Introduction to Credit Risk Modeling

Second Edition

CHAPMAN & HALL/CRC
Financial Mathematics Series

Aims and scope:
The field of financial mathematics forms an ever-expanding slice of the financial sector. This series aims to capture new developments and summarize what is known over the whole spectrum of this field. It will include a broad range of textbooks, reference works and handbooks that are meant to appeal to both academics and practitioners. The inclusion of numerical code and concrete real-world examples is highly encouraged.

Series Editors

M.A.H. Dempster
Centre for Financial Research
Department of Pure
Mathematics and Statistics
University of Cambridge

Dilip B. Madan
Robert H. Smith School
of Business
University of Maryland

Rama Cont
Center for Financial
Engineering
Columbia University
New York

Published Titles

American-Style Derivatives; Valuation and Computation, *Jerome Detemple*
Analysis, Geometry, and Modeling in Finance: Advanced Methods in Option Pricing, *Pierre Henry-Labordère*
Credit Risk: Models, Derivatives, and Management, *Niklas Wagner*
Engineering BGM, *Alan Brace*
Financial Modelling with Jump Processes, *Rama Cont and Peter Tankov*
Interest Rate Modeling: Theory and Practice, *Lixin Wu*
Introduction to Credit Risk Modeling, Second Edition, *Christian Bluhm, Ludger Overbeck, and Christoph Wagner*
Introduction to Stochastic Calculus Applied to Finance, Second Edition, *Damien Lamberton and Bernard Lapeyre*
Monte Carlo Methods and Models in Finance and Insurance, *Ralf Korn, Elke Korn, and Gerald Kroisandt*
Numerical Methods for Finance, *John A. D. Appleby, David C. Edelman, and John J. H. Miller*
Portfolio Optimization and Performance Analysis, *Jean-Luc Prigent*
Quantitative Fund Management, *M. A. H. Dempster, Georg Pflug, and Gautam Mitra*
Robust Libor Modelling and Pricing of Derivative Products, *John Schoenmakers*
Stochastic Financial Models, *Douglas Kennedy*
Structured Credit Portfolio Analysis, Baskets & CDOs, *Christian Bluhm and Ludger Overbeck*
Understanding Risk: The Theory and Practice of Financial Risk Management, *David Murphy*
Unravelling the Credit Crunch, *David Murphy*

Proposals for the series should be submitted to one of the series editors above or directly to:
CRC Press, Taylor & Francis Group
4th, Floor, Albert House
1-4 Singer Street
London EC2A 4BQ
UK

Chapman & Hall/CRC FINANCIAL MATHEMATICS SERIES

Introduction to Credit Risk Modeling

Second Edition

Christian Bluhm
Ludger Overbeck
Christoph Wagner

CRC Press
Taylor & Francis Group
Boca Raton London New York

CRC Press is an imprint of the
Taylor & Francis Group, an **informa** business
A CHAPMAN & HALL BOOK

CRC Press
Taylor & Francis Group
6000 Broken Sound Parkway NW, Suite 300
Boca Raton, FL 33487-2742

© 2010 by Taylor & Francis Group, LLC
CRC Press is an imprint of Taylor & Francis Group, an Informa business

No claim to original U.S. Government works
Version Date: 20140514

International Standard Book Number-13: 978-1-58488-993-9 (eBook - PDF)

This book contains information obtained from authentic and highly regarded sources. Reasonable efforts have been made to publish reliable data and information, but the author and publisher cannot assume responsibility for the validity of all materials or the consequences of their use. The authors and publishers have attempted to trace the copyright holders of all material reproduced in this publication and apologize to copyright holders if permission to publish in this form has not been obtained. If any copyright material has not been acknowledged please write and let us know so we may rectify in any future reprint.

Except as permitted under U.S. Copyright Law, no part of this book may be reprinted, reproduced, transmitted, or utilized in any form by any electronic, mechanical, or other means, now known or hereafter invented, including photocopying, microfilming, and recording, or in any information storage or retrieval system, without written permission from the publishers.

For permission to photocopy or use material electronically from this work, please access www.copyright.com (http://www.copyright.com/) or contact the Copyright Clearance Center, Inc. (CCC), 222 Rosewood Drive, Danvers, MA 01923, 978-750-8400. CCC is a not-for-profit organization that provides licenses and registration for a variety of users. For organizations that have been granted a photocopy license by the CCC, a separate system of payment has been arranged.

Trademark Notice: Product or corporate names may be trademarks or registered trademarks, and are used only for identification and explanation without intent to infringe.

Visit the Taylor & Francis Web site at
http://www.taylorandfrancis.com

and the CRC Press Web site at
http://www.crcpress.com

Contents

Preface to Second Edition

The first edition of this book appeared eight years ago. Since then the banking industry experienced a lot of change and challenges. The most recent financial crisis which started around May 2007 and lasted in its core period until early 2009 gave rise to a lot of scepticism about whether credit risk models are appropriate to capture the true nature of risks inherent in credit portfolios in general and structured credit products in particular. In a recent article we discuss common credit risk modeling approaches in light of the most recent crisis and invite readers to participate in the discussion; see [26].

A key observation in a discussion like the one in [26] is that the universe of available models and tools is sufficiently rich for doing a good job even in a severe crisis scenario as banks recently experienced it. What seems to be more critical is an appropriate *model choice*, *parameterization of models*, dealing with *uncertainties*, e.g., based on insufficient data, and *communication* of model outcomes to decision makers and executive senior management. These are the four main areas of challenge where we think that a lot of work and rethinking needs to be done in a "post-crisis" reflection of credit risk models.

In the first edition of this book we focused on the description of common mathematical approaches to model credit portfolios. We did not change this philosophy for the second edition. Therefore, we left large parts of the book unchanged in its core message but supplemented the exposition with new model developments and with details we omitted in the first edition. The aforementioned four areas of challenge in a "post-crisis" reflection of credit risk models would justify another separate exposition in book form and it should be clear that a second edition of an existing book cannot be an appropriate substitute for that. However, we included a few comments in the text where appropriate.

A brief outline of the track record of changes and updates is as follows.

Chapter 1 is updated in line with developments over the last eight years. We also included more details in the case of some topics. For instance, the sections on the probability of default (PD), exposure-at-default (EAD) and loss-given-default (LGD) are still brief compared to what could be said but have been extended in comparison to the first edition of this book. The brief section on regulatory capital has also been changed and updated.

Chapter 2 has been updated where necessary but is left unchanged in large parts. What is new is a longer section on techniques for the generation of loss distributions. Because such techniques rely on a bunch of tools from probability theory we included for the convenience of the reader an introductory section on "prerequisites from probability theory" in order to keep the survey on calculation and simulation techniques self-contained as much as possible. Readers will also benefit from the probability toolkit in Chapter 4 on CreditRisk$^+$ where generating functions play a major role.

Chapter 3 is left unchanged except for the correction of typos.

Chapter 4 is enhanced by a new section on technical details regarding the calculation of the loss distribution in CreditRisk$^+$.

Chapter 5 has been updated and some new developments are included now. For instance, spectral risk measures and an axiomatic approach to capital allocation are introduced.

Chapter 6 is left unchanged except for a brief section on term structures of default probabilities based on non-homogeneous Markov chains. This new approach has been included because it fits models far better than the time-homogeneous Markov chain approach.

Chapter 7 is left unchanged except for the correction of typos.

Chapter 8 is updated to some extent as well as enhanced. We kept the presentation of cash flow structures because the basic principles and structures remain unchanged. During the most recent crisis, structured products came under pressure and markets dried up. However, we are convinced that securitization as well as portfolio structuring remain a core competence and major tool of banks' financial engineering departments. Therefore, we extended Chapter 8 by a section on multi-period models and a brief section on recent developments. However, we keep the presentation short because we have dedicated a separate book [24] to the topic of CDO modeling and it would not have made much sense to carry sections from [24] over to this book.

We need to make a disclaimer regarding data in examples. In all cases we used data for illustrative purposes only. Therefore, we decided not to re-run all examples with more recent data. Interested readers can find in [24] many more examples with up-to-date data.

Altogether we can say that doing math in the context of credit risk modeling still means a lot of personal satisfaction to us. Credit risk in particular and finance in general are great fields to apply mathematical concepts to real life situations.

However, when doing this one should never forget that senior management, regulators, investors, etc. rely in their decision making on models and valuation outcomes. The most recent crisis showed that it is important to appropriately communicate model outcomes and to make sure that the variation in results is made transparent to decision makers in ways that they can understand.

A last remark we want to make concerns model choice and model risk. It is not recommended to use one and only one model for a particular problem. Instead, we recommend using various models to shed some light on different aspects of the true nature of a credit risk problem. In this way, the problem is viewed from different angles. The most recent crisis showed that more modeling and more analysis are superior to just one model relying on various simplifying assumptions. Regulators, for instance, talk a lot about stress testing. From our perspective, stress testing should already be part of the model. What people consider as stress, for example a market scenario where banks lose several billion euros, is historically seen not as stress but as a 10-year (give or take) regular event. It should be treated as such.

Munich and Giessen, April 2010

Christian Bluhm, Ludger Overbeck, Christoph Wagner

Acknowledgements

Christian Bluhm would like to thank his wife Tabea, and his children Sarah and Noa for their continuous support. Various book projects in the last years consumed a significant amount of time and it is just great to have a patient and understanding family. Ludger Overbeck is grateful to his wife Bettina and his children Leonard, Daniel, Clara, and Benjamin for their ongoing support.

We had great feedback, support, and comments on the first edition of this book by many colleagues, friends and readers from all over the world. We are grateful they let us know of typos, mistakes, and errors and we are happy about input on how the exposition can be improved. We hope that readers will continue to let us know if they find errors or unclear passages in the book and we apologize for still undiscovered shortfalls of the manuscript. Feedback and input can be sent to the contact email addresses at

http://www.christian-bluhm.net
http://www.uni-giessen.de/∼gc1156.

Disclaimer

This book reflects the view of the authors and not the opinion of their current or former employers. The content of the book has been written for educational purposes only. The authors are by no means liable for any damage arising from any application of the theory, examples or data presented in this book.

Preface

In banking, especially in risk management, portfolio management, and structured finance, solid quantitative know-how becomes more and more important. We had a two-fold intention when writing this book.

First, this book is designed to help mathematicians and physicists leaving the academic world and starting a profession as risk or portfolio managers to get quick access to the world of credit risk management. Second, our book is aimed at being helpful to risk managers looking for a more quantitative approach to credit risk.

Following this intention on one side, our book is written in a *lecture notes* style very much reflecting the keyword “introduction” already used in the title of the book. We consequently avoid elaborating on technical details not really necessary for understanding the underlying idea. On the other side, we kept the presentation mathematically precise and included some proofs as well as many references for readers interested in diving deeper into the mathematical theory of credit risk management.

The main focus of the text is on *portfolio* rather than *single obligor* risk. Consequently, correlations and factors play a major role. Moreover, most of the theory in many aspects is based on probability theory. We, therefore, recommend that the reader consult some standard text on this topic before going through the material presented in this book. Nevertheless, we tried to keep it as self-contained as possible.

Summarizing our motivation for writing an introductory text on credit risk management one could say that *we tried to write the book we would have liked to read before starting a profession in risk management* some years ago.

Munich and Frankfurt, August 2002

Christian Bluhm, Ludger Overbeck, Christoph Wagner

About the Authors

Christian Bluhm worked for Deutsche Bank in their risk methodology department, for McKinsey as an associate in their risk management practice, for HypoVereinsbank's Group Credit Portfolio Management in Munich where he headed the Structured Finance Analytics team which was responsible for the evaluation of structured credit assets (short and long side) and for Credit Suisse where he worked as a Managing Director heading the Credit Portfolio Management unit in Zurich. At the end of 2009 Christian left Credit Suisse to enjoy a sabbatical break with much time for his family.

Christian holds a Ph.D. degree in mathematics from the University of Erlangen-Nürnberg and was a one-year post-doctoral member of the mathematics department of Cornell University, Ithaca, New York. He has authored several papers and research articles on harmonic and fractal analysis of random measures, stochastic processes, and random fields. Since he started to work in risk management more than 10 years ago, he continuously publishes in this field and regularly speaks at risk management conferences and workshops. He also lectures at universities.

Ludger Overbeck worked for Deutsche Bundesbank in their supervision department and headed the Research and Development team in the Risk Analytics and Instruments department of Deutsche Bank's credit risk management unit until 2003. Since then he has been Professor for Probability Theory and Quantitative Finance and Risk Management at the Institute of Mathematics of the University of Giessen. Besides this he is serving in different positions in credit portfolio management and securitizaton units for HypoVereinsbank/UniCredit, DZ-Bank, and recently for Commerzbank. His main responsibilities include quantitative credit risk models, bank portfolio steering, valuation of structured credit products, economic capital and capital allocation.

Ludger holds a Ph.D. in probability theory from the University of Bonn. After two post-doctoral years in Paris and Berkeley, from 1995 to 1996, he finished his Habilitation in applied mathematics during his affiliation with the Bundesbank. From the University of Frankfurt, Ludger received a Habilitation in Business and Economics in 2001. He has published papers in mathematical and statistical journals as well as in journals on finance and economics. Ludger frequently speaks at academic and practitioner conferences.

Christoph Wagner worked for Deutsche Bank in their risk methodology department, the risk methodology team of Allianz Group Center in Munich, UniCredit/HypoVereinsbank in Munich, and Allianz Risk Transfer in Zurich. His main responsibilities in his positions were (and continue to be) credit risk and insurance-linked securities, securitizations and alternative risk transfer as well as the valuation of all kinds of structured credit products.

Christoph holds a Ph.D. in statistical physics from the Technical University of Munich. Before joining Deutsche Bank, which was his first position in risk management, he spent several years in postdoctoral positions, both at the Center of Nonlinear Dynamics and Complex Systems, Brussels, and at the Siemens Research Department in Munich. He has published several articles on nonlinear dynamics and stochastic processes, as well as on risk modeling. He regularly speaks at conferences on quantitative topics in risk management.

List of Figures

Chapter 1

The Basics of Credit Risk Management

Why is credit risk management an important issue in banking? To answer this question let us construct an example which is, although simplified, nevertheless not too unrealistic: Assume a major building company is asking its house bank for a loan in the size of 100 million Euro. Somewhere in the bank's credit department a senior analyst has the difficult job of deciding if the loan will be given to the customer or if the credit request will be rejected. Let us further assume that the analyst knows that the bank's chief credit officer has known the chief executive officer of the building company for many years, and to make things even worse, the credit analyst knows from recent default studies that the building industry is under hard pressure and that the *bank-internal rating*[1] of this particular building company is just on the way down to a low *subinvestment grade* (low credit quality).

What should the analyst do? Well, the most natural answer would be that the analyst should reject the deal based on the information she or he has about the company and the current market situation. An alternative would be to grant the loan to the customer but to *insure* the loss potentially arising from the engagement by means of some credit risk management instrument (e.g., a so-called *credit derivative*).

Admittedly, we intentionally exaggerated in our description, but situations like the one just constructed happen from time to time and it is never easy for a credit officer to make a decision under such difficult circumstances. A brief look at any typical banking portfolio will be sufficient to convince people that defaulting obligors belong to the daily business of banking the same way as credit applications or ATM machines. Banks therefore started to think about ways of *loan insurance*

[1] A rating is an indication of creditworthiness; see Section 1.1.1.1.

many years ago, and the insurance paradigm will now directly lead us to the first central building block of credit risk management.

1.1 Expected Loss

Situations as the one described in the introduction suggest the need of a *loss protection* in terms of an *insurance*, as one knows it from car or health insurances. Moreover, history shows that even good customers have a potential to default on their financial obligations, such that an insurance for not only the critical but all loans in the bank's credit portfolio makes much sense.

The basic idea behind insurance is always the same. For example, in health insurance the costs of a few sick customers are covered by the total sum of revenues from the fees paid to the insurance company by all customers. Therefore, the fee that a man at the age of thirty has to pay for health insurance protection somehow reflects the insurance company's experience regarding *expected costs* arising from this particular group of clients.

For bank loans one can argue exactly the same way: Charging an appropriate *risk premium* for every loan and collecting these risk premiums in an internal bank account called *expected loss reserve* will create a capital cushion for covering losses arising from defaulted loans.

1.1.1 Remark Note that for many banks the paradigm of an expected loss reserve in the sense of *saving money in good times for spending it in bad times* is just a theoretical concept. For instance, US-GAAP[2] banks like Deutsche Bank or Credit Suisse who are both exchange-listed at Wall Street need to build loss reserves like, for instance, the so-called FAS-5 reserve, in a *period-conform* manner which means they can not be used in the afore-mentioned sense of an expected loss reserve.

But the paradigm of a reserve for expected losses is still used as a theoretical concept even in US-GAAP banks and, as it will be explained in a moment, expected loss is then applied as part of the risk premium charged to the borrower.

[2]US-GAAP stands for United States Generally Accepted Accounting Principles.

In probability theory the attribute *expected* always refers to an *expectation* or *mean value*, and this is also the case in risk management. The basic idea is as follows: The bank assigns to every customer a *probability of default* (PD), a loss fraction called the *loss given default* (LGD), describing the fraction of the loan's exposure expected to be lost in case of default, and the *exposure at default* (EAD) subject to be lost in the considered time period. The loss of any obligor is then defined by a *loss variable*

$$\tilde{L} \;=\; \text{EAD} \times \text{LGD} \times L \qquad \text{with} \qquad L = \mathbf{1}_D, \quad \mathbb{P}(D) = \text{PD}, \tag{1.1}$$

where D denotes the *event* that the obligor defaults in a certain period of time (most often one year), and $\mathbb{P}(D)$ denotes the probability of D.

The constituents of formula (1.1) are random variables. Although we will not go too much into technical details, we should mention here that underlying our model is some probability space $(\Omega, \mathcal{F}, \mathbb{P})$, consisting of a *sample space* Ω, a σ-*Algebra* $\mathcal{F}$, and a probability measure $\mathbb{P}$. The elements of $\mathcal{F}$ are the *measurable events* of the model, and intuitively it makes sense to claim that the event of default should be measurable. Moreover, it is common to identify $\mathcal{F}$ with the *information* available, and the information if an obligor defaults or survives should be included in the set of measurable events.

Note that the quantities PD, LGD, EAD and all quantities derived from those three are measured w.r.t. a specified time horizon. We drop the time aspect for now but will come back to it later in the text.

In the setup we just described it is very natural to define the *expected loss* (EL) of any customer or, more general, credit-risky asset as follows.

1.1.2 Definition *Given a loss variable $\tilde{L}$ as in (1.1), its expectation*

$$\text{EL} \;=\; \mathbb{E}[\tilde{L}]$$

is called the expected loss of the underlying credit-risky asset.

A common well-known formula for the EL appears in the following special situation.

1.1.3 Proposition *If the constituents of $\tilde{L}$ in (1.1) are independent, the expected loss can be written as*

$$\mathrm{EL} \;=\; \mathbb{E}[\mathrm{EAD}] \times \mathbb{E}[\mathrm{LGD}] \times \mathrm{PD}. \tag{1.2}$$

Moreover, if EAD and LGD are constant values the formula reads as

$$\mathrm{EL} \;=\; \mathrm{EAD} \times \mathrm{LGD} \times \mathrm{PD}. \tag{1.3}$$

Proof. The expectation of any *Bernoulli* random variable like $\mathbf{1}_D$ is its event probability. If the three factors in (1.1) are independent the expectation of their product is the product of their expectation. □

Note that making the assumption that EAD and LGD are constant values can be a good starting point for a back-of-the-envelope calculation to assign fixed values to EAD and LGD. However, in realistic situations EAD has to be modeled as a random variable due to uncertainties in payment profiles like, for instance, amortization, usage, and other drivers of EAD up to the chosen planning horizon.

In Section 1.1.4 we will briefly touch on the question of independence of PD, LGD and EAD. In fact, the independence assumption indeed is rather questionable and very much simplifying. Altogether one can say that (1.3) is the most simple representation formula for the expected loss one can have. The more simplifying assumptions are dropped the more one moves away from closed formulas like (1.3).

Although our focus in the book is pretty much on *portfolio risk* rather than on *single obligor risk* we briefly describe the three constituents of Formula (1.3) in the following paragraphs.

1.1.1 Probability of Default (PD)

The derivation of default probabilities is the "bread and butter" of a credit risk analytics team. There are situations where the assignment of default probabilities is straightforward and there are situations where it seems almost impossible to come up with a reasonable approach.

First of all we want to mention that later in Chapter 6 we will find that it is not sufficient to have default probabilities w.r.t. one particular time horizon. For instance, it is not enough to know for each credit-risky asset in the portfolio what its likelihood is to default within one year. Instead, one needs a whole *term structure*

$$(p_t)_{t\geq 0}$$

of default probabilities where t denotes time and for each point t in time the likelihood p_t is the default probability of the considered asset or client w.r.t. the time interval $[0, t]$. Note that in the literature PD term structures are often called *credit curves*. We dedicated the whole Chapter 6 in this book to this topic. In this section we focus on some basic remarks regarding PDs w.r.t. a fixed time horizon, say, one year. We also speak of 'clients' mostly in this section but what we essentially mean is any kind of credit-risky asset.

1.1.1.1 Ratings

Let us start with an upfront remark. Originally ratings were not developed for the derivation of PDs but only for the discrimination of credit quality on an ordinal scale. And in case of rating agencies which we will introduce later it still is the case that they do not assign PDs directly to rated clients but assign ratings in the sense of Table 1.1. So one has to be careful to put ratings and PDs in one bucket without keeping in mind that they are in fact different objects, as we will point out in a moment. However, because PDs are assigned to ratings and PDs are a main driver of the portfolio loss as well as all kinds of important ratios in banking, including regulatory capital related quantities, it is a common pattern that ratings and PDs are associated. Having said that, we continue our presentation from the viewpoint of the practitioner who uses ratings in the sense explained in the sequel.

The assignment of default probabilities to clients typically functions via so-called *rating systems*. A rating system can be thought of as a discretization of PDs on an *ordinal scale* which is called the *rating scale*. Discretization of a continuous metric quantity like a PD to an ordinal scale makes life in large organizations easier although one could argue that discretization seems a bit artificial and in the context of pricing introduces unnecessary jumps in pricing grids.

Well-known discretizations of PDs are the rating scales by the rating agencies *Moody's*, *Standard & Poor's*, and *Fitch*. Readers unfamiliar with the term "rating agency" can access background information on rating agencies, their work and their publications via their websites.

- For Moody's Investors Service go to: www.moodys.com;
- for Standard & Poor's go to: www.standardandpoors.com;
- for Fitch Ratings go to: www.fitchratings.com.

Rating scales of rating agencies look as follows. Standard & Poor's and Fitch use AAA, AA, A, BBB, BB, B, CCC, CC, C as a rating scale for rating best credit quality (AAA), 2nd-best credit quality (AA), and so on, until worst credit quality (C). The default state indicating that a company already failed in some payment obligation is denoted by D. Moody's uses Aaa, Aa, A, Baa, Ba, B, Caa, etc. to denote a comparable rating scale, again in decreasing order of credit quality.
Each of the rating agencies has a finer rating scale in place to allow for a finer distinction of credit quality among obligors. Standard & Poor's and Fitch, for instance, refine AA in AA+, AA and AA− where AA+ and AA− have lower respectively higher PDs than AA. Later in Section 1.1.1.2 we will work with the fine rating scale from Moody's. As an example and to underline what we just explained, Table 1.1 shows a definition of rating grades as it is used by Standard & Poor's. The wording in the table makes explicit that a rating grade and its assigned default probability address the *creditworthiness* of a client. The table in the upper half of Figure 1.1 shows a discretization of PDs to rating grades, this time w.r.t. Moody's ratings and their data history. The procedure of discretization of PDs, namely the assignment of a PD to every rating grade in the given rating scale is called a *rating calibration*; see Section 1.1.1.2. But before we come to that we want to briefly discuss rating systems in general.

One can divide the universe of rating systems into four broad categories which we will briefly describe in the sequel. It is important to note *that the rating type categories as we introduce them are not fully disjoint. In many cases a rating system has a main flavor but combines it with technology from some other rating model type. For instance, a rating model could be causal in principal but also use elements from scoring theory and regression.*

TABLE 1.1: S&P Rating Categories [172].

AAA	*best credit quality* *extremely reliable with regard to financial obligations*
AA	*very good credit quality* *very reliable*
A	*more susceptible to economic conditions* *still good credit quality*
BBB	*lowest rating in investment grade*
BB	*caution is necessary* *best sub-investment credit quality*
B	*vulnerable to changes in economic conditions* *currently showing the ability to meet its financial obligations*
CCC	*currently vulnerable to non-payment* *dependent on favorable economic conditions*
CC	*highly vulnerable to a payment default*
C	*close to or already bankrupt* *payments on the obligation currently continued*
D	*payment default on some financial obligation has actually occured*

Causal Rating Systems

We consider this type of rating system as superior to all other approaches. Whenever possible, this should be the way to proceed. As the name indicates, *causal rating systems* rely in their mechanism on a causal relationship between underlying credit risk drivers and the default event of an asset or borrower. To mention an example, ratings assigned to tranches in collateralized debt obligations (CDO; see [24]) typically are of causal type because the CDO model derives scenarios where the considered tranche is hit by a loss as well as the loss severity of the tranche as a direct consequence of "turbulences" in the underlying reference portfolio of credit-risky instruments. The model-derived hitting probability, for instance, can then be mapped onto a rating scale such that, for instance, a tranche with low hitting probability might have a letter combination like AAA or AA whereas a tranche with a low capital cushion below might get a rating letter combination of B or even in the C-range.

Why do we think that causal rating models are the best way to think about ratings? The reason is that a causal model approach forces the modeler to extensively analyze, understand and model the "true" mechanism of default. This is under all circumstances the best a modeler can do. For instance, a CDO model requires a fully-fledged model for both the cash flow structure of the CDO as well as the credit risk of the underlying reference portfolio. Causal models force the modeling team to really understand how defaults can happen and how losses will accumulate under certain circumstances.

As another important example let us briefly touch on causal ratings for public companies. The most famous representative of this type of rating systems is the concept of *Expected Default Frequencies* (EDF) from Moody's KMV[3]. An example of how Moody's KMV proceeds in their model is summarized in Section 1.2.3 and in Chapter 3. Note that Moody's KMV is continuously updating and improving their model framework and, therefore, our outline of their approach is indicative and illustrative only.

As a last example for default probabilities with a causal background based on market data we mention spread-implied default probabilities for companies with public debt outstanding. Spread-implied means

[3]See: www.moodyskmv.com; see also www.creditedge.com.

that default probabilities are derived from credit spreads of traded products like corporate bonds and credit derivatives; see Chapter 7.

Balance Sheet Scorings

In some situations a causal approach is rather difficult to follow or maybe it is even impossible to directly model the default mechanism. In such cases one can switch to *scoring systems* which are a good choice and well-established in rating units in banks across the globe. For instance, whereas for stock exchange-listed corporate clients a causal modeling of PDs is market standard as mentioned before, it is hardly thinkable to follow a causal approach for private corporates. Moreover, there are many companies which do not have a so-called external rating, which is a rating assigned by the afore-mentioned rating agencies. In such cases, a balance sheet scoring model is the usual approach to assign a bank-internal[4] rating to such companies. This is typically done by the credit analysts of the bank based on the rating tools developed by the rating quant team. For rating assignment the credit analysts consider various different quantitative and qualitative *drivers* of the considered firm's economic future like, for instance,

- Future *earnings* and *cashflows*,
- *debt*, *short- and long-term liabilities*, and *financial obligations*,
- *capital structure* (e.g., *leverage*),
- *liquidity* of the firm's assets,
- situation (e.g., political, social, etc.) of the firm's home *country*,
- situation of the *market* (e.g., *industry*), in which the company has its main activities,
- *management quality*, company *structure*, etc.

[4]Without going into details we would like to add that banks always should base the decision about creditworthiness on their bank-internal rating systems. As a main reason one could argue that banks know their customers best. Moreover, it is well known that external ratings do not react quickly enough to changes in the economic health of a company. Banks should be able to do it better based on their long-term relationship with their customers.

From this by no means exhaustive list it can be read-off that *rating drivers* can be *quantitative* as well as *qualitative*. To mention another important example, succession planning can be important for smaller firms but can not be captured as a solid quantitative figure like, for instance, a debt-equity relation.

It is best practice in banking that ratings as an outcome of a statistical tool are always re-evaluated by the credit analyst who makes the credit decision which leads to "approved" or "rejected". Credit analysts typically have, in line with their credit decision competence, the right to *overrule* or *override* the calculated rating. In most of the cases this will be an override to a better or worse rating grade by not more than one or two notches. The *overruling quote* which measures the relation of overruled ratings compared to overall assigned ratings is a good measure of the acceptance of a rating system by the practitioners in the credit unit, namely, the credit analysts who distinguish themselves from the rating quants who developed the rating system. An example for a "no-concern" value of an overruling quote is $5-10\%$ give or take, depending on the considered client segment. Overruling competence is crucial because especially for smaller firms one can expect that certain aspects driving the ability to pay of the client might not be captured by a standardized statistical tool.

The afore-mentioned *quantitative drivers* of the rating are taken from the balance sheet and annual report of the borrowing company. These sources of information are important for the lending credit institute because it is pretty much all one can get if a company is not listed at an exchange and has no public debt outstanding. Because the balance sheet is the primary source of information the name of the approach, *balance sheet scoring*, does not come much as a surprise. The afore-mentioned rating drivers are then grouped and set in relation to form a list of *balance sheet ratios* which are mapped into so-called *scores* as a metric-scale measure of default remoteness. The total score of a client is then based on a weighted sum of ratio transformation functions, often involving a lot of *regression analysis* in the development process. The score of a client is then *calibrated* to a PD based on the history of default frequencies; see Section 1.1.1.2. A typical calibration function in this context could look as follows,

$$\mathrm{PD}_{client} \;=\; \frac{1}{1+\exp(-\mathrm{SCORE}_{client})}\,, \tag{1.4}$$

where SCORE_{client} represents the final score based on the afore-mentioned sum of transformed ratios. Readers interested in a deeper dive into internal rating systems should read the the article from FRITZ, LUXENBURGER AND MIEHE [72]. Equation (1.4) is a representative of the class of so-called *logit* calibration functions which is a common transformation approach to get PDs out of scores in balance sheet scorings.

An industry example for an off-the-shelf model to obtain ratings for private companies is the RiskCalc[5] model by Moody's KMV.

Private Client Scorings

In the same way as one can build scoring systems for private companies one can derive scoring systems for private individuals, for instance, for clients which borrow money in the context of a residential mortgage. The basic mechanism is exactly the same but the rating score drivers are different. For instance, personal wealth, income situation, family context, etc. are typical drivers for a private client scoring. Moreover, practitioners know that the main drivers for default in a residential mortgage lending context and, more general, in any lending to private individuals are unemployment, divorce, and poor health.

Expert Rating Systems

There are portfolios where over many years hardly any default occurred. For example, municipalities in Switzerland can be grouped into a portfolio where almost no defaults occurred. In such situations it is difficult to work with balance sheet scorings because the number of defaults in the portfolio is too low for deriving statistically sound conclusions. This deficiency gave such portfolios a name. They are called *low default portfolios*; see WILDE and LEE [188] for more information and for ideas about how one could treat such portfolios. When objective data is missing, expert opinion is requested to come up with at least something. A common approach then is to overcome the problem of missing defaults by involving groups of experts in the considered segment in the bank to assign manual ratings to test cases. The modeling team then can apply techniques like *ordinal response methods* to establish a ranking of default remoteness (in the same way as rating grades) and later also a calibration to PDs for the client segment. Because of

[5]See: www.moodyskmv.com

the expert involvement, such rating approaches are called *expert rating systems*.

In general, and for all four approaches which we briefly outlined, one can measure the quality of a rating system by means of its so-called *discriminatory power*, a concept which goes beyond the purpose of this exposition. A beginner's crash course can be found in [16], pages 36 to 41. Readers interested in the topic should ask some web search engine with the keyword "discriminatory power" which will give them a large list of papers on the predictive power of rating systems, how predictive power is measured, how it can be improved and even what it means for a pricing system if discriminatory power can be increased due to a rating revision.

1.1.1.2 Calibration of Default Probabilities to Ratings

As mentioned several times before, the process of assigning a default probability to a rating grade (say, a letter combination in the sense of the rating agencies) is called a *calibration*. In this paragraph we will demonstrate how such a calibration works in principal. In a true-life situation the process is a bit more complex but for illustrative purposes our outline will be sufficient.

The end product of a calibration of default probabilities to ratings is a mapping of letter combinations (ratings) to default probabilities,

$$R \mapsto \mathrm{PD}(R),$$

such that to every rating R a certain default probability $\mathrm{PD}(R)$ is assigned. For example, the domain of such a mapping in case of Standard & Poor's letter ratings would be $\{AAA, AA, ..., C\}$ and, because we are dealing with probabilities, the range of the calibration function is in all cases the unit interval $[0, 1]$.

In the following, we explain by means of Moody's data how a calibration of default probabilities to external ratings can be done. From Moody's website or from other resources it is easy to get access to their study [141] of *historic corporate bond defaults*. There one can find a table like the one shown in Table 1.2 (see [141] Exhibit 40) showing historic default frequencies for the years 1983 up to 2000. The same exercise we are doing now can be done with Moody's most recent default study [167] or with comparable data from Standard & Poor's [171].

TABLE 1.2: Moody's Historic Corporate Bond Default Frequencies.

Rating	1983	1984	1985	1986	1987	1988
Aaa	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%
Aa1	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%
Aa2	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%
Aa3	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%
A1	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%
A2	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%
A3	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%
Baa1	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%
Baa2	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%
Baa3	0.00%	1.06%	0.00%	4.82%	0.00%	0.00%
Ba1	0.00%	1.16%	0.00%	0.88%	3.73%	0.00%
Ba2	0.00%	1.61%	1.63%	1.20%	0.95%	0.00%
Ba3	2.61%	0.00%	3.77%	3.44%	2.95%	2.59%
B1	0.00%	5.84%	4.38%	7.61%	4.93%	4.34%
B2	10.00%	18.75%	7.41%	16.67%	4.30%	6.90%
B3	17.91%	2.90%	13.86%	16.07%	10.37%	9.72%

Rating	1989	1990	1991	1992	1993	1994
Aaa	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%
Aa1	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%
Aa2	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%
Aa3	1.40%	0.00%	0.00%	0.00%	0.00%	0.00%
A1	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%
A2	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%
A3	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%
Baa1	0.00%	0.00%	0.76%	0.00%	0.00%	0.00%
Baa2	0.80%	0.00%	0.00%	0.00%	0.00%	0.00%
Baa3	1.07%	0.00%	0.00%	0.00%	0.00%	0.00%
Ba1	0.79%	2.67%	1.06%	0.00%	0.81%	0.00%
Ba2	1.82%	2.82%	0.00%	0.00%	0.00%	0.00%
Ba3	4.71%	3.92%	9.89%	0.74%	0.75%	0.59%
B1	6.24%	8.59%	6.04%	1.03%	3.32%	1.90%
B2	8.28%	22.09%	12.74%	1.54%	4.96%	3.66%
B3	19.55%	28.93%	28.42%	24.54%	11.48%	8.05%

Rating	1995	1996	1997	1998	1999	2000
Aaa	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%
Aa1	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%
Aa2	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%
Aa3	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%
A1	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%
A2	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%
A3	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%
Baa1	0.00%	0.00%	0.00%	0.00%	0.00%	0.29%
Baa2	0.00%	0.00%	0.00%	0.32%	0.00%	0.00%
Baa3	0.00%	0.00%	0.00%	0.00%	0.34%	0.98%
Ba1	0.00%	0.00%	0.00%	0.00%	0.47%	0.91%
Ba2	0.00%	0.00%	0.00%	0.61%	0.00%	0.66%
Ba3	1.72%	0.00%	0.47%	1.09%	2.27%	1.51%
B1	4.35%	1.17%	0.00%	2.13%	3.08%	3.25%
B2	6.36%	0.00%	1.50%	7.57%	6.68%	3.89%
B3	4.10%	3.36%	7.41%	5.61%	9.90%	9.92%

As said earlier, note that in our illustrative example we choose the *fine ratings scale* of Moody's, making finer differences regarding the creditworthiness of obligors.

Now, an important observation is that for best ratings no defaults at all have been observed. This is not as surprising as it looks at first sight: For example rating class Aaa is often calibrated with a default probability of 2 bps ("bp" stands for *"basispoint"* and means 0.01%), essentially meaning that one expects a Aaa-default on average twice in $10,000$ years. This is a long time to go; so, one should not be surprised that quite often best rating grades are lacking any default history. Nevertheless we believe that it would not be correct to take the historical zero-balance as an indication that these rating classes are risk-free opportunities for credit investment. Therefore, we have to find a way to assign small but positive default probabilities to those ratings.

Figure 1.1 shows our "quick-and-dirty working solution" of the problem, where we use the attribute "quick-and-dirty" because, as mentioned before, in a true life situation one would try to do the calibration a little more sophisticatedly[6].

Summarized in a recipe-like style, the calibration has three steps:

1. Denote by $h_i(R)$ the historic default frequency of rating class R for year i, where i ranges from 1983 to 2000. For example, $h_{1993}(Ba1) = 0.81\%$. Then compute the mean value and the standard deviation of these frequencies over the years, where the rating is fixed, namely

$$m(R) = \frac{1}{18}\sum_{i=1983}^{2000} h_i(R) \qquad \text{and}$$

$$s(R) = \sqrt{\frac{1}{17}\sum_{i=1983}^{2000}\big(h_i(R) - m(R)\big)^2}\ .$$

The mean value $m(R)$ for rating R is our first guess of the potential default probability assigned to rating R. The standard deviation $s(R)$ gives us some insight about the volatility and therefore about the error we eventually make when believing that $m(R)$

[6] For example, one could look at investment and sub-investment grades separately.

is a good estimate of the default probability of R-rated obligors. Figure 1.1 shows the values $m(R)$ and $s(R)$ for the considered rating classes. Because even best rated obligors are not free of default risk, we write "not observed" in the cells corresponding to $m(R)$ and $s(R)$ for ratings R=Aaa,Aa1,Aa2,A1,A2,A3 (ratings where no defaults have been observed) in Figure 1.1.

2. Next, we plot the mean values $m(R)$ into a coordinate system, where the x-axis refers to the rating classes (here numbered from 1 (Aaa) to 16 ($B3$)). One can see in the chart in Figure 1.1 that on a *logarithmic scale* the mean default frequencies $m(R)$ can be fitted by a regression line. Here we should add a comment that there is strong evidence from various empirical default studies that default frequencies grow exponentially with decreasing creditworthiness. For this reason we have chosen an exponential fit (linear on logarithmic scale). Using standard regression theory, see, e.g., [155] Chapter 4, or by simply using any software providing basic statistical functionality, one can easily obtain the following exponential function fitting our data:
$$\mathrm{PD}(x) \;=\; 3\times 10^{-5}\, e^{0.5075\,x} \qquad (x = 1, ..., 16).$$

3. As a last step, we use our regression equation for the estimation of default probabilities $\mathrm{PD}(x)$ assigned to rating classes x ranging from 1 to 16. Figure 1.1 shows our result, which we now call *a calibration of default probabilities to Moody's ratings*. Note that based on our regression even the best rating Aaa has a small but positive default probability. Moreover, we can hope that our regression analysis has smoothed out sampling errors from the historically observed data.

Although there would be much more to say about default probabilities, we stop the discussion for now and turn our attention to EAD and LGD from Formula (1.3).

1.1.2 The Exposure at Default

EAD is the quantity in Equation (1.3) specifying the exposure the bank does have to its borrower. In practice, banks grant to obligors so-called *credit lines* which function like a credit limit for the single-obligor exposure.

Rating	Mean	Standard-Deviation	Default Probability
Aaa	not observed	not observed	0.005%
Aa1	not observed	not observed	0.008%
Aa2	not observed	not observed	0.014%
Aa3	0.08%	0.33%	0.023%
A1	not observed	not observed	0.038%
A2	not observed	not observed	0.063%
A3	not observed	not observed	0.105%
Baa1	0.06%	0.19%	0.174%
Baa2	0.06%	0.20%	0.289%
Baa3	0.46%	1.16%	0.480%
Ba1	0.69%	1.03%	0.797%
Ba2	0.63%	0.86%	1.324%
Ba3	2.39%	2.35%	2.200%
B1	3.79%	2.49%	3.654%
B2	7.96%	6.08%	6.070%
B3	12.89%	8.14%	10.083%

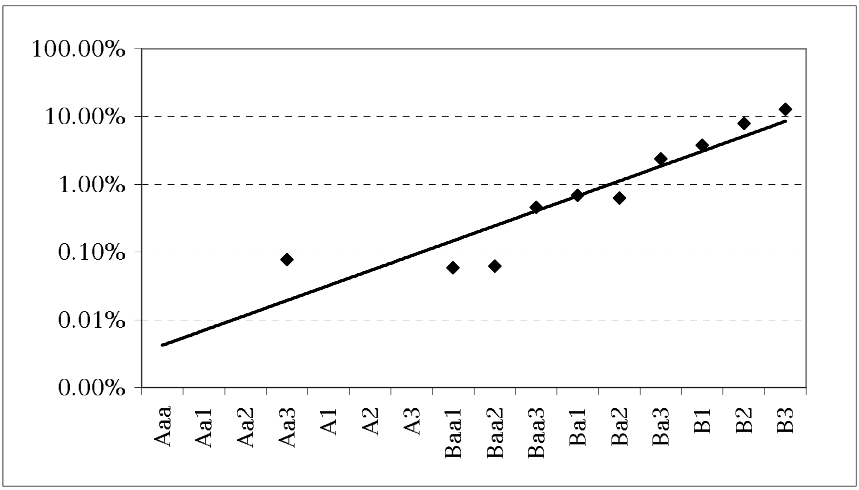


FIGURE 1.1: Calibration of ratings to default probabilities.

For the sake of a better understanding let us introduce a working example which will accompany us through this whole section on EAD. Let us assume that a credit analyst assigns to a borrower, say, a medium-sized firm, a credit line with a total limit of EUR 20m. Let us assume that the credit line is structured in the following way:

- Total credit line is EUR 20m.
- The borrower can draw EUR 12m as cash and can use the remaining EUR 8m of the credit line for so-called *contingent liabilities*, e.g., guarantees or comparable credit constructs but not for cash.

Now let us assume the borrower has drawn EUR 10m already. This part of the credit line is then called the *outstandings* of the client's exposure. The remaining open EUR 10m of the credit line are called *commitments*. In other words, the outstandings refer to the portion of the overall client exposure the obligor is already using. There is no randomness involved, drawn is drawn, and if the obligor defaults then the outstandings are subject to recovery and in a worst case situation could potentially be lost in total.

Of course, there is some time dynamics involved in outstandings. For instance, if the obligor pays back borrowed amounts over time then it makes a big difference whether an obligor defaults today or sometime in the future. Especially in mortgages where one often finds pre-determined amortization schemes the timing of default has a direct impact on the EAD. In our example one would need to accurately evaluate incoming cash from repayments versus newly opened parts of the credit line of the obligor which are subject to be drawn again, depending on the lending contract framework the bank and the obligor agreed to and signed.

The commitments, i.e., the remaining open EUR 10m of the borrower's credit line, are rather tricky to take into account. There is no other way than considering the exposure arising from the open part of the credit line as a random variable. So in our particular example we have EUR 10m open in the credit line but only EUR 2m can be drawn as cash. The other 8m can only be used for contingent liabilities. The two parts of the open line address different random effects:

- The EUR 2m which can be drawn as cash are driven by the likelihood that the borrower draws on them as well as by the fraction quantifying how much of the 2m she or he draws. Describing the situation by a simple equation we could write

$$\text{EAD}_{cash} \;=\; \mathbf{1}_D \times X \times [2\text{m}] \text{ (EUR)} \tag{1.5}$$

for the random exposure adding to current outstandings. Here, D describes the event (in the σ-field $\mathcal{F}$) that the obligor draws on the open cash credit line and X is a random variable defined on the underlying probability space $(\Omega, \mathcal{F}, \mathbb{P})$ with $X(\omega) \in [0, 1]$ for each $\omega \in \Omega$ quantifying the random fraction describing how much of the open 2m line is drawn. Altogether we are dealing with two random variables here. The equation could be made significantly more complex if one wants to take a stepwise drawing behavior into account, say, the obligor draws a partial amount in the future and another amount even later, and so on.

- The remaining EUR 8m which can be used for contingent liabilities are also subject to various random effects. First of all, there are again one or more indicator variables reflecting the optionality of usage of free parts of the credit line. Second, there is randomness in the fact that contingent liabilities not necessarily lead to cash exposure. A guarantee has no real exposure as of today but might converge into exposure in the future. Such random effects are typically treated by so-called *conversion factors*.

Let us put the pieces together for EAD calculation. We assume that the bank has a huge loss database useful for the calibration of exposure parameters. One common exposure parameter is the so-called *drawdown factor* (DDF). In our example it could be the case that the bank is able to say that the given type of obligor tends to draw on the free part of the credit line (EUR 2m) in 80% of the cases and on average uses 60% of the available cash. In other words, based on historic experience the bank obtains parameters in (1.5) like

$$\mathbb{P}(D) \;=\; 80\% \quad \text{and} \quad \mathbb{E}[X] \;=\; 60\%.$$

Assuming independence of $\mathbf{1}_D$ and X, this leads to an expected cash exposure for the unused part of the cash credit line of

$$\mathbb{E}[\text{EAD}_{cash}] \;=\; \mathbb{P}(D) \times \mathbb{E}[X] \times [2\text{m}] \text{ (EUR)} \;=\; 48\% \times [2\text{m}] \text{ (EUR)}.$$

The 48% would then be used as the DDF for this particular situation. Note that the DDF is one particular common example for conversion factors. For the contingent liability part of the credit line we assume again the existence of a rich database which allows for the calibration of a DDF of, say, 40% for the contingent liability part and a so-called *cash equivalent exposure factor* (CEEF) of 80% which is another conversion factor quantifying the conversion of the specific contingent liability, say, a guarantee, into a cash exposure. Altogether we obtain (assuming independence) the following representation for the EAD in our example:

$$\begin{aligned} \mathbb{E}[\mathrm{EAD}] &= [10m] + 48\% \times [2\mathrm{m}] + 32\% \times [8\mathrm{m}] \text{ (EUR)} \qquad (1.6) \\ &= [10m + 0.96m + 2.56m] \text{ (EUR)} \\ &= [13.52\mathrm{m}] \text{ (EUR)} \end{aligned}$$

where $32\% = 40\% \times 80\%$. So altogether our (expected) EAD is between the already utilized 10m and the overall committed 20m but higher than the committed cash line of 12m.

Our example provided some flavor on how complicated EAD calculations can be and in real life it actually is even more complex. For example, commitments of banks to clients often include various so-called *covenants*, which are *embedded options* which, for example, may force an obligor in times of financial distress to provide more collateral[7] or to renegotiate the terms of the loan.

A problem is that often the obligor has some informational advantage in that the bank recognizes financial distress of its borrowers with some delay. In case of covenants allowing the bank to close committed lines triggered by some early default indication, it really is a matter of timing whether the bank picks up such indications early enough to react before the customer has drawn on her or his committed lines. Bankers here often speak of a *race to default* which addresses the problem that distressed clients tend to exhaust their lines just before they default as much as possible.

The *Basel Committee on Banking Supervision*[8] provides conversion factors for banks who are unable or not allowed by their regulator to

[7] Collateral means assets securing a loan, e.g., mortgages, bonds, guarantees, etc. In case a loan defaults, the value of the collateral reduces the realized loss.

[8] The Basel Commitee coordinates the rules and guidelines for banking supervision. Its members are central banks and other national offices or government agencies responsible for banking supervision.

calibrate their own internal conversion factors like DDFs and CEEFs; see [149].

We stop here and come to the last of the three EL-relevant quantities, namely, the *loss-given-default* (LGD).

1.1.3 The Loss Given Default

A first distinction we need to make when it comes to LGDs is that of LGD as an amount of money and LGD as a percentage quote. The first mentioned is often denoted as \$LGD which means loss given default in monetary units. The concept of LGD is best demonstrated by means of an example in the same way as we proceeded for EAD.

Let us assume that a client has m credit products with the bank and pledged n collateral securities to the bank which can in case of default be used for recovery purposes in order to mitigate the realized loss arising from the client's default. Each credit product gets assigned an EAD such that for m credit products we get EAD_1, ..., EAD_m as well as expected recovery proceeds from the n collateral securities. We denote such recovery proceeds by $\$\mathrm{REC}_1$, ..., $\$\mathrm{REC}_n$. Such a constellation, having m credit products and n collateral securities is called an m-to-n situation. It can be difficult to get the interdependence and relation between products and collateral right, especially in cases where we have to deal with *dedicated collateral* which can be used for certain purposes under certain circumstances only. Here we assume that we can simply collect "good cash" (recovery proceeds) and "bad cash" (loss exposure) together in two separate buckets which we then compare to obtain our net balance with the defaulted client. What we get from that approach is the following:

$$\begin{aligned} \$\mathrm{LGD} \;&=\; \max\Big(0, (\mathrm{EAD}_1 + \cdots + \mathrm{EAD}_m) \\ &-\; (\$\mathrm{REC}_1 + \cdots + \$\mathrm{REC}_n)\Big) \end{aligned} \tag{1.7}$$

which leads to a percentage LGD of

$$\mathrm{LGD} \;=\; \frac{\$\mathrm{LGD}}{\mathrm{EAD}_1 + \cdots + \mathrm{EAD}_m}. \tag{1.8}$$

Note that we easily wrote down the quantities $\$\mathrm{REC}_i$ but, in fact, their derivation can be quite complex and needs a rich database storing historic proceeds from collateral security categories, collected with sufficient granularity. A typical discussion point in such calculations is, for instance, the time value of money. Recovery proceeds coming in later in time should be discounted in order to reflect the time value of money. The determination of an appropriate discount rate is just one out of many questions one has to solve in this context.

Summarizing one can say that LGD calibration is a long story and far from being trivial. The current regulatory framework forces banks with approval to use their internal PD, EAD and LGD calibrations to come up with good ideas on LGD calibration but we believe there is still a lot of ground to cover.

1.1.4 A Remark on the Relation between PD, EAD, LGD

Proposition 1.1.3 is based on the assumption that PD, EAD and LGD are independent. In real life this assumption is not realistic. When making such an assumption one should keep awareness that every calculation based on such an assumption creates a laboratory-like model environment in the same way as physicists make simplifying assumptions and create simplified environments in their laboratory.

Why is the independence assumption questionable? Let us focus on LGD and PD first. A fundamental principle in any market is the principle of *supply and demand*. Put in the context of defaults and losses this means the following. In a recession scenario one can expect that default rates increase. As a consequence, banks will be forced to sell collateral securities related to defaulted loans or assets to the market. This will increase supply for certain goods, for instance, in a residential mortgage crisis the market will be swampled with private homes offered for sale. A situation like this occurred in the subprime mortgage crisis in the US two years ago. Now, the principle of *supply and demand* leads to a price drop of such collateral securities which are now over-supplied in the market. But this in turn reduces recovery proceeds achievable by selling collateral to the market. Formula (1.7) shows that this leads to higher LGDs. So altogether we found that higher PDs in a recession can lead to higher LGDs which, neglecting time lagging, means that default rates and realized losses are positively related. One could also reformulate this statement and simply say that *defaults and recoveries to some extent are influenced by the same underlying systematic risk drivers* so that they can not be independent. A study by Altman et al. [3] shows empirical evidence for the comments we just made. What about EAD and PD? We mentioned before that in times of financial distress firms tend to draw on their open credit lines. This increases EADs in times where default rates are going high systematically. So even EAD can not safely be considered as independent from default rates and default rates are the basis for PD estimation.

For a nice approach to dependent LGD modeling we refer to the paper by HILLEBRAND [92].

1.2 Unexpected Loss

At the beginning of this chapter we introduced the EL of a transaction and imagined it as an insurance or loss reserve in order to cover losses the bank expects from historical default experience. But a focus on *expected* losses is not enough. In fact, the bank should in addition to the expected loss also make sure that thay have a good understanding on how much money would be necessary for covering *unexpected* losses where the attribute 'unexpected' addresses losses exceeding the historic average observed in the past. As a measure for the magnitude of the deviation of losses from the EL, the standard deviation of the loss variable $\tilde{L}$ as defined in (1.1) is a natural first choice.

1.2.1 Definition *The standard deviation*

$$\text{UL} \;=\; \sqrt{\mathbb{V}[\tilde{L}]} \;=\; \sqrt{\mathbb{V}\big[\text{EAD} \times \text{LGD} \times L\big]}$$

of the loss variable $\tilde{L}$ from (1.1) is called the unexpected loss of the underlying loan or asset.

In analogy to Proposition 1.1.3 one can prove the following representation formula for the UL of a loan.

1.2.2 Proposition *Under the assumption that EAD is deterministic and that LGD and the default event D are independent, the unexpected* *loss of a loan is given by*

$$\mathrm{UL} \;=\; \mathrm{EAD} \times \sqrt{\mathbb{V}[\mathrm{LGD}] \times \mathrm{PD} + \mathbb{E}[\mathrm{LGD}]^2 \times \mathrm{PD}(1-\mathrm{PD})}\;.$$

Proof. First, we square UL and get

$$\begin{aligned} \mathrm{UL}^2 &= \mathbb{V}\big[\mathrm{EAD} \times \mathrm{LGD} \times L\big] \\ &= \mathrm{EAD}^2 \times \mathbb{V}\big[\mathrm{LGD} \times L\big]. \end{aligned}$$

From the identity $\mathbb{V}[X] = \mathbb{E}[X^2] - \mathbb{E}[X]^2$ we get

$$\begin{aligned} \mathbb{V}\big[\mathrm{LGD} \times L\big] &= \mathbb{E}\big[\mathrm{LGD}^2 \times L^2\big] - \mathbb{E}\big[\mathrm{LGD} \times L\big]^2 \\ &= \mathbb{E}\big[\mathrm{LGD}^2\big] \times \mathbb{E}\big[L^2\big] - \mathbb{E}[\mathrm{LGD}]^2 \times \mathbb{E}[L]^2 \end{aligned}$$

because LGD and L are independent by assumption. Because $L = \mathbf{1}_D$ is a Bernoulli variable we have $\mathbb{E}[L^2] = \mathbb{E}[L] = \mathbb{P}[D] = \mathrm{PD}$ such that

$$\mathbb{V}\big[\mathrm{LGD} \times L\big] \;=\; \mathbb{E}\big[\mathrm{LGD}^2\big] \times \mathrm{PD} - \mathbb{E}[\mathrm{LGD}]^2 \times \mathrm{PD}^2.$$

Now we add $0 = \mathrm{PD} \times \mathbb{E}[\mathrm{LGD}]^2 - \mathrm{PD} \times \mathbb{E}[\mathrm{LGD}]^2$ and find

$$\mathbb{V}\big[\mathrm{LGD} \times L\big] \;=\; \mathrm{PD} \times \mathbb{V}[\mathrm{LGD}] + \mathbb{E}[\mathrm{LGD}]^2 \times \mathrm{PD}(1-\mathrm{PD}).$$

Collecting the pieces together we have our proof. □

We are now ready for a major step forward. So far we always looked at the credit risk of a single facility although banks have to manage *large portfolios* consisting of many different products with different risk characteristics. We therefore will now indicate how one can model the total loss of a credit portfolio.

For this purpose we consider a family of m loans

$$\begin{aligned} \tilde{L}_i &= \mathrm{EAD}_i \times \mathrm{LGD}_i \times L_i\;, \qquad (1.9) \\ \text{with} \quad L_i &= \mathbf{1}_{D_i}\;, \quad \mathbb{P}(D_i) = \mathrm{PD}_i\;. \end{aligned}$$

which we call a *portfolio* from now on.

1.2.3 Definition *A portfolio is a collection of loss variables $\tilde{L}_i$ as in (1.9). The portfolio loss is then defined as the random variable*

$$\tilde{L}_{PF} \;=\; \sum_{i=1}^{m} \tilde{L}_i \;=\; \sum_{i=1}^{m} \mathrm{EAD}_i \times \mathrm{LGD}_i \times L_i\;. \qquad (1.10)$$

Similar to the “standalone” quantities EL and UL we now obtain portfolio quantities EL_{PF} and UL_{PF}, defined as follows.

1.2.4 Definition *Given a portfolio of m loss variables as in (1.9), the expected and unexpected loss of the portfolio are given by*

$$\mathrm{EL}_{PF} \;=\; \mathbb{E}[\tilde{L}_{PF}] \quad \textit{and} \quad \mathrm{UL}_{PF} \;=\; \sqrt{\mathbb{V}[\tilde{L}_{PF}]}.$$

We briefly call them portfolio EL and portfolio UL in the sequel.

For the portfolio EL one has the following representation.

1.2.5 Proposition *Given a portfolio of m loss variables as in (1.9), the portfolio EL is always given by*

$$\mathrm{EL}_{PF} \;=\; \sum_{i=1}^{m} \mathrm{EL}_i \;=\; \sum_{i=1}^{m} \mathrm{EAD}_i \times \mathrm{LGD}_i \times \mathrm{PD}_i \tag{1.11}$$

where EL_i denotes the EL of the single loss $\tilde{L}_i$.

Proof. The assertion follows directly from the linearity of $\mathbb{E}[\,\cdot\,]$. $\square$

In case of the portfolio UL, linearity in general holds only if the loss variables $\tilde{L}_i$ are pairwise uncorrelated (see BIENAYMÉ’s Theorem in [12] Chapter 8). If the loss variables are correlated we can no longer expect that variance behaves linearly. Unfortunately, correlated loss variables are the standard case and the modeling of correlated variables is what this book is all about. So the portfolio UL is the first risk quantity we meet where correlations (say, covariances) between single-name risks play a fundamental role.

1.2.6 Proposition *Given a portfolio of m loss variables as in (1.9) with deterministic* EAD*’s, the portfolio UL is given by*

$$\mathrm{UL}_{PF} = \sqrt{\sum_{i=1}^{m}\sum_{j=1}^{m} \mathrm{EAD}_i \times \mathrm{EAD}_j \times \mathrm{Cov}[\mathrm{LGD}_i \times L_i\,,\, \mathrm{LGD}_j \times L_j]}\;. \tag{1.12}$$

Proof. The proposition is a direct consequence of the formula

$$\mathbb{V}\Big[\sum_{i=1}^{m} c_i X_i\Big] \;=\; \sum_{i=1}^{m}\sum_{j=1}^{m} c_i c_j \mathrm{Cov}[X_i, X_j]$$

for square-integrable random variables $X_1, ..., X_m$ and arbitrary constants $c_1, ..., c_m$. $\square$

1.2.7 Proposition *Given a portfolio of* m *loss variables as in (1.9) with deterministic* EADs *and deterministic* LGDs *we have*

$$\mathrm{UL}_{PF}^2 \;=\; \sum_{i,j=1}^{m} \mathrm{EAD}_i \times \mathrm{EAD}_j \times \mathrm{LGD}_i \times \mathrm{LGD}_j \times$$

$$\times \sqrt{\mathrm{PD}_i(1-\mathrm{PD}_i)\mathrm{PD}_j(1-\mathrm{PD}_j)}\,\rho_{ij}$$

where $\rho_{ij} = \mathrm{Corr}[L_i, L_j] = \mathrm{Corr}[\mathbf{1}_{D_i}, \mathbf{1}_{D_j}]$ *denote the so-called default correlation between counterparties (or assets)* i *and* j.

Proof. The proof follows from the representation of the portfolio UL according to Equation 1.12, from

$$\mathrm{Cov}[L_i, L_j] \;=\; \sqrt{\mathbb{V}[L_i]\,\mathbb{V}[L_j]}\,\mathrm{Corr}[L_i, L_j]$$

and from $\mathbb{V}[L_i] = \mathrm{PD}_i(1-\mathrm{PD}_i)$ for each $i = 1, ..., m$. $\square$

Before continuing we want to spend a moment with thinking about the meaning and interpretation of correlation. For simplicity let us consider a portfolio consisting of two loans with LGD= 100% and EAD= 1. We then only deal with L_i for $i = 1, 2$, and we set $\rho = \mathrm{Corr}[L_1, L_2]$ and $p_i = \mathrm{PD}_i$. Then, the squared UL of our portfolio is given by

$$\mathrm{UL}_{PF}^2 \;=\; p_1(1-p_1)+p_2(1-p_2)+2\rho\sqrt{p_1(1-p_1)}\sqrt{p_2(1-p_2)}\;. \quad (1.13)$$

We consider three possible cases regarding the default correlation ρ:

- $\rho = 0$. In this case, the third term in (1.13) vanishes. Although unusual in this context, one could say that $\rho = 0$ stands for *optimal diversification*. The concept of diversification is easily explained. Investing in many different assets generally reduces the overall portfolio risk, because usually it is very unlikely to see a large number of loans defaulting all at once. The less the loans in the portfolio have in common, the higher the chance that default of one obligor does not mean a lot to the economic future of other loans in the portfolio. The case $\rho = 0$ refers to a situation where the loans in the portfolio are completely unrelated. Interpreting the UL as a substitute[9] for portfolio risk, we see that this case minimizes the risk of joint defaults.

- $\rho > 0$. In this case our two counterparties are interrelated in that default of one counterparty increases the likelihood that the other counterparty will also default. We can make this precise by looking at the conditional default probability of counterparty 2 under the condition that obligor 1 already defaulted:

$$\mathbb{P}[L_2 = 1 \mid L_1 = 1] = \frac{\mathbb{P}[L_1 = 1, L_2 = 1]}{\mathbb{P}[L_1 = 1]} = \frac{\mathbb{E}[L_1 L_2]}{p_1} \tag{1.14}$$
$$= \frac{p_1 p_2 + \mathrm{Cov}[L_1, L_2]}{p_1} = p_2 + \frac{\mathrm{Cov}[L_1 L_2]}{p_1} \; .$$

So we see that positive correlation respectively covariance leads to a conditional default probability higher (because $\mathrm{Cov}[L_1, L_2] > 0$) than the unconditional default probability p_2 of obligor 2. In other words, in case of positive correlation any default in the portfolio has an important implication on other facilities in the portfolio, namely that there might be more losses to be encountered. The extreme case in this scenario is the case of *perfect correlation* ($\rho = 1$). In the case of $p = p_1 = p_2$, Equation (1.13) shows that in the case of perfect correlation we have $\mathrm{UL}_{PF} = 2\sqrt{p(1-p)}$, essentially meaning that our portfolio contains the risk of only one obligor but with double intensity (*concentration risk*). In this situation it follows immediately from (1.14) that default of one obligor makes the other obligor defaulting almost surely.

- $\rho < 0$. This is the mirrored situation of the case $\rho > 0$. We therefore only discuss the extreme case of perfect anti-correlation ($\rho = -1$). One then can view an investment in asset 2 as an

[9] Note that in contrast to the EL, the UL is the "true" uncertainty the bank faces when investing in a portfolio because it captures the deviation from the expectation.

almost *perfect hedge* against an investment in asset 1, if (additionally to $\rho = -1$) the characteristics (exposure, rating, etc.) of the two loans match. Admittedly, this terminology makes much more sense when following a *marked-to-market*[10] approach to loan valuation, where an increase in market value of one of the loans immediately (under the assumption $\rho = -1$) would imply a decrease in market value of the other loan. However, from (1.13) it follows that in the case of a perfect hedge the portfolio's UL completely vanishes ($\mathrm{UL}_{PF} = 0$). This means that our perfect hedge (investing in asset 2 with correlation -1 w.r.t. a comparable and already owned asset 1) completely eliminates (neutralizes) the risk of asset 1.

We now turn to the important notion of economic capital.

1.2.1 Economic Capital

We have learned so far that banks should hold some capital cushion against unexpected losses. However, defining the UL of a portfolio as the *risk capital* saved for cases of financial distress is not the best choice, because there might be a significant likelihood that losses will exceed the portfolio's EL by more than one standard deviation of the portfolio loss. Therefore one seeks other ways to quantify risk capital, hereby taking a *target level* of *statistical confidence* into account.

The most common way to quantify risk capital is the concept of *economic capital*[11]

1.2.8 Definition *Let a portfolio* $(\tilde{L}_i)_{i=1,\ldots,m}$ *be given. The economic capital (EC) w.r.t. a prescribed level of confidence* α *is defined as the* α*-quantile of the portfolio loss* $\tilde{L}_{PF}$ *minus the EL of the portfolio:*

$$\mathrm{EC}_\alpha \;=\; q_\alpha - \mathrm{EL}_{PF} \tag{1.15}$$

where q_α *is the* α*-quantile of* $\tilde{L}_{PF}$ *given as*

$$q_\alpha \;=\; \inf\{q > 0 \mid \mathbb{P}[\tilde{L}_{PF} \le q] \ge \alpha\}\ . \tag{1.16}$$

[10] In a marked-to-market framework loans do not live in a two-state world (default or survival) but rather are evaluated w.r.t. their *market value.*

[11] Synonymously called *Capital at Risk* (CaR) in the literature; the quantile q_α from Definition 1.2.8 sometimes is called the (credit) *Value-at-Risk* (VaR).

For example, if the level of confidence is set to $\alpha = 99.98\%$, then the risk capital EC_α will (on average) be sufficient to cover unexpected losses in 9,998 out of 10,000 years, hereby assuming a planning horizon of one year. Unfortunately, under such a calibration one can on the other side expect that in 2 out of 10,000 years the economic capital $\mathrm{EC}_{99.98\%}$ will not be sufficient to protect the bank from insolvency. This is the downside when calibrating risk capital by means of quantiles. However, today most major banks use an EC framework for their internal credit risk model.

The reason for reducing the quantile q_α by the EL is due to the "best practice" of decomposing the total risk capital (i.e., the quantile) into a first part covering expected losses and a second part meant as a cushion against unexpected losses. Altogether the pricing of a loan typically takes several cost components into account. First of all, the price of the loan should include the costs of administrating the loan and maybe some kind of upfront fees. Second, expected losses are charged to the customer, hereby taking the creditworthiness captured by the customer's rating into account. More risky customers have to pay a higher risk premium than customers showing high credit quality. Third, the bank will also ask for some compensation for taking the risk of unexpected losses coming with the new loan into the bank's credit portfolio. The charge for unexpected losses is often calculated as the *contributory* EC of the loan in reference to the lending bank's portfolio; see Chapter 5. In contrast to the EL which is priced in completely, the EC often is only partially charged in form of

$$\text{EC-charge} \;=\; \mathrm{EC}_{contributory} \times \mathrm{HR}\ [\%]$$

where HR denotes some *hurdle rate*, e.g., 25%.

Note that there is an important difference between the EL and the EC charges: The EL charge is independent from the composition of the reference portfolio, whereas the EC charge strongly depends on the current composition of the portfolio in which the new loan will be included. For example, if the portfolio is already well diversified, then the EC charge as a price for taking unexpected risk does not have to be as high as it would be in the case for a portfolio in which, for example, the new loan would induce or increase some concentration risk. Summarizing one can say the EL charges are *portfolio independent*, but EC charges are *portfolio dependent*. This makes the calculation

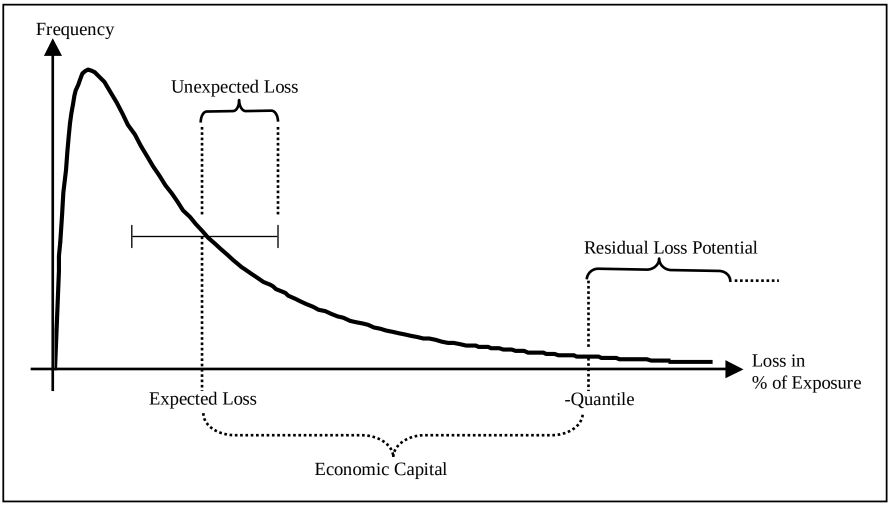


FIGURE 1.2: The portfolio loss distribution.

of the contributory EC in pricing tools more complicated, because one always has to take the complete reference portfolio into account. *Risk contributions* will be discussed in Chapter 5.

An alternative to EC is a risk capital based on *Expected Shortfall* (ES). A capital definition according to ES very much reflects an insurance point of view of the credit risk business. Today it is known that ES is superior to EC as a risk capital measure for various reasons. We will come back to ES and its properties in Chapter 5.

1.2.2 The Loss Distribution

All risk quantities on a portfolio level are based on the portfolio loss variable $\tilde{L}_{PF}$. Therefore it does not come much as a surprise that the distribution of $\tilde{L}_{PF}$, the so-called *loss distribution* of the portfolio, plays a central role in credit risk management. In Figure 1.2 it is illustrated that all risk quantities of the credit portfolio can be identified by means of the loss distribution of the portfolio. This is an important observation, because it shows that in cases where the distribution of the portfolio loss can only be determined in an empirical way one can use empirical statistical quantities as a proxy for the respective "true" risk quantities.

In practice there are various ways to generate a loss distribution; see Section 2.8. The first method is based on *Monte Carlo simulation*; the second is based on a so-called *analytical approximation*. We describe both methods in short in the following section but come back to the topic of generating loss distributions in greater detail in Section 2.8.

1.2.2.1 Monte Carlo Simulation of Losses

In a Monte Carlo simulation, losses are simulated and tabulated in form of a *histogram* in order to obtain an *empirical loss distribution* of the underlying portfolio. The *empirical distribution function* can be determined as follows:

Assume we have simulated n potential portfolio losses $\tilde{L}_{PF}^{(1)}, ..., \tilde{L}_{PF}^{(n)}$, hereby taking the driving distributions of the single loss variables and their correlations[12] into account. Then the empirical loss distribution function is given by

$$F(x) \;=\; \frac{1}{n}\sum_{j=1}^{n} \mathbf{1}_{[0,x]}(\tilde{L}_{PF}^{(j)}) \,. \tag{1.17}$$

Figure 1.3 shows the shape of the density (histogram of the randomly generated numbers $(\tilde{L}_{PF}^{(1)}, ..., \tilde{L}_{PF}^{(n)})$) of the empirical loss distribution of some test portfolio.

From the empirical loss distribution we can derive all of the portfolio risk quantities introduced in the previous paragraphs. For example, the α-quantile of the loss distribution can directly be obtained from our simulation results $\tilde{L}_{PF}^{(1)}, ..., \tilde{L}_{PF}^{(n)}$ as follows:

Starting with the order statistics of $\tilde{L}_{PF}^{(1)}, ..., \tilde{L}_{PF}^{(n)}$, say

$$\tilde{L}_{PF}^{(i_1)} \le \tilde{L}_{PF}^{(i_2)} \le \cdots \le \tilde{L}_{PF}^{(i_n)} \,,$$

the α-quantile $\hat{q}_\alpha$ of the empirical loss distribution for any confidence level α is given by

$$\hat{q}_\alpha \;=\; \begin{cases} \alpha\tilde{L}_{PF}^{(i_{[n\alpha]})} + (1-\alpha)\tilde{L}_{PF}^{(i_{[n\alpha]+1})} & \text{if} \quad n\alpha \in \mathbb{N} \\ \tilde{L}_{PF}^{(i_{[n\alpha]})} & \text{if} \quad n\alpha \notin \mathbb{N} \end{cases} \tag{1.18}$$

[12]We will later see that correlations are incorporated by means of a *factor model*.

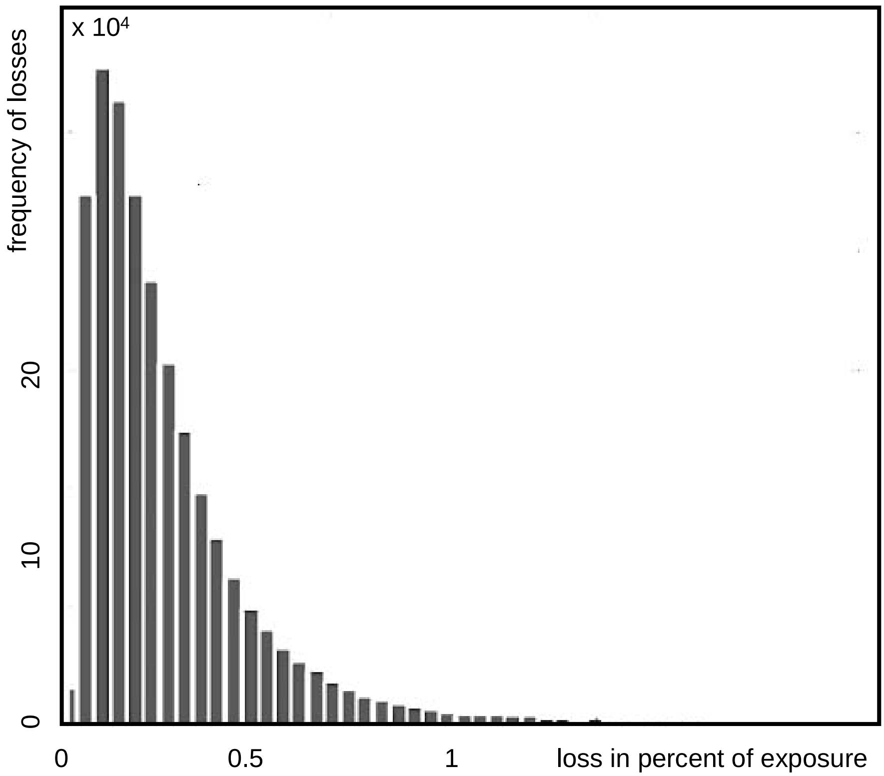


FIGURE 1.3: An empirical portfolio loss distribution obtained by Monte Carlo simulation. The histogram is based on a portfolio of 2.000 middle-size corporate loans.

$$\text{where} \qquad [n\alpha] \;=\; \min\Big\{k \in \{1, ..., n\} \mid n\alpha \le k\Big\} \;.$$

The economic capital can then be estimated by

$$\widehat{\mathrm{EC}}_\alpha \;=\; \hat{q}_\alpha - \frac{1}{n}\sum_{j=1}^{n} \tilde{L}_{PF}^{(j)} \;. \tag{1.19}$$

In an analogous manner, any other risk quantity can be obtained by calculating the corresponding empirical statistics.

