demand curve from AD_1 to AD_2 moves the economy from point 1 to point 1′. Because $Y_{1'} > Y_n$, the aggregate supply curve begins to shift leftward, eventually reaching AS_2, where output returns to Y_n and the price level has risen to P_2.

SUMMARY Table 1 Factors That Shift the Aggregate Demand Curve

Factor	Change	Shift in the Aggregate Demand Curve
Money supply M	↑	
Government spending G	↑	
Taxes T	↑	
Net exports NX	↑	
Consumer optimism C	↑	
Business optimism I	↑	

Note: Only increases (↑) in the factors are shown. The effect of decreases in the factors would be the opposite of those indicated in the "Shift" column. Note that monetarists view only the money supply as an important cause of shifts in the aggregate demand curve.

optimistic ($C\uparrow$, $I\uparrow$). The figure has been drawn so that initially the economy is in long-run equilibrium at point 1, where the initial aggregate demand curve AD_1 intersects the aggregate supply AS_1 curve at Y_n. When the aggregate demand curve shifts rightward to AD_2, the economy moves to point 1′, and both output and the price level rise. However, the economy will not remain at point 1′, because output at $Y_{1'}$ is above the natural rate level. Wages will rise, eventually shifting the aggregate supply curve leftward to AS_2, where it finally comes to rest. The economy thus slides up the aggregate demand curve from point 1′ to point 2, which is the point of long-run equilibrium at the intersection of AD_2 and Y_n. ***Although the initial short-run effect of the rightward shift in the aggregate demand curve is a rise in both the price level and output, the ultimate long-run effect is only a rise in the price level***.

Shifts in Aggregate Supply

www.census.gov/statab/www/
Statistics on the U.S. economy in an easy-to-understand format.

Not only can shifts in aggregate demand be a source of fluctuations in aggregate output (the business cycle), but so can shifts in aggregate supply. Factors that cause the aggregate supply curve to shift are the ones that affect the costs of production: (1) tightness of the labor market, (2) expectations of inflation, (3) workers' attempts to push up their real wages, and (4) changes in the production costs that are unrelated to wages (such as energy costs). The first three factors shift the aggregate supply curve by affecting wage costs; the fourth affects other costs of production.

Tightness of the Labor Market. Our analysis of the approach to long-run equilibrium has shown us that when the labor market is tight ($Y > Y_n$), wages and hence production costs rise, and when the labor market is slack ($Y < Y_n$), wages and production costs fall. The effects on the aggregate supply curve are as follows: ***When aggregate output is above the natural rate level, the aggregate supply curve shifts to the left; when aggregate output is below the natural rate level, the aggregate supply curve shifts to the right.***

Expected Price Level. Workers and firms care about wages in real terms; that is, in terms of the goods and services that wages can buy. When the price level increases, a worker earning the same nominal wage will be able to buy fewer goods and services. A worker who expects the price level to rise will thus demand a higher nominal wage in order to keep the real wage from falling. For example, if Chuck the Construction Worker expects prices to increase by 5%, he will want a wage increase of at least 5% (more if he thinks he deserves an increase in real wages). Similarly, if Chuck's employer knows that the houses he is building will rise in value at the same rate as inflation (5%), his employer will be willing to pay Chuck 5% more. An increase in the expected price level leads to higher wages, which in turn raise the costs of production, lower the profit per unit of output at each price level, and shift the aggregate supply curve to the left (see Figure 2). Therefore, ***a rise in the expected price level causes the aggregate supply curve to shift to the left; the greater the expected increase in price level (that is, the higher the expected inflation), the larger the shift***.

Wage Push. Suppose that Chuck and his fellow construction workers decide to strike and succeed in obtaining higher real wages. This wage push will then raise the costs of production, and the aggregate supply curve will shift leftward. ***A successful wage push by workers will cause the aggregate supply curve to shift to the left.***

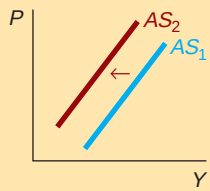
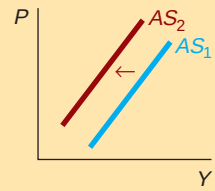
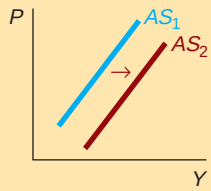
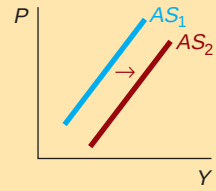
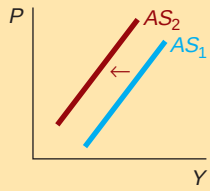
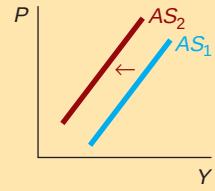
Changes in Production Costs Unrelated to Wages. Changes in technology and in the supply of raw materials (called **supply shocks**) can also shift the aggregate supply curve. A negative supply shock, such as a reduction in the availability of raw materials (like

oil), which raises their price, increases production costs and shifts the aggregate supply curve leftward. A positive supply shock, such as unusually good weather that leads to a bountiful harvest and lowers the cost of food, will reduce production costs and shift the aggregate supply curve rightward. Similarly, the development of a new technology that lowers production costs, perhaps by raising worker productivity, can also be considered a positive supply shock that shifts the aggregate supply curve to the right.

The effect on the aggregate supply curve of changes in production costs unrelated to wages can be summarized as follows: ***A negative supply shock that raises production costs shifts the aggregate supply curve to the left; a positive supply shock that lowers production costs shifts the aggregate supply curve to the right.***[4]

Study Guide As a study aid, factors that shift the aggregate supply curve are listed in Table 2.

SUMMARY **Table 2 Factors That Shift the Aggregate Supply Curve**

Factor	Shifts in the Aggregate Supply Curve	Factor	Shifts in the Aggregate Supply Curve
$Y > Y_n$		Wage push	
$Y < Y_n$		Positive supply shock	
Rise in expected price level		Negative supply shock	

[4]Developments in the foreign exchange market can also shift the aggregate supply curve by changing domestic production costs. As discussed in more detail in Chapter 19, an increase in the value of the dollar makes foreign goods cheaper in the United States. The decline in prices of foreign goods and hence foreign factors of production lowers U.S. production costs and thus raises the profit per unit of output at each price level in the United States. An increase in the value of the dollar therefore shifts the aggregate supply curve to the right. Conversely, a decline in the value of the dollar, which makes foreign factors of production more expensive, shifts the aggregate supply curve to the left.

Now that we know what factors can affect the aggregate supply curve, we can examine what occurs when they cause the aggregate supply curve to shift leftward, as in Figure 6. Suppose that the economy is initially at the natural rate level of output at point 1 when the aggregate supply curve shifts from AS_1 to AS_2 because of a negative supply shock (a sharp rise in energy prices, for example). The economy will move from point 1 to point 2, where the price level rises but aggregate output *falls*. A situation of a rising price level but a falling level of aggregate output, as pictured in Figure 6, has been labeled *stagflation* (a combination of words *stagnation* and *inflation*). At point 2, output is below the natural rate level, so wages fall and shift the aggregate supply curve back to where it was initially at AS_1. The result is that the economy slides down the aggregate demand curve AD_1 (assuming that the aggregate demand curve remains in the same position), and the economy returns to the long-run equilibrium at point 1. ***Although a leftward shift in the aggregate supply curve initially raises the price level and lowers output, the ultimate effect is that output and price level are unchanged (holding the aggregate demand curve constant).***

Shifts in the Long-Run Aggregate Supply Curve: Real Business Cycle Theory and Hysteresis

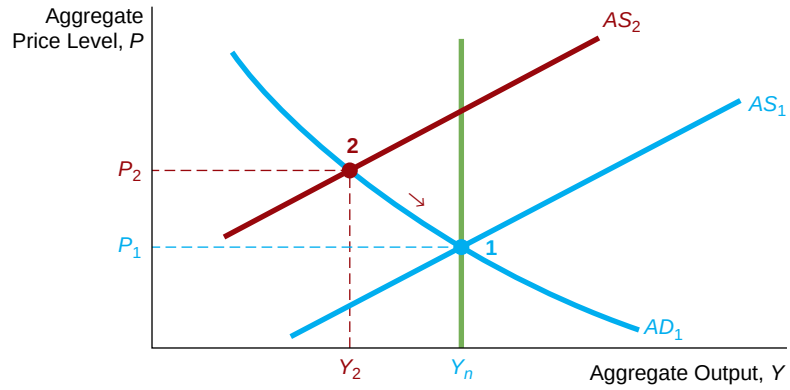
To this point, we have assumed that the natural rate level of output Y_n and hence the long-run aggregate supply curve (the vertical line through Y_n) are given. However, over time, the natural rate level of output increases as a result of economic growth. If the productive capacity of the economy is growing at a steady rate of 3% per year, for example, this means that every year, Y_n will grow by 3% and the long-run aggregate supply curve at Y_n will shift to the right by 3%. To simplify the analysis when Y_n grows at a steady rate, Y_n and the long-run aggregate supply curve are drawn as fixed in the aggregate demand and supply diagrams. Keep in mind, however, that the level of aggregate output pictured in these diagrams is actually best thought of as the level of aggregate output relative to its normal rate of growth (trend).

The usual assumption when conducting aggregate demand and supply analysis is that shifts in either the aggregate demand or aggregate supply curve have no effect on the natural rate level of output (which grows at a steady rate). Movements of aggregate output around the Y_n level in the diagram then describe short-run (business cycle) fluctuations in aggregate output. However, some economists take issue with the assumption that Y_n is unaffected by aggregate demand and supply shocks.

www.fgn.unisg.ch/eumacro/IntrTutor/SGEadas.html
Work with an animated interactive *AD/AS* graph.

FIGURE 6 Response of Output and the Price Level to a Shift in Aggregate Supply

A shift in the aggregate supply curve from AS_1 to AS_2 moves the economy from point 1 to point 2. Because $Y_2 < Y_n$, the aggregate supply curve begins to shift back to the right, eventually returning to AS_1, where the economy is again at point 1.



www.whitehouse.gov/fsbr/esbr.html
The White House sponsors an economic statistics briefing room that reports a wide variety of interesting data dealing with the state of the economy.

One group, led by Edward Prescott of the University of Minnesota, has developed a theory of aggregate economic fluctuations called **real business cycle theory**, in which aggregate supply (real) shocks do affect the natural rate level of output Y_n. This theory views shocks to tastes (workers' willingness to work, for example) and technology (productivity) as the major driving forces behind short-run fluctuations in the business cycle, because these shocks lead to substantial short-run fluctuations in Y_n. Shifts in the aggregate demand curve, perhaps as a result of changes in monetary policy, by contrast are not viewed as being particularly important to aggregate output fluctuations. Because real business cycle theory views most business cycle fluctuations as resulting from fluctuations in the natural rate level of output, it does not see much need for activist policy to eliminate high unemployment. Real business cycle theory is highly controversial and is the subject of intensive research.[5]

Another group of economists disagrees with the assumption that the natural rate level of output Y_n is unaffected by aggregate demand shocks. These economists contend that the natural rate level of unemployment and output are subject to **hysteresis**, a departure from full employment levels as a result of past high unemployment.[6] When unemployment rises because of a reduction of aggregate demand that shifts the *AD* curve inward, the natural rate of unemployment is viewed as rising above the full employment level. This could occur because the unemployed become discouraged and fail to look hard for work or because employers may be reluctant to hire workers who have been unemployed for a long time, seeing it as a signal that the worker is undesirable. The outcome is that the natural rate of unemployment shifts upward after unemployment has become high, and Y_n falls below the full employment level. In this situation, the self-correcting mechanism will be able to return the economy only to the natural rate levels of output and unemployment, not to the full employment level. Only with expansionary policy to shift the aggregate demand curve to the right and raise aggregate output can the natural rate of unemployment be lowered (Y_n raised) to the full employment level. Proponents of hysteresis are thus more likely to promote activist, expansionary policies to restore the economy to full employment.

Study Guide

Aggregate supply and demand analysis are best learned by practicing applications. In this section, we have traced out what happens to aggregate output when there is an increase in the money supply or a negative supply shock. Make sure you can also draw the appropriate shifts in the aggregate demand and supply curves and analyze what happens when other variables such as taxes or the expected price level change.

Conclusions

Aggregate demand and supply analysis yields the following conclusions (under the usual assumption that the natural rate level of output is unaffected by aggregate demand and supply shocks):

1. A shift in the aggregate demand curve—which can be caused by changes in monetary policy (the money supply), fiscal policy (government spending or taxes),

[5] See Charles Plosser, "Understanding Real Business Cycles," *Journal of Economic Perspectives* (1989): 51–77, for a nontechnical discussion of real business cycle theory.

[6] For a further discussion of hysteresis, see Olivier Blanchard and Lawrence Summers, "Hysteresis in the European Unemployment Problem," *NBER Macroeconomics Annual*, 1986, 1, ed. Stanley Fischer (Cambridge, Mass.: M.I.T. Press, 1986), pp. 15–78.

international trade (net exports), or "animal spirits" (business and consumer optimism)—affects output only in the short run and has no effect in the long run. Furthermore, the initial change in the price level is less than is achieved in the long run, when the aggregate supply curve has fully adjusted.

2. A shift in the aggregate supply curve—which can be caused by changes in expected inflation, workers' attempts to push up real wages, or a supply shock—affects output and prices only in the short run and has no effect in the long run (holding the aggregate demand curve constant).

3. The economy has a self-correcting mechanism, which will return it to the natural rate levels of unemployment and aggregate output over time.

Application

Explaining Past Business Cycle Episodes

Aggregate supply and demand analysis is an extremely useful tool for analyzing aggregate economic activity; we will apply it to several business cycle episodes. To simplify our analysis, we always assume in all three examples that aggregate output is initially at the natural rate level.

Vietnam War Buildup, 1964–1970

America's involvement in Vietnam began to escalate in the early 1960s, and after 1964, the United States was fighting a full-scale war. Beginning in 1965, the resulting increases in military expenditure raised government spending, while at the same time the Federal Reserve increased the rate of money growth in an attempt to keep interest rates from rising. What does aggregate supply and demand analysis suggest should have happened to aggregate output and the price level as a result of the Vietnam War buildup?

The rise in government spending and the higher rate of money growth would shift the aggregate demand curve to the right (shown in Figure 5). As a result, aggregate output would rise, unemployment would fall, and the price level would rise. Table 3 demonstrates that this is exactly what happened: The unemployment rate fell steadily from 1964 to 1969, remaining well below what economists now think was the natural rate of unemployment during that period (around 5%), and inflation began to rise. As Figure 5 predicts, unemployment would eventually begin to return to the natural rate level because of the economy's self-correcting mechanism. This is exactly what we saw occurring in 1970, when the inflation rate rose even higher and unemployment increased.

Negative Supply Shocks, 1973–1975 and 1978–1980

In 1973, the U.S. economy was hit by a series of negative supply shocks. As a result of the oil embargo stemming from the Arab-Israeli war of 1973, the Organization of Petroleum Exporting Countries (OPEC) was able to engineer a quadrupling of oil prices by restricting oil production. In addition, a series of crop failures throughout the world led to a sharp increase in food prices. Another factor was the termination of wage and price controls in 1973 and 1974, which led to a push by workers to obtain wage increases that had been prevented by the controls. The triple thrust of these events caused the aggregate supply curve to shift sharply leftward, and as the aggregate demand and

Table 3 Unemployment and Inflation During the Vietnam War Buildup, 1964–1970

Year	Unemployment Rate (%)	Inflation (Year to Year) (%)
1964	5.0	1.3
1965	4.4	1.6
1966	3.7	2.9
1967	3.7	3.1
1968	3.5	4.2
1969	3.4	5.5
1970	4.8	5.7

Source: Economic Report of the President.

supply diagram in Figure 6 predicts, both the price level and unemployment began to rise dramatically (see Table 4).

The 1978–1980 period was almost an exact replay of the 1973–1975 period. By 1978, the economy had just about fully recovered from the 1973–1974 supply shocks, when poor harvests and a doubling of oil prices (as a result of the overthrow of the Shah of Iran) again led to another sharp leftward shift of the aggregate supply curve. The pattern predicted by Figure 6 played itself out again—inflation and unemployment both shot upward (see Table 4).

Favorable Supply Shocks, 1995–1999

In February 1994, the Federal Reserve began to raise interest rates, because it believed the economy would be reaching the natural rate of output and unemployment in 1995 and might become overheated thereafter. As we can see in Table 5, however, the economy continued to grow rapidly, with unem-

Table 4 Unemployment and Inflation During the Negative Supply Shock Periods, 1973–1975 and 1978–1980

Year	Unemployment Rate (%)	Inflation (Year to Year) (%)	Year	Unemployment Rate (%)	Inflation (Year to Year) (%)
1973	4.8	6.2	1978	6.0	7.6
1974	5.5	11.0	1979	5.8	11.3
1975	8.3	9.1	1980	7.0	13.5

Source: Economic Report of the President.

Table 5 Unemployment and Inflation During the Favorable Supply Shock Period, 1995–2000

Year	Unemployment Rate (%)	Inflation (Year to Year) (%)
1995	5.6	2.8
1996	5.4	3.0
1997	4.9	2.3
1998	4.5	1.6
1999	4.2	2.2

Source: Economic Report of the President; ftp://ftp.bls.gov/pub/special.requests/cpi/cpiai.txt.

ployment falling to below 5%, well below what many economists believed to be the natural rate level, and yet inflation continued to fall, declining to around 2%. Can aggregate demand and supply analysis explain what happened?

The answer is yes. Two favorable supply shocks hit the economy in the late 1990s. First, changes in the health care industry, such as the movements to health maintenance organizations (HMOs), reduced medical care costs substantially relative to other goods and services. Second, the computer revolution finally began to have a favorable impact on productivity, raising the potential growth rate of the economy (which journalists have dubbed the "new economy"). The outcome was a rightward shift in the aggregate supply curve, producing the opposite result depicted in Figure 6: Aggregate output rose, and unemployment fell, while inflation also declined.

Summary

1. The aggregate demand curve indicates the quantity of aggregate output demanded at each price level, and it is downward-sloping. Monetarists view changes in the money supply as the primary source of shifts in the aggregate demand curve. Keynesians believe that not only are changes in the money supply important to shifts in the aggregate demand curve, but so are changes in fiscal policy (government spending and taxes), net exports, and the willingness of consumers and businesses to spend ("animal spirits").
2. In the short run, the aggregate supply curve slopes upward, because a rise in the price level raises the profit earned on each unit of production, and the quantity of output supplied rises. Four factors can cause the aggregate supply curve to shift: tightness of the labor market as represented by unemployment relative to the natural rate, expectations of inflation, workers' attempts to push up their real wages, and supply shocks unrelated to wages that affect production costs.
3. Equilibrium in the short run occurs at the point where the aggregate demand curve intersects the aggregate supply curve. Although this is where the economy heads temporarily, it has a self-correcting mechanism, which leads it to settle permanently at the long-run equilibrium where aggregate output is at its natural rate level. Shifts in either the aggregate demand or the aggregate supply curve can produce changes in aggregate output and the price level.

Key Terms

activists, p. 592
aggregate demand, p. 582
aggregate demand curve, p. 582
aggregate supply, p. 582
aggregate supply curve, p. 588
"animal spirits," p. 586
complete crowding out, p. 587
consumer expenditure, p. 585
equation of exchange, p. 583
government spending, p. 585
hysteresis, p. 597
Keynesians, p. 582
long-run aggregate supply curve, p. 590
modern quantity theory of money, p. 584
monetarists, p. 582
natural rate level of output, p. 590
natural rate of unemployment, p. 590
net exports, p. 585
nonaccelerating inflation rate of unemployment (NAIRU), p. 590
nonactivists, p. 592
partial crowding out, p. 587
planned investment spending, p. 585
real business cycle theory, p. 597
self-correcting mechanism, p. 591
supply shocks, p. 594
velocity of money, p. 582

Questions and Problems

Questions marked with an asterisk are answered at the end of the book in an appendix, "Answers to Selected Questions and Problems."

1. Given that a monetarist predicts velocity to be 5, graph the aggregate demand curve that results if the money supply is $400 billion. If the money supply falls to $50 billion, what happens to the position of the aggregate demand curve?

*2. Milton Friedman states, "Money is all that matters to nominal income." How is this statement built into the aggregate demand curve in the monetarist framework?

3. Suppose that government spending is raised at the same time that the money supply is lowered. What will happen to the position of the Keynesian aggregate demand curve? The monetarist aggregate demand curve?

*4. Why does the Keynesian aggregate demand curve shift when "animal spirits" change, but the monetarist aggregate demand curve does not?

5. If the dollar increases in value relative to foreign currencies so that foreign goods become cheaper in the United States, what will happen to the position of the aggregate supply curve? The aggregate demand curve?

*6. "Profit-maximizing behavior on the part of firms explains why the aggregate supply curve is upward-sloping." Is this statement true, false, or uncertain? Explain your answer.

7. If huge budget deficits cause the public to think that there will be higher inflation in the future, what is likely to happen to the aggregate supply curve when budget deficits rise?

*8. If a pill were invented that made workers twice as productive but their wages did not change, what would happen to the position of the aggregate supply curve?

9. When aggregate output is below the natural rate level, what will happen to the price level over time if the aggregate demand curve remains unchanged? Why?

*10. Show how aggregate supply and demand analysis can explain why both aggregate output and the price level fell sharply when investment spending collapsed during the Great Depression.

11. "An important difference between monetarists and Keynesians rests on how long they think the long run actually is." Is this statement true, false, or uncertain? Explain your answer.

Using Economic Analysis to Predict the Future

*12. Predict what will happen to aggregate output and the price level if the Federal Reserve increases the money supply at the same time that Congress implements an income tax cut.

13. Suppose that the public believes that a newly announced anti-inflation program will work and so lowers its expectations of future inflation. What will

happen to aggregate output and the price level in the short run?

*14. Proposals have come before Congress that advocate the implementation of a national sales tax. Predict the effect of such a tax on both the aggregate supply and demand curves and on aggregate output and the price level.

15. When there is a decline in the value of the dollar, some experts expect this to lead to a dramatic improvement in the ability of American firms to compete abroad. Predict what would happen to output and the price level in the United States as a result.

Web Exercises

1. As this book goes to press, the U.S. economy is still suffering from slow growth and relatively high unemployment. Go to www.whitehouse.gov/fsbr/esbr.html and follow the link to unemployment statistics. What has happened to unemployment since the last reported figure in Table 5?

2. As the economy stalled toward the end of 2002, Fed policymakers were beginning to be concerned about deflation. Go to www.whitehouse.gov/fsbr/esbr.html and follow the link to prices. What has happened to prices since the last reported figure in Table 5? Does deflation still appear to be a threat?

appendix to chapter 25

Aggregate Supply and the Phillips Curve

In this appendix, we examine how economists' view of aggregate supply has evolved over time and how the concept called the *Phillips curve*, which described the relationship between unemployment and inflation, fits into the analysis of aggregate supply.

The classical economists, who predated Keynes, believed that wages and prices were extremely flexible, so the economy would always adjust quickly to the natural rate level of output Y_n. This view is equivalent to assuming that the aggregate supply curve is vertical at an output level of Y_n even in the short run.

With the advent of the Great Depression in 1929 and the subsequent long period of high unemployment, the classical view of an economy that adjusts quickly to the natural rate level of output became less tenable. The teachings of John Maynard Keynes emerged as the dominant way of thinking about the determination of aggregate output, and the view that aggregate supply is vertical was abandoned. Instead, Keynesians in the 1930s, 1940s, and 1950s assumed that for all practical purposes, the price level could be treated as fixed. They viewed aggregate supply as a horizontal curve along which aggregate output could increase without an increase in the price level.

In 1958, A. W. Phillips published a famous paper that outlined a relationship between unemployment and inflation.[1] This relationship was popularized by Paul Samuelson and Robert Solow of the Massachusetts Institute of Technology in the early 1960s, and naturally enough, it became known as the *Phillips curve*, after its discoverer. The Phillips curve indicates that the rate of change of wages $\Delta w/w$, called *wage inflation*, is negatively related to the difference between the actual unemployment rate U and the natural rate of unemployment U_n:

$$\Delta w/w = -h(U - U_n)$$

where h is a constant that indicates how much wage inflation changes for a given change in $U - U_n$. If h were 2, for example, a 1 percent increase in the unemployment rate relative to the natural rate would result in a 2 percent decline in wage inflation.

The Phillips curve provides a view of aggregate supply because it indicates that a rise in aggregate output that lowers the unemployment rate will raise wage inflation and thus lead to a higher level of wages and the price level. In other words, the Phillips curve implies that the aggregate supply curve will be upward-sloping. In addition, it indicates that when $U > U_n$ (the labor market is slack), $\Delta w/w$ is negative

[1] A. W. Phillips, "The Relationship Between Unemployment and the Rate of Change of Money Wages in the United Kingdom, 1861–1957," *Economica* 25 (1958): 283–299.

and wages decline over time. Hence the Phillips curve supports the view of aggregate supply in Chapter 24 that when the labor market is slack, production costs will fall and the aggregate supply curve will shift to the right.[2]

Figure 1 shows what the Phillips curve relationship looks like for the United States. As we can see from panel (a), the relationship works well until 1969 and seems to indicate an apparent trade-off between unemployment and wage inflation: If the public wants to have a lower unemployment rate, it can "buy" this by accepting a higher rate of wage inflation.

In 1967, however, Milton Friedman pointed out a severe flaw in the Phillips curve analysis: It left out an important factor that affects wage changes: workers' expectations of inflation.[3] Friedman noted that firms and workers are concerned with

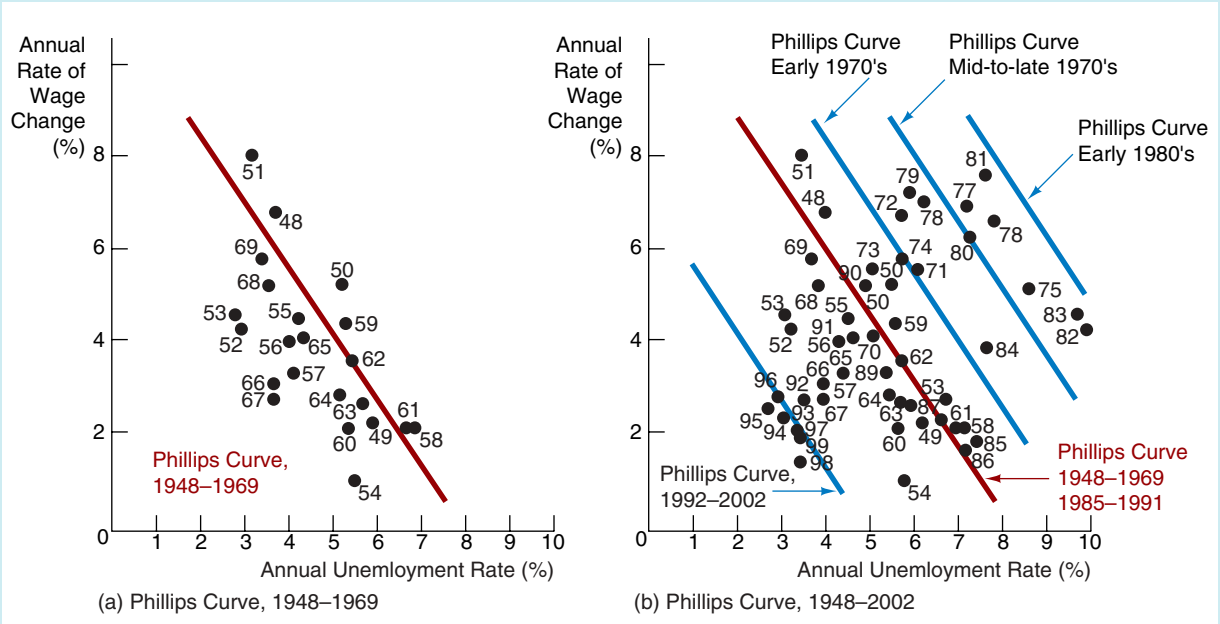


FIGURE 1 Phillips Curve in the United States

Although the Phillips curve relationship worked fairly well from 1948 to 1969, after this period it appeared to shift upward, as is clear from panel (b). Looking at the whole period after World War II, there is no apparent trade-off between unemployment and inflation.

Source: Economic Report of the President. http://w3.access.gpo.gov/usbudget/

[2] Because workers normally become more productive over time as a result of new technology and increases in physical capital, their real wages grow over time even when the economy is at the natural rate of unemployment. To reflect this, the Phillips curve should include a term that reflects the growth in real wages due to higher worker productivity. We have left this term out of the equation in the text, because higher productivity that results in higher real wages will not cause the aggregate supply curve to shift. If, for example, workers become 3 percent more productive every year and their real wages grow at 3 percent per year, the effective cost of workers to the firm (called *unit labor costs*) remains unchanged, and the aggregate supply curve does not shift. Thus the $\Delta w/w$ term in the Phillips curve is more accurately thought of as the change in the unit labor costs.

[3] This criticism of the Phillips curve was outlined in Milton Friedman's famous presidential address to the American economic Association: Milton Friedman, "The Role of Monetary Policy," *American Economic Review* 58 (1968): 1–17.

real wages, not nominal wages; they are concerned with the wage adjusted for any expected increase in the price level—that is, they look at the rate of change of wages minus expected inflation. When unemployment is high relative to the natural rate, real (not nominal) wages should fall ($\Delta w/w - \pi^e < 0$); when unemployment is low relative to the natural state, real wages should rise ($\Delta w/w - \pi^e > 0$). The Phillips curve relationship thus needs to be modified by replacing $\Delta w/w$ by $\Delta w/w - \pi^e$. This results in an *expectations-augmented Phillips curve*, expressed as:

$$\Delta w/w - \pi^e = -h(U - U_n) \quad \text{or} \quad \Delta w/w = -h(U - U_n) + \pi^e$$

The expectations-augmented Phillips curve implies that as expected inflation rises, nominal wages will be increased to prevent real wages from falling, and the Phillips curve will shift upward. The resulting rise in production costs will then shift the aggregate supply curve leftward. The conclusion from Friedman's modification of the Phillips curve is therefore that the higher inflation is expected to be, the larger the leftward shift in the aggregate supply curve; this conclusion is built into the analysis of the aggregate supply curve in the chapter.

Friedman's modifications of the Phillips curve analysis was remarkably clairvoyant: As inflation increased in the late 1960s, the Phillips curve did indeed begin to shift upward, as we can see from panel (b). An important feature of panel (b) is that a trade-off between unemployment and wage inflation is no longer apparent; there is no clear-cut relationship between unemployment and wage inflation—a high rate of wage inflation does not mean that unemployment is low, nor does a low rate of wage inflation mean that unemployment is high. This is exactly what the expectations-augmented Phillips curve predicts: A rate of unemployment permanently below the natural rate of unemployment cannot be "bought" by accepting a higher rate of inflation because no long-run trade-off between unemployment and wage inflation exists.[4]

A further refinement of the concept of aggregate supply came from research by Milton Friedman, Edmund Phelps, and Robert Lucas, who explored the implications of the expectations-augmented Phillips curve for the behavior of unemployment. Solving the expectations-augmented Phillips curve for U leads to the following expression:

$$U = U_n - (\Delta w/w - \pi^e)/h$$

Because wage inflation and price inflation are closely tied to each other, π can be substituted for $\Delta w/w$ in this expression to obtain:

$$U = U_n - (\pi - \pi^e)/h$$

[4] This prediction can be derived from the expectations-augmented Phillips curve as follows. When wage inflation is held at a constant level, inflation and expected inflation will eventually equal wage inflation. Thus in the long run, $\pi^e = \Delta w/w$. Substituting the long-run value of π^e into the expectations-augmented Phillips curve gives:

$$\Delta w/w = -h(U - U_n) + \Delta w/w$$

Subtracting $\Delta w/w$ from both sides of the equation gives $0 = -h(U - U_n)$, which implies that $U = U_n$. This tells us that in the long run, for any level of wage inflation, unemployment will settle to its natural rate level; hence the long-run Phillips curve is vertical, and there is no long-run trade-off between unemployment and wage inflation.

This expression, often referred to as *Lucas supply function*, indicates that deviations of unemployment and aggregate output from the natural rate levels respond to unanticipated inflation (actual inflation minus expected inflation, $\pi - \pi^e$). When inflation is greater than anticipated, unemployment will be below the natural rate (and aggregate output above the natural rate). When inflation is below its anticipated value, unemployment will rise above the natural rate level. The conclusion from this view of aggregate supply is that only unanticipated policy can cause deviations from the natural rate of unemployment and output. The implications of this view are explored in detail in Chapter 28.

Chapter 26

Transmission Mechanisms of Monetary Policy: The Evidence

PREVIEW

Since 1980, the U.S. economy has been on a roller coaster, with output, unemployment, and inflation undergoing drastic fluctuations. At the start of the 1980s, inflation was running at double-digit levels, and the recession of 1980 was followed by one of the shortest economic expansions on record. After a year, the economy plunged into the 1981–1982 recession, the most severe economic contraction in the postwar era—the unemployment rate climbed to over 10%, and only then did the inflation rate begin to come down to below the 5% level. The 1981–1982 recession was then followed by a long economic expansion that reduced the unemployment rate to below 6% in the 1987–1990 period. With Iraq's invasion of Kuwait and a rise in oil prices in the second half of 1990, the economy again plunged into recession. Subsequent growth in the economy was sluggish at first but eventually sped up, lowering the unemployment rate to below 5% in the late 1990s. In March 2001, the economy slipped into recession, with the unemployment rate climbing to around 6%. In light of large fluctuations in aggregate output (reflected in the unemployment rate) and inflation, and the economic instability that accompanies them, policymakers face the following dilemma: What policy or policies, if any, should be implemented to reduce output and inflation fluctuations in the future?

To answer this question, monetary policymakers must have an accurate assessment of the timing and effect of their policies on the economy. To make this assessment, they need to understand the mechanisms through which monetary policy affects the economy. In this chapter, we examine empirical evidence on the effect of monetary policy on economic activity. We first look at a framework for evaluating empirical evidence and then use this framework to understand why there are still deep disagreements on the importance of monetary policy to the economy. We then go on to examine the transmission mechanisms of monetary policy and evaluate the empirical evidence on them to better understand the role that monetary policy plays in the economy. We will see that these monetary transmission mechanisms emphasize the link between the financial system (which we studied in the first three parts of this book) and monetary theory, the subject of this part.

Framework for Evaluating Empirical Evidence

To develop a framework for understanding how to evaluate empirical evidence, we need to recognize that there are two basic types of empirical evidence in economics and other scientific disciplines: **Structural model evidence** examines whether one

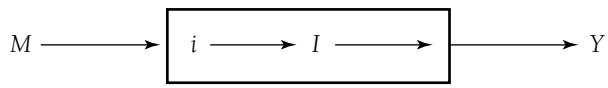
variable affects another by using data to build a model that explains the channels through which this variable affects the other; **reduced-form evidence** examines whether one variable has an effect on another simply by looking directly at the relationship between the two variables.

Suppose that you were interested in whether drinking coffee leads to heart disease. Structural model evidence would involve developing a model that analyzed data on how coffee is metabolized by the human body, how it affects the operation of the heart, and how its effects on the heart lead to heart attacks. Reduced-form evidence would involve looking directly at whether coffee drinkers tend to experience heart attacks more frequently than non–coffee drinkers.

How you look at the evidence—whether you focus on structural model evidence or reduced-form evidence—can lead to different conclusions. This is particularly true for the debate between monetarists and Keynesians. Monetarists tend to focus on reduced-form evidence and feel that changes in the money supply are more important to economic activity than Keynesians do; Keynesians, for their part, focus on structural model evidence. To understand the differences in their views about the importance of monetary policy, we need to look at the nature of the two types of evidence and the advantages and disadvantages of each.

Structural Model Evidence

The Keynesian analysis discussed in Chapter 25 is specific about the channels through which the money supply affects economic activity (called the **transmission mechanisms of monetary policy**). Keynesians typically examine the effect of money on economic activity by building a **structural model**, a description of how the economy operates using a collection of equations that describe the behavior of firms and consumers in many sectors of the economy. These equations then show the channels through which monetary and fiscal policy affect aggregate output and spending. A Keynesian structural model might have behavioral equations that describe the workings of monetary policy with the following schematic diagram:



The model describes the transmission mechanism of monetary policy as follows: The money supply M affects interest rates i, which in turn affect investment spending I, which in turn affects aggregate output or aggregate spending Y. The Keynesians examine the relationship between M and Y by looking at empirical evidence (structural model evidence) on the specific channels of monetary influence, such as the link between interest rates and investment spending.

Reduced-Form Evidence

Monetarists do not describe specific ways in which the money supply affects aggregate spending. Instead, they examine the effect of money on economic activity by looking at whether movements in Y are tightly linked to (have a high correlation with) movements in M. Using reduced-form evidence, monetarists analyze the effect of M on Y as if the economy were a black box whose workings cannot be seen. The monetarist way of looking at the evidence can be represented by the following schematic diagram, in which the economy is drawn as a black box with a question mark:

Now that we have seen how monetarists and Keynesians look at the empirical evidence on the link between money and economic activity, we can consider the advantages and disadvantages of their approaches.

Advantages and Disadvantages of Structural Model Evidence

The structural model approach, used primarily by Keynesians, has the advantage of giving us an understanding of how the economy works. If the structure is correct—if it contains all the transmission mechanisms and channels through which monetary and fiscal policy can affect economic activity, the structural model approach has three major advantages over the reduced-form approach.

1. Because we can evaluate each transmission mechanism separately to see whether it is plausible, we will obtain more pieces of evidence on whether money has an important effect on economic activity. If we find important effects of monetary policy on economic activity, for example, we will have more confidence that changes in monetary policy actually cause the changes in economic activity; that is, we will have more confidence on the direction of causation between *M* and *Y*.

2. Knowing how changes in monetary policy affect economic activity may help us predict the effect of *M* on *Y* more accurately. For example, expansions in the money supply might be found to be less effective when interest rates are low. Then, when interest rates are higher, we would be able to predict that an expansion in the money supply would have a larger impact on *Y* than would otherwise be the case.

3. By knowing how the economy operates, we may be able to predict how institutional changes in the economy might affect the link between *M* and *Y*. For instance, before 1980, when Regulation Q was still in effect, restrictions on interest payments on savings deposits meant that the average consumer would not earn more on savings when interest rates rose. Since the termination of Regulation Q, the average consumer now earns more on savings when interest rates rise. If we understand how earnings on savings affect consumer spending, we might be able to say that a change in monetary policy, which affects interest rates, will have a different effect today than it would have had before 1980. Because of the rapid pace of financial innovation, the advantage of being able to predict how institutional changes affect the link between *M* and *Y* may be even more important now than in the past.

These three advantages of the structural model approach suggest that this approach is better than the reduced-form approach *if we know the correct structure of the model.* Put another way, structural model evidence is only as good as the structural model it is based on; it is best only if all the transmission mechanisms are fully understood. This is a big *if*, as failing to include one or two relevant transmission mechanisms for monetary policy in the structural model might result in a serious underestimate of the impact of *M* on *Y*.

Monetarists worry that many Keynesian structural models may ignore the transmission mechanisms for monetary policy that are most important. For example, if the most important monetary transmission mechanisms involve consumer spending rather than investment spending, the Keynesian structural model (such as the $M\uparrow \Rightarrow i\downarrow \Rightarrow I\uparrow \Rightarrow Y\uparrow$ model we used earlier), which focuses on investment spending for its monetary transmission mechanism, may underestimate the importance of money to economic activity. In other words, monetarists reject the interpretation of evidence from many Keynesian structural models because they believe that the channels of monetary influence are too narrowly defined. In a sense, they accuse Keynesians of wearing blinders that prevent them from recognizing the full importance of monetary policy.

Advantages and Disadvantages of Reduced-Form Evidence

The main advantage of reduced-form evidence over structural model evidence is that no restrictions are imposed on the way monetary policy affects the economy. If we are not sure that we know what all the monetary transmission mechanisms are, we may be more likely to spot the full effect of M on Y by looking at whether movements in Y correlate highly with movements in M. Monetarists favor reduced-form evidence, because they believe that the particular channels through which changes in the money supply affect Y are diverse and continually changing. They contend that it may be too difficult to identify all the transmission mechanisms of monetary policy.

The most notable objection to reduced-form evidence is that it may misleadingly suggest that changes in M cause changes in Y when that is not the case. A basic principle applicable to all scientific disciplines, including economics, states that ***correlation does not necessarily imply causation.*** That movement of one variable is linked to another doesn't necessarily mean that one variable *causes* the other.

Suppose, for example, you notice that wherever criminal activity abounds, more police patrol the street. Should you conclude from this evidence that police patrols cause criminal activity and recommend pulling police off the street to lower the crime rate? The answer is clearly no, because police patrols do not cause criminal activity; criminal activity causes police patrols. This situation is called **reverse causation** and can produce misleading conclusions when interpreting correlations (see Box 1).

The reverse causation problem may be present when examining the link between money and aggregate output or spending. Our discussion of the conduct of monetary policy in Chapter 18 suggested that when the Federal Reserve has an interest-rate or a free reserves target, higher output may lead to a higher money supply. If most of the correlation between M and Y occurs because of the Fed's interest-rate target, controlling the money supply will not help control aggregate output, because it is actually Y that is causing M rather than the other way around.

Another facet of the correlation–causation question is that an outside factor, yet unknown, could be the driving force behind two variables that move together. Coffee drinking might be associated with heart disease not because coffee drinking causes heart attacks but because coffee drinkers tend to be people who are under a lot of stress and the stress causes heart attacks. Getting people to stop drinking coffee, then, would not lower the incidence of heart disease. Similarly, if there is an unknown outside factor that causes M and Y to move together, controlling M will not improve control of Y. (The perils of ignoring an outside driving factor are illustrated in Box 2.)

Box 1

Perils of Reverse Causation

A Russian Folk Tale. A Russian folk tale illustrates the problems that can arise from reverse causation. As the story goes, there once was a severe epidemic in the Russian countryside and many doctors were sent to the towns where the epidemic was at its worst. The peasants in the towns noticed that wherever doctors went, many people were dying. So to reduce the death rate, they killed all the doctors.

Were the peasants better off? Clearly not.

Box 2

Perils of Ignoring an Outside Driving Factor

How to Lose a Presidential Election. Ever since Muncie, Indiana, was dubbed "Middletown" by two sociology studies over half a century ago, it has produced a vote for president that closely mirrors the national vote; that is, in every election, there has been a very high correlation between Muncie's vote and the national vote. Noticing this, a political adviser to a presidential candidate recommends that the candidate's election will be assured if *all* the candidate's campaign funds are spent in Muncie. Should the presidential candidate promote or fire this adviser? Why?

It is very unlikely that the vote in a small town like Muncie drives the vote in a national election. Rather, it is more likely that national preferences are a third driving factor that determines the vote in Muncie and also determines the vote in the national election. Changing the vote in Muncie will thus only break the relationship between that town's vote and national preferences and will have almost no impact on the election. Spending all the campaign money on this town will therefore be a waste of money.

The presidential candidate should definitely fire the adviser.

Conclusions

No clear-cut case can be made that reduced-form evidence is preferable to structural model evidence or vice versa. The structural model approach, used primarily by Keynesians, offers an understanding of how the economy works. If the structure is correct, it predicts the effect of monetary policy more accurately, allows predictions of the effect of monetary policy when institutions change, and provides more confidence in the direction of causation between M and Y. If the structure of the model is not correctly specified because it leaves out important transmission mechanisms of monetary policy, it could be very misleading.

The reduced-form approach, used primarily by monetarists, does not restrict the way monetary policy affects the economy and may be more likely to spot the full effect of M on Y. However, reduced-form evidence cannot rule out reverse causation, whereby changes in output cause changes in money, or the possibility that an outside factor drives changes in both output and money. A high correlation of money and output might then be misleading, because controlling the money supply would not help control the level of output.

Armed with the framework to evaluate empirical evidence we have outlined here, we can now use it to evaluate the empirical debate between monetarists and Keynesians on the importance of money to the economy.

Early Keynesian Evidence on the Importance of Money

Although Keynes proposed his theory for analyzing aggregate economic activity in 1936, his views reached their peak of popularity among economists in the 1950s and early 1960s, when the majority of economists had accepted his framework. Although Keynesians currently believe that monetary policy has important effects on economic activity, the early Keynesians of the 1950s and early 1960s characteristically held the

view that *monetary policy does not matter at all* to movements in aggregate output and hence to the business cycle.

Their belief in the ineffectiveness of monetary policy stemmed from three pieces of structural model evidence:

1. During the Great Depression, interest rates on U.S. Treasury securities fell to extremely low levels; the three-month Treasury bill rate, for example, declined to below 1%. Early Keynesians viewed monetary policy as affecting aggregate demand solely through its effect on nominal interest rates, which in turn affect investment spending; they believed that low interest rates during the depression indicated that monetary policy was easy (expansionary) because it encouraged investment spending and so could not have played a contractionary role during this period. Seeing that monetary policy was not capable of explaining why the worst economic contraction in U.S. history had taken place, they concluded that changes in the money supply have no effect on aggregate output—in other words, that money doesn't matter.
2. Early empirical studies found no linkage between movements in nominal interest rates and investment spending. Because early Keynesians saw this link as the channel through which changes in the money supply affect aggregate demand, finding that the link was weak also led them to the conclusion that changes in the money supply have no effect on aggregate output.
3. Surveys of businesspeople revealed that their decisions on how much to invest in new physical capital were not influenced by market interest rates. This evidence further confirmed that the link between interest rates and investment spending was weak, strengthening the conclusion that money doesn't matter. The result of this interpretation of the evidence was that most economists paid only scant attention to monetary policy until the mid-1960s.

Study Guide

Before reading about the objections that were raised against early Keynesian interpretations of the evidence, use the ideas on the disadvantages of structural model evidence to see if you can come up with some objections yourself. This will help you learn to apply the principles of evaluating evidence discussed earlier.

Objections to Early Keynesian Evidence

While Keynesian economics was reaching its ascendancy in the 1950s and 1960s, a small group of economists at the University of Chicago, led by Milton Friedman, adopted what was then the unfashionable view that money *does* matter to aggregate demand. Friedman and his disciples, who later became known as *monetarists*, objected to the early Keynesian interpretation of the evidence on the grounds that the structural model used by the early Keynesians was severely flawed. Because structural model evidence is only as good as the model it is based on, the monetarist critique of this evidence needs to be taken seriously.

In 1963, Friedman and Anna Schwartz published their classic monetary history of the United States, which showed that contrary to the early Keynesian beliefs, monetary policy during the Great Depression was not easy; indeed, it had never been more contractionary.[1] Friedman and Schwartz documented the massive bank failures of this

[1]Milton Friedman and Anna Jacobson Schwartz, *A Monetary History of the United States, 1867–1960* (Princeton, N.J.: Princeton University Press, 1963).

period and the resulting decline in the money supply—the largest ever experienced in the United States (see Chapter 16). Hence monetary policy could explain the worst economic contraction in U.S. history, and the Great Depression could not be singled out as a period that demonstrates the ineffectiveness of monetary policy.

A Keynesian could still counter Friedman and Schwartz's argument that money was contractionary during the Great Depression by citing the low level of interest rates. But were these interest rates really so low? Referring to Figure 1 in Chapter 6, you will note that although interest rates on U.S. Treasury securities and high-grade corporate bonds were low during the Great Depression, interest rates on lower-grade bonds, such as Baa corporate bonds, rose to unprecedented high levels during the sharpest part of the contraction phase (1930–1933). By the standard of these lower-grade bonds, then, interest rates were high and monetary policy was tight.

There is a moral to this story. Although much aggregate economic analysis proceeds as though there is only *one* interest rate, we must always be aware that there are *many* interest rates, which may tell different stories. During normal times, most interest rates move in tandem, so lumping them all together and looking at one representative interest rate may not be too misleading. But that is not always so. Unusual periods (like the Great Depression), when interest rates on different securities begin to diverge, do occur. This is exactly the kind of situation in which a structural model (like the early Keynesians') that looks at only the interest rates on a low-risk security such as a U.S. Treasury bill or bond can be very misleading.

There is a second, potentially more important reason why the early Keynesian structural model's focus on nominal interest rates provides a misleading picture of the tightness of monetary policy during the Great Depression. In a period of deflation, when there is a declining price level, low *nominal* interest rates do not necessarily indicate that the cost of borrowing is low and that monetary policy is easy—in fact, the cost of borrowing could be quite high. If, for example, the public expects the price level to decline at a 10% rate, then even though nominal interest rates are at zero, the real cost of borrowing would be as high as 10%. (Recall from Chapter 4 that the real interest rate equals the nominal interest rate, 0, minus the expected rate of inflation, −10%, so the real interest rate equals $0 - (-10\%) = 10\%$.)

You can see in Figure 1 that this is exactly what happened during the Great Depression: Real interest rates on U.S. Treasury bills were far higher during the 1931–1933 contraction phase of the depression than was the case throughout the next 40 years.[2] As a result, movements of *real* interest rates indicate that, contrary to the early Keynesians' beliefs, monetary policy was extremely tight during the Great Depression. Because an important role for monetary policy during this depressed period could no longer be ruled out, most economists were forced to rethink their position regarding whether money matters.

Monetarists also objected to the early Keynesian structural model's view that a weak link between nominal interest rates and investment spending indicates that investment spending is unaffected by monetary policy. A weak link between *nominal*

[2]In the 1980s, real interest rates rose to exceedingly high levels, approaching those of the Great Depression period. Research has tried to explain this phenomenon, some of which points to monetary policy as the source of high real rates in the 1980s. For example, see Oliver J. Blanchard and Lawrence H. Summers, "Perspectives on High World Interest Rates," *Brookings Papers on Economic Activity* 2 (1984): 273–324; and John Huizinga and Frederic S. Mishkin, "Monetary Policy Regime Shifts and the Unusual Behavior of Real Interest Rates," *Carnegie-Rochester Conference Series on Public Policy* 24 (1986): 231–274.

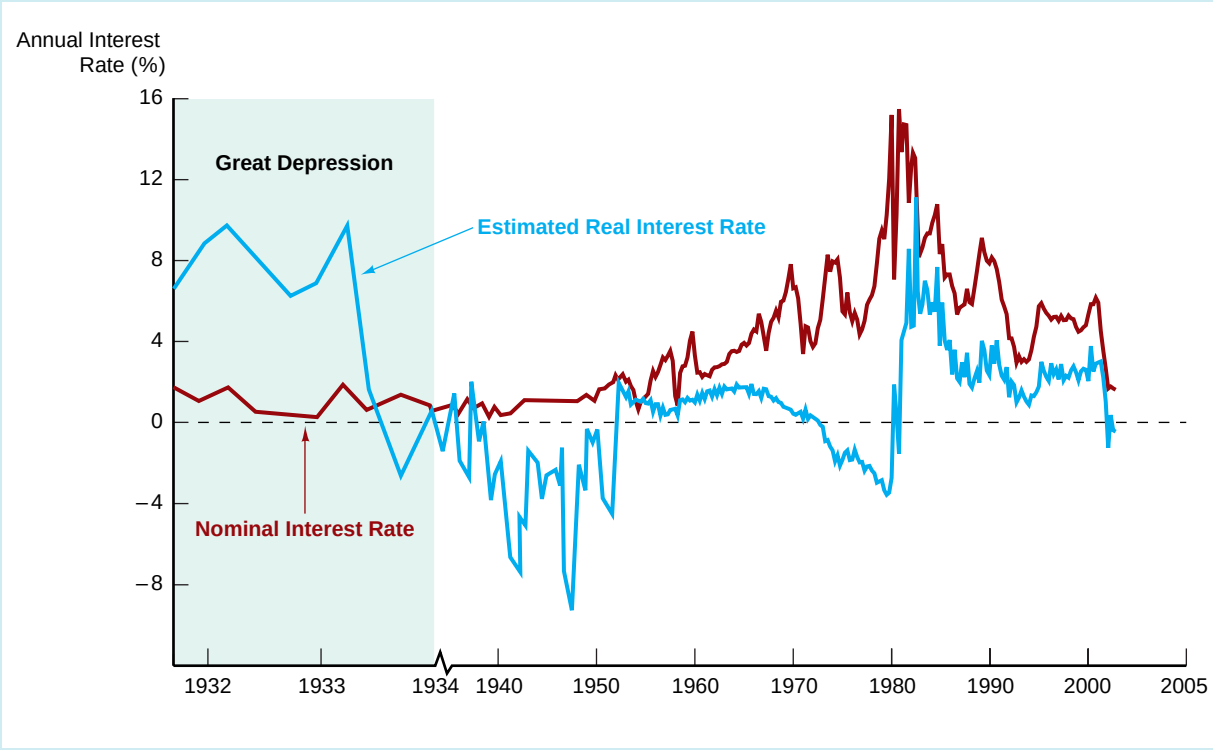


FIGURE 1 Real and Nominal Interest Rates on Three-Month Treasury Bills, 1931–2002

Sources: Nominal rates from www.federalreserve.gov/releases/h15/update/. The real rate is constructed using the procedure outlined in Frederic S. Mishkin, "The Real Interest Rate: An Empirical Investigation," *Carnegie-Rochester Conference Series on Public Policy* 15 (1981): 151–200. This involves estimating expected inflation as a function of past interest rates, inflation, and time trends and then subtracting the expected inflation measure from the nominal interest rate.

www.martincapital.com/

Click on "charts and data," then on "nominal versus real market rates" to find up-to-the-minute data showing the spread between real rates and nominal rates.

interest rates and investment spending does not rule out a strong link between *real* interest rates and investment spending. As depicted in Figure 1, nominal interest rates are often a very misleading indicator of real interest rates—not only during the Great Depression, but in later periods as well. Because real interest rates more accurately reflect the true cost of borrowing, they should be more relevant to investment decisions than nominal interest rates. Accordingly, the two pieces of early Keynesian evidence indicating that nominal interest rates have little effect on investment spending do not rule out a strong effect of changes in the money supply on investment spending and hence on aggregate demand.

Monetarists also assert that interest-rate effects on investment spending might be only one of many channels through which monetary policy affects aggregate demand. Monetary policy could then have a major impact on aggregate demand even if interest-rate effects on investment spending are small, as was suggested by the early Keynesians.

Study Guide

As you read the monetarist evidence presented in the next section, again try to think of objections to the evidence. This time use the ideas on the disadvantages of reduced-form evidence.

Early Monetarist Evidence on the Importance of Money

In the early 1960s, Milton Friedman and his followers published a series of studies based on reduced-form evidence that promoted the case for a strong effect of money on economic activity. In general, reduced-form evidence can be broken down into three categories: *timing evidence*, which looks at whether the movements in one variable typically occur before another; *statistical evidence*, which performs formal statistical tests on the correlation of the movements of one variable with another; and *historical evidence*, which examines specific past episodes to see whether movements in one variable appear to cause another. Let's look at the monetarist evidence on the importance of money that falls into each of these three categories.

Timing Evidence

Monetarist timing evidence reveals how the rate of money supply growth moves relative to the business cycle. The evidence on this relationship was first presented by Friedman and Schwartz in a famous paper published in 1963.[3] Friedman and Schwartz found that in every business cycle over nearly a century that they studied, the money growth rate always turned down before output did. On average, the peak in the rate of money growth occurred 16 months before the peak in the level of output. However, this lead time could vary, ranging from a few months to more than two years. The conclusion that these authors reached on the basis of this evidence is that money growth causes business cycle fluctuations, but its effect on the business cycle operates with "long and variable lags."

Timing evidence is based on the philosophical principle first stated in Latin as *post hoc, ergo propter hoc*, which means that if one event occurs after another, the second event must have been caused by the first. This principle is valid only if we know that the first event is an *exogenous* event, an event occurring as a result of an independent action that could not possibly be caused by the event following it or by some outside factor that might affect both events. If the first event is exogenous, when the second event follows the first we can be more confident that the first event is causing the second.

An example of an exogenous event is a controlled experiment. A chemist mixes two chemicals; suddenly his lab blows up and he with it. We can be absolutely sure that the cause of his demise was the act of mixing the two chemicals together. The principle of *post hoc, ergo propter hoc* is extremely useful in scientific experimentation.

Unfortunately, economics does not enjoy the precision of hard sciences like physics or chemistry. Often we cannot be sure that an economic event, such as a decline in the rate of money growth, is an exogenous event—it could have been caused, itself, by an outside factor or by the event it is supposedly causing. When another event (such as a decline in output) typically follows the first event (a decline in money growth), we cannot conclude with certainty that one caused the other. Timing evidence is clearly of a reduced-form nature because it looks directly at the relationship of the movements of two variables. Money growth could lead output, or both could be driven by an outside factor.

Because timing evidence is of a reduced-form nature, there is also the possibility of reverse causation, in which output growth causes money growth. How can this

[3] Milton Friedman and Anna Jacobson Schwartz, "Money and Business Cycles," *Review of Economics and Statistics* 45, Suppl. (1963): 32–64.

reverse causation occur while money growth still leads output? There are several ways in which this can happen, but we will deal with just one example.[4]

Suppose that you are in a hypothetical economy with a very regular business cycle movement, plotted in panel (a) of Figure 2, that is four years long (four years from peak to peak). Let's assume that in our hypothetical economy, there is reverse causation from output to the money supply, so movements in the money supply and output are perfectly correlated; that is, the money supply M and output Y move upward and downward at the same time. The result is that the peaks and troughs of the M and Y series in panels (a) and (b) occur at exactly the same time; therefore, no lead or lag relationship exists between them.

Now let's construct the rate of money supply growth from the money supply series in panel (b). This is done in panel (c). What is the rate of growth of the money supply at its peaks in years 1 and 5? At these points, it is not growing at all; the rate of growth is zero. Similarly, at the trough in year 3, the growth rate is zero. When the money supply is declining from its peak in year 1 to its trough in year 3, it has a negative growth rate, and its decline is fastest sometime between years 1 and 3 (year 2). Translating to panel (c), the rate of money growth is below zero from years 1 to 3, with its most negative value reached at year 2. By similar reasoning, you can see that the growth rate of money is positive in years 0 to 1 and 3 to 5, with the highest values reached in years 0 and 4. When we connect all these points together, we get the money growth series in panel (c), in which the peaks are at years 0 and 4, with a trough in year 2.

Now let's look at the relationship of the money growth series of panel (c) with the level of output in panel (a). As you can see, the money growth series consistently has its peaks and troughs exactly one year before the peaks and troughs of the output series. We conclude that in our hypothetical economy, the rate of money growth always decreases one year before output does. This evidence does not, however, imply that money growth *drives* output. In fact, by assumption, we know that this economy is one in which causation actually runs from output to the level of money supply, and there is no lead or lag relationship between the two. Only by our judicious choice of using the *growth rate* of the money supply rather than its *level* have we found a leading relationship.

This example shows how easy it is to misinterpret timing relationships. Furthermore, by searching for what we hope to find, we might focus on a variable, such as a growth rate, rather than a level, which suggests a misleading relationship. Timing evidence can be a dangerous tool for deciding on causation.

Stated even more forcefully, "one person's lead is another person's lag." For example, you could just as easily interpret the relationship of money growth and output in Figure 2 to say that the money growth rate lags output by three years—after all, the peaks in the money growth series occur three years after the peaks in the output series. In short, you could say that output leads money growth.

www.economagic.com/bci_97.htm
A site with extensive data on the factors that define business cycles.

We have seen that timing evidence is extremely hard to interpret. Unless we can be sure that changes in the leading variable are exogenous events, we cannot be sure that the leading variable is actually causing the following variable. And it is all too easy to

[4] A famous article by James Tobin, "Money and Income: *Post Hoc, Ergo Propter Hoc,*" *Quarterly Journal of Economics* 84 (1970): 301–317, describes an economic system in which changes in aggregate output cause changes in the growth rate of money but changes in the growth rate of money have no effect on output. Tobin shows that such a system with reverse causation could yield timing evidence similar to that found by Friedman and Schwartz.

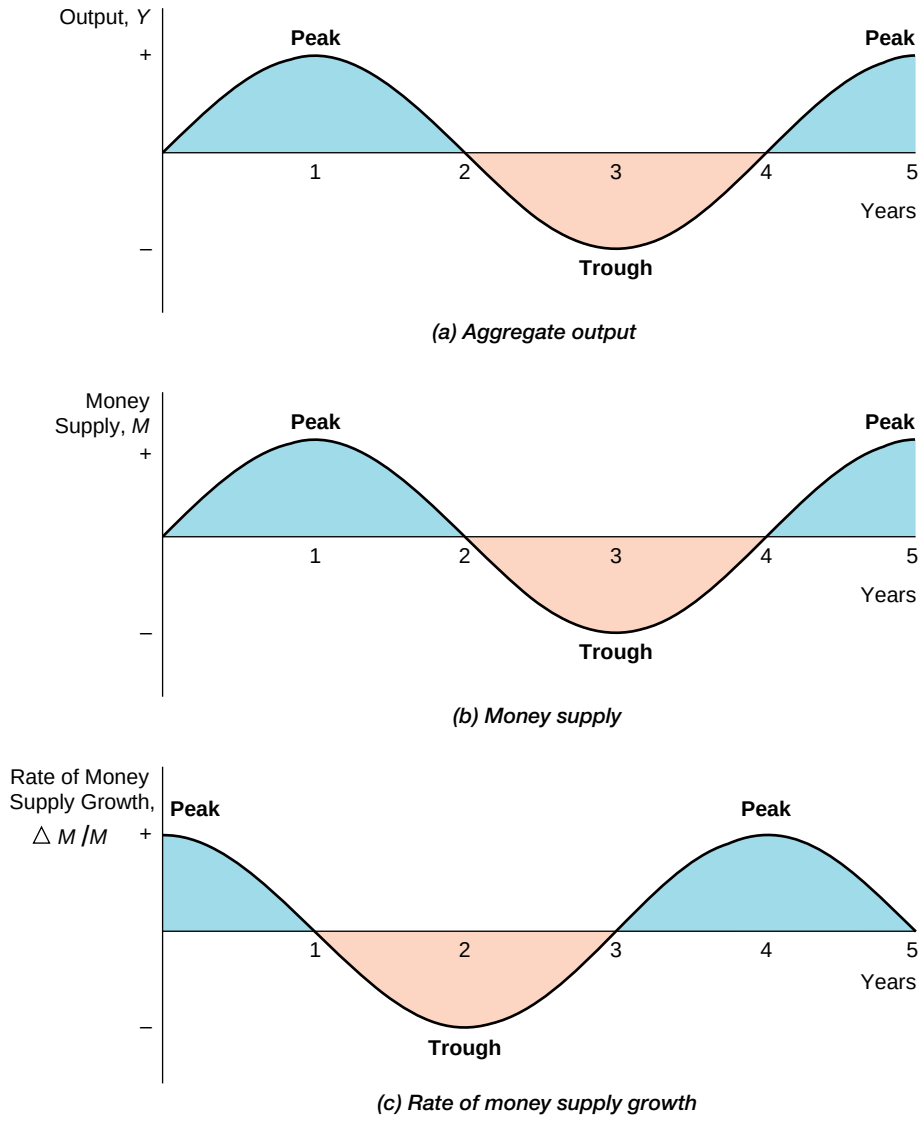


FIGURE 2 Hypothetical Example in Which Money Growth Leads Output

Although neither M nor Y leads the other (that is, their peaks and troughs coincide), $\Delta M/M$ has its peaks and troughs one year ahead of M and Y, thus leading both series. (Note that M and Y in the panels are drawn as movements around a positive average value; a plus sign indicates a value above the average, and a minus sign indicates a value below the average, not a negative value.)

find what we seek when looking for timing evidence. Perhaps the best way of describing this danger is to say that "timing evidence may be in the eyes of the beholder."

Statistical Evidence

Monetarist statistical evidence examines the correlations between money and aggregate output or aggregate spending by performing formal statistical tests. Again in

1963 (obviously a vintage year for the monetarists), Milton Friedman and David Meiselman published a paper that proposed the following test of a monetarist model against a Keynesian model.[5] In the Keynesian framework, investment and government spending are sources of fluctuations in aggregate demand, so Friedman and Meiselman constructed a "Keynesian" autonomous expenditure variable A equal to investment spending plus government spending. They characterized the Keynesian model as saying that A should be highly correlated with aggregate spending Y, while the money supply M should not. In the monetarist model, the money supply is the source of fluctuations in aggregate spending, and M should be highly correlated with Y, while A should not.

A logical way to find out which model is better would be to see which is more highly correlated with Y: M or A. When Friedman and Meiselman conducted this test for many different periods of U.S. data, they discovered that *the monetarist model wins!*[6] They concluded that monetarist analysis gives a better description than Keynesian analysis of how aggregate spending is determined.

Several objections were raised against the Friedman-Meiselman evidence:

1. The standard criticisms of this reduced-form evidence are the ones we have already discussed: Reverse causation could occur, or an outside factor might drive both series.

2. The test may not be fair because the Keynesian model is characterized too simplistically. Keynesian structural models commonly include hundreds of equations. The one-equation Keynesian model that Friedman-Meiselman tested may not adequately capture the effects of autonomous expenditure. Furthermore, Keynesian models usually include the effects of other variables. By ignoring them, the effect of monetary policy might be overestimated and the effect of autonomous expenditure underestimated.

3. The Friedman-Meiselman measure of autonomous expenditure A might be constructed poorly, preventing the Keynesian model from performing well. For example, orders for military hardware affect aggregate demand before they appear as spending in the autonomous expenditure variable that Friedman and Meiselman used. A more careful construction of the autonomous expenditure variable should take account of the placing of orders for military hardware. When the autonomous expenditure variable was constructed more carefully by critics of the Friedman-Meiselman study, they found that the results were reversed: The Keynesian model won.[7] A more recent postmortem on the appropriateness of various ways of determining autonomous expenditure does not give a clear-cut victory to either the Keynesian or the monetarist model.[8]

---

[5]Milton Friedman and David Meiselman, "The Relative Stability of Monetary Velocity and the Investment Multiplier," in *Stabilization Policies,* ed. Commission on Money and Credit (Upper Saddle River, N.J.: Prentice-Hall, 1963), pp. 165–268.

[6]Friedman and Meiselman did not actually run their tests using the Y variable because they felt that this gave an unfair advantage to the Keynesian model in that A is included in Y. Instead, they subtracted A from Y and tested for the correlation of $Y - A$ with M or A.

[7]See, for example, Albert Ando and Franco Modigliani, "The Relative Stability of Monetary Velocity and the Investment Multiplier," *American Economic Review* 55 (1965): 693–728.

[8]See William Poole and Edith Kornblith, "The Friedman-Meiselman CMC Paper: New Evidence on an Old Controversy," *American Economic Review* 63 (1973): 908–917.

Historical Evidence

The monetarist historical evidence found in Friedman and Schwartz's *A Monetary History*, has been very influential in gaining support for the monetarist position. We have already seen that the book was extremely important as a criticism of early Keynesian thinking, showing as it did that the Great Depression was not a period of easy monetary policy and that the depression could be attributed to the sharp decline in the money supply from 1930 to 1933 resulting from bank panics. In addition, the book documents in great detail that the growth rate of money leads business cycles, because it declines before every recession. This timing evidence is, of course, subject to all the criticisms raised earlier.

The historical evidence contains one feature, however, that makes it different from other monetarist evidence we have discussed so far. Several episodes occur in which changes in the money supply appear to be exogenous events. These episodes are almost like controlled experiments, so the *post hoc, ergo propter hoc* principle is far more likely to be valid: If the decline in the growth rate of the money supply is soon followed by a decline in output in these episodes, much stronger evidence is presented that money growth is the driving force behind the business cycle.

One of the best examples of such an episode is the increase in reserve requirements in 1936–1937 (discussed in Chapter 18), which led to a sharp decline in the money supply and in its rate of growth. The increase in reserve requirements was implemented because the Federal Reserve wanted to improve its control of monetary policy; it was not implemented in response to economic conditions. We can thus rule out reverse causation from output to the money supply. Also, it is hard to think of an outside factor that could have driven the Fed to increase reserve requirements and that could also have directly affected output. Therefore, the decline in the money supply in this episode can probably be classified as an exogenous event with the characteristics of a controlled experiment. Soon after this experiment, the very severe recession of 1937–1938 occurred. We can conclude with confidence that in this episode, the change in the money supply due to the Fed's increase in reserve requirements was indeed the source of the business cycle contraction that followed.

A Monetary History also documents other historical episodes, such as the bank panic of 1907 and other years in which the decline in money growth again appears to have been an exogenous event. The fact that recessions have frequently followed apparently exogenous declines in money growth is very strong evidence that changes in the growth rate of the money supply do have an impact on aggregate output. Recent work by Christina and David Romer, both of the University of California, Berkeley, applies the historical approach to more recent data using more sophisticated statistical techniques and also finds that monetary policy shifts have had an important impact on the aggregate economy.[9]

Overview of the Monetarist Evidence

Where does this discussion of the monetarist evidence leave us? We have seen that because of reverse causation and outside-factor possibilities, there are some serious doubts about the conclusions that can be drawn from timing and statistical evidence alone. However, some of the historical evidence in which exogenous declines in

[9] Christina Romer and David Romer, "Does Monetary Policy Matter? A New Test in the Spirit of Friedman and Schwartz," *NBER Macroeconomics Annual, 1989,* 4, ed. Stanley Fischer (Cambridge, Mass.: M.I.T. Press, 1989), 121–170.

Box 3

Real Business Cycle Theory and the Debate on Money and Economic Activity

New entrants to the debate on money and economic activity are advocates of *real business cycle theory*, which states that real shocks to tastes and technology (rather than monetary shocks) are the driving forces behind business cycles. Proponents of this theory are critical of the monetarist view that money matters to business cycles because they believe that the correlation of output with money reflects reverse causation; that is, the business cycle drives money, rather than the other way around. An important piece of evidence they offer to support the reverse causation argument is that almost none of the correlation between money and output comes from the monetary base, which is controlled by the monetary authorities.* Instead, the money–output correlation stems from other sources of money supply movements that, as we saw in Chapters 15 and 16, are affected by the actions of banks, depositors, and borrowers from banks and are more likely to be influenced by the business cycle.

*Robert King and Charles Plosser, "Money, Credit and Prices in a Real Business Cycle," *American Economic Review* 74 (1984): 363–380; Charles Plosser, "Understanding Real Business Cycles," *Journal of Economic Perspectives* 3 (Summer 1989): 51–78.

money growth are followed by business cycle contractions does provide stronger support for the monetarist position. When historical evidence is combined with timing and statistical evidence, the conclusion that money does matter seems warranted.

As you can imagine, the economics profession was quite shaken by the appearance of the monetarist evidence, as up to that time most economists believed that money does not matter at all. Monetarists had demonstrated that this early Keynesian position was probably wrong, and it won them a lot of converts. Recognizing the fallacy of the position that money does not matter does not necessarily mean that we must accept the position that money is *all* that matters. Many Keynesian economists shifted their views toward the monetarist position, but not all the way. Instead, they adopted an intermediate position compatible with the Keynesian aggregate supply and demand analysis described in Chapter 25: They allowed that money, fiscal policy, net exports, and "animal spirits" all contributed to fluctuations in aggregate demand. The result has been a convergence of the Keynesian and monetarist views on the importance of money to economic activity. However, proponents of a new theory of aggregate fluctuations called *real business cycle theory* are more critical of the monetarist reduced-form evidence that money is important to business cycle fluctuations because they believe there is reverse causation from the business cycle to money (see Box 3).

Transmission Mechanisms of Monetary Policy

After the successful monetarist attack on the early Keynesian position, economic research went in two directions. One direction was to use more sophisticated monetarist reduced-form models to test for the importance of money to economic activity.[10]

[10]The most prominent example of more sophisticated reduced-form research is the so-called St. Louis model, which was developed at the Federal Reserve Bank of St. Louis in the late 1960s and early 1970s. It provided support for the monetarist position, but is subject to the same criticisms of reduced-form evidence outlined in the text. The St. Louis model was first outlined in Leonall Andersen and Jerry Jordan, "Monetary and Fiscal Actions: A Test of Their Relative Importance in Economic Stabilization," Federal Reserve Bank of St. Louis *Review* 50 (November 1968): 11–23.

The second direction was to pursue a structural model approach and to develop a better understanding of channels (other than interest-rate effects on investment) through which monetary policy affects aggregate demand. In this section we examine some of these channels, or *transmission mechanisms*, beginning with interest-rate channels, because they are the key monetary transmission mechanism in the Keynesian *ISLM* and *AD/AS* models you have seen in Chapters 23, 24, and 25.

Traditional Interest-Rate Channels

The traditional Keynesian view of the monetary transmission mechanism can be characterized by the following schematic showing the effect of a monetary expansion:

$$M\uparrow \Rightarrow i_r\downarrow \Rightarrow I\uparrow \Rightarrow Y\uparrow \tag{1}$$

where $M\uparrow$ indicates an expansionary monetary policy leading to a fall in real interest rates ($i_r\downarrow$), which in turn lowers the cost of capital, causing a rise in investment spending ($I\uparrow$), thereby leading to an increase in aggregate demand and a rise in output ($Y\uparrow$).

Although Keynes originally emphasized this channel as operating through businesses' decisions about investment spending, the search for new monetary transmission mechanisms recognized that consumers' decisions about housing and **consumer durable expenditure** (spending by consumers on durable items such as automobiles and refrigerators) also are investment decisions. Thus the interest-rate channel of monetary transmission outlined in Equation 1 applies equally to consumer spending, in which I represents residential housing and consumer durable expenditure.

An important feature of the interest-rate transmission mechanism is its emphasis on the *real* rather than the nominal interest rate as the rate that affects consumer and business decisions. In addition, it is often the real *long*-term interest rate and not the short-term interest rate that is viewed as having the major impact on spending. How is it that changes in the short-term nominal interest rate induced by a central bank result in a corresponding change in the real interest rate on both short- and long-term bonds? The key is the phenomenon known as *sticky prices*, the fact that the aggregate price level adjusts slowly over time, meaning that expansionary monetary policy, which lowers the short-term nominal interest rate, also lowers the short-term *real* interest rate. The expectations hypothesis of the term structure described in Chapter 6, which states that the long-term interest rate is an average of expected future short-term interest rates, suggests that the lower real short-term interest rate leads to a fall in the real long-term interest rate. These lower real interest rates then lead to rises in business fixed investment, residential housing investment, inventory investment, and consumer durable expenditure, all of which produce the rise in aggregate output.

The fact that it is the real interest rate rather than the nominal rate that affects spending provides an important mechanism for how monetary policy can stimulate the economy, even if nominal interest rates hit a floor of zero during a deflationary episode. With nominal interest rates at a floor of zero, an expansion in the money supply ($M\uparrow$) can raise the expected price level ($P^e\uparrow$) and hence expected inflation ($\pi^e\uparrow$), thereby lowering the real interest rate ($i_r = [i - \pi^e]\downarrow$) even when the nominal interest rate is fixed at zero and stimulating spending through the interest-rate channel:

$$M\uparrow \Rightarrow P^e\uparrow \Rightarrow \pi^e\uparrow \Rightarrow i_r\downarrow \Rightarrow I\uparrow \Rightarrow Y\uparrow \tag{2}$$

This mechanism thus indicates that monetary policy can still be effective even when nominal interest rates have already been driven down to zero by the monetary authorities. Indeed, this mechanism is a key element in monetarist discussions of why the U.S. economy was not stuck in a liquidity trap (in which increases in the money supply

might be unable to lower interest rates, discussed in Chapter 22) during the Great Depression and why expansionary monetary policy could have prevented the sharp decline in output during that period.

Some economists, such as John Taylor of Stanford University, take the position that there is strong empirical evidence for substantial interest-rate effects on consumer and investment spending through the cost of capital, making the interest-rate monetary transmission mechanism a strong one. His position is highly controversial, and many researchers, including Ben Bernanke of Princeton University and Mark Gertler of New York University, believe that the empirical evidence does not support strong interest-rate effects operating through the cost of capital.[11] Indeed, these researchers see the empirical failure of traditional interest-rate monetary transmission mechanisms as having provided the stimulus for the search for other transmission mechanisms of monetary policy.

These other transmission mechanisms fall into two basic categories: those operating through asset prices other than interest rates and those operating through asymmetric information effects on credit markets (the so-called **credit view**). (All these mechanisms are summarized in the schematic diagram in Figure 3.)

Other Asset Price Channels

As we have seen earlier in the chapter, a key monetarist objection to the Keynesian analysis of monetary policy effects on the economy is that it focuses on only one asset price, the interest rate, rather than on many asset prices. Monetarists envision a transmission mechanism in which other relative asset prices and real wealth transmit monetary effects onto the economy. In addition to bond prices, two other asset prices receive substantial attention as channels for monetary policy effects: foreign exchange and equities (stocks).

Exchange Rate Effects on Net Exports. With the growing internationalization of economies throughout the world and the advent of flexible exchange rates, more attention has been paid to how monetary policy affects exchange rates, which in turn affect net exports and aggregate output.

This channel also involves interest-rate effects, because, as we have seen in Chapter 19, when domestic real interest rates fall, domestic dollar deposits become less attractive relative to deposits denominated in foreign currencies. As a result, the value of dollar deposits relative to other currency deposits falls, and the dollar depreciates (denoted by $E\downarrow$). The lower value of the domestic currency makes domestic goods cheaper than foreign goods, thereby causing a rise in net exports ($NX\uparrow$) and hence in aggregate output ($Y\uparrow$). The schematic for the monetary transmission mechanism that operates through the exchange rate is:

$$M\uparrow \Rightarrow i_r\downarrow \Rightarrow E\downarrow \Rightarrow NX\uparrow \Rightarrow Y\uparrow \tag{3}$$

Recent research has found that this exchange rate channel plays an important role in how monetary policy affects the domestic economy.[12]

---

[11] See John Taylor, "The Monetary Transmission Mechanism: An Empirical Framework," *Journal of Economic Perspectives* 9 (Fall 1995): 11–26, and Ben Bernanke and Mark Gertler, "Inside the Black Box: The Credit Channel of Monetary Policy Transmission," *Journal of Economic Perspectives* 9 (Fall 1995): 27–48.

[12] For example, see Ralph Bryant, Peter Hooper, and Catherine Mann, *Evaluating Policy Regimes: New Empirical Research in Empirical Macroeconomics* (Washington, D.C.: Brookings Institution, 1993), and John B. Taylor, *Macroeconomic Policy in a World Economy: From Econometric Design to Practical Operation* (New York: Norton, 1993).

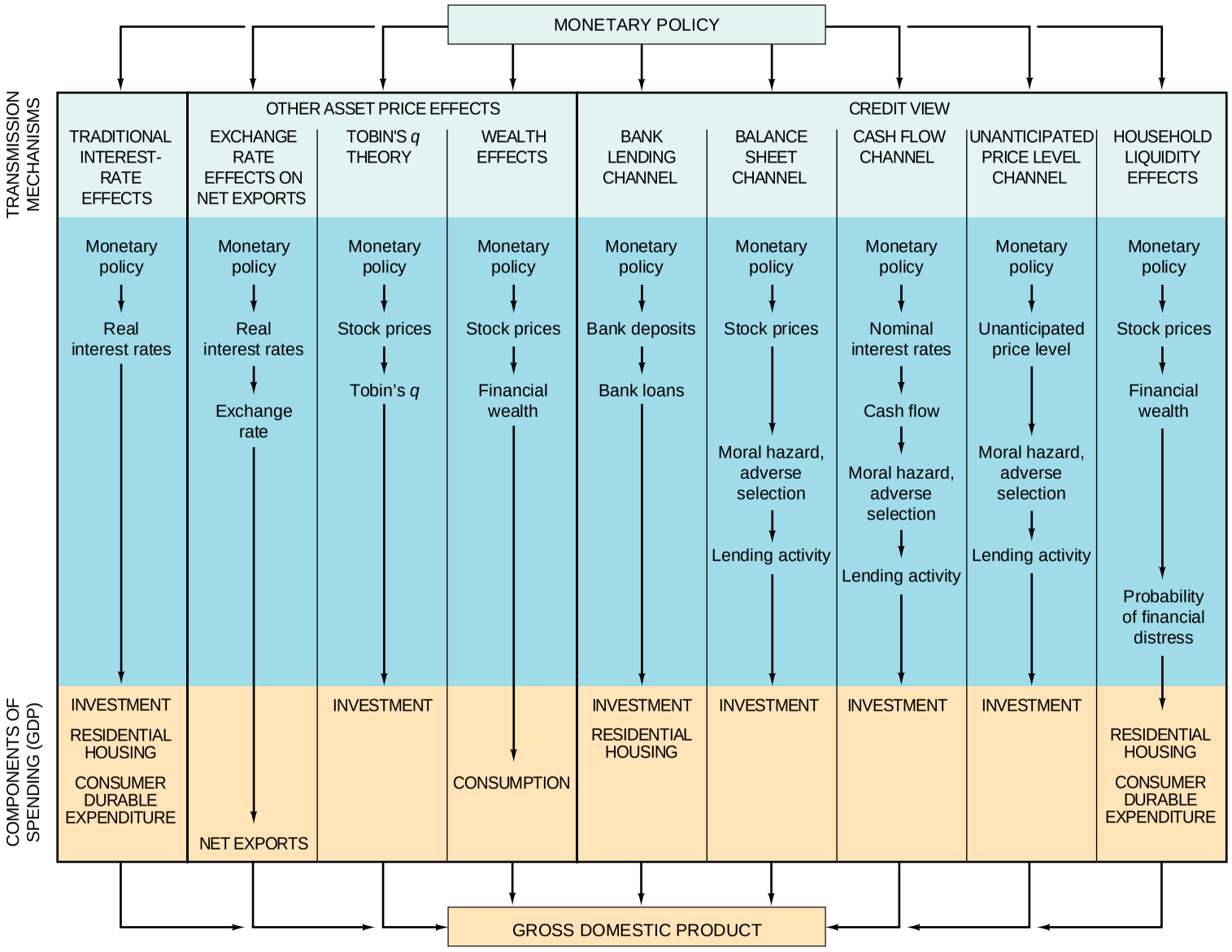


FIGURE 3 The Link Between Monetary Policy and GDP: Monetary Transmission Mechanisms

Tobin's *q* Theory. James Tobin developed a theory, referred to as *Tobin's q Theory*, that explains how monetary policy can affect the economy through its effects on the valuation of equities (stock). Tobin defines q as the market value of firms divided by the replacement cost of capital. If q is high, the market price of firms is high relative to the replacement cost of capital, and new plant and equipment capital is cheap relative to the market value of firms. Companies can then issue stock and get a high price for it relative to the cost of the facilities and equipment they are buying. Investment spending will rise, because firms can buy a lot of new investment goods with only a small issue of stock.

Conversely, when q is low, firms will not purchase *new* investment goods because the market value of firms is low relative to the cost of capital. If companies want to acquire capital when q is low, they can buy another firm cheaply and acquire old capital instead. Investment spending, the purchase of new investment goods, will then be very low. Tobin's q theory gives a good explanation for the extremely low rate of investment spending during the Great Depression. In that period, stock prices collapsed, and by 1933, stocks were worth only one-tenth of their value in late 1929; q fell to unprecedented low levels.

The crux of this discussion is that a link exists between Tobin's q and investment spending. But how might monetary policy affect stock prices? Quite simply, when monetary policy is expansionary, the public finds that it has more money than it wants and so gets rid of it through spending. One place the public spends is in the stock market, increasing the demand for stocks and consequently raising their prices.[13] Combining this with the fact that higher stock prices (P_s) will lead to a higher q and thus higher investment spending I leads to the following transmission mechanism of monetary policy:[14]

$$M\uparrow \Rightarrow P_s\uparrow \Rightarrow q\uparrow \Rightarrow I\uparrow \Rightarrow Y\uparrow \tag{4}$$

Wealth Effects. In their search for new monetary transmission mechanisms, researchers also looked at how consumers' balance sheets might affect their spending decisions. Franco Modigliani was the first to take this tack, using his famous life cycle hypothesis of consumption. **Consumption** is spending by consumers on nondurable goods and services.[15] It differs from *consumer expenditure* in that it does not include spending on consumer durables. The basic premise of Modigliani's theory is that consumers smooth out their consumption over time. Therefore, what determines consumption spending is the lifetime resources of consumers, not just today's income.

---

[13]See James Tobin, "A General Equilibrium Approach to Monetary Theory," *Journal of Money, Credit, and Banking* 1 (1969): 15–29. A somewhat more Keynesian story with the same outcome is that the increase in the money supply lowers interest rates on bonds so that the yields on alternatives to stocks fall. This makes stocks more attractive relative to bonds, so demand for them increases, raises their price, and thereby lowers their yield.

[14]An alternative way of looking at the link between stock prices and investment spending is that higher stock prices lower the yield on stocks and reduce the cost of financing investment spending through issuing equity. This way of looking at the link between stock prices and investment spending is formally equivalent to Tobin's q approach; see Barry Bosworth, "The Stock Market and the Economy," *Brookings Papers on Economic Activity* 2 (1975): 257–290.

[15]Consumption also includes another small component, the services that a consumer receives from the ownership of housing and consumer durables.

An important component of consumers' lifetime resources is their financial wealth, a major component of which is common stocks. When stock prices rise, the value of financial wealth increases, thereby increasing the lifetime resources of consumers, and consumption should rise. Considering that, as we have seen, expansionary monetary policy can lead to a rise in stock prices, we now have another monetary transmission mechanism:

$$M\uparrow \Rightarrow P_s\uparrow \Rightarrow \text{wealth} \uparrow \Rightarrow \text{consumption} \uparrow \Rightarrow Y\uparrow \tag{5}$$

Modigliani's research found this relationship to be an extremely powerful mechanism that adds substantially to the potency of monetary policy.[16]

The wealth and Tobin's q channels allow for a general definition of equity, so the Tobin q framework can also be applied to the housing market, where housing is equity. An increase in house prices, which raises their prices relative to replacement cost, leads to a rise in Tobin's q for housing, thereby stimulating its production. Similarly, housing and land prices are extremely important components of wealth, and so rises in these prices increase wealth, thereby raising consumption. Monetary expansion, which raises land and housing prices through the Tobin's q and wealth mechanisms described here, thus leads to a rise in aggregate demand.

Credit View

Dissatisfaction with the conventional stories that interest-rate effects explain the impact of monetary policy on expenditures on durable assets has led to a new explanation based on the problem of asymmetric information in financial markets (see Chapter 8). This explanation, referred to as the *credit view*, proposes that two types of monetary transmission channels arise as a result of information problems in credit markets: those that operate through effects on bank lending and those that operate through effects on firms' and households' balance sheets.[17]

Bank Lending Channel. The bank lending channel is based on the analysis in Chapter 8, which demonstrated that banks play a special role in the financial system because they are especially well suited to solve asymmetric information problems in credit markets. Because of banks' special role, certain borrowers will not have access to the credit markets unless they borrow from banks. As long as there is no perfect substitutability of retail bank deposits with other sources of funds, the bank lending channel of monetary transmission operates as follows. Expansionary monetary policy, which increases bank reserves and bank deposits, increases the quantity of bank loans available. Because many borrowers are dependent on bank loans to finance their activities, this increase in loans will cause investment (and possibly consumer) spending to rise. Schematically, the monetary policy effect is:

$$M\uparrow \Rightarrow \text{bank deposits} \uparrow \Rightarrow \text{bank loans} \uparrow \Rightarrow I\uparrow \Rightarrow Y\uparrow \tag{6}$$

---

[16]See Franco Modigliani, "Monetary Policy and Consumption," in *Consumer Spending and Money Policy: The Linkages* (Boston: Federal Reserve Bank, 1971), pp. 9–84.

[17]Surveys of the credit view can be found in Ben Bernanke, "Credit in the Macroeconomy," Federal Reserve Bank of New York *Quarterly Review*, Spring 1993, pp. 50–70; Ben Bernanke and Mark Gertler, "Inside the Black Box: The Credit Channel of Monetary Policy Transmission," *Journal of Economic Perspectives* 9 (Fall 1995): 27–48; Stephen G. Cecchetti, "Distinguishing Theories of the Monetary Transmission Mechanism," Federal Reserve Bank of St. Louis *Review* 77 (May–June 1995): 83–97; and R. Glenn Hubbard, "Is There a 'Credit Channel' for Monetary Policy?" Federal Reserve Bank of St. Louis *Review* 77 (May–June 1995): 63–74.

An important implication of the credit view is that monetary policy will have a greater effect on expenditure by smaller firms, which are more dependent on bank loans, than it will on large firms, which can access the credit markets directly through stock and bond markets (and not only through banks).

Though this result has been confirmed by researchers, doubts about the bank lending channel have been raised in the literature, and there are reasons to suspect that the bank lending channel in the United States may not be as powerful as it once was.[18] The first reason this channel is not as powerful is that the current U.S. regulatory framework no longer imposes restrictions on banks that hinder their ability to raise funds (see Chapter 9). Prior to the mid-1980s, certificates of deposit (CDs) were subjected to reserve requirements and Regulation Q deposit rate ceilings, which made it hard for banks to replace deposits that flowed out of the banking system during a monetary contraction. With these regulatory restrictions abolished, banks can more easily respond to a decline in bank reserves and a loss of retail deposits by issuing CDs at market interest rates that do not have to be backed up by required reserves. Second, the worldwide decline of the traditional bank lending business (see Chapter 10) has rendered the bank lending channel less potent. Nonetheless, many economists believe that the bank lending channel played an important role in the slow recovery in the U.S. from the 1990–91 recession.

Balance Sheet Channel. Even though the bank lending channel may be declining in importance, it is by no means clear that this is the case for the other credit channel, the balance sheet channel. Like the bank lending channel, the balance sheet channel also arises from the presence of asymmetric information problems in credit markets. In Chapter 8, we saw that the lower the net worth of business firms, the more severe the adverse selection and moral hazard problems in lending to these firms. Lower net worth means that lenders in effect have less collateral for their loans, and so potential losses from adverse selection are higher. A decline in net worth, which raises the adverse selection problem, thus leads to decreased lending to finance investment spending. The lower net worth of businesses also increases the moral hazard problem because it means that owners have a lower equity stake in their firms, giving them more incentive to engage in risky investment projects. Since taking on riskier investment projects makes it more likely that lenders will not be paid back, a decrease in businesses' net worth leads to a decrease in lending and hence in investment spending.

Monetary policy can affect firms' balance sheets in several ways. Expansionary monetary policy ($M\uparrow$), which causes a rise in stock prices ($P_s\uparrow$) along lines described earlier, raises the net worth of firms and so leads to higher investment spending ($I\uparrow$) and aggregate demand ($Y\uparrow$) because of the decrease in adverse selection and moral hazard problems. This leads to the following schematic for one balance sheet channel of monetary transmission:

$$M\uparrow \Rightarrow P_s\uparrow \Rightarrow \text{adverse selection}\downarrow, \text{moral hazard}\downarrow \Rightarrow \text{lending}\uparrow \Rightarrow I\uparrow \Rightarrow Y\uparrow \qquad (7)$$

Cash Flow Channel. Another balance sheet channel operates through its effects on *cash flow*, the difference between cash receipts and cash expenditures. Expansionary

[18]For example, see Valerie Ramey, "How Important Is the Credit Channel in the Transmission of Monetary Policy?" *Carnegie-Rochester Conference Series on Public Policy* 39 (1993): 1–45, and Allan H. Meltzer, "Monetary, Credit (and Other) Transmission Processes: A Monetarist Perspective," *Journal of Economic Perspectives* 9 (Fall 1995): 49–72.

monetary policy, which lowers nominal interest rates, also causes an improvement in firms' balance sheets because it raises cash flow. The rise in cash flow causes an improvement in the balance sheet because it increases the liquidity of the firm (or household) and thus makes it easier for lenders to know whether the firm (or household) will be able to pay its bills. The result is that adverse selection and moral hazard problems become less severe, leading to an increase in lending and economic activity. The following schematic describes this additional balance sheet channel:

$$M\uparrow \Rightarrow i\downarrow \Rightarrow \text{cash flow}\uparrow \Rightarrow \text{adverse selection}\downarrow, \text{moral hazard}\downarrow \Rightarrow \text{lending}\uparrow \Rightarrow I\uparrow \Rightarrow Y\uparrow \quad (8)$$

An important feature of this transmission mechanism is that it is *nominal* interest rates that affect firms' cash flow. Thus this interest-rate mechanism differs from the traditional interest-rate mechanism discussed earlier, in which it is the real rather than the nominal interest rate that affects investment. Furthermore, the short-term interest rate plays a special role in this transmission mechanism, because it is interest payments on short-term rather than long-term debt that typically have the greatest impact on households' and firms' cash flow.

A related mechanism involving adverse selection through which expansionary monetary policy that lowers interest rates can stimulate aggregate output involves the credit-rationing phenomenon. As we discussed in Chapter 9, credit rationing occurs in cases where borrowers are denied loans even when they are willing to pay a higher interest rate. This is because individuals and firms with the riskiest investment projects are exactly the ones who are willing to pay the highest interest rates, for if the high-risk investment succeeds, they will be the primary beneficiaries. Thus higher interest rates increase the adverse selection problem, and lower interest rates reduce it. When expansionary monetary policy lowers interest rates, less risk-prone borrowers make up a higher fraction of those demanding loans, and so lenders are more willing to lend, raising both investment and output, along the lines of parts of the schematic in Equation 8.

Unanticipated Price Level Channel. A third balance sheet channel operates through monetary policy effects on the general price level. Because in industrialized countries debt payments are contractually fixed in nominal terms, an unanticipated rise in the price level lowers the value of firms' liabilities in real terms (decreases the burden of the debt) but should not lower the real value of the firms' assets. Monetary expansion that leads to an unanticipated rise in the price level ($P\uparrow$) therefore raises real net worth, which lowers adverse selection and moral hazard problems, thereby leading to a rise in investment spending and aggregate output as in the following schematic:

$$M\uparrow \Rightarrow \text{unanticipated } P\uparrow \Rightarrow \text{adverse selection}\downarrow, \text{moral hazard}\downarrow \Rightarrow \text{lending}\uparrow \Rightarrow I\uparrow \Rightarrow Y\uparrow \quad (9)$$

The view that unanticipated movements in the price level have important effects on aggregate demand has a long tradition in economics: It is the key feature in the debt-deflation view of the Great Depression we outlined in Chapter 8.

Household Liquidity Effects. Although most of the literature on the credit channel focuses on spending by businesses, the credit view should apply equally well to consumer spending, particularly on consumer durables and housing. Declines in bank

Box 4

Consumers' Balance Sheets and the Great Depression

The years between 1929 and 1933 witnessed the worst deterioration in consumers' balance sheets ever seen in the United States. The stock market crash in 1929, which caused a slump that lasted until 1933, reduced the value of consumers' wealth by $692 billion (in 1996 dollars), and as expected, consumption dropped sharply (by over $100 billion). Because of the decline in the price level in that period, the level of real debt consumers owed also increased sharply (by over 20%). Consequently, the value of financial assets relative to the amount of debt declined sharply, increasing the likelihood of financial distress. Not surprisingly, spending on consumer durables and housing fell precipitously: From 1929 to 1933, consumer durable expenditure declined by over 50%, while expenditure on housing declined by 80%.*

*For further discussion of the effect of consumers' balance sheets on spending during the Great Depression, see Frederic S. Mishkin, "The Household Balance Sheet and the Great Depression," *Journal of Economic History* 38 (1978): 918–937.

lending induced by a monetary contraction should cause a decline in durables and housing purchases by consumers who do not have access to other sources of credit. Similarly, increases in interest rates cause a deterioration in household balance sheets, because consumers' cash flow is adversely affected.

Another way of looking at how the balance sheet channel may operate through consumers is to consider liquidity effects on consumer durable and housing expenditures—found to have been important factors during the Great Depression (see Box 4). In the liquidity effects view, balance sheet effects work through their impact on consumers' desire to spend rather than on lenders' desire to lend. Because of asymmetric information about their quality, consumer durables and housing are very illiquid assets. If, as a result of a bad income shock, consumers needed to sell their consumer durables or housing to raise money, they would expect a big loss because they could not get the full value of these assets in a distress sale. (This is just a manifestation of the lemons problem described in Chapter 8.) In contrast, if consumers held financial assets (such as money in the bank, stocks, or bonds), they could easily sell them quickly for their full market value and raise the cash. Hence if consumers expect a higher likelihood of finding themselves in financial distress, they would rather be holding fewer illiquid consumer durable or housing assets and more liquid financial assets.

A consumer's balance sheet should be an important influence on his or her estimate of the likelihood of suffering financial distress. Specifically, when consumers have a large amount of financial assets relative to their debts, their estimate of the probability of financial distress is low, and they will be more willing to purchase consumer durables or housing. When stock prices rise, the value of financial assets rises as well; consumer durable expenditure will also rise because consumers have a more secure financial position and a lower estimate of the likelihood of suffering financial distress. This leads to another transmission mechanism for monetary policy, operating through the link between money and stock prices:[19]

$$\begin{aligned} M\uparrow \Rightarrow P_s\uparrow \Rightarrow \text{financial assets} \uparrow \Rightarrow \text{likelihood of financial distress} \downarrow \\ \Rightarrow \text{consumer durable and housing expenditure} \uparrow \Rightarrow Y\uparrow \end{aligned} \tag{10}$$

[19]See Frederic S. Mishkin, "What Depressed the Consumer? The Household Balance Sheet and the 1973–1975 Recession," *Brookings Papers on Economic Activity* 1 (1977): 123–164.

The illiquidity of consumer durable and housing assets provides another reason why a monetary expansion, which lowers interest rates and thereby raises cash flow to consumers, leads to a rise in spending on consumer durables and housing. A rise in consumer cash flow decreases the likelihood of financial distress, which increases the desire of consumers to hold durable goods or housing, thus increasing spending on them and hence aggregate output. The only difference between this view of cash flow effects and that outlined in Equation 8 is that it is not the willingness of lenders to lend to consumers that causes expenditure to rise but the willingness of consumers to spend.

Why Are Credit Channels Likely to Be Important?

There are three reasons to believe that credit channels are important monetary transmission mechanisms. First, a large body of evidence on the behavior of individual firms supports the view that credit market imperfections of the type crucial to the operation of credit channels do affect firms' employment and spending decisions.[20] Second, there is evidence that small firms (which are more likely to be credit-constrained) are hurt more by tight monetary policy than large firms, which are unlikely to be credit-constrained.[21] Third, and maybe most compelling, the asymmetric information view of credit market imperfections at the core of the credit channel analysis is a theoretical construct that has proved useful in explaining many other important phenomena, such as why many of our financial institutions exist, why our financial system has the structure that it has, and why financial crises are so damaging to the economy (all topics discussed in Chapter 8). The best support for a theory is its demonstrated usefulness in a wide range of applications. By this standard, the asymmetric information theory supporting the existence of credit channels as an important monetary transmission mechanism has much to recommend it.

Application

Corporate Scandals and the Slow Recovery from the March 2001 Recession

The collapse of the tech boom and the stock market slump led to a decline in investment spending that triggered a recession starting in March 2001. Just as the recession got under way, the Fed rapidly lowered the federal funds rate. At first it appeared that the Fed's actions would keep the recession mild and stimulate a recovery. However, the economy did not bounce back as quickly as the Fed had hoped. Why was the recovery from the recession so sluggish?

One explanation is that the corporate scandals at Enron, Arthur Andersen, and several other large firms caused investors to doubt the quality of the information about corporations. Doubts about the quality of corporate information meant that asymmetric information problems worsened, so that it became harder for an investor to screen out good firms from bad firms when making investment decisions. Because of the potential for increased adverse selection, as described in the credit view, individuals and financial

[20] For a survey of this evidence, see Hubbard, "Is There a 'Credit Channel' for Monetary Policy?" (note 17).

[21] See Mark Gertler and Simon Gilchrist, "Monetary Policy, Business Cycles, and the Behavior of Small Manufacturing Firms," *Quarterly Journal of Economics* 109 (May 1994): 309–340.

institutions were less willing to lend. This reluctance to lend in turn led to a decline in investment and aggregate output.

In addition, as we saw in Chapter 7, the corporate scandals caused investors to be less optimistic about earnings growth and to think that stocks were riskier, an effect leading to a further drop in the stock market. The decline in the stock market also weakened the economy, because it lowered household wealth. In turn, the decrease in household wealth led not only to restrained consumer spending, but also to weaker investment, because of the resulting drop in Tobin's q. In addition, the stock market decline weakened corporate balance sheets. This weakening increased asymmetric information problems and decreased lending and investment spending.

Corporate scandals have not only decreased our confidence in business leaders, but have also created a drag on the economy that has hindered the recovery from recession.

Lessons for Monetary Policy

What useful implications for central banks' conduct of monetary policy can we draw from the analysis in this chapter? There are four basic lessons to be learned.

1. ***It is dangerous always to associate the easing or tightening of monetary policy with a fall or a rise in short-term nominal interest rates.*** Because most central banks use short-term nominal interest rates—typically, the interbank rate—as the key operating instrument for monetary policy, there is a danger that central banks and the public will focus too much on short-term nominal interest rates as an indicator of the stance of monetary policy. Indeed, it is quite common to see statements that always associate monetary tightenings with a rise in the interbank rate and monetary easings with a decline in the rate. This view is highly problematic, because—as we have seen in our discussion of the Great Depression period—movements in nominal interest rates do not always correspond to movements in real interest rates, and yet it is typically the real and not the nominal interest rate that is an element in the channel of monetary policy transmission. For example, we have seen that during the contraction phase of the Great Depression in the United States, short-term interest rates fell to near zero and yet real interest rates were extremely high. Short-term interest rates that are near zero therefore do not indicate that monetary policy is easy if the economy is undergoing deflation, as was true during the contraction phase of the Great Depression. As Milton Friedman and Anna Schwartz have emphasized, the period of near-zero short-term interest rates during the contraction phase of the Great Depression was one of highly contractionary monetary policy rather than the reverse.

2. ***Other asset prices besides those on short-term debt instruments contain important information about the stance of monetary policy because they are important elements in various monetary policy transmission mechanisms.*** As we have seen in this chapter, economists have come a long way in understanding that other asset prices besides interest rates have major effects on aggregate demand. The view in Figure 3 that other asset prices, such as stock prices, foreign exchange rates, and housing and land prices, play an important role in monetary transmission mechanisms is held by both monetarists and Keynesians. Furthermore, the discussion of such additional channels as those operating through the exchange rate, Tobin's q, and

wealth effects provides additional reasons why other asset prices play such an important role in the monetary transmission mechanisms. Although there are strong disagreements among economists about which channels of monetary transmission are the most important—not surprising, given that economists, particularly those in academia, always like to disagree—they do agree that other asset prices play an important role in the way monetary policy affects the economy.

The view that other asset prices besides short-term interest rates matter has important implications for monetary policy. When we try to assess the stance of policy, it is critical that we look at other asset prices besides short-term interest rates. For example, if short-term interest rates are low or even zero and yet stock prices are low, land prices are low, and the value of the domestic currency is high, monetary policy is clearly tight, *not* easy.

3. ***Monetary policy can be highly effective in reviving a weak economy even if short-term interest rates are already near zero.*** We have recently entered a world where inflation is not always the norm. Japan, for example, recently experienced a period of deflation, when the price level was actually falling. One common view is that when a central bank has driven down short-term nominal interest rates to near zero, there is nothing more that monetary policy can do to stimulate the economy. The transmission mechanisms of monetary policy described here indicate that this view is false. As our discussion of the factors that affect the monetary base in Chapter 15 indicated, expansionary monetary policy to increase liquidity in the economy can be conducted with open market purchases, which do not have to be solely in short-term government securities. For example, purchases of foreign currencies, like purchases of government bonds, lead to an increase in the monetary base and in the money supply. This increased liquidity helps revive the economy by raising general price-level expectations and by reflating other asset prices, which then stimulate aggregate demand through the channels outlined here. Therefore, monetary policy can be a potent force for reviving economies that are undergoing deflation and have short-term interest rates near zero. Indeed, because of the lags inherent in fiscal policy and the political constraints on its use, expansionary monetary policy is the key policy action required to revive an economy experiencing deflation.