Approaching the loss distribution of a large portfolio by Monte Carlo simulation always requires a sound *factor model*; see Section 1.2.3. The classical statistical reason for the existence of factor models is the wish to explain the variance of a variable in terms of underlying factors. Despite the fact that in credit risk we also wish to explain the variability of a firm's economic success in terms of global underlying influences, the necessity for factor models comes from two major reasons.

First of all, the correlation between single loss variables should be made interpretable in terms of *economic variables*, such that large losses can be explained in a sound manner. For example, a large portfolio loss might be due to the *downturn* of an industry common to many counterparties in the portfolio. Along this line, a factor model can also be used as a tool for *scenario analysis*. For example, by setting an industry factor to a particular fixed value and then starting the Monte Carlo simulation again, one can study the impact of a down- or upturn of the respective industry.

The second reason for the need of factor models is a reduction of the computational effort. For example, for a portfolio of 100,000 transactions, $\frac{1}{2} \times 100,000 \times 99,999$ correlations have to be calculated. In contrast, modeling the correlations in the portfolio by means of a factor model with 100 indices reduces the number of involved correlations by a factor of 1,000,000. We will come back to factor models in 1.2.3 and also in later chapters.

1.2.2.2 Analytical Approximation

Another approach to the portfolio loss distribution is by analytical approximation. Roughly speaking, the analytical approximation maps an actual portfolio with unknown loss distribution to an equivalent portfolio with known loss distribution. The loss distribution of the equivalent portfolio is then taken as a substitute for the "true" loss distribution of the original portfolio.

In practice this is often done as follows. Choose a family of distributions characterized by its first and second moment, showing the typical shape (i.e., right-skewed with fat tails[13]) of loss distributions as illustrated in Figure 1.2.

From the known characteristics of the original portfolio (e.g., rating distribution, exposure distribution, maturities, etc.) calculate the first moment (EL) and estimate the second (centered) moment (UL^2).

Note that the EL of the original portfolio usually can be calculated based on the information from the rating, exposure, and LGD distributions of the portfolio.

Unfortunately the second moment can not be calculated without any assumptions regarding the default correlations in the portfolio; see Equation (1.13). Therefore, one now has to make an assumption regarding an *average default correlation* ρ. Note that in case one thinks in terms of *asset value models*, see Section 2.4.1, one would rather guess an average *asset correlation* instead of a default correlation and then calculate the corresponding default correlation by means of applying Proposition 2.5.1 to the definition of the default correlation. However, applying Equation (1.13) by setting all default correlations ρ_{ij} equal to ρ will provide an estimated value for the original portfolio's UL.

Now one can choose from the parametrized family of loss distribution the distribution best matching the original portfolio w.r.t. first and second moments. This distribution is then interpreted as the loss distribution of an equivalent portfolio which was selected by a *moment matching* procedure.

Obviously the most critical part of an analytical approximation is the determination of the average asset correlation. Here one has to rely on practical experience with portfolios where the average asset correlation is known. For example, one could compare the original portfolio with a set of typical bank portfolios for which the average asset correlations are known. In some cases there is empirical evidence regarding a reasonable range in which one would expect the unknown correlation to be

[13] In our terminology, a distribution has *fat tails*, if its quantiles at high confidence are higher than those of a normal distribution with matching first and second moments.

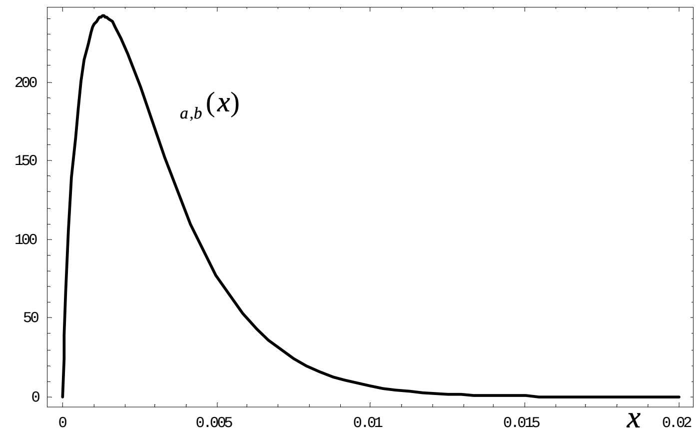


FIGURE 1.4: Analytical approximation by some beta distribution.

located. For example, if the original portfolio is a retail portfolio, then one would expect the average asset correlation of the portfolio to be a small number, maybe contained in the interval $[1\%, 5\%]$. If the original portfolio contains loans given to large firms, then one would expect the portfolio to have a high average asset correlation, maybe somewhere between 40% and 60%. Just to give another example, the new *Basel Capital Accord* (see Section 1.3) assumes an average asset correlation of 20% for corporate loans; see [148]. In Section 2.7 we estimate the average asset correlation in Moody's universe of rated corporate bonds to be around 25%. Summarizing, we can say that calibrating[14] an average correlation is on one hand a typical source of *model risk*, but on the other hand nevertheless often supported by some practical experience.

As an illustration of how the moment matching in an analytical approximation works, assume that we are given a portfolio with an EL of 30 bps and an UL of 22.5 bps, estimated from the information we have about some credit portfolio combined with some assumed average correlation.

Now, in Section 2.5 we will introduce a typical family of two-parameter loss distributions used for analytical approximation. Here, we want to approximate the loss distribution of the original portfolio by a beta

[14]The calibration might be more honestly called a "guestimate", a mixture of a guess and an estimate.

distribution, matching the first and second moments of the original portfolio. In other words, we are looking for a random variable

$$X \sim \beta(a, b) \ ,$$

representing the percentage portfolio loss, such that the parameters a and b solve the following equations:

$$0.003 = \mathbb{E}[X] = \frac{a}{a+b} \qquad \text{and} \tag{1.20}$$

$$0.00225^2 = \mathbb{V}[X] = \frac{ab}{(a+b)^2(a+b+1)} \ .$$

Hereby recall that the probability density φ_X of X is given by

$$\varphi_X(x) \ = \ \beta_{a,b}(x) \ = \ \frac{\Gamma(a+b)}{\Gamma(a)\Gamma(b)}\, x^{a-1}(1-x)^{b-1} \tag{1.21}$$

($x \in [0,1]$) with first moment and second (centered) moment

$$\mathbb{E}[X] = \frac{a}{a+b} \qquad \text{and} \qquad \mathbb{V}[X] = \frac{ab}{(a+b)^2(a+b+1)} \ .$$

Equations (1.20) represent the moment matching addressing the "correct" beta distribution matching the first and second moments of our original portfolio. It turns out that $a = 1.76944$ and $b = 588.045$ solve equations (1.20). Figure 1.4 shows the probability density of the so calibrated random variable X.

The analytical approximation takes the random variable X as a proxy for the unknown loss distribution of the portfolio we started with. Following this assumption, the risk quantities of the original portfolio can be approximated by the respective quantities of the random variable X. For example, quantiles of the loss distribution of the portfolio are calculated as quantiles of the beta distribution. Because the "true" loss distribution is substituted by a closed-form, analytical, and well-known distribution, all necessary calculations can be done in fractions of a second. The price we have to pay for such convenience is that all calculations are subject to significant *model risk*. Admittedly, the beta distribution as shown in Figure 1.4 has the shape of a loss distribution, but there are various two-parameter families of probability densities having the typical shape of a loss distribution. For example, some gamma distributions, the F-distribution, and also the distributions introduced in Section 2.5 have such a shape. Unfortunately they all have different tails, such that in case one of them would approximate really well the unknown loss distribution of the portfolio, the others automatically would be the wrong choice. Therefore, the selection of an appropriate family of distributions for an analytical approximation is a remarkable source of model risk. Nevertheless there are some families of distributions that are established as *best practice* choices for particular cases. For example, the distributions in Section 2.5 are a very natural choice for analytical approximations, because they are limit distributions of a well understood model.

In practice, analytical approximation techniques can be applied quite successfully to so-called *homogeneous portfolios*. These are portfolios where all transactions in the portfolio have comparable risk characteristics, for example, no exposure concentrations, default probabilities in a band with moderate bandwidth, only a few (better: one single!) industries and countries, and so on. There are many portfolios satisfying such constraints. For example, many retail banking portfolios and also many portfolios of smaller banks can be evaluated by analytical approximations with sufficient precision.

In contrast, a full Monte Carlo simulation of a large portfolio can last several hours, depending on the number of counterparties and the number of scenarios necessary to obtain sufficiently rich tail statistics for the chosen level of confidence.

The main advantage of a Monte Carlo simulation is that it accurately captures the dependencies inherent in the portfolio instead of relying on a whole bunch of assumptions. Moreover, a Monte Carlo simulation takes into account all the different risk characteristics of the loans in the portfolio. However, in Section 2.8 we also touch on valuable alternatives to the Monte Carlo approach.

1.2.3 Modeling Correlations by Means of Factor Models

Factor models are a well established technique from multivariate statistics, applied in credit risk models, for identifying underlying drivers of correlated defaults and for reducing the computational effort regarding the calculation of correlated losses. We start by discussing the basic meaning of a *factor*.

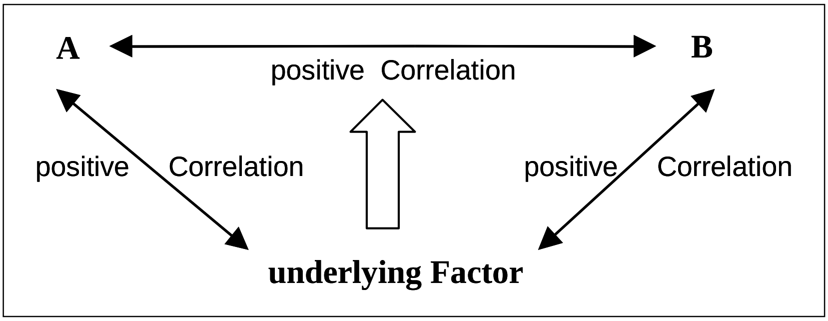


FIGURE 1.5: Correlation induced by an underlying factor.

Assume we have two firms A and B which are positively correlated. For example, let A be DaimlerChrysler and B stand for BMW. Then, it is quite natural to explain the positive correlation between A and B by the correlation of A and B with an *underlying factor*; see Figure 1.5. In our example we could think of the *automotive industry* as an underlying factor having significant impact on the economic future of the companies A and B. Of course there are probably some more underlying factors driving the riskiness of A and B. For example, DaimlerChrysler is to a certain extent also influenced by a factor for *Germany*, the *United States*, and eventually by some factors incorporating *Aero Space* and *Financial Companies*. BMW is certainly correlated with a country factor for *Germany* and probably also with some other factors. However, the crucial point is that factor models provide a way to express the correlation between A and B exclusively by means of their correlation with common factors. As already mentioned in the previous section, we additionally wish underlying factors to be *interpretable* in order to identify the reasons why two companies experience a down- or upturn at about the same time. For example, assume that the automotive industry gets under pressure. Then we can expect that companies A and B also get under pressure, because their fortune is related to the automotive industry. The part of the *volatility* of a company's financial success (e.g., incorporated by its asset value process) related to systematic factors like industries or countries is called the *systematic risk* of the firm. The part of the firm's asset volatility that can not be explained by systematic influences is called the *specific* or *idiosyncratic risk* of the firm. We will make both notions precise later on in this section.

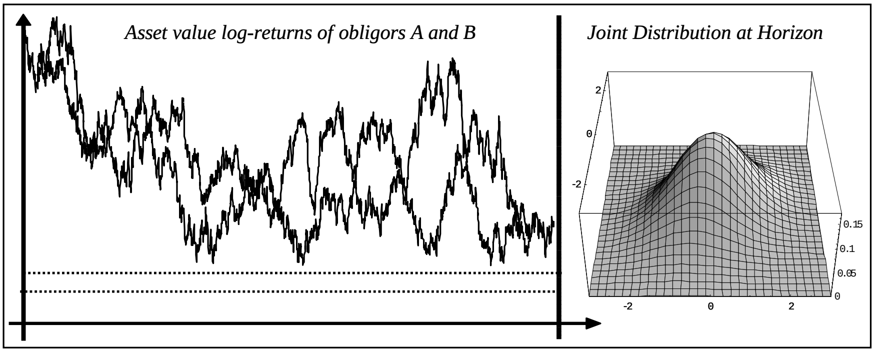


FIGURE 1.6: Correlated processes of obligor's asset value log-returns.

In the sequel we introduce an example of a typical factor model as it has been used by two industry leaders, namely, Moody's KMV and the RiskMetrics Group in their models for years. The companies behind the models continue to contribute in the area of credit risk research and modeling. Readers interested in information about the two firms can consult their websites

- www.moodyskmv.com
- www.riskmetrics.com

and will find a lot of information including research articles. Both firms continuously develop and improve their models so that one has to read through their most recent documentation in order to get a fresh taste on the current state of their models. However, our exposition of a typical factor model is meant as an *illustrative example only* which has the sole purpose to demonstrate how such a model works in principal.

Both models incorporate the idea that every firm admits a process of *asset values*, such that default or survival of the firm depends on the state of the asset values at a certain planning horizon. If the process has fallen below a certain critical threshold, called the *default point* of the firm, then the company has defaulted. If the asset value process is above the critical threshold, the firm survives. Asset value models have their roots in Merton's seminal paper [137] and will be explained in detail in Chapter 3 and also to some extent in Section 2.4.1.

Figure 1.6 illustrates the asset value model for two counterparties. Two correlated processes describing two obligor's asset values are shown. The correlation between the processes is called the *asset correlation*. In case the asset values are modeled by *geometric Brownian motions* (see Chapter 3), the asset correlation is just the correlation of the driving Brownian motions. At the planning horizon, the processes induce a bivariate asset value distribution. In the classical Merton model, where asset value processes are correlated geometric Brownian motions, the log-returns of asset values are normally distributed, so that the joint distribution of two asset value log-returns at the considered horizon is bivariate normal with a correlation equal to the asset correlation of the processes, see also Proposition 2.5.1. The dotted lines in Figure 1.6 indicate the critical thresholds or default points for each of the processes. Regarding the calibration of these default points we refer to CROSBIE [36] for an introduction.

Now let us start with the model used for years by Moody's KMV. They named it the *Global Correlation Model*^{TM}. A highly readable summary of the model can be found in CROUHY, GALAI, and MARK [38]. Our approach to describe the model is slightly different than other presentations because we want to have the relevant formulas in a way supporting a convenient algorithm for the calculation of asset correlations.

Following Merton's model[15], the *Global Correlation Model*^{TM} focuses on the asset value log-returns r_i of counterparties $(i = 1, ..., m)$ at a certain planning horizon (typically 1 year), admitting a representation

$$r_i = \beta_i \Phi_i + \varepsilon_i \qquad (i = 1, ..., m). \tag{1.22}$$

Here, Φ_i is called the *composite factor* of firm i, because in multi-factor models Φ_i typically is a weighted sum of several factors. Equation (1.22) is nothing but a standard *linear regression* equation, where the sensitivity coefficient, β_i, captures the linear correlation of r_i and Φ_i. In analogy to the *capital asset pricing model* (CAPM) (see, e.g., [38]) β is called the *beta* of counterparty i. The variable ε_i represents the *residual* part of r_i, essentially meaning that ε_i is the error one makes

[15] Actually, although the *Global Correlation Model*^{TM} in principal follows Merton's model, it does not really work with Gaussian distributions but rather relies on an empirically calibrated framework; see CROSBIE [36] and also Chapter 3.

when substituting r_i by $\beta_i\Phi_i$. Merton's model lives in a log-normal world, so that $\boldsymbol{r} = (r_1, ..., r_m) \sim N(\boldsymbol{\mu}, \Gamma)$ is multivariate Gaussian with a correlation matrix Γ. The composite factors Φ_i and ε_i are accordingly also normally distributed. Another basic assumption is that ε_i is independent of the Φ_i's for every i. Additionally the residuals ε_i are assumed to be uncorrelated[16]. Therefore, the returns r_i are exclusively correlated by means of their composite factors. This is the reason why Φ_i is thought of as the *systematic* part of r_i, whereas ε_i due to its independence from all other involved variables can be seen as a random effect just relevant for counterparty i. Now, in regression theory one usually decomposes the *variance* of a variable in a systematic and a specific part. Taking variances on both sides of Equation (1.22) yields

$$\mathbb{V}[r_i] \;=\; \underbrace{\beta_i^2\mathbb{V}[\Phi_i]}_{systematic} \;+\; \underbrace{\mathbb{V}[\varepsilon_i]}_{specific} \qquad (i = 1, ..., m). \tag{1.23}$$

Because the variance of r_i captures the risk of unexpected movements of the asset value of counterparty i, the decomposition (1.23) can be seen as a splitting of total risk of firm i in a *systematic* and a *specific* risk. The former captures the variability of r_i coming from the variability of the composite factor, which is $\beta_i^2\mathbb{V}[\Phi_i]$; the latter arises from the variability of the residual variable, $\mathbb{V}[\varepsilon_i]$. Note that some people say *idiosyncratic* instead of *specific*.

Alternatively to the beta of a firm one could also look at the *coefficient of determination* of the regression Equation (1.22). The coefficient of determination quantifies how much of the variability of r_i can be explained by Φ_i. This quantity is usually called the *R-squared*, R^2, of counterparty i and constitutes an important input parameter in all credit risk models based on asset values. It is usually defined as the systematic part of the variance of the standardized[17] returns $\tilde{r}_i = (r_i - \mathbb{E}[r_i])/\sqrt{\mathbb{V}[r_i]}$, namely

$$R_i^2 \;=\; \frac{\beta_i^2\mathbb{V}[\Phi_i]}{\mathbb{V}[r_i]} \qquad (i = 1, ..., m). \tag{1.24}$$

The residual part of the total variance of the standardized returns $\tilde{r}_i$ is then given by $1 - R_i^2$, thereby quantifying the percentage value of the specific risk of counterparty i.

[16] Recall that in the Gaussian case *uncorrelated* is equivalent to *independent*.

[17] That is, normalized in order to have mean zero and variance one.

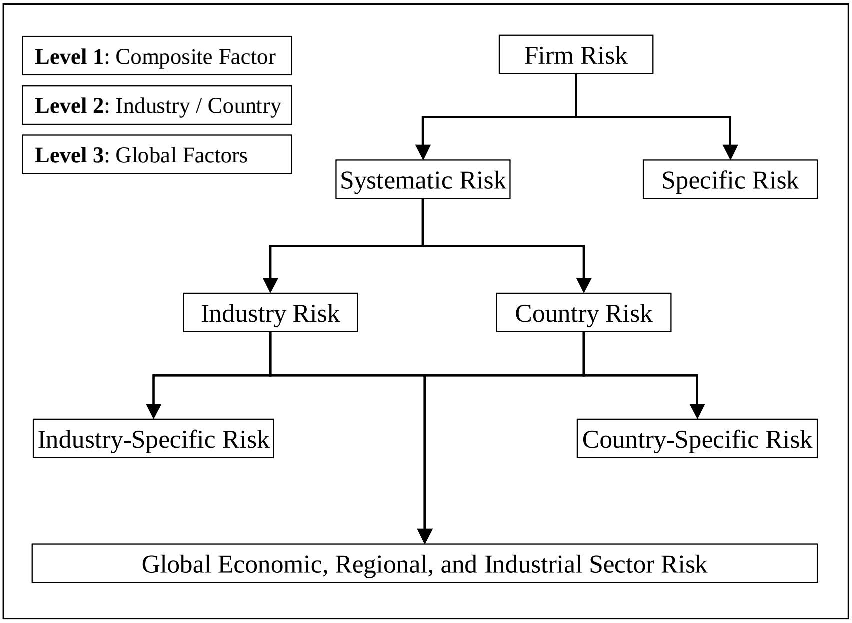


FIGURE 1.7: Three-level factor structure in the *Global Correlation Model*^{TM}; see also comparable presentations in the literature, e.g., Figure 9.9. in [38] and Figure 1.7 in [24].

Now we will look more carefully at the composite factors. The decomposition of a firm's variance in a systematic and a specific part is the first out of three levels in the *Global Correlation Model*^{TM}; see Figure 1.7. The subsequent level is the decomposition of the firm's composite factor Φ in industry and country indices.

Before writing down the level-2 decomposition, let us rewrite Equation (1.22) in vector notation[18] which is more convenient for further calculations. For this purpose denote by $\boldsymbol{\beta} = (\beta_{ij})_{1\leq i,j\leq m}$ the diagonal matrix in $\mathbb{R}^{m\times m}$ with $\beta_{ij} = \beta_i$ if $i = j$ and $\beta_{ij} = 0$ if $i \neq j$. Equation (1.22) then can be rewritten in vector notation as follows:

$$\boldsymbol{r} \;=\; \boldsymbol{\beta}\boldsymbol{\Phi} \;+\; \boldsymbol{\varepsilon}\;, \tag{1.25}$$

$$\boldsymbol{\Phi}^T = (\Phi_1, ..., \Phi_m)\;, \qquad \boldsymbol{\varepsilon}^T = (\varepsilon_1, ..., \varepsilon_m)\;.$$

[18]Note that in the sequel we write vectors as column vectors.

For the second level, the *Global Correlation Model*^{TM} decomposes every Φ_i w.r.t. an *industry* and *country* breakdown,

$$\Phi_i = \sum_{k=1}^{K} w_{i,k}\Psi_k \qquad (i = 1, ..., m), \tag{1.26}$$

where $\Psi_1, ..., \Psi_{K_0}$ are industry indices and $\Psi_{K_0+1}, ..., \Psi_K$ are country indices. The coefficients $w_{i,1}, ..., w_{i,K_0}$ are called the *industry weights* and the coefficients $w_{i,K_0+1}, ..., w_{i,K}$ are called the *country weights* of counterparty i. It is assumed that $w_{i,k} \geq 0$ for all i and k, and that

$$\sum_{k=1}^{K_0} w_{i,k} \;=\; \sum_{k=K_0+1}^{K} w_{i,k} \;=\; 1 \qquad (i = 1, ..., m).$$

In vector notation, (1.25) combined with (1.26) can be written as

$$\boldsymbol{r} \;=\; \boldsymbol{\beta W \Psi} \;+\; \boldsymbol{\varepsilon}\;, \tag{1.27}$$

where $\boldsymbol{W}=(w_{i,k})_{i=1,...,m;\; k=1,...,K}$ denotes the matrix of industry and country weights for the counterparties in the portfolio, and $\boldsymbol{\Psi}^T = (\Psi_1, ..., \Psi_K)$ means the vector of industry and country indices. This constitutes the second level of the *Global Correlation Model*^{TM}.

At the third and last level, a representation by a weighted sum of *independent global factors* is constructed for representing industry and country indices,

$$\Psi_k \;=\; \sum_{n=1}^{N} b_{k,n}\Gamma_n \;+\; \delta_k \qquad (k = 1, ..., K), \tag{1.28}$$

where δ_k denotes the Ψ_k-specific residual. Such a decomposition is typically done by a *principal components analysis* (PCA) of the industry and country indices. In vector notation, (1.28) becomes

$$\boldsymbol{\Psi} \;=\; \boldsymbol{B\Gamma} \;+\; \boldsymbol{\delta} \tag{1.29}$$

where $\boldsymbol{B}=(b_{k,n})_{k=1,...,K;\; n=1,...,N}$ denotes the matrix of *industry* and *country betas*, $\boldsymbol{\Gamma}^T = (\Gamma_1, ..., \Gamma_N)$ is the global factor vector, and $\boldsymbol{\delta}^T = (\delta_1, ..., \delta_K)$ is the vector of industry and country residuals. Combining (1.27) with (1.29), we finally obtain

$$\boldsymbol{r} \;=\; \boldsymbol{\beta W}(\boldsymbol{B\Gamma} + \boldsymbol{\delta}) \;+\; \boldsymbol{\varepsilon}\;. \tag{1.30}$$

So in the *Global Correlation Model™* the vector of the portfolio's returns $\boldsymbol{r}^T = (r_1, ..., r_m)$ can conveniently be written by means of underlying factors. Note that for computational purposes Equation (1.30) is the most convenient one, because the underlying factors are independent. In contrast, for an economic interpretation and for scenario analysis one would rather prefer Equation (1.27), because the industry and country indices are easier to interpret than the global factors constructed by PCA. In fact, the industry and country indices have a clear economic meaning, whereas the global factors arising from a PCA are of synthetic type. Although they admit some vague interpretation as shown in Figure 1.7, their meaning is not as clear as is the case for the industry and country indices.

As already promised, the calculation of asset returns in the model as introduced above is straightforward now. First of all, we standardize the asset value log-returns,

$$\tilde{r}_i \;=\; \frac{r_i - \mathbb{E}[r_i]}{\sigma_i} \qquad (i = 1, ..., m)$$

where σ_i denotes the volatility of the asset value log-return of counterparty i. From Equation (1.30) we then obtain a representation of standardized log-returns,

$$\tilde{r}_i \;=\; \frac{\beta_i}{\sigma_i}\,\tilde{\Phi}_i \;+\; \frac{\tilde{\varepsilon}_i}{\sigma_i} \qquad \text{where} \qquad \mathbb{E}[\tilde{\Phi}_i] = \mathbb{E}[\tilde{\varepsilon}_i] = 0\;. \tag{1.31}$$

Now, the asset correlation between two counterparties is given by

$$\mathrm{Corr}[\tilde{r}_i, \tilde{r}_j] \;=\; \mathbb{E}\big[\tilde{r}_i\tilde{r}_j\big] \;=\; \frac{\beta_i}{\sigma_i}\frac{\beta_j}{\sigma_j}\,\mathbb{E}\big[\tilde{\Phi}_i\tilde{\Phi}_j\big] \tag{1.32}$$

because the *Global Correlation Model™* assumes the residuals $\tilde{\varepsilon}_i$ to be uncorrelated and independent of the composite factors. For calculation purposes it is convenient to get rid of the volatilities σ_i and the betas β_i in Equation (1.32). This can be achieved by replacing the betas by the R-squared parameters of the involved firms. From Equation (1.24) we know that

$$R_i^2 \;=\; \frac{\beta_i^2}{\sigma_i^2}\,\mathbb{V}[\Phi_i] \qquad (i = 1, ..., m). \tag{1.33}$$

Therefore, Equation (1.32) combined with (1.33) yields

$$\mathrm{Corr}[\tilde{r}_i, \tilde{r}_j] \;=\; \frac{R_i}{\sqrt{\mathbb{V}[\Phi_i]}}\frac{R_j}{\sqrt{\mathbb{V}[\Phi_j]}}\,\mathbb{E}\big[\tilde{\Phi}_i\tilde{\Phi}_j\big] \tag{1.34}$$

$$= \frac{R_i}{\sqrt{\mathbb{V}[\tilde{\Phi}_i]}} \frac{R_j}{\sqrt{\mathbb{V}[\tilde{\Phi}_j]}} \mathbb{E}\big[\tilde{\Phi}_i \tilde{\Phi}_j\big]$$

because by construction we have $\mathbb{V}[\Phi_i] = \mathbb{V}[\tilde{\Phi}_i]$.

Based on Equation (1.30) we can now easily compute asset correlations according to (1.34). After standardization, (1.30) changes to

$$\tilde{\boldsymbol{r}} = \tilde{\boldsymbol{\beta}} \boldsymbol{W} (\boldsymbol{B} \tilde{\boldsymbol{\Gamma}} + \tilde{\boldsymbol{\delta}}) + \tilde{\varepsilon} \; , \tag{1.35}$$

where $\tilde{\boldsymbol{\beta}} \in \mathbb{R}^{m \times m}$ denotes the matrix obtained by scaling every diagonal element in $\boldsymbol{\beta}$ by $1/\sigma_i$, and

$$\mathbb{E}\big[\tilde{\boldsymbol{\Gamma}}\big] = 0, \quad \mathbb{E}\big[\tilde{\varepsilon}\big] = 0, \quad \mathbb{E}\big[\tilde{\boldsymbol{\delta}}\big] = 0 \; .$$

Additionally, the residuals $\tilde{\boldsymbol{\delta}}$ and $\tilde{\boldsymbol{\varepsilon}}$ are assumed to be uncorrelated and independent of $\tilde{\boldsymbol{\Gamma}}$. We can now calculate asset correlations according to (1.34) just by computing the matrix

$$\mathbb{E}\big[\tilde{\boldsymbol{\Phi}} \tilde{\boldsymbol{\Phi}}^T\big] \;=\; \boldsymbol{W} \Big[\boldsymbol{B} \mathbb{E}\big[\tilde{\boldsymbol{\Gamma}} \tilde{\boldsymbol{\Gamma}}^T\big] \boldsymbol{B}^T + \mathbb{E}\big[\tilde{\boldsymbol{\delta}} \tilde{\boldsymbol{\delta}}^T\big] \Big] \boldsymbol{W}^T \tag{1.36}$$

because the matrix of standardized composite factors is given by $\tilde{\boldsymbol{\Phi}} = \boldsymbol{W}(\boldsymbol{B}\tilde{\boldsymbol{\Gamma}} + \tilde{\boldsymbol{\delta}})$. Let us quickly prove that (1.36) is true. By definition, we have

$$\begin{aligned}
\mathbb{E}\big[\tilde{\boldsymbol{\Phi}} \tilde{\boldsymbol{\Phi}}^T\big] \;&=\; \mathbb{E}\Big[\boldsymbol{W}(\boldsymbol{B}\tilde{\boldsymbol{\Gamma}} + \tilde{\boldsymbol{\delta}}) \Big(\boldsymbol{W}(\boldsymbol{B}\tilde{\boldsymbol{\Gamma}} + \tilde{\boldsymbol{\delta}}) \Big)^T \Big] \\
&=\; \boldsymbol{W}\, \mathbb{E}\Big[(\boldsymbol{B}\tilde{\boldsymbol{\Gamma}} + \tilde{\boldsymbol{\delta}})(\boldsymbol{B}\tilde{\boldsymbol{\Gamma}} + \tilde{\boldsymbol{\delta}})^T \Big] \boldsymbol{W}^T \\
&=\; \boldsymbol{W} \Big(\boldsymbol{B} \mathbb{E}\big[\tilde{\boldsymbol{\Gamma}} \tilde{\boldsymbol{\Gamma}}^T\big] \boldsymbol{B}^T + \underbrace{\boldsymbol{B} \mathbb{E}\big[\tilde{\boldsymbol{\Gamma}} \tilde{\boldsymbol{\delta}}^T\big]}_{=\,0} + \underbrace{\mathbb{E}\big[\tilde{\boldsymbol{\delta}} (\boldsymbol{B}\tilde{\boldsymbol{\Gamma}})^T\big]}_{=\,0} + \mathbb{E}\big[\tilde{\boldsymbol{\delta}} \tilde{\boldsymbol{\delta}}^T\big] \Big) \boldsymbol{W}^T \; .
\end{aligned}$$

The two expectations above vanish due to our orthogonality assumptions. This proves (1.36). Note that in equation (1.36), $\mathbb{E}\big[\tilde{\boldsymbol{\Gamma}} \tilde{\boldsymbol{\Gamma}}^T\big]$ is a diagonal matrix (because we are dealing with orthogonal global factors) with diagonal elements $\mathbb{V}[\Gamma_n]$ $(n = 1, ..., N)$, and $\mathbb{E}\big[\tilde{\boldsymbol{\delta}} \tilde{\boldsymbol{\delta}}^T\big]$ is a diagonal matrix with diagonal elements $\mathbb{V}[\delta_k]$ $(k = 1, ..., K)$. Therefore, the calculation of asset correlations according to (1.36) can conveniently be implemented in case one knows the variances of global factors, the variances of industry and country residuals, and the beta of the industry and country indices w.r.t. the global factors.

The factor model used by the RiskMetrics Group is quite similar to the *Global Correlation Model*^{TM} just described. So there is no need to start all over again, and we refer to the documentation which, in case of RiskMetrics, is called the *CreditMetrics*^{TM} Technical Document [88]. However, there are two fundamental differences between the models which are worthwhile and important to be mentioned.

First, the *Global Correlation Model*^{TM} is calibrated w.r.t. *asset* value processes, whereas the factor model of *CreditMetrics*^{TM} uses *equity* processes instead of asset value processes, thereby taking *equity correlations as a proxy for asset correlations*; see [88], page 93. We consider this difference to be fundamental, because a very important feature of the model world owned by Moody's KMV is that it really manages the admittedly difficult process of translating equity and market information into asset values; see Chapter 3.

Second, the framework *CreditMetrics*^{TM} uses indices[19] referring to a *combination* of some industry in some particular country, whereas the *Global Correlation Model*^{TM} considers industries and countries separately. So a German automotive company in the *CreditMetrics*^{TM} factor model would get a 100%-weight w.r.t. an index describing the *German automotive* industry, whereas in the *Global Correlation Model*^{TM} this company would have industry *and* country weights equal to 100% w.r.t. an *automotive index* and a *country index* representing Germany. Both approaches are quite different and have their own advantages and disadvantages.

1.3 Regulatory Capital and the Basel Initiative

It is worthwhile to mention that in the first edition of this book we started with the remark that the regulatory capital approach currently is under review. Today, eight years later, this statement is true again. The most recent crisis gave rise to uncountably many discussions on the current regulatory approach. Because regulation is an ongoing issue and frameworks are subject to change it does not make sense in a

[19]MSCI indices; see www.msci.com.

book like this to spend too much time with a topic like regulatory capital. However, it does make sense to provide at least some remarks and some flavor on how regulatory capital is calculated. We concentrate on the calculation aspects of regulatory capital. Topics like disclosure or reporting are left out intentionally. The currently valid regulatory framework can be found in the document [149] and its supplementary papers on www.bis.org. So let us start with a bit of history and then let us briefly present examples of capital formulas in the current framework.

In 1983 the banking supervision authorities of the main industrialized countries (G7) agreed on rules for banking regulation, which should be incorporated into national regulation laws. Since the national regulators discussed these issues, hosted and promoted by the *Bank of International Settlement* (www.bis.org) located in *Basel* in Switzerland, these rules were called *The Basel Capital Accord*.

The best known rule therein is the *8-percent rule.* Under this rule, banks have to prove that the capital they hold is larger than 8% of their so-called *risk-weighted assets* (RWA), calculated for all balance sheet positions. This rule implied that the capital basis for banks was mainly driven by the exposure of the loans to their customers. The RWA were calculated by a simple *weighting scheme.* Roughly speaking, for loans to any government institution the *risk weight* was set to 0%, reflecting the broad opinion that the governments of the world's industrial nations are likely to meet their financial obligations. The risk weight for loans to OECD banks was fixed at 20%. Regarding corporate loans, the committee agreed on a *standard risk weight* of 100%, no matter if the borrowing firm is a more or less risky obligor. The RWA were then calculated by adding up all of the bank's weighted credit exposures, yielding a *regulatory capital* of 8% $\times$ RWA.

The main weakness of this capital accord was that it made no distinction between obligors with different creditworthiness. In 1988 an amendment to this Basel Accord opened the door for the use of internal models to calculate the regulatory capital for off-balance sheet positions in the *trading book.* The trading book was mostly seen as containing deals bearing *market risk*, and therefore the corresponding internal models captured solely the market risk in the trading business. Still, corporate bonds and derivatives contributed to the RWA, since the default risk was not captured by the market risk models.

In 1997 the *Basel Committee on Banking Supervision* allowed the banks to use so-called *specific risk models*, and the eligible instruments no longer fell under the 8%-rule. Around that time regulators recognized that banks already internally used sophisticated models to handle the credit risk for their balance sheet positions with an emphasis on default risk. These models were quite different from the standard specific risk models. In particular, they produced a loss distribution of the entire portfolio and did not so much focus on the volatility of the spreads as in most of the specific risk models.

At the end of the 20th century, the Basel Committee started to look intensively at the models presented in this book. However, in the finally agreed regulatory framework [149], shortly called *Basel II*, they do not allow the use of internal credit risk models for the calculation of regulatory capital. Instead, they use a more or less complicated risk-weighting scheme for bank's credit risk positions. The Basel II approach was switched live in most banks worldwide on January 1st in 2007. In the sequel, we briefly outline the currently used approach.

A major improvement of Basel II compared to the former approach (shortly called Basel I) is that the new capital rules are much more risk sensitive. As already mentioned, the standard risk weight under Basel 1 was 100%, which led to a regulatory capital of

$$[\text{risk weight}] \times [\text{solvability coefficient}] \;=\; 100\% \times 8\% \;=\; 8\%$$

for various assets originated (and later often securitized) by banks. In the Basel II framework, often called the *new capital accord*, risk weights are working in the way they are supposed to work, namely, by weighting positions w.r.t. their credit risk. Depending on the level of sophistication a bank operates, the Basel II accord offers different approaches to regulatory capital. The most sophisticated approach a bank can implement is the so-called *internal ratings-based approach* (IRB). In this approach, banks calculate the risk weight of an asset in the following way[20] (see [149], §271-272):

$$\begin{aligned}
\text{RWA} \;&=\; 12.5 \times \text{EAD} \times \text{LGD} \times K(\text{PD}) \times M(\text{PD}, \text{MAT}) \\
K(\text{PD}) \;&=\; N\left[\frac{N^{-1}[\text{PD}] + \sqrt{\varrho(\text{PD})}\, q_{99.9\%}(Y)}{\sqrt{1-\varrho(\text{PD})}}\right] - \text{PD}
\end{aligned}$$

[20] Note that the function K is the quantile function (here, with respect to a confidence level of 99.9%) of the limit distribution in Formula (2.54).

$$\varrho(\mathrm{PD}) \;=\; 0.12 \times \frac{1-e^{-50\times \mathrm{PD}}}{1-e^{-50}} \;+\; 0.24 \times \left(1 - \frac{1-e^{-50\times \mathrm{PD}}}{1-e^{-50}}\right).$$

The meaning of the parameters in the formula is as follows:

- $M(\mathrm{PD}, \mathrm{MAT})$ is an adjustment factor depending on the effective maturity MAT of the asset and its PD.
- $N[\,\cdot\,]$ is the standard normal distribution function and $N^{-1}[\,\cdot\,]$ is its inverse.
- The quantity $q_{99.9\%}(Y)$ is the 99.9%-quantile of a standard normal random variable Y.
- The quantity ϱ has the meaning of a correlation parameter; see Formula (2.54).
- The formula for the correlation parameter ϱ is an interpolation between 12% and 24% quantifying the systematic risk ("R-squared"; see Formula 1.23 and the discussion thereafter) of an asset as a function of its PD.

Different asset classes get different parameterizations in the Basel II world. For instance, for SMEs[21] some firm size adjustment is applied ([149], §273-274); for retail exposures, e.g., residential mortgages or revolving retail exposures, other correlation parameterizations are prescribed ([149], §327-330), and so on. The RWA formula as presented by us refers to a *standard corporate loan.*

A good question one could ask is why the correlation parameter $\varrho = \varrho(\mathrm{PD})$ is chosen in dependence on the PD. A mathematical answer to this question is that there really are no good reasons for introducing such a functional relation between the parameters. A practitioner would probably argue that one expects on average to have better ratings for larger firms (e.g., multi-nationals like Nestle, Novartis, Deutsche Bank, IBM, etc.) and "large" often is associated with higher systematic risk so that lower PD corresponds to higher systematic risk and, therefore, a higher correlation parameter. However, reducing a two-parameter distribution model (see Formula (2.54)) to a one-parameter model by making one of the parameters a function of the other parameter is a questionable approach, like it or dislike it.

[21] Small- and medium-sized enterprises.

The Basel II approach has its strengths and weaknesses. Clearly a strength is that IRB-banks can use their internal ratings, LGDs and EADs as an input into the RWA-function. This is a huge progress in Basel II compared to Basel I. It basically means that *regulatory and economic approaches exhibit high convergence for single-name risks.*

A true weakness is that Basel II does neither penalize concentration nor award diversification. The approach is based on a simple but fully-fledged portfolio model but does not use the model itself and instead relies on a risk-weighting scheme which is *not portfolio context sensitive*. As an illustration imagine a loan in a first scenario in the credit portfolio of Deutsche Bank and in a second scenario the same loan in the credit portfolio of Credit Suisse. Although both credit books are fundamentally different the RWA-formula does not reflect those differences, the capital calculation according to Basel II is not sensitive to the surrounding credit portfolio.

This disadvantage becomes even more dramatic when one considers the way Basel II treats securitizations. Here one can safely say that Basel II simply fails to capture effects and risks of such so-called *correlation products*. We will briefly touch on the topic of *regulatory arbitrage* in structured credit products in Chapter 8.

What comes next from Basel? Nobody really knows yet. But it is certain that the accord will be revised again although the timing of a new accord and its final content are still open. However, there are many smaller changes and addenda which are published on continuous base on the website www.bis.org. Visiting this website and scanning available documentation and press releases is a must for every credit risk professional. We can only recommend to visit it.

Further Reading

As a general guide to quantitative risk management we recommend the book by McNeil, Frey and Embrechts [136]. Their book is a rich source for quantitative model techniques, not only for credit but also for other risks. Another book on risk management which contains also non-quantitative aspects is the book by Crouhy, Galai and Mark [38]. A book dealing with the integration of credit and interest rate risk is [45] by van Deventer, Imai and Mesler [45]. There are many other books on credit risk modeling available where each book has its own focus and flavor. To mention a few examples we refer to Ammann [5], Duffie and Singleton [51], Lando [120] and Schmid [161]

As a reference to Section 1.3 we refer to the Basel capital accord in its original form [149]. There are some books in the market where the Basel II standards are not only explained but also illustrated by means of examples. For this we refer to the book by Ong [151]. Another book which provides guidance regarding all quantitative aspects of the Basel II capital accord is the book by Engelmann and Rauhmeier [57].

Chapter 2

Modeling Correlated Defaults

In this chapter we will look at default models from a more abstract point of view, hereby providing a framework in which today's industry models can be embedded. Let us start with some general remarks.

Regarding random variables and probabilities we repeat our remark from the beginning of the previous chapter by saying that we always assume that an appropriate probability space $(\Omega, \mathcal{F}, \mathbb{P})$ has been chosen, reflecting the "probabilistic environment" necessary to make the respective statement.

Without loss of generality we will always assume a valuation horizon of one year. Let's say we are looking at a credit portfolio with m counterparties. Every counterparty in the portfolio admits a rating R_i as of today, and by means of some rating calibration as explained in Section 1.1.1.1 we know the default probability p_i corresponding to rating R_i. One year from today the rating of the considered counterparty may have changed due to a change in its creditworthiness. Such a rating change is called a *rating migration*. More formally we denote the range of possible ratings by $\{0, ..., d\}$, where $d \in \mathbb{N}$ means the *default state*,

$$R_i \in \{0, ..., d\} \qquad \text{and} \qquad p_i = \mathbb{P}[R_i \to d] \ ,$$

where the notation $R \to R'$ denotes a rating migration from rating R to rating R' within one year. In this chapter we will focus on a *two-state* approach, essentially meaning that we restrict ourselves to a setting where

$$d = 1, \qquad L_i = R_i \in \{0, 1\}, \qquad p_i = \mathbb{P}[L_i = 1].$$

Two-state models neglect the possibility of rating changes; only *default* or *survival* is considered. However, generalizing a two-state to a multi-state model is straightforward and will be done frequently in subsequent chapters.

In Chapter 1 we defined loss variables as indicators of default events; see Section 1.1. In the context of two-state models, an approach by means of *Bernoulli* random variables is most natural. In fact, most common industry models (e.g., the models by Moody's KMV and the RiskMetrics Group mentioned in Chapter 1) as well as most of the bank-internal models follow this approach. An alternative approach is the modeling of defaults by *Poisson* random variables. CreditRisk$^+$ (see Section 2.4.2) is based on this approach. There is a minority among the major banks who use CreditRisk$^+$ as their internal model but this does not mean that the Poisson approach is not a reasonable way to go. As we will later see the Poisson approach has its strengths and weaknesses and so does the Bernoulli approach. The question of which approach is better for a bank is based on individual judgement and depends on various criteria, for instance, the kind of business a bank is doing, the composition of their credit book, strategic preferences, conditions on the robustness of the model as well as conditions on the ease of portfolio calculations, etc.

There are attempts to bring Bernoulli and Poisson models in a common mathematical framework (see, e.g., Gordy [83] and Hickman and Koyluoglu [115]) and to some extent there are indeed relations and common roots of the two approaches; see Section 2.3. However, in [25] it is shown that the models are not really compatible, because the corresponding *mixture models* (Bernoulli respectively Poisson variables have to be *mixed* in order to introduce correlations into the models) generate loss distributions with significant tail differences. See Section 2.5.3.

Today we can access a rich literature investigating general frameworks for modeling correlated defaults and for embedding the existing industry models in a more abstract framework. See, e.g., Crouhy, Galai and Mark [37], Gordy [83], Frey and McNeil [68], and Hickman and Koyluoglu [115], just to mention a few references. For the sequel we make a notational convention. Bernoulli random variables will always be denoted by L, whereas Poisson variables will be denoted by L'. In the following section we first look at the Bernoulli[1] model, but then also turn to the case of Poissonian default variables. In Section 2.3 we briefly compare both approaches.

[1]Note that the Bernoulli model benefits from the convenient property that the mixture of Bernoulli variables again yields a Bernoulli-type random variable.

2.1 The Bernoulli Model

A vector of random variables $\boldsymbol{L} = (L_1, ..., L_m)$ is called a (Bernoulli) *loss statistics*, if all marginal distributions of $\boldsymbol{L}$ are Bernoulli:

$$L_i \sim B(1; p_i), \qquad \text{i.e.,} \qquad L_i = \begin{cases} 1 \text{ with probability } p_i \\ 0 \text{ with probability } 1 - p_i \end{cases} .$$

The *loss* resp. *percentage loss* of $\boldsymbol{L}$ is defined[2] as

$$L = \sum_{i=1}^{m} L_i \qquad \text{resp.} \qquad \frac{L}{m} .$$

The probabilities $p_i = \mathbb{P}[L_i = 1]$ are called *default probabilities* of $\boldsymbol{L}$.

The reasoning underlying our terminology is as follows:

According to Definition 1.2.3, a credit portfolio is nothing but a collection of, say, m, transactions or deals with certain counterparties or, more general, a collection of m credit-risky assets. Whatever is said in the sequel applies to both types of credit risks and we use "counterparty" or "asset" as exchangeable keywords. Every counterparty involved creates basically (in a two-state model) two future scenarios: Either the counterparty defaults[3], or the counterparty survives[4]. In the case of default of obligor i the indicator variable L_i equals 1; in the case of survival we have $L_i = 0$. In this way, every portfolio generates a natural loss statistics w.r.t. the particular valuation horizon (here, one year). The variable L defined above is then called the *portfolio loss*, no matter if quoted as an absolute or percentage value.

Before we come to more interesting cases we should for the sake of completeness briefly discuss the quite unrealistic case of independent defaults.

[2]Note that in the sequel we sometimes write L for denoting the *gross* loss as well as the *percentage* loss of a loss statistics. But from the context the particular meaning of L will always be clear.

[3]Note that there exist various default definitions in the banking world; as long as nothing different is said, we always mean by default a *payment default on any financial obligation*.

[4]Meets the financial expectations of the bank regarding contractually promised cash flows.

The simplest type of a loss statistic can be obtained by assuming a *uniform default probability* p and the lack of dependence between counterparties. More precisely, under these assumptions we have

$$L_i \sim B(1;p) \qquad \text{and} \qquad (L_i)_{i=1,\dots,m} \quad \text{independent.}$$

In this case, the absolute portfolio loss L is a *convolution* of i.i.d. Bernoulli variables and therefore follows a *binomial* distribution with parameters m and p, $L \sim B(m;p)$.

If the counterparties are still assumed to be independent, but this time admitting different default probabilities,

$$L_i \sim B(1;p_i) \qquad \text{and} \qquad (L_i)_{i=1,\dots,m} \quad \text{independent,}$$

we again obtain the portfolio loss L as a convolution of the single loss variables, but this time with first and second moments

$$\mathbb{E}[L] = \sum_{i=1}^m p_i \qquad \text{and} \qquad \mathbb{V}[L] = \sum_{i=1}^m p_i(1-p_i)\ . \tag{2.1}$$

This follows from $\mathbb{E}[L_i] = p_i$, $\mathbb{V}[L_i] = p_i(1-p_i)$, and the additivity of expectations resp. variances[5].

Now, it is well known that in probability theory *independence* makes things easy. For example the *strong law of large numbers* works well with independent variables and the *central limit theorem* in its most basic version lives from the assumption of independence. If in credit risk management we could assume independence between counterparties in a portfolio, we could – due to the central limit theorem – assume that the portfolio loss (approximable) is a Gaussian variable, at least for large portfolios. In other words, we would never be forced to work with Monte Carlo simulations, because the portfolio loss would conveniently be given in a closed (namely Gaussian) form with well-known properties.

Unfortunately in credit risk modeling we can not expect to find independency of losses. Moreover, it will turn out that *correlation is the central challenge* in credit portfolio risk. Therefore, we turn now to more realistic elaborations of loss statistics.

[5]For having additivity of variances it would be sufficient that the involved random variables are pairwise uncorrelated and integrable (see [12], Chapter 8).

One basic idea for modeling correlated defaults (by mixing) is the randomization of the involved default probabilities in a correlated manner. We start with a so-called standard *binary mixture model*; see JOE [107] for an introduction to this topic.

2.1.1 A General Bernoulli Mixture Model

Following our basic terminology, we obtain the loss of a portfolio from a loss statistics $\boldsymbol{L} = (L_1, ..., L_m)$ with Bernoulli variables $L_i \sim B(1; P_i)$. But now we think of the loss probabilities as random variables $\boldsymbol{P} = (P_1, ..., P_m) \sim \boldsymbol{F}$ with some distribution function $\boldsymbol{F}$ with support in $[0,1]^m$. Additionally, we assume that conditional on a realization $\boldsymbol{p} = (p_1, ..., p_m)$ of $\boldsymbol{P}$ the variables $L_1, ..., L_m$ are independent. In more mathematical terms we express the *conditional independence* of the losses by writing

$$L_i|_{P_i=p_i} \sim B(1; p_i), \qquad (L_i|_{\boldsymbol{P}=\boldsymbol{p}})_{i=1,...,m} \quad \text{independent.}$$

The (unconditional) joint distribution of the L_i's is then determined by the probabilities

$$\mathbb{P}[L_1 = l_1, ..., L_m = l_m] \tag{2.2}$$

$$= \int_{[0,1]^m} \prod_{i=1}^{m} p_i^{l_i}(1-p_i)^{1-l_i} d\boldsymbol{F}(p_1, ..., p_m)\,,$$

where $l_i \in \{0, 1\}$. The first and second moments of the single losses L_i $(i = 1, ..., m)$ are given by

$$\mathbb{E}[L_i] \;=\; \mathbb{E}[P_i], \qquad \mathbb{V}[L_i] \;=\; \mathbb{E}[P_i]\,(1 - \mathbb{E}[P_i]) \tag{2.3}$$

The first equality is obvious from (2.2). The second identity can be seen as follows:

$$\begin{aligned} \mathbb{V}[L_i] &= \mathbb{V}\big[\mathbb{E}[L_i|\boldsymbol{P}]\big] + \mathbb{E}\big[\mathbb{V}[L_i|\boldsymbol{P}]\big] \\ &= \mathbb{V}[P_i] + \mathbb{E}[P_i(1-P_i)] \\ &= \mathbb{E}[P_i]\,(1-\mathbb{E}[P_i])\,, \end{aligned} \tag{2.4}$$

Note that the first line of the equation for $\mathbb{V}[L_i]$ is the well-known *law of total variance* which is a variance decomposition formula based on conditioning. Readers not familiar with this formula can look it up in any book on probability. The covariance between single losses equals

$$\mathrm{Cov}[L_i, L_j] = \mathbb{E}[L_i L_j] - \mathbb{E}[L_i]\mathbb{E}[L_j] = \mathrm{Cov}[P_i, P_j] \ . \tag{2.5}$$

Therefore, the *default correlation* in a Bernoulli mixture model is

$$\mathrm{Corr}[L_i, L_j] \ = \ \frac{\mathrm{Cov}[P_i, P_j]}{\sqrt{\mathbb{E}[P_i]\,(1 - \mathbb{E}[P_i])}\sqrt{\mathbb{E}[P_j]\,(1 - \mathbb{E}[P_j])}} \ . \tag{2.6}$$

Equation (2.5) respectively and Equation (2.6) show that the dependence between losses in the portfolio is fully captured by the covariance structure of the multivariate distribution $\boldsymbol{F}$ of $\boldsymbol{P}$. Section 2.4 presents some examples for a meaningful specification of $\boldsymbol{F}$.

2.1.2 Uniform Default Probability and Uniform Correlation

For portfolios where all exposures are of approximately the same size and type in terms of risk, it makes sense to assume a uniform default probability and a uniform correlation among transactions in the portfolio. As already mentioned in Section 1.2.2.2, retail portfolios and some portfolios of smaller banks are often of a quite homogeneous structure, such that the assumption of a uniform default probability and a simple correlation structure does not harm the outcome of calculations with such a model. In the literature, portfolios with uniform default probability and uniform default correlation are called *uniform portfolios*. Uniform portfolio models generate perfect candidates for analytical approximations. For example, the distributions in Section 2.5 establish a typical family of two-parameter loss distributions used for analytical approximations.

The assumption of *uniformity* yields *exchangeable*[6] Bernoulli variables $L_i \sim B(1; P)$ with a random default probability $P \sim F$, where F is a distribution function with support in $[0, 1]$. We assume conditional independence of the L_i's just as in the general case. The joint distribution of the L_i's is then determined by the probabilities

$$\mathbb{P}[L_1 = l_1, ..., L_m = l_m] \ = \ \int_0^1 p^k (1-p)^{m-k} dF(p), \tag{2.7}$$

[6]That is, $(L_1, ..., L_m) \sim (L_{\pi(1)}, ..., L_{\pi(m)})$ for any permutation π.

$$\text{where} \qquad k = \sum_{i=1}^{m} l_i \quad \text{and} \qquad l_i \in \{0,1\}.$$

The probability that exactly k defaults occur is given by

$$\mathbb{P}[L = k] \;=\; \binom{m}{k} \int_0^1 p^k (1-p)^{m-k} dF(p)\;. \tag{2.8}$$

Of course, Equations (2.3) and (2.6) have their counterparts in this special case of Bernoulli mixtures: The uniform default probability of borrowers in the portfolio obviously equals

$$\overline{p} \;=\; \mathbb{P}[L_i = 1] \;=\; \mathbb{E}[L_i] \;=\; \int_0^1 p\, dF(p) \tag{2.9}$$

and the uniform default correlation of two different counterparties is given by

$$\begin{aligned} \rho \;&=\; \mathrm{Corr}[L_i, L_j] \\ &=\; \frac{\mathbb{P}[L_i = 1, L_j = 1] - \overline{p}^2}{\overline{p}(1-\overline{p})} \\ &=\; \frac{\int_0^1 p^2 dF(p) - \overline{p}^2}{\overline{p}(1-\overline{p})}.. \end{aligned} \tag{2.10}$$

Note that in the course of this book we typically use "ρ" to denote *default* correlations and "ϱ" for denoting *asset* correlations.

We now want to briefly discuss some immediate consequences of Equation (2.10). First of all it implies that

$$\mathrm{Corr}[L_i, L_j] \;=\; \frac{\mathbb{V}[P]}{\overline{p}(1-\overline{p})} \qquad (\text{recall: } P \sim F).$$

This shows that the higher the volatility of P, the higher the default correlation inherent in the corresponding Bernoulli loss statistics. Additionally, it implies that the dependence between the L_i's is either positive or zero, because variances are nonnegative. In other words, in this model we can not implement some negative dependencies between the default risks of obligors.

The case $\mathrm{Corr}[L_i, L_j] = 0$ happens if and only if the variance of F vanishes to zero, essentially meaning that there is no randomness at all regarding P. In such a case, F is a *Dirac measure* $\varepsilon_{\overline{p}}$, concentrated in $\overline{p}$, and the absolute portfolio loss L follows a binomial distribution with default probability $\overline{p}$.

The other extreme case regarding (2.10), $\text{Corr}[L_i, L_j] = 1$, implies a "rigid" behavior of single losses in the portfolio: Either all counterparties default or all counterparties survive simultaneously. The corresponding distribution F of P is then a Bernoulli distribution, such that $P = 1$ with probability $\overline{p}$ and $P = 0$ with probability $1-\overline{p}$. This means that sometimes (such events occur with probability $\overline{p}$), all counterparties default and the total portfolio exposure is lost. In other scenarios (occurring with probability $1-\overline{p}$), all obligors survive and not even one dollar is lost. The rigidity of loss statistics is 'perfect' in this situation.

Realistic scenarios live somewhere between the two discussed extreme cases $\text{Corr}[L_i, L_j] = 0$ and $\text{Corr}[L_i, L_j] = 1$.

2.2 The Poisson Model

In the case of the Poisson approach, defaults of counterparties $i = 1, ..., m$ are modeled by Poisson-distributed random variables

$$L'_i \sim Pois(\lambda_i), \qquad L'_i \in \{0, 1, 2, ...\}, \qquad p_i = \mathbb{P}[L'_i \geq 1] \ , \tag{2.11}$$

where p_i again denotes the default probability of obligor i. Note that (2.11) allows for multiple defaults of a single obligor. The likelihood of the event that obligor i defaults more than once is given by

$$\mathbb{P}[L'_i \geq 2] \ = \ 1 - e^{-\lambda_i}(1 + \lambda_i) \ ,$$

which is typically a small number. For example, in the case of $\lambda_i = 0.01$ we would obtain $\mathbb{P}[L'_i \geq 2] = 0.5$ basispoints. In other words, when simulating a Poisson-distributed default variable with $\lambda_i = 0.01$ we can expect that only 1 out of 20,000 scenarios is not applicable because of a multiple default. On the other side, for obligors with good credit quality (for example, a AAA-borrower with a default probability of 2 basispoints), a multiple-default probability of 0.5 basispoints is a relatively high number.

The intensity λ_i is typically quite close to the default probability p_i which is due to

$$p_i = \mathbb{P}[L'_i \geq 1] = 1 - e^{-\lambda_i} \approx \lambda_i \tag{2.12}$$

for small values of λ_i. Equation (2.12) shows that the one-year default probability equals the probability that an *exponential waiting time* with intensity λ_i takes place in the first year.

In general, the sum of independent variables $L'_1 \sim Pois(\lambda_1)$, $L'_2 \sim Pois(\lambda_2)$ has distribution[7] $Pois(\lambda_1 + \lambda_2)$. Assuming independence, the portfolio's total number of losses would be given by

$$L' \;=\; \sum_{i=1}^{m} L'_i \;\sim\; Pois\Big(\sum_{i=1}^{m} \lambda_i\Big)\ . \tag{2.13}$$

Correlation is introduced into the model by again following a mixture approach, this time with Poisson variables (see Joe [107], Section 7.2).

2.2.1 A General Poisson Mixture Model

Now the loss statistics is a random vector $\boldsymbol{L}' = (L'_1, ..., L'_m)$ of Poisson random variables $L'_i \sim Pois(\Lambda_i)$, where $\boldsymbol{\Lambda} = (\Lambda_1, ..., \Lambda_m)$ is a random vector with some distribution function $\boldsymbol{F}$ with support in $[0,\infty)^m$. Additionally, we assume that conditional on a realization $\boldsymbol{\lambda} = (\lambda_1, ..., \lambda_m)$ of $\boldsymbol{\Lambda}$ the variables $L'_1, ..., L'_m$ are independent:

$$L'_i|_{\Lambda_i=\lambda_i} \sim Pois(\lambda_i), \qquad (L'_i|_{\boldsymbol{\Lambda}=\boldsymbol{\lambda}})_{i=1,\ldots,m} \quad \text{independent}\ .$$

The (unconditional) joint distribution of the variables L'_i is given by

$$\mathbb{P}[L'_1 = l'_1, ..., L'_m = l'_m] \tag{2.14}$$

$$= \int_{[0,\infty)^m} e^{-(\lambda_1+\cdots+\lambda_m)} \prod_{i=1}^{m} \frac{\lambda_i^{l'_i}}{l'_i\,!}\, d\boldsymbol{F}(\lambda_1, ..., \lambda_m)\ ,$$

where $l'_i \in \{0, 1, 2, ...\}$. Analogously to the Bernoulli case we obtain

$$\mathbb{E}[L'_i] \;=\; \mathbb{E}[\Lambda_i] \qquad (i = 1, ..., m) \tag{2.15}$$

for the expectation. For the variance we get

$$\begin{aligned} \mathbb{V}[L'_i] \;&=\; \mathbb{V}\big[\mathbb{E}[L'_i|\boldsymbol{\Lambda}]\big] + \mathbb{E}\big[\mathbb{V}[L'_i|\boldsymbol{\Lambda}]\big] \\ &=\; \mathbb{V}[\Lambda_i] + \mathbb{E}[\Lambda_i]. \end{aligned}$$

[7] More generally, $(Pois(\lambda))_{\lambda\ge 0}$ is a *convolution semigroup*; see, e.g., [12].

Again we have $\text{Cov}[L'_i, L'_j] = \text{Cov}[\Lambda_i, \Lambda_j]$, and the correlation between defaults is given by

$$\text{Corr}[L'_i, L'_j] \;=\; \frac{\text{Cov}[\Lambda_i, \Lambda_j]}{\sqrt{\mathbb{V}[\Lambda_i] + \mathbb{E}[\Lambda_i]}\sqrt{\mathbb{V}[\Lambda_j] + \mathbb{E}[\Lambda_j]}} \; . \tag{2.16}$$

In the same manner as in the Bernoulli model this shows that correlation is exclusively induced by means of the distribution function $\boldsymbol{F}$ of the *random intensity vector* $\boldsymbol{\Lambda}$.

2.2.2 Uniform Default Intensity and Uniform Correlation

Analogously to the Bernoulli model, one can introduce a *Poisson uniform portfolio model* by restriction to one uniform intensity and one uniform correlation among transactions in the portfolio. More explicitly, the uniform portfolio model in the Poisson case is given by Poisson variables $L'_i \sim Pois(\Lambda)$ with a random intensity $\Lambda \sim F$, where F is a distribution function with support in $[0, \infty)$, and the L_i's are assumed to be conditionally independent. The joint distribution of the L_i's is given by

$$\mathbb{P}[L'_1 = l'_1, ..., L'_m = l'_m] \;=\; \int_0^\infty e^{-m\lambda}\, \frac{\lambda^{(l'_1 + \cdots + l'_m)}}{l'_1! \cdots l'_m!}\, dF(\lambda) \; . \tag{2.17}$$

Because (see the beginning of Section 2.2) conditional on $\Lambda = \lambda$ the portfolio loss is again a Poisson distribution with intensity $m\lambda$, the probability of exactly k defaults equals

$$\begin{aligned} \mathbb{P}[L' = k] \;&=\; \int_0^\infty \mathbb{P}[L' = k \mid \Lambda = \lambda]\, dF(\lambda) \qquad (2.18) \\ &=\; \int_0^\infty e^{-m\lambda}\, \frac{m^k \lambda^k}{k\,!}\, dF(\lambda).. \end{aligned}$$

Again, note that due to the unbounded support of the Poisson distribution the absolute loss L' can exceed the number of “physically” possible defaults. As already mentioned at the beginning of this section, the probability of a multiple-defaults event is small for typical parametrizations. In the Poisson framework, the uniform default probability of borrowers in the portfolio is defined by

$$\begin{aligned}\overline{p} &= \mathbb{P}[L'_i \geq 1] \\ &= \int_0^\infty \mathbb{P}[L'_i \geq 1 \mid \Lambda = \lambda]\, dF(\lambda) \\ &= \int_0^\infty (1 - e^{-\lambda})\, dF(\lambda).\end{aligned} \tag{2.19}$$

The counterpart of Formula (2.16) is

$$\mathrm{Corr}[L'_i, L'_j] = \frac{\mathbb{V}[\Lambda]}{\mathbb{V}[\Lambda] + \mathbb{E}[\Lambda]} \qquad (i \neq j). \tag{2.20}$$

Formula (2.20) is especially intuitive if seen in the context of *dispersion*, where the dispersion of a distribution is its *variance to mean ratio*

$$D_X = \frac{\mathbb{V}[X]}{\mathbb{E}[X]} \qquad \text{for any random variable } X. \tag{2.21}$$

The dispersion of the Poisson distribution is equal to 1. Therefore, the Poisson distribution is kind of a benchmark when deciding about *overdispersion* ($D_X > 1$) respectively *underdispersion* ($D_X < 1$). In general, nondegenerate[8] *Poisson mixtures are overdispersed* due to (2.15). This is a very important property of Poisson mixtures, because before using such a model for credit risk measurement one has to make sure that overdispersion can be observed in the data underlying the calibration of the model. Formula (2.20) can be interpreted by saying that the correlation between the number of defaults of different counterparties increases with the dispersion of the random intensity Λ. For proving this statement we write Formula (2.20) in the form

$$\mathrm{Corr}[L'_i, L'_j] = \frac{D_\Lambda}{D_\Lambda + 1} \qquad (i \neq j). \tag{2.22}$$

From (2.22) it follows that *an increase in dispersion increases the mixture effect, which, in turn, strengthens the dependence between obligor's defaults.*

[8]The random intensity Λ is not concentrated in a single point, $P_\Lambda \neq \varepsilon_\lambda$.

2.3 Bernoulli versus Poisson Mixture

The *law of small numbers*[9] implies that for large m and small p

$$B(m;p) \;\approx\; Pois(pm).$$

Setting $\lambda = pm$, this shows that under the assumption of independent defaults the portfolio absolute gross loss $L = \sum L_i$ of a Bernoulli loss statistics $(L_1, ..., L_m)$ with a uniform default probability p can be approximated by a Poisson variable $L' \sim Pois(\lambda)$. But the law of small numbers is by no means an argument strong enough to support the unfortunately widespread opinion that Bernoulli and Poisson approaches are more or less compatible. In order to show that both approaches have significant systematic differences, we turn back to the default correlations induced by the models; see (2.6), combined with (2.4), and (2.16). In the Bernoulli case we have

$$\mathrm{Corr}[L_i, L_j] \;= \tag{2.23}$$

$$= \;\frac{\mathrm{Cov}[P_i, P_j]}{\sqrt{\mathbb{V}[P_i] + \mathbb{E}[P_i(1-P_i)]}\sqrt{\mathbb{V}[P_j] + \mathbb{E}[P_j(1-P_j)]}}\;,$$

whereas in the Poisson case we obtain

$$\mathrm{Corr}[L'_i, L'_j] \;=\; \frac{\mathrm{Cov}[\Lambda_i, \Lambda_j]}{\sqrt{\mathbb{V}[\Lambda_i] + \mathbb{E}[\Lambda_i]}\sqrt{\mathbb{V}[\Lambda_j] + \mathbb{E}[\Lambda_j]}}\;. \tag{2.24}$$

Looking only at the driving random variables P_i, P_j respectively Λ_i, Λ_j, we see that in the denominators of (2.23) and (2.24) we compare

$$\mathbb{V}[P_i] + \mathbb{E}[P_i(1-P_i)] = \mathbb{V}[P_i] + \mathbb{E}[P_i] - \mathbb{E}[P_i^2] \tag{2.25}$$

$$\text{with} \qquad \mathbb{V}[\Lambda_i] + \mathbb{E}[\Lambda_i]\;.$$

Now, analogous to the deterministic case (2.12), we can – even in the random case – expect P_i and Λ_i to be of the same order of magnitude. To keep things simple, let us for a moment assume that P_i and Λ_i have the same first and second moments. In this case Equation (2.25) combined with (2.23) and (2.24) shows that the Bernoulli model always induces a higher default correlation than the Poisson model. But

[9] That is, approximation of binomial distributions by means of Poisson distributions.

higher default correlations result in fatter tails of the corresponding loss distributions. In other words, one could say that given equal first and second moments of P_i and Λ_i, the expectations of L_i and L_i' will match, but the variance of L_i' will always exceed the variance of L_i, thereby inducing lower default correlations.

So there is a *systematic* difference between the Bernoulli and Poisson mixture models. In general one can expect that for a given portfolio the Bernoulli model yields a loss distribution with a fatter tail than a comparably (e.g., by a first and second moment matching) calibrated Poisson model. In Section 2.5.3 we come back to this issue.

2.4 An Overview of Common Model Concepts

In over a decade of time, various industry models for measuring credit portfolio risk have been developed. In the first edition of this book we spent some time on explaining the differences of industry models as they come "off-the-shelf". Today most banks have their *internal model* implemented and do not solely rely on an externally developed model framework for the valuation of their credit book. Moreover, in Section 1.3 we outlined "Basel II" as the current regulatory regime banks have to comply with today. Practical experience makes very clear that for regulators not only the "standalone quality" of a model but also its embedding into the bank and its systems are major criteria. In case of large and sophisticated banks, off-the-shelf models are often mainly used for benchmarking internal against external model calculations. When dealing with the question "make or buy" banks often find out that the efforts to embed an external bought credit risk model are not significantly lower than just developing their own model right within the bank's system environment, tailor-made for the situation of the bank. For these reasons we changed the presentation of this section in this new edition and focus now more on the concepts underlying common models rather than on the way they come to the client off-the-shelf.

However, for most of the industry models it is easy to find some kind of technical documentation describing the mathematical framework of the model and giving some idea about the underlying data and the

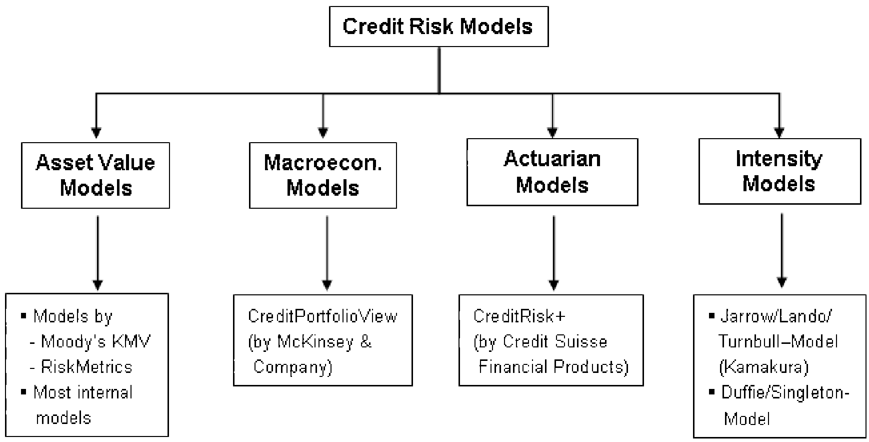


FIGURE 2.1: Common industry models.

calibration of the model to the data. Readers interested in such details can approach the company behind the model. It is our experience over the last 10+ years that companies like Moody's KMV, RiskMetrics or Kamakura will do everything to be helpful in providing support ranging from technical papers to tailor-made inhouse training and education. It is in the interest of these companies to communicate their models to potentially new users and very often the quant team inside the bank and the quant specialists working for the model company get along with each other very well and have mutual benefit from sharing their insights. Before presenting four common but different concepts for modeling credit risk we want to make some brief upfront remarks. Figure 2.1 shows the four main types of industry models and indicates the companies behind them.

CreditRisk$^+$ could alternatively be placed in the group of intensity models, because it is based on a Poisson mixture model incorporating random intensities. Nevertheless, in Figure 2.1 we prefer to stress the difference between CreditRisk$^+$ and the *dynamic* intensity models, based on intensity *processes* instead of on a *static* intensity.

Dynamic intensity models will be briefly discussed in Section 2.4.4 and to some extent in the context of *securitizations*. From a mathematician's point of view they provide a 'mathematically beautiful' approach to credit risk modeling, but from the introductory point of view we adopted for writing this book, we must say that an appropriate presentation of dynamic intensity models is beyond the scope of the book. We therefore decided to provide the reader only with some references to the literature combined with introductory remarks about the approach; see Section 2.4.4

2.4.1 Moody's KMV's and RiskMetrics' Model Approach

In Section 1.2.3 we already presented some material regarding the topic of this section. Note that in our presentation we focus on the "default-only mode", hereby ignoring the fact that the two models incorporate a *mark-to-model* approach. In the default-only mode, both models are of Bernoulli type, deciding about default or survival of a firm by comparing the firm's asset value at a certain horizon with some critical threshold. If the firm value at the horizon is below this threshold, then the firm is considered to be in default. If the firm value is above the threshold, the firm survived the considered time period. In more mathematical terms, for m counterparties denote their asset value at the considered valuation horizon $t = T$ by $A_T^{(i)}$. It is assumed that for every company i there is a critical threshold C_i such that the firm defaults in the period $[0, T]$ if and only if $A_T^{(i)} < C_i$. In the framework of Bernoulli loss statistics A_T can be viewed as a latent variable driving the default event. This is realized by defining

$$L_i \;=\; \mathbf{1}_{\{A_T^{(i)} < C_i\}} \sim B\left(1; \mathbb{P}[A_T^{(i)} < C_i]\right) \qquad (i = 1, ..., m). \tag{2.26}$$

In both models it is assumed that the asset value process is dependent on underlying factors reflecting industrial and regional influences, thereby driving the economic future of the firm. For the convenience of the reader we now recall some formulas from Section 1.2.3. The parametrization w.r.t. underlying factors typically is implemented at the standardized[10] log-return level, i.e., the asset value log-returns $\log(A_T^{(i)}/A_0^{(i)})$ after standardization admit a representation[11]

$$r_i \;=\; R_i \Phi_i + \varepsilon_i \qquad (i = 1, ..., m). \tag{2.27}$$

[10] Shifted and scaled in order to obtain a random variable with mean zero and standard deviation one.

[11] Note that for reasons of a simpler notation we here write r_i for the standardized log-returns, in contrast to the notation in Section 1.2.3, where we wrote $\tilde{r}_i$.

Here R_i is defined as in (1.33), Φ_i denotes the firm's *composite factor*, and ε_i is the *firm-specific effect* or (as it is also often called) the *idiosyncratic* part of the firm's asset value log-return. In both models, the factor Φ_i is a *superposition* of many different *industry* and *country* indices. *Asset correlations* between counterparties are exclusively captured by the correlation between the respective composite factors. The specific effects are assumed to be independent among different firms and independent of the composite factors. The quantity R_i^2 reflects how much of the volatility of r_i can be explained by the volatility of the composite factor Φ_i. Because the composite factor is a superposition of *systematic* influences, namely industry and country indices, R_i^2 quantifies the *systematic risk* of counterparty i.

In both models the basic assumption is that asset value log-returns are assumed to be normally distributed although we already mentioned that in the real-life implementation of the Moody's KMV model one finds empirically best-fitting distributions beyond the Gaussian world. The distribution assumption just made and standardization yield

$$r_i \sim N(0,1), \qquad \Phi_i \sim N(0,1), \qquad \text{and} \qquad \varepsilon_i \sim N\big(0, 1 - R_i^2\big)\,.$$

We are now in a position to rewrite (2.26) in the following form:

$$L_i \;=\; \mathbf{1}_{\{r_i < c_i\}} \sim B\,(1; \mathbb{P}[r_i < c_i]) \qquad (i = 1, ..., m), \tag{2.28}$$

where c_i is the threshold corresponding to C_i after exchanging $A_T^{(i)}$ by r_i. Applying (2.27), the condition $r_i < c_i$ can be written as

$$\varepsilon_i \;<\; c_i - R_i \Phi_i \qquad (i = 1, ..., m). \tag{2.29}$$

Now, in both models, the standard valuation horizon is $T = 1$ year. Denoting the one-year default probability of obligor i by p_i, we naturally have $p_i = \mathbb{P}[r_i < c_i]$. Because $r_i \sim N(0,1)$ we immediately obtain

$$c_i \;=\; N^{-1}[p_i] \qquad (i = 1, ..., m), \tag{2.30}$$

where $N[\cdot]$ denotes the cumulative standard normal distribution function. Scaling the idiosyncratic component towards a standard deviation of one, this changes (2.29) into

$$\tilde{\varepsilon}_i \;<\; \frac{N^{-1}[p_i] - R_i \Phi_i}{\sqrt{1 - R_i^2}}, \qquad \tilde{\varepsilon}_i \sim N(0,1). \tag{2.31}$$

Taking into account that $\tilde{\varepsilon}_i \sim N(0,1)$, we altogether obtain the following representation for the one-year default probability of obligor i conditional on the factor Φ_i:

$$p_i(\Phi_i) \;=\; N\left[\frac{N^{-1}[p_i] - R_i\Phi_i}{\sqrt{1-R_i^2}}\right] \qquad (i=1,...,m). \tag{2.32}$$

The only random part of (2.32) is the composite factor Φ_i. Conditional on $\Phi_i = z$, we obtain the conditional one-year default probability by

$$p_i(z) \;=\; N\left[\frac{N^{-1}[p_i] - R_i z}{\sqrt{1-R_i^2}}\right]. \tag{2.33}$$

Combined with (2.28) this shows that we are in a Bernoulli mixture setting exactly the same way as elaborated in Section 2.1.1. More formally we can – analogously to (2.2) – specify the portfolio loss distribution by the probabilities (here we assume again $l_i \in \{0,1\}$)

$$\begin{aligned} &\mathbb{P}[L_1 = l_1, ..., L_m = l_m] \qquad (2.34)\\ &= \int_{[0,1]^m} \prod_{i=1}^{m} q_i^{l_i}(1-q_i)^{1-l_i}\, d\boldsymbol{F}(q_1,...,q_m), \end{aligned}$$

where the distribution function $\boldsymbol{F}$ is now explicitly given by

$$\boldsymbol{F}(q_1,...,q_m) \;=\; N_m\left[p_1^{-1}(q_1),...,p_m^{-1}(q_m);\Gamma\right], \tag{2.35}$$

where $N_m[\,\cdot\,;\Gamma]$ denotes the cumulative multivariate centered Gaussian distribution with correlation matrix Γ, and $\Gamma = (\varrho_{ij})_{1\le i,j\le m}$ means the asset correlation matrix of the log-returns r_i according to (2.27).

In the case that the composite factors are represented by a weighted sum of industry and country indices $(\Psi_j)_{j=1,...,J}$ of the form

$$\Phi_i \;=\; \sum_{j=1}^{J} w_{ij}\Psi_j \tag{2.36}$$

(see Section 1.2.3), the conditional default probabilities (2.33) equal

$$p_i(z) \;=\; N\left[\frac{N^{-1}[p_i] - R_i(w_{i1}\psi_1 + \cdots + w_{iJ}\psi_J)}{\sqrt{1-R_i^2}}\right], \tag{2.37}$$

with industry and country index realizations $(\psi_j)_{j=1,...,J}$. By varying these realizations and then recalculating the conditional probabilities (2.37) one can perform simple *scenario stress testing*, in order to study the impact of certain changes of industry or country indices on the default probability of some obligor.

2.4.2 Model Approach of CreditRisk$^+$

CreditRisk$^+$ is a credit risk model developed by Credit Suisse Financial Products (CSFP). It is more or less based on a typical insurance mathematics approach, which is the reason for its classification as an *actuarial model*. Regarding its mathematical background, the main reference is the CreditRisk$^+$ Technical Document [35]. In light of this chapter one could say that CreditRisk$^+$ is a typical representative of the group of Poisson mixture models. In this paragraph we only summarize the model, focusing on *defaults only* and not on *losses* in terms of money, but in Chapter 4 a more comprehensive introduction (taking exposure distributions into account) is presented.

The mixture distribution adopted by CreditRisk$^+$ incorporates the *gamma distribution*. Recall that the gamma distribution is defined by the probability density

$$\gamma_{\alpha,\beta}(x) = \frac{1}{\beta^\alpha \Gamma(\alpha)}\, e^{-x/\beta} x^{\alpha-1} \qquad (x \geq 0),$$

where $\Gamma(\cdot)$ denotes[12] the gamma function. The first and second moments of a gamma-distributed random variable Λ are

$$\mathbb{E}[\Lambda] = \alpha\beta, \qquad \mathbb{V}[\Lambda] = \alpha\beta^2\ ; \tag{2.38}$$

see Figure 2.2 for an illustration of gamma densities.

Instead of incorporating a factor model (as we have seen it in the case of the models by Moody's KMV and RiskMetrics in Section 1.2.3), CreditRisk$^+$ implements a so-called *sector model*. However, somehow one can think of a sector as a 'factor-inducing' entity, or – as the CreditRisk$^+$ Technical Document [35] says it – every sector could be

[12]We will also write $X \sim \Gamma(\alpha, \beta)$ for any gamma-distributed random variable X with parameters α and β. Additionally, we use Γ to denote the correlation matrix of a multivariate normal distribution. However, it should be clear from the context which current meaning the symbol Γ has.

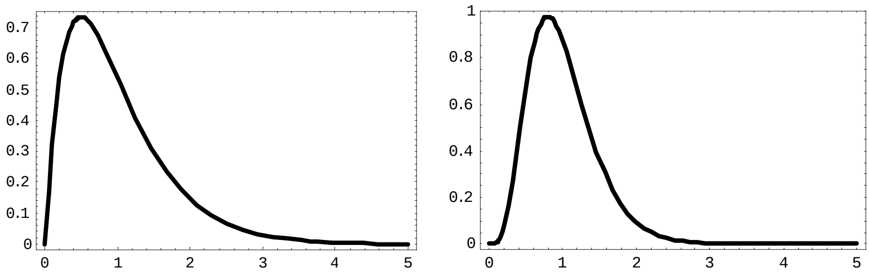


FIGURE 2.2: Shape of gamma distributions for parameters $(\alpha, \beta) \in \{(2, 1/2), (5, 1/5)\}$.

thought of as generated by a single underlying factor. In this way, sectors and factors are somehow comparable objects. From an interpretational point of view, sectors can be identified with industries, countries, or regions, or any other systematic influence on the economic performance of counterparties with a positive weight in this sector. Each sector $s \in \{1, ..., m_S\}$ has its own *gamma-distributed random intensity* $\Lambda^{(s)} \sim \Gamma(\alpha_s, \beta_s)$, where the variables $\Lambda^{(1)}, ..., \Lambda^{(m_S)}$ are assumed to be *independent*.

Now let us assume that a credit portfolio of m loans to m different obligors is given. In the *sector model* of CreditRisk$^+$, every obligor i admits a breakdown into sector weights $w_{is} \geq 0$ with $\sum_{s=1}^{m_S} w_{is} = 1$, such that w_{is} reflects the sensitivity of the default intensity of obligor i to the systematic default risk arising from sector s. The risk of sector s is captured by two parameters: The first driver is the *mean default intensity* of the sector,

$$\lambda_{(s)} \;=\; \mathbb{E}[\Lambda^{(s)}] \;=\; \alpha_s \beta_s \; ;$$

see also (4.18) in Chapter 4. The second driver is the *default intensity's volatility*

$$\sigma_{(s)} \;=\; \mathbb{V}[\Lambda^{(s)}] \;=\; \alpha_s \beta_s^2 \; .$$

In Section 4.3.2 we indicate some possible approaches for calibrating the sector parameters $\lambda_{(s)}$ and $\sigma_{(s)}$. Every obligor i admits a random default intensity Λ_i with mean value $\mathbb{E}[\Lambda_i] = \lambda_i$, which could be calibrated to the obligor's one-year default probability by means of Formula

(2.12). The sector parametrization of Λ_i is as follows:

$$\Lambda_i \;=\; \sum_{s=1}^{m_S} w_{is}\lambda_i \,\frac{\Lambda^{(s)}}{\lambda_{(s)}} \qquad (i = 1, ..., m); \tag{2.39}$$

see also Formula (4.29). This shows that two obligors are correlated if and only if there is at least one sector such that both obligors have a positive sector weight with respect to this sector. Only in such cases two obligors admit a common source of *systematic default risk*. Note that (2.39) is consistent with the assumption that λ_i equals the expected default intensity of obligor i. Default risk of obligor i is then modeled by a mixed Poisson random variable L'_i with random intensity Λ_i.

Note that in accordance with (2.12) any conditional default intensity of obligor i arising from realizations $\theta_1, ..., \theta_{m_S}$ of the sector's default intensities $\Lambda^{(1)}, ..., \Lambda^{(m_S)}$ generates a conditional one-year default probability $p_i(\theta_1, ..., \theta_{m_S})$ of obligor i by setting

$$\begin{aligned} p_i(\theta_1, ..., \theta_{m_S}) \;&=\; \mathbb{P}[L'_i \geq 1 \mid \Lambda_1 = \theta_1, ..., \Lambda_{m_S} = \theta_{m_S}] \\ &=\; 1 - e^{-\lambda_i \sum_{s=1}^{m_S} w_{is}\theta_s/\lambda_{(s)}} . \end{aligned} \tag{2.40}$$

Let L' denote the random variable representing the number of defaults in the portfolio. We already mentioned that CreditRisk$^+$ is a Poisson mixture model. More explicitly, it is assumed that L' is a Poisson variable with random intensity $\Lambda^{(1)} + \cdots + \Lambda^{(m_S)}$. Additionally, it is naturally required to obtain the portfolio's defaults as the sum of single obligor defaults, and indeed (2.39) obviously is consistent with $L' = L'_1 + \cdots + L'_m$ when defining the sector's mean intensity by

$$\lambda_{(s)} \;=\; \sum_{i=1}^{m} w_{is}\lambda_i \;;$$

see also Formula (4.30) in Section 4.3.2.

Now, on the portfolio level, the "trick" CreditRisk$^+$ uses in order to obtain a nice closed-form distribution of portfolio defaults is *sector analysis*. Given that we know distribution of defaults in every single sector, the portfolio's default distribution then just turns out to be the *convolution* of the sector distributions due to the independence of the sector variables $\Lambda^{(1)}, ..., \Lambda^{(m_S)}$. So we only have to find the sector's default distributions.

When focusing on single sectors, it is a standard result from elementary statistics (see, e.g., [160] Section 8.6.1) that any gamma-mixed Poisson distribution follows a *negative binomial distribution*. See Figure 2.7 in Section 2.5.2. Therefore, every sector has its own individually parametrized negative binomial distribution of sector defaults, such that the portfolio's default distribution indeed can be obtained as a convolution of negative binomial distributions. As a consequence, the *generating function* of the portfolio loss can be explicitly written in a closed form; see Formula (4.35) in Chapter 4.

So far we have only discussed the distribution of defaults. The corresponding loss distributions for a single sector are given as the *compound* distribution arising from two independent random effects, where the first random effect is due to the uncertainty regarding the number of defaults (negative binomially distributed) in the sector and the second random effect arises from the uncertainty regarding the exposures affected by the sector defaults; see Section 4.3.2. On the portfolio level the loss distribution again is the convolution of sector loss distributions. The final formula for the generating function of the portfolio loss is presented in (4.36).

2.4.3 CreditPortfolioView

We keep our exposition in a summarizing style. Readers interested in more details should study the papers by Wilson [189, 190], and the technical documentation [135] of CPV in its recent version[13]. Both sources and the overview in Crouhy et al. [38], Section 8.10, have been valuable references for writing this paragraph. Until now we restricted the discussion in this chapter to default modeling. For our summary of CreditPortfolioView (CPV) we will now also include rating migrations.

CPV has its roots in two seminal papers by Wilson [189, 190]. Based on these two papers, McKinsey & Company[14] developed CreditPortfolioView during the years since then as a tool for supporting consulting projects in credit risk management. Summarizing, one could say that CPV is a ratings-based portfolio model incorporating the dependence of default and migration probabilities on the *economic cycle*. Con-

[13] We are grateful to McKinsey & Company for sending us the technical documentation [135] of CPV as a source for writing this section.

[14] McKinsey & Company is an international management consulting firm.

sequently default probabilities and migration matrices are subject to *random fluctuations.*

Let us start with some general remarks regarding *migration matrices.* Mathematically speaking a migration matrix is a *stochastic*[15] *matrix* in $\mathbb{R}^{n\times n}$, where n depends on the number of rating classes incorporated. For example, the rating agencies (Moody's or S&P) typically publish migration matrices w.r.t. two different dimensions, namely $n = 8$ (standard case) and n substantially larger than 8, reflecting the finer rating scale as shown in Figure 1.2. Migration matrices will be extensively studied later on when discussing the *term structure of default probabilities*; see Section 6.3.3.

The basic observation underlying CPV is that migration probabilities show random fluctuations due to the volatility of the economic cycle. Very much reflecting the terminology in this chapter, CPV calls any migration matrix observed in a particular year a *conditional* migration matrix, because it is sampled conditional on the economic conditions of the considered year. Calculating the average of conditional migration matrices sampled over a series of years will give us an *unconditional* migration matrix reflecting *expected migration paths.* Such average migration matrices can be found in the rating agency reports, or can be calculated from bank-internal data.

Now let us assume that an unconditional migration matrix has been chosen. We denote this matrix by $\overline{M} = (\overline{m}_{ij})$ where i, j range from 1 to 8. Compatible to the notation at the beginning of this chapter we denote rating classes by R_i. Rating class R_1 stands for the best possible credit quality, whereas R_8 is the *default state*, such that $\overline{m}_{i8} = \mathbb{P}[R_i \to R_8]$ denotes the probability that obligors with rating R_i at the beginning of a year go into default until the end of that year. In general it is assumed that R_i is more creditworthy than R_j if and only if $i < j$; compare also to Figure 1.2. Because the default state is *absorbing*[16], we additionally have $\overline{m}_{8j} = 0$ for $j = 1, ..., 7$ and $\overline{m}_{88} = 1$. Note that in this notation $\overline{m}_{i8}$ takes over the role of p_i in previous paragraphs, where p_i denoted the one-year default probability of some customer i. Also recall that default probabilities are often *rating-driven* so that there is

[15] A matrix (m_{ij}) is called *stochastic* if $\sum_j m_{ij} = 1$ for every row i.

[16] Absorbing means that the default state is a trap with no escape.

no need to distinguish between two obligors with the same rating when interested in their default probabilities; see also Section 1.1.1.

CPV assumes that there are several *risk segments* differently reacting to the overall economic conditions. For example, typical risk segments refer to industry groups. In our presentation we will not be bothered about the interpretation of risk segments; so, we just assume that there are m_S such segments. Moreover, to keep our presentation free from index-overloading we restrict ourselves to a one-year view. For each segment CPV simulates a conditional migration matrix based on the average migration matrix $\overline{M}$ and a so-called *shift algorithm*. The shift algorithm works in three steps:

1. A segment-specific conditional default probability p_s is simulated for every segment $s = 1, ..., m_S$. The probability p_s is the same for all rating classes, and we will later explain the simulation method CPV uses for generating those probabilities. Any simulated vector $(p_1, ..., p_{m_S})$ can be considered as an aggregated (second-level) *scenario* in a Monte Carlo simulation of CPV. Underlying the generation of such a scenario is the simulation of macroeconomic factors driving $(p_1, ..., p_{m_S})$.

2. A so-called *risk index* r_s representing the state of the economy seen in light of segment s is calculated by means of the ratio

$$r_s \;=\; \frac{p_s}{\overline{p}_s} \qquad (s = 1, ..., m_S), \tag{2.41}$$

   where $\overline{p}_s$ denotes the unconditional default probability of segment s, incorporating the average default potential of segment s.

3. Last, a conditional migration matrix $M^{(s)} = (m_{ij}^{(s)})$ for segment s w.r.t. a scenario $(p_1, ..., p_{m_S})$ is defined by

$$m_{ij}^{(s)} \;=\; \alpha_{ij}(r_s - 1) + \overline{m}_{ij} \qquad (s = 1, ..., m_S). \tag{2.42}$$

   The *shift coefficients* α_{ij} have to be calibrated by the user of CPV, although CPV contains some standard values that can be chosen in case a user does not want to specify the shift factors individually. The shift factors depend on the considered migration path $R_i \to R_j$, hereby expressing the sensitivity of $\mathbb{P}[R_i \to R_j]$ w.r.t. a change in the segment's risk index r_s. Because they are intended to reflect rating class behavior rather than a segment's reaction to macroeconomic conditions, they are uniform w.r.t. different segments. Calculating the row sums of the shifted migration matrices we obtain

$$\sum_{j=1}^{8} m_{ij}^{(s)} \;=\; (r_s - 1)\sum_{j=1}^{8}\alpha_{ij} + \sum_{j=1}^{8}\overline{m}_{ij}.$$

Therefore, in order to guarantee that the shifted matrix $M^{(s)}$ is stochastic, CPV assumes $\sum_{j=1}^{8}\alpha_{ij} = 0$. If a concrete realization of (2.42) results in a migration probability $m_{ij}^{(s)}$ shifted to the negative, CPV performs a correction by setting such negative values to zero. In such cases a renormalization of the rows of $M^{(s)}$ is necessary in order to obtain a stochastic matrix.

For some reasons to be explained later the *shift matrix* (α_{ij}) is supposed to satisfy some more conditions, namely

$$\alpha_{ij} \geq 0 \quad \text{for} \quad i < j \qquad \text{and} \qquad \alpha_{ij} \leq 0 \quad \text{for} \quad i > j\;. \tag{2.43}$$

This assumption is compatible to the condition that the shift matrix has row sums equal to zero. Because the upper (lower) triangle matrix of a migrating matrix contains the probabilities for a rating downgrade (upgrade), the conditions on the shift matrix are not as arbitrary as it seems at first glance. Just below we come back to this issue.

Any conditional migration matrix $M^{(s)}$ is relevant to all obligors in segment s. Thinking about Formula (2.42) and applying the conditions in (2.43) we see that one can distinguish between three different types of scenarios:

- $\boldsymbol{r_s < 1}$ **:**
  In such situations the simulation suggests an *expansion* of the economy, admitting a potential for a lower number of downgrades and a higher number of upgrades, reflecting favourable economic conditions.

- $\boldsymbol{r_s = 1}$ **:**
  This is the *average* macroeconomic scenario. Formula (2.42) shows that in such cases the impact of the shift coefficients vanishes to zero such that the shifted migration probability $m_{ij}^{(s)}$ agrees with the unconditional migration probability $\overline{m}_{ij}$ for all combinations of i and j.

- $\boldsymbol{r_s > 1}$ **:**
  This scenario refers to a *recession*. Downgrades are more likely and the potential for upgrades is reduced when compared to average conditions.

Note that because CPV treats different segments differently, the concept of segment-specific risk indices allows for a great flexibility. In CPV terminology the process of generating macro-scenarios in a Monte-Carlo simulation is called the *systematic risk model*.

Based on any outcome of the systematic risk model, CPV constructs a conditional loss distribution for the considered portfolio. In a last step all conditional loss distributions are aggregated to an unconditional portfolio loss distribution. The details of how CPV tabulates losses for obtaining scenario-specific and unconditional distributions are rather technical and can be found in the technical documentation [135] of CPV. There one also finds information about other special aspects of the model, for example the implementation of country risk or the method for discounting cash flows to a present value. A nice and CPV-unique feature is the ability to incorporate *stressed tails* in the systematic risk model (to be used in the *CPV Direct* mode, see our discussion on CPV-Macro and CPV-Direct later in this section) in order to study the impact of heavy recessions.

Remaining to be done in this section is a brief description of how CPV manages to simulate the segment-specific conditional default probabilities p_s. Here CPV supports two distinct modes of calibration:

- *CPV Macro*
  If CPV is run in the macro mode, default and rating migration shifts are explained by a *macroeconomic regression model*. The macroeconomic model underlying systematic influences on the economic future of obligors is calibrated by means of time series of empirical data; see the original papers by Wilson [189, 190]. The calibration of CPV Macro is more complicated than the alternative CPV Direct. The difficulties in calibrating CPV Macro are mainly due to the many parameters that have to be estimated; see Formula (2.44) and (2.45).

- *CPV Direct*
  In this mode of CPV, the segment-specific conditional default probabilities p_s are directly drawn from a *gamma distribution*. In

other words, the conditional probability determining a segment's risk index is not implied by some underlying macroeconomic factor model. Working with CPV Direct, the user can avoid all the difficulties some macroeconomic regression model incorporates. The effort in sector calibration is reduced to the calibration of two parameters of a gamma distribution for each risk segment.

Originally CPV contained only the macro approach. CPV Direct was developed later on in order to make the calibration of the model easier.

2.4.3.1 CPV Macro

In CPV Macro, macroeconomic variables drive the distribution of default probabilities and migration matrices for each risk segment. Typical candidates for macroeconomic factors are the *unemployment rate*, the growth rate of the *Gross Domestic Product* (GDP), *interest* or *currency exchange rates*, and other variables reflecting the macroeconomy of a country. The regression model underlying CPV Macro can be described as follows.

Let us again assume we work with m_S risk segments. Every risk segment s is represented by a *macroeconomic index* $Y_{s,t}$ where t refers to the particular time the index is considered. The index $Y_{s,t}$ itself is represented by a weighted sum of *macroeconomic variables*,

$$Y_{s,t} = w_{s,0} + \sum_{k=1}^{K} w_{s,k} X_{s,k,t} + \varepsilon_{s,t} \ , \tag{2.44}$$

where $X_{s,k,t}$ are macroeconomic variables at time t, relevant to the economic performance of segment s, $(w_{s,k})_{k=0,...,K}$ are coefficients that have to be calibrated w.r.t. segment s, and $\varepsilon_{s,t}$ describes the *residual* random fluctuation of $Y_{s,t}$ not explainable by the fluctuation of the $X_{s,k,t}$'s. For every segment s, a calibration analogous to (2.44) has to be done. Typically for such regression models the residual variables $\varepsilon_{s,t}$, $s = 1, ..., m_S$, are assumed to be i.i.d. normally distributed and independent of the variables $X_{s,k,t}$.

The macroeconomic variables $X_{s,k,t}$ are parametrized by an autoregressive model with time lag t_0, where t_0 has to be specified in the model. More explicitly, it is assumed that the macroeconomic variables can be written as

$$X_{s,k,t} = \theta_{k,0} + \sum_{j=1}^{t_0} \theta_{k,j} X_{s,k,t-j} + \gamma_{s,k,t} \ . \tag{2.45}$$

In his original papers Wilson used an *AR(2)-process* ($t_0 = 2$). The conditional default probability for obligors in segment s is specified in terms of the segment-specific macroeconomic index conditioned on its respective realization $Y_{s,t} = y_{s,t}$ and is given by the *logit function*

$$p_{s,t} \ = \ \frac{1}{1 + \exp(y_{s,t})} \ . \tag{2.46}$$

Once the logit function has been calculated for a scenario, the systematic risk model will calculate the shifted migration matrix as a preparation for the tabulation of the conditional loss distribution.

2.4.3.2 CPV Direct

CPV Direct models the overall macroeconomy (represented by the probabilities p_s) by a multivariate gamma distribution

$$\mathbf{\Gamma} = \Big(\Gamma(\gamma_{1,1}, \gamma_{1,2}), ..., \Gamma(\gamma_{m_S,1}, \gamma_{m_S,2})\Big),$$

where the parameter pairs $(\gamma_{s,1}, \gamma_{s,2})$ have to be calibrated to each segment. A main issue is the calibration of the *correlation* matrix of $\mathbf{\Gamma}$. In general these challenges are much easier to master than calibrating the macroeconomic indices by means of an autoregression as it is suggested by CPV Macro.

The parameters of the gamma distribution of a segment are calibrated by specifying the mean and the volatility of the random variable generating the segment's default probability p_s. The parameters $(\gamma_{s,1}, \gamma_{s,2})$ are then determined by a moment matching based on (2.38).

Note that the support of the gamma distribution is $\mathbb{R}_+$, so that it theoretically can happen to draw a number $p_s > 1$, which obviously can not be interpreted as a probability. This is an unpleasant side effect when drawing a random number that is supposed to be a probability from a gamma distribution. However, to some extent this very much reminds us of the comparably unpleasant possibility of obtaining a multiple default of a single obligor in the CreditRisk$^+$ framework; see Section 2.4.2 and Chapter 4. In practice such scenarios are not very likely and will be 'thrown away' by the Monte Carlo engine of CPV.

2.4.4 Basic Remarks on Dynamic Intensity Models

We already mentioned that this section is intended to be not more than a brief "remark" with some references. Dynamic intensity models have been extensively studied by Duffie and Singleton [49, 50]. In Duffie and Gârleanu [47] intensity models are applied to the valuation of *collateralized debt obligations*. The theory underlying intensity models has much in common with interest rate term structure models, which are mathematically complex and beyond the scope of this book. For readers interested in the theory we refer to the already mentioned papers by Duffie et al. and also to Jarrow, Lando, and Turnbull [103] and Lando [119].

In the sequel we briefly summarize some basics of one representative of the group of intensity models. First of all, the basic assumption is that every obligor admits a *default time* such that default happens in a time interval $[0, T]$ if and only if the default time of the considered obligor appears to be smaller than the planning horizon T. The default times are driven by an *intensity process*, a so-called *basic affine process*, whose evolution is described by the *stochastic differential equation*

$$d\lambda(t) = \kappa(\theta - \lambda(t))dt + \sigma\sqrt{\lambda(t)}dB(t) + \Delta J(t)\ , \tag{2.47}$$

where $(B(t))_{t\geq 0}$ is a standard Brownian motion and $\Delta J(t)$ denotes the jump that occurs – if it occurs – at time t. Hereby J is a pure jump process, independent of B, whose jump sizes are independent, positive and exponentially distributed with mean μ and admitting jump times according to an independent Poisson process with mean jump arrival rate l. The parameter set $(\kappa, \theta, \sigma, \mu, l)$ can be adjusted to control the manner in which default risk changes over time, e.g., one can vary the mean reversion rate κ, the long-run mean $\bar{m} = \theta + l\mu/\kappa$, or the relative contributions to the total variance of $\lambda(t)$ that are attributed to the jump risk and diffusion volatility.

Conditional on a realization $(\lambda(t))_{t\geq 0}$ of the stochastic process solving (2.47), the default times of obligors are independent *Poisson arrivals* with intensities $\lambda(t)$. From this point of view the dynamic intensity model can be considered as a time-continuous extension of the CreditRisk$^+$ framework.

The unconditional survival probability $q(t)$ is given by

$$q(t) = \mathbb{E}\left[e^{-\int_0^t du\lambda(u)}\right].$$

The conditional survival probability for a time interval of length $s \geq 0$ given survival up to time t can be calculated by

$$q(s+t \mid t) \;=\; \mathbb{E}_t \left[e^{-\int_t^{t+s} du \lambda(u)} \right] = e^{\alpha(s)+\beta(s)\lambda(s)},$$

where the explicit solutions for the coefficients $\alpha(s)$ and $\beta(s)$ can be found in the above-mentioned papers by Duffie et al.

Because the sum of independent basic affine processes with common parameters κ, σ, and μ, governing respectively the mean-reversion rate, diffusive volatility, and mean jump size, again yields a basic affine process, one can introduce dependencies between the default times of the counterparties in a considered portfolio. Each obligor's default intensity can thus be represented by means of a one-factor *Markov model*

$$\lambda_i = X_c + X_i \ ,$$

where X_c and $X_1, \ldots, X_N$ are independent basic affine processes with respective parameters $(\kappa, \theta_c, \sigma, \mu, l_c)$ and $(\kappa, \theta_i, \sigma, \mu, l_i)$. The so constructed process λ_i then again is a basic affine process with parameters $(\kappa, \theta, \sigma, \mu, l)$, where $\theta = \theta_c + \theta_i$ and $l = l_c + l_i$. One could interpret X_c as a *state variable* governing the *common aspects of performance*, whereas X_i, seen as a state variable, contributes the *obligor-specific* or *idiosyncratic risk*. Obviously, this can be extended to handle multi-factor models by introducing additional basic affine processes for each of a collection of *sectors*. Each obligor then admits an intensity process

$$\lambda_i = X_c + X_i + X_{s(i)} \ ,$$

where the *sector factor* $X_{s(i)}$ is common to all obligors in that sector. Here $s(i)$ denotes the sector in which obligor i takes place

A possible simulation algorithm to generate default times $\tau_1, \ldots, \tau_n$ up to some time horizon T with given intensities $\lambda_i, \ldots, \lambda_n$ is the *multi-compensator* method. In this method it is assumed that the *compensator* $\Lambda_i(t) = \int_0^t \lambda_i(u)du$ can be simulated for all i and t. Then n independent unit-mean exponentially distributed random variables $Z_1, \ldots, Z_n$ are drawn. For each i one has $\tau_i > T$ if $\Lambda_i(T) < Z_i$. In this case the obligor survived the time interval $[0, T]$. Otherwise, the default time of the obligor is given by $\tau_i = \min\{t : \Lambda_i(t) = Z_i\}$ (see the references for more details).

Here we stop our discussion and hope that, although we kept our introduction to the basics of dynamic intensity models rather short, the reader nevertheless got some flavor of what these models are all about.

2.5 One-Factor/Sector Models

In Sections 2.1.2 respectively 2.2.2 we discussed portfolios admitting a uniform default probability respectively intensity and a uniform default correlation for Bernoulli respectively Poisson mixture models. In this paragraph we look in more detail at portfolios with uniform dependency structure, namely *one-factor* respectively *one-sector* models.

2.5.1 One-Factor Models in the Asset Value Model Setup

The one-factor model in the context of the models by Moody's KMV and the RiskMetrics Group is completely described by specializing equations (2.27) and (2.32) to the case of only one single factor common to all counterparties, hereby assuming that the asset correlation between obligors is uniform. More explicitly, this means that the composite factors Φ_i of all obligors are equal to one single factor, usually denoted by $Y \sim N(0,1)$. Moreover, instead of (2.27) one can write[17]

$$r_i \;=\; \sqrt{\varrho}\,Y + \sqrt{1-\varrho}\,Z_i \qquad (i = 1, ..., m), \tag{2.48}$$

where $\sqrt{1-\varrho}\,Z_i$, with $Z_i \sim N(0,1)$, takes over the role of the residual ε_i and ϱ is the uniform asset correlation between the asset value log-returns $r_i \sim N(0,1)$. In one-factor models, the uniform asset correlation ϱ equals the *R-squared* as described in (1.24), uniform to all obligors. As before, it is assumed that the residuals Z_i constitute an independent family, also independent of the factor Y.

Under the assumption of a single factor and a uniform ϱ, Equation (2.32) turns into

$$p_i(Y) \;=\; N\left[\frac{N^{-1}[p_i] - \sqrt{\varrho}\,Y}{\sqrt{1-\varrho}}\right] \qquad (i = 1, ..., m). \tag{2.49}$$

[17]Note that here one could more generally work with ϱ_i instead of ϱ. Note also that the term $\sqrt{\varrho}$ takes over the role of R_i in Equation (2.27).

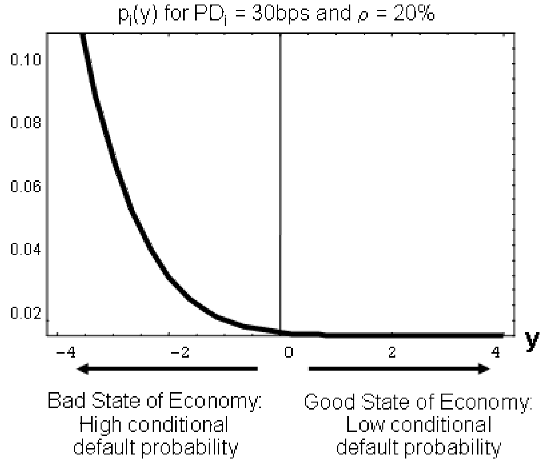


FIGURE 2.3: Asset Value One-Factor Model: Conditional default probability as a function of the factor realizations $Y = y$.

Figure 2.3 illustrates the dependence of the conditional default probability $p_i(y)$ on realizations y of the single factor Y.

Figure 2.4 shows for three fixed *states of economy* $Y = -3, 0, 3$ the conditional default probability $p_i(y)$ as a function of the average one-year default probability p_i arising in formula (2.49), which we denoted by PD_i in the introductory Chapter 1. Figures 2.3 and 2.4 also give an interpretation of the behavior of conditional default probabilities in terms of the *economic cycle* captured by the single factor Y.

Before proceeding, we calculate the *joint default probability* (JDP) of two obligors.

2.5.1 Proposition *In a one-factor portfolio model with uniform asset correlation ϱ and loss statistics $(L_1, ..., L_m)$ with $L_i \sim B(1; p_i(Y))$, where $p_i(Y)$ is defined as in (2.49), the joint default probability (JDP) of two obligors is given by the bivariate normal integral*

$$JDP_{ij} \;=\; \mathbb{P}[L_i = 1, L_j = 1] \;=\; N_2\big[N^{-1}[p_i], N^{-1}[p_j]; \varrho\big] \;\; ,$$

where $N_2\,[\,\cdot, \cdot\,; \varrho]$ denotes the cumulative bivariate normal distribution function with correlation ϱ.

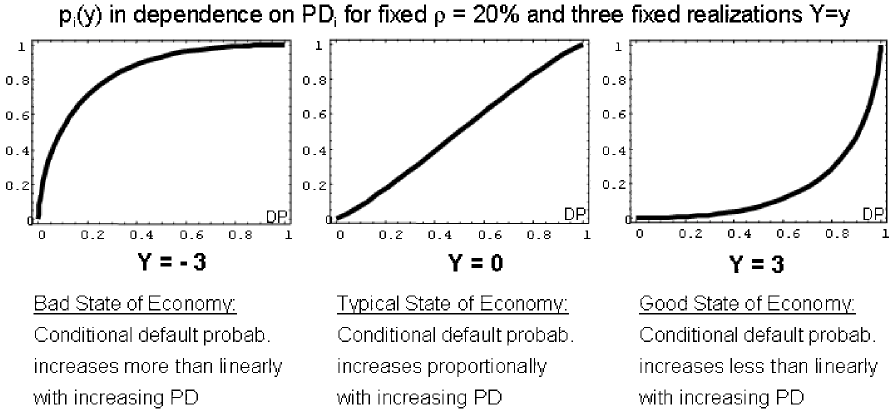


FIGURE 2.4: Asset Value One-Factor Model: Conditional default probability as a function of the average 1-year default probability PD_i.

Proof. The joint default probability can be calculated as

$$\mathbb{P}[L_i = 1, L_j = 1] \ = \ \mathbb{P}(r_i < N^{-1}[p_i], r_j < N^{-1}[p_j]) \ .$$

By construction, the correlation between the asset value log-returns $r_i, r_j \sim N(0,1)$ is ϱ. This proves the proposition. □

We now want to prove that with increasing portfolio size m ("size" in terms of the number of loans in the portfolio) the portfolio loss distribution converges to a closed-form *limit distribution*. References for the sequel are FINGER [61], GORDY [84], SCHÖNBUCHER [163], VASICEK [182], and ONG [150], Example 9.2. In the following we denote by

$$E_i \ = \ \mathrm{EAD}_i \times \mathrm{LGD}_i$$

the exposure that is lost in case obligor i defaults; see Chapter 1 for the meaning of EAD_i and LGD_i. Here we allow for *random* LGDs but deterministic (i.e., fixed) EADs. Moreover, we will not exclude that the LGDs also depend on the state of economy Y in some (not necessarily more detailed specified) way. Such a dependence of the default *and* the recovery rates on the same underlying factor is certainly reasonable, because historic observations show that recovery rates tend to decrease in times where default rates rise up sharply; see, e.g., ALTMAN ET AL. [3] and FRYE [74, 75] for a more detailed discussion of recoveries and their relation to default rates.

Altogether we are looking at a Bernoulli mixture model[18], such that the counterparties are modeled by random variables

$$E_i L_i \,, \qquad L_i \sim B(1; p_i(Y)) \,, \qquad Y \sim N(0,1) \,, \tag{2.50}$$
$$\Big((\mathrm{LGD}_i \times L_i)|_{Y=y}\Big)_{i=1,\dots,m} \quad \text{independent} \,,$$

where we assume that all involved random variables are defined on a common probability space; see Remark 2.5.6. The last condition in (2.50) means that we assume conditional independence of *losses* rather than independence of *default indicators*. For reasons of a shorter and more convenient notation we write in the sequel η_i for LGD_i,

$$\eta_i \;=\; \mathrm{LGD}_i \;.$$

For a portfolio of m obligors[19], the portfolio loss relative to the portfolio's total exposure is given by

$$L \;=\; L^{(m)} \;=\; \sum_{i=1}^{m} w_i \eta_i L_i \qquad \text{where} \quad w_i = \frac{\mathrm{EAD}_i}{\sum_{j=1}^{m} \mathrm{EAD}_j} \;. \tag{2.51}$$

We now want to prove that with increasing number of obligors m some limit behavior of the portfolio loss $L^{(m)}$ can be established. For this we first of all need some technical assumption, essentially taking care that in the limit the portfolio is free of any dominating single exposures.

2.5.2 Assumption *In the following we consider an infinite number of loans with exposures EAD_i. We assume that the following holds:*

$$\sum_{i=1}^{m} EAD_i \;\uparrow\; \infty \qquad (m \to \infty),$$

$$\sum_{m=1}^{\infty} \Big(\frac{EAD_m}{\sum_{i=1}^{m} EAD_i}\Big)^2 \;<\; \infty.$$

[18]Note that the following notation, although intuitive, is not mathematically rigorous. Later on in the proof of Proposition 2.5.4 we will follow a mathematically more precise notation.

[19]Here we make the simplifying assumption that the number of loans in the portfolio equals the number of obligors involved. This can be achieved by aggregating different loans of a single obligor into one loan. Usually the PD, EAD, and LGD of such an aggregated loan are exposure-weighted average numbers.

The first condition says that the total exposure of the portfolio strictly increases to infinity with increasing number of obligors. The second condition implies that the exposure weights shrink very rapidly with increasing number of obligors. Altogether this makes sure that the exposure share of each loan in the portfolio tends to zero.

Condition 2.5.2 is by no means a strict assumption. As an example consider the following situation:

2.5.3 Example *Assuming $a \leq$EAD$_i \leq b$ for some $0 < a \leq b$ and all i, we obtain*

$$\sum_{i=1}^{m} EAD_i \;\geq\; m\,a \;\uparrow\; \infty \quad (m \to \infty),$$

$$\sum_{m=1}^{\infty} \Big(\frac{EAD_m}{\sum_{i=1}^{m} EAD_i}\Big)^2 \;\leq\; \sum_{m=1}^{\infty} \frac{b^2}{m^2 a^2} \;=\; \frac{b^2}{a^2} \sum_{m=1}^{\infty} \frac{1}{m^2} \;<\; \infty,$$

such that Assumption 2.5.2 is fulfilled in this case.

Now we are in a position to prove the following statement.

2.5.4 Proposition *Assumption 2.5.2 is sufficient to guarantee that in the limit the percentage portfolio loss $L^{(m)}$ defined in (2.51) and the conditional expectation $\mathbb{E}[L^{(m)}|Y]$ are equal almost surely, such that*

$$\mathbb{P}\Big[\lim_{m\to\infty}\Big(L^{(m)} - \mathbb{E}[L^{(m)}|Y]\Big) = 0\Big] \;=\; 1.$$

Proof. Fix $y \in \mathbb{R}$. Define the conditional probability measure $\mathbb{P}_y$ by

$$\mathbb{P}_y(\cdot) \;=\; \mathbb{P}[\,\cdot \mid Y = y].$$

Consider the random variable

$$X_k \;=\; \mathrm{EAD}_k(\eta_k L_k - \mathbb{E}[\eta_k L_k|Y]).$$

With respect to $\mathbb{P}_y$, the random sequence $(X_k)_{k\geq 1}$ is independent due to (2.50) and centered by definition. We now define $\tau_m = \sum_{i=1}^{m} \mathrm{EAD}_i$, such that $(\tau_m)_{m\geq 1}$ is a positive sequence strictly increasing to infinity due to Assumption 2.5.2. If we could prove that

$$\sum_{k=1}^{\infty} \frac{1}{\tau_k^2}\,\mathbb{E}\left[X_k^2\right] \;<\; \infty\,, \tag{2.52}$$

then a version[20] of the *strong law of large numbers* (see [12]) would yield

$$\lim_{m\to\infty} \frac{1}{\tau_m} \sum_{k=1}^{m} X_k \;=\; 0 \qquad \mathbb{P}_y - \text{almost surely.} \tag{2.53}$$

We therefore prove (2.52) next. From Assumption 2.5.2 we get

$$\sum_{k=1}^{\infty} \frac{1}{\tau_k^2} \mathbb{E}\left[X_k^2\right] \;\leq\; \sum_{k=1}^{\infty} \frac{4 \times \mathrm{EAD}_k^2}{\tau_k^2} \;<\; \infty$$

due to the uniform boundedness of $(\eta_k L_k - \mathbb{E}[\eta_k L_k|Y])$. So we have established (2.53) for every $y \in \mathbb{R}$. We can now write

$$\mathbb{P}\Big[\lim_{m\to\infty}(L^{(m)} - \mathbb{E}[L^{(m)}|Y]) = 0 \;\Big|\; Y = y\Big] \;=\; 1 \qquad \text{for every } y \in \mathbb{R}.$$

But then almost sure convergence also holds unconditionally,

$$\mathbb{P}\Big[\lim_{m\to\infty}(L^{(m)} - \mathbb{E}[L^{(m)}|Y]) = 0\Big] \;=$$

$$= \;\int \mathbb{P}\Big[\lim_{m\to\infty}(L^{(m)} - \mathbb{E}[L^{(m)}|Y]) = 0 \;\Big|\; Y = y\Big] dP_Y(y) \;=\; 1.$$

Therefore the proposition is proved. □

2.5.5 Corollary *In the case that $(\eta_i L_i)_{i\geq 1}$ are not only conditionally independent but also identically distributed, Proposition 2.5.4 can be reformulated as follows:*
There exists some measurable function $p : \mathbb{R} \to \mathbb{R}$ such that for $m \to \infty$ the portfolio loss $L^{(m)}$ converges to $p \circ Y$ almost surely. Moreover, $p \circ Y$ equals $\mathbb{E}[\eta_1 L_1|Y]$ almost surely.

Proof. Because the conditional expectation $\mathbb{E}[L^{(m)}|Y]$ is by definition $\sigma(Y)$-measurable, where $\sigma(Y)$ denotes the σ-Algebra generated by Y, there exists some measurable function $p : \mathbb{R} \to \mathbb{R}$ with $\mathbb{E}[L^{(m)}|Y] = p \circ Y$; see [111], Lemma 1.13. Combined with Proposition 2.5.4 and the

[20] This version of the LLNs is based on Kronecker's Lemma (see [12]), saying that whenever $(x_k)_{k\geq 1}$ and $(\tau_k)_{k\geq 1}$ are sequences with the latter being positive and strictly increasing to infinity, such that $\sum_{k=1}^{\infty} x_k/\tau_k$ converges, we obtain $\lim_{m\to\infty} \tau_m^{-1} \sum_{k=1}^{m} x_k = 0$.

assumption that all losses are identically distributed this concludes the proof of Corollary 2.5.5. □

The important conclusion from the convergence results above is that in the limit the randomness of the portfolio loss $L^{(m)}$ solely depends on the randomness of the factor Y. By increasing the number of obligors in the portfolio, the specific risk is completely removed, such that in the limit only *systematic risk* arising from the volatility of the factor Y remains in the portfolio.

2.5.6 Remark The proof of Proposition 2.5.4 does not rely on the particular distribution we use for the factor Y. To make this more precise let us look at a probability space suitable for uniform portfolios. We need a factor Y and residual variables $Z_1, Z_2, ...$, which are random variables in $\mathbb{R}$ defined on some not necessarily more specified probability spaces $(\Omega_Y, \mathcal{F}_Y, P_Y)$, $(\Omega_1, \mathcal{F}_1, P_1)$, $(\Omega_2, \mathcal{F}_2, P_2)$, and so on. A suitable probability space for Proposition 2.5.4 is the product space

$$(\Omega, \mathcal{F}, P) \;=\; (\Omega_Y, \mathcal{F}_Y, P_Y) \otimes (\Omega_1, \mathcal{F}_1, P_1) \otimes (\Omega_2, \mathcal{F}_2, P_2) \otimes \cdots \;,$$

because we always assume the variables $Y, Z_1, Z_2, ...$ to be independent. For every $\omega = (y, z_1, z_2, ...) \in \Omega$ the loss variables $L_i(\omega)$ are given by *latent variable indicators* evaluated w.r.t. the realization ω,

$$L_i(\omega) \;=\; \mathbf{1}_{\{\sqrt{\varrho}\, y + \sqrt{1-\varrho}\, z_i < c_i\}} \;.$$

It is not difficult to argue that the proof of Proposition 2.5.4 only relies on the conditional independence of the variables Z_i w.r.t. Y and the asymptotics of the portfolio weights according to Assumption 2.5.2. In the case that the factor Y and the residuals Z_i are normally distributed, $(\Omega, \mathcal{F}, P)$ turns out to be an *infinite dimensional Gaussian space*, but due to the more generally applicable proof we can use the same convergence argument for distributions other than normal. For example, the t-distribution is a natural candidate to replace the normal distribution; see Section 2.6.1.

Now let us apply our findings to *uniform portfolios* by assuming that $p_i = p$ for all obligors i, such that the assumptions of Proposition 2.5.5 are fulfilled. In the asset value model as we introduced it, the factor Y and the residual variables $Z_1, Z_2, ...$ follow a standard normal distribution. For reasons of simplicity we assume constant LGDs ($\eta_i =$

100%). In this framework, the function p from Proposition 2.5.5 can be explicitly derived by applying Equation (2.49), and taking into account that we are in a Bernoulli framework,

$$\mathbb{E}[L^{(m)}|Y] \;=\; \sum_{i=1}^{m} w_i \mathbb{E}[L_i|Y] \;=\; N\left[\frac{N^{-1}[p] - \sqrt{\varrho}\, Y}{\sqrt{1-\varrho}}\right] =: p(Y)\;,$$

such that Proposition 2.5.4 guarantees that

$$L^{(m)} \;\overset{m\to\infty}{\longrightarrow}\; p(Y) \;=\; N\left[\frac{N^{-1}[p] - \sqrt{\varrho}\, Y}{\sqrt{1-\varrho}}\right] \quad \text{almost surely.} \tag{2.54}$$

So for portfolios with a sufficiently large portfolio size m satisfying Assumption 2.5.2, the percentage quote of defaulted loans for a given state of economy $Y = y$ is approximately equal to the conditional default probability $p(y)$. In the limit we obtain a portfolio loss variable $p(Y)$ describing the fraction of defaulted obligors in an *infinitely fine-grained* credit portfolio.

We now want to derive the cumulative distribution function and the probability density of the limit loss variable $p(Y)$, $Y \sim N(0,1)$, with $p(\cdot)$ as in (2.54). Denote the portfolio's percentage number of defaults in an infinitely fine-grained portfolio (again assuming constant LGDs of 100%) by L. We then have for every $0 \le x \le 1$

$$\begin{aligned}
\mathbb{P}[L \le x] \;&=\; \mathbb{P}[p(Y) \le x] \qquad\qquad (2.55)\\
&=\; \mathbb{P}\left[-Y \le \frac{1}{\sqrt{\varrho}}\Big(N^{-1}[x]\sqrt{1-\varrho} - N^{-1}[p]\Big)\right]\\
&=\; N\left[\frac{1}{\sqrt{\varrho}}\Big(N^{-1}[x]\sqrt{1-\varrho} - N^{-1}[p]\Big)\right].
\end{aligned}$$

In the sequel we will denote this distribution function by

$$F_{p,\varrho}(x) \;=\; \mathbb{P}[L \le x] \qquad (x \in [0,1]). \tag{2.56}$$

The corresponding probability density can be derived by calculating the derivative of $F_{p,\varrho}(x)$ w.r.t. x, which is

$$\begin{aligned}
f_{p,\varrho}(x) &= \frac{\partial F_{p,\varrho}(x)}{\partial x}\\
&= \sqrt{\frac{1-\varrho}{\varrho}}\exp\bigg(-\frac{1}{2\varrho}\Big((1-2\varrho)\left(N^{-1}[x]\right)^2\\
&\qquad\qquad -2\sqrt{1-\varrho}N^{-1}[x]N^{-1}[p] + \left(N^{-1}[p]\right)^2\Big)\bigg)\\
&= \sqrt{\frac{1-\varrho}{\varrho}}\,\exp\bigg(\frac{1}{2}\left(N^{-1}[x]\right)^2 - \frac{1}{2\varrho}\Big(N^{-1}[p] - \sqrt{1-\varrho}\,N^{-1}[x]\Big)^2\bigg). \qquad (2.57)
\end{aligned}$$

Figure 2.5 shows the loss densities $f_{p,\varrho}$ for different values of p and ϱ.

2.5.7 Proposition *The density $f_{p,\varrho}$ admits four extreme cases induced by the extreme values of the parameters p and ϱ, namely*

1. $\boldsymbol{\varrho = 0}$ **:**
   This is the correlation-free *case with loss variables*

   $$L_i = \mathbf{1}_{\{r_i = Z_i < N^{-1}[p]\}} \sim B(1;p) \ ,$$

   taking (2.48) into account. In this case, the absolute (size-m) portfolio loss $\sum L_i$ follows a binomial distribution, $\sum_{i=1}^m L_i \sim B(m;mp)$, and the percentage portfolio loss L_m converges by arguments analogous to Proposition 2.5.4 (or just by an application of the Law of Large Numbers*) to p almost surely. Therefore, $f_{p,0}$ is the density of a degenerate distribution (i.e., a Dirac measure) concentrated in p. This is illustrated by the first plot in Figure 2.5, where an almost vanishing correlation (ϱ=1 bps) yields an $f_{p,\varrho}$, which is almost just a peak in p=30 bps.*

2. $\boldsymbol{\varrho = 1}$ **:**
   In this case one has perfect correlation *between all loss variables in the portfolio (see also Section 1.2, where the term "perfect correlation" was mentioned the first time). In this case we can replace the percentage portfolio loss L_m by $L_1 \sim B(1;p)$, which is no longer dependent on m. Therefore, the limit ($m \to \infty$) percentage portfolio loss L is also Bernoulli $B(1;p)$, such that $\mathbb{P}[L = 1] = p$ and $\mathbb{P}[L = 0] = 1 - p$. The case of (almost) perfect correlation is illustrated in the fourth plot (p=30 bps, ϱ=99.99%) of Figure 2.5, clearly showing the shape of a distribution concentrated in only two points, yielding an "all or nothing" loss.*

3. $\boldsymbol{p = 0}$ **:**
   All obligors survive almost surely, such that $\mathbb{P}[L = 0] = 1$.

4. $\boldsymbol{p = 1}$ **:**
   All obligors default almost surely, such that $\mathbb{P}[L = 1] = 1$.

Proof. A proof is straightforward. □

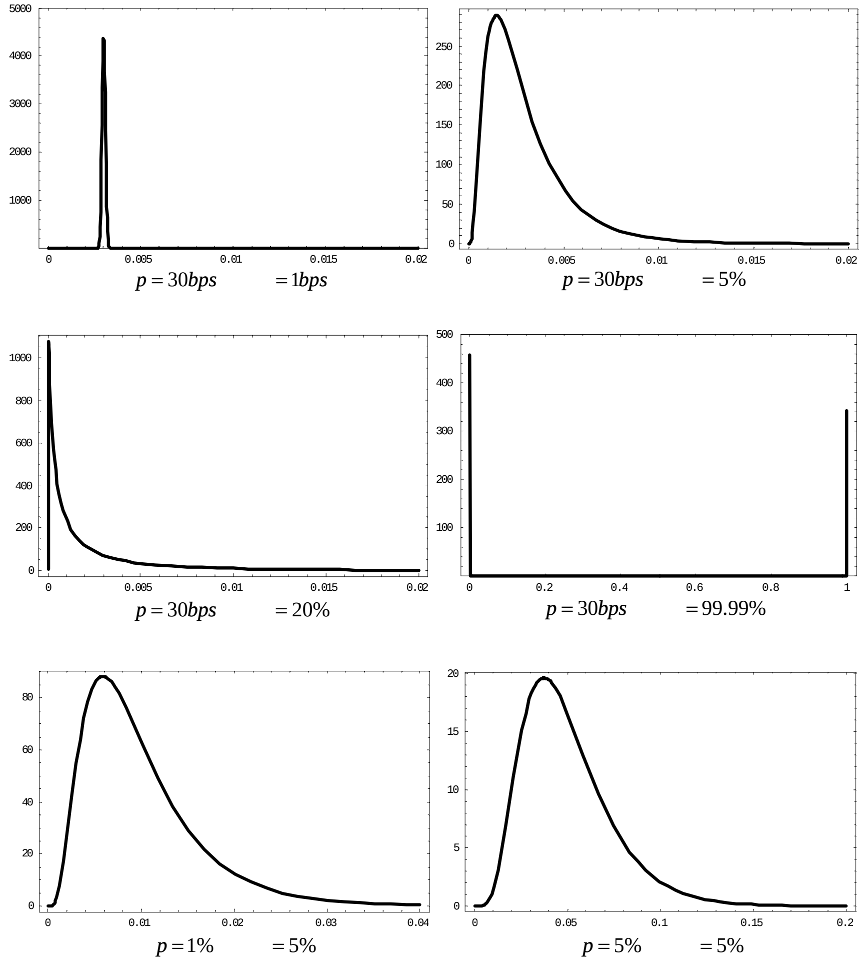


FIGURE 2.5: The probability density $f_{p,\varrho}$ for different combinations of p and ϱ (note that the x-axes of the plots are differently scaled).

For the infinitely fine-grained limit portfolio (encoded by the portfolio's percentage loss variable L) it is very easy to calculate quantiles at any given level of confidence.

2.5.8 Proposition *For any given level of confidence α, the α-quantile $q_\alpha(L)$ of a random variable $L \sim F_{p,\varrho}$ is given by*

$$q_\alpha(L) \;=\; p\Big(-q_\alpha(Y)\Big) \;=\; N\left[\frac{N^{-1}[p] + \sqrt{\varrho}\, q_\alpha(Y)}{\sqrt{1-\varrho}}\right]$$

where $Y \sim N(0,1)$ and $q_\alpha(Y)$ denotes the α-quantile of the standard normal distribution.

Proof. The function $p(\cdot)$ is strictly decreasing, as illustrated by Figure 2.3. Therefore it follows that

$$\mathbb{P}[L \le p(-q_\alpha(Y))] \;=\; \mathbb{P}[p(Y) \le p(-q_\alpha(Y))]$$

$$=\; \mathbb{P}[Y \ge -q_\alpha(Y)] \;=\; \mathbb{P}[-Y \le q_\alpha(Y)]\;,$$

taking (2.55) into account. This proves the proposition. □

By definition (see Section 1.2) the *Unexpected Loss* (UL) is the standard deviation of the portfolio loss distribution. In the following proposition the UL of an infinitely fine-grained uniform portfolio is calculated.

2.5.9 Proposition *The first and second moments of a random variable $L \sim F_{p,\varrho}$ are given by*

$$\mathbb{E}[L] = p \qquad \text{and} \qquad \mathbb{V}[L] = N_2\big[N^{-1}[p], N^{-1}[p]; \varrho\big] - p^2\;,$$

where N_2 is defined as in Proposition 2.5.1.

Proof. That the first moment equals p follows just by construction of $F_{p,\varrho}$. Regarding the second moment, we write $\mathbb{V}[L] = \mathbb{E}[L^2] - \mathbb{E}[L]^2$. We already know $\mathbb{E}[L]^2 = p^2$. So it only remains to show that $\mathbb{E}[L^2] = N_2[N^{-1}[p], N^{-1}[p]; \varrho]$. For proving this, we use a typical "conditioning trick." For this purpose, let $X_1, X_2 \sim N(0,1)$ denote two independent standard normal random variables, independent of the variable

$$X \;=\; \frac{N^{-1}[p] - \sqrt{\varrho}\,Y}{\sqrt{1-\varrho}} \;\sim\; N(\mu,\sigma^2)$$

$$\text{with} \qquad \mu = \frac{N^{-1}[p]}{\sqrt{1-\varrho}}\,, \quad \sigma^2 = \frac{\varrho}{1-\varrho}\,.$$

We write g_{μ,σ^2} for the density of X. Then, we can write $\mathbb{E}[L^2]$ as

$$\begin{aligned}
\mathbb{E}[L^2] \;&=\; \mathbb{E}[p(Y)^2] \;=\; \mathbb{E}[N(X)^2] \\
&=\; \int_{\mathbb{R}} \mathbb{P}[X_1 \le X \mid X = x]\,\mathbb{P}[X_2 \le X \mid X = x]\,dg_{\mu,\sigma^2}(x) \\
&=\; \int_{\mathbb{R}} \mathbb{P}[X_1 \le X,\; X_2 \le X \mid X = x]\,dg_{\mu,\sigma^2}(x) \\
&=\; \mathbb{P}[X_1 - X \le 0,\; X_2 - X \le 0]\,.
\end{aligned}$$

The variables $X_i - X$ are normally distributed with

$$\mathbb{E}[X_i - X] \;=\; -\frac{N^{-1}[p]}{\sqrt{1-\varrho}} \quad \text{and} \quad \mathbb{V}[X_i - X] \;=\; 1 + \frac{\varrho}{1-\varrho} \;=\; \frac{1}{1-\varrho}\,.$$

Moreover, the vector $(X_1 - X, X_2 - X)$ is bivariate Gaussian. Because X_1 and X_2 are independent and independent of X, we obtain

$$\mathrm{Corr}[X_1 - X, X_2 - X] \;=\; \mathbb{V}[X] \;=\; \frac{\varrho}{1-\varrho}\,.$$

Denote by $\tilde{X}_1$ and $\tilde{X}_2$ the standardation of $X_1 - X$ and $X_2 - X$:

$$\tilde{X}_1 \;=\; \frac{(X_1 - X) - \left(-\frac{N^{-1}[p]}{\sqrt{1-\varrho}}\right)}{\sqrt{\frac{1}{1-\varrho}}} \quad \text{and} \quad \tilde{X}_2 \;=\; \frac{(X_2 - X) - \left(-\frac{N^{-1}[p]}{\sqrt{1-\varrho}}\right)}{\sqrt{\frac{1}{1-\varrho}}}\,.$$

Due to standardization we obtain the equivalence

$$X_1 - X \;\le\; 0 \quad \Longleftrightarrow \quad \tilde{X}_1 \;\le\; N^{-1}[p]$$

and analogously for $X_2 - X$ and $\tilde{X}_2$. From this we can conclude that

$$\mathbb{P}_{(X_1 - X, X_2 - X)}(0,0) \;=\; \mathbb{P}_{(\tilde{X}_1, \tilde{X}_2)}\big(N^{-1}[p]\,,\, N^{-1}[p]\big)\,.$$

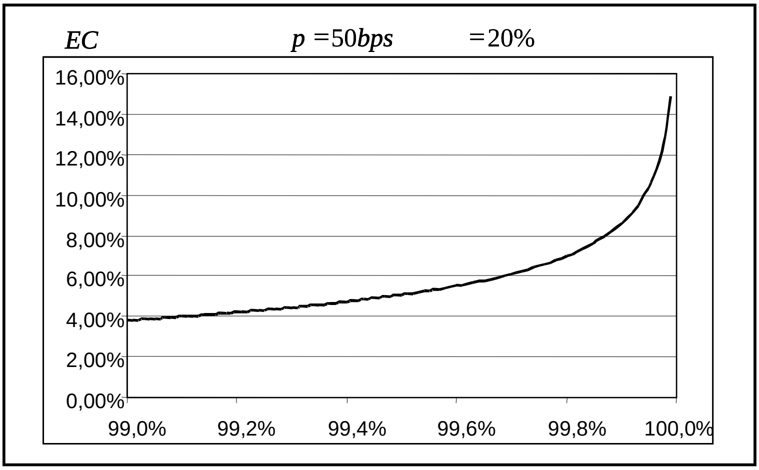


FIGURE 2.6: Dependence of economic capital EC_α on the chosen level of confidence α.

Putting pieces together we obtain

$$\mathbb{E}[L^2] \;=\; \mathbb{P}_{(\tilde{X}_1,\tilde{X}_2)}\big(N^{-1}[p]\,,\,N^{-1}[p]\big)$$

with $(\tilde{X}_1, \tilde{X}_2)$ Gaussian. By construction, $\tilde{X}_1$ and $\tilde{X}_2$ are standardized. What remains to be considered is the correlation between the two:

$$\mathrm{Corr}[\tilde{X}_1, \tilde{X}_2] \;=\; (1-\varrho)\mathrm{Corr}[X_1 - X, X_2 - X] \;=\; \varrho.$$

This finally proves that $\mathbb{E}[L^2] = N_2[N^{-1}[\overline{p}], N^{-1}[\overline{p}]; \varrho]$. $\square$

2.5.10 Proposition *The higher moments of $L \sim F_{p,\varrho}$ are given by*

$$\mathbb{E}\left[L^m\right] \;=\; N_m\big[(N^{-1}[p], ..., N^{-1}[p]), \Gamma_\varrho\big]$$

where $N_m[\cdots]$ denotes the m-dimensional normal distribution function and $\Gamma_\varrho \in \mathbb{R}^{m\times m}$ is a matrix with 1 on the diagonal and ϱ off-diagonal.

Proof. The proof relies on the same argument as the proof of Proposition 2.5.9. A generalization to $m > 2$ is straightforward. $\square$

TABLE 2.1: Economic Capital EC_α for an infinitely fine-grained portfolio (portfolio loss $L \sim F_{p,\varrho}$) w.r.t. p and ϱ, for $\alpha = 99.5\%$.

99.50%	ρ in %							
p in bps	**1%**	**5%**	**10%**	**15%**	**20%**	**30%**	**40%**	**50%**
10	0.12%	0.39%	0.72%	1.06%	1.41%	2.14%	2.86%	3.54%
20	0.22%	0.71%	1.28%	1.87%	2.48%	3.77%	5.14%	6.55%
30	0.32%	0.99%	1.78%	2.58%	3.42%	5.20%	7.13%	9.21%
40	0.41%	1.26%	2.24%	3.24%	4.28%	6.50%	8.93%	11.60%
50	0.49%	1.51%	2.67%	3.85%	5.07%	7.69%	10.58%	13.80%
60	0.57%	1.75%	3.08%	4.42%	5.82%	8.80%	12.11%	15.83%
70	0.65%	1.98%	3.47%	4.97%	6.52%	9.85%	13.55%	17.73%
80	0.73%	2.20%	3.84%	5.49%	7.20%	10.84%	14.90%	19.50%
90	0.81%	2.42%	4.20%	5.99%	7.84%	11.79%	16.18%	21.18%
100	0.88%	2.63%	4.55%	6.48%	8.46%	12.69%	17.40%	22.76%
150	1.23%	3.60%	6.15%	8.67%	11.25%	16.71%	22.75%	29.60%
200	1.55%	4.47%	7.57%	10.60%	13.67%	20.11%	27.18%	35.12%
250	1.85%	5.28%	8.86%	12.33%	15.82%	23.08%	30.96%	39.73%
300	2.14%	6.03%	10.05%	13.90%	17.76%	25.72%	34.26%	43.65%
350	2.41%	6.74%	11.16%	15.36%	19.53%	28.08%	37.17%	47.03%
400	2.67%	7.41%	12.19%	16.70%	21.16%	30.23%	39.76%	49.98%
450	2.92%	8.04%	13.16%	17.96%	22.67%	32.19%	42.09%	52.57%
500	3.16%	8.64%	14.07%	19.13%	24.08%	33.99%	44.19%	54.86%
600	3.62%	9.76%	15.76%	21.28%	26.62%	37.17%	47.82%	58.69%
700	4.04%	10.80%	17.29%	23.20%	28.87%	39.91%	50.84%	61.75%
800	4.44%	11.75%	18.68%	24.93%	30.86%	42.28%	53.38%	64.20%

TABLE 2.2: Economic capital EC_α for an infinitely fine-grained portfolio (portfolio loss $L \sim F_{p,\varrho}$) w.r.t. p and ϱ, for $\alpha = 99.98\%$.

99.98%	ρ in %							
p in bps	**1%**	**5%**	**10%**	**15%**	**20%**	**30%**	**40%**	**50%**
10	0.20%	0.82%	1.79%	3.01%	4.50%	8.34%	13.49%	20.22%
20	0.36%	1.41%	2.99%	4.91%	7.18%	12.88%	20.27%	29.60%
30	0.51%	1.94%	4.01%	6.47%	9.35%	16.39%	25.27%	36.18%
40	0.65%	2.41%	4.91%	7.84%	11.20%	19.30%	29.29%	41.27%
50	0.78%	2.86%	5.74%	9.07%	12.85%	21.83%	32.68%	45.41%
60	0.90%	3.28%	6.51%	10.19%	14.35%	24.07%	35.62%	48.90%
70	1.03%	3.67%	7.23%	11.24%	15.72%	26.08%	38.21%	51.89%
80	1.14%	4.05%	7.90%	12.22%	17.00%	27.92%	40.52%	54.51%
90	1.26%	4.42%	8.55%	13.14%	18.19%	29.61%	42.61%	56.82%
100	1.37%	4.77%	9.17%	14.01%	19.30%	31.17%	44.51%	58.88%
150	1.90%	6.36%	11.91%	17.81%	24.09%	37.62%	52.05%	66.63%
200	2.38%	7.77%	14.24%	20.95%	27.94%	42.55%	57.45%	71.75%
250	2.83%	9.03%	16.28%	23.65%	31.18%	46.50%	61.57%	75.39%
300	3.25%	10.19%	18.11%	26.02%	33.97%	49.78%	64.81%	78.07%
350	3.64%	11.26%	19.77%	28.13%	36.41%	52.54%	67.43%	80.09%
400	4.02%	12.26%	21.29%	30.03%	38.57%	54.91%	69.58%	81.64%
450	4.38%	13.19%	22.69%	31.75%	40.51%	56.96%	71.36%	82.84%
500	4.73%	14.07%	23.99%	33.33%	42.25%	58.74%	72.85%	83.76%
600	5.38%	15.68%	26.32%	36.10%	45.27%	61.70%	75.15%	85.01%
700	5.98%	17.14%	28.36%	38.48%	47.78%	64.01%	76.78%	85.69%
800	6.54%	18.45%	30.17%	40.53%	49.89%	65.83%	77.92%	85.98%

TABLE 2.3: Unexpected loss UL for an infinitely fine-grained portfolio (portfolio loss $L \sim F_{p,\varrho}$) w.r.t. p and ϱ.

	ρ in %							
p in bps	**1%**	**5%**	**10%**	**15%**	**20%**	**30%**	**40%**	**50%**
10	0.03%	0.08%	0.14%	0.19%	0.24%	0.37%	0.53%	0.73%
20	0.06%	0.16%	0.25%	0.33%	0.43%	0.63%	0.88%	1.17%
30	0.09%	0.22%	0.35%	0.47%	0.59%	0.86%	1.18%	1.54%
40	0.12%	0.29%	0.45%	0.59%	0.75%	1.07%	1.44%	1.87%
50	0.15%	0.35%	0.54%	0.71%	0.89%	1.27%	1.69%	2.17%
60	0.17%	0.41%	0.63%	0.83%	1.03%	1.46%	1.93%	2.45%
70	0.20%	0.47%	0.72%	0.94%	1.17%	1.64%	2.15%	2.72%
80	0.22%	0.53%	0.80%	1.05%	1.30%	1.81%	2.36%	2.98%
90	0.25%	0.58%	0.88%	1.15%	1.42%	1.98%	2.57%	3.22%
100	0.27%	0.64%	0.96%	1.26%	1.55%	2.14%	2.77%	3.46%
150	0.38%	0.90%	1.34%	1.74%	2.12%	2.88%	3.68%	4.52%
200	0.49%	1.14%	1.70%	2.18%	2.65%	3.56%	4.49%	5.47%
250	0.59%	1.37%	2.03%	2.60%	3.14%	4.18%	5.23%	6.32%
300	0.69%	1.59%	2.34%	2.99%	3.60%	4.76%	5.92%	7.11%
350	0.78%	1.80%	2.65%	3.37%	4.04%	5.31%	6.57%	7.85%
400	0.87%	2.00%	2.94%	3.73%	4.45%	5.83%	7.18%	8.55%
450	0.95%	2.20%	3.21%	4.07%	4.85%	6.33%	7.77%	9.21%
500	1.04%	2.38%	3.48%	4.40%	5.24%	6.81%	8.32%	9.84%
600	1.20%	2.74%	3.99%	5.03%	5.97%	7.71%	9.37%	11.02%
700	1.35%	3.08%	4.48%	5.62%	6.65%	8.55%	10.34%	12.11%
800	1.49%	3.41%	4.93%	6.18%	7.30%	9.34%	11.25%	13.13%

Given a uniform one-year average default probability p and a uniform asset correlation ϱ, Tables 2.1 and 2.2 report on the *Economic Capital* (EC) w.r.t. confidence levels of $\alpha = 99,5\%$ and $\alpha = 99,98\%$ for an infinitely fine-grained portfolio (described by the distribution $F_{p,\varrho}$), hereby assuming an LGD of 100% (see Section 1.2.1 for the definition of EC). Analogously, Table 2.3 shows the *Unexpected Loss* for a given pair (p, ϱ).

Figure 2.6 illustrates the sensitivity of the EC w.r.t. the chosen confidence level. It can be seen that at high levels of confidence (e.g., from 99,9% on) the impact of every basispoint increase of α on the portfolio EC is enormous.

Another common portfolio-dependent quantity is the so-called *capital multiplier* (CM_α); see also Chapter 5 on *capital allocation*. It is defined as the *EC w.r.t. confidence α in units of UL* (i.e., in units of the portfolio standard deviation). In pricing tools the CM is sometimes assumed to be constant for a portfolio, even when adding new deals to it. The contribution of the new deal to the total EC of the enlarged portfolio is then given by a multiple of the CM. In general, the CM heavily depends on the chosen level of confidence underlying the EC definition. Because for given p and ϱ the CM is just the EC scaled by the inverse of the UL, Figure 2.6 additionally illustrates the shape of the curve describing the dependency of the CM from the assumed level of confidence.

For example, for p=30 bps (about a BBB-rating) and ϱ=20% (the Basel II suggestion for the asset correlation of the benchmark risk weights for corporate loans) the (rounded!) CM of a portfolio with loss variable $L \sim F_{p,\varrho}$ is given by $\mathrm{CM}_{99\%} \approx 4$, $\mathrm{CM}_{99,5\%} \approx 6$, $\mathrm{CM}_{99,9\%} \approx 10$, and $\mathrm{CM}_{99,98\%} \approx 16$ (in this particular situation we have an UL of 59 bps, as can be read from the Figure 2.3).

Now, as a last remark in this section we want to refer back to Section 1.2.2.2, where the *analytical approximation* of portfolio loss distributions is outlined. The distribution $L_{p,\varrho}$, eventually combined with some modifications (e.g., random or deterministic LGDs), is extremely well suited for analytical approximation techniques in the context of asset value (or more generally latent variable) models.

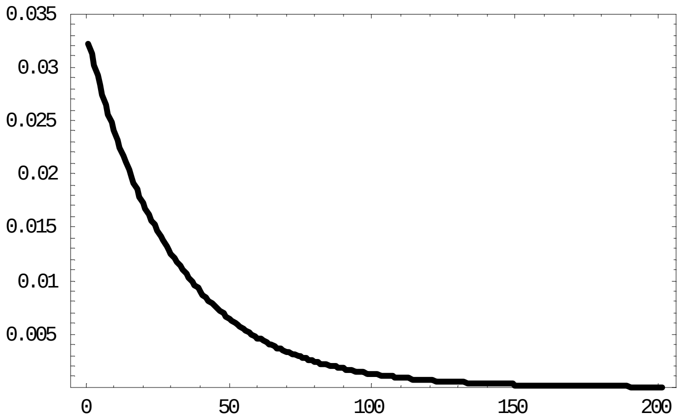


FIGURE 2.7: Negative binomial distribution with parameters $(\alpha, \beta) = (1, 30)$.

2.5.2 The CreditRisk$^+$ One-Sector Model

We already discussed CreditRisk$^+$ in Section 2.4.2 and will come back to it in Chapter 4. Therefore this paragraph is just a brief "warming-up" for the next paragraph where we compare the uniform portfolio loss distributions from an asset value model with the corresponding distribution in the CreditRisk$^+$ world.

Assuming infinitely many obligors and only one sector, we obtain a situation comparable to the uniform portfolio model based on an asset value model.

Under these assumptions, the portfolio loss is distributed according to a *negative binomial distribution* $NB(\alpha, \beta)$ due to a gamma-distributed random intensity. The derivation of the negative binomial distribution in the CreditRisk$^+$ framework is extensively discussed in Chapter 4. Denoting the portfolio loss by $L' \sim NB(\alpha, \beta)$, the loss distribution is determined by

$$\mathbb{P}[L' = n] \;=\; \binom{n + \alpha - 1}{n} \left(1 - \frac{\beta}{1+\beta}\right)^{\alpha} \left(\frac{\beta}{1+\beta}\right)^{n}, \qquad (2.58)$$

where α and β are called the *sector parameters* of the sector; see Formula (4.26). The expectation and the variance of L' are given by

$$\mathbb{E}[L'] \;=\; \alpha\beta \quad \text{and} \quad \mathbb{V}[L'] \;=\; \alpha\beta(1+\beta)\;, \qquad (2.59)$$

as derived in Formula (4.27). Figure 2.7 illustrates the shape of the probability mass function of a negative binomial distribution, here with parameters $\alpha = 1$ and $\beta = 30$. The expected loss in a portfolio admitting such a loss distribution is

$$\mathrm{EL} \;=\; \mathbb{E}[L'] \;=\; 1 \times 30 \;=\; 30,$$

and the unexpected loss (volatility of the portfolio loss) is

$$\mathrm{UL} \;=\; \sqrt{\mathbb{V}[L']} \;=\; \sqrt{1 \times 30 \times (1+30)} \;=\; 30.5.$$

We are now prepared for the next section.

2.5.3 Comparison of One-Factor and One-Sector Models

Recalling the discussion about general mixture models at the beginning of this chapter one could say that in this section we compare Bernoulli and Poisson mixture models by means of a typical example.

As a representative for the Bernoulli mixture models we choose the random variable $L \sim F_{p,\varrho}$ describing the percentage loss of an infinitely fine-grained portfolio with uniform default probability p and uniform asset correlation ϱ; see (2.55). Such portfolios typically arise in analytical approximations in the asset value model framework.