4. ***Avoiding unanticipated fluctuations in the price level is an important objective of monetary policy, thus providing a rationale for price stability as the primary long-run goal for monetary policy.*** As we saw in Chapter 18, central banks in recent years have been putting greater emphasis on price stability as the primary long-run goal for monetary policy. Several rationales have been proposed for this goal, including the undesirable effects of uncertainty about the future price level on business decisions and hence on productivity, distortions associated with the interaction of nominal contracts and the tax system with inflation, and increased social conflict stemming from inflation. The discussion here of monetary transmission mechanisms provides an additional reason why price stability is so important. As we have seen, unanticipated movements in the price level can cause unanticipated fluctuations in output, an undesirable outcome. Particularly important in this regard is the knowledge that, as we saw in Chapter 8, price deflation can be an important factor leading to a prolonged financial crisis, as occurred during the Great Depression. An understanding of the monetary transmission mechanisms thus makes it clear that the goal of price stability is desirable, because it reduces uncertainty about the future price level. Thus the price stability goal implies that a negative inflation rate is at least as undesirable as too high an inflation rate. Indeed, because of the threat of financial crises, central banks must work very hard to prevent price deflation.

Application

Applying the Monetary Policy Lessons to Japan

Until 1990, it looked as if Japan might overtake the United States in per capita income. Since then, the Japanese economy has been stagnating, with deflation and low growth. As a result, Japanese living standards have been falling farther and farther behind those in the United States. Many economists take the view that Japanese monetary policy is in part to blame for the poor performance of the Japanese economy. Could applying the four lessons outlined in the previous section have helped Japanese monetary policy perform better?

The first lesson suggests that it is dangerous to think that declines in interest rates always mean that monetary policy has been easing. In the mid-1990s, when short-term interest rates began to decline, falling to near zero in the late 1990s and early 2000s, the monetary authorities in Japan took the view that monetary policy was sufficiently expansionary. Now it is widely recognized that this view was incorrect, because the falling and eventually negative inflation rates in Japan meant that real interest rates were actually quite high and that monetary policy was tight, not easy. If the monetary authorities in Japan had followed the advice of the first lesson, they might have pursued a more expansionary monetary policy, which would have helped boost the economy.

The second lesson suggests that monetary policymakers should pay attention to other asset prices in assessing the stance of monetary policy. At the same time interest rates were falling in Japan, stock and real estate prices were collapsing, thus providing another indication that Japanese monetary policy was not easy. Recognizing the second lesson might have led Japanese monetary policymakers to recognize sooner that they needed a more expansionary monetary policy.

The third lesson indicates that monetary policy can still be effective even if short-term interest rates are near zero. Officials at the Bank of Japan have frequently claimed that they have been helpless in stimulating the economy, because short-term interest rates had fallen to near zero. Recognizing that monetary policy can still be effective even when interest rates are near zero, as the third lesson suggests, would have helped them to take monetary policy actions that would have stimulated aggregate demand by raising other asset prices and inflationary expectations.

The fourth lesson indicates that unanticipated fluctuations in the price level should be avoided. If the Japanese monetary authorities had adhered to this lesson, they might have recognized that allowing deflation to occur could be very damaging to the economy and would be inconsistent with the goal of price stability. Indeed, critics of the Bank of Japan have suggested that the bank should announce an inflation target in order to promote the price stability objective, but the bank has resisted this suggestion.

Heeding the advice from the four lessons in the previous section might have led to a far more successful conduct of monetary policy in Japan in recent years.

Summary

1. There are two basic types of empirical evidence: reduced-form evidence and structural model evidence. Both have advantages and disadvantages. The main advantage of structural model evidence is that it provides us with an understanding of how the economy works and gives us more confidence in the direction of causation between money and output. However, if the structure is not correctly specified, because it ignores important monetary transmission mechanisms, it could seriously underestimate the effectiveness of monetary policy. Reduced-form evidence has the advantage of not restricting the way monetary policy affects economic activity and so may be more likely to capture the full effects of monetary policy. However, reduced-form evidence cannot rule out the possibility of reverse causation or an outside driving factor, which could lead to misleading conclusions about the importance of money.
2. The early Keynesians believed that money does not matter, because they found weak links between interest rates and investment and because low interest rates on Treasury securities convinced them that monetary policy was easy during the worst economic contraction in U.S. history, the Great Depression. Monetarists objected to this interpretation of the evidence on the grounds that (a) the focus on nominal rather than real interest rates may have obscured any link between interest rates and investment, (b) interest-rate effects on investment might be only one of many channels through which monetary policy affects aggregate demand, and (c) by the standards of real interest rates and interest rates on lower-grade bonds, monetary policy was extremely contractionary during the Great Depression.
3. Early monetarist evidence falls into three categories: timing, statistical, and historical. Because of reverse causation and outside-factor possibilities, some serious doubts exist regarding conclusions that can be drawn from timing and statistical evidence alone. However, some of the historical evidence in which exogenous declines in money growth are followed by recessions provides stronger support for the monetarist position that money matters. As a result of empirical research, Keynesian and monetarist opinion has converged to the view that money does matter to aggregate economic activity and the price level. However, Keynesians do not agree with the monetarist position that money is *all* that matters.
4. The transmission mechanisms of monetary policy include traditional interest-rate channels that operate through the cost of capital and affect investment; other asset price channels such as exchange rate effects, Tobin's *q* theory, and wealth effects; and the credit view channels—the bank lending channel, the balance sheet channel, the cash flow channel, the unanticipated price level channel, and household liquidity effects.
5. Four lessons for monetary policy can be drawn from this chapter: (a) It is dangerous always to associate monetary policy easing or tightening with a fall or a rise in short-term nominal interest rates; (b) other asset prices besides those on short-term debt instruments contain important information about the stance of monetary policy because they are important elements in the monetary policy transmission mechanisms; (c) monetary policy can be highly effective in reviving a weak economy even if short-term interest rates are already near zero; and (d) avoiding unanticipated fluctuations in the price level is an important objective of monetary policy, thus providing a rationale for price stability as the primary long-run goal for monetary policy.

Key Terms

consumer durable expenditure, p. 617
consumption, p. 620
credit view, p. 618
reduced-form evidence, p. 604
reverse causation, p. 606
structural model, p. 604
structural model evidence, p. 603
transmission mechanisms of monetary policy, p. 604

Questions and Problems

Questions marked with an asterisk are answered at the end of the book in an appendix, "Answers to Selected Questions and Problems."

1. Suppose that a researcher is trying to determine whether jogging is good for a person's health. She examines this question in two ways. In method A, she looks to see whether joggers live longer than nonjoggers. In method B, she looks to see whether jogging reduces cholesterol in the bloodstream and lowers blood pressure; then she asks whether lower cholesterol and blood pressure prolong life. Which of these two methods will produce reduced-form evidence and which will produce structural model evidence?
2. If research indicates that joggers do not have lower cholesterol and blood pressure than nonjoggers, is it still possible that jogging is good for your health? Give a concrete example.
3. If research indicates that joggers live longer than nonjoggers, is it possible that jogging is not good for your health? Give a concrete example.

*4. Suppose that you plan to buy a car and want to know whether a General Motors car is more reliable than a Ford. One way to find out is to ask owners of both cars how often their cars go into the shop for repairs. Another way is to visit the factory producing the cars and see which one is built better. Which procedure will provide reduced-form evidence and which structural model evidence?

*5. If the GM car you plan to buy has a better repair record than a Ford, does this mean that the GM car is necessarily more reliable? (GM car owners might, for example, change their oil more frequently than Ford owners.)

*6. Suppose that when you visit the Ford and GM car factories to examine how the cars are built, you have time only to see how well the engine is put together. If Ford engines are better built than GM engines, does that mean that the Ford will be more reliable than the GM car?

7. How might bank behavior (described in Chapter 16) lead to causation running from output to the money supply? What does this say about evidence that finds a strong correlation between money and output?

*8. What operating procedures of the Fed (described in Chapter 18) might explain how movements in output might cause movements in the money supply?

9. "In every business cycle in the past 100 years, the rate at which the money supply is growing always decreases before output does. Therefore, the money supply causes business cycle movements." Do you agree? What objections can you raise against this argument?

*10. How did the research strategies of Keynesian and monetarist economists differ after they were exposed to the earliest monetarist evidence?

11. In the 1973–1975 recession, the value of common stocks in real terms fell by nearly 50%. How might this decline in the stock market have affected aggregate demand and thus contributed to the severity of this recession? Be specific about the mechanisms through which the stock market decline affected the economy.

*12. "The cost of financing investment is related only to interest rates; therefore, the only way that monetary policy can affect investment spending is through its effects on interest rates." Is this statement true, false, or uncertain? Explain your answer.

13. Predict what will happen to stock prices if the money supply rises. Explain why you are making this prediction.

*14. Franco Modigliani found that the most important transmission mechanisms of monetary policy involve consumer expenditure. Describe how at least two of these mechanisms work.

15. "The monetarists have demonstrated that the early Keynesians were wrong in saying that money doesn't matter at all to economic activity. Therefore, we should accept the monetarist position that money is all that matters." Do you agree? Why or why not?

Web Exercises

1. Figure 1 shows the relationship between estimated real interest rates and nominal interest rates. Go to www.martincapital.com/ and click on "charts and data" then on "nominal versus real market rates" to find data showing the spread between real interest and nominal interest rates. Discuss how the current spread differs from that shown most recently in Figure 1. What are the implications of this change?
2. Figure 2 discusses business cycles. While peaks and troughs of economic activity are a normal part of the business cycle, recessions are not. They represent a failure of economic policy. Go to www.econlib.org/library/Enc/Recessions.html and review the material reported on recessions.
   a. What is the formal definition of a recession?
   b. What are the problems with the definition?
   c. What are the three Ds used by the National Bureau of Economic Research (NBER) to define a recession?
   d. Review Chart 1. What trend is apparent about the length of recessions?

Chapter 27

Money and Inflation

PREVIEW

Since the early 1960s, when the inflation rate hovered between 1 and 2%, the economy has suffered from higher and more variable rates of inflation. By the late 1960s, the inflation rate had climbed beyond 5%, and by 1974, it reached the double-digit level. After moderating somewhat during the 1975–1978 period, it shot above 10% in 1979 and 1980, slowed to around 5% from 1982 to 1990, and declined further to around 2% in the late 1990s and early 2000s. Inflation, the condition of a continually rising price level, has become a major concern of politicians and the public, and how to control it frequently dominates the discussion of economic policy.

How do we prevent the inflationary fire from igniting and end the roller-coaster ride in the inflation rate of the past 40 years? Milton Friedman provides an answer in his famous proposition that "inflation is always and everywhere a monetary phenomenon." He postulates that the source of all inflation episodes is a high growth rate of the money supply: Simply by reducing the growth rate of the money supply to low levels, inflation can be prevented.

In this chapter, we use aggregate demand and supply analysis from Chapter 25 to reveal the role of monetary policy in creating inflation. You will find that as long as inflation is defined as the condition of a continually and rapidly rising price level, monetarists and Keynesians both agree with Friedman's proposition that inflation is a monetary phenomenon.

But what *causes* inflation? How does inflationary monetary policy come about? You will see that inflationary monetary policy is an offshoot of other government policies: the attempt to hit high employment targets or the running of large budget deficits. Examining how these policies lead to inflation will point us toward ways of preventing it at minimum cost in terms of unemployment and output loss.

Money and Inflation: Evidence

The evidence for Friedman's statement is straightforward. ***Whenever a country's inflation rate is extremely high for a sustained period of time, its rate of money supply growth is also extremely high***. Indeed, this is exactly what we saw in Figure 6 in Chapter 1, which shows that the countries with the highest inflation rates have also had the highest rates of money growth.

Evidence of this type seems to support the proposition that extremely high inflation is the result of a high rate of money growth. Keep in mind, however, that you are looking at reduced-form evidence, which focuses solely on the correlation of two variables: money growth and the inflation rate. As with all reduced-form evidence, reverse causation (inflation causing money supply growth) or an outside factor that drives both money growth and inflation could be involved.

www.bls.gov/cpi/
The home page of the Bureau of Labor Statistics, which reports inflation numbers.

How might you rule out these possibilities? First, you might look for historical episodes in which an increase in money growth appears to be an exogenous event; a high inflation rate for a sustained period following the increase in money growth would provide strong evidence that high money growth is the driving force behind the inflation. Luckily for our analysis, such clear-cut episodes—hyperinflations (extremely rapid inflations with inflation rates exceeding 50% per month)—have occurred, the most notorious being the German hyperinflation of 1921–1923.

German Hyperinflation, 1921–1923

In 1921, the need to make reparations and reconstruct the economy after World War I caused the German government's expenditures to greatly exceed revenues. The government could have obtained revenues to cover these increased expenditures by raising taxes, but that solution was, as always, politically unpopular and would have taken much time to implement. The government could also have financed the expenditure by borrowing from the public, but the amount needed was far in excess of its capacity to borrow. There was only one route left: the printing press. The government could pay for its expenditures simply by printing more currency (increasing the money supply) and using it to make payments to the individuals and companies that were providing it with goods and services. As shown in Figure 1, this is exactly what the German government did; in late 1921, the money supply began to increase rapidly, and so did the price level.

In 1923, the budgetary situation of the German government deteriorated even further. Early that year, the French invaded the Ruhr, because Germany had failed to make its scheduled reparations payments. A general strike in the region then ensued to protest the French action, and the German government actively supported this "passive resistance" by making payments to striking workers. As a result, government expenditures climbed dramatically, and the government printed currency at an even faster rate to finance this spending. As displayed in Figure 1, the result of the explosion in the money supply was that the price level blasted off, leading to an inflation rate for 1923 that exceeded 1 million percent!

The invasion of the Ruhr and the printing of currency to pay striking workers fit the characteristics of an exogenous event. Reverse causation (that the rise in the price level caused the French to invade the Ruhr) is highly implausible, and it is hard to imagine a third factor that could have been a driving force behind both inflation and the explosion in the money supply. Therefore, the German hyperinflation qualifies as a "controlled experiment" that supports Friedman's proposition that inflation is a monetary phenomenon.

Recent Episodes of Rapid Inflation

Although recent rapid inflations have not been as dramatic as the German hyperinflation, many countries in the 1980s and 1990s experienced rapid inflations in which the high rates of money growth can also be classified as exogenous events. For example, of all Latin American countries in the decade from 1980 to 1990, Argentina, Brazil, and Peru had both the highest rates of money growth and the highest average

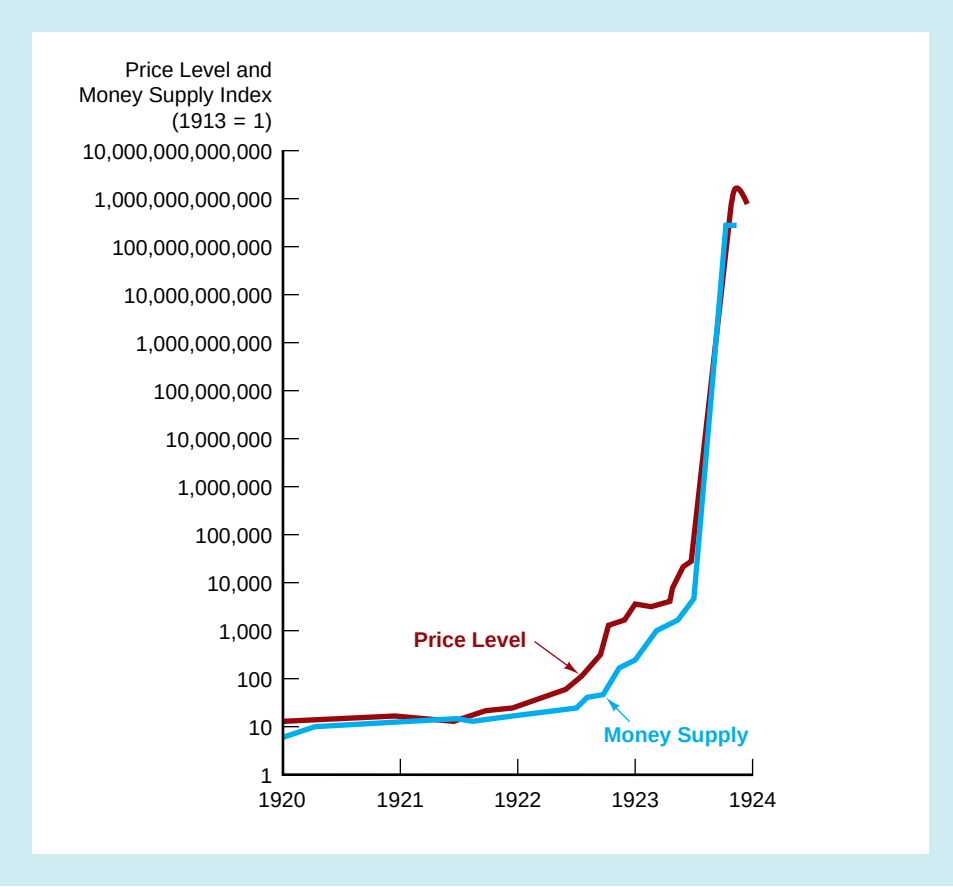


FIGURE 1 Money Supply and Price Level in the German Hyperinflation

Source: Frank D. Graham, *Exchange, Prices and Production in Hyperinflation: Germany, 1920–25* (Princeton, N.J.: Princeton University Press, 1930), pp. 105–106.

inflation rates. However, in the last couple of years, inflation in these countries has been brought down considerably.

The explanation for the high rates of money growth in these countries is similar to the explanation for Germany during its hyperinflation: The unwillingness of Argentina, Brazil, and Peru to finance government expenditures by raising taxes led to large budget deficits (sometimes over 15% of GDP), which were financed by money creation.

That the inflation rate is high in all cases in which the high rate of money growth can be classified as an exogenous event (including episodes in Argentina, Brazil, Peru, and Germany) is strong evidence that high money growth causes high inflation.

Meaning of Inflation

You may have noticed that all the empirical evidence on the relationship of money growth and inflation discussed so far looks only at cases in which the price level is continually rising at a rapid rate. It is this definition of inflation that Friedman and other economists use when they make statements such as "Inflation is always and everywhere a monetary phenomenon." This is not what your friendly newscaster

means when reporting the monthly inflation rate on the nightly news. The newscaster is only telling you how much, in percentage terms, the price level has changed from the previous month. For example, when you hear that the monthly inflation rate is 1% (12% annual rate), this merely indicates that the price level has risen by 1% in that month. This could be a one-shot change, in which the high inflation rate is merely temporary, not sustained. Only if the inflation rate remains high for a substantial period of time (greater than 1% per month for several years) will economists say that inflation has been high.

Accordingly, Milton Friedman's proposition actually says that upward movements in the price level are a monetary phenomenon *only* if this is a sustained process. When *inflation* is defined as a continuing and rapid rise in the price level, most economists, whether monetarist or Keynesian, will agree with Friedman's proposition that money alone is to blame.

Views of Inflation

Now that we understand what Friedman's proposition means, we can use the aggregate supply and demand analysis learned in Chapter 25 to show that large and persistent upward movements in the price level (high inflation) can occur only if there is a continually increasing money supply.

Monetarist View

First, let's look at the outcome of a continually increasing money supply using monetarist analysis (see Figure 2). Initially, the economy is at point 1, with output at the natural rate level and the price level at P_1 (the intersection of the aggregate demand curve AD_1 and the aggregate supply curve AS_1) . If the money supply increases steadily over the course of the year, the aggregate demand curve shifts rightward to AD_2. At first, for a very brief time, the economy may move to point $1'$ and output may increase above the natural rate level to Y', but the resulting decline in unemployment below the natural rate level will cause wages to rise, and the aggregate supply curve will quickly begin to shift leftward. It will stop shifting only when it reaches AS_2, at which time the economy has returned to the natural rate level of output on the long-run aggregate supply curve.[1] At the new equilibrium, point 2, the price level has increased from P_1 to P_2.

If the money supply increases the next year, the aggregate demand curve will shift to the right again to AD_3, and the aggregate supply curve will shift from AS_2 to AS_3; the economy will then move to point $2'$ and then 3, where the price level has risen to P_3. If the money supply continues to grow in subsequent years, the economy will continue to move to higher and higher price levels. As long as the money supply grows, this process will continue, and inflation will occur.

Do monetarists believe that a continually rising price level can be due to any source other than money supply growth? The answer is no. In monetarist analysis, the money supply is viewed as the sole source of shifts in the aggregate demand curve, so

[1] In monetarist analysis, the aggregate supply curve may immediately shift in toward AS_2, because workers and firms may expect the increase in the money supply, so expected inflation will be higher. In this case, the movement to point 2 will be very rapid, and output need not rise above the natural rate level. (Further support for this scenario, from the theory of rational expectations, is discussed in Chapter 28.)

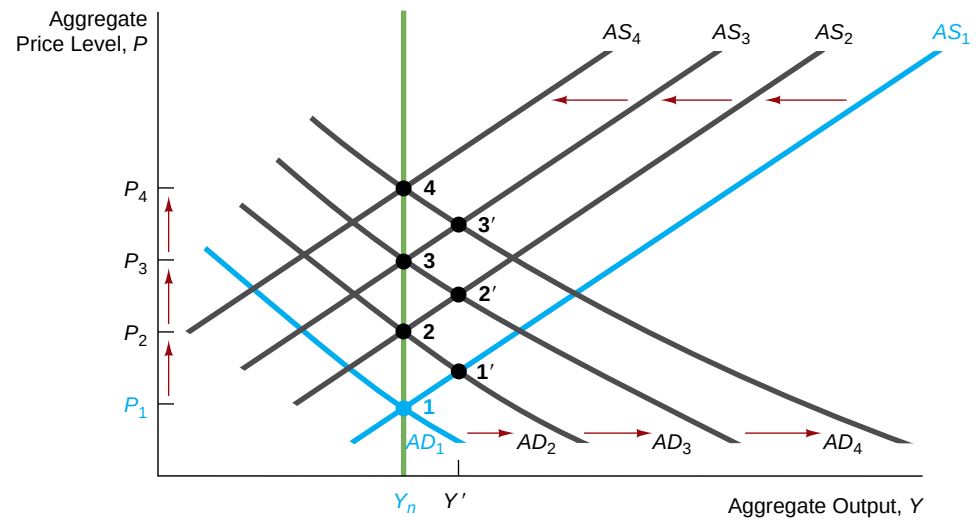


FIGURE 2 Response to a Continually Rising Money Supply

A continually rising money supply shifts the aggregate demand curve to the right from AD_1 to AD_2 to AD_3 to AD_4, while the aggregate supply curve shifts to the left from AS_1 to AS_2 to AS_3 to AS_4. The result is that the price level rises continually from P_1 to P_2 to P_3 to P_4.

there is nothing else that can move the economy from point 1 to 2 to 3 and beyond. ***Monetarist analysis indicates that rapid inflation must be driven by high money supply growth.***

Keynesian View

Keynesian analysis indicates that the continually increasing money supply will have the same effect on the aggregate demand and supply curves that we see in Figure 2: The aggregate demand curve will keep on shifting to the right, and the aggregate supply curve will keep shifting to the left.[2] The conclusion is the same one that the monetarists reach: A rapidly growing money supply will cause the price level to rise continually at a high rate, thus generating inflation.

Could a factor other than money generate high inflation in the Keynesian analysis? The answer is no. This result probably surprises you, for in Chapter 25 you learned that Keynesian analysis allows other factors besides changes in the money supply (such as fiscal policy and supply shocks) to affect the aggregate demand and supply curves. To see why Keynesians also view high inflation as a monetary phenomenon, let's examine whether their analysis allows other factors to generate high inflation in the absence of a high rate of money growth.

Can Fiscal Policy by Itself Produce Inflation? To examine this question, let's look at Figure 3, which demonstrates the effect of a one-shot permanent increase in govern-

[2]The only difference in the two analyses is that Keynesians believe that the aggregate supply curve would shift leftward more slowly than monetarists do. Thus Keynesian analysis suggests that output might tend to stay above the natural rate longer than monetarist analysis does.

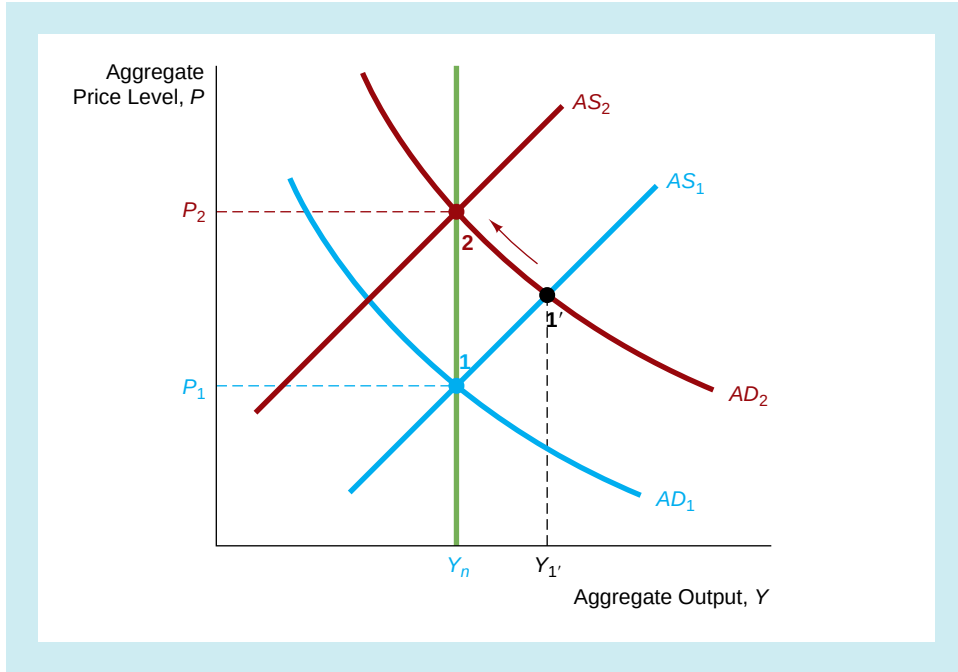


FIGURE 3 Response to a One-Shot Permanent Increase in Government Expenditure

A one-shot permanent increase in government expenditure shifts the aggregate demand curve rightward from AD_1 to AD_2, moving the economy from point 1 to point 1′. Because output now exceeds the natural rate level Y_n, the aggregate supply curve eventually shifts leftward to AS_2, and the price level rises from P_1 to P_2, a one-shot permanent increase but not a continuing increase.

ment expenditure (say, from $500 billion to $600 billion) on aggregate output and the price level. Initially, we are at point 1, where output is at the natural rate level and the price level is P_1. The increase in government expenditure shifts the aggregate demand curve to AD_2, and we move to point 1′, where output is above the natural rate level at $Y_{1'}$. The aggregate supply curve will begin to shift leftward, eventually reaching AS_2, where it intersects the aggregate demand curve AD_2 at point 2, at which output is again at the natural rate level and the price level has risen to P_2.

The net result of a one-shot permanent increase in government expenditure is a one-shot permanent increase in the price level. What happens to the inflation rate? When we move from point 1 to 1′ to 2, the price level rises, and we have a positive inflation rate. But when we finally get to point 2, the inflation rate returns to zero. We see that the one-shot increase in government expenditure leads to only a *temporary* increase in the inflation rate, not to an inflation in which the price level is continually rising.

If, however, government spending increased continually, we *could* get a continuing rise in the price level. It appears, then, that Keynesian analysis could reject Friedman's proposition that inflation is always the result of money growth. The problem with this argument is that a continually increasing level of government expenditure is not a feasible policy. There is a limit on the total amount of possible government expenditure; the government cannot spend more than 100% of GDP. In fact, well before this limit is reached, the political process would stop the increases in government spending. As revealed in the continual debates in Congress over balanced budgets and government spending, both the public and politicians have a particular target level of government spending they deem appropriate; although small deviations from this level might be tolerated, large deviations would not. Indeed, public and political perceptions impose tight limits on the degree to which government expenditures can increase.

What about the other side of fiscal policy—taxes? Could continual tax cuts generate an inflation? Again the answer is no. The analysis in Figure 3 also describes the price and output response to a one-shot decrease in taxes. There will be a one-shot increase in the price level, but the increase in the inflation rate will be only temporary. We can increase the price level by cutting taxes even more, but this process would have to stop—once taxes reach zero, they can't be reduced further. We must conclude, then, that ***Keynesian analysis indicates that high inflation cannot be driven by fiscal policy alone.***[3]

Can Supply-Side Phenomena by Themselves Produce Inflation? Because supply shocks and workers' attempts to increase their wages can shift the aggregate supply curve leftward, you might suspect that these supply-side phenomena by themselves could stimulate inflation. Again, we can show that this suspicion is incorrect.

Suppose that there is a negative supply shock—for example, an oil embargo—that raises oil prices (or workers could have successfully pushed up their wages). As displayed in Figure 4, the negative supply shock shifts the aggregate supply curve from AS_1 to AS_2. If the money supply remains unchanged, leaving the aggregate demand curve at AD_1, we move to point 1′, where output $Y_{1'}$ is below the natural rate level and the price level $P_{1'}$ is higher. The aggregate supply curve will now shift back to AS_1, because unemployment is above the natural rate, and the economy slides down AD_1 from point 1′ to point 1. The net result of the supply shock is that we return to full employment at the initial price level, and there is no continuing inflation. Additional negative supply shocks that again shift the aggregate supply curve leftward will lead to the same outcome: The price level will rise temporarily, but inflation will not result. The conclusion that we have reached is the following: ***Supply-side phenomena cannot be the source of high inflation.***[4]

Summary

Our aggregate demand and supply analysis shows that Keynesian and monetarist views of the inflation process are not very different. Both believe that high inflation can occur only with a high rate of money growth. Recognizing that by inflation we mean a continuing increase in the price level at a rapid rate, most economists agree with Milton Friedman that "inflation is always and everywhere a monetary phenomenon."

Origins of Inflationary Monetary Policy

Although we now know *what* must occur to generate a rapid inflation—a high rate of money growth—we still can't understand *why* high inflation occurs until we have learned how and why inflationary monetary policies come about. If everyone agrees that inflation is not a good thing for an economy, why do we see so much of it? Why

[3]The argument here demonstrates that "animal spirits" also cannot be the source of inflation. Although consumer and business optimism, which stimulates their spending, can produce a one-shot shift in the aggregate demand curve and a temporary inflation, it cannot produce continuing shifts in the aggregate demand curve and inflation in which the price level rises continually. The reasoning is the same as before: Consumers and businesses cannot continue to raise their spending without limit because their spending cannot exceed 100% of GDP.

[4]Supply-side phenomena that alter the natural rate level of output (and shift the long-run aggregate supply curve at Y_n) can produce a permanent one-shot change in the price level. However, this resulting one-shot change results in only a temporary inflation, not a continuing rise in the price level.

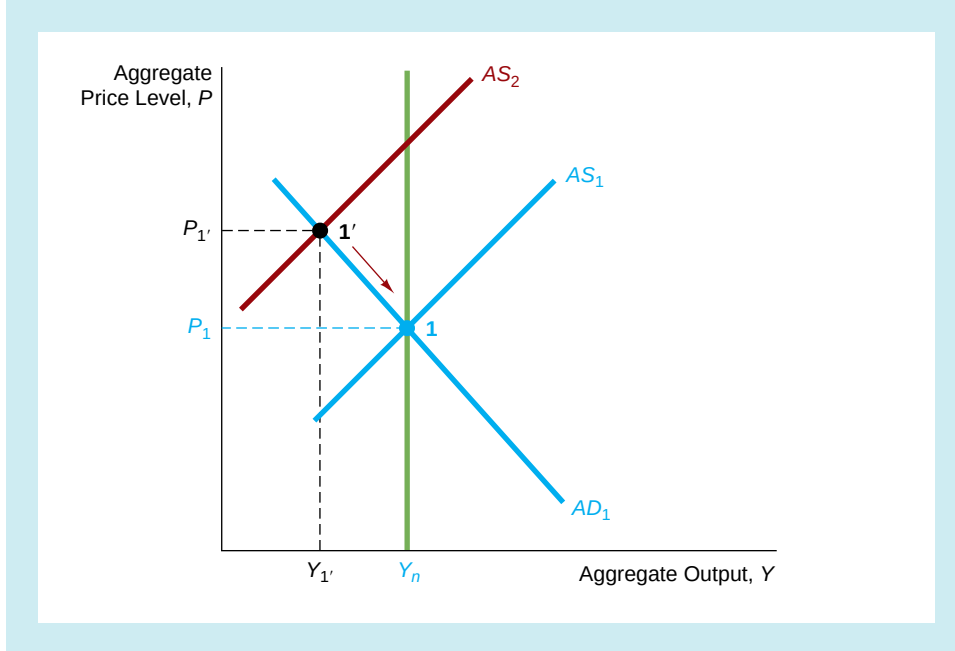


FIGURE 4 Response to a Supply Shock

A negative supply shock (or a wage push) shifts the aggregate supply curve leftward to AS_2 and results in high unemployment at point 1′. As a result, the aggregate supply curve shifts back to the right to AS_1, and the economy returns to point 1, where the price level has returned to P_1.

do governments pursue inflationary monetary policies? Since there is nothing intrinsically desirable about inflation and since we know that a high rate of money growth doesn't happen of its own accord, it must follow that in trying to achieve other goals, governments end up with a high money growth rate and high inflation. In this section, we will examine the government policies that are the most common sources of inflation.

High Employment Targets and Inflation

The first goal most governments pursue that often results in inflation is high employment. The U.S. government is committed by law (the Employment Act of 1946 and the Humphrey-Hawkins Act of 1978) to promoting high employment. Though it is true that both laws require a commitment to a high level of employment consistent with a stable price level, in practice our government has often pursued a high employment target with little concern about the inflationary consequences of its policies. This was true especially in the mid-1960s and 1970s, when the government began to take a more active role in attempting to stabilize unemployment.

Two types of inflation can result from an activist stabilization policy to promote high employment: **cost-push inflation**, which occurs because of negative supply shocks or a push by workers to get higher wages, and **demand-pull inflation**, which results when policymakers pursue policies that shift the aggregate demand curve to the right. We will now use aggregate demand and supply analysis to examine how a high employment target can lead to both types of inflation.

Cost-Push Inflation. In Figure 5, the economy is initially at point 1, the intersection of the aggregate demand curve AD_1 and the aggregate supply curve AS_1. Suppose that workers decide to seek higher wages, either because they want to increase their real wages (wages in terms of the goods and services they can buy) or because they expect inflation to be high and wish to keep up with inflation. The effect of such an increase

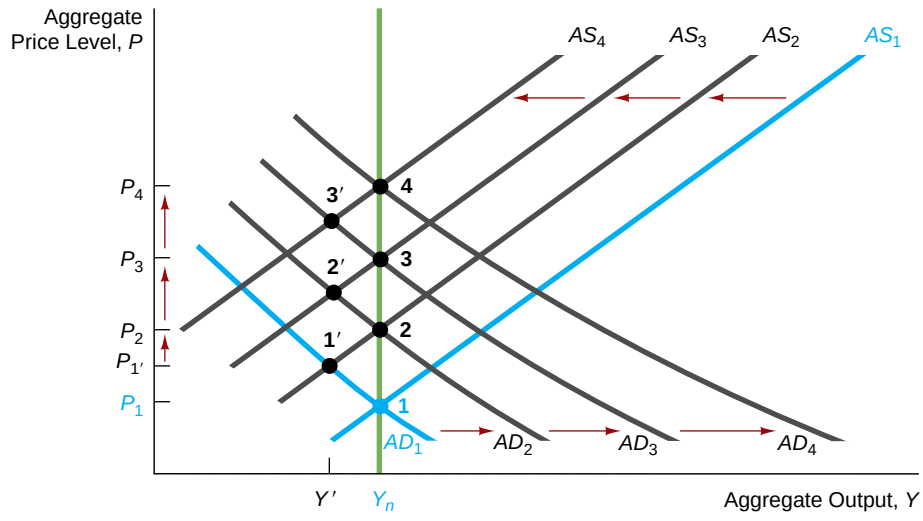


FIGURE 5 Cost-Push Inflation with an Activist Policy to Promote High Employment

In a cost-push inflation, the leftward shifts of the aggregate supply curve from AS_1 to AS_2 to AS_3 and so on cause a government with a high employment target to shift the aggregate demand curve to the right continually to keep unemployment and output at their natural rate levels. The result is a continuing rise in the price level from P_1 to P_2 to P_3 and so on.

(similar to a negative supply shock) is to shift the aggregate supply curve leftward to AS_2.[5] If government fiscal and monetary policy remains unchanged, the economy would move to point 1′ at the intersection of the new aggregate supply curve AS_2 and the aggregate demand curve AD_1. Output would decline to below its natural rate level Y_n, and the price level would rise to $P_{1'}$.

What would activist policymakers with a high employment target do if this situation developed? Because of the drop in output and resulting increase in unemployment, they would implement policies to raise the aggregate demand curve to AD_2, so that we would return to the natural rate level of output at point 2 and price level P_2. The workers who have increased their wages have not fared too badly. The government has stepped in to make sure that there is no excessive unemployment, and they have achieved their goal of higher wages. Because the government has, in effect, given in to the demands of workers for higher wages, an activist policy with a high employment target is often referred to as an **accommodating policy**.

The workers, having eaten their cake and had it too, might be encouraged to seek even higher wages. In addition, other workers might now realize that their wages have fallen relative to their fellow workers', and because they don't want to be left behind, these workers will seek to increase their wages. The result is that the aggregate supply curve shifts leftward again, to AS_3. Unemployment develops again when we move

[5]The cost-push inflation we describe here might also occur as a result either of firms' attempts to obtain higher prices or of negative supply shocks.

to point 2′, and the activist policies will once more be used to shift the aggregate demand curve rightward to AD_3 and return the economy to full employment at a price level of P_3. If this process continues, the result will be a continuing increase in the price level—a cost-push inflation.

What role does monetary policy play in a cost-push inflation? A cost-push inflation can occur only if the aggregate demand curve is shifted continually to the right. In Keynesian analysis, the first shift of the aggregate demand curve to AD_2 could be achieved by a one-shot increase in government expenditure or a one-shot decrease in taxes. But what about the next required rightward shift of the aggregate demand curve to AD_3, and the next, and the next? The limits on the maximum level of government expenditure and the minimum level of taxes would prevent the use of this expansionary fiscal policy for very long. Hence it cannot be used continually to shift the aggregate demand curve to the right. But the aggregate demand curve *can* be shifted continually rightward by continually increasing the money supply, that is, by going to a higher rate of money growth. Therefore, ***a cost-push inflation is a monetary phenomenon because it cannot occur without the monetary authorities pursuing an accommodating policy of a higher rate of money growth.***

Demand-Pull Inflation. The goal of high employment can lead to inflationary monetary policy in another way. Even at full employment, unemployment is always present because of frictions in the labor market, which make it difficult to match workers with employers. An unemployed autoworker in Detroit may not know about a job opening in the electronics industry in California or, even if he or she did, may not want to move or be retrained. So the unemployment rate when there is full employment (the natural rate of unemployment) will be greater than zero. If policymakers set a target for unemployment that is too low because it is less than the natural rate of unemployment, this can set the stage for a higher rate of money growth and a resulting inflation. Again we can show how this can happen using an aggregate supply and demand diagram (see Figure 6).

If policymakers have an unemployment target (say, 4%) that is below the natural rate (estimated to be between $4\frac{1}{2}$ and $5\frac{1}{2}$% currently), they will try to achieve an output target greater than the natural rate level of output. This target level of output is marked Y_T in Figure 6. Suppose that we are initially at point 1; the economy is at the natural rate level of output but below the target level of output Y_T. To hit the unemployment target of 4%, policymakers enact policies to increase aggregate demand, and the effects of these policies shift the aggregate demand curve until it reaches AD_2 and the economy moves to point 1′. Output is at Y_T, and the 4% unemployment rate goal has been reached.

If the targeted unemployment rate was at the natural rate level between $4\frac{1}{2}$ and $5\frac{1}{2}$%, there would be no problem. However, because at Y_T the 4% unemployment rate is below the natural rate level, wages will rise and the aggregate supply curve will shift in to AS_2, moving the economy from point 1′ to point 2. The economy is back at the natural rate of unemployment, but at a higher price level of P_2. We could stop there, but because unemployment is again higher than the target level, policymakers would again shift the aggregate demand curve rightward to AD_3 to hit the output target at point 2′, and the whole process would continue to drive the economy to point 3 and beyond. The overall result is a steadily rising price level—an inflation.

How can policymakers continually shift the aggregate demand curve rightward? We have already seen that they cannot do it through fiscal policy, because of the limits on raising government expenditures and reducing taxes. Instead they will have to

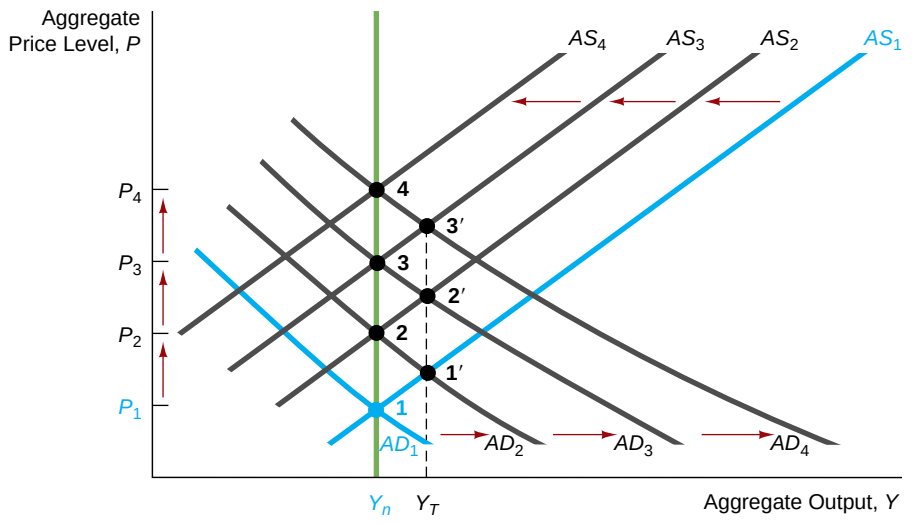


FIGURE 6 Demand-Pull Inflation: The Consequence of Setting Too Low an Unemployment Target

Too low an unemployment target (too high an output target of Y_T) causes the government to shift the aggregate demand curve rightward from AD_1 to AD_2 to AD_3 and so on, while the aggregate supply curve shifts leftward from AS_1 to AS_2 to AS_3 and so on. The result is a continuing rise in the price level known as a demand-pull inflation.

resort to expansionary monetary policy: a continuing increase in the money supply and hence a high money growth rate.

Pursuing too high an output target or, equivalently, too low an unemployment rate is the source of inflationary monetary policy in this situation, but it seems senseless for policymakers to do this. They have not gained the benefit of a permanently higher level of output but have generated the burden of an inflation. If, however, they do not realize that the target rate of unemployment is below the natural rate, the process that we see in Figure 6 will be well under way before they realize their mistake.

Because the inflation described results from policymakers' pursuing policies that shift the aggregate demand curve to the right, it is called a *demand-pull inflation*. In contrast, a *cost-push inflation* occurs when workers push their wages up. Is it easy to distinguish between them in practice? The answer is no. We have seen that both types of inflation will be associated with higher money growth, so we cannot distinguish them on this basis. Yet as Figures 5 and 6 demonstrate, demand-pull inflation will be associated with periods when unemployment is below the natural rate level, whereas cost-push inflation is associated with periods when unemployment is above the natural rate level. To decide which type of inflation has occurred, we can look at whether unemployment has been above or below its natural rate level. This would be easy if economists and policymakers actually knew how to measure the natural rate of unemployment; unfortunately, this very difficult research question is still not fully resolved by the economics profession. In addition, the distinction between cost-push and demand-pull inflation is blurred, because a cost-push inflation can be initiated by a demand-pull inflation: When a demand-pull inflation produces higher inflation rates,

expected inflation will eventually rise and cause workers to demand higher wages so that their real wages do not fall. In this way, demand-pull inflation can eventually trigger cost-push inflation.

Budget Deficits and Inflation

Our discussion of the evidence on money and inflation suggested that budget deficits are another possible source of inflationary monetary policy. To see if this could be the case, we need to look at how a government finances its budget deficits.

Government Budget Constraint. Because the government has to pay its bills just as we do, it has a budget constraint. There are two ways we can pay for our spending: raise revenue (by working) or borrow. The government also enjoys these two options: raise revenue by levying taxes or go into debt by issuing government bonds. Unlike us, however, it has a third option: The government can create money and use it to pay for the goods and services it buys.

The requirement that the government budget deficit equal the sum of the change in the monetary base and the change in government bonds held by the public.

Methods of financing government spending are described by an expression called the **government budget constraint**, which states the following: The government budget deficit *DEF*, which equals the excess of government spending *G* over tax revenue *T*, must equal the sum of the change in the monetary base ΔMB and the change in government bonds held by the public ΔB. Algebraically, this expression can be written as:

$$DEF = G - T = \Delta MB + \Delta B \tag{1}$$

To see what the government budget constraint means in practice, let's look at the case in which the only government purchase is a $100 million supercomputer. If the government convinces the electorate that such a computer is worth paying for, it will probably be able to raise the $100 million in taxes to pay for it, and the budget deficit will equal zero. The government budget constraint then tells us that no issue of money or bonds is needed to pay for the computer, because the budget is balanced. If taxpayers think that supercomputers are too expensive and refuse to pay taxes for them, the budget constraint indicates that the government must pay for it by selling $100 million of new bonds to the public or by printing $100 million of currency to pay for the computer. In either case, the budget constraint is satisfied; the $100 million deficit is balanced by the change in the stock of government bonds held by the public (ΔB = $100 million) or by the change in the monetary base (ΔMB = $100 million).

The government budget constraint thus reveals two important facts: ***If the government deficit is financed by an increase in bond holdings by the public, there is no effect on the monetary base and hence on the money supply. But, if the deficit is not financed by increased bond holdings by the public, the monetary base and the money supply increase.***

There are several ways to understand why a deficit leads to an increase in the monetary base when the public's bond holdings do not increase. The simplest case is when the government's treasury has the legal right to issue currency to finance its deficit. Financing the deficit is then very straightforward: The government just pays for the spending that is in excess of its tax revenues with new currency. Because this increase in currency adds directly to the monetary base, the monetary base rises and the money supply with it through the process of multiple deposit creation described in Chapters 15 and 16.

In the United States, however, and in many other countries, the government does not have the right to issue currency to pay for its bills. In this case, the government must finance its deficit by first issuing bonds to the public to acquire the extra funds to pay its bills. Yet if these bonds do not end up in the hands of the public, the only alternative is that they are purchased by the central bank. For the government bonds not to end up in the hands of the public, the central bank must conduct an open market purchase, which, as we saw in Chapters 15 and 16, leads to an increase in the monetary base and in the money supply. This method of financing government spending is called **monetizing the debt** because, as the two-step process described indicates, government debt issued to finance government spending has been removed from the hands of the public and has been replaced by high-powered money. This method of financing, or the more direct method when a government just issues the currency directly, is also, somewhat inaccurately, referred to as **printing money** because high-powered money (the monetary base) is created in the process. The use of the word *printing* is misleading because what is essential to this method of financing government spending is not the actual printing of money but rather the issuing of monetary liabilities to the public after the money has been printed.

We thus see that a budget deficit can lead to an increase in the money supply if it is financed by the creation of high-powered money. However, earlier in this chapter you have seen that inflation can develop only when the stock of money grows continually. Can a budget deficit financed by printing money do this? The answer is yes, if the budget deficit persists for a substantial period of time. In the first period, if the deficit is financed by money creation, the money supply will rise, shifting the aggregate demand curve to the right and leading to a rise in the price level (see Figure 2). If the budget deficit is still present in the next period, it has to be financed all over again. The money supply will rise again, and the aggregate demand curve will again shift to the right, causing the price level to rise further. As long as the deficit persists and the government resorts to printing money to pay for it, this process will continue. ***Financing a persistent deficit by money creation will lead to a sustained inflation.***

A critical element in this process is that the deficit is persistent. If temporary, it would not produce an inflation because the situation would then be similar to that shown in Figure 3, in which there is a one-shot increase in government expenditure. In the period when the deficit occurs, there will be an increase in money to finance it, and the resulting rightward shift of the aggregate demand curve will raise the price level. If the deficit disappears in the next period, there is no longer a need to print money. The aggregate demand curve will not shift further, and the price level will not continue to rise. Hence the one-shot increase in the money supply from the temporary deficit generates only a one-shot increase in the price level, and no inflation develops.

To summarize, ***a deficit can be the source of a sustained inflation only if it is persistent rather than temporary and if the government finances it by creating money rather than by issuing bonds to the public***.

If inflation is the result, why do governments frequently finance persistent deficits by creating money? The answer is the key to understanding how budget deficits may lead to inflation.

Budget Deficits and Money Creation in Other Countries. Although the United States has well-developed money and capital markets in which huge quantities of its government bonds, both short- and long-term, can be sold, this is not the situation in many

developing countries. If developing countries run budget deficits, they cannot finance them by issuing bonds and must resort to their only other alternative, printing money. As a result, when they run large deficits relative to GDP, the money supply grows at substantial rates, and inflation results.

Earlier we cited Latin American countries with high inflation rates and high money growth as evidence that inflation is a monetary phenomenon. The Latin American countries with high money growth are precisely the ones that have persistent and extremely large budget deficits relative to GDP. The only way to finance the deficits is to print more money, so the ultimate source of their high inflation rates is their large budget deficits.

In all episodes of hyperinflation, huge government budget deficits are also the ultimate source of inflationary monetary policies. The budget deficits during hyperinflations are so large that even if a capital market exists to issue government bonds, it does not have sufficient capacity to handle the quantity of bonds that the government wishes to sell. In this situation, the government must also resort to the printing press to finance the deficits.

Budget Deficits and Money Creation in the United States. So far we have seen why budget deficits in some countries must lead to money creation and inflation. Either the deficit is huge, or the country does not have sufficient access to capital markets in which it can sell government bonds. But neither of these scenarios seems to describe the situation in the United States. True, the United States' deficits were large in the 1980s and early 1990s, but even so, the magnitude of these deficits relative to GDP was small compared to the deficits of countries that have experienced hyperinflations: The U.S. deficit as a percentage of GDP reached a peak of 6% in 1983, whereas Argentina's budget deficit sometimes exceeded 15% of GDP. Furthermore, since the United States has the best-developed government bond market of any country in the world, it can issue large quantities of bonds when it needs to finance its deficit.

Whether the budget deficit can influence the monetary base and the money supply depends critically on how the Federal Reserve chooses to conduct monetary policy. If the Fed pursues a policy goal of preventing high interest rates (a possibility, as we have seen in Chapter 18), many economists contend that a budget deficit will lead to the printing of money. Their reasoning, using the supply and demand analysis of the bond market in Chapter 5, is as follows: When the Treasury issues bonds to the public, the supply of bonds rises (from B_1^s to B_2^s in Figure 7), causing interest rates to rise from i_1 to i_2 and bond prices to fall. If the Fed considers the rise in interest rates undesirable, it will buy bonds to prop up bond prices and reduce interest rates. The net result is that the government budget deficit can lead to Federal Reserve open market purchases, which raise the monetary base (create high-powered money) and raise the money supply. If the budget deficit persists so that the quantity of bonds supplied keeps on growing, the upward pressure on interest rates will continue, the Fed will purchase bonds again and again, and the money supply will continually rise, resulting in an inflation.

Economists such as Robert Barro of Harvard University, however, do not agree that budget deficits influence the monetary base in the manner just described. Their analysis (which Barro named **Ricardian equivalence** after the nineteenth-century British economist David Ricardo) contends that when the government runs deficits and issues bonds, the public recognizes that it will be subject to higher taxes in the future to pay off these bonds. The public then saves more in anticipation of these

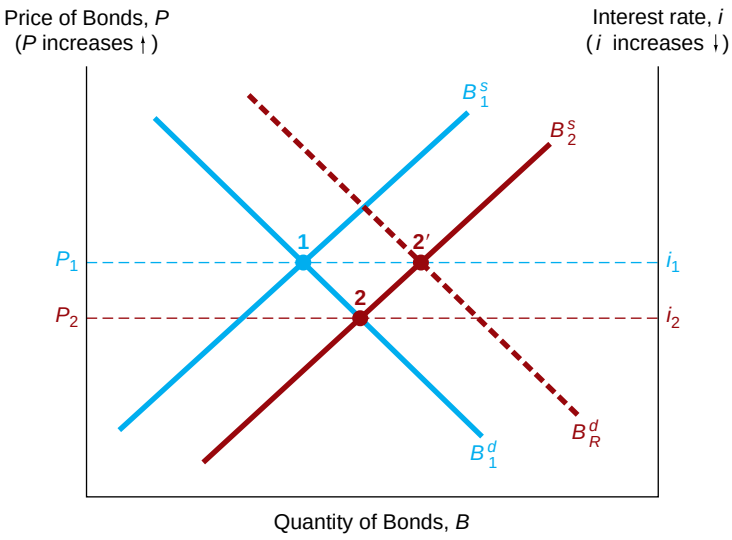


FIGURE 7 Interest Rates and the Government Budget Deficit

When the Treasury issues bonds to finance the budget deficit, the supply curve for bonds shifts rightward from B_1^s to B_2^s. Many economists take the position that the equilibrium moves to point 2 because the bond demand curve remains unchanged, with the result that the bond price falls from P_1 to P_2 and the interest rate rises from i_1 to i_2. Adherents of Ricardian equivalence, however, suggest that the demand curve for bonds also increases to B_R^d, moving the equilibrium to point 2′, where the interest rate is unchanged at i_1. (Note that P and i increase in opposite directions. P on the left vertical axis increases as we go up the axis, whereas i on the right vertical axis increases as we go down the axis.)

future taxes, with the net result that the public demand for bonds increases to match the increased supply. The demand curve for bonds shifts rightward to B_R^d in Figure 7, leaving the interest rate unchanged at i_1. There is now no need for the Fed to purchase bonds to keep the interest rate from rising.