The one-sector model of CreditRisk$^+$ as described in the previous paragraph will serve as a representative for Poisson mixture models.

A very natural way to calibrate the two models on a common basis is by *moment matching*. One problem we face here is that L takes place in the unit interval and L' generates random integers. We overcome this problem by fixing some large m, say 20,000, such that the tail probability $\mathbb{P}[L' > m]$ is negligibly small, and transforming L' into a variable

$$\tilde{L}' \;=\; \frac{L'}{m}\;.$$

So we take $\tilde{L}'$ as a proxy for the percentage portfolio loss in the one-sector model in CreditRisk$^+$. The moment matching procedure is based on the conditions

$$\mathbb{E}[L] \;=\; \mathbb{E}[\tilde{L}'] \qquad \text{and} \qquad \mathbb{V}[L] \;=\; \mathbb{V}[\tilde{L}']\;.$$

Hereby we always start with some p and ϱ specifying the distribution of L. We then set

$$\mathbb{E}[\tilde{L}'] \;=\; p\;, \qquad \mathbb{V}[\tilde{L}'] \;=\; N_2\big[N^{-1}[p], N^{-1}[p]; \varrho\big] - p^2\;,$$

applying Proposition 2.5.9. As a last step we solve (2.59) for α and β. One always has

$$\alpha = \frac{m \times \mathbb{E}[\tilde{L}']^2}{m \times \mathbb{V}[\tilde{L}'] - \mathbb{E}[\tilde{L}']}\;, \quad \beta = \frac{m \times \mathbb{V}[\tilde{L}'] - \mathbb{E}[\tilde{L}']}{\mathbb{E}[\tilde{L}']}\;, \tag{2.60}$$

e.g., for p=30 bps, ϱ=20%, and m=20,000 we apply 2.5.9 for

$$\mathbb{V}[L] \;=\; N_2\big[N^{-1}[0.003], N^{-1}[0.003]; 0.2\big] - 0.003^2 \;=\; 0.000035095.$$

The unexpected loss of L therefore turns out to be UL=59 bps. Applying Formulas (2.60), we get

$$\alpha = 0.26 \qquad \text{and} \qquad \beta = 232.99,$$

so that the distribution of $\tilde{L}'$ is finally determined.

In Table 2.4 high-confidence quantiles of one-factor respectively one-sector models with different parameter settings are compared. It turns out that the Bernoulli mixture model always yields fatter tails than the Poisson mixture model, hereby confirming our theoretical results from Section 2.3. A more detailed comparison of the two model variants can be found in [25].

2.6 Loss Dependence by Means of Copula Functions

Copula functions have been used as a statistical tool for constructing multivariate distributions long before they were rediscovered as a valuable technique in risk management. Currently, the literature on the application of copulas to credit risk is growing every month, so that tracking every single paper on this issue starts being difficult if not impossible. A small and by no means exhaustive selection of papers providing the reader with a good introduction as well as with a valuable source of ideas about how to apply the copula concept to

TABLE 2.4: Comparison of Bernoulli and Poisson mixture models by means of one-factor respectively one-sector models.

portfolio	p	rho	sigma
1	0.01%	10%	0.02%
2	0.01%	20%	0.04%
3	0.01%	30%	0.06%
4	0.30%	10%	0.35%
5	0.30%	20%	0.59%
6	0.30%	30%	0.86%
7	1.00%	10%	0.96%
8	1.00%	20%	1.55%
9	1.00%	30%	2.14%

portfolio	alpha	beta	Q_[99.98%] (KMV)	Q_[99.98%] (CR+)
1	0.37	5.38	0.31%	0.19%
2	0.08	25.35	0.85%	0.59%
3	0.03	78.17	1.67%	1.42%
4	0.75	80.25	4.30%	3.14%
5	0.26	232.99	9.65%	6.84%
6	0.12	496.04	16.69%	12.20%
7	1.09	184.32	10.17%	8.11%
8	0.42	476.85	20.30%	15.76%
9	0.22	911.68	32.17%	25.69%

standard problems in credit risk is Li [124, 125], Frey and McNeil [68], Frey, McNeil, and Nyfeler [70], Frees and Valdez [67], and Wang [185].

In a recent book [24] two of us explain the concept of copula functions in detail and apply different copulas to various modeling problems in the field of *structured credit portfolios*. It turns out that especially structured credit products like collateralized debt obligations (CDOs) and other asset-backed securities are very *sensitive* to the choice of the particular underlying copula function.

When analyzing the most recent crisis (see also [26]) one finds that the *mis-specification of dependency models* is one reason why banks were caught more or less by surprise regarding the large losses accumulated by structured credit products. For credit risk modelers who applied different copula functions in their models for testing what could happen to structured credit markets, it was at all times clear that there was a huge potential for losses in certain products. However, because the model results that made it into official reports and infokits for executive management were based in many cases on the standard copula we introduce in a moment (namely, the Gaussian copula), such loss potential remained unrevealed to decision makers in many cases. It is important that credit risk modeling in the future does consider certain types of copulas not only as a "modeling toy" for quants but uses them to demonstrate what can happen to a tail of a portfolio loss distribution under certain eventually unfortunate circumstances; see also the discussion on copula choice impact included in Chapter 8.

We start our exposition with the following definition.

2.6.1 Definition *A copula (function) is a multivariate distribution (function) such that its marginal distributions are standard uniform. A common notation for copulas we will adopt is*

$$C(u_1, ..., u_m) : [0,1]^m \to [0,1]$$

if considered in R^m (e.g., m obligors, m assets, m latent variables, etc.).

The most commonly applied copula function (e.g., in common asset value models) is the *normal* or *Gaussian copula*, defined by

$$C(u_1, ..., u_m) \;=\; N_m\big[N^{-1}[u_1], ..., N^{-1}[u_m]; \Gamma\big] \;, \tag{2.61}$$

with $N_m(\,\cdot\,;\Gamma)$ as in Section 2.4.1. In this section we also elaborated that asset value models in the classical original setup implicitly incorporate copula functions based on the multivariate Gaussian distribution of asset value processes. For example, Proposition 2.5.9 says that the *bivariate normal copula* determines the second moment of the loss distribution of an infinitely fine-grained portfolio. So we implicitly already met copulas in previous paragraphs. Before continuing, we need to quote a Theorem by SKLAR [168, 169], saying that copulas are a universal tool for studying multivariate distributions.

2.6.2 Theorem (SKLAR [168]) *Let $\boldsymbol{F}$ be a multivariate m-dimensional distribution function with marginals $F_1, ..., F_m$. Then there exists a copula C such that*

$$\boldsymbol{F}(x_1, ..., x_m) = C\Big(F_1(x_1), ..., F_m(x_m)\Big) \qquad (x_1, ..., x_m \in \mathbb{R}).$$

Moreover, if the marginal distributions $F_1, ..., F_m$ are continuous, then C is unique.

Proof. For a comprehensive proof see SKLAR [168], or, alternatively, the textbook [144] by Nelsen. However, the basic idea (which is already the heart of the proof) of deriving a copula from a given multivariate distribution $\boldsymbol{F}$ with marginals $F_1, ..., F_m$ is by imitating what we previously have seen in case of the normal copulas, namely

$$C(u_1, ..., u_m) = \boldsymbol{F}\Big(F_1^{-1}(u_1), ..., F_m^{-1}(u_m)\Big) \,. \tag{2.62}$$

Now one only has to confirm that C defined by (2.62) does the job. □

The converse of Proposition 2.6.2 is also true:

2.6.3 Proposition *For any copula C and (marginal) distribution functions $F_1, ..., F_m$, the function*

$$\boldsymbol{F}(x_1, ..., x_m) = C\Big(F_1(x_1), ..., F_m(x_m)\Big) \qquad (x_1, ..., x_m \in \mathbb{R})$$

defines a multivariate distribution function with marginals $F_1, ..., F_m$.

Proof. The proof is straightforward. One just has to apply the defining properties of copulas. □

Summarizing Theorem 2.6.2 and Proposition 2.6.3, one can say that every multivariate distribution with continous marginals admits a unique *copula representation*. Moreover, copulas and distribution functions are the building blocks to derive new multivariate distributions with prescribed correlation structure and marginal distributions.

The copula representation of multivariate distributions according to SKLAR's theorem indicates why copulas became so prominent and omnipresent in credit risk modeling and finance in general. One could argue that the overall purpose of fitting a credit model to data remains the old fashioned problem of finding a best fitting multivariate distribution, no matter if one looks at copula and marginals in joint or separate ways. However, *separating marginals and copula of a multivariate* *distribution is especially useful in situations where the marginals are known (which is often the case) but the multivariate dependence is still to be determined.* In such cases it makes a lot of sense to focus on the copula part of a multivariate distribution; see also [26].

2.6.1 Copulas: Variations of a Scheme

In this section we are mainly interested in giving some examples of how the copula approach can be used for constructing loss distributions with fatter tails than it would be for normally distributed asset value log-returns. For this section we focus on normal and t-copulas, because they are most common in the credit risk context. Many more classes of copulas can be found in the book by Nelsen [144].

For our example we look at a Bernoulli mixture model but replace the multivariate normal asset value log-return vector by a multivariate *t-distributed* log-return vector. For the convenience of the reader we first recall some basic test distributions from statistics (see, e.g., [155]):

The Chi-Square Distribution

The χ^2-distribution can be constructed as follows: Start with an i.i.d. sample $X_1, ..., X_n \sim N(0,1)$. Then, $X_1^2 + \cdots + X_n^2$ is said to be χ^2-distributed with n degrees of freedom. The first and second moments of a random variable $X \sim \chi^2(n)$ are

$$\mathbb{E}[X] = n \qquad \text{and} \qquad \mathbb{V}[X] = 2n \; .$$

In some sense the χ^2 distribution is a "derivate" of the gamma-distribution (see 2.4.2), because the $\chi^2(n)$-distribution equals the gamma-distribution with parameters $\alpha = n/2$ and $\beta = 2$. Therefore we already know the shape of χ^2-densities from Figure 2.2.

The (Student's) t-Distribution

The building blocks of the t-distribution are a standard normal variable $Y \sim N(0,1)$ and a χ^2-distributed variable $X \sim \chi^2(n)$, such that Y and X are independent. Then the variable Z defined by $Z = Y/\sqrt{X/n}$ is said to be t-distributed with n degrees of freedom. The density of Z is given by

$$t_n(x) \;=\; \frac{\Gamma((n+1)/2)}{\sqrt{\pi n}\,\Gamma(n/2)}\left(1+\frac{x^2}{n}\right)^{-(n+1)/2} \qquad (x \in \mathbb{R}).$$

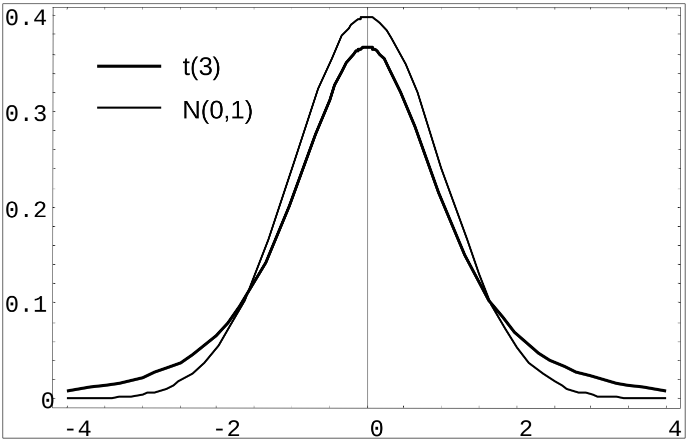


FIGURE 2.8: $t(3)$-density versus $N(0,1)$-density.

The first and second moments of a variable $Z \sim t(n)$ are given by

$$\mathbb{E}[Z] = 0 \quad (n \geq 2) \qquad \text{and} \qquad \mathbb{V}[Z] = \frac{n}{n-2} \quad (n \geq 3).$$

For large n, the t-distribution is close to the normal distribution. More precisely, if F_n denotes the distribution function of a random variable $Z_n \sim t(n)$, then one can show that F_n converges in distribution to the distribution function of a standard normal random variable $Z \sim N(0,1)$; see [155].

This convergence property is a nice result, because it enables us to start in the following modification of the classical asset value model with a parameterization close to the normal case by looking at a large n. By systematically decreasing the degrees of freedom we can transform the model step-by-step towards a model with fatter and fatter tails.

In general the t-distribution has more mass in the tails than a normal distribution. Figure 2.8 illustrates this by comparing a standard normal density with the density of a t-distribution with 3 degrees of freedom.

The Multivariate t-Distribution:

Given a multivariate Gaussian vector $\boldsymbol{Y} = (Y_1, ..., Y_m) \sim N(0, \Gamma)$ with correlation matrix Γ, the scaled vector $\Theta \boldsymbol{Y}$ is said to be multivariate t-distributed with n degrees of freedom if $\Theta = \sqrt{n/X}$ with

$X \sim \chi^2(n)$ and Θ is independent of $\boldsymbol{Y}$. We denote the distribution of such a variable $\Theta\boldsymbol{Y}$ by $t(n,\Gamma)$. The matrix Γ is explicitly addressed as the second parameter, because $\Theta\boldsymbol{Y}$ inherits the correlation structure from $\boldsymbol{Y}$:

$$\mathrm{Corr}[\Theta Y_i, \Theta Y_j] \;=\; \mathrm{Corr}[Y_i, Y_j]\;.$$

This can be easily verified by a conditioning argument (w.r.t. Θ).

Now let $n \geq 3$ and F_n be a t-distribution function with n degrees of freedom. Denote the inverse or quantile function of F_n by F_n^{-1}. For the multivariate t-distribution function with n degrees of freedom and correlation matrix Γ we write $F_{n,\Gamma} \sim t(n,\Gamma)$. Then we can define a *t-copula* function $C_{n,\Gamma}$ by applying formula (2.62),

$$C_{n,\Gamma}(u_1,...,u_m) \;=\; F_{n,\Gamma}\Big(F_n^{-1}(u_1),...,F_n^{-1}(u_m)\Big) \tag{2.63}$$

$$\text{where} \qquad u_1,...,u_m \in [0,1].$$

The copula $C_{n,\Gamma}$ incorporates a *multivariate t-dependency* that we can now combine with any marginal distributions we like. For example, a multivariate distribution function with t-dependency and Gaussian marginals can be defined (for $x_1,...,x_m \in \mathbb{R}$) by

$$\Phi_{n,\Gamma}(x_1,...,x_m) \;=\; C_{n,\Gamma}(N[x_1],...,N[x_m]) \tag{2.64}$$

where $N[\cdot]$ denotes the standard normal distribution function. That indeed $\Phi_{n,\Gamma}$ defines a multivariate distribution function with standard normal marginals is a direct consequence of Proposition 2.6.3. Replacing a normal by a t-dependency will – in accordance with the fact that t-tails are fatter than Gaussian tails – significantly shift mass into the tails of the loss distribution arising from a corresponding asset value model. The fatness of tails strongly depends on the chosen degrees of freedom, so that the calibration of an appropriate n in Formula (2.64) is an essential challenge when dealing with t-copulas. Although there is much literature about the calibration of non-normal distributions to financial time series in general (see, e.g., EBERLEIN [54]), so far we do not know about an established standard calibration methodology for fitting t-copulas to a credit portfolio. Here we believe that some further research is necessary.

The impact of different dependency structures can be best illustrated by means of a scatterplot. In Figure 2.9 we look at four different variations:

- **Bivariate Gaussian copula with normal marginals:**
  We randomly generated points (X_1, X_2) with

$$X_i = \sqrt{\varrho}\, Y + \sqrt{1-\varrho} Z_i \qquad (i = 1, 2)$$

  with $Y, Z_1, Z_2 \sim N(0, 1)$ i.i.d. and $\varrho = 40\%$.

- **Bivariate t-copula with t-distributed marginals:**
  Here we plotted randomly generated pairs (X_1, X_2) with

$$X_i = \sqrt{3}(\sqrt{\varrho}\, Y + \sqrt{1-\varrho} Z_i)/\sqrt{W} \qquad (i = 1, 2)$$

  with $Y, Z_1, Z_2 \sim N(0, 1)$ i.i.d., $W \sim \chi^2(3)$, and $\varrho = 40\%$.

- **Bivariate t-copula with normal marginal distributions:**
  The points (X_1, X_2) are generated according to

$$X_i = N^{-1}[F_3(\sqrt{3}(\sqrt{\varrho}\, Y + \sqrt{1-\varrho} Z_i)/\sqrt{W})] \qquad (i = 1, 2)$$

  with $Y, Z_1, Z_2 \sim N(0, 1)$ i.i.d., $W \sim \chi^2(3)$, W independent of Y, Z_1, Z_2, $\varrho = 40\%$, and F_3 denoting the t-distribution function with 3 degrees of freedom. Generalizing for m instead of 2 dimensions we obtain a multivariate distribution function F with

$$F(x_1, ..., x_m) = \mathbb{P}[X_1 \leq x_1, ..., X_m \leq x_m]$$

$$= P\left[T_1 \leq F_3^{-1}(N[x_1]), ..., T_m \leq F_3^{-1}(N[x_m])\right]$$

  with $(T_1, ..., T_m) \sim t(3, \Gamma_\varrho)$ and Γ_ϱ denoting the ϱ-uniform correlation matrix in $\mathbb{R}^{m\times m}$. Therefore, we finally see that

$$F(x_1, ..., x_m) = C_{3,\Gamma_\varrho}(N[x_1], ..., N[x_m]) = \Phi_{3,\Gamma_\varrho}(x_1, ..., x_m).$$

  This shows that indeed we simulated copula (2.64) for $n = 3$.

- **Independence copula with normal marginal distributions:**
  We randomly generated points (X_1, X_2) with

$$X_1, X_2 \sim N(0, 1) \quad \text{i.i.d.}$$

  The *independence copula* is defined by $C(u_1, ..., u_m) = u_1 \cdots u_m$.

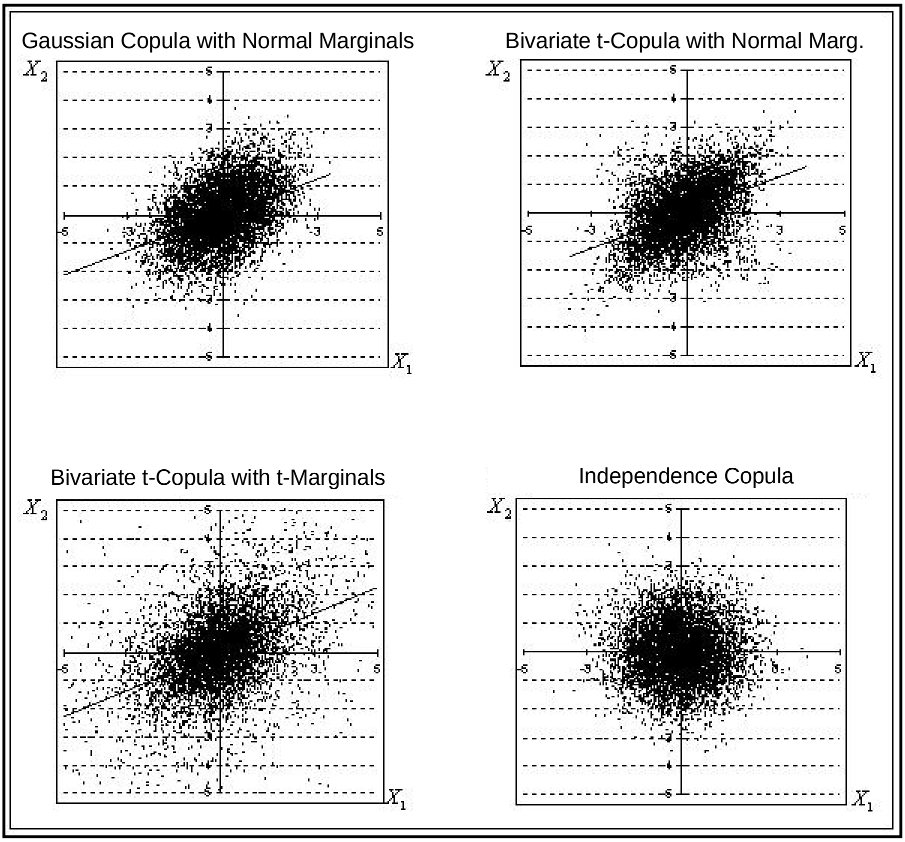


FIGURE 2.9: Normal versus t-dependency with same linear correlation.

TABLE 2.5: Uniform portfolio calculations with t-copulas w.r.t. default probabilities of 50, 80, and 150 basispoints, correlations of 5% and 20%, and degrees of freedom of 10,000, 40, and 10. Quantiles are calculated w.r.t. a confidence of 99%.

Gaussian Copula (not simul.)		
not simulated		
Mean	**5%**	**20%**
0.5%	0.5000%	0.5000%
0.8%	0.8000%	0.8000%
1.5%	1.5000%	1.5000%
Quantile	**5%**	**20%**
0.5%	1.7470%	4.3017%
0.8%	2.6323%	6.2997%
1.5%	4.5250%	10.3283%
Std.Dev.	**5%**	**20%**
0.5%	0.3512%	0.8926%
0.8%	0.5267%	1.2966%
1.5%	0.8976%	2.1205%

T-Copula with df = 10,000		
simulated, 100,000 scenarios		
Mean	**5%**	**20%**
0.5%	0.5002%	0.4983%
0.8%	0.8028%	0.8037%
1.5%	1.5034%	1.4944%
Quantile	**5%**	**20%**
0.5%	1.7605%	4.3060%
0.8%	2.6351%	6.3342%
1.5%	4.5287%	10.1864%
Std.Dev.	**5%**	**20%**
0.5%	0.3522%	0.8946%
0.8%	0.5283%	1.3045%
1.5%	0.8964%	2.0988%

T-Copula with df = 40		
simulated, 100,000 scenarios		
Mean	**5%**	**20%**
0.5%	0.4959%	0.4992%
0.8%	0.8006%	0.8009%
1.5%	1.5030%	1.4970%
Quantile	**5%**	**20%**
0.5%	2.9674%	5.3814%
0.8%	4.2611%	7.6405%
1.5%	6.6636%	11.9095%
Std.Dev.	**5%**	**20%**
0.5%	0.6145%	1.1201%
0.8%	0.8802%	1.5726%
1.5%	1.3723%	2.4653%

T-Copula with df = 10		
simulated, 100,000 scenarios		
Mean	**5%**	**20%**
0.5%	0.4990%	0.4973%
0.8%	0.7999%	0.8051%
1.5%	1.5023%	1.5003%
Quantile	**5%**	**20%**
0.5%	6.0377%	7.9295%
0.8%	8.0921%	11.2434%
1.5%	11.7042%	16.5620%
Std.Dev.	**5%**	**20%**
0.5%	1.2535%	1.7135%
0.8%	1.6574%	2.3104%
1.5%	2.3889%	3.3475%

Figure 2.9 clearly illustrates the impact of different copula variations. For example, the t-copula with normal marginals more or less keeps the Gaussian "tightness" of the point cloud but tends to have greater tail dependency.

We now want to investigate the impact of a change from a normal to a t-copula on the loss distribution of a Bernoulli mixture model based on a uniform asset value model. For this purpose we choose a default probability p, an asset correlation ϱ, and fix n degrees of freedom. Our starting point is Formula (2.48),

$$r_i \;=\; \sqrt{\varrho}\, Y + \sqrt{1-\varrho}\, Z_i \qquad (i = 1, ..., m).$$

By scaling this equation with $\sqrt{n/W}$, $W \sim \chi^2(n)$, W independent of $Y, Z_1, ..., Z_m$, we transform the normal copula into a t-copula yielding t-distributed asset value log-returns $(\tilde{r}_1, ..., \tilde{r}_m) \sim t(n, \Gamma_\varrho)$,

$$\tilde{r}_i \;=\; \sqrt{n/W}\; r_i \;=\; \sqrt{n/W}\sqrt{\varrho}\, Y + \sqrt{n/W}\sqrt{1-\varrho}\, Z_i \;\sim\; t(n)$$

for $i = 1, ..., m$. Again denoting the t-distribution function of $t(n)$ by F_n we can write the default point of the model as $F_n^{-1}(p)$. The Bernoulli loss variables are given by $L_i = \mathbf{1}_{\{\tilde{r}_i \le F_n^{-1}(p)\}}$. The uniform default probability conditional on $Y = y$ and $W = w$ has the following representation:

$$\begin{aligned}
p(y, w) \;&=\; \mathbb{P}[L_i = 1 \mid Y = y, W = w] \\
&=\; \mathbb{P}[\tilde{r}_i \le F_n^{-1}(p) \mid Y = y, W = w] \\
&=\; P\Big[\sqrt{n/W}\sqrt{\varrho}\, Y + \sqrt{n/W}\sqrt{1-\varrho}\, Z_i \le F_n^{-1}(p) \mid Y = y, W = w\Big] \\
&=\; N\left[\frac{\sqrt{W/n}\, F_n^{-1}(p) - \sqrt{\varrho}\, Y}{\sqrt{1-\varrho}} \mid Y = y, W = w\right] \\
&=\; N\left[\frac{\sqrt{w/n}\, F_n^{-1}(p) - \sqrt{\varrho}\, y}{\sqrt{1-\varrho}}\right] .
\end{aligned}$$

Analogous to the conclusion in (2.55), not only the single obligor's conditional default probability but in the limit also, the portfolio's percentage loss is described by $p(y, w)$ given $Y = y$ and $W = w$. We therefore can simulate the portfolio loss in a t-copula model by looking at the distribution of, say, 100,000 samples

$$N\left[\frac{F_n^{-1}[p]\sqrt{W_i/n}-\sqrt{\varrho}\,Y_i}{\sqrt{1-\varrho}}\right], \tag{2.65}$$

$$W_1,...,W_{100{,}000}\sim\chi^2(n),\quad Y_1,...,Y_{100{,}000}\sim N(0,1)\quad\text{independent.}$$

We have done this exercise for different p's, ϱ's, and n's, and the result is shown in Table 2.5. In the table one can see that for 10,000 degrees of freedom the difference of the portfolio statistics compared to a normal copula is very small and just due to stochastic fluctuations in the simulation. But with decreasing n the portfolio statistics significantly changes. For example, there is a multiplicative difference of almost a factor of 2 between the 99%-quantiles w.r.t. $(p,\varrho)=(0.8\%,5\%)$ and degrees of freedom of 40 and 10. If we would calculate the quantiles in Table 2.5 w.r.t. higher levels of confidence, the differences would be even higher. Therefore one can easily increase the potential for extreme losses in a uniform t-copula portfolio model by just decreasing the degrees of freedom of the underlying multivariate t-distribution. Unfortunately, a decision of how fat the tails really should be is never easy and sometimes purely subjective. Maybe this is the reason why people very often rely on asset value models based on the Gaussian copula. Gaussian distributions are uniquely determined by their expectation vector and their covariance matrix, such that more complicated calibrations are not necessary. Moreover, as we already indicated, often even the estimation of linear correlations is a great challenge and far from being obvious. We believe that more research combined with empirical evidence is necessary before other than normal copulas will become "best practice" in credit risk management.

Our last point in this section is the following proposition.

2.6.4 Proposition *Given a Bernoulli loss statistics $(L_1,...,L_m)$ based on an asset value respectively (more general) latent variables model in the form $L_i=\mathbf{1}_{\{r_i\le c_i\}}$, the gross loss distribution of $(L_1,...,L_m)$, defined as the distribution of the variable $L=\sum L_i$, is uniquely determined by the set of one-year default probabilities $p_i=\mathbb{P}[r_i\le c_i]$ and the respective copula function C of $(r_1,...,r_m)$.*

Proof. The distribution of gross losses arising from the loss statistics $(L_1, ..., L_m)$ is determined by the *joint default probabilities*

$$\mathbb{P}[L_{i_1} = 1, ..., L_{i_k} = 1] \;=\; \mathbb{P}[r_{i_1} \leq c_{i_1}, ..., r_{i_k} \leq c_{i_k}] \;=$$

$$=\; C_{i_1,...,i_k}(p_{i_1}, ..., p_{i_k}), \qquad \text{with} \quad \{i_1, ..., i_k\} \subseteq \{1, ..., m\},$$

where $C_{i_1,...,i_k}$ denotes the respective k-dimensional marginal distribution of the copula C of $(r_1, ..., r_m)$. $\square$

Given the standard case of a one-year time horizon asset value model, Proposition 2.6.4 says that besides the one-year default probabilities the used copula function completely determines the portfolio loss distribution. In a Gaussian world, the asset correlation as a second parameter in uniform asset value models is the main driver of fat tails. For people allowing for other than normal copulas, changing the copula has an even stronger impact than just increasing the asset correlation in a Gaussian model.

We will come back to that point in Chapter 8.

2.7 Working Example on Asset Correlations

We conclude this chapter with a working example regarding the estimation of a proxy for asset correlations from historic default frequencies. It remains difficult in practical situations to get some indication about the order of magnitude of asset correlations. Here, we illustrate a back-of-the-envelope kind of way to gain some idea on how asset correlations can look like. The procedure we are going to elaborate is not a substitute for a full statistical evaluation of data which we would recommend in a real-life situation. However, it nicely illustrates how one can obtain *quick estimates from easily accessible and public data*.

We already saw in Table 1.2 rating agency data from Moody's reporting on historic default frequencies of corporate bond defaults. In the same report [141], Exhibit 39, we also find Table 2.6, showing one-year default rates by year and letter rating from 1970–2000.

What one can clearly see is that observed default frequencies are quite volatile, and a natural interpretation of such volatility is the existence

TABLE 2.6: Moody's Historic Corporate Bond Default Frequencies from 1970 to 2000.

Rating	1970	1971	1972	1973	1974	1975	1976	1977	1978	1979	
Aaa	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	
Aa	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	
A	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	
Baa	0.27%	0.00%	0.00%	0.45%	0.00%	0.00%	0.00%	0.28%	0.00%	0.00%	
Ba	4.12%	0.42%	0.00%	0.00%	0.00%	1.02%	1.01%	0.52%	1.08%	0.49%	
B	22.78%	3.85%	7.14%	3.77%	10.00%	5.97%	0.00%	3.28%	5.41%	0.00%	
	1980	1981	1982	1983	1984	1985	1986	1987	1988	1989	
Aaa	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	
Aa	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.61%	
A	0.00%	0.00%	0.26%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	
Baa	0.00%	0.00%	0.31%	0.00%	0.36%	0.00%	1.33%	0.00%	0.00%	0.60%	
Ba	0.00%	0.00%	2.72%	0.91%	0.83%	1.75%	2.04%	2.71%	1.24%	2.98%	
B	4.94%	4.49%	2.41%	6.31%	6.72%	8.22%	11.80%	6.25%	6.04%	9.21%	
	1990	1991	1992	1993	1994	1995	1996	1997	1998	1999	2000
Aaa	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%
Aa	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%
A	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%
Baa	0.00%	0.28%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.12%	0.10%	0.38%
Ba	3.34%	5.30%	0.30%	0.55%	0.24%	0.67%	0.00%	0.19%	0.61%	1.14%	1.05%
B	16.16%	14.71%	9.03%	5.79%	3.82%	4.80%	1.44%	2.11%	4.26%	5.88%	5.14%

of an *economic cycle*. Although it is the simplest approach, the uniform portfolio model as introduced in Section 2.5.1 already provides us with a useful *parametric framework* in order to estimate the *systematic risk* inherent in Moody's corporate bond portfolio. As a reference for the sequel we mention [20], where several approaches for estimating asset correlations are elaborated.

As already indicated above, we use the uniform portfolio model (see Proposition 8 and the discussion thereafter) as a parametric framework. Table 2.6 includes R_1=Aaa, R_2=Aa, ..., and R_6=B, altogether six rating grades. For every rating class R_i we can calculate the mean $\overline{p}_i$ and the corresponding volatility from the historic default frequencies of class R_i over the years from 1970 to 2000. The result is shown in Tables 2.7 and 2.8 in the mean and standard deviation column.

With Table 2.6 we have the same problem we already faced in Section 1.1.1: There is no default history for upper investment grade bonds. We therefore again fit the historic data by a *linear regression on logarithmic scale*. Hereby we distinguish two regression methods:

- *Regression I:*
  Here we just set R_1=Aaa to "not observed" and fit the mean default frequencies $\overline{p}_2, ..., \overline{p}_6$ by an exponential function yielding fitted default probabilities $\mu_1, ..., \mu_6$ for all rating classes (class Aaa is extrapolated). After that we repeat the same procedure with the volatilities of the default frequency time series of rating classes $R_2, ..., R_6$, this time yielding volatilities $\sigma_1, ..., \sigma_6$ (class Aaa again extrapolated). The results are shown in Table 2.7.

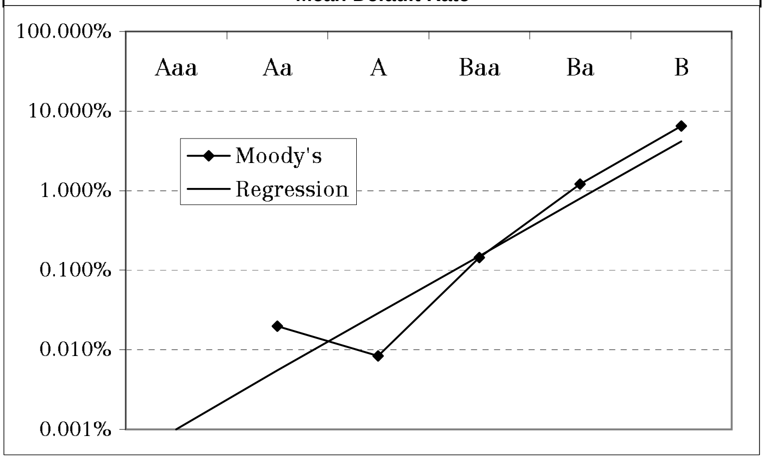
- *Regression II:*
  Regression method II is motivated by the observation that class Aa possibly constitutes an outlier, due to the spike arising from just one observed default frequency in the year 1989. So here we decide to exclude not only Aaa but also Aa from the regression. At the end the default probabilities for Aaa and Aa are extrapolated. Table 2.8 shows the result.

We could continue in this way for rating class A, because in this class we also have only one observation different from zero, namely in year 1982. However, our example is purely illustrative, such that two different regression approaches should be enough to demonstrate the effect. The reason for presenting two approaches is that it shows very clearly

TABLE 2.7: Calibration Results due to Regression I.

Rating	Mean	Stand.Dev.		Σ	
Aaa	0.000%	0.000%	0.0010%	0.0149%	35%
Aa	0.020%	0.110%	0.0055%	0.0442%	32%
A	0.008%	0.047%	0.0288%	0.1313%	28%
Baa	0.145%	0.277%	0.1510%	0.3902%	23%
Ba	1.201%	1.330%	0.7916%	1.1597%	18%
B	6.507%	4.762%	4.1496%	3.4469%	12%
Mean	1.31%	1.09%	0.85%	0.86%	25%

Mean Default Rate



Default Rate Volatility

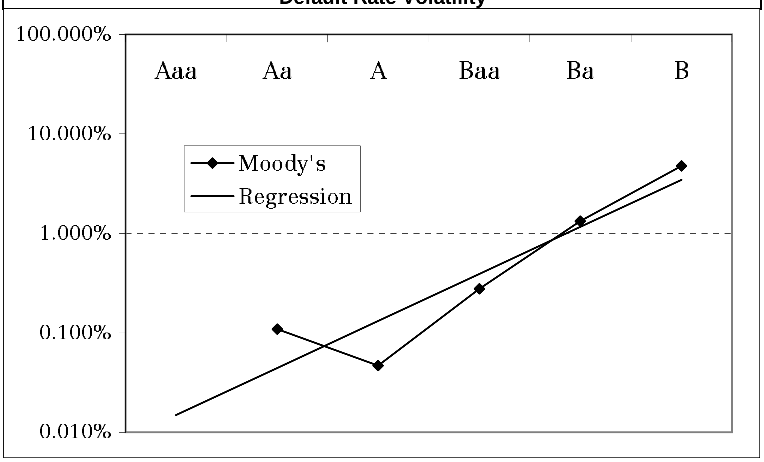
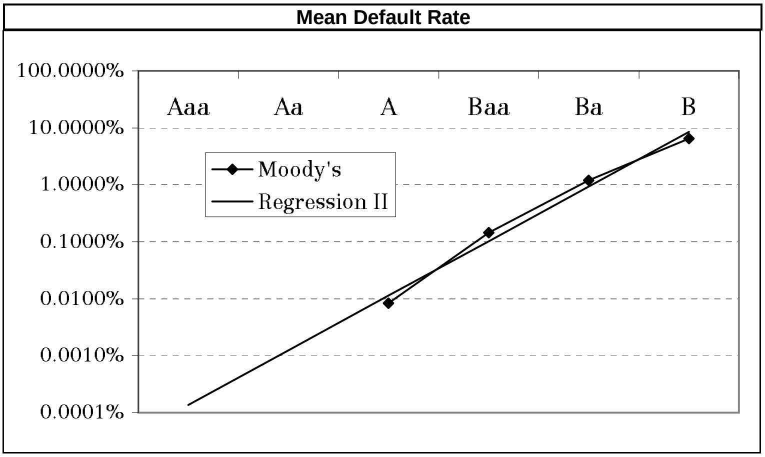
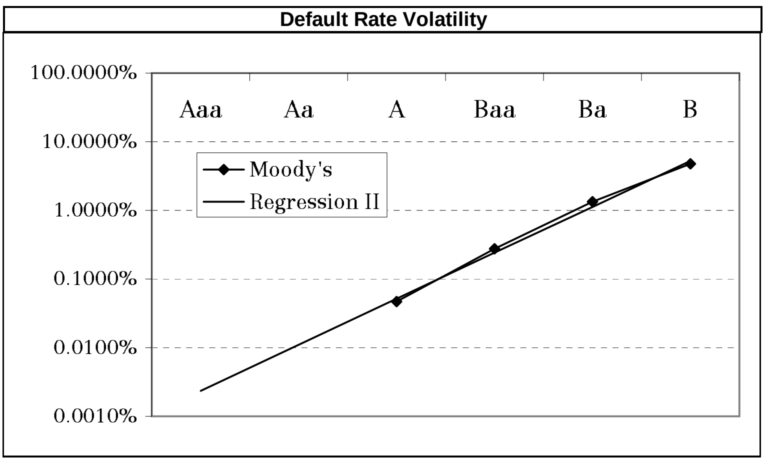


TABLE 2.8: Calibration Results due to Regression II.

Rating	Mean	Stand.Dev.		Σ	
Aaa	0.000%	0.000%	0.0001%	0.0023%	34%
Aa	0.020%	0.110%	0.0012%	0.0110%	28%
A	0.008%	0.047%	0.0113%	0.0514%	24%
Baa	0.145%	0.277%	0.1027%	0.2406%	19%
Ba	1.201%	1.330%	0.9348%	1.1270%	14%
B	6.507%	4.762%	8.5040%	5.2788%	10%
Mean	1.31%	1.09%	1.59%	1.12%	22%





that subjective opinions very often play a crucial role in bank-internal calibrations. In fact, there are various ways in which a regression could have been done in order to obtain reasonable default probabilities for every rating class. So people have to make a decision as to which method best reflects their "expert opinion" and their "analytical honesty." The ρ-columns in Tables 2.7 and 2.8 contain estimated average asset correlations for the considered rating classes, and one can see that the different regression approaches are reflected by differences in estimated asset correlations. For example, people relying on *Regression I* would believe in an overall average asset correlation of 25%, whereas people relying on *Regression II* would believe that the overall average asset correlation in Moody's corporate bond universe is at the lower level of 22%.

Now, it remains to explain how we came up with the asset correlation columns in Tables 2.7 and 2.8. For this purpose let us fix a rating class, such that we can drop the index i referring to rating class i. For the chosen rating class, we know that in year j some default frequency p_j has been observed. The time series $p_1, ..., p_{31}$, addressing the historically observed default frequencies for the chosen rating class in the years 1970 up to 2000, is given by the respective row in Table 2.6. In the uniform portfolio model as we introduced it in (2.55) as a consequence of Proposition 2.5.4, it is assumed that for every year j some realization y_j of a global factor Y drives the realized conditional default probability observed in year j. According to Equation (2.49) we can write

$$p_j \;=\; p(y_j) \;=\; N\left[\frac{N^{-1}[p] - \sqrt{\varrho}\, y_j}{\sqrt{1-\varrho_i}}\right] \qquad (i = 1, ..., m)$$

where p denotes the "true" default probability of the chosen rating class, and ϱ means the unknown asset correlation of the considered rating class, which will be estimated in the following. The parameter p we do not know exactly, but after a moment's reflection it will be clear that the observed historic mean default frequency $\overline{p}$ provides us with a good proxy of the "true" mean default rate. Just note that if $Y_1, ..., Y_n$ are i.i.d.[21] copies of the factor Y, then the law of large

[21] Here we make the simplifying assumption that the economic cycle, represented by $Y_1, ..., Y_n$, is free of *autocorrelation*. In practice one would rather prefer to work with a process incorporating some *intertemporal dependency*, e.g., an AR(1)-process.

numbers guarantees that

$$\frac{1}{n}\sum_{j=1}^{n} p(Y_j) \quad \overset{n\to\infty}{\longrightarrow} \quad \mathbb{E}\big[p(Y)\big] = p \qquad \text{a.s.}$$

Replacing the term on the left side by

$$\overline{p} \;=\; \frac{1}{n}\sum_{j=1}^{n} p_j \;,$$

we see that $\overline{p}$ should be reasonably close to the "true" default probability p. Now, a similar argument applies to the sample variances, because we naturally have

$$\frac{1}{n-1}\sum_{j=1}^{n}\Big(p(Y_j) - \overline{p(Y)}\Big)^2 \quad \overset{n\to\infty}{\longrightarrow} \quad \mathbb{V}\big[p(Y)\big] \qquad \text{a.s.}$$

where $\overline{p(Y)} = \sum p(Y_j)/n$. This shows that the sample variance

$$s^2 \;=\; \frac{1}{n-1}\sum_{j=1}^{n}(p_j - \overline{p})^2$$

should be a reasonable proxy for the "true" variance $\mathbb{V}\big[p(Y)\big]$. Recalling Proposition 2.5.9, we obtain

$$\mathbb{V}\big[p(Y)\big] \;=\; N_2\big[N^{-1}[p], N^{-1}[p]; \varrho\big] - p^2 \;, \tag{2.66}$$

and this is all we need for estimating ϱ. Due to our discussion above we can replace the "true" variance $\mathbb{V}\big[p(Y)\big]$ by the sample variance σ^2 and the "true" default probability p by the sample mean $\overline{p}$. After replacing the unknown parameters p and $\mathbb{V}\big[p(Y)\big]$ by their corresponding estimated values $\overline{p}$ and s^2, the asset correlation ϱ is the only "free parameter" in (2.66). It only remains to solve (2.66) for ϱ. The ϱ-values in Tables 2.7 and 2.8 have been calculated by exactly this procedure, hereby relying on the regression-based estimated values μ_i and σ_i^2. Summarizing, one could say that we estimated asset correlations based on the volatility of historic default frequencies.

As a last calculation we want to infer the economic cycle $y_1, ..., y_n$ for *Regression I*. For this purpose we used an L^2-solver for calculating

$y_1, ..., y_n$ with

$$\sqrt{\sum_{j=1}^{n}\sum_{i=1}^{6}|p_{ij} - p_i(y_j)|^2} \;=\; \min_{(v_1,...,v_n)} \sqrt{\sum_{j=1}^{n}\sum_{i=1}^{6}|p_{ij} - p_i(v_j)|^2}\;,$$

where p_{ij} refers to the observed historic loss in rating class R_i in year j, and $p_i(v_j)$ is defined by

$$p_i(v_j) \;=\; N\left[\frac{N^{-1}[\overline{p}_i] - \sqrt{\varrho_i}\, v_j}{\sqrt{1-\varrho_i}}\right] \qquad (i = 1, ..., 6;\; j = 1, ..., 31).$$

Here, ϱ_i refers to the just estimated asset correlations for the respective rating classes. Figure 2.10 shows the result of our estimation of $y_1, ..., y_n$. In fact, the result is very intuitive: Comparing the economic cycle $y_1, ..., y_n$ with the historic mean default path, one can see that any economic downturn corresponds to an increase of default frequencies.

We conclude our example by a brief remark. Looking at Tables 2.7 and 2.8, we find that estimated asset correlations decrease with decreasing credit quality. Going back to Section 1.3 one finds that this is the assumption of the Basel 2 regulatory capital model when they enforce a strict relationship between PDs and the correlation (or, systematic risk parameter) in the risk-weighting function for RWA. Based on our calculations here, can one say that the regulatory assumption that ϱ should depend on credit quality is reasonable after all?

We think still no, the uniform portfolio model as we introduced it in this chapter truly is a *two-parameter* model without dependencies between p and ϱ. All possible combinations of p and ϱ can be applied in order to obtain a corresponding loss distribution. The rough calculations we made in this section do not support a function between PD and ϱ as it is implemented in the current capital accord because under such strict relationship one gives up the (important!) flexibility to parameterize firms where PD and ϱ are both high or both low.

2.8 Generating the Portfolio Loss Distribution

So far, we mainly talked about how to set up a model for correlated defaults and have spoken only little, apart from simple cases like the

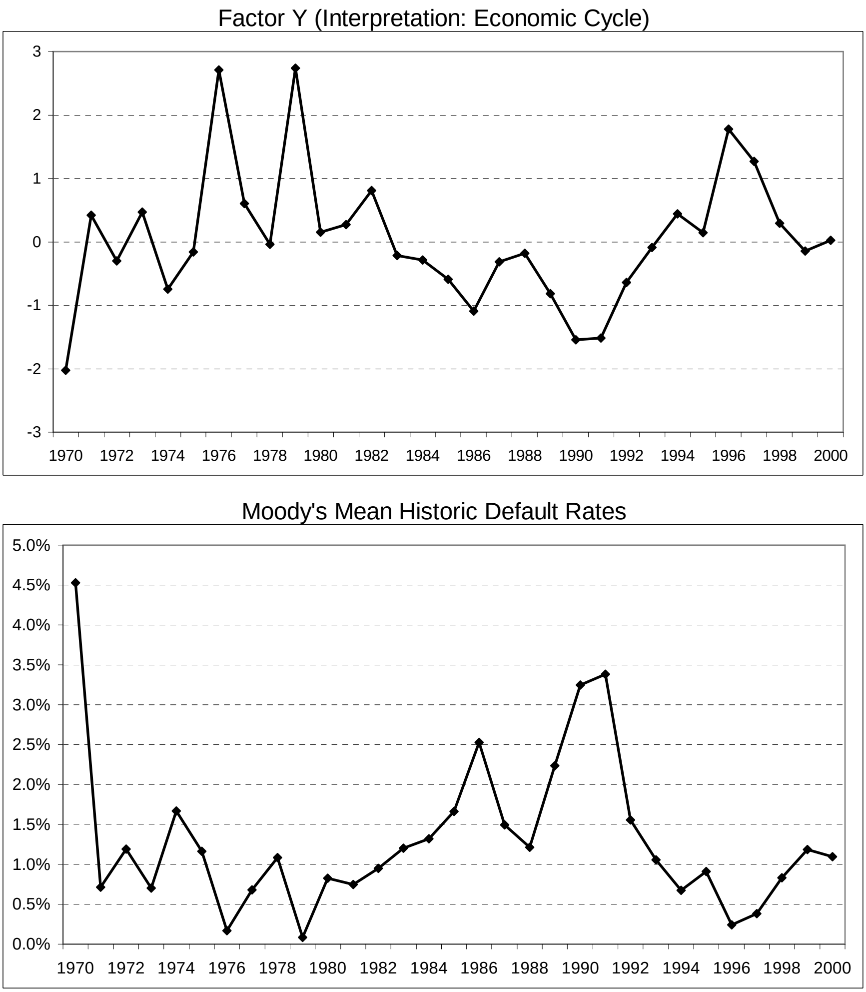


FIGURE 2.10: Estimated economic cycle (top) compared to Moody's average historic default frequencies (bottom).

brief discussion in Chapter 1 on Monte Carlo simulation and analytical approximation, about how to construct an actual loss distribution of a portfolio based on these correlation models. Eventually, we are interested in the portfolio risk in form of a loss distribution and functions on the distribution like risk measures or, e.g., a call loss function like $\max(L-K,0)$ for a certain threshold K as in Chapter 8. Furthermore, before actively managing a credit portfolio it is a major prerequisite to know about single-name contributions to the overall portfolio risk. In other words, an allocation of the portfolio risk to the portfolio constituents is required before one can make decisions regarding portfolio optimization; see also Chapter 5. In this section we discuss some techniques regarding the generation of the portfolio loss distribution of a given portfolio. Each of the techniques has advantages and disadvantages and the best choice depends very much on the objective one has in mind. Note that some techniques used by practitioners allow for a direct computation of some function of losses without having to make the detour of constructing the whole loss distribution first, but we will only treat the generation of whole loss distributions here.

2.8.1 Some Prerequisites from Probability Theory

As a prerequisite for reading this book we mentioned knowledge of concepts from probability. For the sequel we need some techniques used in probability theory to transfer a problem into another space, solving the problem there and then doing the transfer back into the original space. For instance, problems in probability are sometimes translated into the language of *Fourier transforms*, then they are solved in so-called *Fourier space* and, thereafter, they are translated back from Fourier space into the original space where the problem was formulated. A necessary precondition for doing such a problem shift is that the way from a function or measure to its Fourier transform and back again is *unique* so that one can be sure that if two objects are the same in Fourier space they are also the same in the original space. This principle is crucial for the sequel.

For the convenience of the reader and for the sake of introducing the proper notation for our presentation we briefly recall some facts from probability theory before we start with with our topic of this section. It is beyond the scope of this section to provide proofs for all results in the sequel. Readers not familiar with the collection of facts we are going to present are asked to consult an introductory textbook on probability theory for proofs and more details.

The Probability Generating Function

Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space and let X be a random variable in $\mathbb{N}_0$. The (probability) *generating function* G_X of X is defined as

$$G_X(z) \;=\; \mathbb{E}[z^X] \;=\; \sum_{n=0}^{\infty} \mathbb{P}[X = n] z^n$$

for values z for which the series converges. Everything one knows about *power series* can be carried over to generating functions. For instance, there is a radius of convergence such that the series converges absolutely if z is within this radius. In case of generating functions the series has only nonnegative coefficients and their sum equals 1 because they represent the probability mass function of a random variable. As a consequence, the series defining the generating function converges at least for all $z \in \mathbb{R}$ with $|z| \leq 1$.

In accordance with our introductory remarks for this section the generating function is a concept for translating a random variable into an object (a function) in some other space. The question is: can we get X back from its generating function? The answer is yes, and again it is the general theory about power series which gives the answer. In the same way as for power series one gets back X from G_X via

$$\mathbb{P}[X = n] \;=\; \frac{1}{n!}\, G_X^{(n)}(0) \tag{2.67}$$

where $f^{(n)}$ denotes the n-th derivative of a function f. So we have a unique and reversable way to represent a discrete random variable x by its (unique!) generating function which is a convenient and compact form of X carrying exactly the same information as X.

What can we do with generating functions, why is it worth the effort to introduce the concept? A typical answer to that question is that in the other (function) space some operations are easier to carry out. Take two independent random variables X and Y defined in $(\Omega, \mathcal{F}, \mathbb{P})$. The distribution of their sum $X + Y$ is the convolution $\mathbb{P}_X * \mathbb{P}_Y$ of their individual distributions $\mathbb{P}_X$ and $\mathbb{P}_Y$. But $\mathbb{P}_X * \mathbb{P}_Y$ is a more or less complicated object. Well, the following theorem shows that life will be much easier if one switches into the space of generating functions.

2.8.1 Theorem *Let X and Y be independent random variables in $(\Omega, \mathcal{F}, \mathbb{P})$ with values in N_0. Then,*

$$G_{X+Y}(z) \;=\; G_X(z)G_Y(z)$$

for all z for which the generating functions are defined.

Proof. If X, Y are independent then so are z^X and z^Y for any fixed z. For independent random variables the expectation of the product is the product of the individual expectations. Therefore,

$$G_X(z)G_Y(z) \;=\; \mathbb{E}[z^X]\,\mathbb{E}[z^Y] \;=\; \mathbb{E}[z^X z^Y] \;=\; \mathbb{E}[z^{X+Y}] \;=\; G_{X+Y}(z)$$

which is the assertion of the theorem. □

So the generating function of $X + Y$ (in case of independence) is the product of the individual generating functions of the summands. Products are easy to calculate so life is easier when considered from the point of view of generating functions. From (2.67) we know how to calculate the distribution $\mathbb{P}_X * \mathbb{P}_Y$ from G_{X+Y}.

Theorem 2.8.1 can be easily generalized to the sum of finitely many random variables by induction.

We intentionally discussed the concept of generating functions in detail to make the concept of transformation of variables into another space and vice versa as explicit as possible. We conclude our section on generating functions with the two examples which are most important to us given the fact that many credit portfolio models have either the Bernoulli or the Poisson distribution as their basic building block.

2.8.2 Example *Let $L \sim B(1; p)$ be a Bernoulli random variable. Then its generating function is given by*

$$G_L(z) \;=\; 1 + p(z-1).$$

Proof. By definition we have

$$\mathbb{E}[z^L] \;=\; \mathbb{P}[L=0]z^0 + \mathbb{P}[L=1]z\ .$$

But $\mathbb{P}[L=0] = (1-p)$ and $\mathbb{P}[L=1] = p$. □

2.8.3 Example *Let $L' \sim Pois(\lambda)$ be a Poisson random variable. Then its generating function is given by*

$$G_{L'}(z) \;=\; e^{\lambda(z-1)}\;.$$

Proof. We have

$$\mathbb{E}[z^{L'}] \;=\; \sum_{n=0}^{\infty}\mathbb{P}[L'=n]z^n \;=\; \sum_{n=0}^{\infty}\frac{e^{-\lambda}\lambda^n}{n!}\,z^n \;=\; e^{-\lambda}\sum_{n=0}^{\infty}\frac{(\lambda z)^n}{n!}$$

which proves the formula. □

The Moment Generating Function

Another very useful transform of random variables for solving problems in a transform function space is the *moment generating function* of a random variable X defined in $(\Omega, \mathcal{F}, \mathbb{P})$. It is defined by

$$M_X(z) \;=\; \mathbb{E}[e^{zX}]$$

for all $z \in \mathbb{R}$ where the expectation on the right-hand side exists. This is a kind of sloppy formulation which we already used in case of the generating function G_X. It is time to be a bit more formal here. For this we introduce the domain $\mathbb{D}_X$ of M_X for given X as

$$\mathbb{D}_X \;=\; \{z \in \mathbb{R}\;:\;\mathbb{E}[e^{zX}] < \infty\}.$$

Let us see what we can find out about $\mathbb{D}_X$. Because of

$$M_X(0) \;=\; \mathbb{E}[1] \;=\; 1$$

we immediately find that $0 \in \mathbb{D}_X$ for any random variable X. For general X this is it already because one can easily construct an example where $\mathbb{D}_X = \{0\}$. Just take a random variable X in $\mathbb{Z}\backslash\{0\}$ with a (symmetric) probability mass function as follows:

$$p_n = \frac{c\,e^{-z|n|}}{|n|} \quad \text{with} \quad \sum_{n\in\mathbb{Z}\backslash\{0\}} p_n = 1$$

where c is a normalizing constant. For every $z \in \mathbb{R}\backslash\{0\}$ we obtain

$$\mathbb{E}[e^{zX}] \;=\; c\sum_{n=1}^{\infty}\Big(\frac{1}{n}+\frac{e^{-2zn}}{n}\Big) \;>\; c\sum_{n=1}^{\infty}\frac{1}{n} \;=\; +\infty$$

because the harmonic series diverges. So for this particular random variable X we have $\mathbb{D}_X = \{0\}$. In Proposition 2.8.4 we will see that the ideal case is $(-r, r) \subset \mathbb{D}_X$ for some $r > 0$.

For X with range in $\mathbb{N}_0$ we find that the discrete version

$$M_X(z) \;=\; \sum_{n=0}^{\infty} \mathbb{P}[X = n] e^{nz}$$

of the moment generating function is related to the generating function G_X of X via

$$M_X(z) \;=\; G_X(e^z). \tag{2.68}$$

If X is continuous and f_X is its density (w.r.t. Lebesgue measure) then the moment generating function is given by

$$M_X(z) \;=\; \int_D e^{zx} f_X(x) dx \tag{2.69}$$

where D denotes the domain of f_X.

The moment generating function M_X is related to the (two-sided) *Laplace transform*

$$L_X(z) \;=\; \mathbb{E}[e^{-zX}]$$

of X by the simple relation

$$L_X(z) \;=\; M_X(-z) \tag{2.70}$$

for all z where $L_X(z)$ and $M_X(z)$ are defined.

The attribute "moment generating" of the name of M_X is due to the fact that one can obtain all moments (if existent) of X from M_X. The proof is not difficult and is included here.

2.8.4 Proposition *Let X be a random variable in $(\Omega, \mathcal{F}, \mathbb{P})$ and assume that there exists some $r > 0$ such that $(-r, r) \subset \mathbb{D}_X$ which means that M_X is finite in an open interval around 0. Then, the q-th derivative of M_X for all orders $q \in \mathbb{N}$ exists and one has*

$$\mathbb{E}[X^q] \;=\; M_X^{(q)}(0).$$

Recall that $f^{(q)}$ denotes the q-th derivative of a function f.

Proof. We start with a heuristic argument. We would have

$$M_X^{(1)}(z) \;=\; \frac{d}{dz}\,\mathbb{E}[e^{zX}] \;=\; \mathbb{E}\Big[\frac{d}{dz}\,e^{zX}\Big] \;=\; \mathbb{E}\Big[Xe^{zX}\Big]$$

if we knew that expectation and differentiation operators can be interchanged. For $z = 0$ we would then obtain $M_X^{(1)}(0) = \mathbb{E}[X]$. We could then continue this procedure inductively, for instance

$$M_X^{(2)}(z) \;=\; \frac{d}{dz}\,M_X^{(1)}(z) \;=\; \mathbb{E}\Big[\frac{d}{dz}\,(Xe^{zX})\Big] \;=\; \mathbb{E}\Big[X^2e^{zX}\Big]$$

which shows $M_X^{(2)}(0) = \mathbb{E}[X^2]$, and so on. Now, the good news is that interchanging expectation and differentiation can be justified by an integration theorem based on the theorem of dominated convergence; see, e.g., [13], 16.2. So our heuristic argument becomes a true proof. □

We could have followed an alternative route in the proof of Proposition 2.8.4 which we will briefly sketch now. We can use the power series expansion of the exponential function to get

$$M_X(z) \;=\; \mathbb{E}[e^{zX}] \;=\; \mathbb{E}\Big[\sum_{n=0}^{\infty}\frac{z^n}{n!}\,X^n\Big] \;=\; \sum_{n=0}^{\infty}\frac{z^n}{n!}\,\mathbb{E}[X^n] \qquad (2.71)$$

which is a *power series representation of the moment generating function*. Again we need to justify the interchange of the expectation operator and the limit of partial sums of the exponential series but, again, this can be solved by an application of the theorem of dominated convergence. Applying the theory of power series we again obtain

$$M_X^{(q)}(0) \;=\; \mathbb{E}[X^q]$$

for all $q \in \mathbb{N}$ (Taylor series expansion).

In analogy to generating functions we obtain a convolution representation for moment generating functions. The proof is essentially the same as the one for Theorem 2.8.1 and uses just the property $e^{a+b} = e^a e^b$ of the exponential function plus the product formula for the expectation of independent random variables.

2.8.5 Theorem *Let X and Y be independent random variables in $(\Omega, \mathcal{F}, \mathbb{P})$. Then,*

$$M_{X+Y}(z) \;=\; M_X(z)M_Y(z)$$

for all $z \in \mathbb{D}_X \cap \mathbb{D}_Y$.

Proof. We can write

$$M_{X+Y}(z) \;=\; \mathbb{E}[e^{z(X+Y)}] \;=\; \mathbb{E}[e^{zX}]\,\mathbb{E}[e^{zY}] \;=\; M_X(z)M_Y(z)$$

for $z \in \mathbb{D}_X \cap \mathbb{D}_Y$ due to independence of X and Y. □

2.8.6 Proposition *Let X be a random variable in $(\Omega, \mathcal{F}, \mathbb{P})$ and set $Y = aX + b$ for constants $a, b \in \mathbb{R}$. Then,*

$$M_Y(z) \;=\; e^{bz}M_X(az)$$

for all $z \in \mathbb{D}_Y = \{x/a : x \in \mathbb{D}_X\}$.

Proof. We can write

$$M_Y(z) \;=\; \mathbb{E}[e^{z(aX+b)}] \;=\; e^{bz}\mathbb{E}[e^{azX}] \;=\; e^{bz}M_X(az).$$

We have $M_Y(z) < \infty$ if $z = x/a$ and $x \in \mathbb{D}_X$. □

In the section on generating functions we emphasized the usefulness of solving a problem not in "random variable space" but in the function space based on the generating function. Here we have the same in mind, e.g., by application of Theorem 2.8.5 where we saw how easy it is to obtain the convolution of two independent random variables in the function space based on moment generating functions. However, the mechanism to translate a problem from one environment into another environment is only useful as the relationship between random variables and moment generating functions is *unique*. The next theorem makes this precise and confirms our approach.

The proof makes use of the following fact which is worth mentioning in the floating text outside of the proof too.

2.8.7 Lemma *Let X be a random variable in $(\Omega, \mathcal{F}, \mathbb{P})$ and assume $\mathbb{E}[e^{r|X|}] < \infty$ for some $r > 0$. Then, $(-r, r) \subset \mathbb{D}_X$.*

Proof. Based on the inequality

$$e^{a|x|} \;\leq\; e^{r|x|} \quad \text{for all } a \in (-r, r)$$

we can write

$$\mathbb{E}[e^{aX}] \;\leq\; \mathbb{E}[e^{r|X|}] \;<\; \infty \quad \text{for all } a \in (-r, r).$$

This yields $(-r, r) \subset \mathbb{D}_X$. □

2.8.8 Theorem *Let X be a random variable in $(\Omega, \mathcal{F}, \mathbb{P})$ and assume that there exists some $r > 0$ such that $\mathbb{E}[e^{r|X|}] < \infty$. Then, M_X uniquely determines the distribution $\mathbb{P}_X$ of X and the distribution $\mathbb{P}_X$ is uniquely determined by the moments $(\mathbb{E}[X^n])_{n\in\mathbb{N}}$ of X.*

Sketch of Proof. Collecting everything said so far on moment generating functions we find that under the conditions of the theorem M_X is an *analytic function* on $(-r, r)$ with power series expansion

$$M_X(z) \;=\; \sum_{n=0}^{\infty} \frac{z^n}{n!}\,\mathbb{E}[X^n]$$

(see (2.71)) whose coefficients are determined by the moments of X. Combining this with the fact that the Laplace transform (see (2.70)) is *injective* we can conclude the assertion of the theorem. □

In the proof of Theorem 2.8.8 we mentioned that the Laplace transform is injective. In fact, under certain conditions there are inversion formulas for integral transforms. For instance, if X takes only nonnegative values and is continuous (w.r.t. Lebesgue measure) with a density f_X we can write the (one-sided) Laplace transform (cp. also (2.69)) as

$$L_X(z) \;=\; \int_0^{+\infty} e^{-zx} f_X(x)dx. \tag{2.72}$$

Recall from Equation (2.70) that the moment generating function and the Laplace transform are essentially the same objects so that in our example we can focus on Laplace transforms. Under suitable conditions we can then recover f_X from L_X via the *inversion formula*

$$f_X(x) \;=\; \frac{1}{2\pi i}\int\limits_{\gamma-i\infty}^{\gamma+i\infty} e^{xt}L_X(t)dt \tag{2.73}$$

always assuming that every object involved exists and is suitably defined, etc. The constant γ in (2.73) has to be chosen appropriately and it is beyond the scope of our discussion to go into detail here. In convenient cases we can choose $\gamma = 0$.

We stop the discussion at this point and conclude this section by continuing Examples 2.8.2 and 2.8.3.

2.8.9 Example *Let $L \sim B(1;p)$ be a Bernoulli random variable. Then its moment generating function is given by*

$$M_L(z) \;=\; 1 + p(e^z - 1).$$

Proof. We know that $G_L(z) = 1 + p(z-1)$ from Example 2.8.2. We also know from Equation 2.68 that $M_X(z) = G_X(e^z)$ for discrete X. This proves the statements. □

2.8.10 Example *Let $L' \sim Pois(\lambda)$ be a Poisson random variable. Then its generating function is given by*

$$M_{L'}(z) \;=\; e^{\lambda(e^z-1)}\;.$$

Proof. We have $G_{L'}(z) = e^{\lambda(z-1)}$ and $M_{L'}(z) = G_{L'}(e^z)$. □

The cumulant generating function

Consider a random variable X in a probability space $(\Omega, \mathcal{F}, \mathbb{P})$. The *cumulant generating function* K_X of X is defined as the (natural) logarithm of the moment generating function M_X of X:

$$K_X(z) \;=\; \ln M_X(z) \quad \text{for all } z \in \mathbb{D}_X.$$

The cumulant generating function is convenient if in case of independent random variables one prefers a sum at function space side instead of a product because Proposition 2.8.5 immediately yields

$$K_{X+Y}(z) \;=\; \ln M_{X+Y}(z) \;=\; \ln(M_X(z)M_Y(z)) \;=\; K_X(z) + K_Y(z)$$

for independent variables X, Y and $z \in \mathbb{D}_X \cap \mathbb{D}_Y$. In the same way as for M_X we find that K_X is an analytic function under certain conditions (see the proof of Theorem 2.8.8) with a power series representation

$$K_X(z) \;=\; \sum_{n=1}^{\infty} \frac{z^n}{n!} K_X^{(n)}(0).$$

Here, the derivatives $K_X^{(n)}$ are called *cumulants*. The first cumulant is always the expectation of X. To see this, just write

$$\frac{d}{dz} \ln M_X(z) \;=\; \frac{\frac{d}{dz} M_X(z)}{M_X(z)} \;=\; \frac{\mathbb{E}[X e^{zX}]}{\mathbb{E}[e^{zX}]}$$

which equals $\mathbb{E}[X]$ for $z = 0$. Of course, this again needs certain conditions to be fulfilled. We stop here and continue with our next topic.

The Characteristic Function (Fourier Transform)

The most commonly used integral transform in probability theory is the *Fourier transform*. For a bounded positive measure μ on $\mathbb{R}$ it is defined as

$$\hat{\mu}(z) \;=\; \int_{\mathbb{R}} e^{izx} d\mu(x).$$

Now let X be a random variable in a probability space $(\Omega, \mathcal{F}, \mathbb{P})$. Then, the Fourier transform

$$\varphi_X(z) \;=\; \hat{P_X}(z) \;=\; \mathbb{E}[e^{izX}] \;=\; \int_{\mathbb{R}} e^{izx} d\mathbb{P}_X(x) \qquad (2.74)$$

of the distribution $\mathbb{P}_X$ of X is called the *characteristic function* of X. The notation φ_X for the characteristic function of a random variable X is very common and we adopt this notation for the sequel. Obviously, the characteristic function φ_X of a random variable X and its moment generating function M_X are related via the equation

$$\varphi_X(z) \;=\; M_X(iz).$$

In this way we can, for instance, carry over the results from Examples 2.8.9 and 2.8.10 to Examples 2.8.17 and 2.8.18.

Characteristic functions and Fourier transforms have many convenient properties which are widely known. Therefore, we make only a few remarks here and add more properties "on the fly" in the sequel whenever they are needed.

As mentioned at the beginning of this section, Fourier transforms are the role model for the kind of space switching we indicated in the sections on generating functions and moment generating functions. In many situations, *probabilists solve problems in Fourier space and later carry the result over to the original environment of random variables or stochastic processes.*

For doing so one again needs uniqueness of the characteristic function and inversion formulas. The convolution formula which we encountered in the previous two sections remains also valid. We briefly go through a few results in the sequel without going too much into the details. First let us mention that in case X in $\mathbb{R}$ has a density f_X the characteristic function equals

$$\varphi_X(z) \;=\; \int_{-\infty}^{+\infty} e^{izx} f_X(x)dx \tag{2.75}$$

and its Stieltjes integral form is given by

$$\varphi_X(z) \;=\; \int_{-\infty}^{+\infty} e^{izx} dF_X(x) \tag{2.76}$$

where F_X denotes the (cumulative) distribution function of X.

For the moment generating function M_X we talked about its domain $\mathbb{D}_X$. What about characteristic functions, do they always exist? The answer is simple: because of

$$e^{izx} \;=\; \cos(zx) + i\sin(zx) \tag{2.77}$$

and because $\sin(\cdot)$ and $\cos(\cdot)$ are bounded functions we find that the characteristic function always exists.

We start our collection of results with the convolution formula.

2.8.11 Theorem *Let X and Y be independent random variables in $(\Omega, \mathcal{F}, \mathbb{P})$. Then,*

$$\varphi_{X+Y}(z) \;=\; \varphi_X(z)\,\varphi_Y(z) \quad \textit{for all } z \in \mathbb{R}.$$

Proof. Again it is the functional property of the exponential function which enables the convolution formula. One has

$$\varphi_{X+Y}(z) \;=\; \mathbb{E}[e^{iz(X+Y)}] \;=\; \mathbb{E}[e^{izX}]\,\mathbb{E}[e^{izY}] \;=\; \varphi_X(z)\,\varphi_Y(z)$$

due to independence of X and Y. □

2.8.12 Theorem *Let X and Y be random variables in $(\Omega, \mathcal{F}, \mathbb{P})$. If for their characteristic functions φ_X and φ_Y we have*

$$\varphi_X(z) \;=\; \varphi_Y(z) \quad \textit{for all } z \in \mathbb{R}$$

then the distributions of X and Y coincide: $\mathbb{P}_X = \mathbb{P}_Y$.

Proof. Can be found in any book on probability. We skip it here because it needs advanced results beyond the scope of this book. □

Interestingly, the converse of Proposition 2.8.11 also holds. For the proof we need the just presented uniqueness theorem.

2.8.13 Theorem *Let X and Y be random variables in $(\Omega, \mathcal{F}, \mathbb{P})$ with*

$$\varphi_{X+Y}(z) \;=\; \varphi_X(z)\,\varphi_Y(z) \quad \textit{for all } z \in \mathbb{R}.$$

Then, X and Y are independent.

Proof. Choose *independent* random variables $\tilde{X}$ and $\tilde{Y}$ with

$$\mathbb{P}_X \;=\; \mathbb{P}_{\tilde{X}} \quad \text{and} \quad \mathbb{P}_Y \;=\; \mathbb{P}_{\tilde{Y}}.$$

Because $\tilde{X}$ and $\tilde{Y}$ are independent we obtain from Proposition 2.8.11

$$\varphi_{\tilde{X}+\tilde{Y}} \;=\; \varphi_{\tilde{X}}\,\varphi_{\tilde{Y}}. \tag{2.78}$$

The characteristic function of a random variable Z does (by definition) only depend on its distribution $\mathbb{P}_Z$. Therefore,

$$\varphi_X \;=\; \varphi_{\tilde{X}} \quad \text{and} \quad \varphi_Y \;=\; \varphi_{\tilde{Y}}.$$

Equation (2.78) then implies

$$\varphi_{\tilde{X}+\tilde{Y}} = \varphi_{\tilde{X}}\,\varphi_{\tilde{Y}} = \varphi_X\,\varphi_Y = \varphi_{X+Y}$$

where the last equality is the assumption of the theorem. Theorem 2.8.12 then yields that $\mathbb{P}_{X+Y} = \mathbb{P}_{\tilde{X}+\tilde{Y}}$. From this we can conclude that the (joint) distributions of (X, Y) and $(\tilde{X}, \tilde{Y})$ coincide. Because $\tilde{X}$ and $\tilde{Y}$ are independent the same must be true for X and Y. □

2.8.14 Proposition *Let X be a random variable in $(\Omega, \mathcal{F}, \mathbb{P})$ and set $Y = aX + b$ for constants $a, b \in \mathbb{R}$. Then,*

$$\varphi_Y(z) = e^{ibz}\varphi_X(az) \quad \textit{for all } z \in \mathbb{R}.$$

Proof. As in the proof of Proposition 2.8.6 we write

$$\varphi_Y(z) = \mathbb{E}[e^{iz(aX+b)}] = e^{ibz}\mathbb{E}[e^{iazX}] = e^{ibz}\varphi_X(az)$$

which proves the assertion. □

The characteristic function has nice properties. One is the following:

2.8.15 Theorem *Let X be a random variable in $(\Omega, \mathcal{F}, \mathbb{P})$ and denote the characteristic function of X by φ_X. Then, φ_X is uniformly continuous.*

Proof. Writing the characteristic function as

$$\varphi_X(z) = \mathbb{E}[\cos(zX)] + i\mathbb{E}[\sin(zX)]$$

by an application of the Euler identity (2.77) we immediately find by an application of the theorem of dominated convergence that the characteristic function φ_X is continuous. But more is true. Writing

$$\begin{aligned}
|\varphi_X(z+\delta) - \varphi_X(z)| &= \left|\mathbb{E}[e^{i(z+\delta)X}] - \mathbb{E}[e^{izX}]\right| \\
&= \left|\mathbb{E}[e^{i(z+\delta)X} - e^{izX}]\right| \\
&= \left|\mathbb{E}[e^{izX}(e^{i\delta X} - 1)]\right| \\
&\leq \mathbb{E}[|e^{i\delta X} - 1|]
\end{aligned}$$

and the right-hand side converges to zero (uniformly in z) for $\delta \to 0$ due to the theorem of dominated convergence. □

As a last result before we revisit our two main examples (Bernoulli and Poisson distributions) we state a Fourier inversion theorem.

2.8.16 Theorem *Let X be a random variable in $(\Omega, \mathcal{F}, \mathbb{P})$ with density f_X and denote its characteristic function by φ_X. If φ_X is integrable on $\mathbb{R}$ (w.r.t. Lebesgue measure) one can recover the density f_X from φ_X by means of the following inversion formula:*

$$f_X(x) \;=\; \frac{1}{2\pi}\int\limits_{\mathbb{R}} e^{-ixt}\varphi_X(t)dt \quad \text{for almost every } x \in \mathbb{R}.$$

Proof. See the results on *spectral synthesis* in Section III.2 in [129]. □

2.8.17 Example *Let $L \sim B(1;p)$ be a Bernoulli random variable. Then its characteristic function is given by*

$$\varphi_L(z) \;=\; 1 + p(e^{iz} - 1).$$

Proof. Consider Example 2.8.9 and $\varphi_X(z) = M_X(iz)$. □

2.8.18 Example *Let $L' \sim Pois(\lambda)$ be a Poisson random variable. Then its generating function is given by*

$$\varphi_{L'}(z) \;=\; e^{\lambda(e^{iz}-1)}\,.$$

Proof. Consider Example 2.8.10 and $\varphi_X(z) = M_X(iz)$. □

Remark: Extension to the Multivariate Case

In the previous four sections we briefly introduced four transformations for random variables. We always restricted the exposition to the one-dimensional case, say, X being a random variable in $\mathbb{R}$. However, the results presented so far can be generalized for random variables $\boldsymbol{Y}$ in $\mathbb{R}^m$ without difficulties. One just has to do the usual modifications, for instance, instead of

$$e^{zX} \quad \text{with} \quad z \in \mathbb{R},\ X \in \mathbb{R}$$

one has to write

$$\exp(\boldsymbol{z}^T\boldsymbol{Y})\text{with} \quad \boldsymbol{z} \in \mathbb{R}^m, \ \boldsymbol{Y} \in \mathbb{R}^m \ .$$

Recall from Section 1.2.3 that we denote vectors in $\mathbb{R}^m$ as column vectors such that $\boldsymbol{z}^T\boldsymbol{Y}$ is the usual *scalar product* in the Euclidean vector space $\mathbb{R}^m$.

We are now ready and well prepared for our discussion on techniques for generating loss distributions. We will keep the presentation in a style such that the topics mentioned provide an overview but do not cover details, proofs or fully-fledged calculations. Intention of the following discussion is that readers get a first impression on relevant techniques, e.g., for reducing the variance of outcomes in simulations, etc.

2.8.2 Conditional Independence

Mixture models as we introduced them in this chapter rely on a *conditional independence* framework. Going back to Sections 2.1.1 and 2.2.1 we find that involved Bernoulli and Poisson random variables are independent *conditional on a realization of the mixture distribution* for the default probabilities or default intensities. We denoted the mixture distribution by $\boldsymbol{F}$. Let us assume that a portfolio with m obligors is given. In the Bernoulli model the mixing worked via a random vector

$$\boldsymbol{P} \ = \ (P_1, ..., P_m) \ \sim \ \boldsymbol{F}$$

of default probabilities (jointly) drawn w.r.t. the mixture distribution $\boldsymbol{F}$. In the Poisson model the mixing was based on a random vector

$$\boldsymbol{\Lambda} \ = \ (\Lambda_1, ..., \Lambda_m) \ \sim \ \boldsymbol{F}$$

of default intensities with (joint) law $\boldsymbol{F}$. Obviously, $\boldsymbol{F}$ is a different object in the two different models.

In the sequel we follow a slightly different notation. In most cases we find that mixing is not directly done at the level of default probabilities or default intensities but *at the level of some latent variables* which are then transformed via some mapping into random default probabilities or default intensities. For instance, in the Bernoulli mixture one-factor model (see Equation (2.49)) we have a latent variable Y which is transformed into a random default probability via

$$Y \ \mapsto \ N\Big[\frac{N^{-1}[p] - \sqrt{\varrho}\,Y}{\sqrt{1-\varrho}}\Big].$$

So the randomness occurs at the level of the latent variable Y and not directly at the level of the random default probability.

Along these lines, we assume for the rest of this section that the mixing distribution $\boldsymbol{F}$ corresponds to a vector $\boldsymbol{Y}$ of latent variables which are transformed into random default probabilities or random default intensities by some suitable transformations

$$\boldsymbol{Y} \;\mapsto\; p_i(\boldsymbol{Y}) \in [0,1] \quad (i = 1, ..., m)$$

in case of the Bernoulli model and

$$\boldsymbol{Y} \;\mapsto\; \lambda_i(\boldsymbol{Y}) \in [0,\infty) \quad (i = 1, ..., m).$$

in case of the Poisson model. The Bernoulli mixture model can then be written down as

$$L_i \;\sim\; B\big(1; p_i(\boldsymbol{Y})\big) \quad (i = 1, ..., m) \quad \text{with} \quad \boldsymbol{Y} \;\sim\; \boldsymbol{F}$$

and the Poisson mixture model reads as

$$L'_i \;\sim\; Pois\big(\lambda_i(\boldsymbol{Y})\big) \quad (i = 1, ..., m) \quad \text{with} \quad \boldsymbol{Y} \;\sim\; \boldsymbol{F}$$

where we again keep in mind that $\boldsymbol{Y}$ and $\boldsymbol{F}$ in the Bernoulli framework differ from $\boldsymbol{Y}$ and $\boldsymbol{F}$ in the Poisson framework.

For the sake of an easier presentation we collect EAD and LGD in the sequel into one constant factor E_i for each obligor $(i = 1, ..., m)$. For instance, in a Bernoulli model the loss of obligor i is then given by $E_i L_i$ with $L_i \sim B\big(1; p_i(\boldsymbol{Y})\big)$. One can think of E_i as the loss amount in monetary units in case of default which corresponds to $L_i = 1$.

Now, when it comes to the generation of the loss distribution we can read the formulas in Sections 2.1.1 and 2.2.1 as follows.

- **Step 1**. Randomly draw realizations $\boldsymbol{y}$ of $\boldsymbol{Y} \sim \boldsymbol{F}$.

- **Step 2**. Generate the portfolio loss distribution

$$L|_{\boldsymbol{Y}=\boldsymbol{y}} \;=\; \sum_{i=1}^{m} E_i(L_i|_{\boldsymbol{Y}=\boldsymbol{y}}) \qquad (2.79)$$

  conditional on $\boldsymbol{Y} = \boldsymbol{y}$. Because L is a sum of independent variables we find that $\mathbb{P}_{L|_{\boldsymbol{Y}=\boldsymbol{y}}}$ is a convolution. For instance, in a uniform exposure Bernoulli model we get a binomial distribution.

- **Step 3**. Integrate the conditional portfolio loss distributions $L|_{\boldsymbol{Y}=\boldsymbol{y}}$ w.r.t. $\boldsymbol{F}$ over the space of all realizations $\boldsymbol{y}$ of $\boldsymbol{Y}$.

In the sequel we will not spend much time on Step 3. Integration over the latent factors is a numerical problem and we refer to the rich literature on **quadrature** techniques for this topic; see, e.g., [156].

So we will focus on Step 2, say, for a given realization of latent variables. This means that in the following **we will typically assume independence of single-name risks**. We ask readers to keep in mind that the discussion that follows is only one out of three steps in the generation procedure of the loss distribution.

2.8.3 Technique I: Recursive Generation

Let us start with a portfolio with m homogeneous exposures

$$E_i \;=\; E \quad \text{for all } i = 1, ..., m. \tag{2.80}$$

In line with the remark at the end of the previous section we assume independence of single-name risks in the portfolio which basically means that we are studying a *convolution* problem.

Let us look at the Bernoulli framework first. In case of

$$p_i \;=\; p \quad \text{for all } i = 1, ..., m$$

we already pointed out in Section 2.1 that in this case the convolution is a binomial distribution $B(n; p)$ which is easy to handle.

For the case of heterogeneous exposures we are at least able to calculate the portfolio EL and portfolio UL as we did in Equation (2.1). The distribution itself is a bit more tricky and we postpone the discussion of a very useful recursive algorithm for a moment.

For the Poisson framework we have no problem calculating the convolution even in case of heterogeneous intensities λ_i because of the convolution property of independent Poisson random variables:

$$L'_i \;\sim\; Pois(\lambda_i),\; L'_j \;\sim\; Pois(\lambda_j)$$

$$L'_i, L'_j \text{ independent } \Longrightarrow\; L'_i + L'_j \;\sim\; Pois(\lambda_i + \lambda_j).$$

This is the justification for the so-called *Poisson convolution semigroup* which is the basis for the definition of Poisson processes. We used this convolution property already in Section 2.2 but can here now quickly prove it based on our results about generating functions. From Example 2.8.3 we know that

$$G_{L'_i}(z) \ = \ e^{\lambda_i(z-1)} \quad \text{and} \quad G_{L'_j}(z) \ = \ e^{\lambda_j(z-1)} \, ..$$

From Theorem 2.8.1 we know that

$$G_{L'_i+L'_j}(z) \ = \ G_{L'_i}(z)G_{L'_j}(z) \ = \ e^{(\lambda_i+\lambda_j)(z-1)}$$

which is the generating function of a variable $L \sim Pois(\lambda_i + \lambda_j)$. Because the correspondence between distributions and generating functions is one-on-one we found a simple proof for the convolution formula.

Altogether we get for a portfolio with homogeneous exposures but heterogeneous default intensities a Poisson convolution as follows:

$$L' \ = \ \sum_{i=1}^{m} E_i L'_i \ = \ E \times \tilde{L} \quad \text{with} \quad \tilde{L}' \ \sim \ Pois\big(\sum_{i=1}^{m} \lambda_i\big).$$

The CreditRisk$^+$ model (see Section 4) makes use of this convolution property and combines it with an exposure grouping algorithm.

Now let us come back to the Bernoulli model ($L_i \sim B(1;p_i)$) with heterogeneous default probabilities which we postponed until now. In this situation the portfolio loss distribution can be obtained by use of the following *recursive algorithm* where the uniform exposure E has been factored out such that we can without loss of generality assume that all single-name risks have unit exposure. Again, it is crucial that we assumed independence of assets in line with our remark at the end of the previous section.

- **Step 1**. We start with $m = 1$ which means that we have just one obligor or asset in the portfolio. We denote the portfolio loss distribution by $\mathbb{P}^{(1)}$. Then,

  $$\mathbb{P}^{(1)}[L = 0] = (1 - p_1)\,, \quad \mathbb{P}^{(1)}[L = 1] = p_1.$$

  This constitutes the start of the recursion.

- **Step 2.** We extend the portfolio to $m = 2$ by including another asset with default probability p_2. The portfolio loss distribution is now $\mathbb{P}^{(2)}$. We have the following cases.

$$\begin{aligned}
\mathbb{P}^{(2)}[L=0] &= (1-p_1)(1-p_2) \\
&= \mathbb{P}^{(1)}[L=0](1-p_2) \\
\mathbb{P}^{(2)}[L=1] &= p_1(1-p_2)+(1-p_1)p_2 \\
&= \mathbb{P}^{(1)}[L=1](1-p_2)+\mathbb{P}[L=0]\,p_2 \\
\mathbb{P}^{(2)}[L=2] &= p_1p_2 \\
&= \mathbb{P}^{(1)}[L=1]\,p_2\,.
\end{aligned}$$

In this form we can generalize it to arbitrary m as follows.

- **Step m**. For $m>2$ we have the following recursion formula:

$$\mathbb{P}^{(m)}[L=k] \;=\; \mathbb{P}^{(m-1)}[L=k](1-p_m)+\mathbb{P}^{(m-1)}[L=k-1]\,p_m$$

with $k\in\{0,1,2,...,m\}$. A proof is straightforward (induction).

Based on this recursion formula one can construct the portfolio loss distribution starting with one asset and then adding assets successively up to the target number m of assets.

If we want to drop the assumption (2.80) of homogeneous exposures the recursion formula can still be applied in *approximative* terms. One can introduce a *loss unit* u and postulate that the loss risk of obligor (or asset) i refers to n_i units of loss where n_i is called a *loss unit multiple.* In other words, we replace the original E_i's by

$$\tilde{E}_i \;=\; n_i\times u \quad (i=1,...,m;\; n_i\in\mathbb{N}_0).$$

The challenge in this approach is to determine u in a way such that the rounding error from replacing the true exposure vector $(E_1,...,E_m)$ by $(\tilde{E}_1,...,\tilde{E}_m)$ is in an acceptable order of magnitude. Factoring out the loss unit u we get the same recursion formula as before,

$$\mathbb{P}^{(m)}[L=k] \;=\; \mathbb{P}^{(m-1)}[L=k](1-p_m)+\mathbb{P}^{(m-1)}[L=k-n_m]\,p_m\,,$$

but now with states $k\in\{0,1,2,...,n_1+\cdots+n_m\}$ and boundary conditions $\mathbb{P}^{(0)}[L=0]=1$, $\mathbb{P}^{(0)}[L\neq 0]=0$.

Note that our writing is a bit sloppy here because L can only move along a grid generated by partial sums of loss unit multiples. However, because the probability that L falls in-between such effective grid points is zero, there is no harm in writing the recursion a bit sloppy but in turn less complex.

The calculation effort of the recursion formula is roughly proportional to m^2 and also depends on the chosen granularity of the exposure discretization (i.e., the choice of the loss unit u). Therefore, it does not make sense to apply the recursion formula to large portfolios. A typical field of application of the recursive algorithm are collateralized debt obligations (CDOs) with a moderate number of underying credit risks, say, not more than 200 assets. Another advantage of the recursive approach is that it can also be directly formulated for some functions of the portfolio loss, e.g., a *tranche attachment point exceedance function* like $\max(L-K,0)$ where K is the attachment point (lower threshold) of a CDO tranche; see, e.g., [29].

We now turn our attention to the Poisson model again. Unfortunately, the convolution of Poisson distributed loss variables with inhomogeneous exposures can not be treated by a simple summation of Poisson intensities. However, with help of the aforementioned exposure coarsening one can find a useful recursion formula again.

After exposure coarsening and factoring out the loss unit u the Poisson loss variable of obligor i equals $n_i L_i$. Applying Equation (2.68), Proposition 2.8.6 and Example 2.8.3 we obtain

$$G_{n_i L'_i}(z) \;=\; M_{n_i L'_i}(\ln z) \;=\; M_{L'_i}(n_i \ln z) \;=\; G_{L'_i}(z^{n_i}) \;=\; e^{\lambda_i (z^{n_i}-1)}.$$

Application of the convolution property (Theorem 2.8.1) yields

$$G_{L'}(z) \;=\; \exp\Big(\sum_{i=1}^{m} \lambda_i (z^{n_i}-1)\Big).$$

for the generating function of the portfolio loss. We can now apply Formula (2.67) which translates into our context as

$$\mathbb{P}[L' = k] \;=\; \frac{1}{k!}\, G_{L'}^{(k)}(0). \tag{2.81}$$

Note again that the loss unit multiples define a grid of possible loss states but we can (without doing wrong) simply write k in (2.81). In the following we use the notation $\frac{d^k}{dz^k}$ for the order-n differentiation operator. One obtains

$$\begin{aligned}\mathbb{P}(L' = k) &= \frac{1}{k!}\frac{d^k G_{L'}(z)}{dz^k}\bigg|_{z=0} \\ &= \frac{1}{k!}\frac{d^{k-1}}{dz^{k-1}}\left(G_{L'}(z)\frac{d}{dz}\Big(\sum_{i=1}^m \lambda_i z^{n_i}\Big)\right)\bigg|_{z=0} \\ &= \frac{1}{k!}\sum_{j=0}^{k-1}\binom{k-1}{j}\frac{d^{k-j-1}}{dz^{k-j-1}}G_{L'}(z)\frac{d^{j+1}}{dz^{j+1}}\left(\sum_{i=1}^m \lambda_i z^{n_i}\right)\bigg|_{z=0}.\end{aligned}$$

But one also has

$$\frac{d^{j+1}}{dz^{j+1}}\sum_{i=1}^m \lambda_i z^{n_i} = \begin{cases}\lambda_i(j+1) & \text{if } j+1 = n_i \text{ for some } i \\ 0 & \text{otherwise}\end{cases},$$

such that altogether

$$\begin{aligned}\mathbb{P}(L' = k) &= \sum_{\substack{j\le k-1 \\ \exists\, i:\, j=n_i-1}}^{k-1} \frac{(k-j-1)!}{k!}\binom{k-1}{j}\,\mathbb{P}(L' = k-j-1)\lambda_i(j+1)! \\ &= \sum_{i:n_i\le k}\frac{n_i\lambda_i}{k}\,\mathbb{P}[L' = k-n_i]. \qquad (2.82)\end{aligned}$$

Equation (2.82) now allows to recursively construct the portfolio loss distribution starting from the zero loss probability

$$\mathbb{P}[L = 0] = \exp\Big(-\sum_{i=1}^m \lambda_i\Big).$$

The recursion formula we just developed and discussed is in fact a version of the well known *Panjer-recursion* and is also the basic idea for constructing the portfolio loss distribution in the CreditRisk$^+$ model (see Chapter 4).

2.8.4 Technique II: Fourier Transformation

In our section on prerequisites from probability theory we discussed the characteristic function of a random variable X as the Fourier transform of its distribution $\mathbb{P}_X$. The mapping

$$\mathrm{F}:\ \mathbb{P}_X \mapsto \hat{\mathbb{P}_X}$$

of a distribution (or, more general, a bounded positive measure) to its Fourier transform can be introduced as an operator. We will not go into detail here regarding its precise definition, domain, range and properties but will use the Fourier transform in the sequel as an operator (denoted by F) which can be inverted in line with the Fourier inversion theorem 2.8.16.

We start by going back to the unconditional loss

$$L \;=\; \sum_{i=1}^{m} E_i L_i \; .$$

Using the convolution theorem for Fourier transforms (Theorem 2.8.11) we can write

$$\begin{aligned}
\mathbb{P}[L=x] &= \int_D \mathrm{F}^{-1}\left\{\mathrm{F}\left(\bigotimes_{i=1}^{m}\mathbb{P}\left[E_iL_i|\sum_i E_iL_i = x, \boldsymbol{Y}=y\right]\right)\right\} d\boldsymbol{F}(y)\\
&= \int_D \mathrm{F}^{-1}\left\{\prod_{i=1}^{m}\mathrm{F}\left(\mathbb{P}\left[E_iL_i|\sum_i E_iL_i = x, \boldsymbol{Y}=y\right]\right)\right\} d\boldsymbol{F}(y)\\
&= \mathrm{F}^{-1}\left\{\int_D \prod_{i=1}^{m}\mathrm{F}\left(\mathbb{P}\left[E_iL_i|\sum_i E_iL_i = x, \boldsymbol{Y}=y\right]\right) d\boldsymbol{F}(y)\right\}.
\end{aligned}$$

The last step is due to the linearity of the Fourier transformation with the consequence that the calculation of the inverse Fourier transformation has to be performed only once.

The Fourier transfoms and its inverse are usually computed numerically using the *Fast Fourier Transformation* algorithm [156]. As this algorithm works on a lattice, we are left with the problem of discretizing the distribution function where concessions have to be made when balancing accuracy and performance.

In case of conditionally independent Bernoulli loss variables $E_iL_i|_{\boldsymbol{Y}=\boldsymbol{y}}$ the Fourier transform of the conditional portfolio loss

$$L \;=\; \sum_{i=1}^{m} E_i L_i|_{\boldsymbol{Y}=\boldsymbol{y}}$$

(Example 2.8.17, Theorem 2.8.11 and Proposition 2.8.14) is given by

$$\varphi_L(z)|_{\boldsymbol{Y}=\boldsymbol{y}} \;=\; \prod_{i=1}^{m}\Big(1+p_i(\boldsymbol{y})(e^{iE_iz}-1)\Big).$$

In case of Poisson losses $E_i L'_i|_{\boldsymbol{Y}=\boldsymbol{y}}$ we get (Example 2.8.18)

$$\varphi_{L'}(z)|_{\boldsymbol{Y}=\boldsymbol{y}} \;=\; \exp\Big(\sum_{i=1}^{m} \lambda_i(\boldsymbol{y})(e^{iE_i z} - 1)\Big).$$

This concludes our discussion of Fourier techniques.

2.8.5 Technique III: Saddle-Point Approximation

Although introduced not too long ago to the field of credit risk modeling by Martin et al. [132] the saddle point method has quite a long history through its applications in statistics by Daniels [39, 40] and its use for asymptotic expansions of integrals in physics. For a more fundamental reading we refer to Jensen [106] or Dembo and Zeitouni [42].

The basic idea of the saddle-point approximation is an expansion of the cumulant generating function. For the sequel it is useful to recall what has been said in Section 2.8.1 about cumulants, moment generating functions and Laplace transforms.

The cumulant generating function of the portfolio loss $L = \sum_i E_i L_i$ is given by

$$\begin{aligned} K_L(z) &= \ln M_L(z) = \ln \mathbb{E}[e^{zL}] \\ &= \ln \prod_{i=1}^{m} M_{E_i L_i}(z) = \sum_{i=1}^{m} \ln M_{E_i L_i}(z) \end{aligned} \tag{2.83}$$

In case of Bernoulli loss variables $K_L(z)$ has the simple form

$$K_L(z) \;=\; \sum_{i=1}^{m} \ln(1 - p_i + p_i e^{E_i z}) \tag{2.84}$$

(see Example 2.8.9). The probability density function can then be recovered via inversion of the Laplace transformation

$$\mathbb{P}[L = x] = \frac{1}{2\pi i} \int_{\gamma - i\infty}^{\gamma + i\infty} e^{K_L(s) - sx} ds, \tag{2.85}$$

(see also Formula (2.73) and take into account that the cumulant generating function is a logarithm so that the integrand in (2.73) needs an exponential (outer) function). Expanding $K_L(z)$ up to second order around $z = 0$ we get the Gaussian approximation of the loss distribution which matches the first two moments. However, $z = 0$ is not the optimal point to expand the cumulant generating function. Analyzing the leading contribution to the inverse Laplace transform integral one finds that a better choice for the origin of the expansion is the real stationary point of the exponent, i.e., the root z_0 with

$$K'_L(z_0) - x \;=\; 0. \tag{2.86}$$

As K is convex and analytic the stationary point z_0 is in fact a local minimum of $K_L(z) - zx$ on the real axis and a local maximum on the imaginary axis, hence a saddle point in the complex plane. Note that we are still in the conditional independence framework, i.e., the root of Equation (2.86) is a function of the latent factor, $z_0(\boldsymbol{Y})$. Expanding now the integrand of Equation (2.85) around z_0 followed by integration while keeping only the leading term and the first order corrections eventually leads to the approximation

$$\mathbb{P}[L = x] \approx \frac{e^{K_L(z_0) - z_0 x}}{\sqrt{2\pi K''_L(z_0)}} \left[1 + \left(\frac{K^{(4)}(z_0)}{8K''(z_0)^2} - \frac{5K^{(3)}(z_0)^2}{24K''(z_0)^3}\right)\right]. \tag{2.87}$$

For the integration it is again necessary that K is analytic as the integration contour has to be deformed in the complex plane. As it usually is the case with expansions of higher order, corrections will not necessarily yield to improvements in accuracy. In the case of Bernoulli loss variables the cumulants can conveniently be computed: the first two read

$$K'_L(z) = \sum_{i=1}^{m} \frac{E_i p_i e^{zE_i}}{1 - p_i + p_i e^{zE_i}} \quad \text{and}$$
$$K''_L(z) = \sum_{i=1}^{m} \frac{E_i^2 p_i (1 - p_i) e^{zE_i}}{(1 - p_i + p_i e^{zE_i})^2}. \tag{2.88}$$

The great advantage of the saddle-point approximation is the fact that for some functions of interest of the loss distribution like $\mathbb{P}[L \geq l]$, $\mathbb{E}[L|L \geq l]$ or $\mathbb{E}[(L - l)^+]$ we do not need to approximate the density function first and then integrate appropriately, but we can rather apply the saddle-point method directly to the required objects, [8, 192]. Note that the saddle-point approximation remains a continuous approximation, hence it will fail to capture any discrete features of the true loss distribution.

For a portfolio of independent Poisson-distributed loss variables with integer-valued exposures n_i the cumulant generating function is

$$K_{L'}(z) \;=\; \ln\left(G_{L'}(e^z)\right) = \sum_i^m \lambda_i (e^{n_i z} - 1)$$

with derivatives

$$K_L^{(j)}(z) \;=\; \sum_i^m \lambda_i n_i^j e^{n_i z}.$$

But in case of gamma-distributed random intensities as latent variables one can do even better since here the unconditional probability generating function is given in closed form (see Chapter 4). So, one can apply the saddle-point method directly to the unconditional distribution[22], [82].

Finally, we would like to touch briefly on another derivation of the saddle-point approximation since it is related to the technical trick presented in the next section. For estimating the distribution of independent random variables the best known approximation is the *Central Limit Theorem* and beyond that the *Edgeworth expansion*. But both approximations perform poorly in the tail of the distribution and are only a good fit at the center of the distribution. The idea now is to change the measure by "tilting" the density to the region of interest. Denote by $f_L(x)$ the density of the portfolio loss variable L, then we change it by multiplication with e^{sx}

$$\tilde{f}_L(x) \;=\; \frac{f_L(x) e^{zx}}{M_L(z)},$$

where the division by $M_L(z)$ just serves the purpose of renormalizing the new density $\tilde{f}$. It can be easily verified that under the new measure $\widetilde{\mathbb{P}}$ the loss variable L has mean $\widetilde{\mathbb{E}}[L] = K_L'(z)$ and $\widetilde{\mathbb{V}}[L] = K_L''(z)$. If we now choose a z_0 (z_0 is unique since K is convex) such that the tilted mean is at the loss level of interest x, i.e.,

$$K_L'(z_0) \;=\; x,$$

we can argue that L (due to the Central Limit Theorem) is roughly normal under the tilted measure $\widetilde{\mathbb{P}}$ and hence its probability density

[22] MARTIN and ORDOVÁS [131] call this the direct approximation, as opposed to the indirect method when conditioned.

can be approximated at its mean by $\tilde{f}_L(x) \approx 1/\sqrt{2\pi K_L''(z_0)}$. Tilting back, we obtain

$$f_L(x) \;=\; e^{K_L(z_0)-z_0 x}\tilde{f}_L(x) \approx \frac{e^{K_L(z_0)-z_0 x}}{\sqrt{2\pi K_L''(z_0)}},$$

which is the leading term in Equation (2.87). We stop here and refer to the rich literature if readers are interested in details and applications.

2.8.6 Technique IV: Importance Sampling

The Monte Carlo approach (see also the corresponding section in Chapter 1) in financial engineering is often considered as the method of last resort for problems that are not tractable otherwise. This approach essentially boils down to simulating the risk factors (random variables) of the problem and reading off the required results from the output statistics. Advances in computing power, but also in Monte Carlo methods (see e.g. [79, 101]) have made it by now a widely accepted approach to pricing and risk modeling. But it is by no means as simple as it seems, sophisticated techniques have been developed over the last years to enhance convergence of the simulation and improve the reliability of the sampling results. We will not elaborate on general techniques of the Monte Carlo method here but rather turn to an aspect specific to credit risk. As single defaults, and even more, multiple defaults are typically *low probability events* we are faced with the problem of a rare-event simulation. This is even more severe if we are interested in the tail statistic where we usually obtain only a couple of events of interest. This challenge makes the problem a very good candidate for *importance sampling* which is a technique that tries to *sample particularly in the region of interest* by means of a change of measure, hereby reducing the variance of the estimator. Before we go into details we briefly explain the underlying idea by means of the one-factor model of Section 2.5.1. Here, the conditional default probability as a function of the latent factor realization $Y = y$ is given by the function

$$y \;\mapsto\; N\left[\frac{N^{-1}[p_i] - \sqrt{\varrho}\, y}{\sqrt{1-\varrho}}\right]..$$

This function decreases with increasing y and increases with decreasing y; see Figure 2.3. This means that if we would sample the factor Y not from $N(0,1)$ but rather from a *shifted* normal distribution $N(\mu, 1)$

with some suitably chosen $\mu < 0$ then we would generate more adverse scenarios y of Y and hence more portfolio events with large losses.

Obviously, such an artificial shift will result in a bias of the final loss statistics. To correct such bias we have to attach to each scenario a new *likelihood ratio* and use this as a weight when calculating the loss statistics. The likelihood ratio[23] is simply the ratio of probabilities of the sampled scenario y before and after the shifting, i.e. $\Phi(y)/\Phi(y-\mu)$.

In the sequel we follow GLASSERMAN and LI [81] in a more systematic exposition of importance sampling for Gaussian mixed Bernoulli loss variables. Denote by $F_L(x)$ our quantity of interest in the tail loss, say, $F_L(x) = \mathbf{1}_{\{L>x\}}$ or $F_L(x) = L\mathbf{1}_{\{L>x\}}$, and $\hat{F}_L(x) = \mathbb{E}[F_{\hat{L}}(x)]$ the estimator of F on the loss statistics. Then we can decompose the variance of the estimator according to

$$\mathbb{V}[\hat{F}_L(x)] \;=\; \mathbb{E}[\mathbb{V}[\hat{F}_L(x)|\boldsymbol{Y}]] + \mathbb{V}[\mathbb{E}[\hat{F}_L(x)|\boldsymbol{Y}]]. \tag{2.89}$$

This fits nicely into the conditional independence framework because we can reduce the variance of the estimator in two steps.

Let us begin with the first term. Conditional on the common factors $\boldsymbol{Y}$ the Bernoulli loss variables E_iL_i are independent, hence by exponentially twisting the conditional default probabilities according to

$$p_{i,s}(\boldsymbol{Y}) = \frac{p_i(\boldsymbol{Y})e^{sE_i}}{1 + p_i(\boldsymbol{Y})(e^{sE_i} - 1)} \tag{2.90}$$

we arrive at a likelihood ratio of

$$\prod_{i=1}^{m} \left(\frac{p_i(\boldsymbol{Y})}{p_{i,s}(\boldsymbol{Y})}\right)^{L_i} \left(\frac{1-p_i(\boldsymbol{Y})}{1-p_{i,s}(\boldsymbol{Y})}\right)^{1-L_i} = \exp(-sL + K_L(s)), \tag{2.91}$$

where $K_L(\theta)$ is the cumulant generating function of L, see Equation (2.84), conditional on $\boldsymbol{Y}$. From Equation (2.90) we see that for $s > 0$ the exponentially twisted default probabilities are increased and Equation (2.91) shows that this is equivalent to exponentially twisting the portfolio loss variable L itself. The unbiased estimator for $F_L(x)$ under the twisted measure is now

$$F_L(x)e^{-sL+K_L(s)}.$$

[23] This more heuristic approach has been put forward by XIAO [191] and extended to several common factors by KALKBRENER et al. [110].

It remains to choose a suitable value for s such that the variance, or equivalently the second moment, of the estimator is minimized. If F_L is the indicator function, $F_L(x) = \mathbf{1}_{\{L>x\}}$, we can estimate the second moment with an upper bound as

$$\mathbb{E}_s[\mathbf{1}_{\{L>x\}} e^{-2sL+2K_L(s)}] \leq e^{-2sL+2K_L(s)} \tag{2.92}$$

and then minimize the upper bound, i.e., minimize $K_L(s) - sx$ for $s \geq 0$. We know that K is convex and passes through the origin so that a minimum is obtained for

$$K_L'(s_0) \;=\; x.$$

It can easily be verified that $\mathbb{E}_s[L] = K_L'(s)$, i.e., we again shifted the distribution of L to have now a mean of x. Note that L is still conditioned on the factors $\boldsymbol{Y}$ meaning that the optimal shift value is in fact a function of the common factor, namely, $s_0(\boldsymbol{Y})$. Furthermore, note that $x < \mathbb{E}[L]$ entails $s_0 < 0$, hence the exponential twist actually reduces the default probabilities.

Now, we can further reduce the variance in Equation (2.89) through the second term. To this end we consider a shifting of the mean of $\boldsymbol{Y} \sim N(0, \Sigma)$ from the origin to some point $\boldsymbol{\mu}$, say, $\boldsymbol{Y} \sim N(\boldsymbol{\mu}, \Sigma)$. The likelihood ratio of such a change of measure is then

$$\exp\left(-\boldsymbol{\mu}^T\boldsymbol{Y} + \frac{\boldsymbol{\mu}^T\boldsymbol{\mu}}{2}\right).$$

The two-step importance sampling estimator for F_L is

$$F_L \exp\left(-s_0(\boldsymbol{Y})L + K_{L|Y}(s_0(\boldsymbol{Y})) - \boldsymbol{\mu}^T\boldsymbol{Y} + \frac{\boldsymbol{\mu}^T\boldsymbol{\mu}}{2}\right) \tag{2.93}$$

and the importance sampling algorithm proceeds as follows:

1. Sample $\boldsymbol{Y} \sim N(\boldsymbol{\mu}, \Sigma)$.
2. Compute $s_0(\boldsymbol{Y})$ and the twisted conditional default probabilities $p_{i,s_0}(\boldsymbol{Y})$.
3. Generate losses under the twisted conditional distribution.
4. Return the estimator of Equation (2.93).

It remains to find the right $\boldsymbol{\mu}$ for the common factors $\boldsymbol{Y}$. In case of the quantile function $F_L = \mathbf{1}_{\{L>x\}}$, GLASSERMAN and LI [81] suggest choosing the shift by solving the optimization problem

$$\boldsymbol{\mu} = \max_y \mathbb{P}[L > x | \boldsymbol{Y} = y] e^{-y^T y/2},$$

which then has to be approximated further for tractability. KALKBRENER ET AL. [110] propose to use an infinitely fine-grained portfolio under the one-factor approach (see section 2.5.1), $L^\infty \sim F(p, \varrho)$, with suitably chosen parameters as a proxy to find the optimal value of the shift. In this case, the loss distribution is in analytic form so that we can search the minimum of the second moment of the estimator directly:

$$\mu \;=\; \min_x \int_{-\infty}^{\infty} \left(F_{L^\infty_{p,\varrho}(Y)} e^{-Yx+x^2} \right)^2 dN_{x,0}(Y).$$

The found scalar value for μ is then appropriately lifted up to the higher dimensional shift $\boldsymbol{\mu}$ according to the variance contributions of the factors in the factor space.

Similarly to the Bernoulli case we can also set up an importance sampling algorithm for the Poisson mixture model. Let us again first condition on the mixing variables. Then, we exponentially twist the conditional intensities of the independent Poisson variables as[24]

$$\lambda_i(s) \;=\; \lambda_i e^{w_i s} \tag{2.94}$$

which leads to the likelihood ratio

$$\prod_{i=1}^{m} \frac{e^{-\lambda_i} \lambda_i^{l'_i} / l'_i!}{e^{-\lambda_i(s)} \lambda_i(s)^{l'_i} / l'_i!} = \exp\left(-sL + K_L(s)\right). \tag{2.95}$$

In a second step we could again move to the latent factors and twist their distributions. GLASSERMAN and LI [80] investigated this model

[24] The choice of the exponential twisting of p_i and λ_i is motivated by the fact that Bernoulli and Poisson random variables form an exponential family of distributions. If $K(s)$ is the cumulant generating function of a distribution F then $F_s(x) = \int_{-\infty}^{x} e^{su-K(s)} dF(u)$ are again distributions and form an exponential family in s for all s such that $K(s) < \infty$. If F has a density f then F_s has the density $f_s(x) = e^{sx-K(s)} f(x)$.

type for gamma-distributed mixing variables as in CreditRisk$^+$. Suppose for simplicity that the intensities of all Poisson variables L_i' are driven by only one factor that is gamma-distributed, say, $\lambda_i = c_i + \lambda$, $\lambda \sim \Gamma(\alpha, \beta)$. The cumulant generating function of a gamma-distributed random variable is given by

$$K_\lambda(t) = -\alpha \log(1 - \beta t)$$

Hence the likelihood ratio of that measure change is

$$\exp\left(-t\lambda - \alpha \log(1 - \beta t)\right).$$

The twisted λ_t is again gamma-distributed with $\lambda_t \sim \Gamma(\alpha, \beta/(1 - \beta t))$, with $t < 1/\beta$. The likelihood ratio of the two steps is the product of the individual likelihood ratios:

$$\exp\left(-sL + \sum_{i=1}^{m} \lambda_i (e^{w_i s} - 1) - t\lambda - \alpha \log(1 - \beta t)\right).$$

In principle we have to find the optimal values for the two parameters s and t but by choosing $t = \sum_{i=1}^{m}(e^{w_i s} - 1)$ the combined likelihood ratio can be cast into the form

$$\exp(-sL + K_L(s))$$

where $K_L(s)$ now is the cumulant generating function of the unconditional portfolio loss, i.e. the unconditional loss gets twisted. As in the conditional Bernoulli case we find the optimal shift by solving

$$K_L'(s_0) \;=\; x$$

and then use the appropriate estimator.

We close this whole section with a last remark. We only scratched the surface of the techniques which one can use to generate loss distributions. It can not be more than a first taste. We also had quite some effort in the section on "prerequisites from probability theory" to provide some basics on tools which are necessary to follow the presentation. Altogether one can say that the content of the last four sections really is advanced material which needs a lot of math and a lot of patience when going through the literature. A good and established way to dive deeper into the techniques just explained is to start implementing algorithms in a suitable programming environment. A textbook can never replace self-made experience and we therefore encourage readers to do their own experiments.

Further Reading

A mathematical introduction to mixture models an be found in the book by Joe [107]. Mixture models applied to problems in finance and risk management are intensively discussed in the book [136] by McNeil, Frey and Embrechts. As references for copulas we refer to Nelsen [144], Cherubini, Luciano and Vecciato [31], Rank [158] and Lipton and Rennie [126]. As references for the saddle point method we refer to Dembo and Zeitouni [42], Jensen [106] and Martin [130]. As a source for variance reduction techniques and importance sampling we refer to Glasserman [79] and Jäckel [101]. Another reference is Egloff et al. [55].

Chapter 3

Asset Value Models

The *asset value model* (AVM) is an important contribution to modern finance. In the literature one can find a tremendous amount of books and papers treating the classical AVM or one of its various modifications. See, e.g., Crouhy, Galai and Mark [38] (Chapter 9), Sobehart and Keenan [170], and Bohn [27], just to mention a very small selection of especially nicely written contributions.

As already discussed in Section 1.2.3 and also in Chapter 2, two of the most widely used credit risk models are based on the AVM, namely the model by Moody's KMV as well as the model CreditMetrics™ from the RiskMetrics Group; see also Section 1.2.3.

The roots of the AVM are the seminal papers by Merton [137] and Black and Scholes [19], where the *contingent claims* approach to risky debt valuation by *option pricing* theory is elaborated.

3.1 Introduction and a Brief Guide to the Literature

The AVM in its original form goes back to Merton [137] and Black and Scholes [19]. Their approach is based on option pricing theory, and we will frequently use this theory in the sequel. For readers not familiar with options we will try to keep our course as self-contained as possible, but refer to the book by Hull [94] for a practitioner's approach and to the book by Baxter and Rennie [15] for a highly readable introduction to the mathematical theory of financial derivatives. Another excellent book more focusing on the underlying stochastic calculus is the one by Lamberton and Lapeyre [118]. For readers without any knowledge of stochastic calculus we recommend the book by Mikosch [139], which gives an introduction to the basic concepts of stochastic calculus with finance in view. To readers with a strong background in probability we recommend the books by KARATZAS and SHREVE [112, 113]. Besides these, the literature on derivative pricing is so voluminous that one can be sure that there is the optimal book for any reader's taste. All results presented later on can be found in the literature listed above. We therefore will – for the sake of a more fluent presentation – avoid the quotation of particular references but instead implicitly assume that the reader already made her or his particular choice of reference including proofs and further readings.

3.2 A Few Words about Calls and Puts

Before our discussion of Merton's model we want to briefly prepare the reader by explaining some basics on options. The basic assumption underlying option pricing theory is the nonexistence of *arbitrage*, where the word "arbitrage" essentially addresses the opportunity to make a *risk-free profit*. In other words, the common saying that *"there is no free lunch"* is the fundamental principle underlying the theory of financial derivatives.

In the following we will always and without prior notice assume that we are living in a so-called standard[1] Black-Scholes world. In such a world several conditions are assumed to be fulfilled, for example

- stock prices follow *geometric Brownian motions* with constant *drift* μ and constant *volatility* σ;
- *short selling* (i.e., selling a security without owning it) with full use of proceeds is permitted;
- when buying and selling, *no transaction costs* or *taxes* have to be deducted from proceeds;
- there are *no dividend payments*[2] during the lifetime of a financial instrument;

[1] In mathematical finance, various generalizations and improvements of the classical Black-Scholes theory have been investigated.

[2] This assumption will be kept during the introductory part of this chapter but dropped later on.

- the *no-arbitrage principle* holds;
- security trading is *continuous*;
- some *riskless* instrument, a so-called *risk-free bond*, can be bought and sold in arbitrary amounts at the riskless rate r, such that, e.g., investing x_0 units of money in a bond today (at time $t = 0$) yields $x_t = x_0 e^{rt}$ units of money at time t;
- the *risk-free interest rate* $r > 0$ is constant and independent of the maturity of a financial instrument.

As an illustration of how the *no-arbitrage principle* can be used to derive statements about asset values we want to prove the following proposition.

3.2.1 Proposition *Let $(A_t)_{t\geq 0}$ and $(B_t)_{t\geq 0}$ denote the value of two different assets with $A_T = B_T$ at time $T > 0$. Then, if the no-arbitrage principle holds, the values of the assets today (at time 0) also agree, such that $A_0 = B_0$.*

Proof. Assume without loss of generality $A_0 > B_0$. We will show that this assumption contradicts the no-arbitrage principle. As a consequence we must have $A_0 = B_0$. We will derive the contradiction by a simple *investment strategy*, consisting of three steps:

1. short selling of A today, giving us A_0 units of money today;
2. buying asset B today, hereby spending B_0 units of money;
3. investing the residual $A_0 - B_0 > 0$ in the riskless bond today.

At time T, we first of all receive back the money invested in the bond, so that we collect $(A_0 - B_0)e^{rT}$ units of money. Additionally we have to return asset A, which we sold at time $t = 0$, without possessing it. Returning some asset we do not have means that we have to fund the purchase of A. Fortunately we bought B at time $t = 0$, such that selling B for a price of B_T just creates enough income to purchase A at a price of $A_T = B_T$. So for clearing our accounts we were not forced to use the positive payout from the bond, such that at the end we have made some risk-free profit. $\square$

The investment strategy in the proof of Proposition 3.2.1 is "risk-free" in the sense that the strategy yields some positive profit no matter what the value of the underlying assets at time T might be. The information that the assets A and B will agree at time T is sufficient for locking-in a guaranteed positive net gain if the asset values at time 0 differ.

Although Proposition 3.2.1 and its proof are almost trivial from the content point of view, they already reflect the typical proof scheme in option pricing theory: For proving some result, the opposite is assumed to hold and an appropriate investment strategy is constructed in order to derive a contradiction to the no-arbitrage principle.

3.2.1 Geometric Brownian Motion

In addition to our bond we now introduce some risky asset A whose values are given by a stochastic process $A = (A_t)_{t\geq 0}$. We call A a *stock* and assume that it evolves like a *geometric Brownian motion* (gBm). This means that the process of asset values is the solution of the *stochastic differential equation*

$$A_t - A_0 \;=\; \mu_A \int_0^t A_s\,ds + \sigma_A \int_0^t A_s\,dB_s \qquad (t \geq 0), \tag{3.1}$$

where $\mu_A > 0$ denotes the *drift* of A, $\sigma_A > 0$ addresses the *volatility* of A, and $(B_s)_{s\geq 0}$ is a standard Brownian motion; see also (3.14) where (3.1) is presented in a slightly more general form incorporating *dividend payments*. Readers with some background in stochastic calculus can easily solve Equation (3.1) by an application of *Itô's formula* yielding

$$A_t \;=\; A_0 \exp\Big((\mu_A - \frac{1}{2}\,\sigma_A^2)\,t + \sigma_A\,B_t\Big) \qquad (t \geq 0). \tag{3.2}$$

This formula shows that gBm is a really intuitive process in the context of stock prices respectively asset values. Just recall from elementary calculus that the exponential function $f(t) = f_0\,e^{ct}$ is the unique solution of the differential equation

$$df(t) \;=\; cf(t)dt\;, \qquad f(0) = f_0\;.$$

Writing (3.1) formally in the following way,

$$dA_t \;=\; \mu_A\,A_t\,dt + \sigma_A\,A_t\,dB_t\;, \tag{3.3}$$

shows that the first part of the stochastic differential equation describing the evolution of gBm is just the "classical" way of describing exponential growth. The difference turning the exponential growth function into a stochastic process arises from the *stochastic differential* w.r.t. Brownian motion captured by the second term in (3.3). This differential adds some *random noise* to the exponential growth, such that instead of a smooth function the process evolves as a random walk with almost surely nowhere differentiable paths. If price movements are of exponential growth, then this is a very reasonable model. Figure 1.6 actually shows a simulation of two paths of a gBm.

Interpreting (3.3) in a naive nonrigorous way, one can write

$$\frac{A_{t+dt} - A_t}{A_t} \;=\; \mu_A \, dt + \sigma_A \, dB_t \; .$$

The left side can be identified with the *relative return* of asset A w.r.t. an "infinitesimal" small time interval $[t, t+dt]$. The equation then says that this return has a *linear trend* with "slope" μ_A and some *random fluctuation term* $\sigma_A \, dB_t$. One therefore calls μ_A the *mean rate of return* and σ_A the *volatility* of asset A. For $\sigma_A = 0$ the process would be a deterministic exponential function, smooth and without any fluctuations. In this case any investment in A would yield a riskless profit only dependent on the time until payout. With increasing volatility σ_A, investments in A become more and more risky. The stronger fluctuations of the process bear a potential of higher wins (*upside potential*) but carry at the same time a higher risk of downturns respectively losses (*downside risk*). This is also expressed by the expectation and volatility functions of gBm, which are given by

$$\begin{aligned} \mathbb{E}[A_t] \;&=\; A_0 \exp(\mu_A t) \qquad (3.4) \\ \mathbb{V}[A_t] \;&=\; A_0^2 \exp(2\mu_A t) \big(\exp(\sigma_A^2 t) - 1\big) \, . \end{aligned}$$

As a last remark we should mention that there are various other stochastic processes that could be used as a model for price movements. In fact, in most cases asset values will not evolve like a gBm but rather follow a process yielding fatter tails in their distribution of log-returns (see e.g. [54]).

3.2.2 Put and Call Options

An *option* is a contract written by an *option seller* or *option writer* giving the *option buyer* or *option holder* the *right* but not the obligation to buy or sell some specified asset at some specified time for some specified price. The time where the option can be *exercised* is called the *maturity* or *exercise date* or *expiration date*. The price written in the option contract at which the option can be exercised is called the *exercise price* or *strike price*.

There are two basic types of options, namely a *call* and a *put*. A call gives the option holder the right to *buy* the underlying asset for the strike price, whereas a put guarantees the option holder the right to *sell* the underlying asset for the exercise price. If the option can be exercised only at the maturity of the option, then the contract is called a *European* option. If the option can be exercised at any time until the final maturity, it is called an *American* option.

There is another terminology in this context that we will frequently use. If someone wants to purchase an asset she or he does not possess at present, she or he currently is *short* in the asset but wants to go *long*. In general, every option contract has two sides. The investor who purchases the option takes a *long position*, whereas the option writer has taken a *short position*, because he sold the option to the investor.

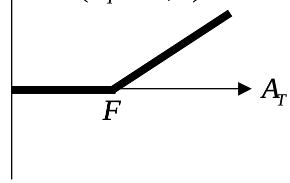
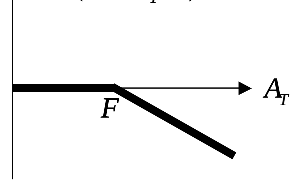
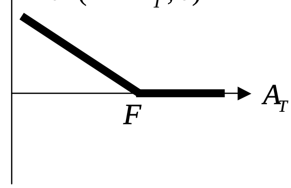
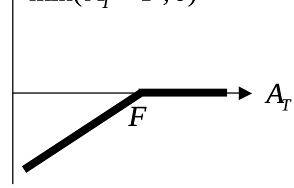
It is always the case that the writer of an option receives cash up front as a compensation for writing the option. But receiving money today includes the potential liabilities at the time where the option is exercised. The question every option buyer has to ask is whether the right to buy or sell some asset by some later date for some price specified today is worth the price she or he has to pay for the option. This question actually is the basic question of *option pricing*.

Let us say the underlying asset of a European call option has price movements $(A_t)_{t\geq 0}$ evolving like a gBm, and the strike price of the call option is F. At the maturity time T one can distinguish between two possible scenarios:

1. Case: $A_T > F$
   In this case the option holder will definitely exercise the option, because by exercising the option he can get an asset worth A_T for the better price F. He will make a net profit in the deal, if the price C_0 of the call is smaller than the price advantage $A_T - F$.

2. Case: $A_T \leq F$
   If the asset is cheaper or equally expensive in the market compared to the exercise price written in the option contract, the

TABLE 3.1: Four different positions are possible in plain-vanilla option trading.

	LONG	SHORT
CALL	▪ buyer/holder of option ▪ payer of option price ▪ option to buy the asset ▪ payoff: $\max(A_T - F, 0)$ 	▪ seller/writer of option ▪ receiver of option price ▪ obligation upon request of option holder to deliver the asset ▪ payoff: $\min(F - A_T, 0)$ 
PUT	▪ buyer/holder of option ▪ payer of option price ▪ option to sell the asset ▪ payoff: $\max(F - A_T, 0)$ 	▪ seller/writer of option ▪ receiver of option price ▪ obligation upon request of option holder to buy the asset ▪ payoff: $\min(A_T - F, 0)$ 

option holder will not exercise the option. In this case, the contract was good for nothing and the price of the option is the investor's loss.

Both cases can be summarized in the *payoff* function of the option, which, in the case of a European call with strike F, is given by

$$\pi : \mathbb{R} \to \mathbb{R}, \quad A_T \mapsto \pi(A_T) = \max(A_T - F, 0) \; .$$

There are altogether four positions in option trading with calls and puts: long call, short call, long put, and short put. Table 3.1 summarizes these four positions and payoffs, clearly showing that for a fixed type of option the payoff of the seller is the reverse of the payoff of the buyer of the option. Note that in the table we have neglected the price of the option, which would shift the payoff diagram along the y-axis, namely into the negative for long positions (because the option price has to be paid) and into the positive for short positions (because the option price will be received as a compensation for writing the option).

It is interesting to mention that long positions have a *limited downside risk*, because the option buyer's worst case is that the money invested in the option is lost in total. The good news for option buyers is the *unlimited upside chance*. Correspondingly, option writers have an *unlimited downside risk*. Moreover, the best case for option writers is that the option holder does not exercise the option. In this case the option price is the net profit of the option writer.

At first glance surprising, European calls and puts are related by means of a formula called the *put-call parity*.

3.2.2 Proposition *Let C_0 respectively P_0 denote the price of a European call respectively put option with strike F, maturity T, and underlying asset A. The risk-free rate is denoted by r. Then,*

$$C_0 + Fe^{-rT} \;=\; P_0 + A_0 \; .$$

This formula is called the put-call parity, connecting puts and calls.

Proof. For proving the proposition we compare two portfolios:

- a long call plus some investment Fe^{-rt} in the risk-free bond;

- a long put plus an investment of one share in asset A.

According to Proposition 3.2.1 we only have to show that the two portfolios have the same value at time $t = T$, because then their values at time $t = 0$ must also agree due to the no-arbitrage principle. We calculate their values at maturity T. There are two possible cases:

$A_T \leq F$: In this case the call option will not be exercised such that the value of the call is zero. The investment Fe^{-rT} in the bond at time $t = 0$ will payout exactly the amount F at $t = T$, such that the value of the first portfolio is F. But the value of the second portfolio is also F, because exercising the put will yield a payout of $F - A_T$, and adding the value of the asset A at $t = T$ gives a total pay out of $F - A_T + A_T = F$.

$A_T > F$: In the same manner as in the first case one can verify that now the value of the first and second portfolio equals A_T.

Altogether the values of the two portfolios at $t = T$ agree. □

The put-call parity only holds for European options, although it is possible to establish some relationships between American calls and puts for a nondividend-paying stock as underlying.

Regarding call options we will now show that it is never optimal to exercise an American call option on a nondividend-paying stock before the final maturity of the option.

3.2.3 Proposition *The price of a European and an American call option are equal if they are written w.r.t. the same underlying, maturity, and strike price.*

Proof. Again we consider two portfolios:

- one American call option plus some cash amount of size Fe^{-rT};
- one share of the underlying asset A.

The value of the cash account at maturity is F. If we forced a payout of cash before expiration of the option, say at time t, then the value of the cash account would be $Fe^{-r(T-t)}$. Because American options can be exercised at any time before maturity, we can exercise the call in portfolio one in order to obtain a portfolio value of

$$A_t - F + Fe^{-r(T-t)} \; < \; A_t \qquad \text{for } t < T \; .$$

Therefore, if the call option is exercised before the expiration date, the second portfolio will in all cases be of greater value than the first portfolio. If the call option is treated like a European option by exercising it at maturity T, then the value of the option is $\max(A_T - F,\, 0)$, such that the total value of the first portfolio equals $\max(A_T,\, F)$. This shows that an American call option on a nondividend-paying stock never should be exercised before the expiration date. □

In 1973 Fischer Black and Myron Scholes found a first analytical solution for the valuation of options. Their method is not too far from the method we used in Propositions 3.2.1 and 3.2.2: By constructing a riskless portfolio consisting of a combination of calls and shares of some underlying stock, an application of the no-arbitrage principle etablished an analytical price formula for European call options on shares of a stock. The pricing formula depends on five parameters:

- the share or asset price A_0 as of today;
- the volatility σ_A of the underlying asset A;
- the strike price F of the option;
- the time to maturity T of the option;
- the risk-free interest rate $r > 0$.

Here we should mention that a key concept leading to the option pricing formulas presented below is the so-called *risk-neutral valuation.* In a world where all investors are *risk-neutral*, all securities earn the risk-free rate. This is the reason why the Black-Scholes formulas do not depend on the drift μ_A of $(A_t)_{t\geq 0}$. *In an arbitrage-free complete market, arbitrage prices of contingent claims equal their discounted expected values under the risk-neutral martingale measure.* Because we will just apply the option pricing formulas without being bothered about their deeper mathematical context, we refer to the literature for further reading. A comprehensive treatment of the mathematical theory of risk-neutral valuation is the book by Bingham and Kiesel [17].

The pricing formula for European calls is then given by

3.2.4 Proposition *The Black-Scholes price of a European call option with parameters (A_0, σ_A, F, T, r) is given by*

$$A_0 N[d_1] - e^{-rT} F N[d_2], \qquad \textit{where}$$

$$d_1 \;=\; \frac{\log(A_0/F) + (r + \sigma_A^2/2)\,T}{\sigma_A \sqrt{T}}\;,$$

$$d_2 \;=\; \frac{\log(A_0/F) + (r - \sigma_A^2/2)\,T}{\sigma_A \sqrt{T}} \;=\; d_1 - \sigma_A \sqrt{T}\;.$$

As usual, $N[\cdot]$ denotes the cumulative standard normal distribution function. In the sequel we write $C_0(A_0, \sigma_A, F, T, r)$ to denote this price.

Proof. A proof can be found in the literature mentioned at the beginning of this chapter. □

Because the prices of a European and an American call option agree due to Proposition 3.2.3, Proposition 3.2.4 also provides the pricing formula for American calls on a nondividend-paying stock. For European put options, the pricing formula follows by an application of the put-call parity.

3.2.5 Proposition *The Black-Scholes price of a European put option with parameters (A_0, σ_A, F, T, r) is given by*

$$e^{-rT} F N[-d_2] - A_0 N[-d_1]\;, \qquad \textit{where}$$

$$d_1 \;=\; \frac{\log(A_0/F) + (r + \sigma_A^2/2)\,T}{\sigma_A \sqrt{T}}\;,$$

$$d_2 \;=\; \frac{\log(A_0/F) + (r - \sigma_A^2/2)\,T}{\sigma_A \sqrt{T}} \;=\; d_1 - \sigma_A \sqrt{T}\;.$$

In the sequel we write $P_0(A_0, \sigma_A, F, T, r)$ to denote this price.

Proof. The put-call parity from Proposition 3.2.2 yields

$$P_0(A_0, \sigma_A, F, T, r) \;=\; C_0(A_0, \sigma_A, F, T, r) + F e^{-rT} - A_0\;.$$

Evaluating the right side of the equation proves the proposition. □

For American put option prices one has to rely on numerical methods, because no closed-form analytic formula is known.

3.3 Merton's Asset Value Model

In this chapter we describe the "classical" asset value model introduced by Merton. As always we assume all involved random variables to be defined on a suitable common probability space. Additionally we make some typical economic assumptions. For example, we assume that markets are *frictionless* with no taxes and without bankruptcy costs. The no-arbitrage principle is assumed to hold. The complete set of conditions necessary for the Merton model can be found in the literature.

3.3.1 Capital Structure: Option-Theoretic Approach

Let's say we consider a firm with risky assets A, such that its asset value process $(A_t)_{t\geq 0}$ follows a gBm. The basic assumption now is that the firm is financed by means of a very simple *capital structure*, namely one *debt obligation* and one type of *equity*. In this case one can write

$$A_0 \;=\; E_0 + D_0 \;, \tag{3.5}$$

where $(E_t)_{t\geq 0}$ is a gBm describing the evolution of equity of the firm, and $(D_t)_{t\geq 0}$ is some stochastic process describing the market value of the debt obligation of the firm, which is assumed to have the cash profile of a *zero coupon bond* with maturity T and interest-adjusted *face value* F. By "interest-adjusted" we mean that F already includes some accrued interest at a rate reflecting the borrowing company's riskiness. The cash profile of debt is then very simple to describe: Debt holders pay a capital of D_0 to the firm at time $t = 0$, and at time $t = T$ they receive an amount equal to F, where F includes the principal D_0 plus the just-mentioned interest payment compensating for the credit risk associated with the credit deal. From the point of view of debt holders, credit risk arises if and only if

$$\mathbb{P}[A_T < F] \;>\; 0 \;,$$

TABLE 3.2: Credit protection by a suitable put option.

	asset value	**debt holder's cash flows**	**debt holder's payout**
$t = 0$	A_0	$-D_0$ (lend money) $-P_0$ (purchase put)	$-D_0 - P_0$
$t = T$	$A_T < F$	A_T (recovery) $F - A_T$ (apply put)	F
$t = T$	$A_T \geq F$	F (receive face value) 0	F

meaning that with positive probability the value of the borrowing company's assets at the debt's maturity is not sufficient for covering the payment F to debt holders. In case this *default probability* is greater than zero, one immediately can conclude that

$$D_0 \; < \; Fe^{-rT} \; ,$$

where r denotes the risk-free interest rate. This inequality must hold because debt holders want some compensation for the credit respectively default risk of its obligor. Such a *risk premium* can be charged implicitly by means of discounting the face value F at a rate higher than the risk-free rate. The payout of debt to the obligor at time $t = 0$ will then be smaller the more risky the obligor's business is.

A typical strategy of debt holders (e.g., a lending bank) is the attempt to *neutralize* the credit risk by purchasing some kind of credit protection. In our case a successful strategy is to buy a suitable derivative. For this purpose, debt holders take a *long position in a put option on A with strike F and maturity T*; see also Figure 3.1. Table 3.2 shows that purchasing the put option guarantees credit protection against the default risk of the borrowing company, because at the maturity date $t = T$ the debt holder's payout equals F no matter if the obligor defaults or not. Therefore, the credit risk of the loan is neutralized and completely hedged. In other words, buying the put transforms the risky corporate loan[3] into a riskless bullet loan with face value F. This brings us to an important conclusion: Taking the hedge into account, the portfolio of debt holders consists of a put option and a loan. Its value at time $t = 0$ is $D_0 + P_0(A_0, \sigma_A, F, T, r)$. The risk-free payout of

[3] Which will be a bond in most cases.

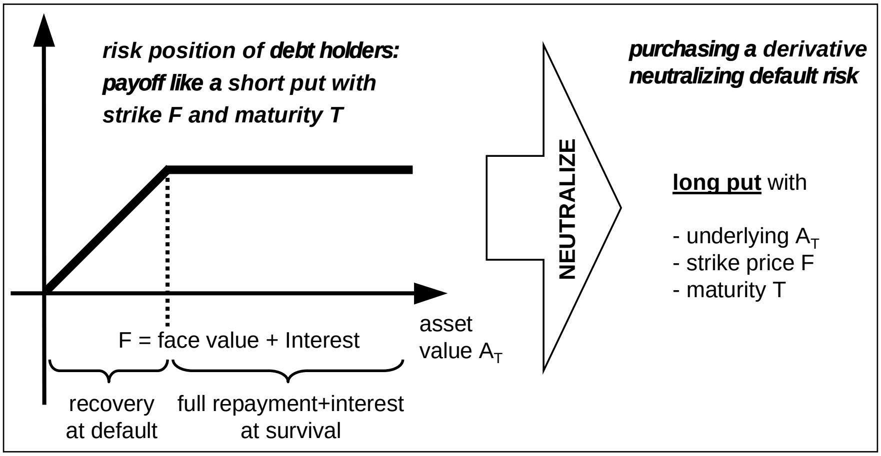


FIGURE 3.1: Hedging default risk by a long put.

this portfolio at time $t = T$ is F. Because we assumed the no-arbitrage principle to hold, the payout of the portfolio has to be discounted to its present value at the risk-free rate r. This implies

$$D_0 + P_0(A_0, \sigma_A, F, T, r) \;=\; Fe^{-rT} \;,$$

so that the present value of debt,

$$D_0 \;=\; Fe^{-rT} - P_0(A_0, \sigma_A, F, T, r) \;, \tag{3.6}$$

is the present value of the face value F discounted at the risk-free rate r corrected by the price for hedging the credit risk by means of a put option.

3.3.1 Corollary [Option-theoretic interpretation of debt]
From the company's point of view, the debt obligation can be described by taking a long position in a put option. From the debt holder's point of view, the debt obligation can be described by writing a put option to the company.

Proof. Using the notation above, at time $t = T$ the company has to pay debt back to debt holders. This yields a cash flow

$$\underbrace{\max(F - A_T, 0)}_{\text{long put payoff}} - F$$

from the company's point of view. From the debt holder's point of view, the cash flow can be written as

$$F + \underbrace{\min(A_T - F, 0)}_{\text{short put payoff}}$$

units of money at time $t = T$. □

So we have found an interpretation of one component of the capital structure of the company in terms of options. But the other component, equity, can also be embedded in an option-theoretic concept: The equity or share holders of the firm have the right to *liquidate* the firm, i.e., paying off the debt and taking over the remaining assets. Let's say equity holders decide to liquidate the firm at the maturity date T of the debt obligation. There are two possible scenarios:

- $A_T < F$:
  This is the default case, where the asset value at maturity is not sufficiently high for paying back debt holders in full. There are no assets left that could be taken over by the equity holders, such that their payoff is zero.

- $A_T \geq F$:
  In this case, there is a net profit to equity holders of $A_T - F$ after paying back the debt.

Summarizing both cases we see that the total payoff to equity holders is $\max(A_T - F, 0)$, which is the payoff of a European call option on A with strike F and maturity T; see Table 3.1. Due to Proposition 3.2.1 the present value of equity therefore is given by

$$E_0 \;=\; C_0(A_0, \sigma_A, F, T, r)\ . \tag{3.7}$$

We conclude as follows:

3.3.2 Corollary [Option-theoretic interpretation of equity]
From the company's point of view, equity can be described by selling a call option to equity holders. Consequently, the position of equity holders is a long call on the firm's asset values.

Proof. The proof follows from the discussion above. □

Combining (3.5) with Conclusions (3.6) and (3.7) we obtain

$$A_0 = E_0 + D_0 = C_0(A_0, \sigma_A, F, T, r) + Fe^{-rT} - P_0(A_0, \sigma_A, F, T, r) \ .$$

Rearranging, we get

$$A_0 + P_0(A_0, \sigma_A, F, T, r) = C_0(A_0, \sigma_A, F, T, r) + Fe^{-rT} \ ,$$

which is nothing but the *put-call parity* we proved in Proposition 3.2.2.

Note that Conclusion 3.3.2 will not be harmed if one allows equity holders to exercise the option before the maturity T. As a justification recall Proposition 3.2.3, saying that the price of a call option is the same no matter if it is European or American.

Our discussion also shows that equity and debt holders have *contrary risk preferences.* To be more explicit, consider

$$C_0(A_0, \sigma_A, F, T, r) \ = \ A_0 - Fe^{-rT} + P_0(A_0, \sigma_A, F, T, r) \ .$$

As can be found in the literature, increasing the riskiness of the investment by choosing some asset A with higher volatility σ_A will also increase the option premium C_0 and P_0 of the call and put options. Therefore, increased volatility (higher risk) is

- good for equity holders, because their natural risk position is a long call, and the value of the call increases with increasing volatility;

- bad for debt holders, because their natural risk position[4] is a short put, whose value decreases with increasing volatility.

Note the unsymmetry in the position of equity holders: Their downside risk is limited, because they can not lose more than their invested capital. In contrast, their upside potential is unlimited. The better the firm performs, the higher the value of the firm's assets, the higher the remaining of assets after a repayment of debt in case the equity holders liquidate the firm.

[4]Which could only be neutralized by a long put.

3.3.2 Asset from Equity Values

The general problem with asset value models is that asset value processes are *not observable.* Instead, what people see every day in the stock markets are equity values. So the big question is how asset values can be derived from market data like equity processes. Admittedly, this is a very difficult question. We therefore approach the problem from two sides. In this section we introduce the classical concept of Merton, saying how one could solve the problem *in principle.* In the next section we then show a way how the problem can be tackled *in practice.* We follow the lines of a paper by Nickell, Perraudin and Varotto [146]. In fact, there are certainly more working approaches for the construction of asset values from market data. For example, in their published papers (see, e.g., Crosbie [36]) Moody's KMV incorporates the *classical* Merton model, but it is well known that in their commercial software (see Section 1.2.3) they have implemented a different, more complicated, and undisclosed algorithm for translating equity into asset values.

The classical approach is as follows: The process of a firm's equity is observable in the market and is given by the company's *market capitalization*, defined by

$$[\text{number of shares}] \ \times \ [\text{value of one share}] \ .$$

Also observable from market data is the *volatility* σ_E of the firm's equity process. Additional information we can get is the *book value* of the firm's *liabilities.* From these three sources,

- equity value of the firm,
- volatility of the firm's equity process, and
- book value of the firm's liabilities,

we now want to infer the *asset value process* $(A_t)_{t\geq 0}$ (as of today). Once more we want to remark that the following is more a "schoolbook model" than a working approach. In contrast, the next paragraph will show a more applicable solution.

Let us assume we consider a firm with the same simple capital structure[5] as introduced in (3.5). From Conclusion 3.3.2 we already know

[5]Actually it is in part due to the assumption of a simple capital structure that the classical Merton model is not really applicable in practice.

that the firm's equity can be seen as a call option on the firm's assets, written by the firm to the equity or share holders of the firm. The strike price F is determined by the book value of the firm's liabilities, and the maturity T is set to the considered planning horizon, e.g., one year. According to (3.7) this option-theoretic intepretation of equity yields the functional relation

$$E_t \;=\; C_t(A_t, \sigma_A, F, (T-t), r) \qquad (t \in [0,T]) \tag{3.8}$$

This functional relation can be locally inverted, due to the *implicit function theorem*, in order to solve (3.8) for A_t. Therefore, the asset value of the firm can be calculated as a function of the firm's equity and the parameters F, t, T, r, and the asset volatility σ_A. If, as we already remarked, asset value processes are not observable, the asset volatility also is not observable. It therefore remains to determine the asset volatility σ_A in order to obtain A_t from (3.8).

Here, we actually need some insights from stochastic calculus, such that for a brief moment we are now forced to use results for which an exact and complete explanation is beyond the scope of the book. However, in the next section we will provide some "heuristic" background on *pathwise stochastic integrals*, such that at least some open questions will be answered later on. As always we assume for the sequel that all random variables respectively processes are defined on a suitable common probability space.

Recall that we assumed that the asset value process $(A_t)_{t\geq 0}$ is assumed to evolve like a *geometric Brownian motion* (see Section 3.2.1), meaning that A solves the stochastic differential equation

$$A_t - A_0 \;=\; \mu_A \int_0^t A_s\, ds + \sigma_A \int_0^t A_s\, dB_s^{(A)}\,.$$

Following almost literally the arguments in Merton's approach, we assume for the equity of the firm that $(E_t)_{t\geq 0}$ solves the stochastic differential equation

$$E_t - E_0 \;=\; \int_0^t \mu_E(s) E_s\, ds + \int_0^t \sigma_E(s) E_s\, dB_s^{(E)}\;. \tag{3.9}$$

Here, $(B_t^{(A)})_{t\geq 0}$ and $(B_t^{(E)})_{t\geq 0}$ denote *standard Brownian motions*. Applying *Itô's lemma* to the function[6]

$$f(t, A_t) \;=\; C_t(A_t, \sigma_A, F, (T-t), r)$$

and comparing[7] the martingale part of the resulting equation with the martingale part of Equation (3.9) yield in informal differential notation

$$\sigma_E E_t dB_t^{(E)} \;=\; f_2(t, A_t)\sigma_A A_t dB_t^{(A)} \;, \tag{3.10}$$

where $f_2(\cdot\,,\cdot)$ denotes the partial derivative w.r.t. the second component. But the coefficients of stochastic differential equations are uniquely determined, such that from (3.10) we can conclude

$$\frac{\sigma_E}{\sigma_A} \;=\; \frac{A_t f_2(t, A_t)}{E_t} \;. \tag{3.11}$$

Solving (3.11) for σ_A and inserting the solution into Equation (3.8) yields A_t for $t \in [0, T]$.

This concludes our discussion of the classical Merton model. We now proceed to a more mathematical as well as more applicable approach. For this purpose, we explicitely define the stochastic integral for a specific class of integrands in Section 3.4.1. Then, in Section 3.4.2, we present a more accurate derivation of the Black-Scholes partial differential equation due to Duffie [46]. Additionally, we introduce a boundary condition going back to Perraudin et al. [146] which specifies a reasonable relation between asset values and equities.

3.4 Transforming Equity into Asset Values: A Working Approach

Let us begin with a few words on pathwise Itô Calculus (see Revuz and Yor [159], and Foellmer [63]). The following treatment is rather self-contained because no difficult prerequisites from measure theory are required. Unfortunately, the pathwise calculus is only valid for a specific type of *trading strategies*, as we will later see.

[6] We refer to the literature for checking that the conditions necessary for applying Itô's lemma are satisfied in our case.

[7] Such a comparison is justified, because the components of so-called *Itô processes* are uniquely determined.

3.4.1 Itô's Formula "Light"

In this paragraph we want to establish the existence of a *pathwise stochastic integral* by an argument based on elementary calculus, thereby avoiding the usual requirements from measure theory.

Let ω be a real-valued continuous function of time t with finite *quadratic variation* $\langle\omega\rangle$, and $F \in \mathcal{C}^2$. Denoting by $\mathcal{Z}_n$ a sequence of partitions of the interval $[0, t)$ with $\mathrm{mesh}(\mathcal{Z}_n) \to 0$, a Taylor expansion up to second order yields

$$\begin{aligned} F(\omega_t) - F(\omega_0) \;&=\; \lim_{n\to\infty}\Bigg(\sum_{(t_i)\in\mathcal{Z}_n^t} F'(\omega_{t_i})(\omega_{t_{i+1}} - \omega_{t_i}) \\ &+ \sum_{(t_i)\in\mathcal{Z}_n^t} \frac{1}{2} F''(\omega_{t_i})(\omega_{t_{i+1}} - \omega_{t_i})^2 + o\left((\Delta\omega)^2\right)\Bigg). \end{aligned} \tag{3.12}$$

From the existence of the quadratic variation of ω we conclude that the second term in (3.12) converges to

$$\frac{1}{2}\int_0^t F''(\omega_s)d\langle\omega\rangle_s.$$

Hence the limit of the first term in (3.12) also exists. It is denoted by

$$\int_0^t F'(\omega_s)d\omega_s$$

and called a *stochastic integral*. In this context, the *Itô formula* is just a by-product of the Taylor expansion (3.12), and can be obtained by writing (3.12) in the limit form

$$F(\omega_t) - F(\omega_0) \;=\; \int_0^t F'(\omega_s)d\omega_s + \frac{1}{2}\int_0^t F''(\omega_s)d\langle\omega\rangle_s\,. \tag{3.13}$$

The just-derived stochastic integral can be interpreted in terms of trading gains. The discrete approximation

$$\sum_{t_i\in\mathcal{Z}_n^t} F'(\omega_{t_i})(\omega_{t_{i+1}} - \omega_{t_i})$$

of the stochastic integral is the *gain* of the following trading strategy:

Buy $F'(\omega_{t_i})$ shares of a financial instrument with value ω_t at time t_i.

The gain over the time interval $[t_i, t_{i+1})$ then equals

$$F'(\omega_{t_i})(\omega_{t_{i+1}} - \omega_{t_i}) \ .$$

The stochastic integral is just the limit of the sum over all these trading gains in the interval $[0, t)$. From these observations it becomes also clear why the stochastic integral as introduced above sometimes is called *non-anticipating.* This terminology just refers to the fact that the investment took place at the beginning of the intervals $[t_i, t_{i+1})$.

For a thorough introduction to the stochastic integral in the more general measure-theoretic setting we refer to the literature mentioned at the beginning of this chapter. However, the intuitive interpretation of the stochastic integral as the gain of a (non-anticipating) trading strategy and the basic structure of the Itô formula remain both valid in the measure-theoretic approach.