To sum up, although high inflation is "always and everywhere a monetary phenomenon" in the sense that it cannot occur without a high rate of money growth, there are reasons why this inflationary monetary policy might come about. The two underlying reasons are the adherence of policymakers to a high employment target and the presence of persistent government budget deficits.

Application

Explaining the Rise in U.S. Inflation, 1960–1980

Now that we have examined the underlying sources of inflation, let's apply this knowledge to understanding the causes of the rise in U.S. inflation from 1960 to 1980.

Figure 8 documents the rise in inflation in those years. At the beginning of the period, the inflation rate is close to 1% at an annual rate; by the late 1970s, it is averaging around 8%. How does the analysis of this chapter explain this rise in inflation?

The conclusion that inflation is a monetary phenomenon is given a fair amount of support by the period from 1960 through 1980. As Figure 8 shows, in this period, there is a close correspondence between movements in the inflation rate and the monetary growth rate from two years earlier. (The money growth rates are from two years earlier, because research indicates

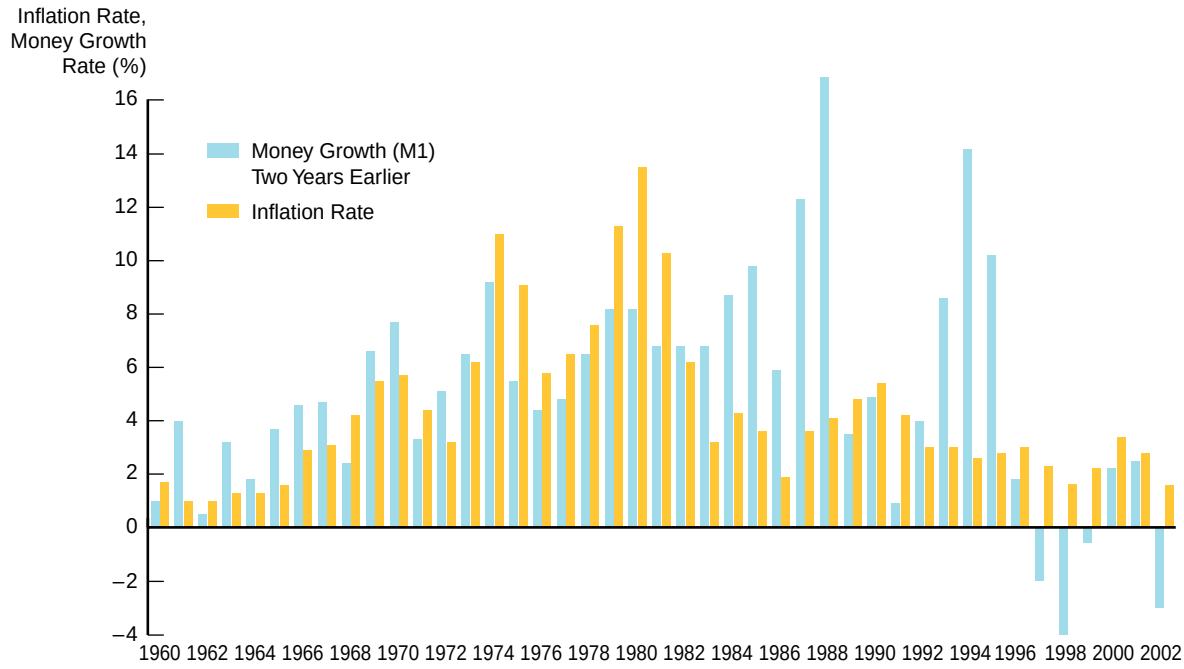


FIGURE 8 Inflation and Money Growth, 1960–2002

Source: Economic Report of the President; www.federalreserve.gov/releases/h6/hist/h6hist1.txt.

ftp://ftp.bls.gov/pub/special.requests/cpi/cpiai.txt

Download historical inflation statistics going back to 1913. This data can easily be moved into Microsoft Excel using the procedure discussed at the end of Chapter 1.

that a change in money growth takes that long to affect the inflation rate.) The rise in inflation from 1960 to 1980 can be attributed to the rise in the money growth rate over this period. But you have probably noticed that in 1974–1975 and 1979–1980, the inflation rate is well above the money growth rate from two years earlier. You may recall from Chapter 25 that temporary upward bursts of the inflation rate in those years can be attributed to supply shocks from oil and food price increases that occurred in 1973–1975 and 1978–1980.

However, the linkage between money growth and inflation after 1980 is not at all evident in Figure 8, and this explains why in 1982 the Fed announced that it would no longer use M1 as a basis to set monetary policy (see Chapter 18). The breakdown of the relationship between money growth and inflation is the result of substantial gyrations in velocity in the 1980s and 1990s (documented in Chapter 22). For example, the early 1980s was a period of rapid disinflation (a substantial fall in the inflation rate), yet the money growth rates in Figure 8 do not display a visible downward trend until after the disinflation was over. (The disinflationary process in the 1980s will be discussed in another application later in this chapter.) Although some economists see the 1980s and 1990s as evidence against the money–inflation link, others view this as an unusual period characterized by large fluctuations

in interest rates and by rapid financial innovation that made the correct measurement of money far more difficult (see Chapter 3). In their view, this period was an aberration, and the close correspondence of money and inflation is sure to reassert itself. However, this has not yet occurred.

What is the underlying cause of the increased rate of money growth that we see occurring from 1960 to 1980? We have identified two possible sources of inflationary monetary policy: government adherence to a high employment target and budget deficits. Let's see if budget deficits can explain the move to an inflationary monetary policy by plotting the ratio of government debt to GDP in Figure 9. This ratio provides a reasonable measure of whether government budget deficits put upward pressure on interest rates. Only if this ratio is rising might there be a tendency for budget deficits to raise interest rates because the public is then being asked to hold more government bonds relative to their capacity to buy them. Surprisingly, over the course of the 20-year period from 1960 to 1980, this ratio was falling, not rising. Thus U.S. budget deficits in this period did not raise interest rates and so could not have encouraged the Fed to expand the money supply by buying bonds. Therefore, Figure 9 tells us that we can rule out budget deficits as a source of the rise in inflation in this period.

Because politicians were frequently bemoaning the budget deficits in this period, why did deficits not lead to an increase in the debt–GDP ratio? The reason is that in this period, U.S. budget deficits were sufficiently small that the increase in the stock of government debt was still slower than the growth in nominal GDP, and the ratio of debt to GDP declined. You can see that interpreting budget deficit numbers is a tricky business.[6]

We have ruled out budget deficits as the instigator; what else could be the underlying cause of the higher rate of money growth and more rapid inflation in the 1960s and 1970s? Figure 10, which compares the actual unemployment rate to the natural rate of unemployment, shows that the economy was experiencing unemployment below the natural rate in all but one year between 1965 and 1973. This suggests that in 1965–1973, the American economy was experiencing the demand-pull inflation described in Figure 6.

Policymakers apparently pursued policies that continually shifted the aggregate demand curve to the right in trying to achieve an output target that was too high, thus causing the continual rise in the price level outlined in Figure 6. This occurred because policymakers, economists, and politicians had become committed in the mid-1960s to a target unemployment rate of 4%, the level of unemployment they thought was consistent with price stability. In hindsight, most economists today agree that the natural rate of unemployment was substantially higher in this period, on the order of 5 to 6%, as shown in Figure 10. The result of the inappropriate 4% unemployment

http://w3.access.gpo.gov/usbudget/

The *Economic Report of the President* reports debt levels and gross domestic product, along with many other economic statistics.

[6]Another way of understanding the decline in the debt–GDP ratio is to recognize that a rise in the price level reduces the value of the outstanding government debt in real terms—that is, in terms of the goods and services it can buy. So even though budget deficits did lead to a somewhat higher nominal amount of debt in this period, the continually rising price level (inflation) produced a lower real value of the government debt. The decline in the real amount of debt at the same time that real GDP was rising in this period then resulted in the decline in the debt–GDP ratio. For a fascinating discussion of how tricky it is to interpret deficit numbers, see Robert Eisner and Paul J. Pieper, "A New View of the Federal Debt and Budget Deficits," *American Economic Review* 74 (1984): 11–29.

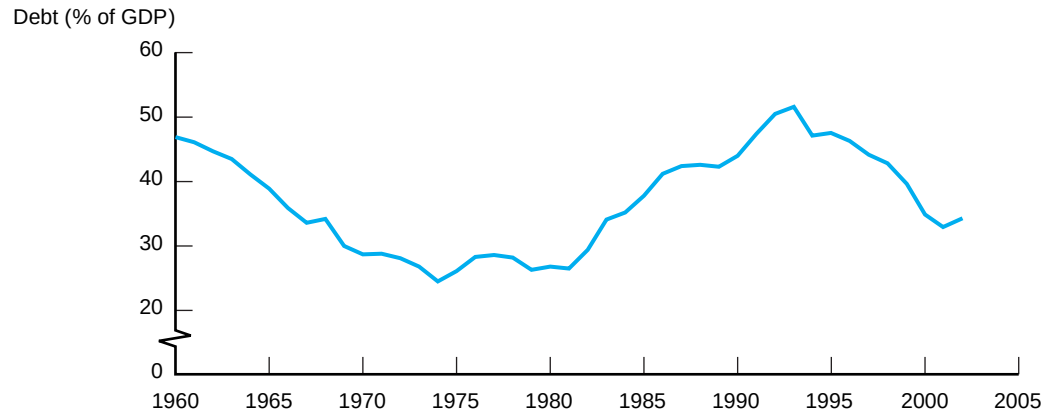


FIGURE 9 Government Debt-to-GDP Ratio, 1960–2002

Source: Economic Report of the President.

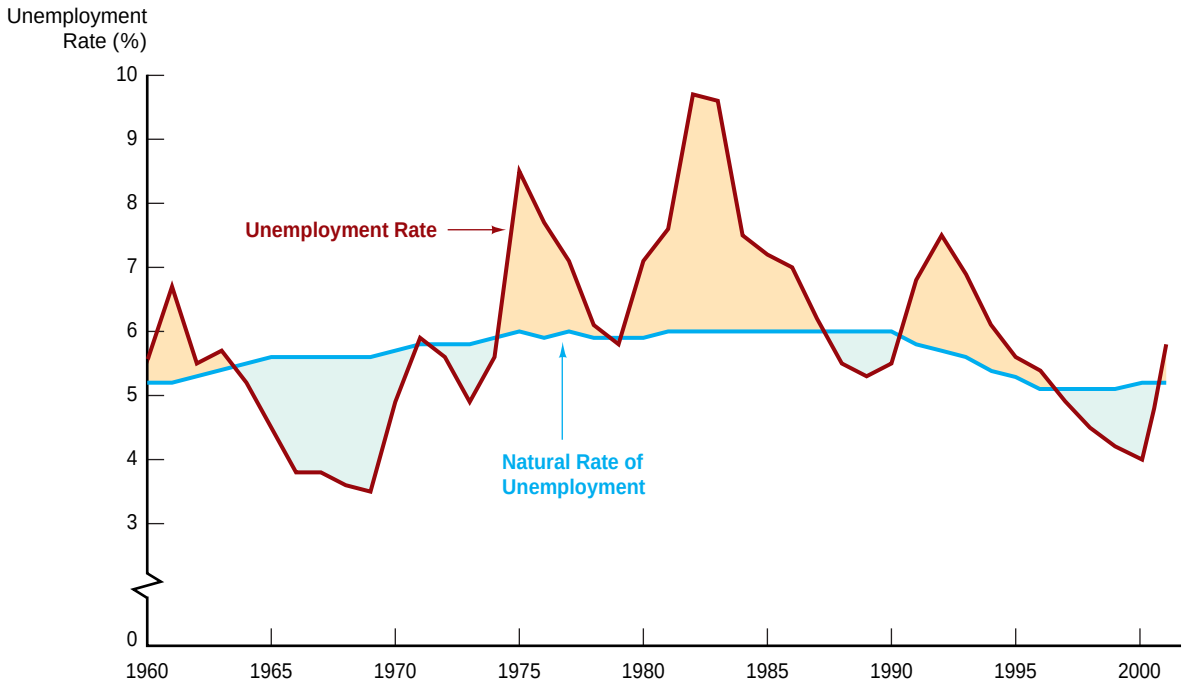


FIGURE 10 Unemployment and the Natural Rate of Unemployment, 1960–2002

Sources: Economic Report of the President and Congressional Budget Office.

target was the beginning of the most sustained inflationary episode in American history.

After 1975, the unemployment rate was regularly above the natural rate of unemployment, yet inflation continued. It appears that we have the phenomenon of a cost-push inflation described in Figure 5 (the impetus for which was the earlier demand-pull inflation). The persistence of inflation can be explained by the public's knowledge that government policy continued to be concerned with achieving high employment. With a higher rate of expected inflation arising initially from the demand-pull inflation, the aggregate supply curve in Figure 5 continued to shift leftward, causing a rise in unemployment that policymakers would try to eliminate by shifting the aggregate demand curve to the right. The result was a continuation of the inflation that had started in the 1960s.

Activist/Nonactivist Policy Debate

All economists have similar policy goals—they want to promote high employment and price stability—and yet they often have very different views on how policy should be conducted. Activists regard the self-correcting mechanism through wage and price adjustment (see Chapter 25) as very slow and hence see the need for the government to pursue active, accommodating, discretionary policy to eliminate high unemployment whenever it develops. Nonactivists, by contrast, believe that the performance of the economy would be improved if the government avoided active policy to eliminate unemployment. We will explore the activist/nonactivist policy debate by first looking at what the policy responses might be when the economy experiences high unemployment.

Responses to High Unemployment

Suppose that policymakers confront an economy that has moved to point 1′ in Figure 11. At this point, aggregate output $Y_{1'}$ is lower than the natural rate level, and the economy is suffering from high unemployment. Policymakers have two viable choices: If they are nonactivists and do nothing, the aggregate supply curve will eventually shift rightward over time, driving the economy from point 1′ to point 1, where full employment is restored. The accommodating, activist alternative is to try to eliminate the high unemployment by attempting to shift the aggregate demand curve rightward to AD_2 by pursuing expansionary policy (an increase in the money supply, increase in government spending, or lowering of taxes). If policymakers could shift the aggregate demand curve to AD_2 instantaneously, the economy would immediately move to point 2, where there is full employment. However, several types of lags prevent this immediate movement from occurring.

1. The *data lag* is the time it takes for policymakers to obtain the data that tell them what is happening in the economy. Accurate data on GDP, for example, are not available until several months after a given quarter is over.
2. The *recognition lag* is the time it takes for policymakers to be sure of what the data are signaling about the future course of the economy. For example, to minimize errors, the National Bureau of Economic Research (the organization that officially

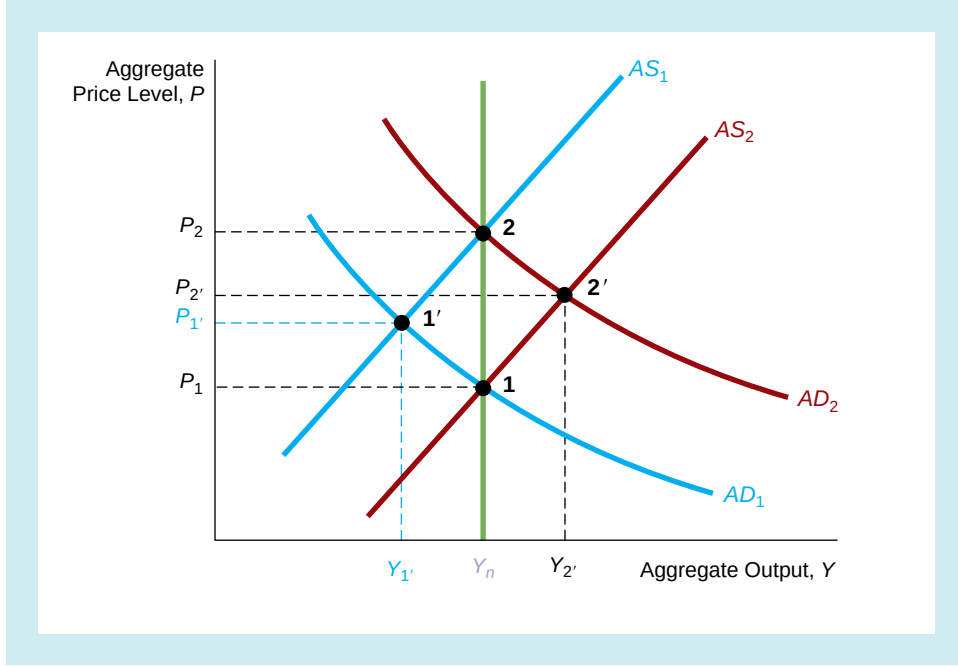


FIGURE 11 The Choice Between Activist and Nonactivist Policy

When the economy has moved to point 1′, the policymaker has two choices of policy: the nonactivist policy of doing nothing and letting the economy return to point 1 or the activist policy of shifting the aggregate demand curve to AD_2 to move the economy to point 2.

dates business cycles) will not declare the economy to be in recession until at least six months after it has determined that one has begun.

3. The *legislative lag* represents the time it takes to pass legislation to implement a particular policy. The legislative lag does not exist for most monetary policy actions such as open market operations. It can, however, be quite important for the implementation of fiscal policy, when it can sometimes take six months to a year to get legislation passed to change taxes or government spending.

4. The *implementation lag* is the time it takes for policymakers to change policy instruments once they have decided on the new policy. Again, this lag is unimportant for the conduct of open market operations because the Fed's trading desk can purchase or sell bonds almost immediately upon being told to do so by the Federal Open Market Committee. Actually implementing fiscal policy may take time, however; for example, getting government agencies to change their spending habits takes time, as does changing tax tables.

5. The *effectiveness lag* is the time it takes for the policy actually to have an impact on the economy. An important element of the monetarist viewpoint is that the effectiveness lag for changes in the money supply is long and variable (from several months to several years). Keynesians usually view fiscal policy as having a shorter effectiveness lag than monetary policy (fiscal policy takes approximately a year until its full effect is felt), but there is substantial uncertainty about how long this lag is.

Activist and Nonactivist Positions

Now that we understand the considerations that affect decisions by policymakers on whether to pursue an activist or nonactivist policy, we can examine when each of these policies would be preferable.

Case for an Activist Policy. Activists, such as the Keynesians, view the wage and price adjustment process as extremely slow. They consider a nonactivist policy costly, because the slow movement of the economy back to full employment results in a large loss of output. However, even though the five lags described result in delay of a year or two before the aggregate demand curve shifts to AD_2, the aggregate supply curve likewise moves very little during this time. The appropriate path for policymakers to pursue is thus an activist policy of moving the economy to point 2 in Figure 11.

Case for a Nonactivist Policy. Nonactivists, such as the monetarists, view the wage and price adjustment process as more rapid than activists do and consider nonactivist policy less costly because output is soon back at the natural rate level. They suggest that an activist, accommodating policy of shifting the aggregate demand curve to AD_2 is costly, because it produces more volatility in both the price level and output. The reason for this volatility is that the time it takes to shift the aggregate demand curve to AD_2 is substantial, whereas the wage and price adjustment process is more rapid. Hence before the aggregate demand curve shifts to the right, the aggregate supply curve will have shifted rightward to AS_2, and the economy will have moved from point 1′ to point 1, where it has returned to the natural rate level of output Y_n. After adjustment to the AS_2 curve is complete, the shift of the aggregate demand curve to AD_2 finally takes effect, leading the economy to point 2′ at the intersection of AD_2 and AS_2. Aggregate output at $Y_{2'}$ is now greater than the natural rate level ($Y_{2'} > Y_n$), so the aggregate supply curve will now shift leftward back to AS_1, moving the economy to point 2, where output is again at the natural rate level.

Although the activist policy eventually moves the economy to point 2 as policymakers intended, it leads to a sequence of equilibrium points—1′, 1, 2′, and 2—at which both output and the price level have been highly variable: Output overshoots its target level of Y_n, and the price level falls from $P_{1'}$ to P_1 and then rises to $P_{2'}$ and eventually to P_2. Because this variability is undesirable, policymakers would be better off pursuing the nonactivist policy, which moved the economy to point 1 and left it there.

Expectations and the Activist/Nonactivist Debate

Our analysis of inflation in the 1970s demonstrated that expectations about policy can be an important element in the inflation process. Allowing for expectations about policy to affect how wages are set (the wage-setting process) provides an additional reason for pursuing a nonactivist policy.

Do Expectations Favor a Nonactivist Approach? Does the possibility that expectations about policy matter to the wage-setting process strengthen the case for a nonactivist policy? The case for an activist policy states that with slow wage and price adjustment, the activist policy returns the economy to full employment at point 2 far more quickly than it takes to get to full employment at point 1 under nonactivist policy. However, the activist argument does not allow for the possibility (1) that expectations about policy matter to the wage-setting process and (2) that the economy might initially have moved from point 1 to point 1′ because an attempt by workers to raise their wages or a negative supply shock shifted the aggregate supply curve from AS_2 to AS_1. We must therefore ask the following question about activist policy: Will the aggregate supply curve continue to shift to the left after the economy has reached point 2, leading to cost-push inflation?

The answer to this question is yes *if* expectations about policy matter. Our discussion of cost-push inflation in Figure 5 suggested that if workers know that policy will be accommodating in the future, they will continue to push their wages up, and

the aggregate supply curve will keep on shifting leftward. As a result, policymakers are forced to accommodate the cost push by continuing to shift the aggregate demand curve to the right to eliminate the unemployment that develops. The accommodating, activist policy with its high employment target has the hidden cost or disadvantage that it may well lead to inflation.[7]

The main advantage of a nonaccommodating, nonactivist policy, in which policymakers do not try to shift the aggregate demand curve in response to the cost push, is that it will prevent inflation. As depicted in Figure 4, the result of an upward push on wages in the face of a nonaccommodating, nonactivist policy will be a period of unemployment above the natural rate level, which will eventually shift the aggregate supply curve and the price level back to their initial positions. The main criticism of this nonactivist policy is that the economy will suffer protracted periods of unemployment when the aggregate supply curve shifts leftward. Workers, however, would probably not push for higher wages to begin with if they knew that policy would be nonaccommodating, because their wage gains will lead to a protracted period of unemployment. A nonaccommodating, nonactivist policy may have not only the advantage of preventing inflation but also the hidden benefit of discouraging leftward shifts in the aggregate supply curve that lead to excessive unemployment.

In conclusion, ***if workers' opinions about whether policy is accommodating or nonaccommodating matter to the wage-setting process, the case for a nonactivist policy is much stronger***.

Do Expectations About Policy Matter to the Wage-Setting Process? The answer to this question is crucial to deciding whether activist or nonactivist policy is preferred and so has become a major topic of current research for economists, but the evidence is not yet conclusive. We can ask, however, whether expectations about policy do affect people's behavior in other contexts. This information will help us know if expectations regarding whether policy is accommodating are important to the wage-setting process.

As any good negotiator knows, convincing your opponent that you will be nonaccommodating is crucial to getting a good deal. If you are bargaining with a car dealer over price, for example, you must convince him that you can just as easily walk away from the deal and buy a car from a dealer on the other side of town. This principle also applies to conducting foreign policy—it is to your advantage to convince your opponent that you will go to war (be nonaccommodating) if your demands are not met. Similarly, if your opponent thinks that you will be accommodating, he will almost certainly take advantage of you (for an example, see Box 1). Finally, anyone who has dealt with a two-year-old child knows that the more you give in (pursue an accommodating policy), the more demanding the child becomes. People's expectations about policy *do* affect their behavior. Consequently, it is quite plausible that expectations about policy also affect the wage-setting process.[8]

---

[7] The issue that is being described here is the time-consistency problem described in Chapter 21.

[8] A recent development in monetary theory, new classical macroeconomics, strongly suggests that expectations about policy are crucial to the wage-setting process and the movements of the aggregate supply curve. We will explore why new classical macroeconomics comes to this conclusion in Chapter 28, when we discuss the implications of the rational expectations hypothesis, which states that expectations are formed using all available information, including expectations about policy.

Box 1

Perils of Accommodating Policy

The Terrorism Dilemma. A major dilemma confronting our foreign policy in recent years is whether to cave in to the demands of terrorists when they are holding American hostages. Because our hearts go out to the hostages and their families, we might be tempted to pursue an accommodating policy of giving in to the terrorists to bring the hostages safely back home. However, pursuing this accommodating policy is likely to encourage terrorists to take hostages in the future.

The terrorism dilemma illustrates the principle that opponents are more likely to take advantage of you in the future if you accommodate them now. Recognition of this principle, which demonstrates the perils of accommodating policy, explains why governments in countries such as the United States and Israel have been reluctant to give in to terrorist demands even though it has sometimes resulted in the death of hostages.

Rules Versus Discretion: Conclusions

The following conclusions can be generated from our analysis: Activists believe in the use of discretionary policy to eliminate excessive unemployment whenever it develops, because they view the wage and price adjustment process as sluggish and unresponsive to expectations about policy. Nonactivists, by contrast, believe that a discretionary policy that reacts to excessive unemployment is counter-productive, because wage and price adjustment is rapid and because expectations about policy can matter to the wage-setting process. Nonactivists thus advocate the use of a policy rule to keep the aggregate demand curve from fluctuating away from the trend rate of growth of the natural rate level of output. Monetarists, who adhere to the nonactivist position and who also see money as the sole source of fluctuations in the aggregate demand curve, in the past advocated a policy rule whereby the Federal Reserve keeps the money supply growing at a constant rate. This monetarist rule is referred to as a **constant-money-growth-rate rule**. Because of the misbehavior of velocity of M1 and M2, monetarists such as Bennett McCallum and Alan Meltzer of Carnegie-Mellon University have advocated a rule for the growth of the monetary base that is adjusted for past velocity changes.

As our analysis indicates, an important element for the success of a nonaccommodating policy rule is that it be *credible*: The public must believe that policymakers will be tough and not accede to a cost push by shifting the aggregate demand curve to the right to eliminate unemployment. In other words, government policymakers need credibility as inflation-fighters in the eyes of the public. Otherwise, workers will be more likely to push for higher wages, which will shift the aggregate supply curve leftward after the economy reaches full employment at a point such as point 2 in Figure 11 and will lead to unemployment or inflation (or both). Alternatively, a credible, nonaccommodating policy rule has the benefit that it makes a cost push less likely and thus helps prevent inflation and potential increases in unemployment. The following application suggests that recent historical experience is consistent with the importance of credibility to successful policymaking.

Application

Importance of Credibility to Volcker's Victory over Inflation

In the period from 1965 through the 1970s, policymakers had little credibility as inflation-fighters—a well-deserved reputation, as they pursued an accommodating policy to achieve high employment. As we have seen, the outcome was not a happy one. Inflation soared to double-digit levels, while the unemployment rate remained high. To wring inflation out of the system, the Federal Reserve under Chairman Paul Volcker put the economy through two back-to-back recessions in 1980 and 1981–1982 (see Chapter 18). (The data on inflation, money growth, and unemployment in this period are shown in Figures 8 and 10.) Only after the 1981–1982 recession—the most severe in the postwar period, with unemployment above the 10% level—did Volcker establish credibility for the Fed's anti-inflation policy. By the end of 1982, inflation was running at a rate of less than 5%.

One indication of Volcker's credibility came in 1983 when the money growth rate accelerated dramatically and yet inflation did not rise. Workers and firms were convinced that if inflation reared its head, Volcker would pursue a nonaccommodating policy of quashing it. They did not raise wages and prices, which would have shifted the aggregate supply curve leftward and would have led to both inflation and unemployment. The success of Volcker's anti-inflation policy continued throughout the rest of his term as chairman, which ended in 1987; unemployment fell steadily, while the inflation rate remained below 5%. Volcker's triumph over inflation was achieved because he obtained credibility the hard way—he earned it.

Summary

1. Milton Friedman's famous proposition that "inflation is always and everywhere a monetary phenomenon" is supported by the following evidence: Every country that has experienced a sustained, high inflation has also experienced a high rate of money growth.
2. Aggregate demand and supply analysis shows that Keynesian and monetarist views of the inflation process are not very different. Both believe that high inflation can occur only if there is a high rate of money growth. As long as we recognize that by inflation we mean a rapid and continuing increase in the price level, almost all economists agree with Friedman's proposition.
3. Although high inflation is "always and everywhere a monetary phenomenon" in the sense that it cannot occur without a high rate of money growth, there are reasons why inflationary monetary policy comes about. The two underlying reasons are the adherence of policymakers to a high employment target and the presence of persistent government budget deficits.
4. Activists believe in the use of discretionary policy to eliminate excessive unemployment whenever it occurs because they view wage and price adjustment as sluggish and unresponsive to expectations about policy. Nonactivists take the opposite view and believe that discretionary policy is counterproductive. In addition, they regard the credibility of a nonaccommodating (nonactivist) anti-inflation policy as crucial to its success.

Key Terms

accommodating policy, p. 640
constant-money-growth-rate rule, p. 654
cost-push inflation, p. 639
demand-pull inflation, p. 639
government budget constraint, p. 643
monetizing the debt, p. 644
printing money, p. 644
Ricardian equivalence, p. 645

Questions and Problems

Questions marked with an asterisk are answered at the end of the book in an appendix, "Answers to Selected Questions and Problems."

1. "There are frequently years when the inflation rate is high and yet money growth is quite low. Therefore, the statement that inflation is a monetary phenomenon cannot be correct." Comment.

*2. Why do economists focus on historical episodes of hyperinflation to decide whether inflation is a monetary phenomenon?

3. "Since increases in government spending raise the aggregate demand curve in Keynesian analysis, fiscal policy by itself can be the source of inflation." Is this statement true, false, or uncertain? Explain your answer.

*4. "A cost-push inflation occurs as a result of workers' attempts to push up their wages. Therefore, inflation does not have to be a monetary phenomenon." Is this statement true, false, or uncertain? Explain your answer.

5. "Because government policymakers do not consider inflation desirable, their policies cannot be the source of inflation." Is this statement true, false, or uncertain? Explain your answer.

*6. "A budget deficit that is only temporary cannot be the source of inflation." Is this statement true, false, or uncertain? Explain your answer.

7. How can the Fed's desire to prevent high interest rates lead to inflation?

*8. "If the data and recognition lags could be reduced, activist policy would more likely be beneficial to the economy." Is this statement true, false, or uncertain? Explain your answer.

9. "The more sluggish wage and price adjustment is, the more variable output and the price level are when an activist policy is pursued." Is this statement true, false, or uncertain? Explain your answer.

*10. "If the public believes that the monetary authorities will pursue an accommodating policy, a cost-push inflation is more likely to develop." Is this statement true, false, or uncertain? Explain your answer.

11. Why are activist policies to eliminate unemployment more likely to lead to inflation than nonactivist policies?

*12. "The less important expectations about policy are to movements of the aggregate supply curve, the stronger the case is for activist policy to eliminate unemployment." Is this statement true, false, or uncertain? Explain your answer.

13. If the economy's self-correcting mechanism works slowly, should the government necessarily pursue an activist policy to eliminate unemployment?

*14. "To prevent inflation, the Fed should follow Teddy Roosevelt's advice: 'Speak softly and carry a big stick.'" What would the Fed's "big stick" be? What is the statement trying to say?

15. In a speech early in the Iraq-Kuwait crisis in 1990, President George Bush stated that although his heart went out to the hostages held by Saddam Hussein, he would not let this hostage-taking deter the United States from insisting on the withdrawal of Iraq from Kuwait. Do you think that Bush's position made sense? Explain why or why not.

Web Exercises

1. Figure 8 reports the inflation rate from 1960 to 2002. As this chapter states, inflation continues to be a major a factor in economic policy. Go to ftp://ftp.bls.gov/pub/special.requests/cpi/cpiai.txt. Move data into Excel using the method described at the end of Chapter 1. Delete all but the first and last column (date and annual CPI). Graph this data and compare it to Figure 8.
   a. Has inflation increased or decreased since the end of 2002?
   b. When was inflation at its highest?
   c. When was inflation at its lowest?
   d. Have we ever had a period of deflation? If so, when?
   e. Have we ever had a period of hyperinflation? If so, when?

2. It can be an interesting exercise to compare the purchasing power of the dollar over different periods in history. Go to www.bls.gov/cpi/ and scroll down to the link to the "inflation calculator." Use this calculator to compute the following:
   a. If a new home cost $125,000 in 2002, what would it have cost in 1950?
   b. The average household income in 2002 was about $37,000. How much would this have been in 1945?
   c. An average new car cost about $18,000 in 2002. What would this have cost in 1945?
   d. Using the results you found in Questions b and c, does a car consume more or less of average household income in 2002 than in 1945?

Chapter 28

Rational Expectations: Implications for Policy

PREVIEW

After World War II, economists, armed with Keynesian models (such as the *ISLM* model) that described how government policies could be used to manipulate employment and output, felt that activist policies could reduce the severity of business cycle fluctuations without creating inflation. In the 1960s and 1970s, these economists got their chance to put their policies into practice (see Chapter 27), but the results were not what they had anticipated. The economic record for that period is not a happy one: Inflation accelerated, the rate often climbing above 10%, while unemployment figures deteriorated from those of the 1950s.[1]

In the 1970s and 1980s, economists, including Robert Lucas of the University of Chicago and Thomas Sargent, now at New York University, used the rational expectations theory discussed in Chapter 7 to examine why activist policies appear to have performed so poorly. Their analysis cast doubt on whether macroeconomic models can be used to evaluate the potential effects of policy and on whether policy can be effective when the public *expects* that it will be implemented. Because the analysis of Lucas and Sargent has such strong implications for the way policy should be conducted, it has been labeled the *rational expectations revolution.*[2]

This chapter examines the analysis behind the rational expectations revolution. We start first with the Lucas critique, which indicates that because expectations are important in economic behavior, it may be quite difficult to predict what the outcome of an activist policy will be. We then discuss the effect of rational expectations on the aggregate demand and supply analysis developed in Chapter 25 by exploring three models that incorporate expectations in different ways.

A comparison of all three models indicates that the existence of rational expectations makes activist policies less likely to be successful and raises the issue of credibility as an important element affecting policy outcomes. With rational expectations, an essential ingredient to a successful anti-inflation policy is the credibility of the policy in the eyes of the public. The rational expectations revolution is now at the center of many of the current debates in monetary theory that have major implications for how monetary and fiscal policy should be conducted.

[1] Some of the deterioration can be attributed to supply shocks in 1973–1975 and 1978–1980.

[2] Other economists who have been active in promoting the rational expectations revolution are Robert Barro of Harvard University, Bennett McCallum of Carnegie-Mellon University, Edward Prescott of the University of Minnesota, and Neil Wallace of Pennsylvania State University.

The Lucas Critique of Policy Evaluation

http://cepa.newschool.edu/het/profiles/lucas.htm
A brief biography of Robert Lucas, including a list of his publications.

In his famous paper "Econometric Policy Evaluation: A Critique," Robert Lucas presented an argument that had devastating implications for the usefulness of conventional **econometric models** (models whose equations are estimated with statistical procedures) for evaluating policy.[3] Economists developed these models for two purposes: to forecast economic activity and to evaluate the effects of different policies. Although Lucas's critique had nothing to say about the usefulness of these models as forecasting tools, he argued that they could not be relied on to evaluate the potential impact of particular policies on the economy.

Econometric Policy Evaluation

To understand Lucas's argument, we must first understand econometric policy evaluation: how econometric models are used to evaluate policy. For example, we can examine how the Federal Reserve uses its econometric model in making decisions about the future course of monetary policy. The model contains equations that describe the relationships among hundreds of variables. These relationships are assumed to remain constant and are estimated using past data. Let's say that the Fed wants to know the effect on unemployment and inflation of a decrease in the fed funds rate from 5% to 4%. It feeds the new, lower fed funds rate into a computer that contains the model, and the model then provides an answer about how much unemployment will fall as a result of the lower fed funds rate and how much the inflation rate will rise. Other possible policies, such as a rise in the fed funds rate by one percentage point, might also be fed into the model. After a series of these policies have been tried out, the policymakers at the Fed can see which policies produce the most desirable outcome for unemployment and inflation.

Lucas's challenge to this procedure for evaluating policies is based on a simple principle of rational expectations theory: *The way in which expectations are formed (the relationship of expectations to past information) changes when the behavior of forecasted variables changes.* So when policy changes, the relationship between expectations and past information will change, and because expectations affect economic behavior, the relationships in the econometric model will change. The econometric model, which has been estimated with past data, is then no longer the correct model for evaluating the response to this policy change and may consequently prove highly misleading.

Example: The Term Structure of Interest Rates

The best way to understand Lucas's argument is to look at a concrete example involving only one equation typically found in econometric models: the term structure equation. The equation relates the long-term interest rate to current and past values of the short-term interest rate. It is one of the most important equations in Keynesian econometric models because the long-term interest rate, not the short-term rate, is the one believed to have an impact on aggregate demand.

In Chapter 6, we learned that the long-term interest rate is related to an average of expected future short-term interest rates. Suppose that in the past, when the short-term rate rose, it quickly fell back down again; that is, any increase was temporary. Because rational expectations theory suggests that any rise in the short-term interest rate is expected to be only temporary, a rise should have only a minimal effect on the

[3] *Carnegie-Rochester Conference Series on Public Policy* 1 (1976): 19–46.

average of expected future short-term rates. It will cause the long-term interest rate to rise by a negligible amount. The term structure relationship estimated using past data will then show only a weak effect on the long-term interest rate of changes in the short-term rate.

Suppose the Fed wants to evaluate what will happen to the economy if it pursues a policy that is likely to raise the short-term interest rate from a current level of 5% to a permanently higher level of 8%. The term structure equation that has been estimated using past data will indicate that there will be just a small change in the long-term interest rate. However, if the public recognizes that the short-term rate is rising to a permanently higher level, rational expectations theory indicates that people will no longer expect a rise in the short-term rate to be temporary. Instead, when they see the interest rate rise to 8%, they will expect the average of future short-term interest rates to rise substantially, and so the long-term interest rate will rise greatly, not minimally as the estimated term structure equation suggests. You can see that evaluating the likely outcome of the change in Fed policy with an econometric model can be highly misleading.

The term structure example also demonstrates another aspect of the Lucas critique. The effects of a particular policy depend critically on the public's expectations about the policy. If the public expects the rise in the short-term interest rate to be merely temporary, the response of long-term interest rates, as we have seen, will be negligible. If, however, the public expects the rise to be more permanent, the response of long-term rates will be far greater. ***The Lucas critique points out not only that conventional econometric models cannot be used for policy evaluation, but also that the public's expectations about a policy will influence the response to that policy.***

The term structure equation discussed here is only one of many equations in econometric models to which the Lucas critique applies. In fact, Lucas uses the examples of consumption and investment equations in his paper. One attractive feature of the term structure example is that it deals with expectations in a financial market, a sector of the economy for which the theory and empirical evidence supporting rational expectations are very strong. The Lucas critique should also apply, however, to sectors of the economy for which rational expectations theory is more controversial, because the basic principle of the Lucas critique is not that expectations are always rational but rather that the formation of expectations changes when the behavior of a forecasted variable changes. This less stringent principle is supported by the evidence in sectors of the economy other than financial markets.

New Classical Macroeconomic Model

We now turn to the implications of rational expectations for the aggregate demand and supply analysis we studied in Chapter 25. The first model we examine that views expectations as rational is the *new classical macroeconomic model* developed by Robert Lucas and Thomas Sargent, among others. In the new classical model, all wages and prices are completely flexible with respect to expected changes in the price level; that is, a rise in the expected price level results in an immediate and equal rise in wages and prices because workers try to keep their *real* wages from falling when they expect the price level to rise.

This view of how wages and prices are set indicates that a rise in the expected price level causes an immediate leftward shift in the aggregate supply curve, which

leaves real wages unchanged and aggregate output at the natural rate (full-employment) level if expectations are realized. This model then suggests that anticipated policy has no effect on aggregate output and unemployment; only unanticipated policy has an effect.

Effects of Unanticipated and Anticipated Policy

First, let us look at the short-run response to an unanticipated (unexpected) policy such as an unexpected increase in the money supply.

In Figure 1, the aggregate supply curve AS_1 is drawn for an expected price level P_1. The initial aggregate demand curve AD_1 intersects AS_1 at point 1, where the realized price level is at the expected price level P_1 and aggregate output is at the natural rate level Y_n. Because point 1 is also on the long-run aggregate supply curve at Y_n, there is no tendency for the aggregate supply to shift. The economy remains in long-run equilibrium.

Suppose the Fed suddenly decides the unemployment rate is too high and so makes a large bond purchase that is unexpected by the public. The money supply increases, and the aggregate demand curve shifts rightward to AD_2. Because this shift is unexpected, the expected price level remains at P_1 and the aggregate supply curve remains at AS_1. Equilibrium is now at point 2′, the intersection of AD_2 and AS_1. Aggregate output increases above the natural rate level to $Y_{2'}$ and the realized price level increases to $P_{2'}$.

If, by contrast, the public expects that the Fed will make these open market purchases in order to lower unemployment because they have seen it done in the past, the expansionary policy will be anticipated. The outcome of such anticipated expansionary policy is illustrated in Figure 2. Because expectations are rational, workers and firms recognize that an expansionary policy will shift the aggregate demand curve

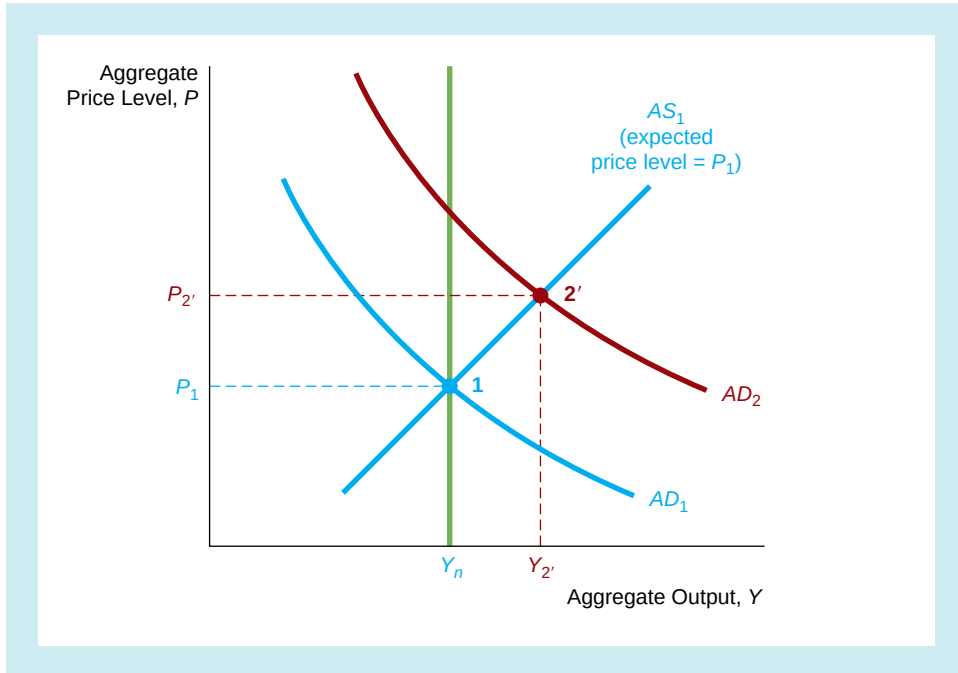


FIGURE 1 Short-Run Response to Unanticipated Expansionary Policy in the New Classical Model

Initially, the economy is at point 1 at the intersection of AD_1 and AS_1 (expected price level = P_1). An expansionary policy shifts the aggregate demand curve to AD_2, but because this is unexpected, the aggregate supply curve remains fixed at AS_1. Equilibrium now occurs at point 2′—aggregate output has increased above the natural rate level to $Y_{2'}$, and the price level has increased to $P_{2'}$.

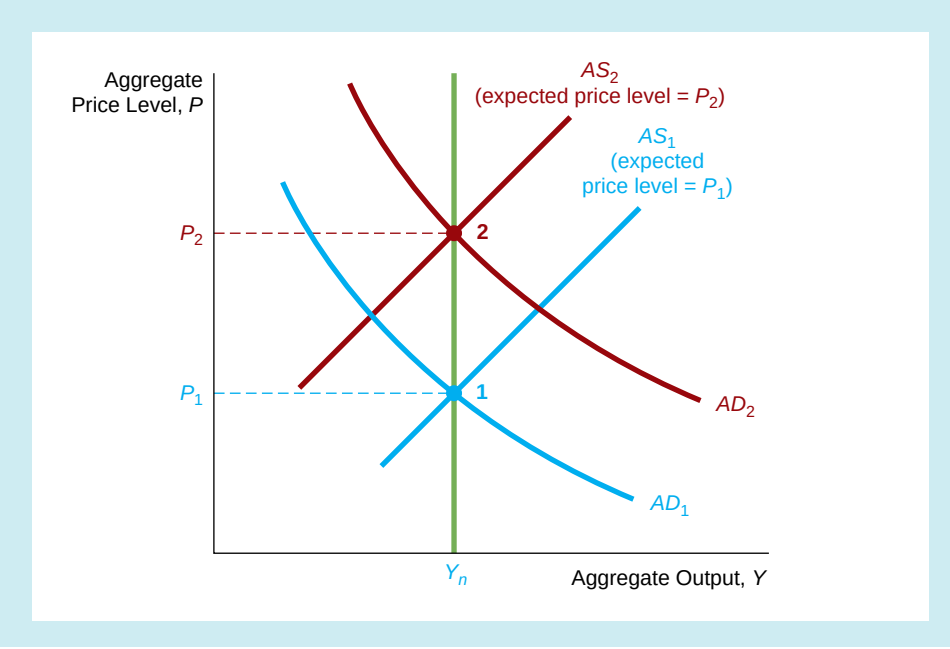


FIGURE 2 Short-Run Response to Anticipated Expansionary Policy in the New Classical Model

The expansionary policy shifts the aggregate demand curve rightward to AD_2, but because this policy is expected, the aggregate supply curve shifts leftward to AS_2. The economy moves to point 2, where aggregate output is still at the natural rate level but the price level has increased to P_2.

to the right and will expect the aggregate price level to rise to P_2. Workers will demand higher wages so that their real earnings will remain the same when the price level rises. The aggregate supply curve then shifts leftward to AS_2 and intersects AD_2 at point 2, an equilibrium point where aggregate output is at the natural rate level Y_n and the price level has risen to P_2.

The new classical macroeconomic model demonstrates that aggregate output does not increase as a result of anticipated expansionary policy and that the economy immediately moves to a point of long-run equilibrium (point 2) where aggregate output is at the natural rate level. Although Figure 2 suggests why this occurs, we have not yet proved why an anticipated expansionary policy shifts the aggregate supply curve to exactly AS_2 (corresponding to an expected price level of P_2) and hence why aggregate output *necessarily* remains at the natural rate level. The proof is somewhat difficult and is dealt with in Box 1.

The new classical model has the word *classical* associated with it because when policy is anticipated, the new classical model has a property that is associated with the classical economists of the nineteenth and early twentieth centuries: Aggregate output remains at the natural rate level. Yet the new classical model allows aggregate output to fluctuate away from the natural rate level as a result of *unanticipated* movements in the aggregate demand curve. The conclusion from the new classical model is a striking one: ***Anticipated policy has no effect on the business cycle; only unanticipated policy matters.***[4]

[4]Note that the new classical view in which anticipated policy has no effect on the business cycle does not imply that anticipated policy has no effect on the overall health of the economy. For example, the new classical analysis does not rule out possible effects of anticipated policy on the natural rate of output Y_n, which can benefit the public.

Box 1

Proof of the Policy Ineffectiveness Proposition

The proof that in the new classical macroeconomic model aggregate output *necessarily* remains at the natural rate level when there is anticipated expansionary policy is as follows. In the new classical model, the expected price level for the aggregate supply curve occurs at its intersection with the long-run aggregate supply curve (see Figure 2). The optimal forecast of the price level is given by the intersection of the aggregate supply curve with the anticipated aggregate demand curve AD_2. If the aggregate supply curve is to the right of AS_2 in Figure 2, it will intersect AD_2 at a price level lower than the expected level (at the intersection of this aggregate supply curve and the Y_n line). The optimal forecast of the price level will then not equal the expected price level, thereby violating the rationality of expectations. A similar argument can be made to show that when the aggregate supply curve is to the left of AS_2, the assumption of rational expectations is violated. Only when the aggregate supply curve is at AS_2 (corresponding to an expected price level of P_2) are expectations rational because the optimal forecast equals the expected price level. As we see in Figure 2, the AS_2 curve implies that aggregate output remains at the natural rate level as a result of the anticipated expansionary policy.

This conclusion has been called the **policy ineffectiveness proposition**, because it implies that one anticipated policy is just like any other; it has no effect on output fluctuations. You should recognize that this proposition does not rule out output effects from policy changes. If the policy is a surprise (unanticipated), it will have an effect on output.[5]

Can an Expansionary Policy Lead to a Decline in Aggregate Output?

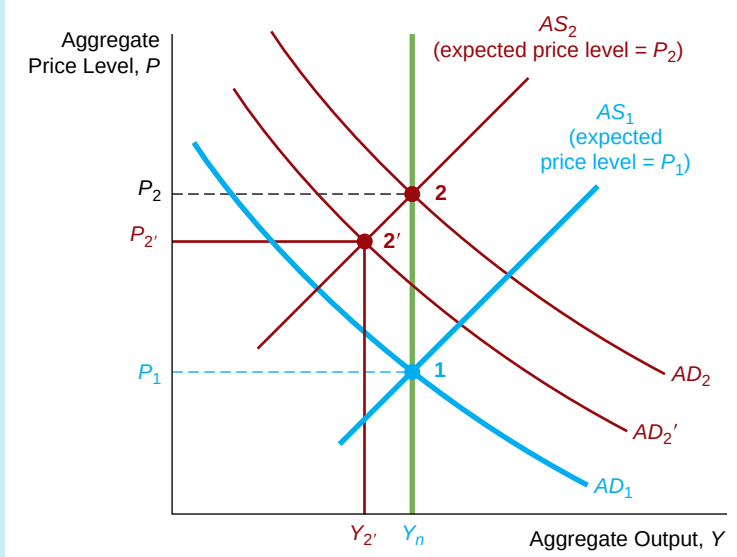
Another important feature of the new classical model is that an expansionary policy, such as an increase in the rate of money growth, can lead to a *decline* in aggregate output if the public expects an even more expansionary policy than the one actually implemented. There will be a surprise in the policy, but it will be negative and drive output down. Policymakers cannot be sure if their policies will work in the intended direction.

To see how an expansionary policy can lead to a decline in aggregate output, let us turn to the aggregate supply and demand diagram in Figure 3. Initially we are at point 1, the intersection of AD_1 and AS_1; output is Y_n, and the price level is P_1. Now suppose that the public expects the Fed to increase the money supply in order to shift the aggregate demand curve to AD_2. As we saw in Figure 2, the aggregate supply curve shifts leftward to AS_2, because the price level is expected to rise to P_2. Suppose that the expansionary policy engineered by the Fed actually falls short of what was expected so that the aggregate demand curve shifts only to $AD_{2'}$. The economy will move to point 2′, the intersection of the aggregate supply curve AS_2 and the aggregate demand curve $AD_{2'}$. The result of the mistaken expectation is that output falls to $Y_{2'}$, while the price level rises to $P_{2'}$ rather than P_2. An expansionary policy that is less expansionary than anticipated leads to an output movement directly opposite to that intended.