3.4.2 Black-Scholes Partial Differential Equation

In this paragraph we follow the approach outlined in DUFFIE [46]. As in the previous paragraphs, we assume that the asset value process $A = (A_t)_{t \geq 0}$ follows a geometric Brownian motion driven by some Brownian motion B. But this time we include *dividend payments*, such that A is the solution of the stochastic differential equation

$$A_t - A_0 \ = \ \int_0^t (\mu_A A_s - C_{A,s}) ds + \sigma_A \int_0^t A_s dB_s \ , \tag{3.14}$$

where $C_{A,s}$ is the dividend paid by the firm at time s. In the literature the following, more intuitive differential notation of (3.14) is also used

$$dA_t \ = \ (\mu_A A_t - C_{A,t}) dt + \sigma_A A_t dB_t \ .$$

In previous paragraphs the capital structure of the considered firm contained one debt obligation. Here we assume that the market value of debt D_t at time t is just a nonstochastic exponential function,

$$D_s \ = \ D_0 e^{\mu_D s} \ .$$

By Itô's formula (3.13), any process $(E_t)_{t \geq 0}$ represented by a smooth function $E(x, y, t)$ applied to the processes A and D,

$$E_t \ = \ E(A_t, D_t, t), \qquad E \in \mathcal{C}^{2,1,1},$$

solves the integral equation

$$
\begin{aligned}
E_t - E_0 \;=\; & \int_0^t [\partial_t E(A_s, D_s, s) + (\mu_A A_s - C_{A,s})\partial_x E(A_s, D_s, s) \\
& + \mu_D D_s \partial_y E(A_s, D_s, s) + \frac{1}{2}\sigma_A^2 A_s^2 \partial_{xx} E(A_s, D_s, s)]ds \\
& + \sigma_A \int_0^t A_s \partial_x E(A_s, D_s, s) dB_s \;.
\end{aligned}
$$

We now want to construct a so-called *self-financing* trading strategy (η_t, θ_t) such that

- $\eta_t A_t + \theta_t K_t \;=\; E_t$
- $K_t \;=\; e^{rt}$,

where K_t denotes the value of a risk-free investment (e.g., some treasury bond) earning interest at the risk-free rate r. The attribute *self-financing* means that the value of the portfolio, at time t consisting of η_t shares of A and θ_t shares of K, has a value equal to the initial investment plus trading gains. More explicitly,

$$
\begin{aligned}
\eta_t A_t + \theta_t K_t \;=\; & \eta_0 A_0 + \theta_0 K_0 + \qquad (3.15) \\
& + \int_0^t \eta_s dA_s + \int_0^t \eta_s C_{A,s} ds + \int_0^t \theta_s dK_s \;.
\end{aligned}
$$

The assumption that there is a self-financing strategy[8] that perfectly replicates E_t then leads to

$$
\begin{aligned}
& \int_0^t [\eta_s \mu_A A_s + \theta_s K_s r]ds + \int_0^t \eta_s \sigma_A A_s dB_s = \qquad (3.16) \\
& = \int_0^t [\partial_t E(A_s, D_s, s) + (\mu_A A_s - C_{A,s})\partial_x E(A_s, D_s, s)
\end{aligned}
$$

[8] A straightforward application of Itô's formula would imply that

$$
\eta_t A_t - \eta_0 A_0 = \int_0^t A_s d\eta_s + \int_0^t \eta_s dA_s + <\eta, A>_t \;.
$$

This result would not lead to the Black-Scholes PDE. "Self-financing" therefore is essential from a mathematical point of view.

$$+ \mu_D D_s \partial_y E(A_s, D_s, s) + \frac{1}{2}\sigma_A^2 A_s^2 \partial_{xx} E(A_s, D_s, s)]ds$$

$$+ \sigma_A \int_0^t \partial_x E(A_s, D_s, s) A_s dB_s \ . \tag{3.17}$$

The unique decomposition of an Itô process into a stochastic integral with respect to B and a drift leads to

$$\eta_t \ = \ \partial_x E(A_t, D_t, t)$$

taking (3.16) and (3.17) into account. Since the trading strategy replicates E_t we necessarily have

$$\theta_t \ = \ \frac{1}{K_t}[E_t - \partial_x E(A_t, D_t, t) A_t] \ .$$

The comparison of the coefficient of dt implies the equation

$$0 \ = \ \partial_t E(A_s, D_s, s) + (rA_s - C_{A,s})\partial_x E(A_s, D_s, s) \tag{3.18}$$

$$+ \mu_D D_s \partial_y E(A_s, D_s, s) + \frac{1}{2}\sigma_A^2 A_s^2 \partial_{xx} E(A_s, D_s, s) - rE_s \ .$$

As in NICKELL, PERRAUDIN and VAROTTO [146], let us now specify the dividend by $C_{A,s} = \delta A_s$. Then, E solves Equation (3.18) if it solves the partial differential equation

$$0 \ = \ \partial_t E(x, y, s) + (rx - \delta x)\partial_x E(x, y, s) + \mu_D D_s \partial_y E(x, y, s) \tag{3.19}$$

$$+ \frac{1}{2}\sigma_A^2 x^2 \partial_{xx} E(x, y, s) - rE(x, y, s) \ .$$

For $\delta = 0$ and $D_0 = 0$, the last equation becomes the celebrated *Black-Scholes formula.* It should be clear that we have to specify boundary conditions for (3.19).

As a first approach, let us assume that the firm has an earnings flow $\delta(A - D)$, which entails $\partial_t E = \delta(A - D)$, and hence, Equation (3.19) becomes an ordinary differential equation.

Analogous to the lines in [146] we now assume that the firm is declared to be in bankruptcy as soon as the ratio of *assets to liabilities* A_t/D_t hits some low level for the very first time. We call this critical threshold γ and assume the equity-holders to receive no money in case of a bankruptcy settlement. Then, the value of the firm's equity, E, satisfies the differential equation (3.19) subject to the following boundary conditions:

$$E(A,D)|_{A/D=\gamma} = 0 \quad \text{and} \quad \lim_{A/D\to\infty} E(A,D) = A - \frac{\delta}{r}D \ .$$

For some background on differential equations, refer to [195].

Now let us present the solution of (3.19) under these boundary conditions. It is given by

$$E(A,D) \ = \ D\Big[\frac{A}{D} - \frac{\delta}{r-\mu_D} - \Big(\gamma - \frac{\delta}{r-\mu_D}\Big)\Big(\frac{A/D}{\gamma}\Big)^{\lambda}\Big] \ , \qquad (3.20)$$

where λ is defined in dependence on σ_A by

$$\lambda \ = \ \lambda(\sigma_A) =$$

$$= \ \frac{1}{\sigma_A^2}\Big[\Big(\frac{\sigma_A^2}{2} + \delta + \mu_D - r\Big) - \sqrt{(r - \frac{\sigma_A^2}{2} - \delta - \mu_D)^2 + 2\sigma_A^2(r-\mu_D)}\Big] \ .$$

In this model, the level of the bankruptcy trigger γ is chosen by the equity holders, since the firm will continue to operate until equity holders are unwilling to absorb more than the already occurred losses. The threshold γ therefore is determined by the first order condition $\partial_\gamma E = 0$, from which it follows that

$$\gamma \ = \ \frac{\lambda}{\lambda - 1}\,\frac{\delta}{r-\mu_D}.$$

Figure 3.2 shows the asset-equity relation from Equation (3.20) for some parameter sets.

Now, if A were an observable variable and σ_A were known, E would be specified by (3.20). But in "option terminology" we observe the price of an option, namely the equity price. Therefore, we can only estimate the volatility of the changes in the price of the option. From this we have to find the value of the underlying instrument A and its volatility σ_A. This means for determining A and σ_A we need a second equation. Since E is an Itô process, its quadratic variation $\langle E\rangle$ can be read off from Equation (3.17) as

$$\langle E\rangle_t \ = \ \int_0^t \sigma_{E,s}^2 ds \ = \ \sigma_A^2 \int_0^t A_s^2\,[\partial_x E(A_s,D_s,s)]^2\,ds \ . \qquad (3.21)$$

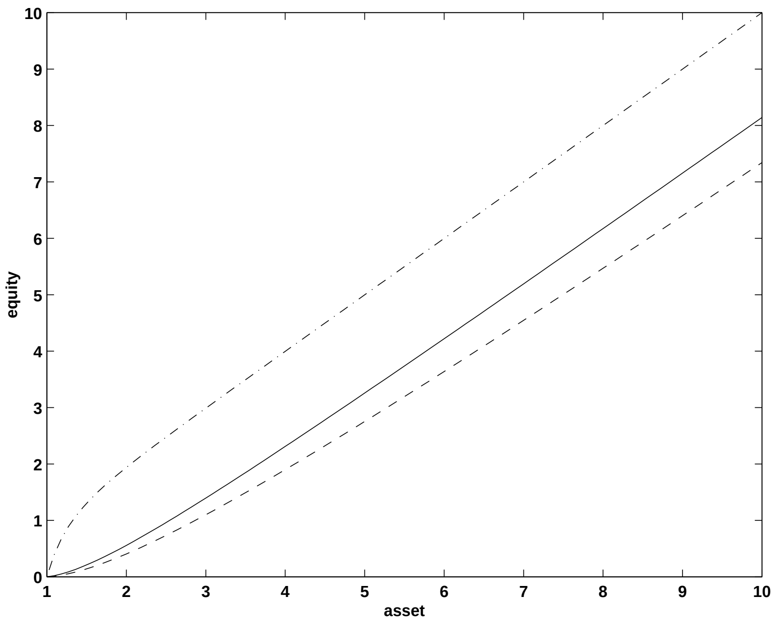


FIGURE 3.2: Asset-Equity relation, Equation (3.20), for parameter sets $(\delta, r, \gamma, \mu, \sigma_A)$ and $D = 1$: (-) solid $(0.1, 0.05, 1, 0.0, 0.1)$, (–) dashed $(0.1, 0.05, 1, 0.03, 0.1)$, (-.) dashed-dotted $(0.0, 0.05, 1, 0.03, 0.1)$.

Therefore A and σ_A have to solve the two equations (3.19 and 3.21), which are strictly speaking *pathwise* equations, since E has a stochastic volatility. Nevertheless Equation (3.21) can be replaced by

$$\sigma_{E,t}^2 \;=\; \sigma_A^2 A_t^2 \left[\partial_x E(A_t, D_t)\right]^2 \;.$$

Let us define

$$E'(A, D, \sigma_A) \;=\; \partial_x E(A, D, \sigma_A) \tag{3.22}$$

$$= 1 - D\Big(\gamma - \frac{\delta}{r - \mu_D}\Big)\Big(\frac{1}{D\gamma}\Big)^{\lambda(\sigma_A)} \lambda(\sigma_A) A^{\lambda(\sigma_A)-1} \;.$$

If we observe E at time t and know the estimate $\sigma_{E,t}$ for the equity volatility, then A and σ_A have to solve the equations

$$E \;=\; D\left[\frac{A}{D} - \frac{\delta}{r - \mu_D} - \left(\gamma - \frac{\delta}{r - \mu_D}\right)\left(\frac{A/D}{\gamma}\right)^{\lambda(\sigma_A)}\right] \tag{3.23}$$

$$\sigma_{E,t} \;=\; \sigma_A A\left[1 - D\left(\gamma - \frac{\delta}{r - \mu_D}\right)\left(\frac{1}{D\gamma}\right)^{\lambda(\sigma_A)} \lambda(\sigma_A) A^{\lambda(\sigma_A)-1}\right] \;. \tag{3.24}$$

As a further simplification it is often assumed that E locally evolves like a geometric Brownian motion, which leads to $\sigma_{E,t} = \sigma_E E$ for some σ_E.

In the implementation one usually starts with some $\sigma_A = \sigma_A^0$. For example, the equity volatility is used to generate two time series $(A_s)_{s\geq 0}$ and $(E_s)_{s\geq 0}$. Then, the volatility of E is estimated, and the parameter σ_A^1 is adjusted to a higher or lower level, trying to best match the estimated volatility of E with the observed equity volatility. One proceeds that way until the σ_E^n, implied by σ_A^n, is close to the observed σ_E. Observe also that the set of equations (3.23) and (3.24) can be generalized to any contingent claim approach for the asset values, once a functional relationship $E = E(A, D, \sigma_A, t)$ is specified between assets A, debt D, and equity E. Conceptually, they look like

$$E \;=\; E(A, D, \sigma_A) \;, \qquad \sigma_E E \;=\; \sigma_A A E'(A, D, \sigma_A) \;.$$

This concludes our discussion of asset value models.

Further Reading

In the classical Merton approach to credit risk the default can only occur at the debt's maturity and the firm's default probability is the likelihood that the firm's asset value at maturity is below the face value of the debt. Several authors extended this simplistic approach by defining the default as the first time the asset value falls below a (possibly stochastic) barrier, which leads to the so-called *first-passage-time* models, Black and Cox [18] and Zhou [194]. Similar to Black and Cox, Leland [121] and Leland and Toft [123] use first-passage-time models to investigate the optimal capital structure by maximizing the equity value. Default risk together with interest rate risk is treated in Longstaff and Schwartz [122]. An important aspect of the structural approach is that it allows to formulate a link between equity correlations, asset-value correlations and default correlations, see, e.g., Zhou [194], Frey and McNeil [68] and Giesecke [77, 78].

Chapter 4

The CreditRisk$^+$ Model

In Section 2.4.2 we already described the CreditRisk$^+$ model as a Poissonian mixture with gamma-distributed random intensities for each sector. In this section we will explain CreditRisk$^+$ in some greater detail. The justification for another and more exhaustive chapter on CreditRisk$^+$ is its broad acceptance by many credit risk managing institutes. Even in the new Capital Accord (some references regarding the Basel II approach are Gordy [84], Wilde [187], and the IRB consultative document [148]), CreditRisk$^+$ was originally applied for the calibration of the so-called *granularity adjustment* in the context of the *Internal Ratings-based Approach* (IRB) of regulatory capital risk weights. The popularity of CreditRisk$^+$ has two major reasons:

- It seems easier to calibrate data to the model than is the case for multi-factor asset value models. Here we intentionally said "it seems" because from our point of view the calibration of bank-internal credit data to a multi-sector model is in general neither easier nor more difficult than the calibration of a multi-factor model on which an asset value model can be based.

- The second and maybe most important reason for the popularity of CreditRisk$^+$ is its *closed-form* loss distribution. Using *probability generating functions*, the CreditRisk$^+$ model offers (even in case of more than one sector) a full analytic description of the portfolio loss of any given credit portfolio. This enables users of CreditRisk$^+$ to compute loss distributions in a quick and still "exact" manner. For many applications of credit risk models, this is a "nice-to-have" feature, e.g., in pricing or ABS structuring.

Before going into the details of the CreditRisk$^+$ model, we like to present a quotation from the CreditRisk$^+$ Technical Document [35] on page 8. There we find that

CreditRisk$^+$ focuses on modeling and managing credit default risk.

In other words, CreditRisk$^+$ helps to quantify the potential risk of defaults and resulting losses in terms of exposure in a given portfolio. Although it incorporates a *term structure of default rates* (more explicitly yearly *marginal* default rates) for implementing multi-year loss distributions (see [35], A5.2), it is not an appropriate choice if one is interested in a mark-to-market model of credit risk.

4.1 The Modeling Framework of CreditRisk$^+$

Crucial in CreditRisk$^+$ is the use of probability-generating functions[1]. Recall that the generating function of a Poisson random variable L' with intensity λ is given by

$$G(z) \;=\; \sum_{k=0}^{\infty} \mathbb{P}[L' = k]\, z^k \;=\; e^{-\lambda} \sum_{k=0}^{\infty} \frac{\lambda^k}{k!}\, z^k \;=\; e^{\lambda(z-1)} \;. \qquad (4.1)$$

In order to reduce the computational effort, CreditRisk$^+$ groups the individual exposures of the obligors in a considered portfolio into *exposure bands*. This is done as follows:

Choose an exposure *unit* amount E. Analogously to Chapter 1, denote for any obligor i its *Expected Loss* by EL_i, its *Exposure At Default* by EAD_i, and its *Loss Given Default* by LGD_i. The exposure that is subject to be lost after an obligor's default is then

$$E_i \;=\; \mathrm{EAD}_i \times \mathrm{LGD}_i \;, \qquad (4.2)$$

assuming a nonrandom LGD. The exposure ν_i respectively the Expected Loss ε_i of obligor i in multiples of the exposure unit E is given by

$$\nu_i \;=\; \frac{E_i}{E}\;, \qquad \varepsilon_i \;=\; \frac{\mathrm{EL}_i}{E}\;.$$

[1] In probability theory there are three concepts of translating a probability distribution into a functional context, namely the *Fourier transform*, the *Laplace transform* (which is in case of distributions on $\mathbb{R}^d_+$ often more convenient), and the *probability-generating function* (often preferably used for distributions on $\mathbb{Z}_+$). The latter is defined by the function $z \mapsto \mathbb{E}[z^X]$ for a random variable X. Regarding basic properties of generating functions we refer to Section 2.8.1.

From this point on, CreditRisk$^+$ "forgets" the exact exposures from the original portfolio and uses an approximation by means of exposure bands by rounding the exposures ν_i to the nearest integer number. In other words, every exposure E_i is replaced by the closest integer multiple of the unit exposure E. Already one can see that an appropriate choice of E is essential in order to end up at an approximation that is on one hand "close" enough to the original exposure distribution of the portfolio in order to obtain a loss distribution applicable to the original portfolio, and on the other hand efficient enough to really partition the portfolio into m_E exposure bands, such that m_E is significantly smaller than the original number of obligors m. An important "rule-of-thumb" for making sure that not too much precision is lost is to at least take care that the width of exposure bands is "small" compared to the average exposure size in the portfolio. Under this rule, *large* portfolios (containing many loans) should admit a good approximation by exposure bands in the described manner.

In the sequel we write $i \in [j]$ whenever obligor[2] i is placed in the exposure band j. After the exposure grouping process, we have a partition of the portfolio into m_E exposure bands, such that obligors in a common band $[j]$ have the common exposure $\nu_{[j]}E$, where $\nu_{[j]} \in \mathbb{N}_0$ is the integer multiple of E representing all obligors i with

$$\min\{|\nu_i - n| \;:\; n \in \mathbb{N}_0\} \;=\; |\nu_i - \nu_{[j]}|$$
$$\text{where} \quad i = 1, ..., m;\; i \in [j];\; j = 1, ..., m_E.$$

In cases where ν_i is an odd-integer multiple of 0.5, the above minimum is not uniquely defined. In such cases (which are obviously not very likely) one has to make a decision, if an up- or down-rounding would be appropriate. In the sequel we only consider $\nu_{[j]} \in \mathbb{N}$, excluding 0.

Now let us discuss how to assign a default intensity to a given exposure band. Because CreditRisk$^+$ plays in a Poissonian world, every obligor in the portfolio has its individual (one-year) *default intensity* λ_i, which can be calibrated from the obligor's one-year default probability PD_i by application of (2.12),

$$\lambda_i \;=\; -\log(1 - \mathrm{PD}_i) \qquad (i = 1, ..., m). \tag{4.3}$$

[2]Here we make the simplifying assumption that the number of loans in the portfolio equals the number of obligors involved. This can be achieved by aggregating different loans of a single obligor into one loan. Usually the PD, EAD, and LGD of such an aggregated loan are exposure-weighted average numbers.

Because the expectation of $L'_i \sim Pois(\lambda_i)$ is $\mathbb{E}[L'_i] = \lambda_i$, the *expected number of defaults* in exposure band $[j]$ (using the additivity of expectations) is given by

$$\lambda_{[j]} \;=\; \sum_{i\in[j]} \lambda_i \; . \tag{4.4}$$

The *Expected Loss* in band $[j]$ will be denoted by $\varepsilon_{[j]}$ and is calculated by multiplying the expected number of defaults in band $[j]$ with the band's exposure,

$$\varepsilon_{[j]} \;=\; \lambda_{[j]}\nu_{[j]} \; . \tag{4.5}$$

Here, the CreditRisk$^+$ Technical Document suggests making an *adjustment of the default intensities* λ_i (which so far have not been affected by the exposure band approximation process) in order to preserve the original value of the obligor's Expected Losses. This could be done by defining an adjustment factor γ_i for every obligor i by

$$\gamma_i \;=\; \frac{E_i}{\nu_{[j]}E} \qquad (i \in [j],\; j = 1, ..., m_E) \; . \tag{4.6}$$

Replacing for every obligor i the original default intensity λ_i by $\gamma_i\lambda_i$ with γ_i as defined in (4.6) preserves the original ELs after approximating the portfolio's exposure distribution by a partition into exposure bands. In the following we assume without loss of generality that the default intensities λ_i already include the adjustment (4.6). From (4.4) respectively (4.5) it is straightforward to write down the *portfolio's expected number of default events* (respectively the *portfolio's overall default intensity*), namely

$$\lambda_{PF} \;=\; \sum_{j=1}^{m_E} \lambda_{[j]} \;=\; \sum_{j=1}^{m_E} \frac{\varepsilon_{[j]}}{\nu_{[j]}} \; . \tag{4.7}$$

After these preparations we are now ready to describe the construction of the CreditRisk$^+$ loss distribution. We will proceed in two steps, starting with a portfolio of independent obligors and then mixing the involved Poisson distributions by means of a sector model as indicated in Section 2.4.2.

4.2 Construction Step 1: Independent Obligors

We begin with a portfolio of m independent obligors whose default risk is modeled by Poisson variables L'_i. As already mentioned in Section 2.2.1, Poisson models allow for multiple defaults of a single obligor. This is an unpleasant, but due to the small occurrence probability, mostly ignored feature of all Poisson approaches to default risk.

Involving the (nonrandom) exposures E_i as defined in (4.2), we obtain loss variables

$$E_i L'_i \qquad \text{where} \quad L'_1 \sim Pois(\lambda_1)\ ,\ \ldots\ ,\ L'_m \sim Pois(\lambda_m) \tag{4.8}$$

are independent Poisson random variables. Grouping the individual exposures E_i into exposure bands $[j]$ and assuming the intensities λ_i to incorporate the adjustments by the factors γ_i as described in the introduction, we obtain new loss variables $\nu_{[j]}L'_i$, where losses are measured in multiples of the exposure unit E. Because obligors are assumed to be independent, the number of defaults L' in the portfolio respectively $L'_{[j]}$ in exposure band j also follow a Poisson distribution, because the convolution of independent Poisson variables yields a Poisson distribution. We obtain

$$L'_{[j]} = \sum_{i\in[j]} L'_i \sim Pois(\lambda_{[j]})\ , \qquad \lambda_{[j]} \ = \ \sum_{i\in[j]} \lambda_i\ , \tag{4.9}$$

for the number of defaults in exposure band $[j]$, $j = 1, ..., m_E$, and

$$L' = \sum_{j=1}^{m_E} \sum_{i\in[j]} L'_i \sim Pois\Big(\sum_{j=1}^{m_E} \lambda_{[j]}\Big) = Pois(\lambda_{PF}) \tag{4.10}$$

(see (4.7)), for the portfolio's number of defaults. The corresponding losses (counted in multiples of the exposure unit E) are given by

$$\tilde{L}'_{[j]} = \nu_{[j]} L'_{[j]} \qquad \text{respectively} \qquad \tilde{L}' = \sum_{j=1}^{m_E} \nu_{[j]} L'_{[j]} = \sum_{j=1}^{m_E} \tilde{L}'_{[j]}\ . \tag{4.11}$$

Due to grouping the exposures $\nu_{[j]} \in \mathbb{N}$ together, we can now conveniently describe the portfolio loss by the probability-generating function of the random variable $\tilde{L}'$ defined in (4.11), applying the convolution theorem[3] for generating functions,

$$G_{\tilde{L}'}(z) \;=\; \prod_{j=1}^{m_E} G_{\tilde{L}'_{[j]}}(z) \;=\; \prod_{j=1}^{m_E}\sum_{k=0}^{\infty}\mathbb{P}[\tilde{L}'_{[j]} = \nu_{[j]}k]\, z^{\nu_{[j]}k} \tag{4.12}$$

$$= \prod_{j=1}^{m_E}\sum_{k=0}^{\infty}\mathbb{P}[L'_{[j]} = k]\, z^{\nu_{[j]}k} = \prod_{j=1}^{m_E}\sum_{k=0}^{\infty} e^{-\lambda_{[j]}}\frac{\lambda_{[j]}^k}{k!}\, z^{\nu_{[j]}k}$$

$$= \prod_{j=1}^{m_E} e^{-\lambda_{[j]}+\lambda_{[j]}z^{\nu_{[j]}}} \;=\; \exp\Big(\sum_{j=1}^{m_E}\lambda_{[j]}(z^{\nu_{[j]}}-1)\Big)\,.$$

So far we assumed independence among obligors and were rewarded by the nice closed formula (4.12) for the generating function of the portfolio loss. In the next section we drop the independence assumption, but the nice feature of CreditRisk$^+$ is that, nevertheless, it yields a closed-form loss distribution, even in the case of correlated defaults.

4.3 Construction Step 2: Sector Model

A key concept of CreditRisk$^+$ is *sector analysis*. The rationale underlying sector analysis is that the volatility of the default intensity of obligors can be related to the volatility of certain underlying factors incorporating a common systematic source of credit risk. Associated with every such background factor is a so-called *sector*, such that every obligor i admits a breakdown into *sector weights* $w_{is} \geq 0$, $\sum_{s=1}^{m_S} w_{is} = 1$, expressing for every $s = 1, ..., m_S$ that sector s contributes with a fraction w_{is} to the default intensity of obligor i. Here m_S denotes the number of involved sectors. Obviously the calibration of sectors and sector weights is the crucial challenge in CreditRisk$^+$. For example, sectors could be constructed w.r.t. industries, countries, or rating classes.

---

[3]For independent variables, the generating function of their convolution equals the product of the corresponding single generating functions.

In order to approach the sector model of CreditRisk$^+$ we rewrite Equation (4.12):

$$G_{\tilde{L}'}(z) \;=\; \exp\Big(\sum_{j=1}^{m_E} \lambda_{[j]}(z^{\nu_{[j]}} - 1)\Big) \tag{4.13}$$

$$=\; \exp\Big(\lambda_{PF}\Big(\sum_{j=1}^{m_E} \frac{\lambda_{[j]}}{\lambda_{PF}}\, z^{\nu_{[j]}} - 1\Big)\Big)\;,$$

where λ_{PF} is defined as in (4.7). Defining functions

$$G_{L'}(z) = e^{\lambda_{PF}(z-1)} \qquad \text{and} \qquad G_N(z) = \sum_{j=1}^{m_E} \frac{\lambda_{[j]}}{\lambda_{PF}}\, z^{\nu_{[j]}}\;, \tag{4.14}$$

we see that the generating function of the portfolio loss variable $\tilde{L}'$ can be written as

$$G_{\tilde{L}'}(z) \;=\; G_{L'} \circ G_N(z) \;=\; e^{\lambda_{PF}(G_N(z)-1)}\;. \tag{4.15}$$

Therefore, the portfolio loss $\tilde{L}'$ has a so-called *compound* distribution, essentially meaning that the randomness inherent in the portfolio loss is due to the compound effect of two independent sources of randomness. The first source of randomness arises from the uncertainty regarding the number of defaults in the portfolio, captured by the Poisson random variable L' with intensity λ_{PF} defined in (4.10). The function $G_{L'}(z)$ is the generating function of L'; recall (4.1). The second source of randomness is due to the uncertainty about the exposure bands affected by the L' defaults. The function $G_N(z)$ is the generating function of a random variable N taking values in $\{\nu_{[1]}, ..., \nu_{[m_E]}\}$ with distribution

$$\mathbb{P}[N = \nu_{[j]}] \;=\; \frac{\lambda_{[j]}}{\lambda_{PF}} \qquad (j = 1, ..., m_E). \tag{4.16}$$

For some more background on compound[4] distributions, refer to the literature. For example in [86] the reader will find theory as well as

[4]Compound distributions arise very naturally as follows: Assume $X_0, X_1, X_2, ...$ be i.i.d. random variables with generating function G_X. Let $N \in \mathbb{N}_0$ be a random variable, e.g., $N \sim Pois(\lambda)$, independent of the sequence $(X_i)_{i\geq 0}$. Denote the generating function of N by G_N. Then, the generating function of $X_1 + \cdots + X_N$ is given by $G = G_N \circ G_X$. In the case where the distribution of N is degenerate, e.g., $\mathbb{P}[N = n] = 1$, we obtain $G_N(z) = z^n$ and therefore $G(z) = [G_X(z)]^n$, confirming the convolution theorem for generating functions in its most basic form.

some interesting examples. Later on we will obtain the generating function of sector losses in form of an equation that, conditional on the sector's default rate, replicates Equation (4.15).

Let us assume that we have parametrized our portfolio by means of m_S sectors. CreditRisk$^+$ assumes that a gamma-distributed random variable

$$\Lambda^{(s)} \sim \Gamma(\alpha_s, \beta_s) \qquad (s = 1, ..., m_S)$$

is assigned to every sector; see Figure 2.2 for an illustration of gamma densities. The number of defaults in any sector s follows a gamma-mixed Poisson distribution with random intensity $\Lambda^{(s)}$; see also Section 2.2.2. Hereby it is always assumed that the *sector variables* $\Lambda^{(1)}, ..., \Lambda^{(m_S)}$ *are independent.*

For a calibration of $\Lambda^{(s)}$ recall from (2.38) that the first and second moment of $\Lambda^{(s)}$ are

$$\mathbb{E}[\Lambda^{(s)}] = \alpha_s\beta_s, \qquad \mathbb{V}[\Lambda^{(s)}] = \alpha_s\beta_s^2 \ . \tag{4.17}$$

We denote the expectation of the random intensity $\Lambda^{(s)}$ by $\lambda_{(s)}$. The volatility of $\Lambda^{(s)}$ is denoted by $\sigma_{(s)}$. Altogether we have from (4.17)

$$\lambda_{(s)} = \alpha_s\beta_s, \qquad \sigma_{(s)} = \sqrt{\alpha_s\beta_s^2} \ . \tag{4.18}$$

Knowing the values of $\lambda_{(s)}$ and $\sigma_{(s)}$ determines the parameters α_s and β_s of the sector variable $\Lambda^{(s)}$.

For every sector we now follow the approach that has taken us to Equation (4.15). More explicitly, we first find the generating function of the number of defaults in sector s, then obtain the generating function for the distribution of default events among the exposures in sector s, and finally get the portfolio-loss-generating function as the product[5] of the compound sector-generating functions.

4.3.1 Sector Default Distribution

Fix a sector s. The defaults in all sectors are gamma-mixed Poisson. Therefore, conditional on $\Lambda^{(s)} = \theta_s$ the sector's conditional generating function is given by (4.1),

$$G_{L'_{(s)}}|_{\Lambda^{(s)}=\theta_s}(z) = e^{\theta_s(z-1)} \ . \tag{4.19}$$

[5]Recall that we assumed independence of sector variables.

The unconditional generating function also is explicitly known, because fortunately it is a standard fact from elementary statistics that gamma-mixed Poisson variables follow a *negative binomial distribution* (see, e.g., [160], 8.6.1.) The negative binomial distribution usually is taken as a suitable model for a counting variable when it is known that the variance of the counts is larger than the mean. Recalling our discussion in Section 2.2.2 we know that the dispersion of Poisson variables is equal to 1 due to the agreement of mean and variance. Mixing Poisson variables with gamma distributions will always result in a distribution with a conditional dispersion of 1 but *unconditionally overdispersed*.

At this point we make a brief detour in order to provide the reader with some background knowledge on negative binomial distributions. There are two major reasons justifying this. First, the negative binomial distribution is probably not as well known to all readers as the (standard) binomial distribution. Second, the negative binomial distribution is differently defined in different textbooks. We therefore believe that some clarification about our view might help to avoid misunderstandings.

One approach to the negative binomial distribution (see, e.g., [86]) is as follows: Start with a sequence of independent Bernoulli default indicators $X_i \sim B(1; p)$. Let T be the *waiting time* until the first default occurs, $T = \min\{i \in \mathbb{N} \mid X_i = 1\}$. We have

$$\begin{aligned} \mathbb{P}[T = k] \;&=\; \mathbb{P}[T > k-1] - \mathbb{P}[T > k] \\ &=\; (1-p)^{k-1} - (1-p)^k \;=\; p(1-p)^{k-1}. \end{aligned}$$

Therefore, T has a *geometric distribution*. If more generally we ask for the waiting time T_q until the q-th default occurs, then we obtain the *negative binomial distribution with parameters p and q*. The mass function of T_q obviously is given by

$$\mathbb{P}[T_q = k] = \binom{k-1}{q-1} p^q (1-p)^{k-q} \qquad (k \geq q). \tag{4.20}$$

For $q = 1$ the negative binomial and the geometric distributions agree. Moreover,

$$T_q = \sum_{i=1}^{q} T_i' \quad \text{where} \quad T_1' = T_1,\; T_i' = T_i - T_{i-1} \text{ for } i = 2, ..., q, \tag{4.21}$$

where $T_1', ..., T_q'$ are *independent* geometric variables with parameter p. For $i \geq 2$ the variable T_i' is the waiting time until the next default following the $(i-1)$-th default. Because the mean and the variance of a geometric random variable T with parameter p are $\mathbb{E}[T] = 1/p$ respectively $\mathbb{V}[T] = (1-p)/p^2$, (4.21) yields

$$\mathbb{E}[T_q] = \frac{q}{p} \qquad \text{and} \qquad \mathbb{V}[T_q] = \frac{q(1-p)}{p^2}\ . \tag{4.22}$$

The generating function of a geometric variable T with parameter p is

$$G_T(z) = \sum_{k=1}^{\infty} p(1-p)^{k-1} z^k = \frac{pz}{1-(1-p)z} \qquad (|z| < 1/(1-p)).$$

Therefore the convolution theorem for generating functions immediately implies

$$G_{T_q}(z) = \Big(\frac{pz}{1-(1-p)z}\Big)^q \qquad (|z| < 1/(1-p)). \tag{4.23}$$

Application of the relation $\binom{x}{k} = (-1)^k \binom{k-x-1}{k}$ $(x \in \mathbb{R},\ k \in \mathbb{N}_0)$ and the symmetry property $\binom{n}{m} = \binom{n}{n-m}$ yields

$$\mathbb{P}[T_q = k] = \binom{-q}{k-q} p^q (p-1)^{k-q}\ ,$$

which explains the name *negative* binomial distribution.

So far, this is what many authors consider to be a negative binomial distribution. Now, some people consider it a technical disadvantage that the (according to our discussion above very naturally arising) negative binomial distribution ranges in $\{k \in \mathbb{N} \mid k \geq q\}$. For reasons also applying to the situation in CreditRisk$^+$ one would rather like to see T_q ranging in $\mathbb{N}_0$. We can adopt this view by replacing T_q by $\tilde{T}_q = T_q - q$, again applying the symmetry property $\binom{n}{m} = \binom{n}{n-m}$, and substituting $n = k - q$ in Equation (4.20):

$$\begin{aligned} \mathbb{P}[\tilde{T}_q = n] &= \mathbb{P}[T_q = n+q] \\ &= \binom{n+q-1}{n} p^q (1-p)^n \qquad (n \geq 0). \end{aligned} \tag{4.24}$$

The variable $\tilde{T}_q$ obviously describes the number of survivals until the q-th default has occurred.

Because CreditRisk$^+$ requires it in this way, we from now on mean by a negative binomial distribution with parameters q and p the distribution of $\tilde{T}_q$ defined by (4.24). It is well known (see, e.g., [160], 8.6.1) that any $\Gamma(\alpha,\beta)$-mixed Poisson variable L' follows a negative binomial distribution with parameters α and $1/(1+\beta)$. This concludes our detour and we return to the actual topic of this section.

The conditional distribution of the sector defaults is given by (4.19). The mixing variable is $\Lambda^{(s)} \sim \Gamma(\alpha_s, \beta_s)$. According to our discussion above, the unconditional distribution of sector defaults (denoted by $L'_{(s)}$) is negative binomial with parameters α_s and $1/(1+\beta_s)$; in short: $L'_{(s)} \sim NB(\alpha_s, 1/(1+\beta_s))$. We can now easily obtain the unconditional generating function $G_{L'_{(s)}}(z) = G_s(z)$ of the sector defaults by evaluating Formula (4.23) with T_q replaced by $\tilde{T}_q = T_q - q$ and taking the parametrization $q = \alpha_s$ and $p = 1/(1+\beta_s)$ into account. Replacing T_q by $\tilde{T}_q = T_q - q$ changes (4.23) to

$$G_{\tilde{T}_q}(z) \;=\; \Big(\frac{pz}{1-(1-p)z}\Big)^q \,\frac{1}{z^q} \;=\; \Big(\frac{p}{1-(1-p)z}\Big)^q \,.$$

Inserting $q = \alpha_s$ and $p = 1/(1+\beta_s)$ finally yields

$$\begin{aligned} G_{L'_{(s)}}(z) \;&=\; \int_0^\infty [G_{L'_{(s)}}|_{\Lambda^{(s)}=\theta_s}](z)\,\gamma_{\alpha_s,\beta_s}(\theta_s)\,d\theta_s \qquad (4.25)\\ &=\; \left(\frac{1-\frac{\beta_s}{1+\beta_s}}{1-\frac{\beta_s}{1+\beta_s}\,z}\right)^{\alpha_s}, \end{aligned}$$

where $\gamma_{\alpha_s,\beta_s}$ denotes the density of $\Gamma(\alpha_s,\beta_s)$ and α_s, β_s are calibrated to the sector by means of (4.18). We included the integral in the center of (4.25) in order to explicitly mention the link to Section 2.2.1.

Formula (4.25) can be found in the CreditRisk$^+$ Technical Document [35] (A8.3, Equation (55)). The probability mass function of $L'_{(s)}$ follows from (4.24),

$$\mathbb{P}[L'_{(s)} = n] \;=\; \binom{n+\alpha_s-1}{n}\Big(1-\frac{\beta_s}{1+\beta_s}\Big)^{\alpha_s}\Big(\frac{\beta_s}{1+\beta_s}\Big)^n \,. \qquad (4.26)$$

The first and second moments of the distribution of defaults in sector s directly follow from the general results on Poisson mixtures; see (2.15) in Section 2.2. They depend on the mixture distribution only and are given by

$$\mathbb{E}[L'_{(s)}] \;=\; \mathbb{E}[\Lambda^{(s)}] \;=\; \alpha_s\beta_s \quad \text{and} \tag{4.27}$$

$$\mathbb{V}[L'_{(s)}] \;=\; \mathbb{V}[\Lambda^{(s)}] + \mathbb{E}[\Lambda^{(s)}] \;=\; \alpha_s\beta_s(1+\beta_s)$$

(see also (4.17)), hereby confirming our previous remark that the unconditional distribution of sector defaults is *overdispersed*. In (4.28) we see that we always have $\beta_s \geq 0$ and that $\beta_s > 0$ if and only if the volatility of the sector's default intensity does not vanish to zero. Figure 2.7 in Section 2.5.2 graphically illustrates (4.26).

Alternatively, the first and second moments of $L'_{(s)}$ could have been calculated by application of (4.22), taking the shift $T_q \to T_q - q$ and the parametrization of q and p into account. It is a straightforward calculation to show that the result of such a calculation agrees with the findings in (4.27).

4.3.2 Sector Compound Distribution

As a preparation for the compound approach on the sector level we begin this section with a remark regarding the calibration of α_s and β_s. Solving Equations (4.18) for α_s and β_s gives us the mixing parameters in terms of the sector parameters $\lambda_{(s)}$ and $\sigma_{(s)}$:

$$\alpha_s \;=\; \frac{\lambda_{(s)}^2}{\sigma_{(s)}^2} \qquad \text{and} \qquad \beta_s \;=\; \frac{\sigma_{(s)}^2}{\lambda_{(s)}}\;. \tag{4.28}$$

Most often $\lambda_{(s)}$ and $\sigma_{(s)}$ will be calibrated from obligor data. As already mentioned before, in the sector model every obligor i admits a breakdown into sector weights w_{is} such that

$$\sum_{s=1}^{m_S} w_{is} = 1 \qquad (w_{is} \geq 0;\; i = 1, ..., m).$$

Moreover, any obligor i admits a random default intensity defined by

$$\Lambda_i \;=\; \sum_{s=1}^{m_S} w_{is}\lambda_i \frac{\Lambda^{(s)}}{\lambda_{(s)}} \qquad (i = 1, ..., m). \tag{4.29}$$

The expected intensity then obviously equals λ_i, which is consistent with the case of independent obligors where λ_i denoted the nonrandom default intensity of obligor i. The expected intensities λ_i can be calibrated to one-year default probabilities by application of Formula (4.3). Due to the additivity of expectations it is then very natural to define the expected sector intensity $\lambda_{(s)}$ by

$$\lambda_{(s)} \;=\; \sum_{i=1}^{m} w_{is}\lambda_i \;=\; \sum_{j=1}^{m_E}\sum_{i\in[j]} w_{is}\lambda_i \;, \tag{4.30}$$

where the right side expresses the grouping into exposure bands. Note that an exposure band j takes part in sector s if and only if there exists some obligor $i \in [j]$ such that $w_{is} > 0$. The sector volatility $\sigma_{(s)}$ can be either calibrated from empirical data related to the meaning of sector s or calculated from the single obligor's default intensity volatilities. An example for the first case would be an industry sector, where the volatility of a historical time series of insolvency quotes for that particular industry could be taken as a proxy for $\sigma_{(s)}$. An example for the latter case would be a portfolio where, in addition to the default rate, the default rate volatility is also known for every obligor. Such estimates usually depend on the creditworthiness of obligors. For example, denoting the default rate volatility for obligor i by σ_i and assuming that the sectors perform a *partition* of the portfolio's set of obligors (more explicitly: $w_{is} = 1$ for a unique sector $s = s(i)$ for every obligor i) one obtains from (4.29) and (4.30) for every sector s the following identity:

$$\sum_{i:w_{is}=1} \sigma_i \;=\; \sum_{i:w_{is}=1} \sqrt{\mathbb{V}\Big[w_{is}\lambda_i \frac{\Lambda^{(s)}}{\lambda_{(s)}}\Big]} \;=\; \sum_{i:w_{is}=1} w_{is}\lambda_i \,\frac{\sigma_{(s)}}{\lambda_{(s)}} \;=\; \sigma_{(s)} \;,$$

where the sum takes all obligors i in sector s into account. So in this particular example, the volatility of the sector default intensity can be directly estimated from the volatility of the default intensities of obligors collected into that sector. The calibration of a sector variable $\Lambda^{(s)}$ can then be finalized by applying (4.28).

For the general case where obligors are allowed to be influenced by more than one sector, the CreditRisk$^+$ Technical Document [35] (A12.2) suggests an analogous approach by estimating the sector volatility $\sigma_{(s)}$ by the weighted contribution of the default rate volatilities of obligors influenced by the sector, namely $\hat{\sigma}_{(s)} = \sum_{i=1}^{m} w_{is}\sigma_i$. Again note that only obligors i with $w_{is} > 0$ contribute to sector s.

Based on our calculations above we can now just follow the lines of (4.14), (4.15), and (4.16). Analogously to (4.16) we first of all define a random variable N_s by

$$\mathbb{P}[N_s = \nu_{[j]}] \;=\; \frac{1}{\lambda_{(s)}} \sum_{i\in[j]} w_{is}\lambda_i \qquad (j = 1, ..., m_E). \tag{4.31}$$

Equation (4.30) shows that (4.16) really defines a probability distribution on the set of exposures $\{\nu_{[1]}, ..., \nu_{[m_E]}\}$. The generating function of N_s is given by a polynomial analogous to (4.14),

$$G_{N_s}(z) \;=\; \sum_{j=1}^{m_E} \Big(\frac{1}{\lambda_{(s)}} \sum_{i\in[j]} w_{is}\lambda_i\Big) z^{\nu_{[j]}} \; . \tag{4.32}$$

Instead of $G_{L'}$ as in (4.14) we now use the generating function $G_{L'_{(s)}}$ of the sector defaults as described in (4.25). Because the generating function G_{N_s} does not depend on realizations of the random intensity $\Lambda^{(s)}$, the arguments leading to Formula (4.25) are not affected when replacing the variable z in (4.25) by $G_{N_s}(z)$. We therefore obtain the compound generating function of the distribution of losses $\tilde{L}'_{(s)}$ in sector s by writing

$$G_{\tilde{L}'_{(s)}}(z) \;=\; G_{L'_{(s)}} \circ G_{N_s}(z) \;=\; \left(\frac{1 - \frac{\beta_s}{1+\beta_s}}{1 - \frac{\beta_s}{1+\beta_s}\, G_{N_s}(z)}\right)^{\alpha_s} \tag{4.33}$$

$$= \left(\frac{1 - \frac{\beta_s}{1+\beta_s}}{1 - \frac{\beta_s}{1+\beta_s}\, \frac{1}{\lambda_{(s)}} \sum_{j=1}^{m_E} \sum_{i\in[j]} w_{is}\lambda_i z^{\nu_{[j]}}}\right)^{\alpha_s} .$$

Note that this is the same two-step randomness we previously derived in the case of independent obligors leading to Formula (4.15). The bridge between the independent case and (4.33) is just a simple conditioning argument. Conditional on a sector intensity realization θ_s, the conditional compound probability-generating function analogous to (4.15) is given by

$$[G_{\tilde{L}'_{(s)}}|_{\Lambda^{(s)}=\theta_s}](z) \;=\; e^{\theta_s (G_{N_s}(z)-1)} \; . \tag{4.34}$$

Integration of the right side w.r.t. the mixing gamma distribution gives (4.33), taking into account that $G_{N_s}(z)$ does not depend on the integration variable θ_s. Therefore the integral can be calculated by means of exactly the same argument as in (4.25).

4.3.3 Sector Convolution

The portfolio loss $L' = L'_1 + \cdots + L'_m$ is a mixed Poisson variable with random intensity $\Lambda = \Lambda_1 + \cdots + \Lambda_m$. Grouped into sectors, the intensity Λ of L' can also be written as the sum of sector intensities,

$$\Lambda \;=\; \Lambda^{(1)} + \cdots + \Lambda^{(m_S)}.$$

This follows from Formulas (4.29) and (4.30). Because sectors are assumed to be independent, the distribution of defaults in the portfolio is the convolution of the sector's default distributions. Therefore, due to (4.25) the generating function of L' is given by

$$G_{L'}(z) \;=\; \prod_{s=1}^{m_S} \left(\frac{1 - \frac{\beta_s}{1+\beta_s}}{1 - \frac{\beta_s}{1+\beta_s}\, z} \right)^{\alpha_s} . \tag{4.35}$$

The generating function of the portfolio losses is determined by the convolution of compound sector distributions as elaborated in (4.33),

$$\begin{aligned} G_{\tilde{L}'}(z) &= \prod_{s=1}^{m_S} G_{\tilde{L}'_{(s)}}(z) \\ &= \prod_{s=1}^{m_S} \left(\frac{1 - \frac{\beta_s}{1+\beta_s}}{1 - \frac{\beta_s}{1+\beta_s} \frac{1}{\lambda_{(s)}} \sum_{j=1}^{m_E} \sum_{i \in [j]} w_{is} \lambda_i z^{\nu_{[j]}}} \right)^{\alpha_s} . \end{aligned} \tag{4.36}$$

So we see that in CreditRisk$^+$ the portfolio loss distribution can be described in an analytical manner by means of a closed-form generating function. Remarkable is the fact that this nice property even holds in the most general case of a sector model where complex dependence structures are allowed. In the general sector model of CreditRisk$^+$ leading to Formula (4.36), obligors i_1 and i_2 will be correlated if and only if there exists at least one sector s such that $w_{i_1 s} > 0$ and $w_{i_2 s} > 0$.

4.3.4 Calculating the Loss Distribution

Because probability distributions and generating functions are uniquely associated with each other, Formula (4.36) allows for a direct computation of loss distributions. For completeness, we will present here the original derivation of the loss distribution as shown in the CreditRisk$^+$ Technical Document [35]. The calculation scheme for the loss amount distribution of portfolios is again a variation of the well known Panjer recurrence relation, similar to the one presented in Section 2.8.3.

From Section 2.8 we know that the generating function of a discrete probability distribution is

$$G_L(z) = \mathbb{E}[z^L] = \sum_{n=0}^{\infty} \mathbb{P}[L = n] z^n = \sum_{n=0}^{\infty} A_n z^n. \tag{4.37}$$

Suppose now that the generating function $G(z)$ also satisfies the differential equation

$$\frac{d}{dz}\left(\log(G(z))\right) = \frac{1}{G(z)} \frac{G(z)}{dz} = \frac{A(z)}{B(z)}, \tag{4.38}$$

where $A(z) = a_0 + \cdots + a_r z^r$ and $B(z) = b_o + \cdots + b_s z^s$ are polynomials. Then, for the coefficients A_n of the powerseries expansion (4.37) holds the following recurrence relation:

$$A_{n+1} = \frac{1}{b_0(n+1)} \left(\sum_{i=0}^{\min(r,n)} a_i A_{n-i} - \sum_{j=0}^{\min(s-1,n-1)} b_{j+1} A_{n-j}(n-j) \right). \tag{4.39}$$

Proof. Rearranging Equation (4.38) as

$$B(z)G'(z) = A(z)G(z)$$

and differentiating G leads to

$$\left(\sum_{i=0}^{s} b_i z^i\right)\left(\sum_{n=0}^{\infty}(n+1)A_{n+1}z^n\right) = \left(\sum_{i=0}^{r} a_i z^i\right)\left(\sum_{n=0}^{\infty} A_n z^n\right)$$

Equating coefficients of z^n, $n \geq 0$ on both sides yields

$$\sum_{j=0}^{\min(s,n)} b_j A_{n+1-j}(n+1-j) = \sum_{i=0}^{\min(r,n)} a_i A_{n-i}$$

and after rearranging, we obtain

$$b_0(n+1)A_{n+1} = \sum_{i=0}^{\min(r,n)} a_i A_{n-i} - \sum_{j=0}^{\min(s-1,n-1)} b_{j+1} A_{n-j}(n-j).$$

It remains to express the logarithmic derivative of $G_{\tilde{L}'}(z)$ from Equation (4.36) as a rational function. To this end we write

$$\frac{d}{dz}\big(\log(G_{\tilde{L}'}(z))\big) = \sum_{s=1}^{m_S} \frac{G'_{\tilde{L}'_{(s)}}(z)}{G_{\tilde{L}'_{(s)}}(z)}$$

$$= \sum_{s=1}^{m_S} \frac{\frac{\beta_s\alpha_s}{1+\beta_s}\,\frac{1}{\lambda_{(s)}}\sum_{j=1}^{m_E}\sum_{i\in[j]} w_{is}\lambda_i z^{\nu_{[j]}}}{1-\frac{\beta_s}{1+\beta_s}\,\frac{1}{\lambda_{(s)}}\sum_{j=1}^{m_E}\sum_{i\in[j]} w_{is}\lambda_i z^{\nu_{[j]}}}, \tag{4.40}$$

where the summation has to be performed explicitly by adding up the rational summands for each sector. Once the coefficients of the polynomials $A(z)$, $B(z)$ are obtained, we can use the recurrence relation (4.39) to calculate the probability distribution.

Unfortunately, for practical cases the algorithm turns out to be numerically fragile. Its numerical instability arises from an accumulation of numerical round-off errors due to the summation of numbers of similar magnitude but opposite sign.

Further Reading

There would be more to say about CreditRisk$^+$ but due to the introductory character of this book we will not go any further. The Technical Document [35] contains some more detailed information on the management of credit portfolios, the calibration of the model, and the technical implementation. Additionally, we would like to point to the collection of papers on CreditRisk$^+$ in [87], where, among others, alternatives to the recurrence relation for generating the loss distribution are discussed. Both [35] and [87] address, for example, the question of risk contributions (see also Chapter 5) and the role of correlations in CreditRisk$^+$. Risk contributions in CreditRisk$^+$ are also extensively studied in Tasche [175]. In [35], A12.3, the introduction of a sector for incorporating *specific risk* is discussed. As a last remark we should mention that because the sector distributions are Poisson mixtures, the general results from Section 2.2 can also be applied.

Chapter 5

Risk Measures and Capital Allocation

The definition of economic capital as introduced in Chapter 1 appears fully satisfactory at first glance. Starting with the path-breaking paper by Artzner et al. [11], several studies revealed a number of methodological weaknesses in the VaR concept by drawing up a catalog of mathematical and material attributes that a risk measure should fulfill, and proving that the VaR concept only partly fulfills them. Risk measures that satisfy these axioms are called *coherent*. Before we describe their basic features in the next section, we briefly reiterate the main notations (cf. Chapters 1 and 2):

The portfolio loss variable (compare Equation (2.51)) is given by

$$L = \sum_{i=1}^{m} w_i \eta_i L_i,$$

where

$$w_i \;=\; \frac{\mathrm{EAD}_i}{\sum_{i=1}^{m} \mathrm{EAD}_i}$$

are the exposure weights, so all portfolio quantities are calculated in percent of the portfolio's total exposure. The severity is abbreviated to $\eta_i = \mathrm{LGD}_i$; $L_i = \mathbf{1}_{D_i}$ are Bernoulli random variables with default events D_i. Accordingly, the default probability (e.g., for a one-year horizon) for obligor i is given by

$$\mathrm{PD}_i = \mathbb{P}(D_i) \ .$$

Default correlations are denoted as

$$\rho_{ij} = \mathrm{Corr}[L_i, L_j] = \mathrm{Corr}[\mathbf{1}_{D_i}, \mathbf{1}_{D_j}] \ .$$

In this chapter we assume a deterministic severity, η_i, but, if severities are assumed to be independent of defaults, then it is straightforward to extend the following to random severities. For notational convenience we denote by

$$\hat{L}_i := \eta_i L_i$$

the loss variable of some obligor i.

5.1 Coherent Risk Measures and Expected Shortfall

Denote by $L^\infty = L^\infty(\Omega, \mathcal{F}, \mathbb{P})$ the space of bounded real random variables, defined on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$. The definition of a *coherent risk measure* as suggested by Artzner et al. [10, 11] can then be stated in the following way:

5.1.1 Definition *A mapping $\rho : L^\infty(\Omega, \mathcal{F}, \mathbb{P}) \to \mathbb{R}$ is called a coherent risk measure if the following properties hold:*

(i) **Subadditivity** $\forall\, X, Y \in L^\infty$: $\rho(X + Y) \leq \rho(X) + \rho(Y)$

(ii) **Monotonicity** $\forall\, X, Y \in L^\infty$ *with* $X \leq Y$ *a.s.*: $\rho(X) \leq \rho(Y)$

(iii) **Positive homogeneity** $\forall\, \lambda > 0, \forall\, X \in L^\infty$: $\rho(\lambda X) = \lambda\rho(X)$

(iv) **Translation invariance** $\forall\, x \in \mathbb{R}, \forall\, X \in L^\infty$:
$\rho(X + x) = \rho(X) + x$.

Note that the definition here slightly differs from the original set of axioms as they were introduced by Artzner et al. [10, 11]. Our definition equals the one given by Frey and McNeil in [69], because we want to think about X in terms of a portfolio loss and about $\rho(X)$ as the amount of capital required as a cushion against the loss X, according to the credit management policy of the bank. In the original approach by Artzner et al., X was interpreted as the future value of the considered portfolio. Let us now briefly explain the four axioms in an intuitive manner:

Subadditivity: Axiom (i) reflects the fact that due to diversification effects the risk inherent in the union of two portfolios should be less than the sum of risks of the two portfolios considered separately. We will later see that quantiles are not subadditive in general, such that the economic capital (EC) as introduced in Chapter 1 turns out to be not coherent.

Monotonicity: Let us say we are considering portfolios A and B with losses X_A and X_B. If almost surely the losses of portfolio A are lower than the losses of portfolio B, i.e., $X_A \leq X_B$ a.s., then the required risk capital $\rho(X_A)$ for portfolio A should be less than the required risk capital $\rho(X_B)$ of portfolio B. Seen from this perspective, monotonicity is a very natural property.

Homogeneity: Axiom (iii) can best be illustrated by means of the following example. Consider a credit portfolio with loss X and scale all exposures by a factor λ. Then, of course, the loss X changes to a scaled loss λX. Accordingly, the originally required risk capital $\rho(X)$ will also change to $\lambda\rho(X)$.

Translation invariance: If x is some capital which will be lost/gained on a portfolio with certainty at the considered horizon, then the risk capital required for covering losses in this portfolio can be increased/reduced accordingly. Translation invariance implies the natural property $\rho(X - \rho(X)) = 0$ for every loss $X \in L^\infty$.

From convex analysis follows that a sub-additive positive homogeneous function ρ can be point-wise written as the maximal value of all linear function which are below ρ, i.e.

$$\rho(X) = \max\{l(X) \mid l < \rho,\ l \text{ linear function}\}$$

Conceptually this is similar to the gradient of the function ρ evaluated at the point X as the best linear approximation of ρ which coincides with ρ at X. We will later see that this gives rise to a sensible capital allocation.

Based on this thread of thoughts the proposition below provides another interpretation of coherency and can be found in [11].

5.1.2 Proposition *Let the underlying probability space Ω be finite. Then a risk measure ρ is coherent if and only if there exists a family $\mathcal{P}$ of probability measures such that*

$$\rho \;=\; \rho_{\mathcal{P}} \;,$$

where $\rho_{\mathcal{P}}$ is defined by

$$\rho_{\mathcal{P}}(X) \;=\; \sup_{Q \in \mathcal{P}} \mathbb{E}_Q[X] \qquad \text{for all } X \in L^{\infty} \;.$$

The probability measures in $\mathcal{P}$ are called generalized scenarios. [1]

The challenge underlying Proposition 5.1.2 is to find a suitable set $\mathcal{P}$ of probability distributions matching a given coherent risk measure ρ such that $\rho = \rho_{\mathcal{P}}$. We will recall the set of scenarios in the example on Expected Shortfall.

Typical risk measures discussed by Artzner et al. are the *value-at-risk* and *expected shortfall capital*, which will be briefly discussed in the sequel.

Value-at-Risk Value-at-risk (VaR) has already been mentioned in Section 1.2.1 as a synonymous name for EC. Here, VaR will be defined for a probability measure $\mathbb{P}$ and some confidence level α as the α-quantile of a loss random variable X,

$$\mathrm{VaR}_{\alpha}(X) = \inf\{x \geq 0 \mid \mathbb{P}[X \leq x] \geq \alpha\} \;.$$

VaR as a risk measure defined on L^{∞} is

- translation invariant, because shifting a loss distribution by a fixed amount will shift the quantile accordingly,
- positively homogeneous, because scaling a loss variable will scale the quantile accordingly,
- monotone, because quantiles preserve monotonicity, but
- not subadditive, as the following example will show.

Because VaR is not subadditive, it is *not coherent*. Now let us give a simple example showing that VaR is not subadditive.

Consider two independent loans, represented by two loss indicator variables $\mathbf{1}_{D_A}, \mathbf{1}_{D_B} \sim B(1;p)$ with, e.g., $0.006 \leq p < 0.01$. Assume LGDs equal to 100% and exposures equal to 1. Define two portfolios A and B, each consisting of one single of the above introduced loans. Then, for the portfolio losses $X_A = \mathbf{1}_{D_A}$ and $X_B = \mathbf{1}_{D_B}$ we have

$$\mathrm{VaR}_{99\%}(X_A) \;=\; \mathrm{VaR}_{99\%}(X_B) \;=\; 0\;,$$

Now consider a portfolio C defined as the union of portfolios A and B, and denote by $X_C = X_A + X_B$ the corresponding portfolio loss. Then

$$\mathbb{P}[X_C = 0] \;=\; (1-p)^2 \;<\; 99\%\;.$$

Therefore, $\mathrm{VaR}_{99\%}(X_C) > 0$, so that

$$\mathrm{VaR}_{99\%}(X_A + X_B) \;>\; \mathrm{VaR}_{99\%}(X_A) + \mathrm{VaR}_{99\%}(X_B)\;.$$

This shows that in general VaR is not a coherent risk measure.

Tail Conditional Expectation The *tail conditional expectation*, or *expected shortfall*, w.r.t. a confidence level α is defined as

$$\mathrm{TCE}_\alpha(X) = \mathbb{E}[X \mid X > \mathrm{VaR}_\alpha(X)]\;.$$

In the literature one can find several slightly different versions of TCE definitions. Tasche [176] showed that Tail Conditional Expectation to a great extent enjoys the coherence properties; for further reading see also [11, 105]. For example, Tail Conditional Expectation is coherent when restricted to loss variables X with continuous distribution function. It then coincides with *expected shortfall* measure, [2],

$$\mathrm{ES}_\alpha(X) = \frac{1}{1-\alpha}\Big(\mathbb{E}\left[X\ \mathbf{1}_{\{X > \mathrm{VaR}_\alpha(X)\}}\right] \\ + \mathrm{VaR}_\alpha(X)\left(\mathbb{P}\left[X \leq \mathrm{VaR}_\alpha(X)\right] - \alpha\right)\Big). \quad (5.1)$$

The second term in the above equation takes care of a possible positive probability mass at the quantile itself. If the distribution of X is continuous, the second term in the definition of ES_α vanishes, and the expected shortfall coincides with TCE in this case. It is shown in [2] that expected shortfall ES (5.1) is a coherent risk measure.

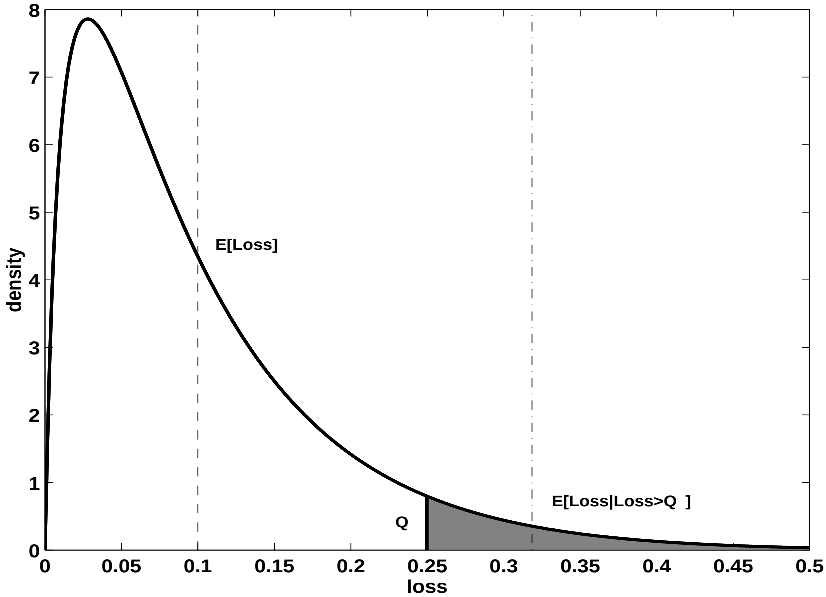


FIGURE 5.1: Tail conditional expectation $\mathbb{E}[X \mid X > \mathrm{VaR}_\alpha(X)]$.

5.1.1 Expected Shortfall

The most useful representation of ES is as a mixture of VaR measures.

$$\mathrm{ES}_\alpha(X) \;=\; \frac{1}{1-\alpha}\int_\alpha^1 \mathrm{VaR}_\beta(X)d\beta \qquad (5.2)$$

Instead of fixing a quantile at a particular confidence level α, the Expected Shortfall averages VaR across the entire tail specified by α. The mixing or weight function has a very simple form, it is zero from the 0-quantile up to the (1-α)-quantile and has equal positive weight from $1-\alpha$ to 1. This weight function can be viewed as a risk aversion function and gives rise to a much wider class of coherent risk measures, the so-called spectral risk measures which we will introduce in the next section.

But let us first give a simple interpretation of ES. Figure 5.1 illustrates the definition of expected shortfall capital. From an insurance point of view, expected shortfall is a very reasonable measure: Defining by $c = \mathrm{VaR}_\alpha(X)$ a *critical loss threshold* corresponding to some confidence level α, expected shortfall capital provides a cushion against the mean value of losses exceeding the critical threshold c. In other words, ES focusses on the expected loss in the tail, starting at c, of the portfolio's loss distribution. The critical threshold c, driven by the confidence level α, has to be fixed by the senior management of the bank and is part of the bank's credit management policy.

Economic capital, based on shortfall risk, can be defined as the mean loss above a threshold c minus the expected loss:

$$\mathrm{EC}_{\mathrm{TCE}}(c) = \mathbb{E}[X \mid X > c] - \mathbb{E}[X].$$

This calculation method for risk capital also includes events above the critical loss threshold c, e.g., $c = \mathrm{VaR}_\alpha$, and answers the question "how bad is bad" from a mean value's perspective. If $c = \mathrm{VaR}_\alpha(X)$, we write

$$\mathrm{EC}_{\mathrm{TCE}_\alpha} \;=\; \mathrm{EC}_{\mathrm{TCE}}(\mathrm{VaR}_\alpha(X))$$

in the sequel in order to keep the notation simple.

Quantiles The following table shows a comparison of expected shortfall EC and VaR-EC over the 99%-quantile for different distributions. For example, one can see that if X had a normal distribution, the expected shortfall EC would not much differ from the VaR-EC. In contrast, for a t-distribution and the loss distribution $F_{p,\varrho}$ defined by a uniform (limit) portfolio as introduced right before Proposition 2.5.7, the difference between the ECs is quite significant.

	t(3)	N(0,1.73)	LN(0,1)	N(1.64,2.16)	Weil(1,1)	N(1,1)
std	1.73	1.73	2.16	2.16	1	1
$\mathrm{EC}_{\mathrm{VaR}}(0.99)$	4.54	4.02	8.56	5.02	3.6	2.32
$\mathrm{EC}_{\mathrm{TCE}}(0.99)$	6.99	4.61	13.57	5.76	4.6	2.66
rel.diff. (%)	54	15	58	15	27	14

	$F_{0.003,0.12}$
std	0.0039
$\mathrm{EC}_{\mathrm{VaR}}(0.99)$	0.0162
$\mathrm{EC}_{\mathrm{TCE}}(0.99)$	0.0237
rel.diff. (%)	46

This table highlights the sensitivity of the determination of economic capital w.r.t. its definition (VaR-based or shortfall-based) and the choice of the underlying probability measure. Here, the probability measures are given by the following distributions:

- the t-distribution $t(3)$ with 3 degrees of freedom,
- three normal distributions with different parameter pairs,
- the $(0,1)$-log-normal distribution,
- the $(1,1)$-Weibull distribution, and
- the uniform portfolio distribution $F_{p,\varrho}$ with uniform default probability $p = 30$bps and uniform asset correlation $\varrho = 12\%$.

In the table, the first row contains the respective standard deviations, the second row shows the EC based on VaR at the 99%-confidence level, and the third row shows the EC based on expected shortfall with threshold $c = \text{VaR}_{99\%}$. The fourth row shows the relative difference between the two EC calculations.

5.1.2 Spectral Risk Measures

As already indicated in the Section 5.1.1, one might associate a risk aversion function with ES, which we now introduce formally [1, 2]. The risk aversion weight function associated with ES_α turns out to be

$$w_{\text{ES}_\alpha}(u) = (1-\alpha)^{-1}\mathbf{1}_{\{u>\alpha\}}. \tag{5.3}$$

From a risk management point of view there might be many other weights given to some confidence levels u. If the weight function is increasing, which is reasonable since higher losses should have larger risk aversion weight, then we arrive at spectral risk measures.

5.1.3 Definition *Let w be an increasing function on $[0,1]$ such that $\int_0^1 w(u)du = 1$, then the map ρ_w defined by*

$$\rho_w(X) = \int_0^1 w(u)\, VaR_u(X) du$$

is called a spectral risk measure with weight function w.

The name "spectral risk measure" comes from the representation

$$\begin{aligned}
\rho_w(X) &= \int_0^1 \mathrm{ES}_\alpha(1-\alpha)\mu_w(d\alpha) \qquad (5.4)\\
&= \int_0^1 \int_\alpha^1 \mathrm{VaR}_u(X)du\, \mu_w(d\alpha)\\
&= \int_0^1 \mathrm{VaR}_u(X) \int_0^u \mu_w(d\alpha)\, du\\
&= \int_0^1 \mathrm{VaR}_u(X) w(u) du
\end{aligned}$$

with the spectral measure $\mu((0,b]) = w(b)$.

There might be several pragmatic approaches to calibrate the risk aversion function:

1. As a first step in the application of spectral risk measures one might think to give to different loss probability levels different weight. This is a straightforward extension of expected shortfall. One might view Expected Shortfall at the 99%-level view as a risk aversion which ignores losses below the 99%-quantile and all losses above the 99%-quantile have the same influence. From an investor's point of view this means that only senior debts are cushioned by risk capital. One might, on the other hand, also be aware of losses which occur more frequently, but of course with a lower aversion than those appearing rarely.

   As a concrete example one might set that losses up to the 50% confidence level should have zero weights, losses between 50% and 99% should have a weight w_0 and losses above the 99%-quantile should have a weight of $k_1 w_0$ and above the 99.9% quantile it should have a weight of $k_2 w_0$. The first tranche from 50% to 99% correspond to an investor in junior debt, and the tranche from 99% to 99.9% to a senior investor and above the 99.9% a super senior investor or the regulators are concerned. This gives a step function for w:

   $$w(u) = w_0 \mathbf{1}_{\{0.99>u>0.5\}} + k_1 w_0 \mathbf{1}_{\{0.999>u>0.99\}} + k_2 w_0 \mathbf{1}_{\{1>u>0.999\}}$$

   The parameter w_0 should be chosen such that the integral over w is still 1.

2. A more continuous form of this is an exponential function starting at a point u_0 between 0 and 1 and then increasing up to 1

$$w(u) = \mathbf{1}_{\{u>u_0\}} \exp(\kappa u)$$

with some constant κ.

3. One can also try to derive the risk aversion function from the traded tranches, e.g. from the iTRAXX or CDX credit indices. The value of a tranche can be viewed as the expectation of the loss in Tranche T under the valuation measure E_Q,

$$E_Q[L_T].$$

As explained below the risk aversion can be converted to a density Q_w. All tranches of the credit index are valued under the same measure $Q = Q_w$ with $w = w_{Index}$. Hence the prices reveal information about the traded risk aversion on the Index portfolio. A first attempt to back out a function w_{Index} can be found in [138].

5.1.3 Density of a Risk Measure

In most cases it can actually be proved that the supremum in the representation of a coherent risk measure in terms of generalized scenarios $\mathcal{P}$ (Prop. 5.1.2) is attained. One has the general representation from Proposition 5.1.2

$$\begin{aligned} \rho_{\mathcal{P}}(X) &= \sup_{Q\in\mathcal{P}} \mathbb{E}_Q[X] \\ &= \mathbb{E}[\lambda_{X,\mathcal{P}} X], \end{aligned} \tag{5.5}$$

since the supremum is a maximum and whereby $\lambda_{X,\mathcal{P}}$ denotes the density of the maximal generalized scenario. For an explicit version see theorem 5.1.4 Theorem below.

This is a very useful result if it comes to allocation. Notice however, that the density $\lambda_{X,\mathcal{P}} = \lambda_{\rho,X}$ depends on the risk measure ρ and the loss variable X.

For expected shortfall the dependency of the density λ as a function of X is explicitely given by the representation $\mathrm{ES}_\alpha(X) = \mathbb{E}[X g_\alpha(X)]]$

with the function, cf. (5.1),

$$g_\alpha(X) := (1-\alpha)^{-1}(\mathbf{1}_{\{X>VaR_\alpha(X)\}} + \beta_X \mathbf{1}_{\{X=VaR_\alpha(X)\}}), \tag{5.6}$$

where β_X is

$$\beta_X := \frac{\mathbb{P}(X \leq VaR_\alpha(X)) - \alpha}{\mathbb{P}(X = VaR_\alpha(X))} \quad \text{if} \quad \mathbb{P}(X = VaR_\alpha(X)) > 0.$$

The density of the associated maximal scenario, as a measure in $\mathcal{P}$, is therefore equal to the random variable $g_\alpha(X)$. Note that $ES_\alpha(Y) = \mathbb{E}(Y \cdot g(Y))$ and $ES_\alpha(X) \geq \mathbb{E}(X \cdot g(Y))$ for every pair of random variables X, Y.

Qualitatively, this is the density which assigns equal mass to all events which lead to a loss above the quantile and zero mass to all events below the quantile. The probability mass of all events above the α-quantile equals by definition $1-\alpha$. The set of generalized scenarios consists of all measures whose density is bounded by $\frac{1}{1-\alpha}$, cf. e.g. [1]. This includes the elementary conditional expectations defined $\mathbb{E}[\cdot|A]$ with $\mathbb{P}[A] \leq 1-\alpha$, i.e., the expectation conditioned on a set of probability less than $1-\alpha$.

The density of spectral risk measures can be obtained by their representation as a convex combination of Expected Shortfall Measures as in the following theorem (cf. [153])

5.1.4 Theorem *The density of the scenario associated with the spectral risk measure with risk aversion function w equals*

$$\lambda_w := g_w(X) := \int_0^1 g_\alpha(X)(1-\alpha)\mu(d\alpha). \tag{5.7}$$

Here, $g_\alpha(\cdot)$ is defined in formula (5.6). In particular

$$\rho_w(X) = \mathbb{E}[X\lambda_w]. \tag{5.8}$$

Proof: We have

$$\begin{aligned}
\rho_w(X) &= \int_0^1 \mathrm{ES}_\alpha(X)(1-\alpha)\mu(d\alpha) \\
&= \int_0^1 \mathbb{E}[Xg_\alpha(X)](1-\alpha)\mu(d\alpha) \\
&= \int_0^1 \max\{\mathbb{E}[Xg_\alpha(Y)]|Y \in L^\infty\}\}(1-\alpha)\mu(d\alpha) \\
&\geq \max\{\mathbb{E}[X\int_0^1 g_\alpha(Y)(1-\alpha)\mu(d\alpha)]|Y \in L^\infty\} \\
&= \max\{\mathbb{E}[Xg_w(Y)]|Y \in L^\infty\} \\
&\geq \mathbb{E}[Xg_w(X)]
\end{aligned}$$

From the second line we deduce by exchanging integration and expectation that $\rho_w(X) = \mathbb{E}[Xg_w(X)]$, hence

$$\rho_w(X) = \max\{\mathbb{E}[Xg_w(Y)]|\forall Y \in L^\infty\} = \mathbb{E}[Xg_w(X)]. \quad \square$$

The last equation in the proof also shows how to identify the set of generalized scenarios $\mathcal{P}_\rho$ associated with a spectral risk measure $\rho = \rho_w$ with weight function w, namely

$$\mathcal{P}_{\rho_w} = \{Q|\frac{dQ}{dP} = g_w(Y), Y \in L^\infty\}$$

and

$$\sup\{\mathbb{E}[Xg_w(Y)]|\forall Y \in L_\infty\} = \mathbb{E}[Xg_w(X)].$$

As mentioned in Section 1.1.3, this coincides for Expected Shortfall with

$$\sup\left\{\mathbb{E}_Q[X]|\frac{dQ}{dP} \leq \frac{1}{1-\alpha}\right\},$$

or intuitively with

$$\sup\{\mathbb{E}_P[X|A]|P[A] \leq \alpha\}.$$

5.2 Contributory Capital

If the economic capital EC of the bank is determined, one of the main tasks of risk management is to calculate *contributory economic* *capital* for each business division of the bank. Such a contributory EC could be interpreted as the *marginal capital* a single transaction respectively business unit adds or contributes to the overall required capital. Tasche [174] showed that there is only one definition for *risk contributions* that is suitable for *performance measurement*, namely the *derivative of the underlying risk measure in direction of the asset weight of the considered entity*, (i.e., the gradient or Euler principle). Additionally, Denault [43] showed by arguments borrowed from game theory that in the case of a so-called *1-homogeneous risk measure* its gradient is the only "allocation principle" that satisfies some "coherency" conditions. A third approach, in the spirit of the axioms on risk measures (cf. Def.(5.1.1)), was introduced by Kalkbrener [109], see also [110] and the survey paper [53]. There, it is shown that capital allocation can be viewed as a generalized (sub-)gradient of the risk measure as a convex and positive homogenous function and that the Euler principle is an immediate consequence of the proposed axioms.

5.2.1 Axiomatic Approach to Capital Allocation

A prerequisite of portfolio optimization is the ability to measure the total portfolio risk and attribute this risk to the portfolio constituents on possibly several hierarchical levels. From an abstract point of view assume that a risk measure ρ has been fixed and that X is a portfolio which consists of subportfolios $X_1, \ldots, X_n$ with $X = X_1 + \ldots X_n$. The objective is to distribute the risk capital $k := \rho(X)$ of the portfolio X to its subportfolios, i.e. to compute risk contributions $k_1, \ldots, k_n$ of $X_1, \ldots, X_n$ such that $k = k_1 + \cdots + k_m$.

The basis of the axiomatic approach is the assumption that the risk allocated to subportfolio X_i should only depend on X_i and X but not on the remainder $X - X_i$. Hence, capital allocation can be viewed as a real valued function on the space of random variables identified with the portfolio loss variables, $\Lambda : V \times V \to \mathbb{R}$ and V denoting a subspace of L^∞. $\Lambda(X, Y)$ is then interpreted as the capital allocated to X considered as a subportfolio of Y.

Naturally, the required capital or associated risk of a portfolio viewed as its own subportfolio is the (standalone) risk of the portfolio. Therefore $\Lambda(X, X)$ should be a risk measure in the sense of the previous sections, i.e. the risk measure associated with the capital allocation is

$$\rho_\Lambda(X) = \Lambda(X, X).$$

Of course it is even more natural to look for a capital allocation Λ which is associated with a given risk measure ρ, i.e. find Λ_ρ such that $\Lambda_\rho(X, X) = \rho(X)$.

Let us now start with the intuitive simple requirements that the capital allocated to X, viewed as a subportfolio Y, should fulfill.

1. **Linearity.** For a given overall portfolio Z the capital allocated to a union of subportfolios is equal to the sum of the capital amounts allocated to the individual subportfolios. In particular, the risk capital of a portfolio equals the sum of the risk capital of its subportfolios. More formally, Λ is called linear if

$$\forall a, b \in \mathbb{R}, X, Y, Z \in V \quad \Lambda(aX + bY, Z) = a\Lambda(X, Z) + b\Lambda(Y, Z).$$

2. **Diversification.** The capital allocated to a subportfolio X of a larger portfolio Y never exceeds the risk capital of X considered as a stand-alone portfolio: Λ is called diversifying if

$$\forall X, Y \in V \quad \Lambda(X, Y) \leq \Lambda(X, X).$$

3. **Continuity.** A small increase in a position does only have a small effect on the risk capital allocated to that position: Λ is called continuous at $Y \in V$ if

$$\forall X \in V \quad \lim_{\epsilon \to 0} \Lambda(X, Y + \epsilon X) = \Lambda(X, Y).$$

Now we turn to the relation between the natural risk measure associated with a capital allocation Λ, which is established through the following two theorems:

5.2.1 Theorem *If Λ is a linear, diversifying capital allocation with associated risk measure $\rho = \rho_\Lambda$ defined by*

$$\rho_\Lambda(X) = \Lambda(X, X)$$

then ρ is positively homogeneous and subadditive.

Proof: The subadditivity follows from

$$\begin{aligned}\rho(X+Y) &= \Lambda(X+Y, X+Y) \\ &= \Lambda(X, X+Y) + \Lambda(Y, X+Y) \\ &\leq \Lambda(X, X) + \Lambda(Y, Y),\end{aligned}$$

where we have used linearity in the second equality and diversifying in the third. For the positive homogeneity we see for positive a that on the one-hand side we have

$$\rho(aX) = \Lambda(aX, aX) = a\Lambda(X, aX) \leq a\Lambda(X, X)$$

and on the other hand

$$\rho(aX) = \Lambda(aX, aX) \geq \Lambda(aX, X) = a\rho(X). \quad \square$$

But also the converse can be proved, however, with some convex analysis, cf. [109]:

5.2.2 Theorem *If ρ is positively homogeneous and subadditive then its subgradient Λ_ρ as defined in (5.9) below is a linear, diversifying capital allocation with associated risk measure ρ.*

Sketch of Proof: Consider a fixed (portfolio) loss variable Y. Let $\tilde{l}_Y$ be defined on the linear subspace $\{bY | b \in \mathbb{R}\}$ of V by $\tilde{l}_Y(bY) = b\rho(Y)$. By the positive homogeneity and subadditivity of ρ it can be shown that $\tilde{l}_Y$ is smaller than ρ on its domain of definition $\{bY | b \in \mathbb{R}\}$. By the Hahn-Banach Theorem (cf. [52], Ch II, 3.10) this can be extended to a function l_Y on V which is smaller than ρ on V and satisfies by definition

$$\rho(Y) = l_Y(Y) = \max\{l(U) | l \in \mathcal{L}, l \leq \rho\}$$

where $\mathcal{L}$ is the set of linear functionals on V.

Simplifying this result to elementary school mathematics, this is the statement that for a nice convex function like x^2 we can draw at each point (x_0, x_0^2) of its graph the tangent. The tangent through (x_0, x_0^2) lies always below the graph and all other straight lines below the graph will always be smaller at x_0 than the tangent at x_0. Thus, the linear functional l_X can be viewed as the tangent, which is of course nothing else as the derivative (subgradient) of the function ρ at point Y.

Now, it is obvious that we can define $\Lambda(\cdot, Y)$ to be this function l_Y, in other words:

$$\Lambda_\rho(X, Y) = l_Y(X). \tag{5.9}$$

Linearity is then clear since l_Y is linear and the same line of arguments as in the proof above will give the property "diversifying." □

If Λ satisfies an additional smoothness property (Axiom 3), namely continuity in the second entry,

$$\Lambda(X, Y) = \Lambda(X, Y + \epsilon X), \tag{5.10}$$

then already Gateux-differentiability of the diagonal ρ_Λ, i.e. of the associated risk measures, follows and

$$\Lambda(X, Y) = \lim_{\epsilon \to 0} \frac{\rho_\Lambda(Y + \epsilon X) - \rho_\Lambda(Y)}{\epsilon}. \tag{5.11}$$

5.2.1.1 Expected Shortfall Contribution

In line with the theorem 5.2.2 and formula (5.6) above, we see that the capital allocation rule with respect to expected shortfall equals

$$\begin{aligned}
\Lambda_\alpha(X, Y) &= \mathbb{E}[g_a(Y)X] \qquad (5.12)\\
&= \mathbb{E}\left[\left((1-\alpha)^{-1}\mathbf{1}_{\{Y > VaR_\alpha(Y)\}} + \beta_Y \mathbf{1}_{\{Y = VaR_\alpha(Y)\}}\right) X\right]\\
&= \mathbb{E}\left[\left((1-\alpha)^{-1}\mathbf{1}_{\{Y > VaR_\alpha(Y)\}}\right) X\right],
\end{aligned}$$

where the last equation only holds if there is no point mass at the α-quantile, in particular if Y has a continuous distribution.

5.2.1.2 Spectral Capital Allocation

It follows from theorem 5.1.4 and formula (2.7) that the allocation with respect to a spectral risk measure can be written as a mixture of Expected Shortfall allocations

$$\begin{aligned}
&\Lambda_w(X, Y) = \mathbb{E}[g_w(Y)X] \qquad (5.13)\\
&= \int_0^1 \mathbb{E}\left[\left((1-\alpha)^{-1}\mathbf{1}_{\{Y > VaR_\alpha(Y)\}} + \beta_Y \mathbf{1}_{\{Y = VaR_\alpha(Y)\}}\right) X\right] (1-\alpha)\mu_w(d\alpha)\\
&= \int_0^1 \mathbb{E}\left[\left(\mathbf{1}_{\{Y > VaR_\alpha(Y)\}}\right) X\right] \mu_w(d\alpha),
\end{aligned}$$

where the last equation again only holds if there is no point mass at the α-quantile.

If we assume a continuous distribution of Y and have mixture of three Expected Shortfall, i.e.

$$\rho_w(Y) = ES_{\alpha_1}\tilde{w}(\alpha_1) + ES_{\alpha_2}\tilde{w}(\alpha_2) + ES_{\alpha_3}\tilde{w}(\alpha_3),$$

then the capital allocation with respect to this spectral risk measure, is such the same weighted sum of the capital allocation with respect to the three Expected Shortfall contributions. This enables the risk management of the financial institution to formulate its risk appetite in terms of weights given to some well chosen quantiles of its loss distribution. For example it used to be an ongoing discussion whether to really only allocate capital with respect to tail events and ignore the volatility of the loss distribution for the capital allocation. Spectral Capital Allocation allows the risk manager to take care of the entire loss distribution by considering or approximating volatility by a lower quantile. Thus even relatively frequent events can be used, although with a smaller weight, in the capital allocation process.

Figure 5.2 shows expected shortfall contributions for a set of portfolio assets at increasing confidence levels. These shortfall contributions are then used to build contributions to a spectral risk measure as an equally weighted sum (white triangles). Obviously, the assets contribute in very different percentage amounts of their exposures to the various shortfall measures, which can be composed to a spectral measure according to a given risk appetite.

5.2.2 Capital Allocation in Practice

So when talking about capital allocation, we are therefore faced with the problem of *differentiability* of some considered risk measure.

It turns out that the standard deviation as a risk measure has nice differentiability properties. This observation is the "heart" of Markovitz's classical portfolio theory (var/covar approach). Now, VaR-based EC is a *quantile-based* quantity, and only in case of normal distributions, quantiles and standard deviations are reconcilable from a portfolio optimization point of view. Because we know from previous chapters that credit portfolio loss distributions are typically skewed with fat tails, the

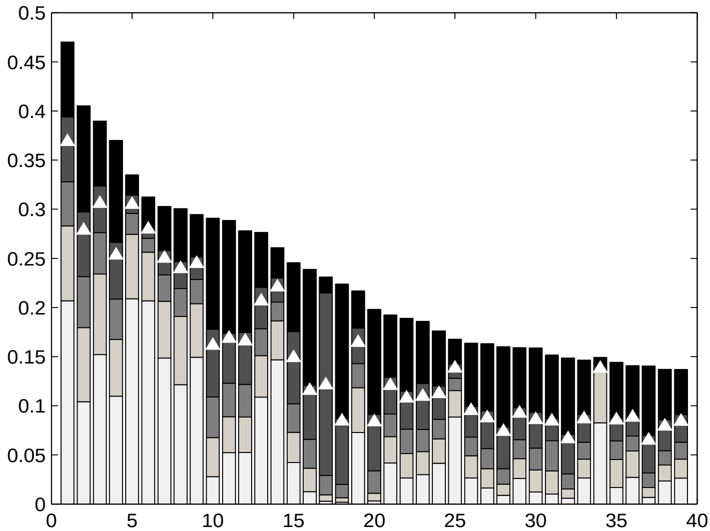


FIGURE 5.2: Expected shortfall contributions (in percent of the exposure) at various confidence levels (EL, 70%, 90%, 99%, 99.9%, with increasing darkness of shading) for a set of assets. The white triangles indicate the contributions to a spectral risk measure, composed of an equally weighted sum of the different ES-contributions.

"classical" approach can not be applied to credit risk in a straightforward manner without "paying" for the convenience of the approach by allowing for inconsistencies.

Fortunately, the var/covar approach for *capital allocation*, adapted to credit risk portfolios, in many cases yields acceptable results and is, due to its simplicity, implemented in most standard software packages. The following section explains the details.

5.2.3 Variance/Covariance Approach

At the core of var/covar is the question of what an individual credit or business unit contributes to the portfolio's standard deviation UL_{PF}. To answer this question, the classical var/covar approach splits the portfolio risk UL_{PF} into risk contributions RC_i in a way such that

$$\sum_{i=1}^{m} w_i \times \mathrm{RC}_i \;=\; \mathrm{UL}_{PF} \;.$$

In this way, the weighted risk contributions sum-up to the total risk of the portfolio, where "risk" is identified with volatility. It follows from Proposition 1.2.7 that

$$\begin{aligned}\mathrm{UL}_{PF} &= \frac{1}{\mathrm{UL}_{PF}} \sum_{i=1}^{m} w_i \cdot \sum_{j=1}^{m} w_j \eta_i \eta_j \sigma_{D_i} \sigma_{D_j} \rho_{ij} \\ &= \frac{1}{\mathrm{UL}_{PF}} \sum_{i=1}^{m} w_i \cdot \sum_{j=1}^{m} w_j \mathrm{UL}_i \mathrm{UL}_j \rho_{ij},\end{aligned}$$

where $\sigma_{D_i} = \sqrt{\mathrm{PD}_i(1-\mathrm{PD}_i)}$ denotes the standard deviation of the default indicator and $\mathrm{UL}_i = \eta_i \sigma_{D_i}$. Thus,

$$\mathrm{RC}_i = \frac{\mathrm{UL}_i}{\mathrm{UL}_{PF}} \sum_{j=1}^{m} w_j \mathrm{UL}_j \rho_{ij}$$

is a plausible quantity measuring the "risk portion" of credit i in a way such that all weighted risks sum-up to the portfolio's UL. It is straightforward to show that the quantity RC_i corresponds to the covariance of credit (business unit) i and the total portfolio loss, divided by the portfolio's volatility respectively UL. The definition of RC_i obviously is in analogy to *beta-factor models* used in market risk. Furthermore,

RC_i is equal to the *partial derivative* of UL_{PF} w.r.t. w_i, the weight of the i-th credit in the portfolio, i.e.,

$$\mathrm{RC}_i \;=\; \frac{\partial \mathrm{UL}_{PF}}{\partial w_i}\,.$$

In other words, an increase in the weight of the considered credit by a small amount h in the portfolio, implies a growth of UL_{PF} by $h \times \mathrm{RC}_i$. Coming from this side, it can in turn be easily shown that the total sum of weighted partial derivatives again equals UL_{PF}.

Regarding the ratio between risk contributions and the standard deviation of the individual exposure, it is true in most cases that

$$\frac{\mathrm{RC}_i}{\mathrm{UL}_i} \;\leq\; 1.$$

This quantity is known as the *retained risk* of unit i. It is the portion of risk of the i-th entity that has not been absorbed by diversification in the portfolio. In contrast, the quantity

$$1 - \frac{\mathrm{RC}_i}{\mathrm{UL}_i}$$

is often called the corresponding *diversification effect*.

Capital multiplier Since the total risk capital is typically determined via quantiles, i.e.,

$$\mathrm{EC}_{\mathrm{VaR}_\alpha} = \mathrm{VaR}_\alpha(L) - \mathbb{E}[L]\ ,$$

the individual risk contributions have to be rescaled with the so-called *capital multiplier*

$$\mathrm{CM}_\alpha = \frac{\mathrm{EC}_{\mathrm{VaR}_\alpha}}{\mathrm{UL}_{PF}}$$

in order to imitate the classical approach from market risk. The contributory capital for credit i then equals

$$\delta_i = \mathrm{CM}_\alpha \times \mathrm{RC}_i, \quad \text{with} \quad \sum_{i=1}^{m} w_i \delta_i = \mathrm{EC}_{\mathrm{VaR}_\alpha}\ .$$

The quantity δ_i is called the *analytic capital contribution* of transaction i to the portfolio capital. For a business unit in charge for credits 1 to

l, where $l < m$, the capital requirement is

$$\mathrm{CM}_\alpha \sum_{j=1}^{l} w_j \, \mathrm{RC}_j .$$

Note that the capital multiplier is an auxiliary quantity depending on the particular portfolio, due to the fact that, in contrast to the normal distribution, the quantiles of credit portfolio loss distributions not only depend on the standard deviation, but also on other influences like correlations, default probabilities, and exposure weights. Therefore it is unrealistic, after changing the portfolio, to obtain the same capital multiplier CM_α as originally calculated.

In the next two sections we discuss capital allocation with respect to EC based on VaR and expected shortfall respectively.

5.2.4 Capital Allocation w.r.t. Value-at-Risk

Calculating risk contributions associated with the VaR risk measure is a natural but difficult attempt, since in general the quantile function will not be differentiable with respect to the asset weights. Under certain continuity assumptions on the joint density function of the random variables X_i, differentiation of $\mathrm{VaR}_\alpha(X)$, where $X = \sum_i w_i X_i$, is guaranteed. One has (see [176])

$$\frac{\partial \mathrm{VaR}_\alpha}{\partial w_i}(X) \;=\; \mathbb{E}[X_i \mid X = \mathrm{VaR}_\alpha(X)]. \tag{5.14}$$

Unfortunately, the distribution of the portfolio loss $L = \sum w_i \hat{L}_i$, as specified at the beginning of this chapter, is purely discontinuous. Therefore the derivatives of VaR_α in the above sense will either not exist or will vanish to zero. In this case we could still define risk contributions via the right-hand side of Equation (5.14) by writing

$$\gamma_i \;=\; \mathbb{E}[\hat{L}_i \mid L = \mathrm{VaR}_\alpha(L)] - \mathbb{E}[\hat{L}_i] \; . \tag{5.15}$$

For a clearer understanding, note that

$$\frac{\partial \mathbb{E}[L]}{\partial w_i} \;=\; \mathbb{E}[\hat{L}_i] \qquad \text{and} \qquad \sum_{i=1}^{m} w_i \gamma_i \;=\; \mathrm{EC}_{\mathrm{VaR}_\alpha}.$$

Additionally observe, that for a large portfolio and on an appropriate scale, the distribution of L will appear to be "close to continuous."

Unfortunately, even in such "approximately good" cases, the loss distribution often is not given in an analytical form in order to allow for differentiations.

Remark For the CreditRisk$^+$ model, an analytical form of the loss distribution can be found; see Section 2.4.2 and Chapter 4 for a discussion of CreditRisk$^+$. Tasche [175] showed that in the CreditRisk$^+$ framework the VaR contributions can be determined by calculating the corresponding loss distributions several times with different parameters. Martin et al. [132] suggested an approximation to the partial derivatives of VaR via the so-called *saddle point method*.

Capital allocation based on VaR is not really satisfying, because in general, although $(\mathrm{RC}_i)_{i=1,\ldots,m}$ might be a reasonable partition of the portfolio's standard deviation, it does not really say much about the tail risks captured by the quantile on which VaR-EC is relying. Even if in general one views capital allocation by means of partial derivatives as useful, the problem remains that the var/covar approach completely neglects the dependence of the quantile on correlations. For example, var/covar implicitely assumes

$$\frac{\partial \mathrm{VaR}_\alpha(X)}{\partial \mathrm{UL}_{PF}} \;=\; \text{const} \;=\; \mathrm{CM}_\alpha,$$

for the specified confidence level α. This is true for (multivariate) normal distributions, but generally is not the case for loss distributions of credit portfolios. As a consequence it can happen that transactions require a contributory EC exceeding the original exposure of the considered transaction. This effect is very unpleasant. Therefore, we now turn to expected shortfall-based EC instead of VaR-based EC.

5.2.5 Capital Allocations w.r.t. Expected Shortfall

At the beginning we must admit that shortfall-based risk contributions bear the same "technical" difficulty as VaR-based measures, namely the quantile function is not differentiable in general. But, we find in Tasche [176] that if the underlying loss distribution is "sufficiently smooth," then TCE_α is partially differentiable with respect to the exposure weights. One finds that

$$\frac{\partial \mathrm{TCE}_\alpha}{\partial w_i}(X) \;=\; \mathbb{E}[X_i \mid X > \mathrm{VaR}_\alpha(X)].$$

In case the partial derivatives do not exist, one again can rely on the right-hand side of the above equation by defining shortfall contributions for, e.g., discontinuous portfolio loss variables $L = \sum w_i \hat{L}_i$ by

$$\zeta_i \;=\; \mathbb{E}[\hat{L}_i \mid L > \mathrm{VaR}_\alpha(L)] - \mathbb{E}[\hat{L}_i] \;, \tag{5.16}$$

which is then also the natural capital allocation rule with respect to Expected Shortfall. The proof that this definition of capital allocation coincides with the theoretical capital allocation derived in Section (5.2.1) has still to be carried out. Analogous to what we saw in case of VaR-EC, we can write

$$\sum_{i=1}^{m} w_i \zeta_i \;=\; \mathrm{EC}_{\mathrm{TCE}_\alpha} \;,$$

such that shortfall-based EC can be obtained as a weighted sum of the corresponding contributions.

Remarks With expected shortfall we have identified a coherent (or close to coherent) risk measure, which overcomes the major drawbacks of classical VaR approaches. Furthermore, shortfall-based measures allow for a consistent definition of risk contributions. We continue with some further remarks:

- The results on shortfall contributions together with the findings on differentiability in [152] indicate that the proposed capital allocation ζ_i can be used as a performance measure, as pointed out in Theorem 4.4 in [176], for example. In particular, it shows that if one increases the exposure to a counterparty having a RAROC above portfolio RAROC, the portfolio RAROC will be improved. Here RAROC is defined as the return over (contributory) economic capital.

- We obtain $\zeta_i < \hat{L}_i$, i.e., by construction the capital is always less than the exposure, a feature that is not shared by risk contributions defined in terms of covariances.

- The definition of shortfall contributions reflects a causality relation. If counterparty i contributes higher to the overall loss than counterparty j in extreme loss scenarios, then, as a consequence, business with i should be more costly (assuming stand-alone risk characteristics are the same).

- Capital allocation rules according to shortfall contributions can easily be extended to the space of all coherent risk measures as shown in this chapter. If the density L_Y of the maximal generalized scenario associated with the portfolio Y is known then the capital allocation to transaction L_i equals

$$E[L_Y L_i].$$

5.2.6 A Simulation Study

In the simulation study we want to compare the two different allocation techniques, namely allocation based on VaR and allocation based on expected shortfall. We first tested it on a transaction base. In a subsequent test case we considered the allocation of capital to business units. There are at least two reasons justifying the efforts for the second test. First, it might not be reasonable to allocate economic capital that is based on extreme loss situations to a single transaction, since the risk in a single transaction might be driven by short-term volatility and not by the long-term view of extreme risks. The second reason is more driven by the computational feasibility of expected shortfall. In the "binary world" of default simulations, too many simulations are necessary in order to obtain a positive contribution conditional on extreme default events for all counterparties.

The basic result of the simulation study is that analytic contributions produce a steeper gradient between risky and less risky loans than tail risk contributions. In particular, loans with a high default probability but moderate exposure concentration require more capital in the analytic contribution method, whereas loans with high concentration require relatively more capital in the shortfall contribution method.

Transaction View The first simulation study is based on a credit portfolio considered in detail in [152]. The particular portfolio consists of 40 counterparties.

As capital definition, the 99% quantile of the loss distribution is used. Within the Monte-Carlo simulation it is straightforward to evaluate risk contributions based on expected shortfall. The resulting risk contributions and its comparison to the analytically calculated risk contributions based on the volatility decomposition are shown in Figure 5.3.

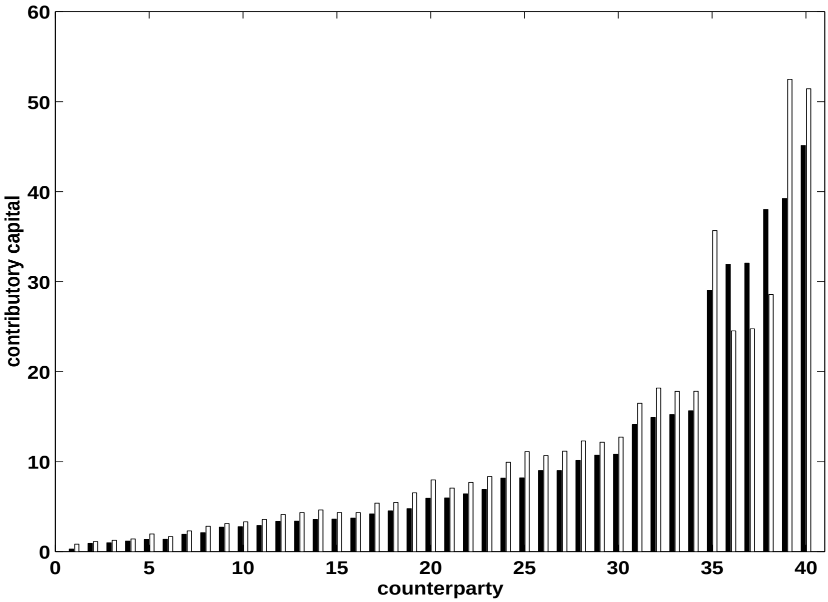


FIGURE 5.3: The bar chart depicts the different risk contributions for every counterparty in the portfolio. The dark bars belong to the counterparty contribution measured by the shortfall; the white ones correspond to the analytic Var/Covar-contribution.

In the present portfolio example the difference between the contributory capital of two different types, namely analytic risk contributions and contributions to shortfall, should be noticed, since even the order of the assets according to their risk contributions changed. The asset with the largest shortfall contribution is the one with the second largest var/covar risk contribution, and the largest var/covar risk contribution goes with the second largest shortfall contribution. A review of the portfolio shows that the shortfall contributions are more driven by the relative asset size. However, it is always important to bear in mind that these results are still tied to the given portfolio.

It should also be noticed that the gradient of the EC is steeper for the analytic approach. Bad loans might be able to breech the hurdle rate in a RAROC-Pricing tool if one uses the expected shortfall approach, but might fail to earn above the hurdle rate if EC is based on var/covar.

Business Unit View The calculation of expected shortfall contributions requires a lot more computational power, which makes it less feasible for large portfolios. However, the capital allocation on the business level can accurately be measured by means of expected shortfall contributions. Figure 5.4 shows an example of a bank with six business units. Again we see that expected shortfall allocation differs from var/covar allocation.

Under var/covar, it sometimes can even happen that the capital allocated to a business unit is larger if considered consolidated with the bank than capitalized standalone. This again shows the non-coherency of VaR measures. Such effects are very unpleasant and can lead to significant misallocations of capital. Here, expected shortfall provides the superior way of capital allocation. We conclude this chapter with a simple remark about how one can calculated EC on VaR-basis but allocate capital shorfall-based.

If a bank calculates its total EC by means of VaR, it still can allocate capital in a coherent way. For this purpose, one just has to determine some threshold $c < \mathrm{VaR}_\alpha$ such that

$$\mathrm{EC}_{\mathrm{TCE}}(c) \;\approx\; \mathrm{EC}_{\mathrm{VaR}_\alpha} \; .$$

This *VaR-matched expected shortfall* is a coherent risk measure preserving the VaR-based overall economic capital. It can be viewed as an approximation to VaR-EC by considering the whole tail of the loss

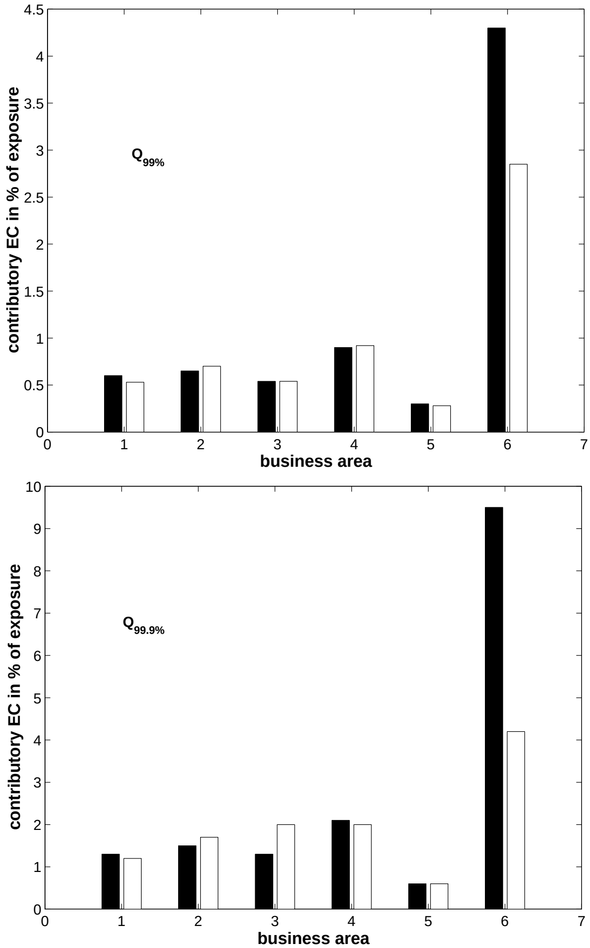


FIGURE 5.4: The bar charts depict the different risk contributions (top: 99% quantile, bottom: 99.9% quantile) of the business areas of a bank. The black bars are based on a Var/Covar approach; the white ones correspond to shortfall risk.

distribution, starting at some threshold below the quantile, such that the resulting mean value matches the quantile. Proceeding in this way, allocation of the total VaR-based EC to business units will reflect the coherency of shortfall-based risk measures.

Further Reading

In order to round off the discussion on risk measures and capital allocation we would like to point to further literature and research. The concept of coherent risk measures is widely accepted by now, but a serious criticism to the necessity of subadditivity and positive homogeneity arises when liquidity risk is taken into account. This risk arises when the market cannot be assumed to be infinitely deep, i.e., when large positions cannot easily be absorbed by the market and a doubling of a position results in more than doubling the risk. To address these problems Heath and Kuhn [91], Föllmer and Schied [64, 65] and Fritelli and Gianin [71] introduce convex risk measures by replacing the axioms on positive homogeneity and subadditivity by the weaker requirement of convexity. A representation theorem for generalized spectral risk measures can be found in Kusuoka [117], Acerbi [1] and Tasche [177], Föllmer and Schied [65] and Weber [186].

Chapter 6

Term Structure of Default Probability

So far, default has mostly been modeled as a binary event (except the intensity model), suited for single-period considerations within the regulatory framework of a fixed planning horizon. However, the choice of a specific period like one year is more or less arbitrary. Even more, default is an inherently time-dependent event. This chapter serves to introduce the idea of a term structure of default probability. This credit curve represents a necessary prerequisite for a time-dependent modeling as in Chapters 7 and 8. In principle, there are three different methods to obtain a credit curve: from historical default information, as implied probabilities from market spreads of defaultable bonds, and through Merton's option theoretic approach. The latter has already been treated in a previous chapter, but before introducing the other two in more detail we first lay out some terminology used in survival analysis (see [28, 33] for a more elaborated presentation).

6.1 Survival Function and Hazard Rate

For any model of default timing, let $S(t)$ denote the probability of surviving until t. With help of the time-until-default τ (or briefly "default time"), a continuous random variable, the *survival function* $S(t)$ can be written as

$$S(t) = \mathbb{P}[\tau > t], \quad t \geq 0\,.$$

That is, starting at time $t = 0$ and presuming no information is available about the future prospects for survival of a firm, $S(t)$ measures the likelihood that it will survive until time t. The probability of default between time s and $t \geq s$ is simply $S(s) - S(t)$. In particular, if $s = 0$, and because $S(0) = 1$, then the probability of default $F(t)$ is

$$F(t) = 1 - S(t) = \mathbb{P}[\tau \leq t], \quad t \geq 0. \tag{6.1}$$

$F(t)$ is the distribution function of the random default time τ. The corresponding probability density function is defined by

$$f(t) = F'(t) = -S'(t) = \lim_{\Delta \to 0^+} \frac{\mathbb{P}[t \leq \tau < t + \Delta]}{\Delta},$$

if the limit exists. Furthermore, we introduce the *conditional* or *forward* default probability

$$p(t|s) = \mathbb{P}[\tau \leq t | \tau > s], \quad t \geq s \geq 0,$$

i.e., the probability of default of a certain obligor between t and s conditional on its survival up to time s, and

$$q(t|s) = 1 - p(t|s) = \mathbb{P}[\tau > t | \tau > s] = S(t)/S(s), \quad t \geq s \geq 0,$$

the forward survival probability. An alternative way of characterizing the distribution of the default time τ is the *hazard function*, which gives the instantaneous probability of default at time t conditional on the survival up to t. The hazard function is defined via

$$\mathbb{P}[t < \tau \leq t + \Delta t | \tau > t] = \frac{F(t + \delta t) - F(t)}{1 - F(t)} \approx \frac{f(t)\Delta t}{1 - F(t)}$$

as

$$h(t) = \frac{f(t)}{1 - F(t)}.$$

Equation (6.1) yields

$$h(t) = \frac{f(t)}{1 - F(t)} = -\frac{S'(t)}{S(t)},$$

and solving this differential equation in $S(t)$ results in

$$S(t) = e^{-\int_0^t h(s)ds}. \tag{6.2}$$

This allows us to express $q(t|s)$ and $p(t|s)$ as

$$q(t|s) = e^{-\int_s^t h(u)du}, \tag{6.3}$$

$$p(t|s) = 1 - e^{-\int_s^t h(u)du}. \tag{6.4}$$

Additionally, we obtain

$$F(t) = 1 - S(t) = 1 - e^{-\int_0^t h(s)ds},$$

and

$$f(t) = S(t)h(t).$$

One could assume the hazard rate to be piecewise constant, i.e., $h(t) = h_i$ for $t_i \le t < t_{i+1}$. In this case, it follows that the density function of τ is

$$f(t) = h_i e^{-h_i t} \mathbf{1}_{[t_i, t_{i+1}[}(t),$$

showing that the survival time is exponentially distributed with parameter h_i. Furthermore, this assumption entails over the time interval $[t_i, t_{i+1}[$ for $0 < t_i \le t < t_{i+1}$

$$q(t|t_i) = e^{-\int_{t_i}^t h(u)du} = e^{-h_i(t-t_i)}.$$

Remark The "forward default rate" $h(t)$ as a basis of a default risk term structure is in close analogy to a forward interest rate, with zero-coupon bond prices corresponding to survival probabilities. The hazard rate function used to characterize the distribution of survival time can also be called a "credit curve" due to its similarity to a yield curve. If h is continuous then $h(t)\Delta t$ is approximately equal to the probability of default between t and $t + \Delta t$, conditional on survival to t. Understanding the first arrival time τ as associated with a Poisson arrival process, the constant mean arrival rate h is then called intensity and often denoted by λ[1]. Changing from a deterministically varying intensity to random variation, and thus closing the link to the stochastic intensity models [50], turns Equation (6.3) into

$$q(t|s) = \mathbb{E}_s\left[e^{-\int_s^t h(u)du}\right],$$

where $\mathbb{E}_s$ denotes expectation given all information available at time s.

[1] Note that some authors explicitly distinguish between the intensity $\lambda(t)$ as the arrival rate of default at t conditional on *all* information available at t, and the forward default rate $h(t)$ as arrival rate of default at t, conditional *only* on survival until t.

6.2 Risk-Neutral vs. Actual Default Probabilities

When estimating the risk and the value of credit-related securities we are faced with the question of the appropriate probability measure, *risk-neutral* or *objective probabilities*. But in fact, the answer depends on the objective we have. If one is interested in estimating the economic capital and risk charges, one adopts an actuarial-like approach by choosing historical probabilities as underlying probability measure. In this case we assume that actual default rates from historical information allow us to estimate a capital quota protecting us against losses in worst case default scenarios. The objective is different when it comes to pricing and hedging of credit-related securities. Here we have to model under the risk-neutral probability measure. In a risk-neutral world all individuals are indifferent to risk. They require no compensation for risk, and the expected return on all securities is the risk-free interest rate. This general principle in option pricing theory is known as risk-neutral valuation and states that it is valid to assume the world is risk-neutral when pricing options. The resulting option prices are correct not only in the risk-neutral world, but in the real world as well. In the credit risk context, risk-neutrality is achieved by calibrating the default probabilities of individual credits with the market-implied probabilities drawn from bond or credit default swap spreads. The difference between actual and risk-neutral probabilities reflects risk-premiums required by market participants to take risks. To illustrate this difference suppose we are pricing a one-year par bond that promises its face value 100 and a 7% coupon at maturity. The one-year risk-free interest rate is 5%. The actual survival probability for one year is $1 - \mathrm{PD} = 0.99$; so, if the issuer survives, the investor receives 107. On the other hand, if the issuer defaults, with actual probability $\mathrm{PD} = 0.01$, the investor recovers 50% of the par value. Simply discounting the expected payoff computed with the actual default probability leads to

$$\frac{(107 \times 0.99 + 50\% \times 100 \times 0.01)}{1 + 5\%} = 101.36\,,$$

which clearly overstates the price of this security. In the above example we have implicitly adopted an actuarial approach by assuming that the price the investor is to pay should exactly offset the expected loss due to a possible default. Instead, it is natural to assume that investors are concerned about default risk and have an aversion to bearing more risk. Hence, they demand an additional risk premium and the pricing should somehow account for this risk aversion. We therefore turn the above pricing formula around and ask which probability results in the quoted price, given the coupons, the risk-free rate, and the recovery value. According to the risk-neutral valuation paradigm, the fact that the security is priced at par implies that

$$100 = \frac{(107 \times (1 - \mathrm{PD}^*) + 50\% \times 100 \times \mathrm{PD}^*)}{1 + 5\%}.$$

Solving for the market-implied risk-neutral default probability yields $\mathrm{PD}^* = 0.0351$. Note that the actual default probability $\mathrm{PD} = 0.01$ is less than PD^*. Equivalently, we can say that the bond is priced as though it were a break-even trade for a “stand-in” investor who is not risk adverse but assumes a default probability of 0.0351. The difference between PD and PD^* reflects the risk premium for default timing risk. Most credit market participants think in terms of spreads rather than in terms of default probabilities, and analyze the shape and movements of the spread curve rather than the change in default probabilities. And, indeed, the link between credit spread and probability of default is a fundamental one, and is analogous to the link between interest rates and discount factors in fixed income markets. If s represents a multiplicative spread over the risk-free rate one gets

$$\mathrm{PD}^* = \frac{1 - \frac{1}{1+s}}{1 - REC} \approx \frac{s}{1 - REC},$$

where the approximation is also valid for additive spreads.

“Actuarial credit spreads” are those implied by assuming that investors are neutral to risk, and use historical data to estimate default probabilities and expected recoveries. Data from Fons [66] suggest that corporate yield spreads are much larger than the spreads suggested by actuarial default losses alone. For example, actuarially implied credit spreads on a A-rated 5-year US corporate debt were estimated by Fons to be six basis points. The corresponding market spreads have been in the order of 100 basis points. Clearly, there is more than default risk behind the difference between “actuarial credit spreads” and actual yield spreads, like liquidity risk, tax-related issues, etc. But even after measuring spreads relative to AAA yields (thereby stripping out treasury effects), actuarial credit spreads are smaller than actual market spreads, especially for high-quality bonds.

6.3 Term Structure Based on Historical Default Information

Multi-year default probabilities can be extracted from historical data on corporate defaults similarly to the one-year default probabilities. But before going into details we first show a "quick and dirty" way to produce a whole term structure if only one-year default probabilities are at hand.

6.3.1 Exponential Term Structure

The derivation of exponential default probability term structure is based on the idea that credit dynamics can be viewed as a two-state time-homogeneous Markov-chain, the two states being survival and default, and the unit time between two time steps being Δ. Suppose a default probability PD_T for a time interval T (e.g., one year) has been calibrated from data; then the survival probability for the time unit Δ (e.g., one day) is given by

$$\mathbb{P}[\tau > t + \Delta | \tau \geq t] = (1 - \mathrm{PD}_T)^{\Delta/T} \ , \tag{6.5}$$

and the default probability for the time t, in units of Δ, is then

$$\mathrm{PD}_t = 1 - (1 - \mathrm{PD}_T)^{t/T} \ . \tag{6.6}$$

In the language of survival analysis we can write for the probability of survival until T

$$1 - \mathrm{PD}_T = q(T|0) = e^{-\int_0^T h(u)du} = e^{-\bar{h}T},$$

where the last equation defines the average default rate $\bar{h}$,

$$\bar{h} = -\log(1 - \mathrm{PD}_T)/T.$$

Assuming a constant default rate over the whole lifetime of the debt, Equation (6.6) reads

$$F(t) = q(t|0) = 1 - p(t|0) = 1 - e^{-\bar{h}t}.$$

6.3.2 Direct Calibration of Multi-Year Default Probabilities

Rating agencies also provide data on multi-year default rates in their reports. For example, Moody's [141] trailing $T+1$-month default rates for month t and rating universe k are defined as

$$D_{k,t} = \frac{\sum_{i=t-T}^{t} Y_{k,i}}{I_{k,t-11}} . \tag{6.7}$$

k, for example, could be all corporate issuers, US speculative grade issuers, or Ba-rated issuers in the telecom sector. The numerator is the sum of defaulters, Y, in month t that were in the rating universe k as of $t-T$. The denominator, $I_{k,t}$, is the number of issuers left in the rating universe k in month $t-T$, adjusted to reflect the withdrawal from the market of some of those issuers for noncredit-related reasons (e.g., maturity of debt). The adjustment for withdrawal is important because the denominator is intended to represent the number of issuers who could potentially have defaulted in the subsequent $T+1$-month period. Underlying Equation (6.7) is the assumption that defaults in a given rating universe are independent and identically distributed Bernoulli random variables, i.e., the number of defaults w.r.t. a certain pool, rating, and year follow a binomial distribution. Note that this assumption is certainly not correct in a strict sense; in fact, correlated defaults are the core issue of credit portfolio models.

Moody's employs a dynamic cohort approach to calculating multi-year default rates. A cohort consists of all issuers holding a given estimated senior rating at the start of a given year. These issuers are then followed through time, keeping track of when they default or leave the universe for noncredit-related reasons. For each cumulation period, default rates based on dynamic cohorts express the ratio of issuers who did default to issuers who were in the position to default over that time period. In terms of Equation (6.7) above, this constitutes lengthening the time horizon T ($T = 11$ in the case of one-year default rates). Since more and more companies become rated over the years, Moody's and S&P use an issuer weighted average to averaged cumulative default rates. To estimate the average risk of default over time horizons longer than one year, Moody's calculates the risk of default in each year since a cohort was formed. The issuer-weighted average of each cohort's one-year default rate forms the average cumulative one-year default rate. The issuer-weighted average of the second-year (marginal) default rates

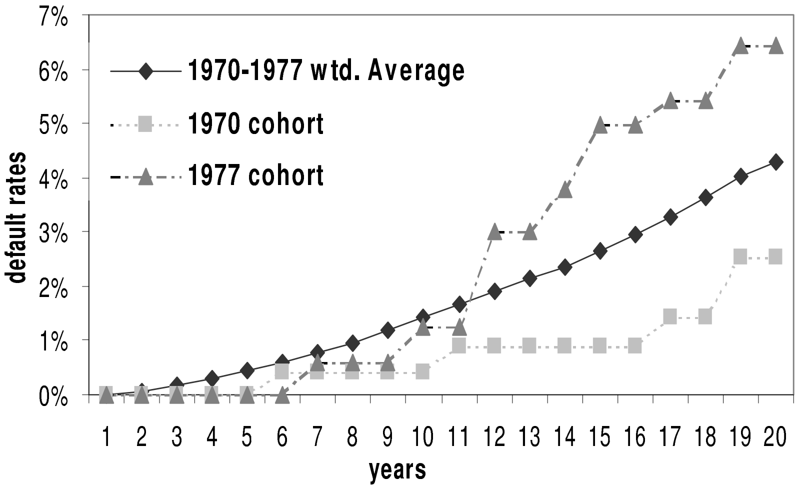


FIGURE 6.1: Cumulative default rate for A-rated issuer. Weighted average default rates smooth economic cycles' effects on individual cohorts, from [141].

(default in exactly the second year) cumulated with that of the first year yields the two-year average cumulative default rate, and so on. Figure 6.1 shows how different cohorts produce a different credit history in response to different economic and market conditions.

Table 6.1 gives the average cumulative default rates as reported by Moody's [141]. A closer look reveals some unpleasant features in this table. For example, one would expect column monotony for each year, i.e., high credit quality should never show a higher default rate than low credit quality, which is violated at various entries. Furthermore, some marginal default rates are zero, even for non-triple A rated corporates, which is unrealistic. Clearly, these problems stem from a lack of sufficient data for a reliable statistical analysis, and, obviously, pooling on a more aggregated level produces more satisfactory results w.r.t. these plausibility conditions; see Table 6.2. In the next section we show a way to avoid these deficiencies by use of migration analysis.

One can argue that the issuer-weighted arithmetic mean is perhaps not the right thing to do. Since more and more corporates are rated during the last years, issuer-weighted averaging means that recent years have much more impact than years further back in history and the result does not reflect a typical year as averaged over economic cycles.

TABLE 6.1: Average cumulative default rates from 1 to 10 years – 1983–2000, from [141].

	1	2	3	4	5	6	7	8	9	10
Aaa	0,00%	0,00%	0,00%	0,06%	0,18%	0,25%	0,34%	0,43%	0,43%	0,43%
Aa1	0,00%	0,00%	0,00%	0,21%	0,21%	0,35%	0,35%	0,35%	0,35%	0,35%
Aa2	0,00%	0,00%	0,06%	0,18%	0,41%	0,49%	0,59%	0,71%	0,85%	1,01%
Aa3	0,06%	0,09%	0,17%	0,26%	0,37%	0,49%	0,49%	0,49%	0,49%	0,49%
A1	0,00%	0,03%	0,30%	0,47%	0,59%	0,73%	0,79%	0,86%	0,86%	0,96%
A2	0,00%	0,02%	0,16%	0,41%	0,62%	0,84%	0,99%	1,35%	1,63%	1,71%
A3	0,00%	0,12%	0,22%	0,30%	0,35%	0,47%	0,68%	0,77%	0,97%	1,09%
Baa1	0,07%	0,30%	0,53%	0,86%	1,19%	1,43%	1,82%	2,05%	2,20%	2,20%
Baa2	0,06%	0,29%	0,61%	1,22%	1,89%	2,54%	2,93%	3,17%	3,46%	3,81%
Baa3	0,39%	1,05%	1,62%	2,47%	3,15%	4,09%	4,99%	5,95%	6,54%	7,03%
Ba1	0,64%	2,10%	3,81%	6,15%	8,12%	10,09%	11,43%	12,75%	13,35%	14,08%
Ba2	0,54%	2,44%	4,95%	7,32%	9,27%	10,88%	12,59%	13,60%	14,27%	14,71%
Ba3	2,47%	6,82%	11,68%	16,18%	20,63%	24,74%	28,39%	32,28%	35,83%	38,22%
B1	3,48%	9,71%	15,59%	20,56%	25,62%	30,78%	36,15%	40,30%	44,16%	48,01%
B2	6,23%	13,70%	20,03%	24,63%	28,24%	31,14%	32,73%	34,33%	35,03%	35,90%
B3	11,88%	20,18%	26,71%	31,95%	36,68%	39,89%	42,81%	46,80%	51,42%	53,53%
Caa1-C	18,85%	28,29%	34,51%	40,23%	43,42%	46,48%	46,48%	49,73%	53,92%	59,04%
Investment-Grade	0,05%	0,17%	0,35%	0,60%	0,84%	1,08%	1,28%	1,47%	1,62%	1,73%
Speculative-Grade	3,69%	8,39%	12,87%	16,80%	20,39%	23,61%	26,44%	29,04%	31,22%	32,89%
All Corporates	1,21%	2,72%	4,12%	5,34%	6,39%	7,30%	8,05%	8,71%	9,23%	9,61%

TABLE 6.2: Average cumulative default rates by letter rating from 1 to 10 years – 1970–2000, from [141].

Rating	1	2	3	4	5	6	7	8	9	10
Aaa	0.00%	0.00%	0.00%	0.04%	0.12%	0.21%	0.31%	0.42%	0.54%	0.67%
Aa	0.02%	0.04%	0.08%	0.20%	0.31%	0.43%	0.55%	0.67%	0.76%	0.83%
A	0.01%	0.05%	0.18%	0.31%	0.45%	0.61%	0.78%	0.96%	1.18%	1.43%
Baa	0.14%	0.44%	0.83%	1.34%	1.82%	2.33%	2.86%	3.39%	3.97%	4.56%
Ba	1.27%	3.57%	6.11%	8.65%	11.23%	13.50%	15.32%	17.21%	19.00%	20.76%
B	6.16%	12.90%	18.76%	23.50%	27.92%	31.89%	35.55%	38.69%	41.51%	44.57%
Investment-Grade	0.05%	0.17%	0.35%	0.59%	0.82%	1.07%	1.34%	1.61%	1.91%	2.21%
Speculative-Grade	4.15%	8.39%	12.19%	15.48%	18.56%	21.26%	23.48%	25.60%	27.54%	29.46%
All Corporates	1.30%	2.61%	3.76%	4.77%	5.67%	6.46%	7.13%	7.76%	8.37%	8.96%

Having now extracted cumulative default probabilities at discrete points in time, PD_i, we might be interested in a continuous version of this term structure. The simplest answer is an adroit linear interpolation of the multi-year default probability table (the interpolation *between* rating classes ought to be done on a logarithmic scale).

A slightly more sophisticated method can be formulated with the help of the forward default rate h. The forward default probability between t_i and t_{i+1} is given by

$$p(t_{i+1}|t_i) = \frac{\mathrm{PD}_{i+1} - \mathrm{PD}_i}{1 - \mathrm{PD}_i} = 1 - \exp\left(-\int_{t_i}^{t_{i+1}} h(u)du\right).$$

Note that $\mathrm{PD}_i = F(t_i)$. Define for the time interval $[t_i, t_{i+1}]$ an average forward default rate by

$$\bar{h}_i = \frac{1}{t_{i+1} - t_i}\int_{t_i}^{t_{i+1}} h(u)du, \quad \text{for} \quad i = 0, \ldots, n.$$

In terms of the multi-year default probabilities the forward default rate for period i is

$$\bar{h}_i = -\frac{1}{t_{i+1} - t_i}\log\left(\frac{1 - \mathrm{PD}_{i+1}}{1 - \mathrm{PD}_i}\right).$$

Two hazard rate functions obtained from the multi-year default probabilities in Table 6.2 are depicted in Figure 6.2 and show a typical feature: investment grade securities tend to have an upward sloping hazard rate term structure, whereas speculative grades tend to have a downward sloping term structure.

The cumulative default probability until time t, $t_i \le t < t_{i+1}$ boils down to

$$\begin{aligned}\mathrm{PD}_t &= F(t) = 1 - q(t_i|0)q(t|t_i) \\ &= 1 - (1 - \mathrm{PD}_i)\left(\frac{1 - \mathrm{PD}_{i+1}}{1 - \mathrm{PD}_i}\right)^{(t-t_i)/(t_{i+1}-t_i)}.\end{aligned}$$

For $0 < t < 1$ we obtain again the exponential term structure. For $t > t_n$ the term structure can be extrapolated by assuming a constant forward default rate $\bar{h}_{n-1}$ beyond t_{n-1},

$$\mathrm{PD}_t = 1 - (1 - \mathrm{PD}_{n-1})\left(\frac{1 - \mathrm{PD}_n}{1 - \mathrm{PD}_{n-1}}\right)^{(t-t_n)/(t_{i+1}-t_i)}.$$

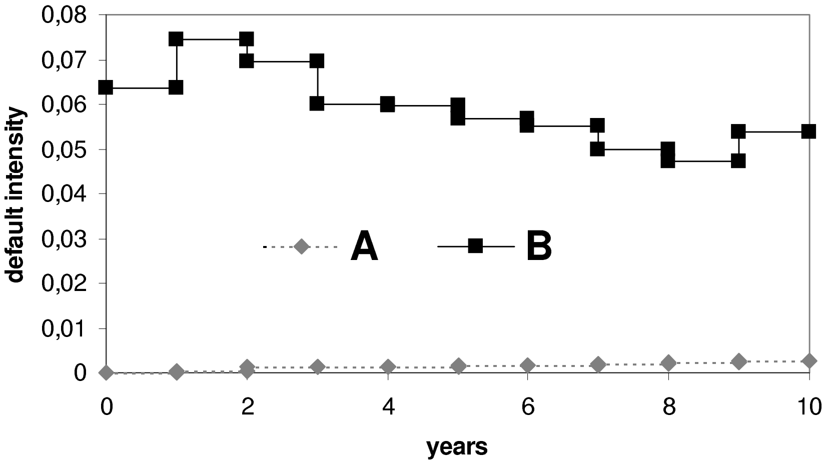


FIGURE 6.2: Hazard rate functions for ratings A and B, calculated from Moody's cumulative default rates.

6.3.3 Migration Technique and Q-Matrices

The reliability of the default frequencies reported in Table 6.1 strongly depends on the quality of the underlying data. In the previous section we pointed out that the amount of data used for the calibration decreases with increasing time horizon. As a consequence, the quality of the calibration suffers from a lack of sufficient data for a reliable statistical analysis, especially at larger time horizons. For this reason, we now discuss a different approach to multi-year default probabilities, whose key idea is the use of migration analysis.

According to migration frequency tables reported by Moody's, S&P, and other rating agencies it is quite likely that corporates experience changes in credit quality over the years. This phenomenon is called *credit migration*, and the likelihoods of transitions from a given rating category to another are collected in *migration matrices.*

The migration technique can be treated best within the mathematical framework of Markov chains, i.e., we assume the existence of a credit migration process controlled solely by the transition probabilities given in the one-year migration matrix. More precisely, we define the finite state space of the chain covering possible bond ratings, e.g., $\Omega = \{AAA, AA, A, BBB, BB, B, CCC, Default\}$ and assign to every pair (i, j) of states a transition or migration probability

$$m_{ij} = P[i \to j] \qquad (i = 1, ..., 7; j = 1, ..., 8),$$

where $P[i \to j]$ denotes the probability of change from rating class i at the beginning of a year to rating class j at year's end. In the present context the Markov property represents the assumption that the evolution of credit migration is independent of the past credit migration history. Through the homogeneity we assume the migration rates to be independent of time, i.e., the probability of a one-year migration $i \to j$ does not depend on the considered year.

Remark Both the assumptions of time-homogeneity and of the Markov property give rise to lively discussions among practitioners and experts regarding how realistic such assumption can possibly be. In Section 6.3.4 we consider a recently developed model where the assumption of time-homogeneity is dropped.

Now we collect the migration probabilities into a one-year migration (or transition) matrix $M = (m_{ij})_{i,j=1,...,8}$ where the 8th row is given by the vector $(0, 0, 0, 0, 0, 0, 0, 1)$. The following properties of M follow immediately:

(i) M has only nonnegative entries: $m_{ij} \geq 0$ for $i, j = 1, ..., 8$.

(ii) All row sums of M are equal to 1: $\sum_{j=1}^{8} m_{ij} = 1$ for $i = 1, ..., 8$.

(iii) The last column contains the 1-year default probabilities: $m_{i,8} = \mathrm{PD}(1, i)$ for $i = 1, ..., 7$.

(iv) The default state is absorbing: $m_{8,j} = 0$ for $j = 1, ..., 7$, and $m_{8,8} = 1$.
This means that there is no escape from the default state.

Rating agencies publish transition matrices for one, two, or more years, usually with an additional column representing the no-longer rated debts. Since this primarily occurs when a company's outstanding debt issue expires, this portion is typically distributed along the rows proportionally to the probability weights in the rated states. Moody's average one-year migration matrix [141], for example, is

$$
\begin{array}{c|c}
 & \begin{array}{ccccccccc} Aaa & Aa & A & Baa & Ba & B & C & Default & WR \end{array} \\
\hline
\begin{array}{c} Aaa \\ Aa \\ A \\ Baa \\ Ba \\ B \\ C \end{array} &
\begin{pmatrix}
0.8617 & 0.0945 & 0.0102 & 0.0000 & 0.0003 & 0.0000 & 0.0000 & 0.0000 & 0.0333 \\
0.0110 & 0.8605 & 0.0893 & 0.0031 & 0.0011 & 0.0001 & 0.0000 & 0.0003 & 00346 \\
0.0006 & 0.0285 & 0.8675 & 0.0558 & 0.0066 & 0.0017 & 0.0001 & 0.0001 & 0.0391 \\
0.0006 & 0.0034 & 0.0664 & 0.8100 & 0.0552 & 0.0097 & 0.0008 & 0.0016 & 0.0523 \\
0.0003 & 0.0006 & 0.0054 & 0.0546 & 0.7550 & 00818 & 0.0053 & 0.0132 & 0.0838 \\
0.0001 & 0.0004 & 0.0020 & 0.0056 & 0.0592 & 0.7593 & 0.0303 & 0.0641 & 0.0790 \\
0.0000 & 0.0000 & 0.0000 & 0.0087 & 0.0261 & 0.0562 & 0.5701 & 0.2531 & 0.0858
\end{pmatrix}
\end{array}
$$

and reads after adjustment for rating withdrawal

$$
M_{Moody's} =
\begin{array}{c|c}
 & \begin{array}{cccccccc} Aaa & Aa & A & Baa & Ba & B & C & D \end{array} \\
\hline
\begin{array}{c} Aaa \\ Aa \\ A \\ Baa \\ Ba \\ B \\ C \\ D \end{array} &
\begin{pmatrix}
0.8914 & 0.0978 & 0.0106 & 0.0000 & 0.0003 & 0.0000 & 0.0000 & 0.0000 \\
0.0114 & 0.8913 & 0.0925 & 0.0032 & 0.0011 & 0.0001 & 0.0000 & 0.0003 \\
0.0006 & 0.0297 & 0.9028 & 0.0581 & 0.0069 & 0.0018 & 0.0001 & 0.0001 \\
0.0006 & 0.0036 & 0.0701 & 0.8547 & 0.0582 & 0.0102 & 0.0008 & 0.0017 \\
0.0003 & 0.0007 & 0.0059 & 0.0596 & 0.8241 & 0.0893 & 0.0058 & 0.0144 \\
0.0001 & 0.0004 & 0.0022 & 0.0061 & 0.0643 & 0.8244 & 0.0329 & 0.0696 \\
0.0000 & 0.0000 & 0.0000 & 0.0095 & 0.0285 & 0.0615 & 0.6236 & 0.2769 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 1
\end{pmatrix}
\end{array}.
$$

A useful consequence of the Markovian property and the time-homogeneity is the fact that the n-year transition matrix is simply given by the nth power of the one-year transition matrix,

$$M_n = M_1^n,$$

where again the cumulative nth year default probabilities for rating classes are given by the last column of M_n.

Properties (i), (ii) make M a *stochastic matrix*. Furthermore, one might want to impose the following plausibility constraints to reflect our intuition.

(v) Low-risk states should never show a higher default probability than high-risk states, i.e., $M_{i8} \le M_{i+1\,8}, \quad i = 1, \ldots, 7$.

(vi) It should be more likely to migrate to closer states than to more distant states (row monotony towards the diagonal),

$$M_{ii+1} \ge M_{ii+2} \ge M_{ii+3} \ldots$$

$$M_{ii-1} \ge M_{ii-2} \ge M_{ii-3} \ldots$$

(vii) The chance of migration into a certain rating class should be greater for more closely adjacent rating categories (column monotony towards the diagonal).

$$M_{i+1i} \geq M_{i+2i} \geq M_{i+3i} \ldots$$

$$M_{i-1i} \geq M_{i-2i} \geq M_{i-3i} \ldots$$

Insofar as a lower rating presents a higher credit risk, JARROW et al. [103] formulated the condition:

(viii) $\sum_{j \geq k} m_{ij}$ is a nondecreasing function of i for every fixed k,

which is equivalent to requiring that the underlying Markov chain be *stochastically monotonic*. Note that row and column monotony towards the diagonal (properties (vi) and (vii)) implies stochastic monotony but not vice versa.

The problem with this wish list is that one cannot expect these properties to be satisfied by transition matrices sampled from historical data; so, the question remains how to best match a transition matrix to sampled data but still fulfill the required properties. ONG [150] proposes to solve this optimization problem, with the plausibility constraints stated as "soft conditions," through a simulated-annealing approach, where perturbed matrices are produced through additional random terms and tested to find an optimal solution. At this point we do not want to dive into the vast world of multidimensional optimization algorithms, but rather turn to another approach for obtaining a suitable migration matrix, namely via generators.

Generator Matrix The shortest time interval from which a transition matrix is estimated is typically one year. Data quality of rating transition observations within a shorter period is too poor to allow for a reliable estimate of a migration matrix. Nevertheless, for valuation purposes or loans that allow for the possibility of under-year intervention, we are interested in transition matrices for shorter time periods. One might be tempted to approach this problem by fractional powers of M, but unfortunately the roots of transition matrices are not stochastic in general nor is it clear which root to choose when more of them exist. The idea is now to try to embed the time-discrete Markov chain in a time-continuous Markov process, the latter being totally controlled by its generator. Collecting the rates of migration in a matrix Q, a time-dependent transition matrix $M(t)$, $t \geq 0$ then satisfies the matrix-valued (backward) differential equation

$$dM(t) = QM(t)dt.$$

Under the boundary condition $M(0) = I$, where I is the identity matrix, the formal solution of the differential equation is the matrix exponential

$$M(t) = e^{tQ} = \sum_{k=0}^{\infty} \frac{(tQ)^k}{k!}. \tag{6.8}$$

Furthermore, the following theorem holds [147]:

6.3.1 Theorem *$M(t)$ defined by (6.8) is a stochastic matrix for all $t \geq 0$ if and only if $Q = (q_{ij})$ satisfies the following properties:*

(i) $0 \leq -q_{ii} < \infty$ *for all* $i = 1, ..., 8$;

(ii) $q_{ij} \geq 0$ *for all* $i \neq j$;

(iii) $\sum_{j=1}^{8} q_{ij} = 0$ *for all* $i = 1, ..., 8$.

In Markov chain theory such matrices are called *Q-matrices* or *generators*. Unfortunately, this theorem is not that much of a help. Since we only have a single-period transition matrix available, the existence of the true generator is not necessarily guaranteed. The problem of finding generators for empirical transition matrices has been comprehensively treated by Israel et al. [99]. They rediscovered some known results and derived some new findings. In the following we freely cite the most useful ones for our purposes and refer for proofs to the literature.

Equation (6.8), nevertheless, can give us some guidance on how to find a valid, or at least construct an approximate, generator, i.e., the matrix logarithm.

6.3.2 Theorem ([99]) *Let $M = (m_{ij})$ be an $n \times n$ strictly diagonally dominant Markov transition matrix, i.e., $m_{ii} > 1/2$ for all i. Then the series*

$$\tilde{Q} = \sum_{k=1}^{l} (-1)^{k+1} \frac{(M-I)^k}{k}$$

converges geometrically quickly for $l \to \infty$, and gives rise to an $n \times n$ matrix $\tilde{Q}$ having row-sums zero, such that $\exp(\tilde{Q}) = M$ exactly.

Note that the condition of strictly diagonal dominance is only a sufficient one. It is usually satisfied by credit migration matrices. For theorems on the (non-)existence of true generators see [99] and the references therein. The main problem of the log-expansion is that the matrix $\tilde{Q}$ is not guaranteed to have nonnegative off-diagonal entries, which we need by the first theorem. However, any negative off-diagonal entries of $\tilde{Q}$ will usually be quite small. Therefore, we try to correct the matrix simply by replacing these negative entries by zero, and redistribute the values by some appropriate ad hoc rules to the other entries to preserve the property of having vanishing row sum, in the hope that the thus obtained Q-matrix yields an, in some sense close, embedding.

One version is to define a Q-matrix Q from $\tilde{Q}$ as (see also Stromquist [173])

$$q_{ij} = \max(\tilde{q}_{ij}, 0), \quad i \neq j; \qquad q_{ii} = \tilde{q}_{ii} + \sum_{i \neq j} \min(\tilde{q}_{ij}, 0), \tag{6.9}$$

i.e., the sum of the negative off-diagonal entries is allotted in full to the diagonal element of the respective row.

A different Q-matrix is obtained by adding the negative values back to all entries of the same row that have the correct sign, proportional to their absolute values (see also Araten [9] for a closely related algorithm), i.e., let

$$g_i = |\tilde{q}_{ii}| + \sum_{i \neq j} \max(\tilde{q}_{ij}, 0); \qquad b_i = \sum_{i \neq j} \max(-\tilde{q}_{ij}, 0)$$

and set

$$\hat{q}_{ij} = \begin{cases} 0, & i \neq j \text{ and } \tilde{q}_{ij} < 0 \\ \tilde{q}_{ij} - b_i |\tilde{q}_{ij}| / g_i, & \text{otherwise if } g_i > 0 \\ \tilde{q}_{ij}, & \text{otherwise if } g_i = 0. \end{cases} \tag{6.10}$$

In both cases the new matrix will still have by construction row sum zero, but now with nonnegative off-diagonals. Clearly, it will no longer satisfy $\exp(Q) = M$. Other, more refined "redistribution" choices are conceivable; however, they would rarely induce substantial difference to the distance of $\exp(Q)$ and M. Note that it is possible that a valid generator exists even if the $\tilde{Q}$ computed by the series expansion is not a valid one. Furthermore, it may be possible that there exist more than one valid generator for a given transition matrix M.

Assuming that there is at most one migration per year JARROW et al. [103] derived the following formula for a suitable generator:

$$\check{q}_{ii} = \log(m_{ii}), \qquad \check{q}_{ij} = m_{ij}\log(m_{ii})/(m_{ii} - 1) \quad \text{for } i \neq j. \tag{6.11}$$

Let us now consider some examples. From the Moody's one-year transition matrix $M_{Moody's}$ we arrive with Equation (6.11) at the Q-matrix

$$\check{Q}_{Moody's} = \begin{pmatrix} -0.1150 & 0.1035 & 0.0112 & 0.0000 & 0.0003 & 0.0000 & 0.0000 & 0.0000 \\ 0.0121 & -0.1150 & 0.0979 & 0.0034 & 0.0012 & 0.0001 & 0.0000 & 0.0003 \\ 0.0007 & 0.0312 & -0.1023 & 0.0611 & 0.0072 & 0.0019 & 0.0001 & 0.0001 \\ 0.0007 & 0.0039 & 0.0757 & -0.1570 & 0.0629 & 0.0111 & 0.0009 & 0.0018 \\ 0.0004 & 0.0007 & 0.0065 & 0.0655 & -0.1935 & 0.0982 & 0.0064 & 0.0158 \\ 0.0001 & 0.0005 & 0.0024 & 0.0067 & 0.0707 & -0.1930 & 0.0362 & 0.0765 \\ 0.0000 & 0.0000 & 0.0000 & 0.0119 & 0.0358 & 0.0771 & -0.4722 & 0.3473 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}$$

with the matrix exponential

$$\exp(\check{Q}_{Moody's}) = \begin{pmatrix} 0.8919 & 0.0925 & 0.0146 & 0.0006 & 0.0004 & 0.0000 & 0.0000 & 0.0000 \\ 0.0108 & 0.8933 & 0.0882 & 0.0057 & 0.0015 & 0.0003 & 0.0000 & 0.0003 \\ 0.0008 & 0.0282 & 0.9063 & 0.0540 & 0.0080 & 0.0023 & 0.0002 & 0.0003 \\ 0.0007 & 0.0045 & 0.0670 & 0.8585 & 0.0535 & 0.0120 & 0.0010 & 0.0028 \\ 0.0003 & 0.0009 & 0.0079 & 0.0556 & 0.8288 & 0.0816 & 0.0060 & 0.0189 \\ 0.0001 & 0.0005 & 0.0026 & 0.0078 & 0.0591 & 0.8284 & 0.0262 & 0.0753 \\ 0.0000 & 0.0001 & 0.0006 & 0.0099 & 0.0282 & 0.0570 & 0.6247 & 0.2797 \\ 0.0000 & 0.0000 & 0.0000 & 0.0000 & 0.0000 & 0.0000 & 0.0000 & 1.0000 \end{pmatrix}.$$

Using instead Equation (6.10) we obtain

$$\hat{Q}_{Moody's} = \begin{pmatrix} -0.1159 & 0.1095 & 0.0061 & 0.0000 & 0.0003 & 0.0000 & 0.0000 & 0.0000 \\ 0.0128 & -0.1175 & 0.1032 & 0.0002 & 0.0009 & 0.0000 & 0.0000 & 0.0003 \\ 0.0005 & 0.0330 & -0.1066 & 0.0660 & 0.0056 & 0.0014 & 0.0001 & 0.0000 \\ 0.0007 & 0.0027 & 0.0797 & -0.1621 & 0.0690 & 0.0084 & 0.0007 & 0.0009 \\ 0.0003 & 0.0005 & 0.0039 & 0.0708 & -0.2003 & 0.1082 & 0.0054 & 0.0111 \\ 0.0001 & 0.0004 & 0.0021 & 0.0042 & 0.0773 & -0.1990 & 0.0456 & 0.0692 \\ 0.0000 & 0.0000 & 0.0000 & 0.0114 & 0.0359 & 0.0834 & -0.4748 & 0.3440 \\ 0.0000 & 0.0000 & 0.0000 & 0.0000 & 0.0000 & 0.0000 & 0.0000 & 0.0000 \end{pmatrix}$$

with

$$\exp(\hat{Q}_{Moody's}) = \begin{pmatrix} 0.8912 & 0.0976 & 0.0105 & 0.0003 & 0.0003 & 0.0000 & 0.0000 & 0.0000 \\ 0.0114 & 0.8913 & 0.0925 & 0.0032 & 0.0011 & 0.0001 & 0.0000 & 0.0003 \\ 0.0006 & 0.0297 & 0.9028 & 0.0581 & 0.0069 & 0.0018 & 0.0001 & 0.0001 \\ 0.0006 & 0.0036 & 0.0701 & 0.8547 & 0.0582 & 0.0102 & 0.0008 & 0.0017 \\ 0.0003 & 0.0007 & 0.0059 & 0.0596 & 0.8241 & 0.0893 & 0.0058 & 0.0144 \\ 0.0001 & 0.0004 & 0.0022 & 0.0061 & 0.0643 & 0.8244 & 0.0329 & 0.0696 \\ 0.0000 & 0.0000 & 0.0005 & 0.0095 & 0.0285 & 0.0614 & 0.6234 & 0.2766 \\ 0.0000 & 0.0000 & 0.0000 & 0.0000 & 0.0000 & 0.0000 & 0.0000 & 1.0000 \end{pmatrix}.$$

Computing the L^1-norm[2] yields $||M_{Moody's} - \exp(\check{Q}_{Moody's})||_1 = 0.10373$ and $||M_{Moody's} - \exp(\hat{Q}_{Moody's})||_1 = 0.00206$, i.e., the generator obtained from the log series expansion seems to be a better approximation in this case. Note that some, but not all entries violating the monotony conditions have been smoothed out.

Kealhofer et al. [114] question that rating changes are a good indicator for credit quality changes. In particular, they claim that rating agencies are too slow in changing ratings and therefore the probability of staying in a given grade overstates the true probability of keeping approximately the same credit quality. Suppose firms are classified according to KMV's respective expected default frequencies (EDF), based upon non-overlapping ranges of default probabilities. Each of these ranges corresponds then to a rating class, i.e., firms with default rates less than or equal to 0.002% are mapped to AAA, 0.002% to 0.04% corresponds to AA, etc. The historical frequencies of changes from one range to another are estimated from the history of changes in default rates as measured by EDFs. This yields the following KMV one-year transition matrix.

$$M_{KMV} = \begin{pmatrix} 0.6626 & 0.2222 & 0.0737 & 0.0245 & 0.0086 & 0.0067 & 0.0015 & 0.0002 \\ 0.2166 & 0.4304 & 0.2583 & 0.0656 & 0.0199 & 0.0068 & 0.0020 & 0.0004 \\ 0.0276 & 0.2034 & 0.4419 & 0.2294 & 0.0742 & 0.0197 & 0.0028 & 0.0010 \\ 0.0030 & 0.0280 & 0.2263 & 0.4254 & 0.2352 & 0.0695 & 0.0100 & 0.0026 \\ 0.0008 & 0.0024 & 0.0369 & 0.2293 & 0.4441 & 0.2453 & 0.0341 & 0.0071 \\ 0.0001 & 0.0005 & 0.0039 & 0.0348 & 0.2047 & 0.5300 & 0.2059 & 0.0201 \\ 0.0000 & 0.0001 & 0.0009 & 0.0026 & 0.0179 & 0.1777 & 0.6995 & 0.1013 \\ 0.0000 & 0.0000 & 0.0000 & 0.0000 & 0.0000 & 0.0000 & 0.0000 & 1.0000 \end{pmatrix}.$$

[2]The L^1-norm of a matrix M is defined as $||M||_1 = \sum_{i,j} |m_{ij}|$.

Estimating an appropriate generator according to Equation (6.10) yields

$$\hat{Q}_{KMV} =$$
$$\begin{pmatrix} -0.4941 & 0.4512 & 0.0043 & 0.0290 & 0.0006 & 0.0087 & 0.0002 & 0.0001 \\ 0.4503 & -1.1663 & 0.6945 & 0.0000 & 0.0180 & 0.0009 & 0.0025 & 0.0001 \\ 0.0000 & 0.5813 & -1.2204 & 0.6237 & 0.0025 & 0.0122 & 0.0000 & 0.0007 \\ 0.0217 & 0.0000 & 0.6583 & -1.3037 & 0.6119 & 0.0000 & 0.0109 & 0.0009 \\ 0.0000 & 0.0314 & 0.0000 & 0.6355 & -1.2157 & 0.5418 & 0.0000 & 0.0070 \\ 0.0013 & 0.0000 & 0.0235 & 0.0000 & 0.4805 & -0.8571 & 0.3463 & 0.0054 \\ 0.0000 & 0.0010 & 0.0000 & 0.0118 & 0.0000 & 0.3004 & -0.4276 & 0.1145 \\ 0.0000 & 0.0000 & 0.0000 & 0.0000 & 0.0000 & 0.0000 & 0.0000 & 0.0000 \end{pmatrix}$$

with $||M_{KMV} - \exp(\hat{Q}_{KMV})||_1 = 0.5717$, whereas Equation (6.11) gives

$$\check{Q}_{KMV} =$$
$$\begin{pmatrix} -0.4116 & 0.2711 & 0.0899 & 0.0299 & 0.0105 & 0.0082 & 0.0018 & 0.0002 \\ 0.3206 & -0.8430 & 0.3823 & 0.0971 & 0.0295 & 0.0101 & 0.0030 & 0.0006 \\ 0.0404 & 0.2976 & -0.8167 & 0.3357 & 0.1086 & 0.0288 & 0.0041 & 0.0015 \\ 0.0045 & 0.0417 & 0.3366 & -0.8547 & 0.3499 & 0.1034 & 0.0149 & 0.0039 \\ 0.0012 & 0.0035 & 0.0539 & 0.3348 & -0.8117 & 0.3582 & 0.0498 & 0.0104 \\ 0.0001 & 0.0007 & 0.0053 & 0.0470 & 0.2765 & -0.6349 & 0.2781 & 0.0272 \\ 0.0000 & 0.0001 & 0.0011 & 0.0031 & 0.0213 & 0.2113 & -0.3574 & 0.1205 \\ 0.0000 & 0.0000 & 0.0000 & 0.0000 & 0.0000 & 0.0000 & 0.0000 & 0.0000 \end{pmatrix}$$

with $||M_{KMV} - \exp(\check{Q}_{KMV})||_1 = 1.289$, again showing that the log-expansion approximation is superior in this case. We therefore conclude that, despite the awkward ad hoc transformation from $\tilde{Q}$ to $\check{Q}$ resp. $\hat{Q}$, the L^1-error based on a log-expansion is smaller than the error based on the method proposed by Jarrow et al. for the given examples.

Adjustment of the default column Clearly, any approximate generator results in a more or less modified default column. Additionally, one might have some exogenously given default master scale and still want to represent the dynamics of rating migration by a given Markov generator. The following property allows adjusting the generator appropriately.

6.3.3 Proposition *Let Q be a 8×8 generator matrix and define $\Lambda \in \mathbb{R}^{8\times 8}$ as*

$$(\Lambda)_{i,j} = \begin{cases} 0 & \text{if } i \neq j \\ \lambda_i > 0 & \text{if } i = j \end{cases}.$$

Then ΛQ is again a generator matrix, i.e., row scaling by constant positive factors is a closed operation in the space of Q-matrices.