[5]Thomas Sargent and Neil Wallace, "'Rational' Expectations, the Optimal Monetary Instrument, and the Optimal Money Supply Rule," *Journal of Political Economy* 83 (1975): 241–254, first demonstrated the full implications of the policy ineffectiveness proposition.

FIGURE 3 Short-Run Response to an Expansionary Policy That Is Less Expansionary Than Expected in the New Classical Model

Because the public expects the aggregate demand curve to shift to AD_2, the aggregate supply curve shifts to AS_2 (expected price level $= P_2$). When the actual expansionary policy falls short of the public's expectation (the aggregate demand curve merely shifts to $AD_{2'}$), the economy ends up at point 2′, at the intersection of $AD_{2'}$ and AS_2. Despite the expansionary policy, aggregate output falls to $Y_{2'}$.



Study Guide

Mastering the new classical macroeconomic model, as well as the new Keynesian model in the next section, requires practice. Make sure that you can draw the aggregate demand and supply curves that explain what happens in each model when there is a contractionary policy that is (1) unanticipated, (2) anticipated, and (3) less contractionary than anticipated.

Implications for Policymakers

The new classical model, with its policy ineffectiveness proposition, has two important lessons for policymakers: It illuminates the distinction between the effects of anticipated versus unanticipated policy actions, and it demonstrates that policymakers cannot know the outcome of their decisions without knowing the public's expectations regarding them.

At first you might think that policymakers can still use policy to stabilize the economy. Once they figure out the public's expectations, they can know what effect their policies will have. There are two catches to such a conclusion. First, it may be nearly impossible to find out what the public's expectations are, given that the public consists of close to 300 million U.S. citizens. Second, even if it were possible, policymakers would run into further difficulties, because the public has rational expectations and will try to guess what policymakers plan to do. Public expectations do not remain fixed while policymakers are plotting a surprise—the public will revise its expectations, and policies will have no predictable effect on output.[6]

[6]This result follows from one of the implications of rational expectations: The forecast error of expectations about policy (the deviation of actual policy from expectations of policy) must be unpredictable. Because output is affected only by unpredictable (unanticipated) policy changes in the new classical model, policy effects on output must be unpredictable as well.

Where does this lead us? Should the Fed and other policymaking agencies pack up, lock the doors, and go home? In a sense, the answer is yes. The new classical model implies that discretionary stabilization policy cannot be effective and might have undesirable effects on the economy. Policymakers' attempts to use discretionary policy may create a fluctuating policy stance that leads to unpredictable policy surprises, which in turn cause undesirable fluctuations around the natural rate level of aggregate output. To eliminate these undesirable fluctuations, the Fed and other policymaking agencies should abandon discretionary policy and generate as few policy surprises as possible.

As we have seen in Figure 2, even though anticipated policy has no effect on aggregate output in the new classical model, it *does* have an effect on the price level. The new classical macroeconomists care about anticipated policy and suggest that policy rules be designed so that the price level will remain stable.

New Keynesian Model

In the new classical model, all wages and prices are completely flexible with respect to expected changes in the price level; that is, a rise in the expected price level results in an immediate and equal rise in wages and prices. Many economists who accept rational expectations as a working hypothesis do not accept the characterization of wage and price flexibility in the new classical model. These critics of the new classical model, called *new Keynesians*, object to complete wage and price flexibility and identify factors in the economy that prevent some wages and prices from rising fully with a rise in the expected price level.

www.federalreserve.gov/pubs/feds/2001/200113/200113pap.pdf

The Federal Reserve recently published a paper discussing the new Keynesian model and price stickiness.

Long-term labor contracts are one source of rigidity that prevents wages and prices from responding fully to changes in the expected price level (called *wage–price stickiness*) . For example, workers might find themselves at the end of the first year of a three-year wage contract that specifies the wage rate for the coming two years. Even if new information appeared that would make them raise their expectations of the inflation rate and the future price level, they could not do anything about it because they are locked into a wage agreement. Even with a high expectation about the price level, the wage rate will not adjust. In two years, when the contract is renegotiated, both workers and firms may build the expected inflation rate into their agreement, but they cannot do so immediately.

Another source of rigidity is that firms may be reluctant to change wages frequently even when there are no explicit wage contracts, because such changes may affect the work effort of the labor force. For example, a firm may not want to lower workers' wages when unemployment is high, because this might result in poorer worker performance. Price stickiness may also occur because firms engage in fixed-price contracts with their suppliers or because it is costly for firms to change prices frequently. All of these rigidities (which diminish wage and price flexibility), even if they are not present in all wage and price arrangements, suggest that an increase in the expected price level might not translate into an immediate and complete adjustment of wages and prices.

Although the new Keynesians do not agree with the complete wage and price flexibility of the new classical macroeconomics, they nevertheless recognize the importance of expectations to the determination of aggregate supply and are willing to accept rational expectations theory as a reasonable characterization of how expectations are

formed. The model they have developed, the *new Keynesian model*, assumes that expectations are rational but does not assume complete wage and price flexibility; instead, it assumes that wages and prices are sticky. Its basic conclusion is that unanticipated policy has a larger effect on aggregate output than anticipated policy (as in the new classical model). However, in contrast to the new classical model, the policy ineffectiveness proposition does not hold in the new Keynesian model: Anticipated policy *does* affect aggregate output and the business cycle.

Effects of Unanticipated and Anticipated Policy

In panel (a) of Figure 4, we look at the short-run response to an unanticipated expansionary policy for the new Keynesian model. The analysis is identical to that of the new classical model. We again start at point 1, where the aggregate demand curve AD_1 intersects the aggregate supply curve AS_1 at the natural rate level of output and price level P_1. When the Fed pursues its expansionary policy of purchasing bonds and raising the money supply, the aggregate demand curve shifts rightward to AD_2. Because the expansionary policy is unanticipated, the expected price level remains unchanged, leaving the aggregate supply curve unchanged. Thus the economy moves to point U, where aggregate output has increased to Y_U and the price level has risen to P_U.

In panel (b), we see what happens when the Fed's expansionary policy that shifts the aggregate demand curve from AD_1 to AD_2 is anticipated. Because the expansionary policy is anticipated and expectations are rational, the expected price level increases, causing wages to increase and the aggregate supply curve to shift to the left. Because of rigidities that do not allow *complete* wage and price adjustment, the aggregate supply curve does not shift all the way to AS_2 as it does in the new classical model. Instead, it moves to AS_A, and the economy settles at point A, the intersection of AD_2 and AS_A. Aggregate output has risen above the natural rate level to Y_A, while the price level has increased to P_A. ***Unlike the new classical model, in the new Keynesian model anticipated policy does have an effect on aggregate output.***

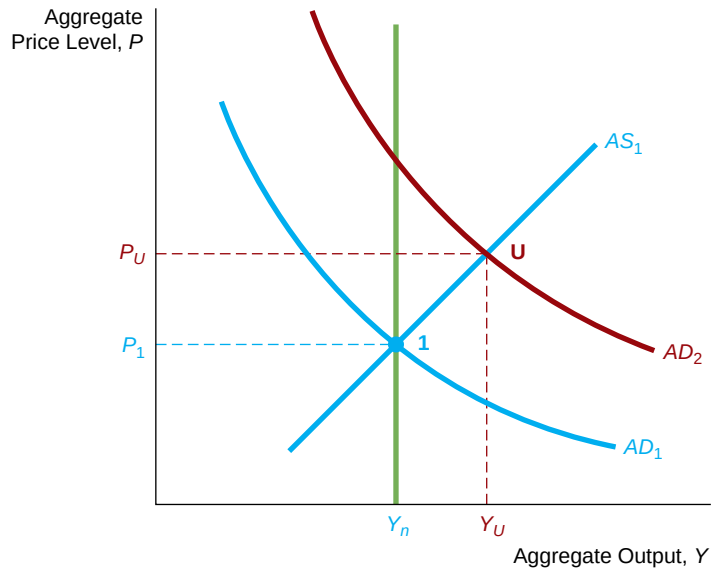
We can see in Figure 4 that Y_U is greater than Y_A, meaning that the output response to unanticipated policy is greater than to anticipated policy. It is greater because the aggregate supply curve does not shift when policy is unanticipated, causing a lower price level and hence a higher level of output. We see that ***like the new classical model, the new Keynesian model distinguishes between the effects of anticipated versus unanticipated policy, with unanticipated policy having a greater effect***.

Implications for Policymakers

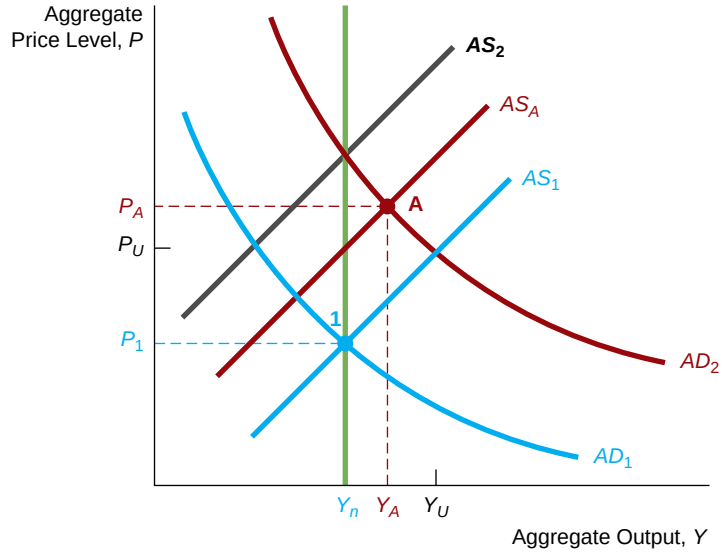
Because the new Keynesian model indicates that anticipated policy has an effect on aggregate output, it does not rule out beneficial effects from activist stabilization policy, in contrast to the new classical model. It does warn the policymaker that designing such a policy will not be an easy task, because the effects of anticipated and unanticipated policy can be quite different. As in the new classical model, to predict the outcome of their actions, policymakers must be aware of the public's expectations about those actions. Policymakers face similar difficulties in devising successful policies in both the new classical and new Keynesian models.

Comparison of the Two New Models with the Traditional Model

To obtain a clearer picture of the impact of the rational expectations revolution on our analysis of the aggregate economy, we can compare the two rational expectations mod-



(a) Response to an unanticipated expansionary policy



(b) Response to an anticipated expansionary policy

FIGURE 4 Short-Run Response to Expansionary Policy in the New Keynesian Model

The expansionary policy that shifts aggregate demand to AD_2 has a bigger effect on output when it is unanticipated than when it is anticipated. When the expansionary policy is unanticipated in panel (a), the short-run aggregate supply curve does not shift, and the economy moves to point U, so that aggregate output increases to Y_U and the price level rises to P_U. When the policy is anticipated in panel (b), the short-run aggregate supply curve shifts to AS_A (but not all the way to AS_2 because rigidities prevent complete wage and price adjustment), and the economy moves to point A so that aggregate output rises to Y_A (which is less than Y_U) and the price level rises to P_A (which is higher than P_U).

els (the new classical macroeconomic model and the new Keynesian model) to a model that we call, for lack of a better name, the *traditional model*. In the traditional model, expectations are *not* rational. That model uses adaptive expectations (mentioned in Chapter 7), expectations based solely on past experience. The traditional model views expected inflation as an average of past inflation rates. This average is not affected by

the public's predictions of future policy; hence predictions of future policy do not affect the aggregate supply curve.

First we will examine the short-run output and price responses in the three models. Then we will examine the implications of these models for both stabilization and anti-inflation policies.

Study Guide

As a study aid, the comparison of the three models is summarized in Table 1. You may want to refer to the table as we proceed with the comparison.

Short-Run Output and Price Responses

Figure 5 compares the response of aggregate output and the price level to an expansionary policy in the three models. Initially, the economy is at point 1, the intersection of the aggregate demand curve AD_1 and the aggregate supply curve AS_1. When the expansionary policy occurs, the aggregate demand curve shifts to AD_2. If the

SUMMARY Table 1 The Three Models

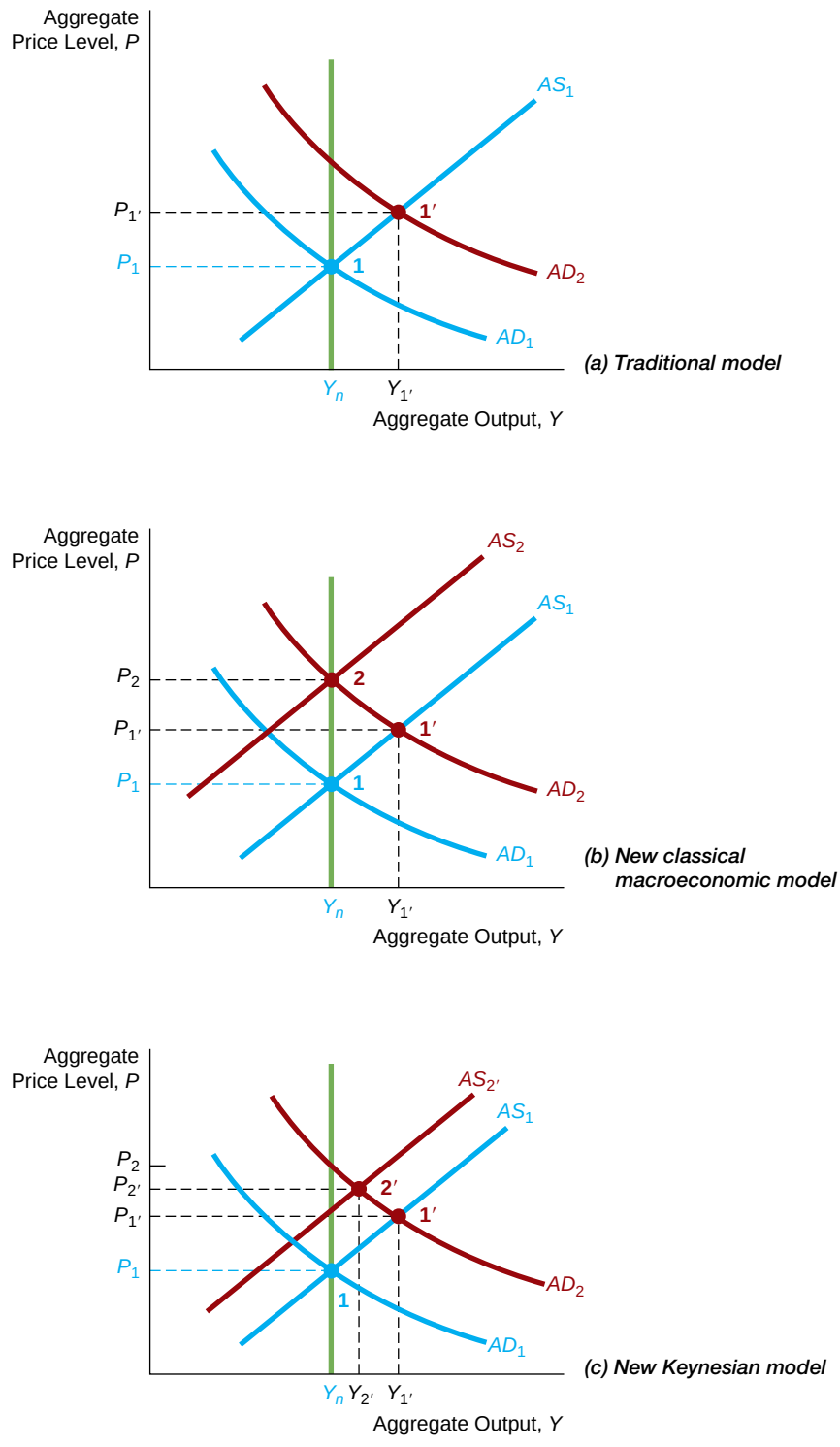
Model	Response to Unanticipated Expansionary Policy	Response to Anticipated Expansionary Policy	Can Activist Policy Be Beneficial?	Response to Unanticipated Anti-inflation Policy	Response to Anticipated Anti-inflation Policy	Is Credibility Important to Successful Anti-inflation Policy?
Traditional model	$Y\uparrow$, $P\uparrow$	$Y\uparrow$, $P\uparrow$ by same amount as when policy is unanticipated	Yes	$Y\downarrow$, $\pi\downarrow$	$Y\downarrow$, $\pi\downarrow$ by same amount as when policy is unanticipated	No
New classical macroeconomic model	$Y\uparrow$, $P\uparrow$	Y unchanged, $P\uparrow$ by more than when policy is unanticipated	No	$Y\downarrow$, $\pi\downarrow$	Y unchanged, $\pi\downarrow$ by more than when policy is unanticipated	Yes
New Keynesian model	$Y\uparrow$, $P\uparrow$	$Y\uparrow$ by less than when policy is unanticipated, $P\uparrow$ by more than when policy is unanticipated	Yes, but designing a beneficial policy is difficult	$Y\downarrow$, $\pi\downarrow$	$Y\downarrow$ by less than when policy is unanticipated, $\pi\downarrow$ by more than when policy is unanticipated	Yes

Note: π represents the inflation rate.

FIGURE 5

Comparison of the Short-Run Response to Expansionary Policy in the Three Models

Initially, the economy is at point 1. The expansionary policy shifts the aggregate demand curve from AD_1 to AD_2. In the traditional model, the expansionary policy moves the economy to point 1′ whether the policy is anticipated or not. In the new classical model, the expansionary policy moves the economy to point 1′ if it is unanticipated and to point 2 if it is anticipated. In the new Keynesian model, the expansionary policy moves the economy to point 1′ if it is unanticipated and to point 2′ if it is anticipated.



expansionary policy is *unanticipated*, all three models show the same short-run output response. The traditional model views the aggregate supply curve as given in the short run, while the other two view it as remaining at AS_1 because there is no change in the expected price level when the policy is a surprise. Hence when policy is *unanticipated*, all three models indicate a movement to point 1′, where the AD_2 and AS_1 curves intersect and where aggregate output and the price level have risen to $Y_{1'}$ and $P_{1'}$ respectively.

The response to the *anticipated* expansionary policy is, however, quite different in the three models. In the traditional model in panel (a), the aggregate supply curve remains at AS_1 even when the expansionary policy is anticipated, because adaptive expectations imply that anticipated policy has no effect on expectations and hence on aggregate supply. It indicates that the economy moves to point 1′, which is where it moved when the policy was unanticipated. The traditional model does not distinguish between the effects of anticipated and unanticipated policy: Both have the same effect on output and prices.

In the new classical model in panel (b), the aggregate supply curve shifts leftward to AS_2 when policy is anticipated, because when expectations of the higher price level are realized, aggregate output will be at the natural rate level. Thus it indicates that the economy moves to point 2; aggregate output does not rise, but prices do, to P_2. This outcome is quite different from the move to point 1′ when policy is unanticipated. The new classical model distinguishes between the short-run effects of anticipated and unanticipated policies: Anticipated policy has no effect on output, but unanticipated policy does. However, anticipated policy has a bigger impact than unanticipated policy on price level movements.

The new Keynesian model in panel (c) is an intermediate position between the traditional and new classical models. It recognizes that anticipated policy affects the aggregate supply curve, but due to rigidities such as long-term contracts, wage and price adjustment is not as complete as in the new classical model. Hence the aggregate supply curve shifts only to $AS_{2'}$ in response to anticipated policy, and the economy moves to point 2′, where output at $Y_{2'}$ is lower than the $Y_{1'}$ level reached when the expansionary policy is unanticipated. But the price level at $P_{2'}$ is higher than the level $P_{1'}$ that resulted from the unanticipated policy. Like the new classical model, the new Keynesian model distinguishes between the effects of anticipated and unanticipated policies: Anticipated policy has a smaller effect on output than unanticipated policy but a larger effect on the price level. However, in contrast to the new classical model, anticipated policy does affect output fluctuations.

Stabilization Policy

The three models have different views of the effectiveness of *stabilization policy*, policy intended to reduce output fluctuations. Because the effects of anticipated and unanticipated policy are identical in the traditional model, policymakers do not have to concern themselves with the public's expectations. This makes it easier for them to predict the outcome of their policy, an essential matter if their actions are to have the intended effect. In the traditional model, it is possible for an activist policy to stabilize output fluctuations.

The new classical model takes the extreme position that activist stabilization policy serves to aggravate output fluctuations. In this model, only unanticipated policy affects output; anticipated policy does not matter. Policymakers can affect output only by surprising the public. Because the public is assumed to have rational expectations, it will always try to guess what policymakers plan to do.

In the new classical model, the conduct of policy can be viewed as a game in which the public and the policymakers are always trying to outfox each other by guessing the other's intentions and expectations. The sole possible outcome of this process is that an activist stabilization policy will have no predictable effect on output and cannot be relied on to stabilize economic activity. Instead, it may create a lot of uncertainty about policy that will increase random output fluctuations around the natural rate level of output. Such an undesirable effect is exactly the opposite of what the activist stabilization policy is trying to achieve. The outcome in the new classical view is that policy should follow a nonactivist rule in order to promote as much certainty about policy actions as possible.

The new Keynesian model again takes an intermediate position between the traditional and the new classical models. Contrary to the new classical model, it indicates that anticipated policy *does* matter to output fluctuations. Policymakers can count on some output response from their anticipated policies and can use them to stabilize the economy.

In contrast to the traditional model, however, the new Keynesian model recognizes that the effects of anticipated and unanticipated policy will not be the same. Policymakers will encounter more uncertainty about the outcome of their actions, because they cannot be sure to what extent the policy is anticipated or not. Hence an activist policy is less likely to operate always in the intended direction and is less likely to achieve its goals. The new Keynesian model raises the possibility that an activist policy could be beneficial, but uncertainty about the outcome of policies in this model may make the design of such a beneficial policy extremely difficult.

Anti-inflation Policies

So far we have focused on the implications of these three models for policies whose intent is to eliminate fluctuations in output. By the end of the 1970s, the high inflation rate (then over 10%) helped shift the primary concern of policymakers to the reduction of inflation. What do these models have to say about anti-inflation policies designed to eliminate upward movements in the price level? The aggregate demand and supply diagrams in Figure 6 will help us answer the question.

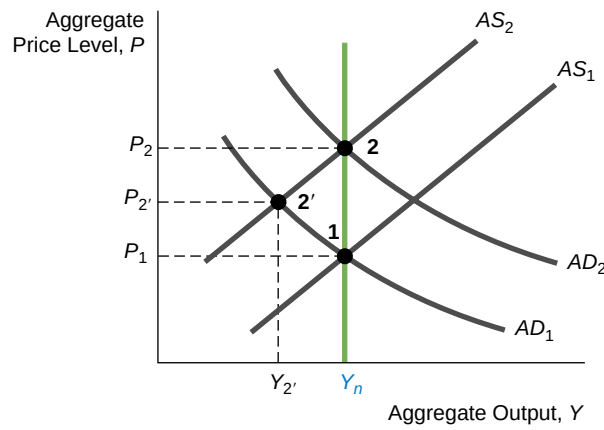
Suppose that the economy has settled into a sustained 10% inflation rate caused by a high rate of money growth that shifts the aggregate demand curve so that it moves up by 10% every year. If this inflation rate has been built into wage and price contracts, the aggregate supply curve shifts so as to rise at the same rate. We see this in Figure 6 as a shift in the aggregate demand curve from AD_1 in year 1 to AD_2 in year 2, while the aggregate supply curve moves from AS_1 to AS_2. In year 1, the economy is at point 1 (intersection of AD_1 and AS_1); in the second year, the economy moves to point 2 (intersection of AD_2 and AS_2), and the price level has risen 10%, from P_1 to P_2. (Note that the figure is not drawn to scale.)

Now suppose that a new Federal Reserve chairman is appointed who decides that inflation must be stopped. He convinces the FOMC to stop the high rate of money growth so that the aggregate demand curve will not rise from AD_1. The policy of halting money growth immediately could be costly if it led to a fall in output. Let's use our three models to explore the degree to which aggregate output will fall as a result of an anti-inflation policy.

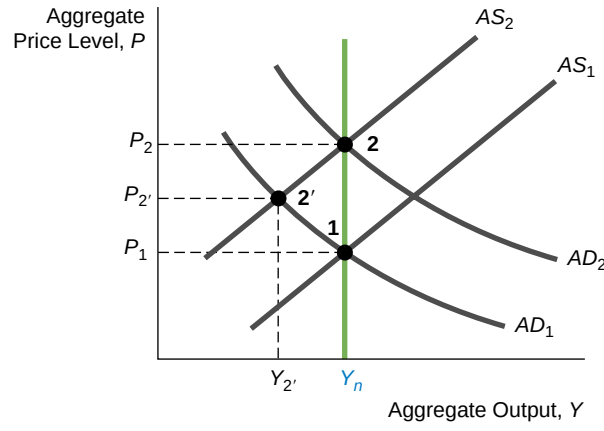
First, look at the outcome of this policy in the traditional model's view of the world in panel (a). The movement of the aggregate supply curve to AS_2 is already set in place and is unaffected by the new policy of keeping the aggregate demand curve at AD_1 (whether the effort is anticipated or not). The economy moves to point 2′ (the

FIGURE 6
Anti-inflation Policy in the Three Models

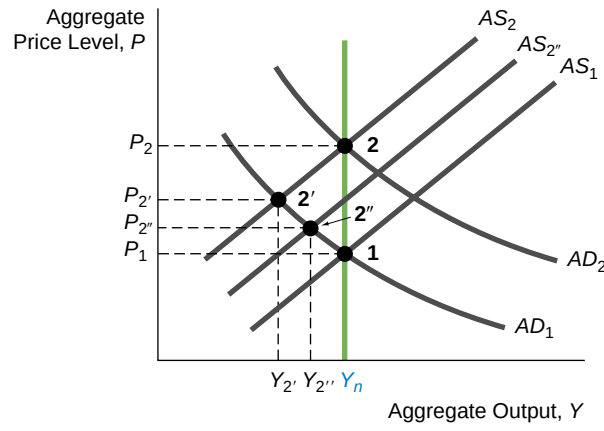
With an ongoing inflation in which the economy is moving from point 1 to point 2, the aggregate demand curve is shifting from AD_1 to AD_2 and the short-run aggregate supply curve from AS_1 to AS_2. The anti-inflation policy, when implemented, prevents the aggregate demand curve from rising, holding it at AD_1. (a) In the traditional model, the economy moves to point 2′ whether the anti-inflation policy is anticipated or not. (b) In the new classical model, the economy moves to point 2′ if the policy is unanticipated and to point 1 if it is anticipated. (c) In the new Keynesian model, the economy moves to point 2′ if the policy is unanticipated and to point 2″ if it is anticipated.



(a) Traditional model



(b) New classical macroeconomic model



(c) New Keynesian model

intersection of the AD_1 and AS_2 curves), and the inflation rate slows down because the price level increases only to $P_{2'}$ rather than P_2. The reduction in inflation has not been without cost: Output has declined to $Y_{2'}$, which is well below the natural rate level.

In the traditional model, estimates of the cost in terms of lost output for each 1% reduction in the inflation rate are around 4% of a year's real GDP. The high cost of reducing inflation in the traditional model is one reason why some economists are reluctant to advocate an anti-inflation policy of the sort tried here. They question whether the cost of high unemployment is worth the benefits of a reduced inflation rate.

If you adhere to the new classical philosophy, you would not be as pessimistic about the high cost of reducing the inflation rate. If the public *expects* the monetary authorities to stop the inflationary process by ending the high rate of money growth, it will occur without any output loss. In panel (b), the aggregate demand curve will remain at AD_1, but because this is expected, wages and prices can be adjusted so that they will not rise, and the aggregate supply curve will remain at AS_1 instead of moving to AS_2. The economy will stay put at point 1 (the intersection of AD_1 and AS_1), and aggregate output will remain at the natural rate level while inflation is stopped because the price level is unchanged.

An important element in the story is that the anti-inflation policy be anticipated by the public. If the policy is *not* expected, the aggregate demand curve remains at AD_1, but the aggregate supply curve continues its shift to AS_2. The outcome of the unanticipated anti-inflation policy is a movement of the economy to point 2′. Although the inflation rate slows in this case, it is not entirely eliminated as it was when the anti-inflation policy was anticipated. Even worse, aggregate output falls below the natural rate level to $Y_{2'}$. An anti-inflation policy that is unanticipated, then, is far less desirable than one that is.

The new Keynesian model in panel (c) also leads to the conclusion that an unanticipated anti-inflation policy is less desirable than an anticipated one. If the policy of keeping the aggregate demand curve at AD_1 is *not* expected, the aggregate supply curve will continue its shift to AS_2, and the economy moves to point 2′ at the intersection of AD_1 and AS_2. The inflation rate slows, but output declines to $Y_{2'}$, well below the natural rate level.

If, by contrast, the anti-inflation policy is *expected*, the aggregate supply curve will not move all the way to AS_2. Instead it will shift only to $AS_{2''}$ because some wages and prices (but not all) can be adjusted, so wages and the price level will not rise at their previous rates. Instead of moving to point 2′ (as occurred when the anti-inflation policy was not expected), the economy moves to point 2″, the intersection of the AD_1 and $AS_{2''}$ curves. The outcome is more desirable than when the policy is unanticipated—the inflation rate is lower (the price level rises only to $P_{2''}$ and not $P_{2'}$), and the output loss is smaller as well ($Y_{2''}$ is higher than $Y_{2'}$).

Credibility in Fighting Inflation

Both the new classical and new Keynesian models indicate that for an anti-inflation policy to be successful in reducing inflation at the lowest output cost, the public must believe (expect) that it will be implemented. In the new classical view of the world, the best anti-inflation policy (when it is credible) is to go "cold turkey." The rise in the aggregate demand curve from AD_1 should be stopped immediately. Inflation would be eliminated at once with no loss of output *if the policy is credible*. In a new Keynesian world, the cold-turkey policy, *even if credible*, is not as desirable, because it will produce some output loss.

John Taylor, a proponent of the new Keynesian model, has demonstrated that a more gradual approach to reducing inflation may be able to eliminate inflation without producing a substantial output loss.[7] An important catch here is that this gradual policy must somehow be made credible, which may be harder to achieve than a cold-turkey anti-inflation policy, which demonstrates immediately that the policymakers are serious about fighting inflation. Taylor's contention that inflation can be reduced with little output loss may be overly optimistic.

Incorporating rational expectations into aggregate supply and demand analysis indicates that a successful anti-inflation policy must be credible. Evidence that credibility plays an important role in successful anti-inflation policies is provided by the dramatic end of the Bolivian hyperinflation in 1985 (see Box 2). But establishing credibility is easier said than done. You might think that an announcement by policymakers at the Federal Reserve that they plan to pursue an anti-inflation policy might do the trick. The public would expect this policy and would act accordingly. However, that conclusion implies that the public will believe the policymakers' announcement. Unfortunately, that is not how the real world works.

Our historical review of Federal Reserve policymaking in Chapter 18 suggests that the Fed has not always done what it set out to do. In fact, during the 1970s, the chairman of the Federal Reserve Board, Arthur Burns, repeatedly announced that the Fed would pursue a vigorous anti-inflation policy. The actual policy pursued, how-

Box 2: Global

Ending the Bolivian Hyperinflation

Case Study of a Successful Anti-inflation Program. The most remarkable anti-inflation program in recent times was implemented in Bolivia. In the first half of 1985, Bolivia's inflation rate was running at 20,000% and rising. Indeed, the inflation rate was so high that the price of a movie ticket often rose while people waited in line to buy it. In August 1985, Bolivia's new president announced his anti-inflation program, the New Economic Policy. To rein in money growth and establish credibility, the new government took drastic actions to slash the budget deficit by shutting down many state-owned enterprises, eliminating subsidies, freezing public sector salaries, and collecting a new wealth tax. The finance ministry was put on a new footing; the budget was balanced on a day-by-day basis. Without exceptions, the finance minister would not authorize spending in excess of the amount of tax revenue that had been collected the day before.

The rule of thumb that a reduction of 1% in the inflation rate requires a 4% loss of a year's aggregate output indicates that ending the Bolivian hyperinflation would have required halving Bolivian aggregate output for 1,600 years! Instead, the Bolivian inflation was stopped in its tracks within one month, and the output loss was minor (less than 5% of GDP).

Certain hyperinflations before World War II were also ended with small losses of output using policies similar to Bolivia's,* and a more recent anti-inflation program in Israel that also involved substantial reductions in budget deficits sharply reduced inflation without any clear loss of output. There is no doubt that credible anti-inflation policies can be highly successful in eliminating inflation.

*For an excellent discussion of the end of four hyperinflations in the 1920s, see Thomas Sargent, "The Ends of Four Big Inflations," in *Inflation: Causes and Consequences,* ed. Robert E. Hall (Chicago: University of Chicago Press, 1982), pp. 41–98.

[7] John Taylor, "The Role of Expectations in the Choice of Monetary Policy," in *Monetary Policy Issues in the 1980s* (Kansas City: Federal Reserve Bank, 1982), pp. 47–76.

ever, had quite a different outcome: The rate of growth of the money supply increased rapidly during the period, and inflation soared. Such episodes have reduced the credibility of the Federal Reserve in the eyes of the public and, as predicted by the new classical and new Keynesian models, have had serious consequences. The reduction of inflation that occurred from 1981 to 1984 was bought at a very high cost; the 1981–1982 recession that helped bring the inflation rate down was the most severe recession in the post–World War II period. Unless some method of restoring credibility to anti-inflation policy is achieved, eliminating inflation will be a costly affair because such policy will be unanticipated.

The U.S. government can play an important role in establishing the credibility of anti-inflation policy. We have seen that large budget deficits may help stimulate inflationary monetary policy, and when the government and the Fed announce that they will pursue a restrictive anti-inflation policy, it is less likely that they will be believed *unless* the federal government demonstrates fiscal responsibility. Another way to say this is to use the old adage, "Actions speak louder than words." When the government takes actions that will help the Fed adhere to anti-inflation policy, the policy will be more credible. Unfortunately, this lesson has sometimes been ignored by politicians in the United States and in other countries.

Application

Credibility and the Reagan Budget Deficits

The Reagan administration was strongly criticized for creating huge budget deficits by cutting taxes in the early 1980s. In the Keynesian framework, we usually think of tax cuts as stimulating aggregate demand and increasing aggregate output. Could the expectation of large budget deficits have helped create a more severe recession in 1981–1982 after the Federal Reserve implemented an anti-inflation monetary policy?

Some economists answer yes, using diagrams like panels (b) and (c) of Figure 6. They claim that the prospect of large budget deficits made it harder for the public to believe that an anti-inflationary policy would actually be pursued when the Fed announced its intention to do so. Consequently, the aggregate supply curve would continue to rise from AS_1 to AS_2 as in panels (b) and (c). When the Fed actually kept the aggregate demand curve from rising to AD_2 by slowing the rate of money growth in 1980–1981 and allowing interest rates to rise, the economy moved to a point like 2′ in panels (b) and (c), and much unemployment resulted. As our analysis in panels (b) and (c) of Figure 6 predicts, the inflation rate did slow substantially, falling below 5% by the end of 1982, but this was very costly: Unemployment reached a peak of 10.7%.

If the Reagan administration had actively tried to reduce deficits instead of raising them by cutting taxes, what might have been the outcome of the anti-inflation policy? Instead of moving to point 2′, the economy might have moved to point 2″ in panel (c)—or even to point 1 in panel (b), if the new classical macroeconomists are right. We would have had an even more rapid reduction in inflation and a smaller loss of output. No wonder some economists were so hostile to Reagan's budget policies!

Reagan is not the only head of state who ran large budget deficits while espousing an anti-inflation policy. Britain's Margaret Thatcher preceded Reagan in this activity, and economists such as Thomas Sargent assert that the reward for her policy was a climb of unemployment in Britain to unprecedented levels.[8]

Although many economists agree that the Fed's anti-inflation program lacked credibility, especially in its initial phases, not all of them agree that the Reagan budget deficits were the cause of that lack of credibility. The conclusion that the Reagan budget deficits helped create a more severe recession in 1981–1982 is controversial.

Impact of the Rational Expectations Revolution

The theory of rational expectations has caused a revolution in the way most economists now think about the conduct of monetary and fiscal policies and their effects on economic activity. One result of this revolution is that economists are now far more aware of the importance of expectations to economic decision making and to the outcome of particular policy actions. Although the rationality of expectations in all markets is still controversial, most economists now accept the following principle suggested by rational expectations: Expectation formation will change when the behavior of forecasted variables changes. As a result, the Lucas critique of policy evaluation using conventional econometric models is now taken seriously by most economists. The Lucas critique also demonstrates that the effect of a particular policy depends critically on the public's expectations about that policy. This observation has made economists much less certain that policies will have their intended effect. An important result of the rational expectations revolution is that economists are no longer as confident in the success of activist stabilization policies as they once were.

Has the rational expectations revolution convinced economists that there is no role for activist stabilization policy? Those who adhere to the new classical macroeconomics think so. Because anticipated policy does not affect aggregate output, activist policy can lead only to unpredictable output fluctuations. Pursuing a nonactivist policy in which there is no uncertainty about policy actions is then the best we can do. Such a position is not accepted by many economists, because the empirical evidence on the policy ineffectiveness proposition is mixed. Some studies find that only unanticipated policy matters to output fluctuations, while other studies find a significant impact of anticipated policy on output movements.[9] In addition, some

[8]Thomas Sargent, "Stopping Moderate Inflations: The Methods of Poincaré and Thatcher," in *Inflation, Debt, and Indexation,* ed. Rudiger Dornbusch and M. H. Simonsen (Cambridge, Mass.: MIT Press, 1983), pp. 54–96, discusses the problems that Thatcher's policies caused and contrasts them with more successful anti-inflation policies pursued by the Poincaré government in France during the 1920s.

[9]Studies with findings that only unanticipated policy matters include Thomas Sargent, "A Classical Macroeconometric Model for the United States," *Journal of Political Economy* 84 (1976): 207–237; Robert J. Barro, "Unanticipated Money Growth and Unemployment in the United States," *American Economic Review* 67 (1977): 101–115; and Robert J. Barro and Mark Rush, "Unanticipated Money and Economic Activity," in *Rational Expectations and Economic Policy,* ed. Stanley Fischer (Chicago: University of Chicago Press, 1980), pp. 23–48. Studies that find a significant impact of anticipated policy are Frederic S. Mishkin, "Does Anticipated Monetary Policy Matter? An Econometric Investigation," *Journal of Political Economy* 90 (1982): 22–51, and Robert J. Gordon, "Price Inertia and Policy Effectiveness in the United States, 1890–1980," *Journal of Political Economy* 90 (1982): 1087–1117.

economists question whether the degree of wage and price flexibility required in the new classical model actually exists.

The result is that many economists take an intermediate position that recognizes the distinction between the effects of anticipated versus unanticipated policy but believe that anticipated policy can affect output. They are still open to the possibility that activist stabilization policy can be beneficial, but they recognize the difficulties of designing it.

The rational expectations revolution has also highlighted the importance of credibility to the success of anti-inflation policies. Economists now recognize that if an anti-inflation policy is not believed by the public, it may be less effective in reducing the inflation rate when it is actually implemented and may lead to a larger loss of output than is necessary. Achieving credibility (not an easy task in that policymakers often say one thing but do another) should then be an important goal for policymakers. To achieve credibility, policymakers must be consistent in their course of action.

The rational expectations revolution has caused major rethinking about the way economic policy should be conducted and has forced economists to recognize that we may have to accept a more limited role for what policy can do for us. Rather than attempting to fine-tune the economy so that all output fluctuations are eliminated, we may have to settle for policies that create less uncertainty and thereby promote a more stable economic environment.

Summary

1. The simple principle (derived from rational expectations theory) that expectation formation changes when the behavior of forecasted variables changes led to the famous Lucas critique of econometric policy evaluation. Lucas argued that when policy changes, expectation formation changes; hence the relationships in an econometric model will change. An econometric model that has been estimated on the basis of past data will no longer be the correct model for evaluating the effects of this policy change and may prove to be highly misleading. The Lucas critique also points out that the effects of a particular policy depend critically on the public's expectations about the policy.
2. The new classical macroeconomic model assumes that expectations are rational and that wages and prices are completely flexible with respect to the expected price level. It leads to the policy ineffectiveness proposition that anticipated policy has no effect on output; only unanticipated policy matters.
3. The new Keynesian model also assumes that expectations are rational but views wages and prices as sticky. Like the new classical model, the new Keynesian model distinguishes between the effects from anticipated and unanticipated policy: Anticipated policy has a smaller effect on aggregate output than unanticipated policy. However, anticipated policy does matter to output fluctuations.
4. The new classical model indicates that activist policy can only be counterproductive, while the new Keynesian model suggests that activist policy might be beneficial. However, since both indicate that there is uncertainty about the outcome of a particular policy, the design of a beneficial activist policy may be very difficult. A traditional model in which expectations about policy have no effect on the aggregate supply curve does not distinguish between the effects of anticipated or unanticipated policy. This model favors activist policy, because the outcome of a particular policy is less uncertain.
5. If expectations about policy affect the aggregate supply curve, as they do in the new classical and new Keynesian models, an anti-inflation policy will be more successful (will produce a faster reduction in inflation with smaller output loss) if it is credible.
6. The rational expectations revolution has forced economists to be less optimistic about the effective use of activist stabilization policy and has made them more aware of the importance of credibility to successful policymaking.

Key Terms

econometric models, p. 659

policy ineffectiveness proposition, p. 663

Questions and Problems

Questions marked with an asterisk are answered at the end of the book in an appendix, "Answers to Selected Questions and Problems."

1. If the public expects the Fed to pursue a policy that is likely to raise short-term interest rates permanently to 12% but the Fed does not go through with this policy change, what will happen to long-term interest rates? Explain your answer.

*2. If consumer expenditure is related to consumers' expectations of their average income in the future, will an income tax cut have a larger effect on consumer expenditure if the public expects the tax cut to last for one year or for ten years?

Use an aggregate supply and demand diagram to illustrate your answer in all the following questions.

3. Having studied the new classical model, the new chairman of the Federal Reserve Board has thought up a surefire plan for reducing inflation and lowering unemployment. He announces that the Fed will lower the rate of money growth from 10% to 5% and then persuades the FOMC to keep the rate of money growth at 10%. If the new classical view of the world is correct, can his plan achieve the goals of lowering inflation and unemployment? How? Do you think his plan will work? If the traditional model's view of the world is correct, will the Fed chairman's surefire plan work?

*4. "The costs of fighting inflation in the new classical and new Keynesian models are lower than in the traditional model." Is this statement true, false, or uncertain? Explain your answer.

5. The new classical model is sometimes characterized as an offshoot of the monetarist model because the two models have similar views of aggregate supply. What are the differences and similarities between the monetarist and new classical views of aggregate supply?

*6. "The new classical model does not eliminate policymakers' ability to reduce unemployment because they can always pursue policies that are more expansionary than the public expects." Is this statement true, false, or uncertain? Explain your answer.

7. What principle of rational expectations theory is used to prove the proposition that stabilization policy can have no predictable effect on aggregate output in the new classical model?

*8. "The Lucas critique by itself casts doubt on the ability of activist stabilization policy to be beneficial." Is this statement true, false, or uncertain? Explain your answer.

9. "The more credible the policymakers who pursue an anti-inflation policy, the more successful that policy will be." Is this statement true, false, or uncertain? Explain your answer.

*10. Many economists are worried that a high level of budget deficits may lead to inflationary monetary policies in the future. Could these budget deficits have an effect on the current rate of inflation?

Using Economic Analysis to Predict the Future

11. Suppose that a treaty is signed limiting armies throughout the world. The result of the treaty is that the public expects military and hence government spending to be reduced. If the new classical view of the economy is correct and government spending does affect the aggregate demand curve, predict what will happen to aggregate output and the price level when government spending is reduced in line with the public's expectations.

12. How would your prediction differ in Problem 11 if the new Keynesian model provides a more realistic description of the economy? What if the traditional model provides the most realistic description of the economy?

*13. The chairman of the Federal Reserve Board announces that over the next year, the rate of money growth will be reduced from its current rate of 10% to a rate of 2%. If the chairman is believed by the public but the Fed actually reduces the rate of money growth to 5%, predict what will happen to the inflation rate and aggregate output if the new classical view of the economy is correct.

*14. How would your prediction differ in Problem 13 if the new Keynesian model provides a more accurate description of the economy? What if the traditional model provides the most realistic description of the economy?

15. If, in a surprise victory, a new administration is elected to office that the public believes will pursue inflationary policy, predict what might happen to the level of output and inflation even before the new administration comes into power. Would your prediction differ depending on which of the three models—traditional, new classical, and new Keynesian—you believed in?

Web Exercises

1. Robert Lucas won the Nobel Prize in Economics. Go to http://www.nobel.se/economics/ and locate the press release on Robert Lucas. What was his Nobel Prize awarded for? When was it awarded?

Applying Theory to the Real World: Applications and Boxes

Applications

Reading the *Wall Street Journal:* The Bond Page, p. 72

Changes in the Equilibrium Interest Rate Due to Expected Inflation or Business Cycle Expansions, p. 99

Explaining Low Japanese Interest Rates, p. 103

Reading the *Wall Street Journal* "Credit Markets" Column, p. 103

Changes in the Equilibrium Interest Rate Due to Changes in Income, the Price Level, or the Money Supply, p. 108

Money and Interest Rates, p. 112

The Enron Bankruptcy and the Baa-Aaa Spread, p. 124

Effects of the Bush Tax Cut on Bond Interest Rates, p. 127

Interpreting Yield Curves, 1980–2003, p. 137

Monetary Policy and Stock Prices, p. 146

The September 11 Terrorist Attacks, the Enron Scandal, and the Stock Market, p. 146

Should Foreign Exchange Rates Follow a Random Walk?, p. 155

Practical Guide to Investing in the Stock Market, p. 158

What Do the Black Monday Crash of 1987 and the Tech Crash of 2000 Tell Us About Rational Expectations and Efficient Markets?, p. 163

Financial Development and Economic Growth, p. 187

Financial Crises in the United States, p. 191

Financial Crises in Emerging-Market Countries: Mexico, 1994–1995; East Asia, 1997–1998; and Argentina, 2001–2002, p. 194

Strategies for Managing Bank Capital, p. 215

Did the Capital Crunch Cause a Credit Crunch in the Early 1990s?, p. 216

Strategies for Managing Interest-Rate Risk, p. 222

Insurance Management, p. 290

Hedging with Interest-Rate Forward Contracts, p. 310

Hedging with Financial Futures, p. 314

Hedging Foreign Exchange Risk, p. 319

Hedging with Futures Options, p. 325

Hedging with Interest-Rate Swaps, p. 329

Explaining Movements in the Money Supply, 1980–2002, p. 384

The Great Depression Bank Panics, 1930–1933, p. 387

Why Have Reserve Requirements Been Declining Worldwide?, p. 406

The Channel/Corridor System for Setting Interest Rates in Other Countries, p. 406

Changes in the Equilibrium Exchange Rate: Two Examples, p. 452

Why Are Exchange Rates So Volatile?, p. 455

The Dollar and Interest Rates, 1973–2002, p. 455

The Euro's First Four Years, p. 457

Reading the *Wall Street Journal:* The "Currency Trading" Column, p. 457

The Foreign Exchange Crisis of September 1992, p. 475

Recent Foreign Exchange Crises in Emerging Market Countries: Mexico 1994, East Asia 1997, Brazil 1999, and Argentina 2002, p. 477

The Collapse of Investment Spending and the Great Depression, p. 545

Targeting Money Supply Versus Interest Rates, p. 571

Explaining Past Business Cycle Episodes, p. 598

Corporate Scandals and the Slow Recovery from the March 2001 Recession, p. 625

Applying the Monetary Policy Lessons to Japan, p. 628

Explaining the Rise in U.S. Inflation, 1960–1980, p. 646

Importance of Credibility to Volcker's Victory over Inflation, p. 655

Credibility and the Reagan Budget Deficits, p. 675

Following the Financial News

Foreign Stock Market Indexes, p. 30

The Monetary Aggregates, p. 54

Bond Prices and Interest Rates, p. 73

The "Credit Markets" Column, p. 104

Forecasting Interest Rates, p. 111

Yield Curves, p. 128

Stock Prices, p. 159

New Securities Issues, p. 304

Financial Futures, p. 312

Futures Options, p. 321

Foreign Exchange Rates, p. 437

The "Currency Trading" Column, p. 458

Aggregate Output, Unemployment, and the Price Level, p. 583

Global Boxes

The Importance of Financial Intermediaries to Securities Markets: An International Comparison, p. 31

Birth of the Euro: Will It Benefit Europe?, p. 49

Negative T-Bill Rates? Japan Shows the Way, p. 69

Barings, Daiwa, Sumitomo, and Allied Irish: Rogue Traders and the Principal–Agent Problem, p. 225

Comparison of Banking Structure in the United States and Abroad, p. 249

Ironic Birth of the Eurodollar Market, p. 255

The Spread of Government Deposit Insurance Throughout the World: Is This a Good Thing?, p. 262

Basel 2: Is It Spinning Out of Control?, p. 266

Foreign Exchange Rate Intervention and the Monetary Base, p. 363

The Growing European Commitment to Price Stability, p. 413

International Policy Coordination: The Plaza Agreement and the Louvre Accord, p. 428

The Euro's Challenge to the Dollar, p. 471

Argentina's Currency Board, p. 494

The European Central Bank's Monetary Policy Strategy, p. 498

Ending the Bolivian Hyperinflation: Case Study of a Successful Anti-inflation Program, p. 674

E-Finance Boxes

Why Are Scandinavians So Far Ahead of Americans in Using Electronic Payments?, p. 50

Are We Headed for a Cashless Society?, p. 52

Venture Capitalists and the High-Tech Sector, p. 183

Will "Clicks" Dominate "Bricks" in the Banking Industry?, p. 236

Information Technology and Bank Consolidation, p. 247

Electronic Banking: New Challenges for Bank Regulation, p. 270

Mutual Funds and the Internet, p. 298

The Internet Comes to Wall Street, p. 306

Inside the Fed Boxes

The Political Genius of the Founders of the Federal Reserve System, p. 336

The Special Role of the Federal Reserve Bank of New York, p. 339

The Role of the Research Staff, p. 342

Green, Blue, and Beige: What Do These Colors Mean at the Fed?, p. 344

The Role of Member Banks in the Federal Reserve System, p. 346

Federal Reserve Transparency, p. 352

Discounting to Prevent a Financial Panic: The Black Monday Stock Market Crash of 1987 and the Terrorist Destruction of the World Trade Center in September 2001, p. 404

Bank Panics of 1930–1933: Why Did the Fed Let Them Happen?, p. 421

A Day at the Federal Reserve Bank of New York's Foreign Exchange Desk, p. 463

Special-Interest Boxes

Helping Investors to Select Desired Interest-Rate Risk, p. 79

With TIPS, Real Interest Rates Have Become Observable in the United States, p. 82

Should You Hire an Ape as Your Investment Adviser?, p. 160

The Enron Implosion and the Arthur Andersen Conviction, p. 178

Case Study of a Financial Crisis, p. 194

Should Social Security Be Privatized?, p. 296

The Long-Term Capital Management Debacle, p. 300

Are Fannie Mae and Freddie Mac Getting Too Big for Their Britches?, p. 302

The Return of the Financial Supermarket?, p. 305

Fed Watching, p. 430

Meaning of the Word *Investment,* p. 540

Perils of Reverse Causation: A Russian Folk Tale, p. 606

Perils of Ignoring an Outside Driving Factor: How to Lose a Presidential Election, p. 607

Real Business Cycle Theory and the Debate on Money and Economic Activity, p. 616

Consumers' Balance Sheets and the Great Depression, p. 624

Perils of Accommodating Policy: The Terrorism Dilemma, p. 654

Proof of the Policy Ineffectiveness Proposition, p. 663

Symbol	Page where Introduced	Term
Δ	369	change in a variable
π^e	80	expected inflation
a	538	autonomous consumer expenditure
AD	578	aggregate demand curve
AS	588	aggregate supply curve
B^d	89	demand for bonds
B^s	90	supply of bonds
c	375	currency ratio
C	66	yearly coupon payment
C	375	currency
C	536	consumer expenditure
D	122	demand curve
D	375	checkable deposits
DL	382	discount loans
e	375	excess reserves ratio
E	442	exchange (spot) rate
$(E^e_{t+1} - E_t)E_t$	444	expected appreciation of domestic currency
EM	214	equity multiplier
ER	370	excess reserves
G	537	government spending
i	62	interest rate (yield to maturity)
i_d	395	discount rate
i^D	443	interest rate on domestic assets
i^F	443	interest rate on foreign assets
i_r	80	real interest rate
I	536	planned investment spending
IS	551	*IS* curve
LM	551	*LM* curve
m	375	money multiplier

Symbol	Page where Introduced	Term
M	375	money supply
M^d	105	demand for money
M^s	105	supply of money
M1	52	M1 monetary aggregate
M2	52	M2 monetary aggregate
M3	53	M3 monetary aggregate
MB	359	monetary base (high-powered money)
MB_n	382	nonborrowed monetary base
mpc	538	marginal propensity to consume
NX	537	net exports
P	577	price level
P^e	617	expected price level
P_s	620	stock prices
P_t	76	price of a security at time t
r	238	reserve requirement ratio for checkable deposits
R	359	reserves
R	76	return
R^e	88	expected return
R^D	444	expected return on domestic deposits
R^F	444	expected return on foreign deposits
ROA	214	return on assets
ROE	214	return on equity
RR	370	required reserves
S	90	supply curve
T	546	taxes
V	518	velocity of money
Y	537	aggregate output (national income)
Y^{ad}	537	aggregate demand
Y_n	575	natural rate level of output

GLOSSARY

accommodating policy An activist policy in pursuit of a high employment target. **640**

activist An economist who views the self-correcting mechanism through wage and price adjustment to be very slow and hence sees the need for the government to pursue active, discretionary policy to eliminate high unemployment whenever it develops. **592**

adaptive expectations Expectations of a variable based on an average of past values of the variable. **147**

adverse selection The problem created by asymmetric information *before* a transaction occurs: The people who are the most undesirable from the other party's point of view are the ones who are most likely to want to engage in the financial transaction. **32**

agency theory The analysis of how asymmetric information problems affect economic behavior. **175, 189**

aggregate demand The total quantity of output demanded in the economy at different price levels. **537, 582**

aggregate demand curve A relationship between the price level and the quantity of aggregate output demanded when the goods and money markets are in equilibrium. **577, 582**

aggregate demand function The relationship between aggregate output and aggregate demand that shows the quantity of aggregate output demanded for each level of aggregate output. **541**

aggregate income The total income of factors of production (land, labor, capital) in the economy. **20**

aggregate output The total production of final goods and services in the economy. **9**

aggregate price level The average price of goods and services in an economy. **10**

aggregate supply The quantity of aggregate output supplied by the economy at different price levels. **582**

aggregate supply curve The relationship between the quantity of output supplied in the short run and the price level. **588**

American option An option that can be exercised at any time up to the expiration date of the contract. **320**

"animal spirits" Waves of optimism and pessimism that affect consumers' and businesses' willingness to spend. **544, 586**

annuities Financial contracts under which a customer pays an annual premium in exchange for a future stream of annual payments beginning at a set age, say 65, and ending when the person dies. **288**

appreciation Increase in a currency's value. **436**

arbitrage Elimination of a riskless profit opportunity in a market. **313**

asset A financial claim or piece of property that is a store of value. **3**

asset management The acquisition of assets that have a low rate of default and diversification of asset holdings to increase profits. **208**

asset transformation The process of turning risky assets into safer assets for investors by creating and selling assets with risk characteristics that people are comfortable with and then using the funds they acquire by selling these assets to purchase other assets that may have far more risk. **32**

asset market approach An approach to determine asset prices using stocks of assets rather than flows. **93**

asymmetric information The unequal knowledge that each party to a transaction has about the other party. **32**

autonomous consumer expenditure The amount of consumer expenditure that is independent of disposable income. **538**

balance of payments A bookkeeping system for recording all payments that have a direct bearing on the movement of funds between a country and foreign countries. **467**

balance-of-payments crisis A foreign exchange crisis stemming from problems in a country's balance of payments. **476**

balance sheet A list of the assets and liabilities of a bank (or firm) that balances: Total assets equal total liabilities plus capital. **201**

bank failure A situation in which a bank cannot satisfy its obligations to pay its depositors and other creditors and so goes out of business. **260**

bank holding companies Companies that own one or more banks. **245**

bank panic The simultaneous failure of many banks, as during a financial crisis. **191**

banks Financial institutions that accept money deposits and make loans (such as commercial banks, savings and loan associations, and credit unions). **8**

bank supervision Overseeing who operates banks and how they are operated. **265**

Basel Accord An agreement that required that banks hold as capital at least 8% of their risk-weighted assets. **265**

Basel Committee on Banking Supervision An international committee of bank supervisors that meets under the auspices of the Bank for International Settlements in Basel, Switzerland. **265**

basis point One one-hundredth of a percentage point. **74**

Board of Governors of the Federal Reserve System A board with seven governors (including the chairman) that plays an essential role in decision making within the Federal Reserve System. **337**

bond A debt security that promises to make payments periodically for a specified period of time. **3**

branches Additional offices of banks that conduct banking operations. **244**

Bretton Woods system The international monetary system in use from 1945 to 1971 in which exchange rates were fixed and the U.S. dollar was freely convertible into gold (by foreign governments and central banks only). **470**

brokerage firms Firms that participate in securities markets as brokers, dealers, and investment bankers. **304**

brokers Agents for investors; they match buyers with sellers. **26**

bubble A situation in which the price of an asset differs from its fundamental market value. **164**

budget deficit The excess of government expenditure over tax revenues. **12**

budget surplus The excess of tax revenues over government expenditures. **12**

business cycles The upward and downward movement of aggregate output produced in the economy. **9**

call option An option contract that provides the right to buy a security at a specified price. **322**

capital account An account that describes the flow of capital between the United States and other countries. **467**

capital adequacy management A bank's decision about the amount of capital it should maintain and then acquisition of the needed capital. **208**

capital market A financial market in which longer-term debt (generally with original maturity of greater than one year) and equity instruments are traded. **27**

capital mobility A situation in which foreigners can easily purchase a country's assets and the country's residents can easily purchase foreign assets. **445**

cash flow The difference between cash receipts and cash expenditures. **141, 190**

central bank The government agency that oversees the banking system and is responsible for the amount of money and credit supplied in the economy; in the United States, the Federal Reserve System. **12, 230**

closed-end fund A mutual fund in which a fixed number of nonredeemable shares are sold at an initial offering, then traded in the over-the-counter market like common stock. **299**

coinsurance A situation in which only a portion of losses are covered by insurance, so that the insured suffers a percentage of the losses along with the insurance agency. **293**

collateral Property that is pledged to the lender to guarantee payment in the event that the borrower is unable to make debt payments. **172**

commodity money Money made up of precious metals or another valuable commodity. **48**

common stock A security that is a claim on the earnings and assets of a company. **5**

compensating balance A required minimum amount of funds that a firm receiving a loan must keep in a checking account at the lending bank. **219**

complete crowding out The situation in which expansionary fiscal policy, such as an increase in government spending, does not lead to a rise in output because there is an exactly offsetting movement in private spending. **571, 587**

consol A perpetual bond with no maturity date and no repayment of principal that periodically makes fixed coupon payments. **67**

constant-money-growth-rate rule A policy rule advocated by monetarists, whereby the Federal Reserve keeps the money supply growing at a constant rate. **654**

consumer durable expenditure Spending by consumers on durable items such as automobiles and household appliances. **617**

consumer expenditure The total demand for (spending on) consumer goods and services. **536, 585**

consumption Spending by consumers on nondurable goods and services (including services related to the ownership of homes and consumer durables). **620**

consumption function The relationship between disposable income and consumer expenditure. **538**

costly state verification Monitoring a firm's activities, an expensive process in both time and money. **182**

cost-push inflation Inflation that occurs because of the push by workers to obtain higher wages. **639**

coupon bond A credit market instrument that pays the owner a fixed interest payment every year until the maturity date, when a specified final amount is repaid. **63**

coupon rate The dollar amount of the yearly coupon payment expressed as a percentage of the face value of a coupon bond. **64**

creditor A holder of debt. **188**

credit rationing A lender's refusing to make loans even though borrowers are willing to pay the stated interest rate or even a higher rate or restricting the size of loans made to less than the full amount sought. **220**

credit risk The risk arising from the possibility that the borrower will default. **208**

credit view Monetary transmission mechanisms operating through asymmetric information effects on credit markets. **618**

currency Paper money (such as dollar bills) and coins. **44**

currency board A monetary regime in which the domestic currency is backed 100% by a foreign currency (say dollars) and in which the note-issuing authority, whether the central bank or the government, establishes a fixed exchange rate to this foreign currency and stands ready to exchange domestic currency at this rate whenever the public requests it. **492**

currency swap The exchange of a set of payments in one currency for a set of payments in another currency. **328**

current account An account that shows international transactions involving currently produced goods and services. **467**

current yield An approximation of the yield to maturity that equals the yearly coupon payment divided by the price of a coupon bond. **70**

dealers People who link buyers with sellers by buying and selling securities at stated prices. **26**

debt deflation A situation in which a substantial decline in the price level sets in, leading to a further deterioration in firms' net worth because of the increased burden of indebtedness. **192**

deductible The fixed amount by which the insured's loss is reduced when a claim is paid off. **292**

default A situation in which the party issuing a debt instrument is unable to make interest payments or pay off the amount owed when the instrument matures. **120**

default-free bonds Bonds with no default risk, such as U.S. government bonds. **121**

default risk The chance that the issuer of a debt instrument will be unable to make interest payments or pay off the face value when the instrument matures. **120**

defensive open market operations Open market operations intended to offset movements in other factors that affect the monetary base (such as changes in Treasury deposits with the Fed or changes in float). **398**

defined-benefit plan A pension plan in which benefits are set in advance. **294**

defined-contribution plan A pension plan in which benefits are determined by the contributions into the plan and their earnings. **294**

demand curve A curve depicting the relationship between quantity demanded and price when all other economic variables are held constant. **87**

demand-pull inflation Inflation that results when policymakers pursue policies that shift the aggregate demand curve. **639**

deposit outflows Losses of deposits when depositors make withdrawals or demand payment. **208**

deposit rate ceiling Restriction on the maximum interest rate payable on deposits. **238**

depreciation Decrease in a currency's value. **436**

devaluation Resetting of the fixed value of a currency at a lower level. **472**

dirty float See *managed float regime*. **462**

discount bond A credit market instrument that is bought at a price below its face value and whose face value is repaid at the maturity date; it does not make any interest payments. Also called a zero-coupon bond. **64**

discount loans A bank's borrowings from the Federal Reserve System; also known as advances. **203**

discount rate The interest rate that the Federal Reserve charges banks on discount loans. **210, 359**

discount window The Federal Reserve facility at which discount loans are made to banks. **400**

discount yield See *yield on a discount basis*. **71**

disintermediation A reduction in the flow of funds into the banking system that causes the amount of financial intermediation to decline. **238**

disposable income Total income available for spending, equal to aggregate income minus taxes. **538**

diversification Investing in a collection (**portfolio**) of assets whose returns do not always move together, with the result that overall risk is lower than for individual assets. **32**

dividends Periodic payments made by equities to shareholders. **26, 142**

dollarization The adoption of a sound currency, like the U.S. dollar, as a country's money. **493**

dual banking system The system in the United States in which banks supervised by the federal government and banks supervised by the states operate side by side. **231**

duration analysis A measurement of the sensitivity of the market value of a bank's assets and liabilities to changes in interest rates. **221**

dynamic open market operations Open market operations that are intended to change the level of reserves and the monetary base. **398**

e-cash Electronic money that is used on the Internet to purchase goods or services. **51**

econometric model A model whose equations are estimated using statistical procedures. **659**

economies of scale The reduction in transaction costs per dollar of transaction as the size (scale) of transactions increases. **30**

economies of scope The ability to use one resource to provide many different products and services. **248**

Edge Act corporation A special subsidiary of a U.S. bank that is engaged primarily in international banking. **255**

effective exchange rate index An index reflecting the value of a basket of representative foreign currencies. **455**

efficient market hypothesis The application of the theory of rational expectations to financial markets. **149**

e-finance A new means of delivering financial services electronically. **8**

electronic money (or **e-money**) Money that exists only in electronic form and substitutes for cash as well. **51**

equation of exchange The equation $MV = PY$, which relates nominal income to the quantity of money. **518, 583**

equities Claims to share in the net income and assets of a corporation (such as common stock). **26**

equity capital See *net worth*. **180**

equity multiplier (EM) The amount of assets per dollar of equity capital. **214**

Eurobonds Bonds denominated in a currency other than that of the country in which they are sold. **28**

Eurocurrencies A variant of the Eurobond, which are foreign currencies deposited in banks outside the home country. **28**

Eurodollars U.S. dollars that are deposited in foreign banks outside the United States or in foreign branches of U.S. banks. **28**

European option An option that can be exercised only at the expiration date of the contract. **320**

excess demand A situation in which quantity demanded is greater than quantity supplied. **90**

excess reserves Reserves in excess of required reserves. **204, 359**

excess supply A situation in which quantity supplied is greater than quantity demanded. **90**

exchange rate The price of one currency in terms of another. **435**

exchange rate overshooting A phenomenon whereby the exchange rate changes by more in the short run than it does in the long run when the money supply changes. **454**

exchanges Secondary markets in which buyers and sellers of securities (or their agents or brokers) meet in one central location to conduct trades. **27**

exercise price The price at which the purchaser of an option has the right to buy or sell the underlying financial instrument. Also known as the strike price. **320**

expectations theory The proposition that the interest rate on a long-term bond will equal the average of the short-term interest rates that people expect to occur over the life of the long-term bond. **129**

expected return The return on an asset expected over the next period. **86**

expenditure multiplier The ratio of a change in aggregate output to a change in investment spending (or autonomous spending). **543**

face value A specified final amount paid to the owner of a coupon bond at the maturity date. Also called par value. **63**

federal funds rate The interest rate on overnight loans of deposits at the Federal Reserve. **393**

Federal Open Market Committee (FOMC) The committee that makes decisions regarding the conduct of open market operations; composed of the seven members of the Board of Governors of the Federal Reserve System, the president of the Federal Reserve Bank of New York, and the presidents of four other Federal Reserve banks on a rotating basis. **337**

Federal Reserve banks The 12 district banks in the Federal Reserve System. **337**

Federal Reserve System (the Fed) The central banking authority responsible for monetary policy in the United States. **12**

fiat money Paper currency decreed by a government as legal tender but not convertible into coins or precious metal. **48**

financial crisis A major disruption in financial markets that is characterized by sharp declines in asset prices and the failures of many financial and nonfinancial firms. **189**

financial derivatives Instruments that have payoffs that are linked to previously issued securities, used as risk reduction tools. **233, 309**

financial engineering The process of researching and developing new financial products and services that would meet customer needs and prove profitable. **232**

financial futures contract A futures contract in which the standardized commodity is a particular type of financial instrument. **312**

financial futures option An option in which the underlying instrument is a *futures contract*. Also called a futures option. **321**

financial intermediaries Institutions (such as banks, insurance companies, mutual funds, pension funds, and finance companies) that borrow funds from people who have saved and then make loans to others. **7**

financial intermediation The process of indirect finance whereby financial intermediaries link lender-savers and borrower-spenders. **29**

financial markets Markets in which funds are transferred from people who have a surplus of available funds to people who have a shortage of available funds. **3**

financial panic The widespread collapse of financial markets and intermediaries in an economy. **39**

fiscal policy Policy that involves decisions about government spending and taxation. **12**

Fisher effect The outcome that when expected inflation occurs, interest rates will rise; named after economist Irving Fisher. **100**

fixed exchange rate regime A regime in which central banks buy and sell their own currencies to keep their exchange rates fixed at a certain level. **470**

fixed investment Spending by firms on equipment (computers, airplanes) and structures (factories, office buildings) and planned spending on residential housing. **539**

fixed-payment loan A credit market instrument that provides a borrower with an amount of money that is repaid by making a fixed payment periodically (usually monthly) for a set number of years. **63**
float Cash items in process of collection at the Fed minus deferred-availability cash items. **365**
foreign bonds Bonds sold in a foreign country and denominated in that country's currency. **28**
foreign exchange intervention An international financial transaction in which a central bank buys or sells currency to influence foreign exchange rates. **462**
foreign exchange market The market in which exchange rates are determined. **5, 435**
foreign exchange rate See *exchange rate*. **5**
forward contract An agreement by two parties to engage in a financial transaction at a future (forward) point in time. **310**
forward exchange rate The exchange rate for a forward transaction. **436**
forward transaction A transaction that involves the exchange of bank deposits denominated in different currencies at some specified future date. **436**
free-rider problem The problem that occurs when people who do not pay for information take advantage of the information that other people have paid for. **176**
fully funded Describing a pension plan in which the contributions to the plan and their earnings over the years are sufficient to pay out the defined benefits when they come due. **294**
futures contract A contract in which the seller agrees to provide a certain standardized commodity to the buyer on a specific future date at an agreed-on price. **233**
futures option See *financial futures option*. **321**

gap analysis A measurement of the sensitivity of bank profits to changes in interest rates, calculated by subtracting the amount of rate-sensitive liabilities from the amount of rate-sensitive assets. **221**
goal independence The ability of the central bank to set the goals of monetary policy. **347**
gold standard A regime under which a currency is directly convertible into gold. **469**
goodwill An accounting entry to reflect value to the firm of its having special expertise or a particularly profitable business line. **275**
government budget constraint The requirement that the government budget deficit equal the sum of the change in the monetary base and the change in government bonds held by the public. **643**
government spending Spending by all levels of government on goods and services. **537, 585**
gross domestic product (GDP) The value of all final goods and services produced in the economy during the course of a year. **12, 20**

hedge To protect oneself against risk. **233, 309**
hedge fund A special type of mutual fund that engages in "market-neutral strategies." **299**
high-powered money The monetary base. **359**
hyperinflation An extreme inflation in which the inflation rate exceeds 50% per month. **47**
hysteresis A departure from full employment levels as a result of past high unemployment. **597**

incentive-compatible Having the incentives of both parties to a contract in alignment. **185**
income The flow of earnings. **45**
indexed bond A bond whose interest and principal payments are adjusted for changes in the price level, and whose interest rate thus provides a direct measure of a real interest rate. **82**
inflation The condition of a continually rising price level. **10**
inflation rate The rate of change of the price level, usually measured as a percentage change per year. **11**
initial public offering (IPO) A stock whose firm is issuing it for the first time. **303**
insolvent A situation in which the value of a firm's or bank's assets has fallen below its liabilities; bankrupt. **192**
instrument independence The ability of the central bank to set monetary policy instruments. **347**
interest parity condition The observation that the domestic interest rate equals the foreign interest rate plus the expected appreciation in the foreign currency. **445**
interest rate The cost of borrowing or the price paid for the rental of funds (usually expressed as a percentage per year). **4**
interest-rate forward contract A forward contract that is linked to a debt instrument. **310**
interest-rate risk The possible reduction in returns associated with changes in interest rates. **78, 208**
interest-rate swap A financial contract that allows one party to exchange (swap) a set of interest payments for another set of interest payments owned by another party. **328**
intermediate target Any of a number of variables, such as monetary aggregates or interest rates, that have a direct effect on employment and the price level and that the Fed seeks to influence. **414**
intermediate-term With reference to a debt instrument, having a maturity of between one and ten years. **26**
international banking facilities (IBFs) Banking establishments in the United States that can accept time deposits from foreigners but are not subject to either reserve requirements or restrictions on interest payments. **225**
International Monetary Fund (IMF) The international organization created by the Bretton Woods agreement whose objective is to promote the growth of world trade by making loans to countries experiencing balance-of-payments difficulties. **470**

international policy coordination Agreements among countries to enact policies cooperatively. **428**
international reserves Central bank holdings of assets denominated in foreign currencies. **462**
inventory investment Spending by firms on additional holdings of raw materials, parts, and finished goods. **539**
inverted yield curve A yield curve that is downward-sloping. **127**
investment banks Firms that assist in the initial sale of securities in the primary market. **26**
***IS* curve** The relationship that describes the combinations of aggregate output and interest rates for which the total quantity of goods produced equals the total quantity demanded (goods market equilibrium). **551**

January effect An abnormal rise in stock prices from December to January. **156**
junk bonds Bonds with ratings below Baa (or BBB) that have a high default risk. **124**

Keynesian A follower of John Maynard Keynes who believes that movements in the price level and aggregate output are driven by changes not only in the money supply but also in government spending and fiscal policy and who does not regard the economy as inherently stable. **582**

large, complex banking organizations (LCBOs) Large companies that provide banking as well as many other financial services. **248**
law of one price The principle that if two countries produce an identical good, the price of this good should be the same throughout the world no matter which country produces it. **439**
lender of last resort Provider of reserves to financial institutions when no one else would provide them in order to prevent a financial crisis. **402**
leverage ratio A bank's capital divided by its assets. **265**
liabilities IOUs or debts. **24**
liability management The acquisition of funds at low cost to increase profits. **208**
liquid Easily converted into cash. **27**
liquidity The relative ease and speed with which an asset can be converted into cash. **47, 86**
liquidity management The decisions made by a bank to maintain sufficient liquid assets to meet the bank's obligations to depositors. **208**
liquidity preference framework A model developed by John Maynard Keynes that predicts the equilibrium interest rate on the basis of the supply of and demand for money. **105**
liquidity preference theory John Maynard Keynes's theory of the demand for money. **521**
liquidity premium theory The theory that the interest rate on a long-term bond will equal an average of short-term interest rates expected to occur over the life of the long-term bond plus a positive term (liquidity) premium. **133**
***LM* curve** The relationship that describes the combinations of interest rates and aggregate output for which the quantity of money demanded equals the quantity of money supplied (money market equilibrium). **551**
load funds Open-end mutual funds sold by salespeople who receive a commission that is paid at the time of purchase and is immediately subtracted from the redemption value of the shares. **299**
loanable funds The quantity of loans. **92**
loanable funds framework Determining the equilibrium interest rate by analyzing the supply of and demand for bonds (loanable funds). **92**
loan commitment A bank's commitment (for a specified future period of time) to provide a firm with loans up to a given amount at an interest rate that is tied to some market interest rate. **219**
loan sale The sale under a contract (also called a secondary loan participation) of all or part of the cash stream from a specific loan, thereby removing the loan from the bank's balance sheet. **223**
long position A contractual obligation to take delivery of an underlying financial instrument. **309**
long-run aggregate supply curve The quantity of output supplied in the long run at any given price level. **590**
long-run monetary neutrality See *monetary neutrality*. **576**
long-term With reference to a debt instrument, having a maturity of ten years or more. **26**

M1 A measure of money that includes currency, traveler's checks, and checkable deposits. **52**
M2 A measure of money that adds to M1: money market deposit accounts, money market mutual fund shares, small-denomination time deposits, savings deposits, overnight repurchase agreements, and overnight Eurodollars. **52**
M3 A measure of money that adds to M2: large-denomination time deposits, long-term repurchase agreements, and institutional money market fund shares. **53**
macro hedge A hedge of interest-rate risk for a financial institution's entire portfolio. **315**
managed float regime The current international financial environment in which exchange rates fluctuate from day to day but central banks attempt to influence their countries' exchange rates by buying and selling currencies. Also known as a dirty float. **462**
marginal propensity to consume The slope of the consumption function line that measures the change in consumer expenditure resulting from an additional dollar of disposable income. **538**

margin requirement A sum of money that must be kept in an account (the margin account) at a brokerage firm. **318**

marked to market Repriced and settled in the margin account at the end of every trading day to reflect any change in the value of the futures contract. **318**

market equilibrium A situation occurring when the quantity that people are willing to buy (demand) equals the quantity that people are willing to sell (supply). **90**

market fundamentals Items that have a direct impact on future income streams of a security. **152**

matched sale–purchase transaction An arrangement whereby the Fed sells securities and the buyer agrees to sell them back to the Fed in the near future; sometimes called a reverse repo. **400**

maturity Time to the expiration date (maturity date) of a debt instrument. **26**

mean reversion The phenomenon that stocks with low returns today tend to have high returns in the future, and vice versa. **157**

medium of exchange Anything that is used to pay for goods and services. **45**

micro hedge A hedge for a specific asset. **315**

modern quantity theory of money The theory that changes in aggregate spending are determined primarily by changes in the money supply. **584**

monetarist A follower of Milton Friedman who sees changes in the money supply as the primary source of movements in the price level and aggregate output and who views the economy as inherently stable. **582**

monetary aggregates The various measures of the money supply used by the Federal Reserve System (M1, M2, and M3). **52**

monetary base The sum of the Fed's monetary liabilities (currency in circulation and reserves) and the U.S. Treasury's monetary liabilities (Treasury currency in circulation, primarily coins). **358**

monetary neutrality A proposition that in the long run, a percentage rise in the money supply is matched by the same percentage rise in the price level, leaving unchanged the real money supply and all other economic variables such as interest rates. **453**

monetary policy The management of the money supply and interest rates. **12**

monetary theory The theory that relates changes in the quantity of money to changes in economic activity. **10, 517**

monetizing the debt A method of financing government spending whereby the government debt issued to finance government spending is removed from the hands of the public and is replaced by high-powered money instead. Also called *printing money*. **644**

money Anything that is generally accepted in payment for goods or services or in the repayment of debts. **8**

money center banks Large banks in key financial centers (New York, Chicago, San Francisco). **212**

money market A financial market in which only short-term debt instruments (generally those with original maturity of less than one year) are traded. **27**

money multiplier A ratio that relates the change in the money supply to a given change in the monetary base. **374**

money supply The quantity of money. **8**

moral hazard The risk that one party to a transaction will engage in behavior that is undesirable from the other party's point of view. **33**

multiple deposit creation The process whereby, when the Fed supplies the banking system with $1 of additional reserves, deposits increase by a multiple of this amount. **366**

NAIRU (nonaccelerating inflation rate of unemployment) The rate of unemployment when demand for labor equals supply, consequently eliminating the tendency for the inflation rate to change. **429, 590**

national banks Federally chartered banks. **231**

natural rate level of output The level of aggregate output produced at the natural rate of unemployment at which there is no tendency for wages or prices to change. **575, 590**

natural rate of unemployment The rate of unemployment consistent with full employment at which the demand for labor equals the supply of labor. **412, 590**

net exports Net foreign spending on domestic goods and services, equal to exports minus imports. **537, 585**

net worth The difference between a firm's assets (what it owns or is owed) and its liabilities (what it owes). Also called *equity capital*. **180**

no-load funds Mutual funds sold directly to the public on which no sales commissions are charged. **299**

nominal anchor A nominal variable such as the inflation rate, an exchange rate, or the money supply that monetary policymakers use to tie down the price level. **487**

nominal interest rate An interest rate that does not take inflation into account. **79**

nonaccelerating inflation rate of unemployment See *NAIRU*. **429, 590**

nonactivist An economist who believes that the performance of the economy would be improved if the government avoided active policy to eliminate unemployment. **592**

nonborrowed monetary base The monetary base minus discount loans. **381**

notional principal The amount on which interest is being paid in a swap arrangement. **328**

off-balance-sheet activities Bank activities that involve trading financial instruments and the generation of income from fees and loan sales, all of which affect bank profits but are not visible on bank balance sheets. **223, 265**
official reserve transactions balance The current account balance plus items in the capital account. **468**
open-end fund A mutual fund in which shares can be redeemed at any time at a price that is tied to the asset value of the fund. **298**
open interest The number of contracts outstanding. **315**
open market operations The Fed's buying or selling of bonds in the open market. **340, 359**
open market purchase A purchase of bonds by the Fed. **359**
open market sale A sale of bonds by the Fed. **359**
operating target Any of a set of variables, such as reserve aggregates or interest rates, that the Fed seeks to influence and that are responsive to its policy tools. **415**
opportunity cost The amount of interest (expected return) sacrificed by not holding an alternative asset. **106**
optimal forecast The best guess of the future using all available information. **148**
option A contract that gives the purchaser the option (right) to buy or sell the underlying financial instrument at a specified price, called the exercise price or strike price, within a specific period of time (the term to expiration). **320**
over-the-counter (OTC) market A secondary market in which dealers at different locations who have an inventory of securities stand ready to buy and sell securities "over the counter" to anyone who comes to them and is willing to accept their prices. **27**

partial crowding out The situation in which an increase in government spending leads to a decline in private spending that does not completely offset the rise in government spending. **587**
par value See *face value*. **63**
payments system The method of conducting transactions in the economy. **48**
pecking order hypothesis The hypothesis that the larger and more established a corporation is, the more likely it will be to issue securities to raise funds. **180**
perpetuity See *consol*. **67**
Phillips curve theory A theory suggesting that changes in inflation are influenced by the state of the economy relative to its production capacity, as well as to other factors. **429**
planned investment spending Total planned spending by businesses on new physical capital (machines, computers, apartment buildings) plus planned spending on new homes. **536, 585**
policy ineffectiveness proposition The conclusion from the new classical model that anticipated policy has no effect on output fluctuations. **663**
political business cycle A business cycle caused by expansionary policies before an election. **353**
preferred habitat theory A theory that is closely related to liquidity premium theory, in which the interest rate on a long-term bond equals an average of short-term interest rates expected to occur over the life of the long-term bond plus a positive term premium. **134**
premium The amount paid for an option contract. **320**
present discounted value See *present value*. **61**
present value Today's value of a payment to be received in the future when the interest rate is *i*. Also called *present discounted value*. **61**
primary dealers Government securities dealers, operating out of private firms or commercial banks, with whom the Fed's open market desk trades. **399**
primary market A financial market in which new issues of a security are sold to initial buyers. **26**
principal–agent problem A moral hazard problem that occurs when the managers in control (the agents) act in their own interest rather than in the interest of the owners (the principals) due to different sets of incentives. **181**
printing money See *monetizing the debt*. **644**
prudential supervision See *bank supervision*. **265**
put option An option contract that provides the right to sell a security at a specified price. **322**

quantity theory of money The theory that nominal income is determined solely by movements in the quantity of money. **519**
quotas Restrictions on the quantity of foreign goods that can be imported. **441**

random walk The movements of a variable whose future changes cannot be predicted (are random) because the variable is just as likely to fall as to rise from today's value. **154**
rate of capital gain The change in a security's price relative to the initial purchase price. **76**
rate of return See *return*. **75**
rational expectations Expectations that reflect optimal forecasts (the best guess of the future) using all available information. **147**
real bills doctrine A guiding principle (now discredited) for the conduct of monetary policy that states that as long as loans are made to support the production of goods and services, providing reserves to the banking system to make these loans will not be inflationary. **420**
real business cycle theory A theory that views real shocks to tastes and technology as the major driving force behind short-run business cycle fluctuations. **597**
real interest rate The interest rate adjusted for expected changes in the price level (inflation) so that it more accurately reflects the true cost of borrowing. **79**

real money balances The quantity of money in real terms. **523**

real terms Terms reflecting actual goods and services one can buy. **80**

recession A period when aggregate output is declining. **9**

reduced-form evidence Evidence that examines whether one variable has an effect on another by simply looking directly at the relationship between the two variables. **604**

regulatory arbitrage A process in which banks keep on their books assets that have the same risk-based capital requirement but are relatively risky, such as a loan to a company with a very low credit rating, while taking off their books low-risk assets, such as a loan to a company with a very high credit rating. **265**

regulatory forbearance Regulators' refraining from exercising their right to put an insolvent bank out of business. **275**

reinsurance An allocation of the portion of the insurance risk to another company in exchange for a portion of the insurance premium. **290**

repurchase agreement (repo) An arrangement whereby the Fed, or another party, purchases securities with the understanding that the seller will repurchase them in a short period of time, usually less than a week. **400**

required reserve ratio The fraction of deposits that the Fed requires be kept as reserves. **204, 359**

required reserves Reserves that are held to meet the Fed's requirement that for every dollar of deposits at a bank, a certain fraction must be kept as reserves. **204, 359**

reserve currency A currency, such as the U.S. dollar, that is used by other countries to denominate the assets they hold as international reserves. **470**

reserve requirements Regulation making it obligatory for depository institutions to keep a certain fraction of their deposits in accounts with the Fed. **204**

reserves Banks' holding of deposits in accounts with the Fed plus currency that is physically held by banks (vault cash). **204, 359**

restrictive covenants Provisions that restrict and specify certain activities that a borrower can engage in. **172**

return The payments to the owner of a security plus the change in the security's value, expressed as a fraction of its purchase price. More precisely called the *rate of return*. **75**

return on assets (ROA) Net profit after taxes per dollar of assets. **214**

return on equity (ROE) Net profit after taxes per dollar of equity capital. **214**

revaluation Resetting of the fixed value of a currency at a higher level. **472**

reverse causation A situation in which one variable is said to cause another variable when in reality the reverse is true. **606**

reverse repo See *matched sale–purchase transaction*. **400**

Ricardian equivalence Named after the nineteenth-century British economist David Ricardo, it contends that when the government runs deficits and issues bonds, the public recognizes that it will be subject to higher taxes in the future in order to pay off these bonds. **645**

risk The degree of uncertainty associated with the return on an asset. **31, 86**

risk premium The spread between the interest rate on bonds with default risk and the interest rate on default-free bonds. **121**

risk sharing The process of creating and selling assets with risk characteristics that people are comfortable with and then using the funds they acquire by selling these assets to purchase other assets that may have far more risk. **31**

risk structure of interest rates The relationship among the various interest rates on bonds with the same term to maturity. **120**

seasoned issue A stock issued for sale for which prior issues currently sell in the market. **303**

secondary market A financial market in which securities that have previously been issued (and are thus second-hand) can be resold. **26**

secondary reserves Short-term U.S. government and agency securities held by banks. **204**

secured debt Debt guaranteed by collateral. **172**

securitization The process of transforming illiquid financial assets into marketable capital market instruments. **237**

security A claim on the borrower's future income that is sold by the borrower to the lender. Also called a financial instrument. **3**

segmented markets theory A theory of term structure that sees markets for different-maturity bonds as completely separated and segmented such that the interest rate for bonds of a given maturity is determined solely by supply of and demand for bonds of that maturity. **132**

seignorage The revenue a govenment receives by issuing money. **493**

self-correcting mechanism A characteristic of the economy that causes output to return eventually to the natural rate level regardless of where it is initially. **591**

short position A contractual obligation to deliver an underlying financial instrument. **309**

short-term With reference to a debt instrument, having a maturity of one year or less. **26**

simple deposit multiplier The multiple increase in deposits generated from an increase in the banking system's reserves in a simple model in which the behavior of depositor and bank plays no role. **369**

simple loan A credit market instrument providing the borrower with an amount of funds that must be repaid to the lender at the maturity date along with an additional payment (interest). **62**

smart card A stored-value card that contains a computer chip that lets it be loaded with digital cash from the owner's bank account whenever needed. **51**

special drawing rights (SDRs) An IMF-issued paper substitute for gold that functions as international reserves. **474**

specialist A dealer-broker operating in an exchange who maintains orderly trading of the securities for which he or she is responsible. **306**

spot exchange rate The exchange rate for a spot transaction. **436**

spot transaction The predominant type of exchange rate transaction, involving the immediate exchange of bank deposits denominated in different currencies. **436**

state banks State-chartered banks. **231**

sterilized foreign exchange intervention A foreign exchange intervention with an offsetting open market operation that leaves the monetary base unchanged. **465**

stock A security that is a claim on the earnings and assets of a company. **5**

stock option An option on an individual stock. **321**

store of value A repository of purchasing power over time. **47**

strike price See *exercise price*. **320**

structural model A description of how the economy operates, using a collection of equations that describe the behavior of firms and consumers in many sectors of the economy. **604**

structural model evidence Evidence that examines whether one variable affects another by using data to build a model illustrating the channels through which this variable affects the other. **603**

superregional banks Bank holding companies similar in size to money center banks, but whose headquarters are not based in one of the money center cities (New York, Chicago, San Francisco). **247**

supply curve A curve depicting the relationship between quantity supplied and price when all other economic variables are held constant. **90**

supply shock Any change in technology or the supply of raw materials that can shift the aggregate supply curve. **594**

swap A financial contract that obligates one party to exchange (swap) a set of payments it owns for a set of payments owned by another party. **328**

sweep account An arrangement in which any balances above a certain amount in a corporation's checking account at the end of a business day are "swept out" of the account and invested in overnight repos that pay the corporation interest. **239**

T-account A simplified balance sheet with lines in the form of a T that lists only the changes that occur in balance sheet items starting from some initial balance sheet position. **205**

tariffs Taxes on imported goods. **441**

Taylor rule Economist John Taylor's monetary policy rule that explains how the federal funds rate target is set. **428**

term structure of interest rates The relationship among interest rates on bonds with different terms to maturity. **120**

theory of asset demand The theory that the quantity demanded of an asset is (1) usually positively related to wealth, (2) positively related to its expected return relative to alternative assets, (3) negatively related to the risk of its return relative to alternative assets, and (4) positively related to its liquidity relative to alternative assets. **87**

theory of purchasing power parity (PPP) The theory that exchange rates between any two currencies will adjust to reflect changes in the price levels of the two countries. **439**

thrift institutions (thrifts) Savings and loan associations, mutual savings banks, and credit unions. **34**

time-consistency problem The problem that occurs when monetary policymakers conduct monetary policy in a discretionary way and pursue expansionary policies that are attractive in the short run but lead to bad long-run outcomes. **488**

trade balance The difference between merchandise exports and imports. **467**

transaction costs The time and money spent trying to exchange financial assets, goods, or services. **29**

transmission mechanisms of monetary policy The channels through which the money supply affects economic activity. **604**

underfunded Describing a pension plan in which the contributions and their earnings are not sufficient to pay out the defined benefits when they come due. **294**

underwriters Investment banks that guarantee prices on securities to corporations and then sell the securities to the public. **303**

underwriting Guaranteeing prices on securities to corporations and then selling the securities to the public. **26**

unemployment rate The percentage of the labor force not working. **9**

unexploited profit opportunity A situation in which an investor can earn a higher than normal return. **152**

unit of account Anything used to measure value in an economy. **46**

unsecured debt Debt not guaranteed by collateral. **172**

unsterilized foreign exchange intervention A foreign exchange intervention in which a central bank allows the purchase or sale of domestic currency to affect the monetary base. **464**

vault cash Currency that is physically held by banks and stored in vaults overnight. **204**

velocity See *velocity of money*. **518, 582**

velocity of money The rate of turnover of money; the average number of times per year that a dollar is spent in buying the total amount of final goods and services produced in the economy. **518, 582**

venture capital firm A financial intermediary that pools the resources of its partners and uses the funds to help entrepreneurs start up new businesses. **182**

virtual bank A bank that has no building but rather exists only in cyberspace. **235**

wealth All resources owned by an individual, including all assets. **45, 86**

World Bank The International Bank for Reconstruction and Redevelopment, an international organization that provides long-term loans to assist developing countries in building dams, roads, and other physical capital that would contribute to their economic development. **470**

yield curve A plot of the interest rates for particular types of bonds with different terms to maturity. **127**

yield on a discount basis The measure of interest rates by which dealers in bill markets quote the interest rate on U.S. Treasury bills; formally defined in Equation 8 of Chapter 4. Also known as the *discount yield*. **71**

yield to maturity The interest rate that equates the present value of payments received from a credit market instrument with its value today. **64**

zero-coupon bond See *discount bond*. **64**

ANSWERS TO SELECTED QUESTIONS AND PROBLEMS

Chapter 1 Why Study Money, Banking, and Financial Markets?

2. The data in Figures 1, 2, 3, and 4 suggest that real output, the inflation rate, and interest rates would all fall.
4. You might be more likely to buy a house or a car because the cost of financing them would fall, or you might be less likely to save because you earn less on your savings.
6. No. It is true that people who borrow to purchase a house or a car are worse off because it costs them more to finance their purchase; however, savers benefit because they can earn higher interest rates on their savings.
8. They channel funds from people who do not have a productive use for them to people who do, thereby resulting in higher economic efficiency.
10. The lower price for a firm's shares means that it can raise a smaller amount of funds, and so investment in facilities and equipment will fall.
12. It makes foreign goods more expensive, so British consumers will buy fewer foreign goods and more domestic goods.
14. In the mid- to late 1970s and in the late 1980s and early 1990s, the value of the dollar was low, making travel abroad relatively more expensive; thus it was a good time to vacation in the United States and see the Grand Canyon. With the rise in the dollar's value in the early 1980s, travel abroad became relatively cheaper, making it a good time to visit the Tower of London.

Chapter 2 An Overview of the Financial System

1. The share of IBM stock is an asset for its owner, because it entitles the owner to a share of the earnings and assets of IBM. The share is a liability for IBM, because it is a claim on its earnings and assets by the owner of the share.
3. Yes, because the absence of financial markets means that funds cannot be channeled to people who have the most productive use for them. Entrepreneurs then cannot acquire funds to set up businesses that would help the economy grow rapidly.
5. This statement is false. Prices in secondary markets determine the prices that firms issuing securities receive in primary markets. In addition, secondary markets make securities more liquid and thus easier to sell in the primary markets. Therefore, secondary markets are, if anything, more important than primary markets.
7. Because you know your family member better than a stranger, you know more about the borrower's honesty, propensity for risk taking, and other traits. There is less asymmetric information than with a stranger and less likelihood of an adverse selection problem, with the result that you are more likely to lend to the family member.
9. Loan sharks can threaten their borrowers with bodily harm if borrowers take actions that might jeopardize their paying off the loan. Hence borrowers from a loan shark are less likely to increase moral hazard.
11. Yes, because even if you know that a borrower is taking actions that might jeopardize paying off the loan, you must still stop the borrower from doing so. Because that may be costly, you may not spend the time and effort to reduce moral hazard, and so the problem of moral hazard still exists.
13. Because the costs of making the loan to your neighbor are high (legal fees, fees for a credit check, and so on), you will probably not be able to earn 5% on the loan after your expenses even though it has a 10% interest rate. You are better off depositing your savings with a financial intermediary and earning 5% interest. In addition, you are likely to bear less risk by depositing your savings at the bank rather than lending them to your neighbor.
15. Increased discussion of foreign financial markets in the U.S. press and the growth in markets for international financial instruments such as Eurodollars and Eurobonds.

Chapter 3 What Is Money?

2. Since the orchard owner likes only bananas but the banana grower doesn't like apples, the banana grower will not want apples in exchange for his bananas, and they will not trade. Similarly, the chocolatier will not be willing to trade with the banana grower because she does not like bananas. The orchard owner will not trade with the chocolatier because he doesn't like chocolate. Hence in a barter economy, trade among these three people may well not take place, because in no case is there a double coincidence of wants. However, if money is introduced into the economy, the orchard owner can sell his apples to the chocolatier and then use the money to buy bananas from the banana grower. Similarly, the banana grower can use the money she receives from the orchard owner to buy chocolate from the chocolatier, and the chocolatier can use the money to buy apples from the orchard owner. The result is that the need for a double coincidence of wants is eliminated, and everyone is better off because all three producers are now able to eat what they like best.
4. Because a check was so much easier to transport than gold, people would frequently rather be paid by check even if there was a possibility that the check might bounce. In other words, the lower transactions costs involved in handling checks made people more willing to accept them.
6. Because money was losing value at a slower rate (the inflation rate was lower) in the 1950s than in the 1970s, it was

then a better store of value, and you would have been willing to hold more of it.

9. Money loses its value at an extremely rapid rate in hyperinflation, so you want to hold it for as short a time as possible. Thus money is like a hot potato that is quickly passed from one person to another.
11. Not necessarily. Although the total amount of debt has predicted inflation and the business cycle better than M1, M2, or M3, it may not be a better predictor in the future. Without some theoretical reason for believing that the total amount of debt will continue to predict well in the future, we may not want to define money as the total amount of debt.
13. M1 contains the most liquid assets. M3 is the largest measure.
15. Revisions are not a serious problem for long-run movements of the money supply, because revisions for short-run (one-month) movements tend to cancel out. Revisions for long-run movements, such as one-year growth rates, are thus typically quite small.

Chapter 4 Understanding Interest Rates

1. Less. It would be worth $1/(1 + 0.20) = \$0.83$ when the interest rate is 20%, rather than $1/(1 + 0.10) = \$0.91$ when the interest rate is 10%.
3. $\$1{,}100/(1 + 0.10) + \$1{,}210/(1 + 0.10)^2 + \$1{,}331/(1 + 0.10)^3 = \$3{,}000$.
5. $\$2{,}000 = \$100/(1 + i) + \$100/(1 + i)^2 + \cdots + \$100/(1 + i)^{20} + \$1{,}000/(1 + i)^{20}$.
7. 14.9%, derived as follows: The present value of the $2 million payment five years from now is $\$2/(1 + i)^5$ million, which equals the $1 million loan. Thus $1 = 2/(1 + i)^5$. Solving for i, $(1 + i)^5 = 2$, so that $i = \sqrt[5]{2} - 1 = 0.149 = 14.9\%$.
9. If the one-year bond did not have a coupon payment, its yield to maturity would be $(\$1{,}000 - \$800)/\$800 = \$200/\$800 = 0.25 = 25\%$. Since it does have a coupon payment, its yield to maturity must be greater than 25%. However, because the current yield is a good approximation of the yield to maturity for a 20-year bond, we know that the yield to maturity on this bond is approximately 15%. Therefore, the one-year bond has a higher yield to maturity.
11. You would rather own the Treasury bill, because it has a higher yield to maturity. As the example in the text indicates, the discount yield's understatement of the yield to maturity for a one-year bond is substantial, exceeding one percentage point. Thus the yield to maturity on the one-year bill would be greater than 9%, the yield to maturity on the one-year Treasury bond.
13. No. If interest rates rise sharply in the future, long-term bonds may suffer such a sharp fall in price that their return might be quite low; possibly even negative.
15. The economists are right. They reason that nominal interest rates were below expected rates of inflation in the late 1970s, making real interest rates negative. The expected inflation rate, however, fell much faster than nominal interest rates in the mid-1980s, so nominal interest rates were above the expected inflation rate and real rates became positive.

Chapter 5 The Behavior of Interest Rates

2. (a) More, because your wealth has increased; (b) more, because it has become more liquid; (c) less, because its expected return has fallen relative to Microsoft stock; (d) more, because it has become less risky relative to stocks; (e) less, because its expected return has fallen.
4. (a) More, because they have become more liquid; (b) more, because their expected return has risen relative to stocks; (c) less, because they have become less liquid relative to stocks; (d) less, because their expected return has fallen; (e) more, because they have become more liquid.
6. When the Fed sells bonds to the public, it increases the supply of bonds, thus shifting the supply curve B^s to the right. The result is that the intersection of the supply and demand curves B^s and B^d occurs at a lower price and a higher equilibrium interest rate, and the interest rate rises. With the liquidity preference framework, the decrease in the money supply shifts the money supply curve M^s to the left, and the equilibrium interest rate rises. The answer from the loanable funds framework is consistent with the answer from the liquidity preference framework.
8. When the price level rises, the quantity of money in real terms falls (holding the nominal supply of money constant); to restore their holdings of money in real terms to their former level, people will want to hold a greater nominal quantity of money. Thus the money demand curve M^d shifts to the right, and the interest rate rises.
11. Interest rates would rise. A sudden increase in people's expectations of future real estate prices raises the expected return on real estate relative to bonds, so the demand for bonds falls. The demand curve B^d shifts to the left, and the equilibrium interest rate rises.
13. In the loanable funds framework, the increased riskiness of bonds lowers the demand for bonds. The demand curve B^d shifts to the left, and the equilibrium interest rate rises. The same answer is found in the liquidity preference framework. The increased riskiness of bonds relative to money increases the demand for money. The money demand curve M^d shifts to the right, and the equilibrium interest rate rises.
15. Yes, interest rates will rise. The lower commission on stocks makes them more liquid than bonds, and the demand for bonds will fall. The demand curve B^d will therefore shift to the left, and the equilibrium interest rate will rise.
17. The interest rate on the AT&T bonds will rise. Because people now expect interest rates to rise, the expected return on long-term bonds such as the $8\frac{1}{8}$s of 2022 will fall, and the demand for these bonds will decline. The demand curve B^d will therefore shift to the left, the price falls, and the equilibrium interest rate will rise.
19. Interest rates will rise. When bond prices become volatile and bonds become riskier, the demand for bonds will fall.

The demand curve B^d will shift to the left, the price falls, and the equilibrium interest rate will rise.

Chapter 6 The Risk and Term Structure of Interest Rates

2. U.S. Treasury bills have lower default risk and more liquidity than negotiable CDs. Consequently, the demand for Treasury bills is higher, and they have a lower interest rate.
4. True. When bonds of different maturities are close substitutes, a rise in interest rates for one bond causes the interest rates for others to rise because the expected returns on bonds of different maturities cannot get too far out of line.
6. (a) The yield to maturity would be 5% for a one-year bond, 6% for a two-year bond, 6.33% for a three-year bond, 6.5% for a four-year bond, and 6.6% for a five-year bond. (b) The yield to maturity would be 5% for a one-year bond, 4.5% for a two-year bond, 4.33% for a three-year bond, 4.25% for a four-year bond, and 4.2% for a five-year bond. The upward-sloping yield curve in (a) would be even steeper if people preferred short-term bonds over long-term bonds, because long-term bonds would then have a positive liquidity premium. The downward-sloping yield curve in (b) would be less steep and might even have a slight positive upward slope if the long-term bonds have a positive liquidity premium.
8. The flat yield curve at shorter maturities suggests that short-term interest rates are expected to fall moderately in the near future, while the steep upward slope of the yield curve at longer maturities indicates that interest rates further into the future are expected to rise. Because interest rates and expected inflation move together, the yield curve suggests that the market expects inflation to fall moderately in the near future but to rise later on.
10. The reduction in income tax rates would make the tax-exempt privilege for municipal bonds less valuable, and they would be less desirable than taxable Treasury bonds. The resulting decline in the demand for municipal bonds and increase in demand for Treasury bonds would raise interest rates on municipal bonds while causing interest rates on Treasury bonds to fall.
12. Lower brokerage commissions for corporate bonds would make them more liquid and thus increase their demand, which would lower their risk premium.
14. You would raise your predictions of future interest rates, because the higher long-term rates imply that the average of the expected future short-term rates is higher.

Chapter 7 The Stock Market, the Theory of Rational Expectations, and the Efficient Market Hypothesis

2. There are two cash flows from stock, periodic dividends and a future sales price. Dividends are frequently changed when firm earnings either rise or fall. The future sales price is also difficult to estimate, since it depends on the dividends that will be paid at some date even farther in the future. Bond cash flows also consist of two parts, periodic interest payments and a final maturity payment. These payments are established in writing at the time the bonds are issued and cannot be changed without the firm defaulting and being subject to bankruptcy. Stock prices tend to be more volatile, since their cash flows are more subject to change.
4. $P_0 = \dfrac{\$3 \times (1.07)}{.18 - .07} = \$29.18$
6. False. Expectations can be highly inaccurate and still be rational, because optimal forecasts are not necessarily accurate: A forecast is optimal if it is the best possible even if the forecast errors are large.
8. No, because he could improve the accuracy of his forecasts by predicting that tomorrow's interest rates will be identical to today's. His forecasts are therefore not optimal, and he does not have rational expectations.
10. No, you shouldn't buy stocks, because the rise in the money supply is publicly available information that will be already incorporated into stock prices. Hence you cannot expect to earn more than the equilibrium return on stocks by acting on the money supply information.
12. No, because this is publicly available information and is already reflected in stock prices. The optimal forecast of stock returns will equal the equilibrium return, so there is no benefit from selling your stocks.
14. No, if the person has no better information than the rest of the market. An expected price rise of 10% over the next month implies over a 100% annual return on IBM stock, which certainly exceeds its equilibrium return. This would mean that there is an unexploited profit opportunity in the market, which would have been eliminated in an efficient market. The only time that the person's expectations could be rational is if the person had information unavailable to the market that allowed him or her to beat the market.
16. False. The people with better information are exactly those who make the market more efficient by eliminating unexploited profit opportunities. These people can profit from their better information.
18. True, in principle. Foreign exchange rates are a random walk over a short interval such as a week, because changes in the exchange rate are unpredictable. If a change were predictable, large unexploited profit opportunities would exist in the foreign exchange market. If the foreign exchange market is efficient, these unexploited profit opportunities cannot exist and so the foreign exchange rate will approximately follow a random walk.
20. False. Although human fear may be the source of stock market crashes, that does not imply that there are unexploited profit opportunities in the market. Nothing in rational expectations theory rules out large changes in stock prices as a result of fears on the part of the investing public.

Chapter 8 An Economic Analysis of Financial Structure

2. Financial intermediaries develop expertise in such areas as computer technology so that they can inexpensively provide liquidity services such as checking accounts that lower transaction costs for depositors. Financial intermediaries can

also take advantage of economies of scale and engage in large transactions that have a lower cost per dollar per transaction.

4. Standard accounting principles make profit verification easier, thereby reducing adverse selection and moral hazard problems in financial markets and hence making them operate better. Standard accounting principles make it easier for investors to screen out good firms from bad firms, thereby reducing the adverse selection problem in financial markets. In addition, they make it harder for managers to understate profits, thereby reducing the principal–agent (moral hazard) problem.
6. Smaller firms that are not well known are the most likely to use bank financing. Since it is harder for investors to acquire information about these firms, it will be hard for the firms to sell securities in the financial markets. Banks that specialize in collecting information about smaller firms will then be the only outlet these firms have for financing their activities.
8. Yes. The person who is putting her life savings into her business has more to lose if she takes on too much risk or engages in personally beneficial activities that don't lead to higher profits. So she will act more in the interest of the lender, making it more likely that the loan will be paid off.
10. True. If the borrower turns out to be a bad credit risk and goes broke, the lender loses less, because the collateral can be sold to make up any losses on the loan. Thus adverse selection is not as severe a problem.
12. The separation of ownership and control creates a principal–agent problem. The managers (the agents) do not have as strong an incentive to maximize profits as the owners (the principals). Thus the managers might not work hard, might engage in wasteful spending on personal perks, or might pursue business strategies that enhance their personal power but do not increase profits.
14. A stock market crash reduces the net worth of firms and so increases the moral hazard problem. With less of an equity stake, owners have a greater incentive to take on risky projects and spend corporate funds on items that benefit them personally. A stock market crash, which increases the moral hazard problem, thus makes it less likely that lenders will be paid back. So lending and investment will decline, creating a financial crisis in which financial markets do not work well and the economy suffers.

Chapter 9 Banking and the Management of Financial Institutions

2. The rank from most to least liquid is (c), (b), (a), (d).
4. Reserves drop by $500. The T-account for the First National Bank is as follows:

FIRST NATIONAL BANK

Assets		Liabilities	
Reserves	−$500	Checkable deposits	−$500

6. The bank would rather have the balance sheet shown in this problem, because after it loses $50 million due to deposit outflow, the bank would still have excess reserves of $5 million: $50 million in reserves minus required reserves of $45 million (10% of the $450 million of deposits). Thus the bank would not have to alter its balance sheet further and would not incur any costs as a result of the deposit outflow. By contrast, with the balance sheet in Problem 5, the bank would have a shortfall of reserves of $20 million ($25 million in reserves minus the required reserves of $45 million). In this case, the bank will incur costs when it raises the necessary reserves through the methods described in the text.
8. No. When you turn a customer down, you may lose that customer's business forever, which is extremely costly. Instead, you might go out and borrow from other banks, corporations, or the Fed to obtain funds so that you can make the customer loans. Alternatively, you might sell negotiable CDs or some of your securities to acquire the necessary funds.
10. It can raise $1 million of capital by issuing new stock. It can cut its dividend payments by $1 million, thereby increasing its retained earnings by $1 million. It can decrease the amount of its assets so that the amount of its capital relative to its assets increases, thereby meeting the capital requirements.
12. Compensating balances can act as collateral. They also help establish long-term customer relationships, which make it easier for the bank to collect information about prospective borrowers, thus reducing the adverse selection problem. Compensating balances help the bank monitor the activities of a borrowing firm so that it can prevent the firm from taking on too much risk, thereby not acting in the interest of the bank.
14. The assets fall in value by $8 million (= $100 million × −2% × 4 years) while the liabilities fall in value by $10.8 million (= $90 million × −2% × 6 years). Since the liabilities fall in value by $2.8 million more than the assets do, the net worth of the bank rises by $2.8 million. The interest-rate risk can be reduced by shortening the maturity of the liabilities to a duration of four years or lengthening the maturity of the assets to a duration of six years. Alternatively, you could engage in an interest-rate swap, in which you swap the interest earned on your assets with the interest on another bank's assets that have a duration of six years.

Chapter 10 Banking Industry: Structure and Competition

2. (a) Office of the Comptroller of the Currency; (b) the Federal Reserve; (c) state banking authorities and the FDIC; (d) the Federal Reserve.
4. New technologies such as electronic banking facilities are frequently shared by several banks, so these facilities are not classified as branches. Thus they can be used by banks to escape limitations on offering services in other states and, in effect, to escape limitations from restrictions on branching.
6. Because restrictions on branching are stricter for commercial banks than for savings and loans. Thus small commercial

banks have greater protection from competition and are more likely to survive than small savings and loans.

8. International banking has been encouraged by giving special tax treatment and relaxed branching regulations to Edge Act corporations and to international banking facilities (IBFs); this was done to make American banks more competitive with foreign banks. The hope is that it will create more banking jobs in the United States.
10. No, because the Saudi-owned bank is subject to the same regulations as the American-owned bank.
12. The rise in inflation and the resulting higher interest rates on alternatives to checkable deposits meant that banks had a big shrinkage in this low-cost way of raising funds. The innovation of money market mutual funds also meant that the banks lost checking account business. The abolishment of Regulation Q and the appearance of NOW accounts did help decrease disintermediation, but raised the cost of funds for American banks, which now had to pay higher interest rates on checkable and other deposits. Foreign banks were also able to tap a large pool of domestic savings, thereby lowering their cost of funds relative to American banks.
14. The growth of the commercial paper market and the development of the junk bond market meant that corporations were now able to issue securities rather than borrow from banks, thus eroding the competitive advantage of banks on the lending side. Securitization has enabled other financial institutions to originate loans, again taking away some of the banks' loan business.

Chapter 11 Economic Analysis of Banking Regulation

2. There would be adverse selection, because people who might want to burn their property for some personal gain would actively try to obtain substantial fire insurance policies. Moral hazard could also be a problem, because a person with a fire insurance policy has less incentive to take measures to prevent a fire.
4. Regulations that restrict banks from holding risky assets directly decrease the moral hazard of risk taking by the bank. Requirements that force banks to have a large amount of capital also decrease the banks' incentives for risk taking, because banks now have more to lose if they fail. Such regulations will not completely eliminate the moral hazard problem, because bankers have incentives to hide their holdings of risky assets from the regulators and to overstate the amount of their capital.
6. The S&L crisis did not occur until the 1980s, because interest rates stayed low before then, so S&Ls were not subjected to losses from high interest rates. Also, the opportunities for risk taking were not available until the 1980s, when legislation and financial innovation made it easier for S&Ls to take on more risk, thereby greatly increasing the adverse selection and moral hazard problems.
8. FIRREA provided funds for the S&L bailout, created the Resolution Trust Corporation to manage the resolution of insolvent thrifts, eliminated the Federal Home Loan Bank Board and gave its regulatory role to the Office of Thrift Supervision, eliminated the FSLIC and turned its insurance role and regulatory responsibilities over to the FDIC, imposed restrictions on thrift activities similar to those in effect before 1982, increased the capital requirements to those adhered to by commercial banks, and increased the enforcement powers of thrift regulators.
10. If political candidates receive campaign funds from the government and are restricted in the amount they spend, they will have less need to satisfy lobbyists to win elections. As a result, they may have greater incentives to act in the interest of taxpayers (the principals), and so the political process might improve.
12. Eliminating or limiting the amount of deposit insurance would help reduce the moral hazard of excessive risk taking on the part of banks. It would, however, make bank failures and panics more likely, so it might not be a very good idea.
14. The economy would benefit from reduced moral hazard; that is, banks would not want to take on too much risk, because doing so would increase their deposit insurance premiums. The problem is, however, that it is difficult to monitor the degree of risk in bank assets because often only the bank making the loans knows how risky they are.

Chapter 12 Nonbank Finance

1. Because there would be more uncertainty about how much they would have to pay out in any given year, life insurance companies would tend to hold shorter-term assets that are more liquid.
3. Because benefits paid out are set to equal contributions to the plan and their earnings.
5. False. Government pension plans are often underfunded. Many pension plans for both federal and state employees are not fully funded.
7. Because the bigger the policy, the greater the moral hazard—the incentive for the policyholder to engage in activities that make the insurance payoff more likely. Because payoffs are costly, the insurance company will want to reduce moral hazard by limiting the amount of insurance.
9. Because interest rates on loans are typically lower at banks than at finance companies.
11. Because you do not have to pay a commission on a no-load fund, it is cheaper than a load fund, which does require a commission.
13. Government loan guarantees may be very costly, because like any insurance, they increase moral hazard. Because the banks and other institutions making the guaranteed loans do not suffer any losses if the loans default, these institutions have little incentive not to make bad loans. The resulting losses to the government can be substantial, as was true in past years.
15. No. Investment banking is a risky business, because if the investment bank cannot sell a security it is underwriting for the price it promised to pay the issuing firm, the investment bank can suffer substantial losses.

Chapter 13 Financial Derivatives

2. You would enter into a contract that specifies that you will sell the $25 million of 8s of 2015 at a price of 110 one year from now.
4. You have a loss of 6 points, or $6,000, per contract.
6. You would buy $100 million worth (1,000 contracts) of the call long-term bond option with a delivery date of one year in the future and with a strike price that corresponds to a yield of 8%. This means that you would have the option to buy the long bond with the 8% interest rate, thereby making sure that you can earn the 8%. The disadvantage of the options contract is that you have to pay a premium that you would not have to pay with a futures contract. The advantage of the options contract is that if the interest rate rises and the bond price falls during the next year, you do not have to exercise the option and so will be able to earn a higher rate than 8% when the funds come in next year, whereas with the futures contract, you have to take delivery of the bond and will only earn 8%.
8. You have a profit of 1 point ($1,000) when you exercise the contract, but you have paid a premium of $1,500 for the call option, so your net profit is −$500, a loss of $500.
10. Because for any given price at expiration, a lower strike price means a higher profit for a call option and a lower profit for a put option. A lower strike price makes a call option more desirable and raises its premium and makes a put option less desirable and lowers its premium.
12. It would swap interest on $42 million of fixed-rate assets for the interest on $42 million of variable-rate assets, thereby eliminating its income gap.
14. You would hedge the risk by buying 80 euro futures contracts that mature 3 months from now.

Chapter 14 Structure of Central Banks and the Federal Reserve System

1. Because of traditional American hostility to a central bank and centralized authority, the system of 12 regional banks was set up to diffuse power along regional lines.
3. Like the U.S. Constitution, the Federal Reserve System, originally established by the Federal Reserve Act, has many checks and balances and is a peculiarly American institution. The ability of the 12 regional banks to affect discount policy was viewed as a check on the centralized power of the Board of Governors, just as states' rights are a check on the centralized power of the federal government. The provision that there be three types of directors (A, B, and C) representing different groups (professional bankers, businesspeople, and the public) was again intended to prevent any group from dominating the Fed. The Fed's independence of the federal government and the setting up of the Federal Reserve banks as incorporated institutions were further intended to restrict government power over the banking industry.
5. The Board of Governors sets reserve requirements and the discount rate; the FOMC directs open market operations. In practice, however, the FOMC helps make decisions about reserve requirements and the discount rate.
7. The Board of Governors has clearly gained power at the expense of the regional Federal Reserve banks. This trend toward ever more centralized power is a general one in American government, but in the case of the Fed, it was a natural outgrowth of the Fed's having been given the responsibility for promoting a stable economy. This responsibility has required greater central direction of monetary policy, the role taken over the years by the Board of Governors and by the FOMC, which the board controls.
9. The threat that Congress will acquire greater control over the Fed's finances and budget.
11. False. Maximizing one's welfare does not rule out altruism. Operating in the public interest is clearly one objective of the Fed. The theory of bureaucratic behavior only points out that other objectives, such as maximizing power, also influence Fed decision making.
13. False. The Fed is still subject to political pressure, because Congress can pass legislation limiting the Fed's power. If the Fed is performing badly, Congress can therefore make the Fed accountable by passing legislation that the Fed does not like.
15. The argument for not releasing the FOMC directives immediately is that it keeps Congress off the Fed's back, thus enabling the Fed to pursue an independent monetary policy that is less subject to inflation and political business cycles. The argument for releasing the directive immediately is that it would make the Fed more accountable.

Chapter 15 Multiple Deposit Creation and the Money Supply Process

2. Reserves are unchanged, but the monetary base falls by $2 million, as indicated by the following T-accounts:

IRVING THE INVESTOR

Assets		Liabilities	
Currency	−$2 million		
Securities	+$2 million		

FEDERAL RESERVE SYSTEM

Assets		Liabilities	
Securities	−$2 million	Currency	−$2 million

3. Reserves increase by $50 million, but the monetary base increases by $100 million, as the T-accounts for the five banks and the Fed indicate:

FIVE BANKS

Assets		Liabilities	
Reserves	+$50 million	Discount loans	+$100 million
		Deposits	−$50 million

FEDERAL RESERVE SYSTEM

Assets		Liabilities	
Discount loans	+$100 million	Reserves	+$50 million
		Currency	+$50 million

5. The T-accounts are identical to those in the sections "Deposit Creation: The Single Bank" and "Deposit Creation: The Banking System" except that all the entries are multiplied by 10,000 (that is, $100 becomes $1 million). The net result is that checkable deposits rise by $10 million.
7. The $1 million Fed purchase of bonds increases reserves in the banking system by $1 million, and the total increase in checkable deposits is $10 million. The fact that banks buy securities rather than make loans with their excess reserves makes no difference in the multiple deposit creation process.
9. Reserves in the banking system fall by $1,000, and a multiple contraction occurs, reducing checkable deposits by $10,000.
11. The level of checkable deposits falls by $50 million. The T-account of the banking system in equilibrium is as follows:

BANKING SYSTEM

Assets		Liabilities	
Reserves	−$5 million	Checkable deposits	−$50 million
Securities	+$5 million		
Loans	−$50 million		

13. The $1 million holdings of excess reserves means that the bank has to reduce its holdings of loans or securities, thus starting the multiple contraction process. Because the required reserve ratio is 10%, checkable deposits must decline by $10 million.
15. The deposit of $100 in the bank increases its reserves by $100. This starts the process of multiple deposit expansion, leading to an increase in checkable deposits of $1,000.

Chapter 16 Determinants of the Money Supply

1. Uncertain. As the formula in Equation 4 indicates, if $r_D + e$ is greater than 1, the money multiplier can be less than 1. In practice, however, e is so small that $r_D + e$ is less than 1 and the money multiplier is greater than 1.
3. The money supply fell sharply because when c rose, there was a shift from one component of the money supply (checkable deposits) with more multiple expansion to another (currency) with less. Overall multiple deposit expansion fell, leading to a decline in the money supply.
5. There is a shift from one component of the money supply (checkable deposits) with less multiple expansion to another (traveler's checks) with more. Multiple expansion therefore increases, and the money supply increases.
7. Yes, because with no reserve requirements on time deposits, a shift from checkable deposits (with less multiple expansion) to time deposits (with more multiple expansion) increases the total amount of deposits and raises M2. However, if reserve requirements were equal for both types of deposits, they would both undergo the same amount of multiple expansion, and a shift from one to the other would have no effect on M2. Thus control of M2 would be better because random shifts from time deposits to checkable deposits or vice versa would not affect M2.
9. Both the Fed's purchase of $100 million of bonds (which raises the monetary base) and the lowering of r (which increases the amount of multiple expansion and raises the money multiplier) lead to a rise in the money supply.
11. The Fed's sale of $1 million of bonds shrinks the monetary base by $1 million, and the reduction of discount loans also lowers the monetary base by another $1 million. The resulting $2 million decline in the monetary base leads to a decline in the money supply.
13. A rise in expected inflation would increase interest rates (through the Fisher effect), which would in turn cause e to fall and the volume of discount loans to rise. The fall in e increases the amount of reserves available to support checkable deposits so that deposits and the money multiplier will rise. The rise in discount loans causes the monetary base to rise. The resulting increase in the money multiplier and the monetary base leads to an increase in the money supply.
15. The money supply would fall, because if the discount window were eliminated, banks would need to hold more excess reserves, making fewer reserves available to support deposits. Moreover, abolishing discounting would reduce

the volume of discount loans, which would also cause the monetary base and the money supply to fall.

Chapter 17 Tools of Monetary Policy

1. The snowstorm would cause float to increase, which would increase the monetary base. To counteract this effect, the manager will undertake a defensive open market sale.
3. As we saw in Chapter 15, when the Treasury's deposits at the Fed fall, the monetary base increases. To counteract this increase, the manager would undertake an open market sale.
5. It suggests that defensive open market operations are far more common than dynamic operations because repurchase agreements are used primarily to conduct defensive operations to counteract temporary changes in the monetary base.
7. A rise in checkable deposits leads to a rise in required reserves at any given interest rate and thus shifts the demand curve to the right. If the federal funds rate is initially below the discount rate, this then leads to a rise in the federal funds rate. If the federal funds rate is initially at the discount rate, then the federal funds rate will just remain at the discount rate.
9. This statement is incorrect. The FDIC would not be effective in eliminating bank panics without Fed discounting to troubled banks in order to keep bank failures from spreading.
11. Most likely not. If the federal funds rate target is initially below the discount rate and the decline in the discount rate still leaves it above the federal funds target, then the shift in the supply curve has no effect on the federal funds rate. Furthermore, the Fed usually moves the discount rate in line with changes in the federal funds rate target, so that changes in the discount rate provide no additional information about the direction of monetary policy.
13. False. As the analysis of the channel/corridor approach to setting interest rates demonstrates, central banks can still tightly control interest rates by putting in place standing facilities where the difference between the interest rate paid on reserves kept at the central bank and the interest rate charged in central bank loans to banks is kept small.
15. Open market operations are more flexible, reversible, and faster to implement than the other two tools. Discount policy is more flexible, reversible, and faster to implement than changing reserve requirements, but it is less effective than either of the other two tools.

Chapter 18 Conduct of Monetary Policy: Goals and Targets

1. Disagree. Some unemployment is beneficial to the economy because the availability of vacant jobs makes it more likely that a worker will find the right job and that the employer will find the right worker for the job.
3. True. In such a world, hitting a monetary target would mean that the Fed would also hit its interest target, or vice versa. Thus the Fed could pursue both a monetary target and an interest-rate target at the same time.
5. The Fed can control the interest rate on three-month Treasury bills by buying and selling them in the open market. When the bill rate rises above the target level, the Fed would buy bills, which would bid up their price and lower the interest rate to its target level. Similarly, when the bill rate falls below the target level, the Fed would sell bills to raise the interest rate to the target level. The resulting open market operations would of course affect the money supply and cause it to change. The Fed would be giving up control of the money supply to pursue its interest-rate target.
7. Disagree. Although *nominal* interest rates are measured more accurately and more quickly than the money supply, the interest-rate variable that is of more concern to policymakers is the *real* interest rate. Because the measurement of real interest rates requires estimates of expected inflation, it is not true that real interest rates are necessarily measured more accurately and more quickly than the money supply. Interest-rate targets are therefore not necessarily better than money supply targets.
9. Because the Fed did not lend to troubled banks during this period, massive bank failures occurred, leading to a decline in the money supply when depositors increased their holdings of currency relative to deposits and banks increased their excess reserves to protect themselves against runs. As the money supply model presented in Chapters 15–16 indicates, these decisions by banks and depositors led to a sharp contraction of the money supply.
11. When the economy enters a recession, interest rates usually fall. If the Fed is targeting interest rates, it tries to prevent a decline in interest rates by selling bonds, thereby lowering their prices and raising interest rates to the target level. The open market sale would then lead to a decline in the monetary base and in the money supply. Therefore, an interest-rate target can sometimes be problematic if it is left unchanged too long because it can lead to a slower rate of money supply growth during a recession, just when the Fed would want money growth to be higher.
13. A borrowed reserves target will produce smaller fluctuations in the federal funds rate. In contrast to what happens when there is a nonborrowed reserves target, when the federal funds rate rises with a borrowed reserves target, the Fed prevents the tendency of discount borrowings to rise by buying bonds to lower interest rates. The result is smaller fluctuations in the federal funds rate with a borrowed reserves target.
15. The Fed may prefer to control interest rates rather than the money supply because it wishes to avoid the conflict with Congress that occurs when interest rates rise. The Fed might also believe that interest rates are actually a better guide to future economic activity.

Chapter 19 The Foreign Exchange Market

2. False. Although a weak currency has the negative effect of making it more expensive to buy foreign goods or to travel abroad, it may help domestic industry. Domestic goods become cheaper relative to foreign goods, and the demand for domestically produced goods increases. The resulting

higher sales of domestic products may lead to higher employment, a beneficial effect on the economy.

4. It predicts that the value of the yen will fall 5% in terms of dollars.
6. Even though the Japanese price level rose relative to the American, the yen appreciated because the increase in Japanese productivity relative to American productivity made it possible for the Japanese to continue to sell their goods at a profit at a high value of the yen.
8. The pound depreciates but overshoots, declining by more in the short run than in the long run. Consider Britain the domestic country. The rise in the money supply leads to a higher domestic price level in the long run, which leads to a lower expected future exchange rate. The resulting expected depreciation of the pound raises the expected return on foreign deposits, shifting R^F to the right. The rise in the money supply lowers the interest rate on pound deposits in the short run, which shifts R^D to the left. The short-run outcome is a lower equilibrium exchange rate. However, in the long run, the domestic interest rate returns to its previous value, and R^D shifts back to its original position. The exchange rate rises to some extent, although it still remains below its initial position.
10. The dollar will depreciate. A rise in nominal interest rates but a decline in real interest rates implies a rise in expected inflation that produces an expected depreciation of the dollar that is larger than the increase in the domestic interest rate. As a result, the expected return on foreign deposits rises by more than the expected return on domestic deposits. R^F shifts rightward more than R^D, so the equilibrium exchange rate falls.
12. The dollar will depreciate. An increased demand for imports would lower the expected future exchange rate and result in an expected appreciation of the foreign currency. The higher resulting expected return on foreign deposits shifts the R^F schedule to the right, and the equilibrium exchange rate falls.
14. The contraction of the European money supply will increase European interest rates and raise the future value of the euro, both of which will shift R^F (with Europe as the foreign country) to the right. The result is a decline in the value of the dollar.

Chapter 20 The International Financial System

2. The purchase of dollars involves a sale of foreign assets, which means that international reserves fall and the monetary base falls. The resulting fall in the money supply causes interest rates to rise and R^D to shift to the right while it lowers the future price level, thereby raising the future expected exchange rate, causing R^F to shift to the left. The result is a rise in the exchange rate. However, in the long run, the R^D curve returns to its original position, and so there is overshooting.
4. Because other countries often intervene in the foreign exchange market when the United States has a deficit so that U.S. holdings of international reserves do not change. By contrast, when the Netherlands has a deficit, it must intervene in the foreign exchange market and buy euros, which results in a reduction of international reserves for the Netherlands and Euroland.
6. Two francs per dollar.
8. A large balance-of-payments surplus may require a country to finance the surplus by selling its currency in the foreign exchange market, thereby gaining international reserves. The result is that the central bank will have supplied more of its currency to the public, and the monetary base will rise. The resulting rise in the money supply can cause the price level to rise, leading to a higher inflation rate.
10. In order to finance the deficits, the central bank in these countries might intervene in the foreign exchange market and buy domestic currency, thereby implementing a contractionary monetary policy. The result is that they sell off international reserves and their monetary base falls, leading to a decline in the money supply.
12. When other countries buy U.S. dollars to keep their exchange rates from changing vis-à-vis the dollar because of the U.S. deficits, they gain international reserves and their monetary base increases. The outcome is that the money supply in these countries grows faster and leads to higher inflation throughout the world.
14. There are no direct effects on the money supply, because there is no central bank intervention in a pure flexible exchange rate regime; therefore, changes in international reserves that affect the monetary base do not occur. However, monetary policy can be affected by the foreign exchange market, because monetary authorities may want to manipulate exchange rates by changing the money supply and interest rates.

Chapter 21 Monetary Policy Strategy: The International Experience

4. First is that the exchange-rate target directly keeps inflation under control by tying the inflation rate for internationally traded goods to that found in the anchor country to which its currency is pegged. Second is that it provides an automatic rule for the conduct of monetary policy that helps mitigate the time-inconsistency problem. Third, it has the advantage of simplicity and clarity.
6. With a pegged exchange rate, speculators are sometimes presented with a one-way bet in which the only direction for a currency to go is down in value. In this case, selling the currency before the likely depreciation gives speculators an attractive profit opportunity with potentially high expected returns. As a result, they jump on board and attack the currency.
8. The long-term bond market can help reduce the time-consistency problem because politicians and central banks will realize that pursuing an overly expansionary policy will lead to an inflation scare in which inflation expectations surge, interest rates rise, and there is a sharp fall in long-term bond prices. Similarly, they will realize that overly expansionary

monetary policy will result in a sharp fall in the value of the currency. Avoiding these outcomes constrains policymakers and politicians so time-consistent monetary policy is less likely to occur.

10. A currency board has the advantage that the central bank no longer can print money to create inflation, and so it is a stronger commitment to a fixed exchange rate. The disadvantage is that it is still subject to a speculative attack, which can lead to a sharp contraction of the money supply. In addition, a currency board limits the ability of the central bank to play a lender-of-last-resort role.
12. Monetary targeting has the advantage that it enables a central bank to adjust its monetary policy to cope with domestic considerations. Furthermore, information on whether the central bank is achieving its target is known almost immediately.
14. Inflation-targeting central banks engage in extensive public information campaigns that include the distribution of glossy brochures, the publication of *Inflation Report*–type documents, making speeches to the public, and continual communication with the elected government.
16. Uncertain. If the relationship between monetary aggregates and the goal variable, say inflation, is unstable, then the signal provided by the monetary aggregates is not very useful and is not a good indicator of whether the stance of monetary policy is correct.
18. With a nominal GDP target, a decline in projected real output growth would automatically imply an increase in the central bank's inflation target. This increase would tend to be stabilizing because it would automatically lead to an easier monetary policy. Nominal GDP targeting does suffer from potential confusion about what nominal GDP is and from the political complications that arise because nominal GDP requires the announcement of a potential GDP growth path.
20. All allow a central bank to pursue an independent monetary policy that can focus on domestic considerations.

Chapter 22 The Demand for Money

1. Velocity is approximately 10 in 2001, 11 in 2002, and 12 in 2003. The rate of velocity growth is approximately 10% per year.
3. Nominal GDP declines by approximately 10%.
5. The price level quadruples.
7. The two largest declines are during the recession in 1920 and the Great Depression of 1929-33. These declines suggest that velocity is procyclical, i.e., it rises in business cycle upturns and falls in business cycle downturns. The data in Figure 1 indicates that it is not reasonable to assume that declines in the quantity of money cause declines in aggregate spending because when aggregate spending declines it could just reflect the fact that velocity declines at that time.
9. The demand for money will decrease. People would be more likely to expect interest rates to fall and therefore more likely to expect bond prices to rise. The increase in the expected return on bonds relative to money will then mean that people would demand less money.
11. Money balances should average one-half of Grant's monthly income, because he would hold no bonds, since holding them would entail additional brokerage costs but would not provide him with any interest income.
13. True. Because bonds are riskier than money, risk-averse people would be likely to want to hold both.
15. In Keynes's view, velocity is unpredictable because interest rates, which have large fluctuations, affect the demand for money and hence velocity. In addition, Keynes's analysis suggests that if people's expectations of the normal level of interest rates change, the demand for money changes. Keynes thought that these expectations moved unpredictably, meaning that money demand and velocity are also unpredictable. Friedman sees the demand for money as stable, and because he also believes that changes in interest rates have only small effects on the demand for money, his position is that the demand for money, and hence velocity, is predictable.

Chapter 23 The Keynesian Framework and the *ISLM* Model

2. Companies cut production when their unplanned inventory investment is greater than zero, because they are then producing more than they can sell. If they continue at current production, profits will suffer because they are building up unwanted inventory, which is costly to store and finance.
4. The equilibrium level of output is 1,500. When planned investment spending falls by 100, the equilibrium level of output falls by 500 to 1,000.
6. Nothing. The $100 billion increase in planned investment spending is exactly offset by the $100 billion decline in autonomous consumer expenditure, and autonomous spending and aggregate output remain unchanged.
8. Equilibrium output of 2,000 occurs at the intersection of the 45° line $Y = Y^{ad}$ and the aggregate demand function $Y^{ad} = C + I + G = 500 + 0.75Y$. If government spending rises by 100, equilibrium output will rise by 400 to 2,400.
10. Taxes should be reduced by $400 billion because the increase in output for a $\$T$ decrease in taxes is $\$T$; that is, it equals the change in autonomous spending $mpc \times T$ times the multiplier $1/(1 - mpc) = (mpc \times T)\,[1/(1 - mpc)] = 0.5T\,[1/(1 - 0.5)] = 0.5T/0.5 = T$.
12. Rise. The fall in autonomous spending from an increase in taxes is always less than the change in taxes because the marginal propensity to consume is less than 1. By contrast, autonomous spending rises one-for-one with a change in autonomous consumer expenditure. So if taxes and autonomous consumer expenditure rise by the same amount, autonomous spending must rise, and aggregate output also rises.
14. When aggregate output falls, the demand for money falls, shifting the money demand curve to the left, which causes the equilibrium interest rate to fall. Because the equilibrium interest rate falls when aggregate output falls, there is a pos-

itive association between aggregate output and the equilibrium interest rate, and the *LM* curve slopes up.

Chapter 24 Monetary and Fiscal Policy in the *ISLM* Model

2. When investment spending collapsed, the aggregate demand function in the Keynesian cross diagram fell, leading to a lower level of equilibrium output for any given interest rate. The fall in equilibrium output for any given interest rate implies that the *IS* curve shifted to the left.
4. False. It can also be eliminated by a fall in aggregate output, which lowers the demand for money and brings it back into equality with the supply of money.
6. The *ISLM* model gives exactly this result. The tax cuts shifted the *IS* curve to the right, while tight money shifted the *LM* curve to the left. The interest rate at the intersection of the new *IS* and *LM* curves is necessarily higher than at the initial equilibrium, and aggregate output can be higher.
8. Because it suggests that an interest-rate target is better than a money supply target. The reason is that unstable money demand increases the volatility of the *LM* curve relative to the *IS* curve, and as demonstrated in the text, this makes it more likely that an interest-rate target is preferred to a money supply target.
10. The effect on the aggregate demand curve is uncertain. A rise in government spending would shift the *IS* curve to the right, raising equilibrium output for a given price level. But the reduction in the money supply would shift the *LM* curve to the left, lowering equilibrium output for a given price level. Depending on which of these two effects on equilibrium output is stronger, the aggregate demand curve could shift either to the right or to the left.
12. No effect. The *LM* curve would be vertical in this case, meaning that a rise in government spending and a rightward shift in the *IS* curve would not lead to higher aggregate output but rather only to a rise in the interest rate. For any given price level, therefore, equilibrium output would remain the same, and the aggregate demand curve would not shift.
14. The increase in net exports shifts the *IS* curve to the right, and the equilibrium level of interest rates and aggregate output will rise.

Chapter 25 Aggregate Demand and Supply Analysis

2. Because the position of the aggregate demand curve is fixed if nominal income ($P \times Y$) is fixed, Friedman's statement implies that the position of the aggregate demand curve is completely determined by the quantity of money. This is built into the monetarist aggregate demand curve because it shifts only when the money supply changes.
4. The Keynesian aggregate demand curve shifts because a change in "animal spirits" causes consumer expenditure or planned investment spending to change, which then causes the quantity of aggregate output demanded to change at any given price level. In the monetarist view, by contrast, a change in "animal spirits" has little effect on velocity, and aggregate spending ($P \times Y$) remains unchanged; hence the aggregate demand curve does not shift.
6. True. Given fixed production costs, firms can earn higher profits by producing more when prices are higher. Profit-maximizing behavior on the part of firms thus leads them to increase production when prices are higher.
8. The aggregate supply curve would shift to the right because production costs would fall.
10. The collapse in investment spending during the Great Depression reduced the quantity of output demanded at any given price level and shifted the aggregate demand curve to the left. In an aggregate demand and supply diagram, the equilibrium price level and aggregate output would then fall, which explains the decline in aggregate output and the price level that occurred during the Great Depression.
12. Both the increase in the money supply and the income tax cut will increase the quantity of output demanded at any given price level and so will shift the aggregate demand curve to the right. The intersection of the aggregate demand and aggregate supply curve will be at a higher level of both output and price level in the short run. However, in the long run, the aggregate supply curve will shift leftward, leaving output at the natural rate level, but the price level will be even higher.
14. Because goods would cost more, the national sales tax would raise production costs, and the aggregate supply curve would shift to the left. The intersection of the aggregate supply curve with the aggregate demand curve would then be at a higher level of prices and a lower level of aggregate output; aggregate output would fall, and the price level would rise.

Chapter 26 Transmission Mechanisms of Monetary Policy: The Evidence

4. Seeing which car is built better produces structural model evidence, because it explains why one car is better than the other (that is, how the car is built). Asking owners how often their cars undergo repairs produces reduced-form evidence, because it looks only at the correlation of reliability with the manufacturer of the car.
5. Not necessarily. If GM car owners change their oil more frequently than Ford owners, GM cars would have better repair records, even though they are not more reliable cars. In this case, it is a third factor, the frequency of oil changes, that leads to the better repair record for GM cars.
6. Not necessarily. Although the Ford engine might be built better than the GM engine, the rest of the GM car might be better made than the Ford. The result could be that the GM car is more reliable than the Ford.
8. If the Fed has interest-rate targets, a rise in output that raises interest rates might cause the Fed to buy bonds and bid up their price in order to drive interest rates back down to their target level (see Chapter 5). The result of these open market purchases would be that the increase in output would cause an increase in the monetary base and hence an increase in

the money supply. In addition, a rise in output and interest rates would cause free reserves to fall (because excess reserves would fall and the volume of discount loans would rise). If the Fed has a free reserves target, the increase in aggregate output will then cause the Fed to increase the money supply because it believes that money is tight.

10. Monetarists went on to refine their reduced-form models with more sophisticated statistical procedures, one outcome of which was the St. Louis model. Keynesians began to look for transmission mechanisms of monetary policy that they may have ignored.
12. False. Monetary policy can affect stock prices, which affect Tobin's *q,* thereby affecting investment spending. In addition, monetary policy can affect loan availability, which may also influence investment spending.
14. There are three mechanisms involving consumer expenditure. First, a rise in the money supply lowers interest rates and reduces the cost of financing purchases of consumer durables, and consumer durable expenditure rises. Second, a rise in the money supply causes stock prices and wealth to rise, leading to greater lifetime resources for consumers and causing them to increase their consumption. Third, a rise in the money supply that causes stock prices and the value of financial assets to rise also lowers people's probability of financial distress, and so they spend more on consumer durables.

Chapter 27 Money and Inflation

2. Because hyperinflations appear to be examples in which the increase in money supply growth is an exogenous event, the fact that hyperinflation occurs when money growth is high is powerful evidence that a high rate of money growth causes inflation.
4. False. Although workers' attempts to push up their wages can lead to inflation if the government has a high employment target, inflation is still a monetary phenomenon, because it cannot occur without accommodating monetary policy.
6. True. If financed with money creation, a temporary budget deficit can lead to a onetime rightward shift in the aggregate demand curve and hence to a onetime increase in the price level. However, once the budget deficit disappears, there is no longer any reason for the aggregate demand curve to shift. Thus a temporary deficit cannot lead to a continuing rightward shift of the aggregate demand curve and therefore cannot produce inflation, a continuing increase in the price level.
8. True. The monetarist objection to activist policy would no longer be as serious. The aggregate demand curve could be quickly moved to AD_2 in Figure 11, and the economy would move quickly to point 2 because the aggregate supply curve would not have as much time to shift. The scenario of a highly variable price level and output would not occur, making an activist policy more desirable.
10. True, if expectations about policy affect the wage-setting process. In this case, workers and firms are more likely to push up wages and prices because they know that if they do so and unemployment develops as a result, the government will pursue expansionary policies to eliminate the unemployment. Therefore, the cost of pushing up wages and prices is lower, and workers and firms will be more likely to do it.
12. True. If expectations about policy have no effect on the aggregate supply curve, a cost-push inflation is less likely to develop when policymakers pursue an activist accommodating policy. Furthermore, if expectations about policy do not matter, pursuing a nonaccommodating, nonactivist policy does not have the hidden benefit of making it less likely that workers will push up their wages and create unemployment. The case for an activist policy is therefore stronger.
14. The Fed's big stick is the ability to let unemployment develop as a result of a wage push by not trying to eliminate unemployment with expansionary monetary policy. The statement proposes that the Fed should pursue a nonaccommodating policy because this will prevent cost-push inflation and make it less likely that unemployment develops because of workers' attempts to push up their wages.

Chapter 28 Rational Expectations: Implications for Policy

2. A tax cut that is expected to last for ten years will have a larger effect on consumer expenditure than one that is expected to last only one year. The reason is that the longer the tax cut is expected to last, the greater its effect on expected average income and consumer expenditure.
4. True, if the anti-inflation policy is credible. As shown in Figure 6, if anti-inflation policy is believed (and hence expected), there is no output loss in the new classical model (the economy stays at point 1 in panel b), and there is a smaller output loss than would otherwise be the case in the new Keynesian model (the economy goes to point 2″ rather than point 2′ in panel c).
6. Uncertain. It is true that policymakers can reduce unemployment by pursuing a more expansionary policy than the public expects. However, the rational expectations assumption indicates that the public will attempt to anticipate policymakers' actions. Policymakers cannot be sure whether expansionary policy will be more or less expansionary than the public expects and hence cannot use policy to make a predictable impact on unemployment.
8. True, because the Lucas critique indicates that the effect of policy on the aggregate demand curve depends on the public's expectations about that policy. The outcome of a particular policy is therefore less certain in Lucas's view than if expectations about it do not matter, and it is harder to design a beneficial activist stabilization policy.
10. Yes, if budget deficits are expected to lead to an inflationary monetary policy and expectations about monetary policy affect the aggregate supply curve. In this case, a large budget deficit would cause the aggregate supply curve to shift more

to the left because expected inflation would be higher. The result is that the increase in the price level (the inflation rate) would be higher.

13. The aggregate supply curve would shift to the left less than the aggregate demand curve shifts to the right; hence at their intersection, aggregate output would rise and the price level would be higher than it would have been if money growth had been reduced to a rate of 2%.
14. Using the traditional model, the aggregate supply curve would continue to shift leftward at the same rate, and the smaller rightward shift of the aggregate demand curve because money supply growth has been reduced would mean a smaller increase in the price level and a reduction of aggregate output. In the new Keynesian model, the effect of this anti-inflation policy on aggregate output is uncertain. The aggregate supply curve would not shift leftward by as much as in the traditional model, because the anti-inflation policy is expected, but it would shift to the left by more than in the new classical model. Hence inflation falls, but aggregate output may rise or fall, depending on whether the aggregate supply curve shifts to the left more or less than the aggregate demand curve shifts to the right.

CREDITS

Page 30. Following the Financial News: Foreign Stock Market Indexes. "International Stock Market Indexes" from *The Wall Street Journal,* January 21, 2003, p. C6. Republished by permission of Dow Jones, Inc. via Copyright Clearance Center, Inc. © 2003 Dow Jones & Co., Inc. All Rights Reserved Worldwide.

Page 54. Following the Financial News: The Monetary Aggregates. "Monetary Aggregates" from *The Wall Street Journal,* January 3, 2003, p. C10. Republished by permission of Dow Jones, Inc. via Copyright Clearance Center, Inc. © 2003 Dow Jones & Co., Inc. All Rights Reserved Worldwide.

Page 73. Following the Financial News: Bond Prices and Interest Rates. "Treasury Bonds, Notes and Bills" from *The Wall Street Journal,* January 23, 2003, p. C11. Republished by permission of Dow Jones, Inc. via Copyright Clearance Center, Inc. © 2003 Dow Jones & Co., Inc. All Rights Reserved Worldwide.

Page 104. Following the Financial News: The "Credit Markets" Column. "Treasurys Drop Ahead of Bush Stimulus Package" from *The Wall Street Journal*, January 7, 2003, p. C14. Republished by permission of Dow Jones, Inc. © 2003 Dow Jones & Co., Inc. All Rights Reserved Worldwide.

Page 113. Following the Financial News: Forecasting Interest Rates. "The Wall Street Journal Forecasting Survey for 2003" from *The Wall Street Journal,* January 2, 2003, p. A2. Republished by permission of Dow Jones, Inc. via Copyright Clearance Center, Inc. © 2003 Dow Jones & Co., Inc. All Rights Reserved Worldwide.

Page 128. Following the Financial News: Yield Curves. "Treasury Yield Curve" from *The Wall Street Journal,* January 22, 2003, p. C2. Republished by permission of Dow Jones, Inc. via Copyright Clearance Center, Inc. © 2003 Dow Jones & Co., Inc. All Rights Reserved Worldwide.

Page 159. Following the Financial News: Stock Prices. Chart from *The Wall Street Journal,* January 3, 2003, p. C4. Republished by permission of Dow Jones, Inc. via Copyright Clearance Center, Inc. © 2003 Dow Jones & Co., Inc. All Rights Reserved Worldwide.

Page 304. Following the Financial News: New Securities Issues. Advertisement from *The Wall Street Journal,* January 12, 2003, p. C5. Republished by permission of Dow Jones, Inc. via Copyright Clearance Center, Inc. © 2003 Dow Jones & Co., Inc. All Rights Reserved Worldwide.

Page 312. Following the Financial News: Financial Futures. "Interest Rate Futures" from *The Wall Street Journal,* January 31, 2003, p. B6. Republished by permission of Dow Jones, Inc. via Copyright Clearance Center, Inc. © 2003 Dow Jones & Co., Inc. All Rights Reserved Worldwide.

Page 316. Following the Financial News: Table 1: Widely Traded Financial Futures Contracts from *The Wall Street Journal,* January 30, 2003, p. C21. Republished by permission of Dow Jones, Inc. via Copyright Clearance Center, Inc. © 2003 Dow Jones & Co., Inc. All Rights Reserved Worldwide.

Page 321. Following the Financial News: Futures Options. "Interest Rate" from *The Wall Street Journal,* February 14, 2003, p. C10. Republished by permission of Dow Jones, Inc. via Copyright Clearance Center, Inc. © 2003 Dow Jones & Co., Inc. All Rights Reserved Worldwide.

Page 437. Following the Financial News: Foreign Exchange Rates. "Exchange Rates" from *The Wall Street Journal,* February 6, 2003, p. C12. Republished by permission of Dow Jones, Inc. via Copyright Clearance Center, Inc. © 2003 Dow Jones & Co., Inc. All Rights Reserved Worldwide.

Page 458. Following the Financial News: The "Currency Trading" Column. "Concerns About War Put Pressure on the Dollar" from *The Wall Street Journal*, January 13, 2003, p. C16. Republished by permission of Dow Jones, Inc. © 2003 Dow Jones & Co., Inc. All Rights Reserved Worldwide.

INDEX

Aaa rated bonds, 124–127
ABM (automated banking machine), 235
Accommodating policy, 640
Activist/nonactivist policy debate, 650–655
Activists, 592
Activities, restrictions on, 39–40
Adaptive expectations, 147
Adjustable-rate mortgages (ARMs), 233
Adverse selection, 32–34, 174–175
  property and casualty companies, 291
  regulations, 263
Advice, investment, 160
Afghanistan, 146–147
Agency theory, 175
Agent problem, 225
Aggregate demand curves, 582–587
  ISLM model, 577–580
Aggregate demand function, 541
Aggregate economic activity, 189–198
Aggregate output, 9, 583. *See also* GDP
  changes in equilibrium level, 566–568
  income, 20
  international trade, 548
  ISLM model, 557–558
  Keynesian framework, 536–551
  money supply, 567
  real terms, 575
Aggregate price level, 10, 21–22
Aggregate supply curve, 587–588
  equilibrium in, 588–600
Agriculture, Department of, 315
Akerlof, George, 175
Alexander, Sidney, 155
Allen, F., 155
Allied Irish, 225
Allstate, 288
American Express, 234, 305
American options, 320
AMEX (American Stock Exchange), 26, 305
Animal spirits, 544, 586
Annuities, 288
Anticipated expansionary policy, 670
Anti-inflation polices, 671
Applications, insurance, 291
Appreciation, 436
Arbitrage, 313
Argentina
  currency board, 494
  hyperinflation, 413
ARMs (adjustable-rate mortgages), 233
Arthur Andersen, 178
Asian financial crises, 194–198
Assets, 3
  costs, 241–242
  determinants of demand, 85–87
  effects on market balance sheets, 189
  Federal Reserve System, 464
  liquidity, 47
  management, 208, 211–212
  market approach, 93
  physical bank capital, 205
  price channels, 618
  restrictions on, 39–40
  ROA, 214
  wealth, 85. *See also* wealth
Asset transformation, 32
Asymmetric information, 32, 174–175, 187
  banking regulation, 260–271
ATMs (automated teller machines), 51, 235, 245
Audits, 315
Austria, euro, 49
Automated banking machine (ABM), 235
Automated teller machines (ATMs), 51, 235, 245
Autonomous consumer expenditure, 538, 562–563
Autonomous spending, 544

Baa-Aaa bond spreads, 124–127
Balance-of-payments
  Bretton Woods system, 483
  crises, 476
  foreign exchange markets, 467–468
Balance sheets
  banks, 201–205
  channels, 622
  effects of asset markets on, 189
  money supply, 358–359
  off-balance-sheet activities, 223–226, 265
Ball, Ray, 154
BankAmericard, 234
Bank for International Settlements (BIS), 339
Bank holding companies, 232, 245
Banking Act of 1935, 345
Banking procedures, 205–208
Banking sector problems, 191
Bank Insurance Fund (BIF), 278
Bank lending channels, 621
Bank loans, 169
Banknotes, 230
Bank of America, 235, 305
Bank of Canada, 349
Bank of Credit and Commerce International (BCCI), 272
Bank of England, 272, 349–350
Bank of Japan, 360
Bank of the United States, 230
Bankruptcy, Enron Corporation, 124–127

Banks, 8. *See also* banking procedures
asymmetric information and regulation, 260–271
balance sheets, 201–205
capital, 204
central bank, 12
charters, 230
commercial banks, 34
competition, restrictions on, 269
consolidation, 245–250, 246
consolidation regulations, 264–271
credit cards, 234
debit cards, 234
decline in traditional banking, 242–243
deposits at other banks, 204
electronic banking, 234–235
Eurocurrencies, 28
European Central Bank, 49
evolution of, 232–243
failures, 260
Federal Reserve banks, 337
foreign central banks, 349
historical development of, 229–232
international banking, 253–257, 272–273, 280–284
investment, 26, 303
involved in the sale of insurance, 290
management, 208–216
mergers, 248
money center, 212
mutual savings banks, 34–35
open market purchases, 360–362
panic, 191
profitability, 242
separation of financial services, 250–252
state banking and insurance commissions, 38
structure of U.S. commercial industry, 243–245
superregional, 247
universal banking, 251
World Bank, 470
Bank supervision, 265
Barings, 225
Barnett Bank, 290
Barro, Robert J., 488, 645
Barter economies, 45
Basel 2, 266
Basel Accord, 265, 272
Basel Committee on Banking Supervision, 265
Basic gap analysis, 221
Basis points, 74
Baumol, William, 524
Baumol-Tobin analysis, 526
BCCI (Bank of Credit and Commerce International), 272
Beige book, FOMC, 344
Belarus, inflation rates, 11
Belgium, EMS, 474
Belgium, euro, 49
Bent, Bruce, 238
Bernanke, Ben, 618, 621
BIF (Bank Insurance Fund), 278
BIS (Bank for International Settlements), 339
Black, Fischer, 157
Black Monday Crash of 1987, 5, 163–164
reserve requirements, 404
Black-Scholes model, 328
Blanchard, Oliver, 597
Bliss, Robert, 137
Blue book, FOMC, 344
Board of Governors of the Federal Reserve System, 337
Bolivia, inflation, 674
Bond market, 4
equilibrium in, 89
excess supply and demand, 90
interest rates, 87–93
shifts in demand curves, 94
Bonds, 3. *See also* debt; loans
Baa-Aaa bond spreads, 124–127
Bush tax cut of 2001, 127
business cycles expansion, 95
capital market, 27
corporate long-term, 121
coupon, 63–64, 65–68
"Credit Markets" column, *Wall Street Journal*, 103–104
default-free, 121
default risk, 120
discount, 64, 68–69
effect of economic expansion for supply, 102
Enron Corporation, bankruptcy of, 124–127
Eurobond, 28
foreign, 28
income tax, 125–126
indexed, 82
junk, 124, 235–236
lemons problem, 175–180
liquidity, 96, 125
liquidity premium theory, 133
low interest rates in Japan, 103
Moody's, 123
mutual funds, 297. *See also* mutual funds
perpetual, 67
preferred habitat theory, 134
risk, 96
segmented markets theory, 132
Standard and Poor's, 123
structure of financial markets, 25–28
T-bonds, 74
Treasury bonds, 12
U.S government, 26
volatility of returns, 78
Wall Street Journal, 72–73
yields, 76. *See also* returns
Borrowing. *See* debt; direct loans; loans
Bosworth, Barry, 620
Branches, restrictions on, 244
Branson, William H., 466
Brazil
hyperinflation, 413
inflation rates, 11

Bretton Woods system, 469, 470
balance-of-payments considerations, 483
fixed exchange rates, 473–478
British-style universal banking system, 252
Brokerage firms, 304
Brokers, 26, 303. *See also* securities
audits, 315
Brown, Henry, 238
Brown, Phillip, 154
Bryant, Ralph, 618
Bubbles, 164
Budget deficits
inflation, 643
surpluses, 12
Bull markets, 5
Burns, Arthur, 424
Bush, George H. W., 348
Bush tax cut of 2001, 127
Business cycles, 9
analyzing past cycles, 598–600
bonds, 95
changes in interest rate equilibrium, 99
interest rates, 102
real business cycle theory, 596, 616
Businesses, financing, 169–172
Business finance companies, 297

Calculations. *See also* valuation
common stocks, 141–144
coupon bonds, 65–68
current yield, 70–75
discount bonds, 68–69
discounting the future, 62
fixed-payment loans, 65
money multiplier, 375–378
multiple deposit creation, 370
nominal interest rates, 79–82
present value, 65
real interest rates, 79–82
returns, 75
simple loans, 64–65
value of final goods and services, 20
yield to maturity, 64–65
Call options, 322
Calvo, Guillermo, 488
CAMELS rating, 266, 268
Campbell, John Y., 136, 137
Canada
Bank of Canada, 349
financial regulation of, 40
inflation targeting, 502
Royal Bank of Canada, 235
Cancellation of insurance, 292
Capital
accounts, 467
banks, 204
management, 215–216
Capital adequacy management, 213–215
Capital controls, 254, 478–479
Capital gains, rate of, 7
Capital markets, 27
theory of efficient capital markets, 149
Capital mobility, 445
Carte Blanche, 234
Cash
e-cash, 51. *See also* payments systems
e-finance, 8
items in process of collection, 204
M2, 52
vault, 204
Cash flow, 141
channels, 622–623
common stock, 142–143. *See also* common stock
financial crises, 190
Cashless societies, predictions of, 52
Cash management account (CMA), 305
Casualty insurance companies, 36, 288–293. *See also* financial intermediaries
CD (certificate of deposit), 203, 622
Cease and desist orders, 266
CEBA (Competitive Equality in Banking Act) of 1987, 276
Cecchetti, Stephen G., 621
Central bank, 12, 230. *See also* banks
behavior, 351–352
foreign, 349
monetary policy, 626–628
targets, 414–416
Certificate of deposit (CD), 203, 622
CFTC (Commodities Futures Trading Commission), 38, 315, 321
Changes in equilibrium, interest rates, 93–104
Channel/corridor system for setting interest rates, 406–408
Channels
asset price, 618
balance sheets, 622
bank lending, 621
credit, 625
interest rates, 617
unanticipated price level, 623
Charles Schwab, 305
Charters
banks, 230
foreign banks, 266
Chase Manhattan Bank, 249
Checkable deposits, 201
Checks, 48–49
M1, 52
Chemical Bank, 249
Chicago Board of Trade, 233, 311, 315
Chicago Mercantile Exchange, 315
Chrysler, 121
Circulation of currency, 358
Citicorp, 251, 305
Citigroup, 297
Clinton, William Jefferson, 348
Closed-end mutual funds, 299
CMA (cash management account), 305
Coins, evolution of payments system, 48–51
Coinsurance, 292
Coldwell Banker, 305

Collateral, 172
compensating balances, 219
net worth, 180
Collection, items in process of, 204
Commercial banks, 34. *See also* financial intermediaries
checking accounts, 52
separation from securities industry, 39–40
Commercial paper, 236–237
Commissions, brokerage firms, 304
Commodities Futures Trading Commission (CFTC), 38, 315, 321
Commodity money, 48
Common stock, 5. *See also* stock
calculations, 141–144
generalized dividend valuation model, 143
Gordon growth model, 143
restrictions on holdings, 40
Community banks, 249
Community Reinvestment Act (CRA) of 1977, 269
Compensating balances, 219
Competition
limits on, 40
restrictions on, 269
Competitive Equality in Banking Act (CEBA) of 1987, 276
Complete crowding out, 587
Consol, 67
Consolidation regulations, 264–271
Consolidation banks, 245–250, 246
Constant-money-growth-rate rule, 654
Consumer durable expenditure, 617
Consumer expenditure, 536, 585
autonomous consumer expenditure, 538
Consumer finance companies, 297
Consumer price index (CPI), 21, 583
Consumer protection, 269
Consumer Protection Act of 1969, 269
Consumption, 620
Contracts. *See also* loans; mortgages
debt, 172, 183
futures, 233, 311–320
moral hazards, 180–184
options. *See* options
Controllability of targets, 418–419
Conversion of currencies, 5
Cootner, Paul, 155
Corporate long-term bonds, 121
Costly state verification, 182
Cost-push inflation, 639
Costs
liabilities, 240–241
opportunity, 106
production, changes in, 594–595
transactions, 173–174, *See also* transaction costs
Counterparties, 311
Coupon bonds, 63, 65–68
Coupon rates, 64
Covenants, 185
CPI (consumer price index), 21, 583
CRA (Community Reinvestment Act) of 1977, 269
Credit. *See also* debt; loans
channels, 625
effect of less capital on markets, 216
factoring, 297
federal credit agencies, 301
management, 217–220
rationing, 220
risk, 208
secondary, 402
Credit cards, 234
Credit Control Act of 1969, 341
"Credit Markets" column, *Wall Street Journal*, 103–104
Credit unions, 35–36. *See also* financial intermediaries
NCI (National Credit Union Administration), 38
Credit view, 618, 621
Crises
balance-of-payments crises, 476
East Asian, 480
emerging market foreign exchange, 477–478
financial, 189–198
foreign exchange crisis of September 1992, 475
Great Depression, 387–390. *See also* Great Depression
international banking, 280–284
slow recovery in March 2001, 625
United States 1980s banking crisis, 273–276
Crowding out, 586, 587
Currencies, 44. *See also* money; money supply
banknotes, 230
conversion of, 5
euro, 49
Eurocurrencies, 28
Eurodollars, 254
evolution of payments system, 48–51
gold standard, 469
M1, 52
Office of the Comptroller of the Currency, 38, 231
reserve currency, 470
swaps, 328
Currency board, 492
Current accounts, 467
Current yield, 70–75

Dai-Ichi Kangyo Bank, 256
Daiwa Bank, 225
Data lag, 650
Dealers, 26, 303. *See also* securities
audits, 315
primary, 399
Dean Witter, 305
Debit cards, 51. *See also* payments systems
banks, 234
De Bondt, Werner, 157

Debt. *See also* interest rates
budget deficits and surpluses, 12
collateral, 172
contracts, 172, 183
coupon bonds, 63, 65–68
deflation, 192
discount bonds, 64, 68–69
evolution of the euro, 49
fixed-payment loans, 63, 65
Great Depression, 193
influences on financial structures, 184–188
maturity of, 26
monetizing the, 644
moral hazards, 180–184
secured debt, 172
simple loans, 52, 63, 64–65
structure of financial markets, 25–28
transaction costs, 30
unsecured debt, 172
yield to maturity, 64
Decline of reserve requirements, 406
Deductibles, insurance, 292
Default-free bonds, 121
Default risk, 120. *See also* risk
Defensive open market operations, 398
Defined-benefit plans, 294
Defined-contribution plans, 294
Deflation
debt, 192
Great Depression, 193
Delayed signaling, 506
Demand
aggregate demand function, 541
for loanable funds, 92
for money, 532–533
supply and demand. *See also* supply and demand
Demand analysis, equilibrium in, 588–600
Demand curves, 87–89
aggregate, 582–587
federal funds rate, 394
loanable funds framework, 91–92
shifts in bond markets, 94
Demand-pull inflation, 639, 641–643
Denmark
EMS, 474
euro, 49
Department of Agriculture, 315
Deposit insurance, 40
Depository institutions, 34. *See also* financial intermediaries
Depository Institutions Deregulation and Monetary Control Act (DIDMCA) of 1980, 274
Deposit rate ceilings, 238
Deposits
checkable, 201
float, 365
foreign exchange markets, 447
large-denomination time deposits, 203
money multiplier, 380
multiple deposit creation, 357, 365–371
nontransaction, 203
at other banks, 204
outflows, 208
securitization, 238
simple deposit multiplier, 369
Treasury, 365
Depreciation, 436
Derivatives, 233–234, 309
futures contracts and markets, 311–320
hedging, 309–310
interest-rate forward contracts, 310–311
interest-rate swaps, 328–330
options, 320–328
Determinants of asset demand, 85–87
Devaluation, 472
DIDMCA (Depository Institutions Deregulation and Monetary Control Act) of 1980, 274
Diners Club, 234
Direct finance, 24
Dirty float, 462
Disclosure, 39, 268
Discount bonds, 64, 68–69
expected returns, 88
Discount lending, 395–397
Discount loans, 203
money supply, 364
Discount policy, 400–403
Discount rate, 359, 420
Discounts, 420
discounting the future, 62
yields, 71
Discount window, 400
Discover Card, 305
Disintermediation, 238
Disposable income, 538
Diversification, 32
Dividends, 26
common stock, 142
generalized dividend valuation model, 143
Gordon growth model, 143
DJIA (Dow Jones Industrial Average), 5
Dollarization, 493
Dollars
effect of value fluctuations, 6
Eurodollars, 28
euro's challenge to, 471
foreign exchange markets, 447
present value, 61–62
transaction costs, 30
Domestic goods, preference for, 441
Domestic interest rates, 452
Domestic open market operations, 398
Dornbusch, Rudiger, 452
Dow Jones Industrial Average (DJIA), 5
Downward-sloping demand curves, 89
Drexel Burnham, 236
Dual banking system, 231

Duration analysis, 221
Dynamic open market operations, 398

East Asia, 480
  banking crises, 284
  financial crises, 194–198
Eastern Europe, banking crises, 282
E-cash, 51. *See also* payments systems
ECB (European Central Bank), 49, 350, 498
Econometric policy evaluation, 659
Economic expansions, 9
Economic growth, 22
  financial development and, 187–188
  inflation targeting, 508–509
  monetary policy, 412
Economies
  effect of expansion for bonds, 102
  of scale, 30, 173
  of scope, 248
ECU (European currency unit), 474
Edge Act corporation, 255
Edwards, Sebastian, 478
Effective exchange rate index, 455
Effectiveness lag, 651
Efficient market hypothesis, 149, 150–152
  evidence on the, 153–162
E-finance, 8
Eisenhower, Dwight D., 423
Electronic banking, 234–235
  regulation, 270
Electronic money (e-money), 51
Electronic payment systems, 50
Eligible paper, 420
EM (equity multiplier), 214
Emerging country financial crises, 194–198
Emerging market foreign exchange crises, 477–478
Empirical evidence
  framework for evaluation, 603–607
  Keynesian evidence, 607–610
  monetarist evidence, 611–616
Employee Retirement Income Security Act (ERISA), 295
Employment Act of 1946, 411
Employment rates, 411
EMS (European Monetary System), 471
Engineering, financial, 232
Engle, Charles, 158
Enron Corporation
  bankruptcy of, 124–127
  collapse after Arthur Andersen conviction, 178
  effect on stock market, 146–147
Equal Credit Opportunity Act of 1974, 269
Equation of exchange, 518, 583
Equilibrium
  in the bond market, 89
  changes in interest rates, 93–104
  IS curve, 552
  Keynesian cross diagram, 540
  markets, 90
Equities, 26. *See also* financial instruments; securities
  ROE, 214
Equity capital, 141, 180. *See also* common stock
Equity multiplier (EM), 214
Ericsson, 50
ERISA (Employee Retirement Income Security Act), 295
ERM (exchange rate mechanism), 474, 490
Errors of expectations, forecasting, 150
ESCB (European System of Central Banks), 350
Estrella, Arturo, 429
Euro, 28
  EMS (European Monetary System), 471
  foreign exchange markets, 457
Eurobond, 28
Eurocurrencies, 28
Eurodollars, 28, 53, 254, 255
Europe
  banking crises, 282
  currency (euro), 49
  electronic payment systems, 50
  financial regulation, 40
  price stability, 413
European Central Bank (ECB), 49, 350, 498
European currency unit (ECU), 474
European Economic Commission, 49
European Monetary System (EMS), 471
European options, 320
European System of Central Banks (ESCB), 350
Evaluation, Lucas critique of policy evaluation, 659–660
Evidence, 603. *See also* empirical evidence
  reduced-form, 604
  statistical, 613
  structural model, 604
  timing, 611
Evolution of banking system, 232–243
Excess demand, 90, 589
Excess reserves, 204
  money supply, 359
Excess supply, 90, 589
Exchange rate. *See also* Foreign exchange markets
  effects on net exports, 618
  overshooting, 454
Exchange rate mechanism (ERM), 474, 490
Exchange-rate targeting, 489–495
Exchanges, 6, 27
Exercise price, 320
Expanded-inflation effect, 113
Expansions, economic, 9
Expectations
  adaptive, 147
  relationship to liquidity premiums, 134

theory, 129
theory of rational, 147–150
Expected deposit outflows, 380
Expected price levels, 594
Expected returns. *See also* returns
discount bonds, 88
equilibrium in interest rates, 95–96
foreign exchange markets, 448–459
interest rate, 86
segmented markets theory, 132
Expenditure multiplier, 542
Expertise, financial intermediaries, 174
Exports, 537. *See also* net exports
changes in, 563–564
effect of dollar value fluctuations, 6
exchange rate effects on, 618
External financing, 170
External funds, 171–172

Face value, 63
Factoring credit, 297
Factors of production, 20
Failures, banks, 260
Fair Credit Billing Act of 1974, 269
Fallen angels, 235
Fama, Eugene F., 137, 154, 155
Farm Credit System, 301
FDIC (Federal Deposit Insurance Corporation), 38, 40, 231, 261
discount policy, 402
FDICIA (Federal Deposit Insurance Corporation Improvement Act), 278
Federal; funds rate, 393
Federal Advisory Council, 345
Federal credit agencies, 301
Federal Deposit Insurance Corporation (FDIC), 38, 40, 231, 261
discount policy, 402
Federal Deposit Insurance Corporation Improvement Act (FDICIA), 278
Federal Deposit Insurance Corporation Improvement Act of 1991, 279
Federal funds rate, reserve markets, 393–398
Federal Home Loan Bank Board, 275
Federal Home Loan Banks, 278
Federal Home Loan Bank System (FHLBS), 252
Federal Home Loan Mortgage Corporation (FHLMC), 301
Federal Insurance Contribution Act (FICA), 295
Federally sponsored agencies (FSEs), 301
Federal Market Open Committee (FMOC), 424
Federal National Mortgage Association (FNMA), 301
Federal Open Market Committee (FOMC), 337, 341, 398, 463
Federal Reserve Bank, 337, 463
Federal Reserve Bank of New York, 339
monetary policy, 424
Federal Reserve System, 12, 38
assets, 464
borrowing, 203
central bank behavior, 351–352
creation of, 231
foreign central banks, 349–351
formal structure of, 336–344
independence of, 346–349, 352–354
informal structure of, 344–346
liabilities, 464
monetary aggregates, 54
monetary policy procedures, 419–428
money measurement, 51
money supply balance sheets, 358–359
origin of, 335–336
Regulation B, 269
Regulation K, 255
Regulation Q, 40, 341
Regulation Z, 269
shifts in money supply, 108
Federal Savings and Loan Insurance Corporation (FSLIC), 265
Fed watchers, 430
FHLBS (Federal Home Loan Bank System), 252
FHLMC (Federal Home Loan Mortgage Corporation), 301
FICA (Federal Insurance Contribution Act), 295
FICO (Financing Corporation), 276
Finance companies, 37, 296–297. *See also* financial intermediaries
Financial contracts, 180–184
Financial crises, 189–198
Financial derivatives, 233–234, 309
futures contracts and markets, 311–320
hedging, 309–310
interest-rate forward contracts, 310–311
interest-rate swaps, 328–330
options, 320–328
Financial engineering, 232
Financial futures contracts, 312. *See also* futures
Financial institutions, 7–8
Financial Institutions Reform, Recovery, and Enforcement Act of 1989, 278
Financial instruments, 24. *See* securities
importance of financial intermediaries, 31
liquidity, 27
Financial intermediaries, 7, 27–28, 34–37, 177–180
deposit insurance, 40
disclosure, 39
expertise, 174

Financial intermediaries *(continued)*
function of, 29–34
importance of, 31
indirect finance, 171
insurance companies, 287–293
interest-rate swaps, 330
limits on competition, 40
restrictions on assets and activities, 39–40
restrictions on entry, 39
restrictions on interest rates, 40–41
Financial markets, 3
foreign stock market indexes, 30
function of, 23–25
internationalization of, 28–29
stability, 413–414
structure of, 25–28
Financial panic, 39
Financial scandals, effect on stock market, 146–147
Financial services industry, separation of banks, 250–252
Financial structures, 169–172
debt, 184–188
financial crises and aggregate economic activity, 189–198
lemons problem, 175–180
Financial systems, regulation of, 37–41
Financial Times-Stock Exchange 100-Share Index (London), 29
Financing Corporation (FICO), 276
Finland
electronic payment systems, 50
euro, 49
Fire and casualty insurance companies, 36. *See also* financial intermediaries; insurance companies
First Boston Corporation, 231
First National Bank of Boston, 231
Fiscal policy, 12
comparing effectiveness to monetary policy, 568–574
inflation, 636
responses to changes in, 567
Fisher, Irving, 518–519
Fisher, Lawrence, 154
Fisher effect, 99. *See also* inflation
Fisher equation, 79
Fixed exchange rates
Bretton Woods system, 473–478
regime, 470
Fixed investment, 539
Fixed-payment loans, 63, 65
Fixed-rate mortgages, 233
Flat money, 48
Float deposits, 365
FMOC (Federal Market Open Committee), 424
FNMA (Federal National Mortgage Association), 301
FOMC (Federal Open Market Committee), 337, 341, 398, 463
Forecasting
common stock valuation, 142
errors of expectations, 150
interest rates, 111–112
optimal forecast, 148
technical analysis, 155
Foreign banking systems, 253–257
Foreign bonds, 28
Foreign central banks, 349–351
Foreign exchange crisis of September 1992, 475
Foreign exchange markets, 5, 435–439
balance of payments, 467–468
equilibrium in, 446
Federal Reserve Bank, 463
intervention in, 462–467
prices, 439–442
productivity, 442
returns, 443–448
shift in expected returns, 448–459
stability, 414
Foreign exchange rates, 5
Foreign exchange risk, hedging, 319
Foreign goods, preference for, 441
Formulas. *See also* calculations
current yield, 70–75
Forward contracts, 310
Forward transactions, 436
Framework for empirical evidence, 603–607
France
EMS, 474
euro, 49
Fraud
electronic payment systems, 52
prevention of, 292
Free-rider problem, 176
Frekel, Jacob A., 466
French, Kenneth R., 157
Frictional unemployment, 411
Friedman, Benjamin, 415
Friedman, Milton, 11, 112, 387, 528–532, 608
Friedman-Meiselman measure of autonomous expenditure, 614
FSEs (federally sponsored agencies), 301
FSLIC (Federal Savings and Loan Insurance Corporation), 265
Fuji Bank, 256
Full Employment and Balanced Growth Act of 1978, 411
Fully amortized loans. *See* fixed-payment loans
Fully funded pension plans, 294
Fundamentals, market, 152
Funds, sources of, 201
Futures
contracts, 233
financial derivatives, 311–320
globalization of, 317
hedging, 314
options, 321. *See also* options

Gap analysis, 221
GDP (gross domestic product), 12, 20
central bank targets, 417–418
GDP deflator, 21–22
nominal GDP, 20–21
real GDP, 21, 583

General Electric (GE), 149
Generalized dividend valuation model, 143
General Motors Acceptance Corporation (GMAC), 297
General Theory of Employment, Interest, and Money, The, 536. *See also* Keynes, John Maynard
Generation X, 298
Germany
  EMS, 474
  euro, 49
  financial regulation, 40–41
  hyperinflation, 633
  hyperinflation after World War I, 47
  monetary targeting, 497–501
  reunification of, 490
Gertler, Mark, 174, 618, 621
Glass Stegall Act of 1933, 231, 250, 269
Globalization of futures, 317
Globex electronic trading system, 317
GMAC (General Motors Acceptance Corporation), 297
GNMA (Government National Mortgage Association), 301
Goal independence, 347
Gold standard, 469
Goldstein, Morris, 466
Goods and services. *See* aggregate output
Goodwill, 275
Gordon, David, 488
Gordon Equation 5, 147
Gordon growth model, 143
  effect of scandals, terrorism on markets, 146–147
Government
  bonds, 26
  budget constraints, 643
  debt, 12
  deposit insurance, 262. *See also* FDIC
  equilibrium in interest rates, 98
  financial intermediation, 301–302
  fiscal imbalances, 191
  increasing information, 177
  moral hazards, 182
  retirement funds, 37
  spending, 537, 563, 585
Government National Mortgage Association (GNMA), 301
Gramm-Leach-Bliley Act of 1999, 251, 290
Great Depression, 40–41. *See also* stock market crash of 1929
  bank failures, 231
  collapse of investment spending, 545
  consumers' balance sheets, 624
  debt deflation, 193
  discount policy, 402
  Federal Reserve System policy procedures, 421
  Keynesian evidence, 609
  money supply, 387
Green book, FOMC, 344
Greenspan, Alan, 343, 429–430
Gross domestic product (GDP), 12, 20
  central bank targets, 417–418
  GDP deflator, 21–22
  nominal GDP, 20–21
  real GDP, 21, 583
Growth
  constant-money-growth-rate rule, 654
  economic and financial development, 187–188
  Gordon growth model, 143
  inflation targeting, 508–509
  monetary policy, 412
  rate of money, 116
Guarantees, arbitrage, 313

Hackers, electronic payment systems, 52
Hamilton, Alexander, 229
Hedge funds, 299
Hedging, 233, 314
  financial derivatives, 309–310
  foreign exchange risk, 319
  interest-rate swaps, 329
  macro hedge, 315
  micro hedge, 315
  options, 325–326
Hicks, John, 536, 554
High employment
  effect on inflation, 639
  rates, 411
High interest rates, 4. *See also* Interest rates
Highly leveraged transaction loans, 274
High-tech sector, venture capitalists, 183
High unemployment, responses to, 650
Historical development of banks, 229–232
Home banking, 235
Hooper, Peter, 618
Hostages, 654
Hot tips, 160
Household liquidity effects, 623
Hubbard, R. Glenn, 31, 621
Huberman, Gur, 157
Huizinga, John, 609
Humphrey-Hawkins Act, 411
Hyperinflation, 47, 413, 633
Hysteresis, 597

IBFs (international banking facilities), 255
IDA (International Development Association), 470
IMF (International Monetary Fund), 470
  role of, 479–482
Implementation lag, 651
Incentive-compatible, 185
Income, 3, 45. *See also* money
  aggregate output, 20
  disposable, 538
  money markets, 113

Income *(continued)*
shifts in demand for money, 107–108
Income taxes
bonds, 125–126
Bush tax cut of 2001, 127
changes in, 563
Independence of the Federal Reserve System, 352–354
Indexes
bonds, 82
effective exchange rate, 455
foreign stock markets, 30
TIPS, 82
Indirect finance, 171
Individual retirement accounts (IRAs), 294
Indonesia, financial crises, 194–198
Industrial production, 583
Inflation, 10
anti-inflation polices, 671
Bolivia, 674
causes of, 11–12
credibility in fighting, 673
equilibrium in interest rates, 98
expanded-inflation effect, 113
growth rates, 22
hyperinflation, 47, 413
meaning of, 634–635
monetary policy, 638–650
money, 632–634
NAIRU, 429, 590
rapid, 633
rates, 11
real interest rates, 79
rise in U.S., 1960-1980, 646–650
targeting, 501–509
views of, 635–638
Volcker, Paul, 655
Inflows, controls on capital, 478–479
Information, sale of, 176–177
Information technology, 234
bank consolidation, 247
economies of scope, 248
Initial public offering (IPO), 303
Insider trading, 39
Insolvency, 192
Institutional investors, 298
Institutional money market mutual fund shares, 53
Instrument independence, 347
Instrument targets, 415
Insurance companies, 27–28, 287–293. *See also* financial intermediaries
fire and casualty insurance companies, 36
life insurance, 287–288
property and casualty insurance, 288–293
state banking and insurance commissions, 38
Interest, 238
Interest-bearing NOW accounts, 202
Interest parity condition, 445
Interest-rate forward contracts, 310–311
Interest rates, 4
business cycle expansions, 102
changes in equilibrium, 93–104, 566–568
changes in investment spending, 562
changes in net exports, 563–564
channel/corridor system for setting, 406–408
channels, 617
control of, 116
"Credit Markets" column, *Wall Street Journal*, 103–104
determinants of asset demand, 85–87
expectations theory, 129–131
forecasting, 111–112
increases in, 189
ISLM model, 557–558
Japan, 103
liquidity, 86, 104–117
liquidity preference analysis, 112–117
liquidity preference framework, 107–117
liquidity premium theory, 133
long-term bonds, 96
lowering, 117
management of, 12
management of risk, 220–223
measuring, 61–75
monetary policy, 413
money demand, 533
negative, 69
nominal interest rates, 79–82
planned investment spending, 552
preferred habitat theory, 134
real interest rates, 79–82
restrictions on, 40–41
returns, 75–79
risk, 78, 120–127, 208
segmented markets theory, 132
supply and demand, 87–93
term structure of, 127–138, 659
TIPS, 82
Wall Street Journal, 72–73
yield curve, 136
Interest-rate swaps
financial derivatives, 328–330
financial intermediaries, 330
hedging, 329
Intermediate targets, 414
Intermediate-term debt, 26
International banking, 253–257
crises, 280–284
regulations, 272–273
International banking facilities (IBFs), 255
International Development Association (IDA), 470
International financial system, evolution of, 468–478
Internationalization of financial markets, 28–29
International Monetary Fund (IMF), 470
role of, 479–482

International monetary policy, 427
exchange-rate targeting, 489–495
inflation targeting, 501–509
monetary targeting, 496–501
nominal anchor, 487–489
International policy coordination, 428
International reserves, 462
International trade, aggregate output, 548
Internet
banking, 234–235
mutual funds, 298
securities market operations, 306
Inventory investment, 539
Inverted yield curve, 127
Investment banks, 26, 303
Investments. *See also* securities; stocks
brokers, 26
changes in spending, 563
collapse during Great Depression, 545
diversification, 32
effect of stock market fluctuations, 5
expectations theory, 129–130
expected profitability of, 98
fixed, 539
hot tips, 160
intermediaries, 37
inventory, 539
life insurance, 287–288
low Japanese interest rates, 103
planned investment spending, 536
Investors, information available to, 39
IPO (initial public offering), 303
IRAs (individual retirement accounts), 294
Ireland
EMS, 474
euro, 49
IS curve, 551
factors that cause to shift, 561–564
shifts in, 578–579
ISLM model, 568, 571, 575–580
aggregate demand curve, 577–580
aggregate output, 557–558
interest rates, 557–558
Keynesian framework, 551–558
Italy
EMS, 474
euro, 49

Jackson, Andrew, 230
January effect, 155–156
Japan
banking crises, 282–284
banking systems, 252
Bank of Japan, 350
Dai-Ichi Kangyo Bank, 256
financial regulation, 40
Fuji Bank, 256
interest rates, 103
internationalization of financial markets, 29
monetary policy, 628
monetary targeting, 497
negative T-bill rates, 69
Jegadeesh, N., 158
Jensen, Michael C., 154
Jobs
effect of dollar value fluctuations, 7
unemployment rates, 9. *See also* unemployment
Junk bonds, 124, 235–236

Kandel, Shumuel, 157
Kane, Edward, 275
Kansas City Board of Trade, 315
Karjalainen, R., 155
Keogh plans, 294
Keynes, John Maynard, 521
Keynesian approach
aggregate demand curve, 582–587
comparing to Friedman, 530–532
developments in, 524–528
inflation, 636
Keynesian cross diagram, equilibrium in, 540
Keynesian evidence, 607–610
Keynesian framework, 536
aggregate output, 536–551
ISLM model, 551–558
Keynesian model, new, 665–666
Keynesian structural models, 605
Krugman, Paul, 467
Kydland, Finn, 388

Labor market, tightness of, 594
Lakonishok, J., 158
Large, complex, banking organizations (LCBOs), 248
Large-denomination time deposits, 53, 203
Latin America, banking crises, 281–282
Law of one price, 439
LCBOs (large, complex, banking organizations), 248
Leeson, Nick, 225
Legislative lag, 651
Lemons problem, 175–180
Lender of last resort, 402
Lending, specialization in, 218. *See also* banks; loans
Leverage ratio, 265
Liabilities, 24
balance sheets, 201–204
costs, 240–241
Federal Reserve System, 464
management, 208, 212–213
Life cycles, Gordon growth model, 143
Life insurance, 287–288
Life insurance companies, 36
Limits of insurance, 292
Limits on competition, 40
Liquidity, 27, 47
bonds, 96, 125
equilibrium in interest rates, 96–97
expectations theory, 134
household liquidity effects, 623

Liquidity *(continued)*
interest rates, 86, 104–117
management, 208
preference analysis, 112–117, 555
services, 31
Liquidity preference framework
interest rates, 107–117
money markets, 105–107
Liquidity preference theory, 521–524
Liquidity premium theory, 133
Lloyd's of London, 290. *See also* insurance companies
LM curve, 551, 556, 564–566
shifts in, 579–580
Load mutual funds, 299
Loanable funds framework, 91–92
Loans, 8, 169. *See also* banks; financial institutions
commitments, 219
discount, 203
finance companies. *See* finance companies
fixed-payment, 63, 65
function of financial markets, 23–25
highly leveraged transaction loans, 274
money supply, 364
moral hazards, 32–33
profitability, 205
sales, 223
securitization, 237–238
simple, 62, 63, 64–65
transaction costs, 30
yield to maturity, 64
London International Financial Futures Exchange, 317
Long positions, 309
Long-run aggregate supply curve, 590
Long-run monetary neutrality, 576
Long-term banking customer relationships, 218
Long-term bonds. *See also* bonds
capital market, 27
interest rates, 96
liquidity premium theory, 133
rates, 12
Long-Term Capital Management, 300, 427
Long-term debt, 26
Loophole mining, 237
Losses, options, 322
Louvre Accord, 428
Lowering interest rates, 117
Low interest rates, Japan, 103
Lucas, Robert, 660
Lucas critique of policy evaluation, 659–660
Luxembourg
EMS, 474
euro, 49

M1, 52
M2, 52, 426
M3, 53
Maastricht Treaty, 49
Macro hedge, 315
Malaysia, financial crises, 194–198
Managed float, 473
Managed float regime, 462
Management
assets, 208, 211–212
banks, 208–216
capital, 215–216
capital adequacy, 213–215
credit, 217–220
insurance, 290. *See also* insurance companies
of interest rate risk, 220–223
liabilities, 208, 212–213
liquidity, 208
Mankiw, N. Gregory, 136
Mann, Catherine, 618
March 2001 recession, slow recovery in, 625
Marginal propensity to consume, 538
Margin requirements, 318
Marked to market, 318
Market equilibrium, federal funds rate, 395
Market fundamentals, 152
Market interest rates, money multiplier, 380
Markets. *See also* Financial markets
Black Monday Crash of 1987, 163–164
capital, 27
debt. *See* debt
efficient market hypothesis, 149, 150–152, 153–162
equilibrium, 90
Eurodollar, 255
exchanges, 27
foreign stock market indexes, 30
futures, 311–320
lemons problem, 175–180
liquidity preference framework, 105–107
money, 27
OTC, 27
overreaction, 157
price-level effect, 108
primary markets, 26
rational expectations, 162–164
reserves, 393–398
secondary markets, 26
securities operations, 302–306
segmented markets theory, 132
setting security prices, 144–147
stability, 413–414
supply and demand, 87–93
technical analysis, 155
theory of efficient capital markets, 149
Martin Jr., William McChesney, 423
MasterCard, 51, 234
Matched sale-purchase transactions, 400
Maturity
of bonds, 129. *See also* bonds
dates, 62
of debt, 26

liquidity premium theory, 133
preferred habitat theory, 134
Maturity bucket approach, 221
Mayer, Colin, 31
McFadden Act of 1927, 244
Mean reversion, 157
Measurability of targets, 418
Medium of exchange, 45
Medium-Term Financial Strategy, 496
Meetings, FOMC, 343
Meiselman, David, 614
Meltzer, Alan, 480, 622
Member banks, Federal Reserve System, 340, 346
Membership in the Federal Reserve System, 231
Mergers, banks, 248
Mexico
financial crises, 194–198
role of IMF, 480
Micro hedge, 315
MidAmerica Commodity Exchange, 315
Milken, Michael, 236
Mishkin, Frederic S., 162, 163, 429, 452, 624
MMDAs (money market deposit accounts), 202
Modern quantity theory of money, 584
Modigliani, Franco, 621
Mondex, 52
Monetarist evidence, 611–616. *See also* Friedman, Milton
Monetarists, 608
view of inflation, 635
Monetary Affairs Division, 343
Monetary aggregates, 52, 54
targets, 426
Monetary base, 358
Monetary neutrality, 453
Monetary policy, 12, 51
central banks' conduct, 626–628
central bank targets, 414–416
comparing effectiveness to fiscal policy, 568–574
comparison of types, 666–676
Federal Reserve System procedures, 419–428
goals of, 411–414
implications for, 658
inflation, 638–650
international, 427, 482–483
Lucas critique of policy evaluation, 659–660
new classical macroeconomic model, 660–665
nominal anchor, 509–511
Phillips curve, 429–430
rational expectations revolution, 676–677
responses to changes in, 566–567
selection of targets, 416–419
stock prices, 146
Taylor rule, 428–430
transmission mechanisms, 604, 616–626
Monetary theory, 10, 517
Monetizing the debt, 644
Money, 8–13
checks, 48–49
commodity money, 48
constant-money-growth-rate rule, 654
creation in foreign countries, 644
demand for, 532–533
electronic money (e-money), 51
evolution of payments system, 48–51
Federal Reserve System, 12
flat money, 48
functions of, 45–48
growth rate, 116
inflation, 632–634
liquidity, 47, 112–117
management of interest rates, 12
meaning of, 44–45
measuring, 51–55
modern quantity theory of money, 584
printing, 644
Quantity Theory of Money: A Restatement, The, 528–532
real money balances, 523
reliability of data, 55–56
seignorage, 493
transaction costs, 29–30
velocity of, 518, 520–521, 582
Money center banks, 212
Money market deposit accounts (MMDAs), 202
Money market mutual funds, 37, 52–53, 238–239, 299. *See also* financial intermediaries
Money markets, 27
income effect, 113
liquidity preference framework, 105–107
M1, 52
price-level effect, 108
shifts in demand for, 107–108
Money multiplier, 374
money supply, 375–378
variables, 378–381
Money supply, 8
aggregate output, 567
determinants of, 374, 381–383
discount loans, 364
discount policy, 400403
excess reserves, 359
Federal Reserve balance sheets, 358–359
inflation, 11
money multiplier, 375–378
movements in, 384
multiple deposit creation, 357, 365–371
process, 357, 383–390
required reserves, 359
reserve requirements, 359, 403–408
unemployment, 453
Money theory, 517–520
Monitoring
credit risk, 218–219
principal-agent problems, 182
Moody's bond ratings, 123

Moral hazards, 32–33, 174–175
debt, 180–184
influences in debt markets, 184–188
regulations, 262–263
Morgan, J. P., 231, 251
Morris, Charles S., 158
Mortgages. *See also* federal credit agencies; loans
ARMs (adjustable-rate mortgages), 233
fixed-rate, 233
structure of financial markets, 25–28
Multiple deposit creation, 357, 365–371
Multipliers, 542
Muncie, Indiana, 607
Municipal bond income tax considerations, 125–126
Muth, John, 147, 150
Mutual funds, 37, 173, 297–301. *See also* financial intermediaries
efficient market hypothesis, 153
money market mutual funds, 238–239
Mutual savings banks, 34–35. *See also* financial intermediaries

NAIRU (nonaccelerarting inflation rate of unemployment), 429, 590
NASDAQ (National Association of Securities Dealers), 26, 306
National Bank Act of 1863, 231
National Credit Union Administration (NCUA), 38, 253
National Credit Union Share Insurance Fund (NCUSIF), 40, 253
Nationwide banking, 245–250
Natural rate level of output, 575, 590
Natural rate of unemployment, 412, 590
NCUA (National Credit Union Administration), 38, 253
NCUSIF (National Credit Union Share Insurance Fund), 40, 253
Negative T-bill rates, 69
Negotiable order of withdrawal (NOW), 202, 425
Net exports, 537
changes in, 563–564
exchange rate effects on, 618
interest rates, 552–554
Netherlands
EMS, 474
euro, 49
Net worth, collateral, 180
New classical macroeconomic model, 660–665
New York Futures Exchange, 315
New York Stock Exchange (NYSE), 26, 27, 305
New Zealand, inflation targeting, 501–502
Nikkei 225 Average (Tokyo), 29
Nokia, 50
No-load mutual funds, 299
Nominal anchor
monetary policy, 509–511
role of, 487–489
Nominal GDP, 20–21
Nominal interest rates, 79–82
Nonaccelerarting inflation rate of unemployment (NAIRU), 590
Nonactivist policy debate, 650–655
Nonactivists, 592
Nonbank finance, 287. *See also* financial intermediaries
finance companies, 296–297
government financial intermediation, 301–302
insurance companies, 287–293
mutual funds, 297–301
pension funds, 294–296
securities market operations, 302–306
Nonbank loans, 169
Nonborrowed monetary base, 381–382
Non-interest bearing checking accounts, 202
Nonpublic banks, open market purchases from, 360–362
Nontransaction deposits, 203
Notional principal, 328
NOW (negotiable order of withdrawal), 202, 425
NYSE (New York Stock Exchange), 26, 27, 305

Obstfeld, Maurice, 467
OCC (Office of the Comptroller of the Currency), 38, 231, 290
OECD (Organization for Electronic Cooperation and Development), 265
Off-balance-sheet activities, 223–226, 265
Office of the Comptroller of the Currency (OCC), 38, 231, 290
Office of Thrift Supervision (OTS), 38, 252, 278
Official reserve transaction balance, 468
One-period valuation model, 142
Open-end mutual funds, 298
Open market
discovery of, 420
federal funds rate, 395
operations, 340, 398–400
purchases form nonpublic banks, 360–362
purchases from banks, 360
sales, 362–364
Operating targets, 415
Opportunity cost, 106
Optimal forecast, 148
Options
call, 322
financial derivatives, 320–328
hedging, 325–326
premiums, 326
SEC, 321

Organization for Electronic Cooperation and Development (OECD), 265
OTC (over-the-counter) markets, 27
OTS (Office of Thrift Supervision), 38, 252, 278
Outflows
  controls on capital, 478–479
  deposits, 208
  money multiplier, 380
Output
  aggregate, 583
  inflation targeting, 507–508
  international trade, 548
  Keynesian framework, 536–551
  natural rate level of, 575
  short-run, 668
Overreaction, markets, 157
Overshooting exchange rates, 454
Over-the-counter (OTC) markets, 27

Panic
  banks, 191
  Federal Reserve System policy procedures, 421
  Great Depression, 387
Panic, financial, 39
Paper currency, 48. *See also* currencies; money
Partial crowding out, 587
Par value, 63
Payments system, evolution of, 48–51
Payoff methods, 261
PCE (personal consumption expenditures), 21
Pension Benefit Guarantee Corporation ("Penny Benny"), 295
Pension funds, 27–28, 37, 294–296. *See also* financial intermediaries
Perfect substitutes, 129
Permanent life insurance, 288
Perpetual bonds, 67
Personal consumption expenditures (PCE), 21
Person-to-Person Finance Company, 297
Pesando, James, 162
Philippines, financial crises, 194–198
Phillips curve, 429–430
Plain vanilla swaps, 328. *See also* interest-rate swaps
Planned investment spending, 536, 552, 585
Plaza Agreement, 428, 483
Plosser, Charles, 597
Policyholders, 287. *See also* insurance companies
Policy ineffectiveness proposition, 663
Political business cycle, 353
Portfolios, 32
Portugal
  EMS, 474
  euro, 49
Positive risk premiums, 123
PPI (producer price index), 583
PPP (theory of purchasing power parity), 439
Predictability of target goals, 419
Predictions. *See* forecasting
Preference analysis, liquidity, 112–117
Preferred habitat theory, 134
Prell, Michael J., 162
Premiums
  expectations theory, 134
  insurance, 288. *See also* insurance companies
  liquidity premium theory, 133
  options, 320, 326
  risk, 121, 123
  risk-based premiums, 291–292
Prescott, Edward, 488
Present value, 61–62
  coupon bonds, 65–68
  discount bonds, 68–69
  fixed-payment loans, 65
  simple loans, 64–65
Price-level effect, 108, 113
Prices. *See also* GDP
  aggregate output, 583
  aggregate price level, 21–22
  arbitrage, 313
  asset price channels, 618
  availability of public information, 154
  Black Monday Crash of 1987, 163–164
  "Credit Markets" column, *Wall Street Journal*, 103–104
  effect of new information, 158
  exercise, 320
  expected price levels, 594
  foreign exchange markets, 439–442
  goods and services (aggregate price level), 10
  monetary policy, 146, 412–413
  random-walk behavior of stock prices, 154
  relative price levels, 441
  setting security prices, 144–147
  settlement, 318
  short-run output, 668
  stock market fluctuations, 5
  strike, 320
  technical analysis, 155
  unanticipated price level channel, 623
Primary dealers, 399
Primary markets, 26
Principal, 62
Principal-agent problem, 181
  monitoring, 182
  S&Ls, 277
Printing money, 644
Privacy
  electronic banking, 270
  electronic payment systems, 52
Private pension plans, 295
Private production, 176–177
Procyclical monetary policy, 423
Producer price index (PPI), 583
Production costs, changes in, 594–595

Productivity, foreign exchange markets, 442
Profitability
  arbitrage, 313
  banks, 242
  equilibrium in interest rates, 98
  loans, 205
  options, 322
  unexploited profit opportunity, 152
Property and casualty insurance, 288–293
Prudential supervision, 265
Public pension funds, 295
Purchase and assumption methods, 261
Put (sell) options, 327

Quantity of assets, 85–87
Quantity of loans demanded, 92
Quantity theory of money, 519
Quantity Theory of Money: A Restatement, The, 528–532

Radford, R. A., 46
Ramey, Valerie, 622
Random-walk behavior of stock prices, 154
Rapid inflation, 633
Rate of capital gain, 76
Rate of return, 75. *See also* returns
Ratings, bonds. *See* bonds
Rational bubbles, 164
Rational expectations
  evidence on markets, 162–164
  theory of, 147–150
  revolution, 676–677
Rationing credit, 220
Ratios, money multiplier, 378–381
Reagan, Ronald, 276, 348, 675–676
Real bills doctrine, 420
Real business cycle theory, 596, 616
Real GDP, 21, 583
Real interest rates, 79–82
Real money balances, 523
Real terms, aggregate output, 575
Recessions, 9
  rate of money growth, 9–10
  slow recovery in March 2001, 625
Recognition lag, 650
Rediscounting, 420
Redlining, 269
Reduced-form evidence, 604
RefCorp (Resolution Funding Corporation), 278
Regulation K, 255
Regulation Q, 40, 238, 341
Regulations
  adverse selection, 263
  asymmetric information and banking, 260–271
  checking accounts, 52
  consolidation, 264–271
  of financial systems, 37–41, 172
  increasing information, 177
  international banking, 272–273
  life insurance companies, 287–288
  moral hazards, 182, 262–263
  securitization, 237–238
  thrift industry, 252–257
Regulation Z, 269
Regulatory arbitrage, 265
Regulatory forbearance, 275
Reinsurance, 290
Relative price levels, 441
Repurchase agreements, 52, 400
Required reserves
  ratio, 359
  ration, 204
Research staff, Federal Reserve System, 342
Reserve currency, 470
Reserve repo, 400
Reserve requirements, 422
Reserves, 204
  federal funds rate, 397
  international, 462
  liquidity management, 208
  markets, 393–398
  money supply, 359
  official reserve transaction balance, 468
  requirements, 403–408
  securitization, 238
  sweep accounts, 239–240
Residual claimant, 141
Resolution Funding Corporation (RefCorp), 278
Resolution Trust Corporation (RTC), 278
Restrictions on assets and activities, 39–40
Restrictions
  on entry, 39
  on interest rates, 40–41
Restrictive covenants, 185
Restrictive provisions, insurance companies, 292
Retirement funds, government, 37
Retirement plans. *See* pension funds
Return on assets (ROA), 214
Return on equity (ROE), 214
Returns. *See also* expected returns
  foreign exchange markets, 443–448, 448–459
  interest rates, 75–79
  segmented markets theory, 132
  volatility of, 78
Revaluation, 472
Reverse causation, 606
Ricardian equivalence, 645
Ricardo, David, 645
Riegle-Neal Interstate Banking and Branching Efficiency Act of 1994, 251
Right axis, 88
Risk, 31
  arbitrage, 313
  assets, 85–87. *See also* assets
  bonds, 96
  credit, 208
  determining desired level of, 79
  equilibrium in interest rates, 96
  hedging foreign exchange, 319
  interest rates, 78, 86, 208
  management of credit, 217–220

management of interest rates, 220–223
moral hazards, 33, 186. *See also* moral hazards
premiums, 121, 123
Risk-based premiums, 291–292
Risk structure of interest rates, 120–127
ROA (return on assets), 214
ROE (return on equity), 214
Roll, Richard, 154, 157
Romania, inflation rates, 11
Romer, Christina, 615
Romer, David, 615
Royal Bank of Canada, 235
RTC (Resolution Trust Corporation), 278
Russia
banking crises, 282
hyperinflation, 413
inflation rates, 11

SAIF (Savings Association Insurance Fund), 278
Sales
of information, 176–177
loans, 223
Sales finance companies, 297
Sargent, Thomas, 660
Savings accounts, 205
Savings and loan associations (S&Ls), 34–35, 252. *See also* financial intermediaries
FSLIC, 275
political economy of, 276–278
Savings Association Insurance Fund (SAIF), 40, 278
Savings deposits, M1, 52
Scales, economies of, 30
Scandals
agent problem, 225
slow recoveries in, 625
effect on stock market, 146–147
Scandinavia
banking crises, 280–281
electronic payment systems, 50
Schoenholtz, Kermit L., 136
Schwartz, Anna, 387, 608
Screening, 217, 291
SDRs (special drawing rights), 474
Sears Roebuck Acceptance Corporation, 297
Seasoned issues, 303
Secondary credit, 402
Secondary markets, 26
Secondary reserves, 204
Second Bank of the United States, 230
SEC (Securities and Exchange Commission), 38, 39, 303, 321
Secured debt, 172
Securities, 3, 24
bank holdings, 204–205
brokers, 26
capital market, 27
dealers, 26
efficient market hypothesis, 150–152
hot tips, 160
importance of financial intermediaries, 31
lemons problem, 175–180
market operations, 302–306
mutual funds, 297. *See also* mutual funds
separation from commercial banks, 39–40
setting prices, 144–147
Treasury, 127
underwriting, 26
Securities Amendment Act of 1975, 306
Securities and Exchange Commission (SEC), 38, 39, 303, 321
Securitization, 237–238
Security of electronic payment systems, 52
Segmented markets theory, 132
Seignorage, 493
Self-correcting mechanism, 591
Sell (put) options, 327
Semistrong-form efficiency, 155
Separation of commercial banking and securities industry, 39–40
September 11, 2001
property and casualty insurance, 290
reserve requirements, 404
effect on stock market, 146–147
Settlement price, futures, 318
Shearson, Loeb, Rhodes, 305
Shifts in demand curves, 94
Shiller, Robert J., 136, 157
Short positions, 309
Short-run output, 668
Short-term debt, 26
Short-term interest rates, 136
Siebert, Horst, 498
Simple deposit multiplier, 369
Simple interest rates, 62. *See also* interest rates
Simple loans, 62, 63, 64–65
Simple model, multiple deposit creation, 371
Slow recovery in March 2001 recession, 625
S&Ls (savings and loan associations), 34–35, 252, 275. *See also* financial intermediaries
FSLIC, 275
political economy of, 276–278
Small-denomination time deposits, 52
Small-firm effect, 156
Smart cards, 51. *See also* payments systems
Social Security, 295, 296
Sources of funds, 201
South Korea, financial crises, 194–198
Spain
EMS, 474
euro, 49
Special drawing rights (SDRs), 474
Specialists, dealer-brokers, 306
Specialization in lending, 218

Spot exchange rate, 436
Spot transactions, 436
Spreads, Baa-Aaa bonds, 124–127
Stabilization policy, 670
Standard and Poor's bond ratings, 123
Standardized gap analysis, 221
State banking and insurance commissions, 38
State banks, 231. *See also* banks
State Farm, 288
Statistical evidence, 613
Sterilized foreign exchange intervention, 465
Stockholders, 141
Stock market crash of 1929, 39
Stock market crash of 1987, 5
Stock markets, 5–6, 29
Stocks, 5
  availability of public information, 154
  Black Monday Crash of 1987, 163–164
  brokers, 26
  capital market, 27
  effect of new information, 158
  effect of scandals, terrorism on prices, 146–147
  efficient market hypothesis, 150–152, 153–162
  lemons problem, 175–180
  monetary policy effect on prices, 146
  mutual funds, 297. *See also* mutual funds
  options, 321
  random-walk behavior of stock prices, 154
  restrictions on holdings, 40
  supply and demand, 93
  technical analysis, 155
  theory of rational expectations, 147–150
Stored-value cards, 51. *See also* payments systems
Store of value, 47
Strike price, 320
Strong-form efficiency, 155
Structural model evidence, 603
Sumitomo, 225
Summers, Lawrence H., 136, 158, 452, 597
Superregional banks, 247
Supply and demand
  "Credit Markets" column, *Wall Street Journal*, 103–104
  effect of expanding economies, 102
  excess, 90
  interest rates, 87–93
  money markets, 105–107
  price-level effect, 108
  shifts in money markets, 107–108
  shifts in supply of bonds, 97
  stocks, 93
Supply curve, 90, 394–395
Supply of loanable funds, 92
Supply-side phenomena, inflation, 638
Swaps, 328. *See also* interest-rate swaps
Sweden
  electronic payment systems, 50
  euro, 49
Sweep accounts, 239–240
Switzerland, 497–501

T-account, 205
Taliban, the, 146–147
Targets
  central bank, 414–416
  exchange-rate targeting, 489–495
  inflation, 501–509, 639
  monetary aggregate, 426
  monetary policy, 496–501
  selection of, 416–419
Tariffs, 441
Taxes. *See also* income tax
  changes in, 563
  effect of budget deficits and surpluses, 12
Taylor, John B., 429, 618
Taylor rule, 428–430
T-bills, 69
T-bonds, 74
Temporary life insurance, 288
Term life insurance, 288
Terms of loanable funds, 92
Term structure of interest rates, 127–138
Terrorism, effect on stock market, 146–147, 654
Thailand, financial crises, 194–198
Thatcher, Margaret, 496
Theory of asset demand, 87. *See also* assets
Theory of efficient capital markets, 149
Theory of purchasing power parity (PPP), 439
Theory of rational expectations, 147–150
Thrift institutions, 34. *See also* financial intermediaries
  Office of Thrift Supervision, 38
  regulations, 252–257
Time-consistency problem, 488
Timing evidence, 611
Timmerman, Sullivan A., 155
TIPS (Treasury Inflation Protection Securities), 82
Titman, Sheridan, 158
Tobin, James, 524, 620
Tobin's q Theory, 620
Tokyo Stock Exchange, 317
Tombstones, securities advertisements, 304
Trade
  aggregate output, 548
  balance, 467
  barriers, 441
Trading Room Automated Processing (TRAPS), 399
Traditional monetary policy models, 667
Transaction costs, 29–30, 32, 45

Transactions
adverse selection and moral hazards, 174–175
costs, 173–174
evolution of payments system, 48–51
forward, 436
highly leveraged transaction loans, 274
Keynesian approach, developments in, 524–528
matched sale-purchase, 400
movement of, 521
nontransaction deposits, 203
official reserve transaction balance, 468
spot, 436
Transmission mechanisms, 604, 616–626
TRAPS (Trading Room Automated Processing), 399
Traveler's checks, 52
Travelers Group, 305
Treasury bonds, 12
Bush tax cut of 2001, 127
income tax considerations, 125–126
yield curve, 127–128
Treasury deposits, 365
Treasury Inflation Protection Securities (TIPS), 82

Unanticipated price level channel, 623
Uncertainty, increases in, 189
Underfunded pension plans, 294
Underwriters, 303
Underwriting securities, 26
Unemployment
aggregate output, 583
money supply, 453
NAIRU, 429, 590
natural rate of, 412, 590
rates, 9, 411, 583
responses to, 650
Unexploited profit opportunity, 152
United Kingdom
Bank of England, 349–350
EMS, 474
gold standard, 469
inflation, 11, 502–506
United States
electronic payment systems, 50
financial crises in, 191–192
financial regulation, 40–41
foreign banks in, 256
inflation rates, 11
nationwide banking, 245–250
1980s banking crisis, 273–276
structure of commercial banking industry, 243–245
TIPS, 82
Unit of accounts, 46
Universal banking, 251
Universal life insurance, 288
Unsecured debt, 172
Unsterilized foreign exchange intervention, 464
U.S. Government bond yield curve, 138

Valuation
devaluation, 472
generalized dividend valuation model, 143
Gordon growth model, 143
one-period valuation model, 142. *See also* common stock
revaluation, 472
Value Line Survey, 157
Values
face, 63
money. *See* money
par, 63
present value, 61–62
store of, 47
Vanguard Group, 79
Variable life insurance, 288
Variables, 378–381
Vault cash, 204
Velocity of money, 518, 520–521, 582
Venture capital firms, 182, 183
Vertical axes, 88
Virtual banks, 236. *See also* banks; electronic banking
Visa, 51, 234
Volatile exchange rates, 455
Volatility of bond returns, 78
Volcker, Paul, 425, 430, 655

Wage push, 594
Wage-setting processes, 653
Wall Street Journal, 72, 142
"Credit Markets" column, 103–104
Currency Trading, 458
Forecasting Survey for 2003, 111–112
foreign exchange rates, 437
futures, 312
options, 321, 326
stock prices, 159
underwriters, 303
War of 1812, 230
Weak-form efficiency, 155
Wealth, 45. *See also* money
effects of, 620–621
equilibrium in interest rates, 95
interest rates, 86
Wells Fargo, 235
White, H., 155
Whole life insurance, 288
World Bank, 470
World War I, hyperinflation after, 47
World War II, 422–423
World Wide Web (WWW), 247. *See also* electronic banking; information technology; Internet banking

Yield
current, 70–75
discounts, 71
returns. *See* returns
Yield curves, 127
1980-2003, 137–138
liquidity premiums, 135

Yield curves *(continued)*
short-term interest rates, 136
U.S. Government bonds, 138
Yield to maturity, 64
Zero-coupon bonds. *See* discount bonds
Zero impact, expanded-inflation effect, 113–114
Zombie S&Ls, 275–276. *See also* S&Ls