The proof is obvious by the properties of Q-matrices. We can now use this proposition to successively adjust the generator to reproduce a given default column according to the following algorithm:

1. Choose $\Lambda^{(1)}$ with $\lambda_1 > 0$ and $\lambda_{i\neq 1} = 1$ such that $\big(\exp\big(\Lambda^{(1)}Q\big)\big)_{1,8} = m_{1,8}$.

2. Choose $\Lambda^{(2)}$ with $\lambda_2 > 0$ and $\lambda_{i\neq 2} = 1$ such that $\big(\exp\big(\Lambda^{(2)}\Lambda^{(1)}Q\big)\big)_{2,8} = m_{2,8}$.

$$\vdots \qquad \vdots \qquad \vdots \qquad \vdots \qquad \vdots$$

7. Choose $\Lambda^{(7)}$ with $\lambda_7 > 0$ and $\lambda_{i\neq 7} = 1$ such that $\big(\exp\big(\Lambda^{(7)}\cdot\dots\cdot\Lambda^{(1)}Q\big)\big)_{7,8} = m_{7,8}$.

8. Scaling a row of a Q-matrix has some impact on every single entry in the exponential of that matrix. This means, for example, that after we achieved the right AAA-PD (step 1), step 2 produces the right AA-PD but slightly changes the just calibrated AAA-PD. Therefore one has to repeat steps 1-7 until the default column of the respective matrix exponential agrees with the default column of M within some error bars.

The factors λ_i as described in the algorithm can be found by a simple trial-and-error method, e.g., using dyadic approximation. The above algorithm converges due to the fact that the mappings $\lambda \mapsto \exp(\Lambda^{(i)}Q)$, $i = 1, ..., 7$, are continuous. This follows from the power series representation (6.8). Furthermore, from (6.8) follows that $\exp(D_{i,\lambda}Q) \approx I + \Lambda^{(i)}Q$ indicating that multiplication of the i-th row of Q by a factor λ_i mainly affects the i-th row of the corresponding matrix exponential. For example, starting from $\hat{Q}_{KMV}$ the modified adjusted generator $\bar{Q}$ reads

$$\bar{Q}_{KMV} =$$
$$\begin{pmatrix} -0.4965 & 0.4535 & 0.0043 & 0.0291 & 0.0006 & 0.0087 & 0.0002 & 0.0001 \\ 0.4151 & -1.0752 & 0.6403 & 0.0000 & 0.0166 & 0.0008 & 0.0023 & 0.0001 \\ 0.0000 & 0.5764 & -1.2100 & 0.6184 & 0.0025 & 0.0121 & 0.0000 & 0.0007 \\ 0.0239 & 0.0000 & 0.7253 & -1.4364 & 0.6742 & 0.0000 & 0.0120 & 0.0010 \\ 0.0000 & 0.0273 & 0.0000 & 0.5526 & -1.0571 & 0.4711 & 0.0000 & 0.0061 \\ 0.0014 & 0.0000 & 0.0253 & 0.0000 & 0.5180 & -0.9239 & 0.3733 & 0.0058 \\ 0.0000 & 0.0010 & 0.0000 & 0.0127 & 0.0000 & 0.3243 & -0.4616 & 0.1235 \\ 0.0000 & 0.0000 & 0.0000 & 0.0000 & 0.0000 & 0.0000 & 0.0000 & 0.0000 \end{pmatrix}$$

with corresponding matrix exponential

$$
\exp(\bar{Q}_{KMV}) =
\begin{pmatrix}
0.6587 & 0.2290 & 0.0693 & 0.0256 & 0.0093 & 0.0063 & 0.0016 & 0.0002 \\
0.2090 & 0.4482 & 0.2420 & 0.0688 & 0.0230 & 0.0064 & 0.0023 & 0.0004 \\
0.0548 & 0.2177 & 0.4301 & 0.2025 & 0.0727 & 0.0171 & 0.0041 & 0.0010 \\
0.0224 & 0.0736 & 0.2378 & 0.3576 & 0.2333 & 0.0589 & 0.0138 & 0.0026 \\
0.0070 & 0.0249 & 0.0716 & 0.1915 & 0.4575 & 0.1974 & 0.0430 & 0.0071 \\
0.0023 & 0.0077 & 0.0232 & 0.0546 & 0.2173 & 0.4754 & 0.1993 & 0.0201 \\
0.0005 & 0.0017 & 0.0050 & 0.0125 & 0.0415 & 0.1732 & 0.6642 & 0.1013 \\
0.0000 & 0.0000 & 0.0000 & 0.0000 & 0.0000 & 0.0000 & 0.0000 & 1.0000
\end{pmatrix}
$$

and $||M_{KMV} - \exp(\bar{Q}_{KMV})||_1 = 0.6855$.

Remark Before closing this section we briefly mention a theorem on the non-existence of a valid generator.

6.3.4 Theorem ([99]) *Let M be a transition matrix and suppose that either*

(i) $\det(M) \leq 0$*; or*

(ii) $\det(M) > \prod_i m_{ii}$*; or*

(iii) there exist states i and j such that j is accessible from i, i.e., there is a sequence of states $k_0 = i, k_1, k_2, \ldots, k_m = j$ such that $m_{k_l k_{l+1}} > 0$ for each l, but $m_{ij} = 0$.

Then there does not exist an exact generator.

Strictly diagonal dominance of M implies $\det(M) > 0$; so, part (i) does usually not apply for credit migration matrices (for a proof, see references). But case (iii) is quite often observed with empirical matrices. For example, $M_{Moody's}$ has zero Aaa default probability, but a transition sequence from Aaa to D is possible. Note that if we adjust a generator to a default column with some vanishing entries the respective states become trapped states due to the above theorem ($\exp(\check{Q}_{Moody's})$ and $\exp(\hat{Q}_{Moody's})$ are only accurate to four decimals), i.e., states with zero default probability and an underlying Markov process dynamic are irreconcilable with the general ideas of credit migration with default as the only trapped state.

Remark Strictly diagonal dominance is a necessary prerequisite for the logarithmic power series of the transition matrix to converge [99]. Now, the default state being the only absorbing state, any transition matrix M risen to the power of some $t > 1$, M^t, loses the property of diagonal dominance, since in the limit $t \to \infty$ only the default state is populated, i.e.,

$$M(t) = M^t \to \begin{pmatrix} 0 & \dots & 0 & 1 \\ \dots & \dots & \dots & \dots \\ 0 & \dots & 0 & 1 \end{pmatrix} \quad \text{as} \quad t \to \infty,$$

which is clearly not strictly diagonally dominant. Kreinin and Sidelnikova [116] proposed regularization algorithms for matrix roots and generators that do not rely on the property of diagonal dominance. These algorithms are robust and computationally efficient, but in the time-continuous case are only slightly advantageous when compared to the weighted adjustment. In the time-discrete case, i.e., transition matrices as matrix-roots, their method seems to be superior for the given examples to other known regularization algorithms.

6.3.4 A Non-Homogeneous Markov Chain Approach

In a recent paper [23] and also in the book [24], Section 2.3.2, a continuous-time Markov chain approach without making the assumption of time-homogeneity is presented. Calculations with rating agency data (see [24], Figure 2.11) nicely illustrate that the non-homogeneous continuous-time Markov chain (NHCTMC) approach results in an almost perfectly fitted model. We do not repeat the calculations from [23] and [24] here but want to briefly explain the NHCTMC approach in principal so that readers can do their own experiments.

The starting point for our construction is again a generator $Q = (q_{ij})_{1 \le i,j \le 8}$. But now we no longer assume that the transition rates q_{ij} are constant over time but replace the time-homogeneous generator Q leading to migration matrices $\exp(tQ)$ for time intervals $[0, t]$ by a time-dependent generator

$$Q_t \;=\; \Phi_t * Q \tag{6.12}$$

where $*$ denotes matrix multiplication and $\Phi_t = (\varphi_{ij}(t))_{1 \le i,j \le 8}$ is a diagonal matrix in $\mathbb{R}^{8 \times 8}$ with

$$\varphi_{ij}(t) \;=\; \begin{cases} 0 & \text{if } i \neq j \\ \varphi_{\alpha_i,\beta_i}(t) & \text{if } i = j \end{cases} \tag{6.13}$$

To see that Q_t is a generator, note that Φ_t is a diagonal matrix, and that scaling the rows of a Q-matrix leads to a Q-matrix again. The functions $\varphi_{\alpha,\beta}$ with respect to parameters α and β are defined in [23] as follows. Set

$$\varphi_{\alpha,\beta} \;:\; [0,\infty) \to [0,\infty),\; t \;\mapsto\; \varphi_{\alpha,\beta}(t) \;=\; \frac{(1-e^{-\alpha t})t^{\beta-1}}{1-e^{-\alpha}}$$

for *nonnegative* constants α and β.

Let us summarize some basic properties of the functions $\varphi_{\alpha,\beta}(t)$.

1. $\varphi_{\alpha,\beta}(1) = 1$ ("normalized" functions).
2. $t \mapsto t\varphi_{\alpha,\beta}(t)$ is increasing in the time parameter $t \geq 0$.
3. The first part of $t\varphi_{\alpha,\beta}$, namely $(1-e^{-\alpha t})$, is the distribution function of an exponentially distributed random variable with intensity α. The second part of $t\varphi_{\alpha,\beta}$, namely t^{β}, can be considered as a time-slowing-down ($\beta < 1$) or time-accelerating ($\beta > 1$) adjustment term. The scaling factor $(1-e^{-\alpha})^{-1}$ is the normalizing multiplier forcing $\varphi_{\alpha,\beta}(1) = 1$.

In the sequel we denote rating letters by R. For instance, $row(R)$ for $R = AA$ refers to the second row in a migration matrix.

In analogous way as we did before we can now define migration matrices for given time periods $[0,t]$ via

$$M_t \;=\; \exp(tQ_t) \qquad (t \geq 0). \tag{6.14}$$

The corresponding PD term structures are defined by

$$\big((M_t)_{row(R),8}\big)_{t\geq 0;\, R=AAA,AA,...,CCC}\,.$$

Since the functional form of $(Q_t)_{t\geq 0}$ is given by Equation (6.13), generators Q_t are solely determined by two vectors $(\alpha_1,...,\alpha_8)$ and $(\beta_1,...,\beta_8)$ in $[0,\infty)^8$. We can now try to find α- and β-vectors such that

$$\text{distance}\left[(\hat{p}_R^{(t)})_{t;R},\; \big((M_t)_{row(R),8}\big)_{t;R}\right] \;\stackrel{!}{=}\; \text{small}. \tag{6.15}$$

Here, $(\hat{p}_R^{(t)})_{t;R}$ denotes the time series of historically observed (empirical) multi-year default frequencies. The optimization task for a good fit is then to find α- and β-vectors such that the distance or approximation error (6.15) attains its minimum value. Note that α_8 and β_8 have no meaning in the approach and can be fixed at some arbitrary value, e.g., $\alpha = \beta = 1$. As distance measure for the optimization problem (6.15) we use the mean-squared distance.

A major difference between homogeneous and non-homogeneous approaches is that the homogeneous approach relies on *observed migration rates only* whereas the non-homogeneous approach relies on *observed migration rates* for the calibration of the generator Q *as well as on observed multi-year default frequencies* for finding α- and β-vectors such that the NHCTMC approach best possible approximates observed multi-year default frequencies.

Arguments from probability theory, why the NHCTMC approach really makes sense can be found in [24], Section 2.3.2. There it is proven that the approach can be embedded in well understood general Markov process theory. The overall conclusion is that *Markov chain techniques can yield excellently fitted models if the very much restricting assumption of time-homogeneity is dropped.*

6.4 Term Structure Based on Market Spreads

Alternatively, we can construct an implied default term structure by using market observable information, such as asset swap spreads or defaultable bond prices.

This approach is commonly used in credit derivative pricing. The extracted default probabilities reflect the market agreed perception today about the future default tendency of the underlying credit; they are by construction risk-neutral probabilities. Yet, in some sense, market spread data presents a classic example of a joint observation problem. Credit spreads imply loss severity given default, but this can only be derived if one is prepared to make an assumption as to what they are simultaneously implying about default likelihoods (or vice versa).

In practice, one usually makes exogenous assumptions on the recovery rate, based on the security's seniority. In any credit-linked product the primary risk lies in the potential default of the reference entity:

absent any default in the reference entity, the expected cash flow will be received in full, whereas if a default event occurs the investor will receive some recovery amount. It is therefore natural to model a risky cash flow as a portfolio of contingent cash flows corresponding to the different scenarios weighted by the probability of these scenarios.

The time origin, $t = 0$, is chosen to be the current date and our time frame is $[0, T]$, i.e., we have market observables for comparison up to time T. Furthermore, assume that the event of default and the default-free discount factor are statistically independent. Then the present value for a risky payment X promised for time t (assuming no recovery) equals

$$B(0, t)S(t)X$$

where $B(0, t)$ is the risk-free discount factor (zero bond prices) and $S(t)$ as usual the cumulative survival probability as of today. Consider a credit bond from an issuer with notional V, fixed coupon c, and maturity T_n, and let the accrual dates for the promised payments be $0 \leq T_1 < T_1 < \cdots < T_n$. We assume that the coupon of the bond to be paid at time T_i is $c\Delta_i$ where Δ_i is the day count fraction for period $[T_{i-1}, T_i]$ according to the given day count convention. When the recovery rate REC is nonzero, it is necessary to make an assumption about the claim made by the bond holders in the event of default.

Jarrow and Turnbull [104] and Hull and White [96] assume that the claim equals the no-default value of the bond. In this case value additivity is given, i.e., the value of the coupon-bearing bond is the sum of the values of the underlying zero bonds. Duffie and Singleton [48] assume that the claim is equal to the value of the bond immediately prior to default. In [97], Hull and White advocate that the best assumption is that the claim made in the event of default equals the face value of the bond plus accrued interests. Whilst this is more consistent with the observed clustering of asset prices during default it makes splitting a bond into a portfolio of risky zeros much harder, and value additivity is no longer satisfied. Here, we define recovery as a fraction of par and suppose that recovery rate is exogenously given (a refinement of this definition is made in Chapter 7), based on the seniority and rating of the bond, and the industry of the corporation. Obviously, in case of default all future coupons are lost.

The net present value of the payments of the risky bond, i.e., the *dirty price*, is then given as

$$\begin{aligned}\text{dirty price} = &\sum_{T_i>0} B(0,T_i)\Delta_i S(T_i) + \\ &+V\left[B(0,T_n)S(T_n) + REC\int_0^T B(0,t)F(dt)\right].\end{aligned} \tag{6.16}$$

The interpretation of the integral is just the recovery payment times the discount factor for time t times the probability to default around t summed up from time zero to maturity.

Similarly, for a classic default swap we have spread payments $\Delta_i s$ at time T_i where s is the spread, provided that there is no default until time T_i. If the market quotes the fair default spread s the present value of the spread payments and the event premium $V(1-REC)$ cancel each other:

$$0 = \sum_{i=1}^{n} B(0,T_i)s\Delta_i S(T_i) - V(1-REC)\int_0^{T_n} B(0,t)F(dt). \tag{6.17}$$

Given a set of fair default spreads or bond prices (but the bonds have to be from the same credit quality) with different maturities and a given recovery rate, one now has to back out the credit curve. To this end we have to specify also a riskless discount curve $B(0,t)$ and an interpolation method for the curve, since it is usually not easy to get a smooth default curve out of market prices. In the following we briefly sketch one method:

Fitting a credit curve Assuming that default is modeled as the first arrival time of a Poisson process we begin by supposing that the respective hazard rate is constant over time. Equations (6.16) and (6.17), together with Equation (6.2) $S(t) = e^{-\int_0^t h(s)ds} = e^{-ht}$, allow us then to back out the hazard rate from market observed bond prices or default spreads. If there are several bond prices or default spreads available for a single name one could in principle extract a term structure of a piece-wise constant hazard rate. In practice, this might lead to inconsistencies due to data and model errors. So, a slightly more sophisticated but still parsimonious model is obtained by assuming a time-varying, but deterministic default intensity $h(t)$. Suppose, for example, that $\int_0^t h(s)ds = \Phi(t)\cdot t$, where the function $\Phi(t)$ captures term structure effects. An interesting candidate for the fit function Φ is the Nelson-Siegel [145] yield curve function:

$$\Phi(t) = a_0 + (a_1 + a_2)\left(\frac{1-\exp(-t/a_3)}{t/a_3}\right) - a_2 \exp\left(-t/a_3\right) . \quad (6.18)$$

This function is able to generate smooth upward sloping, humped and downward sloping default intensity curves with a small number of parameters, and, indeed, we have seen in Figure 6.2 that investment grade bonds tend to have a slowly upward sloping term structure whereas those of speculative grade bonds tend to be downward sloping. Equation (6.18) implies that the default intensity of a given issuer tends towards a long-term mean. Other functions like cubic or exponential spline may also be used in Equation (6.18), although they might lead to fitting problems due to their greater flexibility and the frequency of data errors. The parameter a_0 denotes the long-term mean of the default intensity, whereas a_1 represents its current deviation from the mean. Specifically, a positive value of a_1 implies a downward sloping intensity and a negative value implies an upward sloping term structure. The reversion rate towards the long-term mean is negatively related to $a_3 > 0$. Any hump in the term structure is generated by a nonzero a_2. However, in practice, allowing for a hump may yield implausible term structures due to overfitting. Thus, it is assumed that $a_2 = 0$, and the remaining parameters $\{a_0, a_1, a_3\}$ are estimated from data. The Nelson-Siegel function can yield negative default intensities if the bonds are more liquid or less risky than the default-free benchmark, or if there are data errors.

Using Equations (6.2) and (6.18) the survival function $S(t)$ can then be written as

$$S(t) = \exp\left[-\left(a_0 + a_1\left(\frac{1-\exp(-t/a_3)}{t/a_3}\right)\right)\cdot t\right] . \quad (6.19)$$

Now, we construct default curves from reference bond and default swap prices as follows: Consider a sample of N constituents which can be either bonds or swaps or both. To obtain the values of the parameters of the default intensity curve, $\{a_0, a_1, a_3\}$, we fit equations (6.16, 6.17), and with the help of Equation (6.19), to the market observed prices by use of a nonlinear optimization algorithm under the constraints $a_3 > 0$, $S(0) = 1$, and $S(t) - S(t+1) \geq 0$. Mean-Absolute-Deviation regression seems to be more suitable than Least-Square regression since the former is less sensitive to outliers.

Risk-neutral approach in asset value models (In this section we follow to some extent the presentation in Crouhy et al. [38], 9.5 and 9.6.) Under the Merton-style approach, the actual cumulative default probability from time 0 to time t in a real, risk averse world (cf. Chapter 3) is given by

$$\mathrm{PD}_t^{real} = N\left(-\frac{\log(A_0/C) + (\mu - \sigma^2/2)t}{\sigma\sqrt{t}}\right), \tag{6.20}$$

where A_0 is the market value of the firm's asset at time 0, C is the firm's default point, σ the asset volatility, and μ the expected return of the firm's assets. In a world where investors are neutral to risk, all assets should yield the same risk-free return r. So, the risk-neutral default probabilities are given as

$$\mathrm{PD}_t^{rn} = N\left(-\frac{\log(A_0/C) + (r - \sigma^2/2)t}{\sigma\sqrt{t}}\right), \tag{6.21}$$

where the expected return μ has been replaced by the continuous time risk-free interest rate r. Because investors refuse to hold risky assets with expected return less than the risk-free base rate, μ must be larger than r. It follows that

$$\mathrm{PD}_t^{rn} \geq \mathrm{PD}_t^{real}.$$

Substituting Equation (6.20) into Equation (6.21) and rearranging, we can write the risk-neutral default probability as:

$$\mathrm{PD}_t^{rn} = N\left(N^{-1}(\mathrm{PD}_t^{real}) + \frac{\mu - r}{\sigma}\sqrt{t}\right). \tag{6.22}$$

From the continuous time CAPM we have

$$\mu - r = \beta\pi \quad \text{with} \quad \beta = \frac{\mathrm{Cov}(r_a, r_m)}{\mathbb{V}(r_m)} = \rho_{a,m}\frac{\sigma_a}{\sigma_m}$$

as beta of the asset w.r.t. the market. Here, r_a and r_m denote the continuous time rate of return on the firm's asset and the market portfolio, σ_a and σ_m are the respective volatilities, $\rho_{a,m}$ denotes the correlation between the asset and the market return, and π means the market risk premium, defined by

$$\pi = \mu_m - r \ ,$$

where μ_m denotes the expected return on the market portfolio. Putting all together leads to

$$\mathrm{PD}_t^{rn} = N\left(N^{-1}(\mathrm{PD}_t^{real}) + \rho_{a,m}\frac{\pi}{\sigma_m}\sqrt{t}\right). \tag{6.23}$$

The correlation $\rho_{a,m}$ is estimated from a linear regression of the asset return against the market return. The market risk premium π is time varying, and is much more difficult to estimate statistically. Moody's KMV uses a slightly different mapping from distance-to-default to default probability than the normal distribution. For these reasons, the risk-neutral default probability is estimated by calibrating the market *Sharpe ratio*, $\mathrm{SR} = \pi/\sigma_m$ and some power θ, replacing the power 1/2 (for $\sqrt{\text{time}}$), in the following relation by means of bond data:

$$\mathrm{PD}_t^{rn} = N\left(N^{-1}(\mathrm{PD}_t^{real}) + \rho_{a,m}\,\mathrm{SR}\,t^{\theta}\right). \tag{6.24}$$

From (6.16) we obtain for the credit spread s of a risky zero bond

$$e^{-(r+s)t} = [(1-\mathrm{PD}_t^{rn}) + (1-\mathrm{LGD})\mathrm{PD}_t^{rn}]\, e^{-rt}. \tag{6.25}$$

Combining Equation (6.24) and Equation (6.25) yields

$$s \;=\; -\frac{1}{t}\log\left[1 - N(N^{-1}(\mathrm{PD}_t^{real}) + \rho_{a,m}\,\mathrm{SR}\,t^{\theta})\mathrm{LGD}\right],$$

which then serves to calibrate SR and θ in the least-square sense from market data.

Further Reading

The field of migration modeling is still developing. Very recently, Trueck and Rachev [180] published a collection of various topics in the context of rating-based credit risk modeling. The book contains a very comprehensive exposition on the theory of credit migrations and is highly recommended for readers seeking a complete discussion on the developments of the theory over the last decade. Another very interesting discussion can be found in a recent PhD thesis (in German) by Vogl [183]. There are many research papers in the market dealing with various approaches to credit migrations and PD term structures. Schuermann and Jafry [165, 166], Schuermann and Frydman [73], Höse et al. [93], Kadam and Lenk [108] and Trueck and Oezturkmen [179] are contributions to the field, just to mention a few out of many examples for recent research.

Chapter 7

Credit Derivatives

Credit derivatives are instruments that help banks, financial institutions, and debt security investors to manage their credit-sensitive investments. Credit derivatives *insure* and *protect* against adverse movements in the credit quality of the counterparty or borrower. For example, if a borrower defaults, the investor will suffer losses on the investment, but the losses can be offset by gains from the credit derivative transaction. One might ask why both banks and investors do not utilize the well-established insurance market for their protection. The major reasons are that credit derivatives offer lower transaction cost, quicker payment, and more liquidity. Credit default swaps, for instance, often pay out very soon after the event of default[1]; in contrast, insurances take much longer to pay out, and the value of the protection bought may be hard to determine. Finally, as with most financial derivatives initially invented for hedging, credit derivatives can now be traded speculatively. Like other over-the-counter derivative securities, credit derivatives are privately negotiated financial contracts. These contracts expose the user to operational, counterparty, liquidity, and legal risk. From the viewpoint of quantitative modeling we here are only concerned with *counterparty risk*. One can think of credit derivatives being placed somewhere between traditional credit insurance products and financial derivatives. Each of these areas has its own valuation methodology, but neither is wholly satisfactory for pricing credit derivatives. The insurance techniques make use of historical data, as, e.g., provided by rating agencies, as a basis for valuation (see Chapter 6).

This approach assumes that the future will be like the past, and does not take into account market information about credit quality. In contrast, derivative technology employs market information as a basis for valuation. Derivative securities pricing is based on the assumption

[1]Especially under the ISDA master agreement, cf. [98].

of *risk-neutrality* which assumes arbitrage-free and complete markets, but it is not clear whether these conditions hold for the credit market or not. If a credit event is based on a freely observable property of market prices, such as credit spreads, then we believe that conventional derivative pricing methodology may be applicable.

Credit derivatives are bilateral financial contracts that isolate specific aspects of credit risk from an underlying instrument and *transfer* that *risk* between two counterparties. By allowing credit risk to be freely traded, risk management becomes far more flexible. There are lots of different types of credit derivatives, but we shall only treat the most commonly used ones. They could be classified into two main categories according to valuation, namely the *replication* products, and the *default* products. The former are priced off the capacity to replicate the transaction in the money market, such as credit spread options. The latter are priced as a function of the exposure underlying the security, the default probability of the reference asset, and the expected recovery rate, such as credit default swaps. Another classification could be along their performance as protection-like products, such as credit default options and exchange-like products, such as total return swaps. In the next sections we describe the most commonly used credit derivatives and illustrate simple examples. For a more elaborate introduction to the different types of credit derivatives and their use for risk management see [100, 157]; for documentation and guidelines we refer to [98].

7.1 Total Return Swaps

A total return swap (TRS) [102, 142] is a mean of duplicating the cash flows of either selling or buying a reference asset, without necessarily possessing the asset itself. The TRS seller pays to the TRS buyer the total return of a specified asset and receives a floating rate payment plus a margin. The total return includes the sum of interest, fees, and any change in the value with respect to the reference asset, the latter being equal to any appreciation (positive) or depreciation (negative) in the market value of the reference security. Any net depreciation in value results in a payment to the TRS seller. The margin, paid by the TRS buyer, reflects the cost to the TRS seller of financing and servicing the

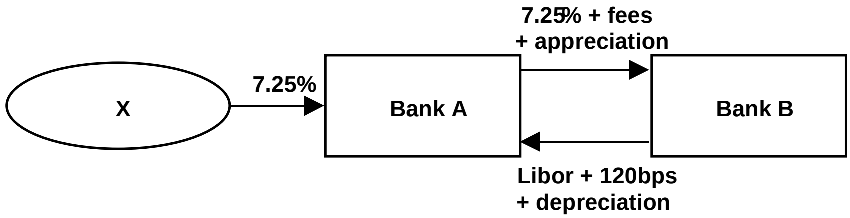


FIGURE 7.1: Total return swap.

reference asset on its own balance sheet. Such a transaction *transfers the entire economic benefit and risk* as well as the reference security to another counterparty.

A company may wish to sell an asset that it holds, but for tax or political reasons may be unable to do so. Likewise, it might hold a view that a specific asset is likely to depreciate in value in the near future, and wish to short it. However, not all assets in the market are easy to short in this way. Whatever the reason, the company would like to receive the cash flows which would result from selling the asset and investing the proceeds. This can be achieved exactly with a total return swap. Let us give an example: Bank A decides to get the economic effect of selling securities (bonds) issued by a German corporation, X. However, selling the bonds would have undesirable consequences, e.g., for tax reasons. Therefore, it agrees to swap with bank B the total return on one million 7.25% bonds maturing in December 2005 in return for a six-month payment of LIBOR plus 1.2% margin plus any decrease in the value of the bonds. Figure 7.1 illustrates the total return swap of this transaction.

Total return swaps are popular for many reasons and attractive to different market segments [102, 100, 157]. One of the most important features is the facility to obtain an almost unlimited amount of leverage. If there is no transfer of physical asset at all, then the notional amount on which the TRS is paid is unconstrained. Employing TRS, banks can diversify credit risk while maintaining confidentiality of their client's financial records. Moreover, total return swaps can also give investors access to previously unavailable market assets. For instance, if an investor can not be exposed to the Latin American market for various reasons, he or she is able to do so by doing a total return swap with a counterparty that has easy access to this market. Investors can also receive cash flows that duplicate the effect of holding an asset while keeping the actual assets away from their balance sheet. Furthermore, an institution can take advantage of another institution's back-office and documentation experience, and get cash flows that would otherwise require infrastructure, which it does not possess.

7.2 Credit Default Products

Credit default swaps [134] are bilateral contracts in which one counterparty pays a fee periodically, typically expressed in basis points on the notional amount, in return for a contingent payment by the protection seller following a credit event of a reference security. The credit event could be either default or downgrade; the credit event and the settlement mechanism used to determine the payment are flexible and negotiated between the counterparties. A TRS is importantly distinct from a CDS in that it exchanges the total economic performance of a specific asset for another cash flow. On the other hand, a credit default swap is triggered by a credit event. Another similar product is a credit default option. This is a binary put option that pays a fixed sum if and when a predetermined credit event (default/downgrade) happens in a given time.

Let us assume that bank A holds securities (swaps) of a low-graded firm X, say BB, and is worried about the possibility of the firm defaulting. Bank A pays to firm X floating rate (Libor) and receives fixed (5.5%). For protection bank A therefore purchases a credit default swap from bank B which promises to make a payment in the event of default. The fee reflects the probability of default of the reference asset, here the low-graded firm. Figure 7.2 illustrates the above transaction.

Credit default swaps are almost exclusively inter-professional transactions, and range in nominal size of reference assets from a few millions to billions of euros. Maturities usually run from one to ten years. The only true limitation is the willingness of the counterparties to act on a credit view. Credit default swaps allow users to reduce credit exposure without physically removing an asset from their balance sheet. Purchasing default protection via a CDS can hedge the credit exposure of

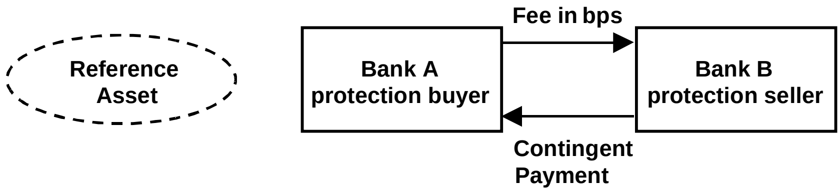


FIGURE 7.2: Credit default swap.

such a position without selling for either tax or accounting purposes. When an investor holds a credit-risky security, the return for assuming that risk is only the net spread earned after deducting the cost of funding. Since there is no up-front principal outlay required for most protection sellers when assuming a CDS position, they take on credit exposure in off-balance sheet positions that do not need to be funded. On the other hand, financial institutions with low funding costs may fund risky assets on their balance sheets and buy default protection on those assets. The premium for buying protection on such securities may be less than the net spread earned over their funding costs.

Modeling For modeling purposes let us reiterate some basic terminology; see [89, 90]. We consider a frictionless economy with finite horizon $[0, T]$. We assume that there exists a unique martingale measure Q making all the default-free and risky security prices martingales, after renormalization by the money market account. This assumption is equivalent to the statement that the markets for the riskless and credit-sensitive debt are complete and arbitrage-free [89]. A filtered probability space $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{(t \geq 0)}, Q)$ is given and all processes are assumed to be defined on this space and adapted to the filtration $\mathcal{F}_t$ ($\mathcal{F}_t$ describes the information observable until time t). We denote the conditional expectation and the probability with respect to the equivalent martingale measure by $\mathbb{E}_t(\cdot)$ and $Q_t(\cdot)$, respectively, given information at time t.

Let $B(t, T)$ be the time t price of a default-free zero-coupon bond paying a sure currency unit at time T. We assume that forward rates of all maturities exist; they are defined in the continuous time by

$$f(t, T) = -\frac{\partial}{\partial T} \log B(t, T).$$

The default free spot rate is defined by

$$r(t) = \lim_{T \to t} f(t,T).$$

Spot rates can be modeled directly as by Cox et al. [34] or via forward rates as in Heath et al. [90]. The money market account that accumulates return at the spot rate is defined as

$$A(t) = e^{\int_0^t r(s)ds}.$$

Under the above assumptions, we can write default-free bond prices as the expected discount value of a sure currency unit received at time T, that is,

$$B(t,T) = \mathbb{E}_t\left[\frac{A(t)}{A(T)}\right] = \mathbb{E}_t\left[e^{-\int_t^T r(s)ds}\right].$$

Now, let $\widetilde{B}(t,T)$ be the time t price of a credit risky zero-coupon bond promising to pay a currency unit at time T. This promised payment may not be made in full if the firm is bankrupt at time T, i.e., only a fraction of the outstanding will be recovered in the event of default. Here we assume that the event premium is the difference of par and the value of a specified reference asset after default. Let again τ represent the random time at which default occurs, with a distribution function $F(t) = \mathbb{P}[\tau \le t]$ and $\mathbf{1}_{\{\tau<T\}}$ as the indicator function of the event. Then the price of the risky zero-coupon can be written in two ways:

$$\widetilde{B}(t,T) = \mathbb{E}_t\left[e^{-\int_t^T r(s)ds}(\mathbf{1}_{\{\tau>T\}} + REC(T)\mathbf{1}_{\{\tau<T\}})\right] \tag{7.1}$$

$$= \mathbb{E}_t\left[e^{-\int_t^T r(s)ds}\mathbf{1}_{\{\tau>T\}} + e^{-\int_t^\tau r(s)ds}REC(\tau)\mathbf{1}_{\{\tau<T\}}\right]. \tag{7.2}$$

In the first expression the recovery rate $REC(T)$ is thought of as a payout received at maturity, whereas in the second expression, we think of $REC(\tau)$ as the payment made at the time of default. Given the existence of the money market account, we can easily translate from one representation of the recovery to the other by

$$REC(T) = REC(\tau)e^{\int_\tau^T r(s)ds}.$$

A credit default swap now has a default leg and a premium leg. The present value of the contingent payment $1 - REC(\tau)$ is then

$$A_{def,t} = \mathbb{E}_t\left[e^{-\int_t^\tau r(u)du}(1 - REC(\tau))\mathbf{1}_{\{\tau<T\}}\right].$$

The present value of the spread payments s is given by:

$$A_{fee,t} = s\mathbb{E}_t \left[e^{-\int_t^T r(u)du} \mathbf{1}_{\{\tau > T\}} \right] .$$

From arbitrage-free arguments the value of the swap should be zero when it is initially negotiated. In the course of time its present value from the protection buyer's point of view is $A_{def,t} - A_{fee,t}$. In order to calculate the value of the CDS, it is required to estimate the survival probability, $S(t) = 1 - F(t)$, and the recovery rates $REC(t)$.

Swap premiums are typically due at prespecified dates and the amount is accrued over the respective time interval. Let $0 \leq T_0 \leq T_1 \leq \ldots T_n$ denote the accrual periods of the default swap, i.e., at time T_i, $i \geq 1$ the protection buyer pays $s\Delta_i$, where Δ_i is the day count fraction for period $[T_{i-1}, T_i]$, provided that there is no default until time T_i. Assuming furthermore a deterministic recovery rate at default, $REC(\tau) = REC$, and no correlation between default and interest rates we arrive at

$$A_{def,t} = (1 - REC) \int_{T_0}^{T_n} B(T_0, u) F(du) \tag{7.3}$$

$$A_{fee,t} = \sum_{i=1}^{n} s\Delta_i B(T_0, T_i)(1 - F(T_i)). \tag{7.4}$$

The integral describes the present value of the payment $(1 - REC)$ at the time of default. For a default "at" time u, we have to discount with $B(T_0, u)$ and multiply with the probability $F(du)$ that default happens "around" u.

In some markets a plain default swap includes the features of paying the accrued premium at default, i.e., if default happens in the period (T_{i-1}, T_i) the protection buyer is obliged to pay the already accrued part of the premium payment. In this case the value of the premium leg changes to

$$A_{fee,t} = \sum_{i=1}^{n} s \left[\Delta B(T_0, T_i)(1 - F(T_i)) + \int_{T_{i-1}}^{T_i} (u - T_{i-1}) B(T_0, u) F(du) \right], \tag{7.5}$$

where the difference $u - T_{i-1}$ is according to the given day count convention.

Both reduced-form models (intensity models) and structural models can in principle be applied to price default swaps. In the reduced-form model framework the relation between the intensity process h_t and the random survival probabilities at future times t provided $\tau > t$ is given by

$$q(t,T) = \mathbb{P}[\tau > T | \mathcal{F}_t] = \mathbb{E}_t \left(e^{-\int_t^T h(s)ds} \right).$$

If we assume a deterministic recovery rate REC and understand the recovery as a fraction of a corresponding riskless zero with the same maturity, we can write the price for a risky zero bond (7.1) as (on $\{\tau > t\}$):

$$\begin{aligned} \widetilde{B}(t,T) = {} & REC\ \mathbb{E}_t \left(e^{-\int_t^T r(s)ds} \right) \\ & + (1 - REC)\ \mathbb{E}_t \left(e^{-\int_t^T (r(s)+h(s))ds} \right). \end{aligned} \tag{7.6}$$

In the case of zero correlation between the short rate and the intensity process both processes in the exponent would factorize when taking the expectation value. But a really sophisticated default swap model would call for correlated default and interest rates, which leads us beyond the scope of this presentation. Instead, we turn in the following section back to correlated defaults and their application to basket swaps.

7.3 Basket Credit Derivatives

Basket default swaps are more sophisticated credit derivatives that are linked to several underlying credits. The standard product is an insurance contract that offers protection against the event of the kth default on a basket of n, $n \geq k$, underlying names. It is similar to a plain default swap but now the credit event to insure against is the event of the kth default and not specified to a particular name in the basket. Again, a premium, or spread, s is paid as an insurance fee until maturity or the event of kth default. We denote by s^{kth} the fair spread in a kth-to-default swap, i.e., the spread making the value of this swap equal to zero at inception.

If the n underlying credits in the basket default swap are independent, the fair spread s^{1st} is expected to be close to the sum of the fair default swap spreads s_i over all underlyings $i = 1, \ldots, n$. If the underlying credits are in some sense "totally" dependent the first default will be the one with the worst spread; therefore $s^{1st} = \max_i(s_i)$.

The question is now how to introduce dependencies between the underlying credits to our model. Again, the concept of copulas as introduced in Section 2.6 can be used, and, to our knowledge, LI [124, 125] was the first to apply copulas to valuing basket swaps by generating correlated default times as random variables via a correlation model and a credit curve. For more on copulas we refer to Section 2.6 and the literature referenced there, but see also EMBRECHTS et al. [56] for possible pitfalls.

Modeling Dependencies via Copulas Denote by τ_i, $i = 1, \ldots, n$ the random default times for the n credits in the basket, and let furthermore $(F_i(t))_{t \geq 0}$ be the curve of cumulative (risk-neutral) default probabilities for credit i:

$$F_i(t) = \mathbb{P}[\tau_i \leq t], \quad t \geq 0\,,$$

with $S_i(t) = \mathbb{P}[\tau_i > t] = 1 - F_i(t)$. $F(t)$ is assumed to be a strictly increasing function of t with $F(0) = 0$ and $\lim_{t \to \infty} F(t) = 1$. This implies the existence of the quantile function $F^{-1}(x)$ for all $0 \leq x \leq 1$. From elementary probability theory we know that for any standard uniformly distributed U,

$$U \sim U(0,1) \quad \Rightarrow \quad F^{-1}(U) \sim F. \tag{7.7}$$

This gives a simple method for simulating random variates with distribution F, i.e., random default times in our case. The cash flows in a basket default swap are functions of the whole random vector $(\tau_1, \ldots, \tau_n)$, but in order to model and evaluate this basket swap we need the joint distribution of the τ_i's:

$$\boldsymbol{F}(t_1, \ldots, t_n) = \mathbb{P}[\tau_1 \leq t_1, \ldots, \tau_n \leq t_n].$$

Similarly, we define the multivariate survival function $\boldsymbol{S}$ by

$$\boldsymbol{S}(t_1, \ldots, t_n) = \mathbb{P}[\tau_1 > t_1, \ldots, \tau_n > t_n].$$

Note that

$$\begin{aligned} S_i(t_i) &= \boldsymbol{S}(0, \ldots, 0, t_i, 0, \ldots, 0), \\ \boldsymbol{S}(t_1, \ldots, t_n) &\neq 1 - \boldsymbol{F}(t_1, \ldots, t_n). \end{aligned}$$

We exploit again the concept of copula function where, for uniform random variables, $U_1, U_2, \ldots, U_n$,

$$C(u_1, u_2, \ldots, u_n) = \mathbb{P}[U_1 \le u_1, U_2 \le u_2, \ldots, U_n \le u_n]$$

defines a joint distribution with uniform marginals. The function $C(u_1, u_2, \ldots, u_n)$ is called a Copula function. Remember that $U_i = F_i(\tau_i)$ admits a uniform distribution on the interval $[0, 1]$; so, the joint distribution of $(\tau_1, \ldots, \tau_n)$ can be written as:

$$\boldsymbol{F}(t_1, \ldots, t_n) = C(F_1(t_1), \ldots, F_n(t_n)) . \tag{7.8}$$

Hence, the Copula function introduces a mutual correlation by linking univariate marginals to their full multivariate distribution thereby separating the dependency structure C, i.e., the ingredients are some credit curve for each credit as marginal distribution functions for the default times and a suitable chosen copula function. Observe that by Sklar's theorem (Section 2.6) any joint distribution can be reduced to a copula and the marginal distributions, although it may be difficult to write down the copula explicitly.

One of the most elementary copula functions is the multivariate normal distribution

$$C(u_1, u_2, \ldots, u_n) = N_n \left[N^{-1}(u_1), N^{-1}(u_2), \ldots, N^{-1}(u_n); \Gamma\right] \tag{7.9}$$

where N_n is as before the cumulative multivariate normal distribution with correlation matrix Γ and N^{-1} is the inverse of a univariate normal distribution. Clearly, there are various different copulas generating all kinds of dependencies, and the choice of the copula entails a significant amount of model risk [68, 70]. The advantage of the normal copula, however, is that, as we have seen in Chapter 2, it relates to the latent variable approach to model dependent default. Assume that the default event of credit i up to time T is driven by a single random variable r_i (ability-to-pay variable) being below a certain threshold $c_i(T)$:

$$\tau_i < T \Leftrightarrow r_i < c_i(T) .$$

If the Z_i's admit a multivariate standard normal distribution with correlation matrix $\widetilde{\Gamma}$, then to be consistent with our given default curve, we set $c_i(T) = N^{-1}(F_i(T))$. The pairwise joint default probabilities are now given in both representations by:

$$\begin{aligned} \mathbb{P}\left[\tau_i \le T, \tau_j \le T\right] &= \mathbb{P}\left[r_i \le c_i(T), r_j \le c_j(T)\right] \\ &= N_2[N^{-1}(F_i(T)), N^{-1}(F_j(T)); \widetilde{\Gamma}_{ij}] \end{aligned} \tag{7.10}$$

We see that these probabilities (7.10) only coincide with those from the normal copula approach (7.8), (7.9), if the asset correlation matrix $\widetilde{\Gamma}$ and the correlation matrix Γ in the normal copula are the same. But note that since the asset value approach can only model defaults up to a single time horizon T, the calibration between the two models can only be done for one fixed horizon. So, we see again that the factor model approach to generate correlated defaults based on standard normal asset returns is tantamount to a normal copula approach.

Remark Analogously to the default distribution we can apply Sklar's theorem to the survival function, i.e., when $\boldsymbol{S}$ is a multivariate survival function with margins $S_1, \ldots, S_n$, then there exists a copula representation

$$\boldsymbol{S}(t_1, \ldots, t_n) = \breve{C}(S_1(t_1), \ldots, S_n(t_n))\,. \tag{7.11}$$

There is an explicit, although rather complex relation between the survival copula $\breve{C}$ and the distribution copula C [76]; in the two-dimensional case we obtain

$$\begin{aligned}
\breve{C}(u_1, u_2) &= \boldsymbol{S}(S_1^{-1}(u_1), S_2^{-1}(u_2)) = \boldsymbol{S}(t_1, t_2) \\
&= 1 - F_1(t_1) - F_2(t_2) + \boldsymbol{F}(t_1, t_2) \\
&= S_1(t_1) + S_2(t_2) - 1 + C(1 - S_1(t_1), 1 - S_2(t_2)) \\
&= u_1 + u_2 - 1 + C(1 - u_1, 1 - u_2),
\end{aligned}$$

where it can easily be shown that $\breve{C}$ is indeed a copula function. At this point let us state that a copula is radially symmetric if and only if $C = \breve{C}$ (proof [76]). The normal copula is radial symmetric; so, e.g., in two dimensions we find indeed that

$$\begin{aligned}
\breve{C}(u_1, u_2) &= u_1 + u_2 - 1 + C(1 - u_1, 1 - u_2) \\
&= u_1 + u_2 - 1 + N_2\left[N^{-1}(1 - u_1), N^{-1}(1 - u_2); \Gamma\right] \\
&= N_2\left[N^{-1}(u_1), +\infty; \Gamma\right] + N_2\left[+\infty, N^{-1}(u_2); \Gamma\right] \\
&\quad - N_2\left[+\infty, +\infty; \Gamma\right] + N_2\left[-N^{-1}(u_1), -N^{-1}(u_2); \Gamma\right] \\
&= N_2\left[N^{-1}(u_1), N^{-1}(u_2); \Gamma\right] \\
&= C(u_1, u_2).
\end{aligned}$$

This property is very interesting for computational purposes, since in the radially symmetric case it is thus equivalent to work with the distribution copula or with the survival copula.

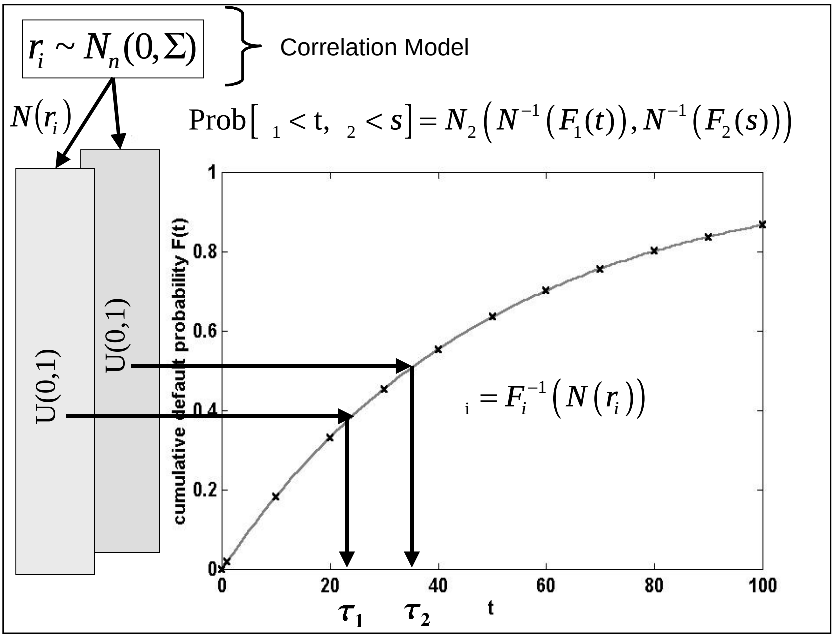


FIGURE 7.3: Generating correlated default times via the copula approach (illustrative: here, $F_1 = F_2$).

Summarizing, the normal copula function approach for modeling correlated default times is as follows (Figure 7.3):

- Specify the cumulative default time distribution F_i (credit curve), such that $F_i(t)$ gives the probability that a given asset i defaults prior to t.

- Assign a standard normal random variable r_i to each asset, where the correlation between distinct r_i and r_j is ρ_{ij}.

- Obtain the default time τ_i for asset i through

$$\tau_i = F_i^{-1}\left(N(r_i)\right) .$$

Note that since $F_i(t)$ is a strictly increasing continuous function with $\lim_{t\to\infty} F_i(t) = 1$ there is always a default time, though it may be very large.

In the one-period case, positively correlated defaults mean that if one asset defaulted it is more likely that the second defaults as well, compared to independent defaults. For default times, a positive correlation means that the time between the two default events is smaller, on average, than if they were uncorrelated. Figure 7.4 depicts the average standard deviation of default times τ_i, $1 \leq i \leq 5$, $< \mathrm{std}_i[\tau_i] >$ (the average is taken over numerous scenarios), in units of the average default time, $< \mathrm{mean}_i[\tau_i] >$, the average first-to-default-time $< \min_i[\tau_i] >$, and the average last-to-default-time $< \max_i[\tau_i] >$, for a uniform basket of five loans in dependence of the asset correlation with cumulative multi-year default probabilities as in Table 7.1.

TABLE 7.1: Term structure of cumulative default probability.

year	1	2	3	4	5	6	7	8	9	10
PD	0.0071	0.0180	0.0320	0.0484	0.0666	0.0859	0.1060	0.1264	0.1469	0.1672

Pricing In order to price basket default swaps, we need the distribution of the time τ^{kth} of the kth default. The kth default time is in fact the order statistic $\tau_{(k:n)}$, $k \leq n$, and in general, we have for the distribution functions

$$S_{(k:n)}(t) = 1 - F_{(k:n)}(t)\,.$$

The distribution of the first order statistic $\tau_{(1:n)}$ is

$$F_{(1:n)}(t) = \mathbb{P}[\tau_{(1:n)} \leq t] = 1 - \mathbb{P}[\tau_1 > t, \ldots, \tau_n > t] = 1 - \boldsymbol{S}(t, \ldots, t),$$

and the one of the last order statistic (the time of the last default) is obviously

$$F_{(n:n)}(t) = \mathbb{P}[\tau_1 \leq t, \ldots, \tau_n \leq t] = \boldsymbol{F}(t, \ldots, t). \tag{7.12}$$

The corresponding formulas for the other distribution function $F_{(k:n)}$ in terms of the copula function are much more involved (see [76]); we only state the special cases $n = 2, 3$:

n=2:

$$\begin{aligned}
\boldsymbol{F}_{(2:2)}(t) &= C(F_1(t), F_2(t)) \\
\boldsymbol{F}_{(1:2)}(t) &= F_1(t) + F_2(t) - C(F_1(t), F_2(t))
\end{aligned}$$

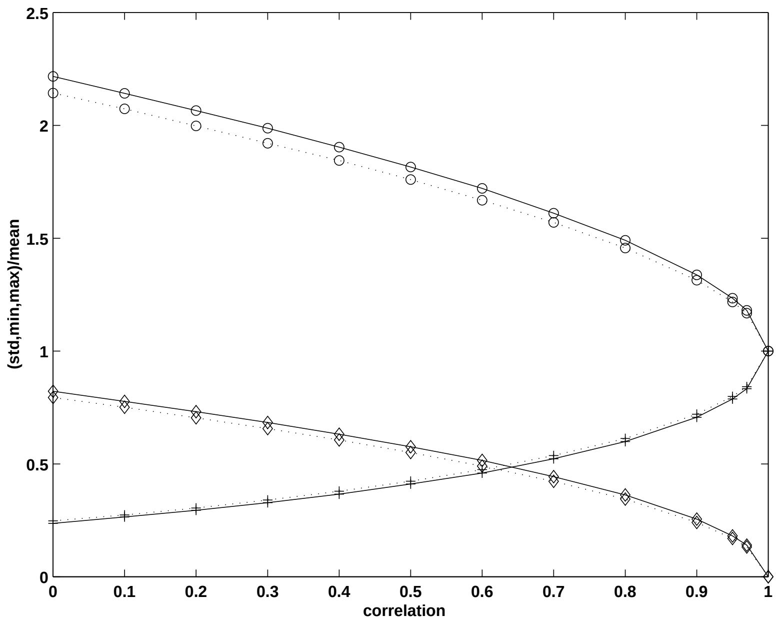


FIGURE 7.4: The average standard deviation of the default times $< \mathrm{std}_i[\tau_i] >$ ($\diamond$), the average first-to-default time $< \min_i[\tau_i] >$ ($+$), and the average last-to-default-time $< \max_i[\tau_i] >$ ($\circ$) in units of the average default time $< \mathrm{mean}_i[\tau_i] >$ for a uniform basket of five loans in dependence of the asset correlation for normal distributed (solid) and t-distributed (dashed) latent variables.

n=3:

$$\begin{aligned}
\boldsymbol{F}_{(3:3)}(t) &= C(F_1(t), F_2(t), F_3(t)) \\
\boldsymbol{F}_{(2:3)}(t) &= C(F_1(t), F_2(t)) + C(F_1(t), F_3(t)) + C(F_2(t), F_3(t)) \\
&\quad -2C(F_1(t), F_2(t), F_3(t)) \\
\boldsymbol{F}_{(1:3)}(t) &= F_1(t) + F_2(t) + F_3(t) - C(F_1(t), F_2(t)) - C(F_1(t), F_3(t)) \\
&\quad -C(F_2(t), F_3(t)) + C(F_1(t), F_2(t), F_3(t))
\end{aligned}$$

The fair spread s^{kth} for maturity T_m is then given by (compare Equations (7.3), (7.4))

$$\begin{aligned}
0 = s^{kth} &\sum_{i=1}^{m} \Delta_i B(T_0, T_i) S_{(k:n)}(T_i) \\
&- \sum_{i=1}^{n} (1 - REC_i) \int_{T_0}^{T_m} B(T_0, u) F_{(k:n)}^{\text{kth}=i}(du).
\end{aligned} \tag{7.13}$$

The first part is the present value of the spread payments, which stops at τ^{kth}. The second part is the present value of the payment at the time of the kth default. Since the recovery rates might be different for the n underlying names, we have to sum up over all names and weights with the probability that the kth default happens around u and that the kth name is just i. (We assume that there are no joint defaults at exactly the same time.) So $F_{(k:n)}^{\text{kth}=i}$ is the probability distribution of the kth order statistic of the default times and that $kth = i$. Figure 7.5 show the kth-to-default spreads for a basket of three underlyings with fair spreads $s_1 = 0.009$, $s_2 = 0.010$, and $s_3 = 0.011$, and pair-wise equal correlation. Schmidt and Ward [162] already observed that the sum of the kth-to-default swap spreads is greater than the sum of the individual spreads, i.e., $\sum_{k=1}^{n} s^{kth} > \sum_{i=1}^{n} s_i$. Both sides insure exactly the same risk; so, this discrepancy is due to a windfall effect of the first-to-default swap. At the time of the first default one stops paying the huge spread s^{1st} on the one side but on the plain-vanilla side one stops just paying the spread s_i of the first default i. Of course this mismatch is only a superficial one, since the sums of the present values of the spreads on both sides are equal. Note also the two extreme cases. For fully correlated underlyings, $\rho = 1$, the first-to-default spread is the

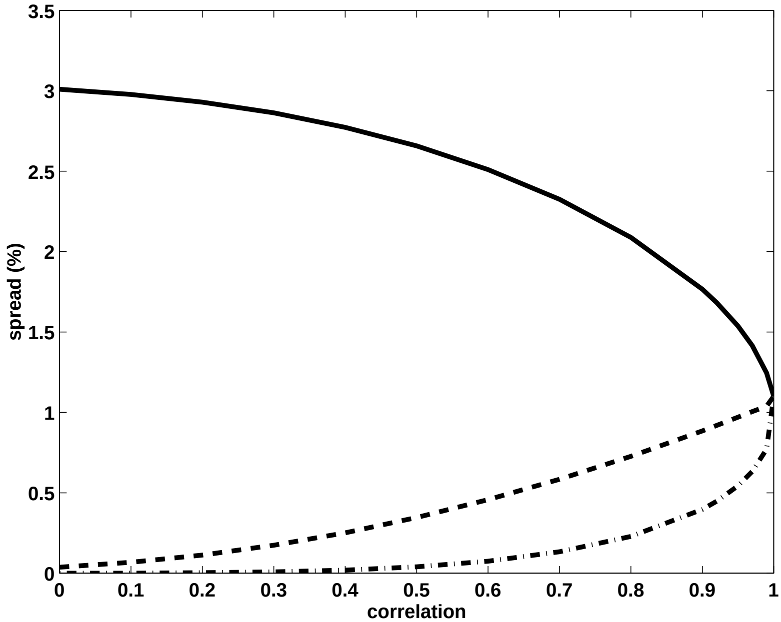


FIGURE 7.5: kth-to-default spread versus correlation for a basket with three underlyings: (solid) s^{1st}, (dashed) s^{2nd}, (dashed-dotted) s^{3rd}.

worst of all underlyings. Of course, in the normal copula framework perfect linear correlation means that the state variables are identical and that the name with the largest default probability dominates all others (assuming the same recovery rates for all underlyings). On the other hand, for $\rho = 0$, from an arbitrage-free argument one can show that the first-to-default spread is close to the sum of the individual spreads. If the correlation is greater than zero the underlying names are dependent, which entails a spread widening of the remaining names as a consequence of the default of credit i. Schmidt and Ward [162] investigated how this implied spread widening is reflected in the copula approach and found that given a flat correlation structure the size of the spread widenings depends on the quality of the credit first defaulting, i.e., the less risky the defaulting name the larger the impact. Also the implied spread widening admits a pronounced term structure: the earlier the first default, the larger the impact on the remaining spreads.

Counterparty Risk So far, we have tacitly ignored the counterparty risk of the protection seller to default. This feature could also be dealt with in the context of the copula approach (but see also Hull and White [95] for another approach). For simplicity we reduce the problem to a single obligor CDS, but the generalization to baskets is straightforward. We now have the additional risk that the protection seller, i.e., the swap counterparty, can default, together with the reference security. So, instead of making the promised payments $1 - REC$ in the event of the reference default, only a fraction δ of that payment is recovered by the protection buyer. The formulas for the default leg (7.3) and the premium leg (7.4) change then, in informal notation, to

$$\begin{aligned}
A_{def,t} &= \int_{T_0}^{T_n} B(T_0,u)(1 - REC(u))(-\boldsymbol{S}(du,u)) \\
&\quad + \int_{T_0}^{T_n} B(T_0,u)(1 - REC(u))\delta(u)\boldsymbol{F}(du,u) \\
&= \int_{T_0}^{T_n} B(T_0,u)(1 - REC(u))\left[(1-\delta(u))(-\boldsymbol{S}(du,u)) \right. \\
&\qquad\qquad \left. +\delta(u)F_{ra}(du)\right] \\
A_{fee,t} &= \sum_{i=1}^{n} s\Delta_i B(T_0,T_i)\boldsymbol{S}(T_i,T_i),
\end{aligned}$$

where F_{ra} denotes the default curve of the reference asset. $-\boldsymbol{S}(du,u) = -\partial_1\boldsymbol{S}(u,u)du$ is the probability that the reference asset defaults between u and $u + du$ while the swap counterparty is still alive, whereas $\boldsymbol{F}(du,u) = \partial_1\boldsymbol{F}(u,u)du$ is the probability that the reference asset defaults between u and $u+du$ and the counterparty has already defaulted. The bivariate survival function $\boldsymbol{S}$ can then again be represented by a copula function $\breve{C}$. Figure 7.6 shows the spread of a single-asset default swap as a function of the correlation of the reference asset to the swap counterparty. The risk-free rate is $r = 4\%$, the hazard rates are supposed to be constant at $\lambda_{ra} = 0.011$ and $\lambda_{cp} = 0.01$, and the recovery rate is at $REC = 0.2$.

Remark Obviously, the normal distribution is only one choice for a possible copula function. See, for example, [56, 68] for possible pitfalls in modeling dependencies via copula functions. For comparison, we also compute default times with t-distributed latent variables for the uniform basket of five loans (see also [133] for a use of t-copulas in

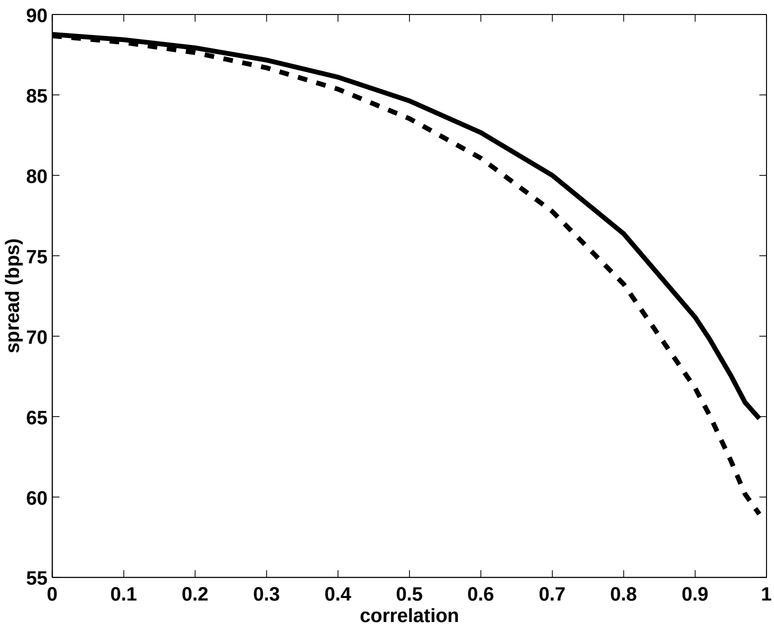


FIGURE 7.6: Default spread versus correlation between reference asset and swap counterparty: (solid) $\delta = 0.2$, (dashed) $\delta = 0$ as fraction of recovery payment made.

modeling default baskets). Choosing identical linear correlations, ρ_n, ρ_t in the normal- and t-copula function ($\nu = 5$ degrees of freedom), produces the results in Figure 7.4. The standard deviations and the maxima (last-to-default) of the default times in the t-distributed case are slightly lower over the full range of linear correlation than the ones based on normal distributed variables. The first-to-default times are slightly higher in the t-distributed case than in the normal distributed one. Next, we calibrated for some cases the linear correlation parameters in the t- and normal distributed case to match the one-year default correlation, based on a one-year default probability of $\mathrm{PD}_1 = 0.0071$. Note that this is not generally possible. The following table shows that now the differences are much larger.

$\mathrm{PD}_1 = 0.0071$	$< \mathrm{std}_i[\tau_i] >$	$< \min_i[\tau_i] >$	$< \max_i[\tau_i] >$
$\rho_t = 0.1,\ \nu = 5$	0.75	0.27	2.07
$\rho_n = 0.455$	0.61	0.38	1.85
$\rho_t = 0.2,\ \nu = 5$	0.71	0.31	2.00
$\rho_n = 0.52$	0.56	0.42	1.79
$\rho_t = 0.1,\ \nu = 10$	0.76	0.27	2.11
$\rho_n = 0.32$	0.67	0.33	1.97
$\rho_t = 0.2,\ \nu = 10$	0.72	0.30	2.03
$\rho = 0.393$	0.64	0.36	1.9

7.4 Credit Spread Products

Credit spread is the difference between the yield on a particular debt security and a benchmark yield, usually on a government bond. Credit spread options (CSO) can be based on various types of credit spreads, the asset-swap spread, the default-swap spread and the yield spread [134], and they allow investors to express a directional view on credit spreads or to hedge risk.

In case of options on CDS spreads one speaks of a credit default swaption. The "put/call" terminology is sometimes confusing here. Credit default swaptions use the lingo "payer" and "receiver" instead (see also interest rate swaptions). A payer option is the right to buy credit default protection at a pre-specified strike level K on a future date with a payoff equal to $\max(S(T) - K, 0)$ at maturity, where $S(T)$

is the credit spread. A payer option is both a put on credit quality – a bet that credit will deteriorate – and a call on spreads – a bet that spreads will widen. Likewise, a receiver option is the right to sell credit default protection at a pre-specified strike level on a future date, $\max(K-S(T),0)$. A receiver option is both a call on credit – the buyer makes money when credit quality improves – and a put on spreads.

One of the main characteristics of these products is that the return is not dependent on a specific credit event. It merely depends on the value of one reference credit spread against another. If the reference asset owner's credit rating goes down, and therefore the default probability increases, the credit spread goes up and vice versa. A debt issuer can make use of payer options (put on credit) to hedge against a rise in the average credit spread. On the other hand, a financial institution that holds debt securities can purchase receiver options to hedge against a fall in the credit spread (call on credit).

Credit spread derivatives are priced by means of a variety of models. One can value them by modeling the spread itself as an asset price. The advantage of this approach is its relative simplicity. Longstaff and Schwartz [128] developed a simple framework for pricing credit spread derivatives, which we will summarize in the following. It captures the major empirical properties of observed credit spreads. They use this framework to derive closed-form solutions for call and put CSOs.

Let x denote the logarithm of the credit spreads, that is $x_t = \log(S(t))$. We assume that x is given by the SDE

$$dx = (a - bx)dt + sdB_1,$$

where a, b, s are parameters and B_1 is a Wiener process. This implies that changes in x are mean-reverting and homoscedastic, which is consistent with the empirical data. We assume that the default-free term structure is determined by a one-factor-model [181], that is

$$dr = (\alpha - \beta r)dt + \sigma dB_2.$$

Again α, β, σ_2 are parameters and B_2 is a Wiener process. The correlation coefficient between dB_1 and dB_2 is $\hat{\rho}$. Let us assume market prices of the risk premium are incorporated into a and α. Thus, both a and α are risk-adjusted parameters rather than empirical ones. This assumption is consistent with Vasicek [181] and Longstaff and Schwartz

[127]. The risk-adjusted process for x is given by [128]

$$dx = \left[a - bx - \frac{\hat{\rho}\sigma s}{\beta}\left(1 - e^{-\beta(T-t)}\right)\right] dt + s dB_1. \tag{7.14}$$

This SDE in (7.14) can be solved by making a change of variables and then integrating. The resulting solution implies that x_T is conditionally normally distributed with respect to (7.14) with mean μ and variance η^2, where

$$\mu = e^{-bT}x + \frac{1}{b}\left(\alpha - \frac{\hat{\rho}\sigma s}{\beta}\right)[1 - e^{-bT}] + \frac{\hat{\rho}\sigma s}{\beta(b+\beta)}[1 - e^{-(b+\beta)T}]$$
$$\eta^2 = \frac{s^2[1 - e^{-2bT}]}{2b}.$$

Note that as $T \to \infty$ the values of μ and η^2 converge to fixed values, and the distribution of x_T converges to a steady-state stationary distribution. With this framework we can find the price of a European call CSO. Let $C(x, r, T)$ denote the value of the option. The payoff function for this option is simply $H(x) = \max(e^x - K, 0)$. The closed-form solution for the call CSO is given by

$$C(x, r, T) = p(r, T)\left[e^{\mu + \eta^2/2} N(d_1) - K N(d_2)\right].$$

Here, $N(\cdot)$ is the cumulative standard normal distribution, $p(r, T)$ is a riskless discount bond, and

$$d_1 = \frac{-\log(K) + \mu + \eta^2}{\eta}, \qquad d_2 = d_1 - \eta.$$

The value of a European put CSO is

$$P(x, r, T) = C(x, r, T) + p(r, T)\left[K - e^{\mu + \eta^2/2}\right].$$

The option formula has some similarities with the Black-Scholes option pricing formula. However, the value of a call option can be less than its intrinsic value even when the call is only slightly in the money. This surprising result is due to the mean reversion of the credit spreads. When the spread is above the long-run mean, it is expected to decline over time. This can not happen in the B-S model because the underlying asset must appreciate like the riskless rate in the risk-neutral

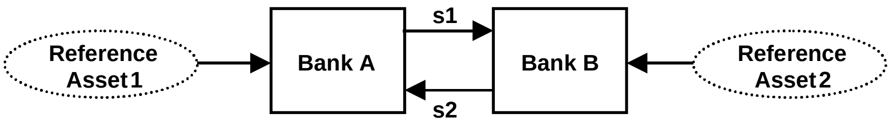


FIGURE 7.7: Credit spread swap.

framework. The delta for a call is always positive, as in the B-S framework, but the delta of a CSO call decreases to zero as the time until the expiration increases. A change in the current credit spread is heavily outweighed by the effects of mean reversion if the expiration date of the call is far in the future.

An investor may combine a payer and a receiver option to create a straddle, which is a bet on spread volatility. The buyer of the straddle makes money if spreads either widen or tighten by more than the breakeven level. Investors can also insure against rising credit spreads by buying a payer option and reduce the cost by selling a receiver option.

In a credit spread forward (CSF), counterparty A pays at time T a pre-agreed fixed payment and receives the credit spread of the reference asset at time T. Conversely, counterparty B receives the fee and pays the credit spread. The fixed payment is chosen at time $t < T$ to set the initial value of the credit spread forward to zero. The credit spread forward can also be structured around the relative credit spread between two different defaultable bonds. Credit spread forwards can be combined to a credit spread swap (Figure 7.7) in which one counterparty pays periodically the relative credit spread, $(S_1(t) - S_2(t))$, to the other.

7.5 Credit-Linked Notes

Credit-linked notes exist in various forms in the credit derivatives market; see [41, 143, 100, 44]. In its most common form, a credit-linked note (CLN) is a *synthetic bond* with an *embedded default swap* as illustrated in Figure 7.8.

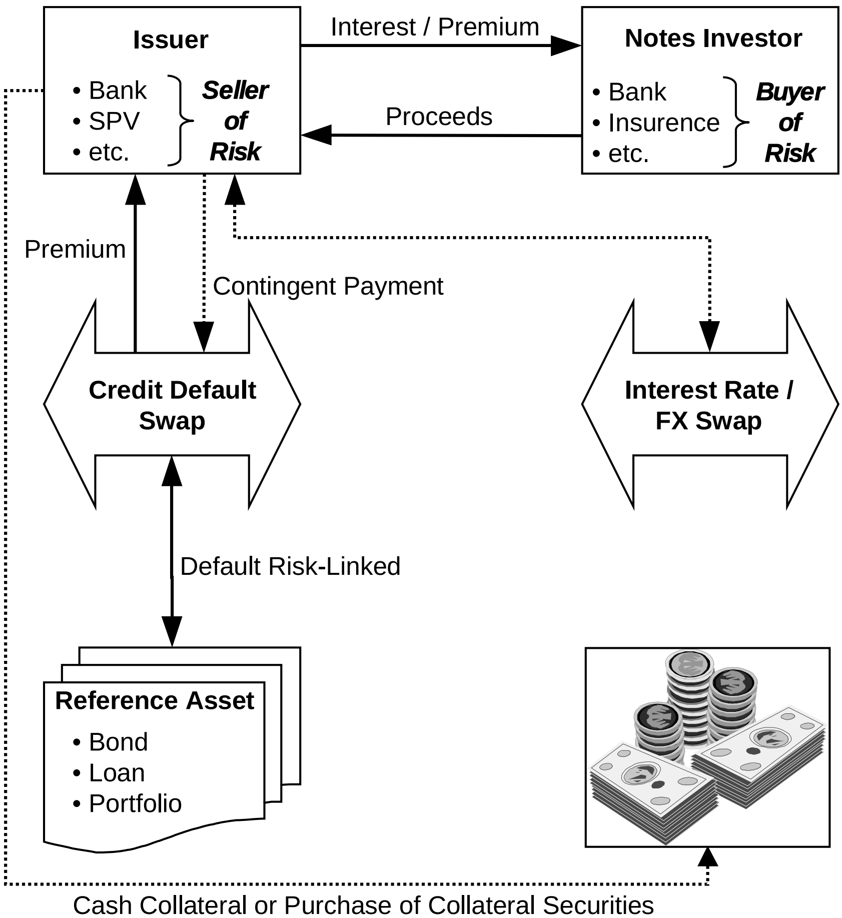


FIGURE 7.8: Example of a credit-linked note.

CLNs are initiated in several ways. In the following we outline four examples of typical CLN structures.

The first case we present is the situation of an (institutional) investor who wants to have access to a credit exposure (the *reference asset*) for which by policy, regulation, or other reasons he has no direct access. In such cases, a CLN issued by another institution (the *issuer*) which has access to this particular credit exposure offers a way to evade the problems hindering the investor to purchase the exposure he is interested in. The issuer sells a note to the investor with underlying exposure equal to the face value of the reference asset. He receives the face value of the reference asset as cash proceeds at the beginning of the transaction and in turn pays interest, including some premium for the default risk, to the investor. In case the reference asset experiences a *credit event*, the issuer pays to the investor the recovery proceeds of the reference asset. The spread between the face value and the recovery value of the reference asset is the investor's *exposure at risk*. In case no credit event occurred during the lifetime of the reference note, the issuer pays the full principal back to the investor. So in this example one could summarize a CLN as a *synthetic bond* with an *embedded default swap*.

In our second example, an investor, who has no access to the credit derivatives market or is not allowed to do off-balance sheet transactions, wants to invest in a credit default swap, selling protection to the owner of some reference asset. This can be achieved by investing in a CLN in the same way as described in our first example. Note that from the investor's point of view the CLN deal differs from a default swap agreement by the cash payment made upfront. In a default swap, no principal payments are exchanged at the beginning.

Another common way to set up a CLN is *protection* buying. Assume that a bank is exposed to the default risk of some reference asset. This could be the case by means of an asset on the balance sheet of the bank or by means of a situation where the bank is the protection seller in a credit default swap. In both cases the bank has to carry the reference asset's default risk; see Figure 7.8. The bank can now issue a CLN to some investor who pays the exposure of the reference asset upfront in cash to the bank and receives interest, including some premium reflecting the riskiness of the reference asset, during the lifetime of the note. If the reference asset defaults, the bank suffers a loss for its balance sheet asset (funded case) or has to make a contingent payment for the default swap (unfunded case). The CLN then compensates the bank for the loss, such that the CLN functions as an *insurance*.

In this example, the difference between a CLN and just another default swap arises from the *cash proceeds* the bank receives *upfront* from the CLN investor. As a consequence, the bank is *not exposed to the counterparty risk* of the protection selling investor. Therefore, the credit quality of the investor is of no relevance[2]. The proceeds from the CLN can be kept as a cash collateral or be invested in high-quality collateral securities, so that losses on the reference asset will be covered *with certainty*.

Our last example refers to Chapter 8, where CLNs will be discussed as notes issued by a *special purpose vehicle* (SPV) in order to set up a *synthetic CDO*. In this case, CLNs are used for the exploitation of *regulatory arbitrage* opportunities and for *synthetic risk transfer*.

Besides the already mentioned reasons, there are certainly more advantages of CLNs worthwhile to be mentioned. For example, CLNs do not require an ISDA master agreement, but rather can contractually rely on the term sheet of the notes. Another advantage of CLNs is that not only the investor's credit quality but also his *correlation* with the reference asset is of no relevance to the CLN, because the money for the protection payment is delivered upfront. This concludes our discussion of credit derivatives.

Further Reading

A reference to credit derivatives is the book by Schönbucher [164] and a very recent book by Wagner [184], just to mention two out of many examples.

[2] Of course, for a short time at the start of the CLN there could be a *settlement risk*.

Chapter 8

Collateralized Debt Obligations

Collateralized debt obligations constitute an important class of so-called *asset backed securities* (ABS), which are securities backed by a pool of assets. Depending on the underlying asset class, ABS include various subclasses, for example *residential* or *commercial mortgage backed securities* (RMBS, CMBS), *trade receivables* ABS, *credit card* ABS (often in the form of so-called *Master Trusts*), *consumer loan* ABS, and so on. Some years ago it started that ABS were also structured based on pools of derivative instruments, like *credit default swaps* and other derivative instruments. In the first-mentioned case such ABS transactions are often called *collateralized swap obligations* (CSO). In general, one can say that ABS can be based on *any* pool of assets generating a cash flow suitable for being structured in order to meet investor's risk and return preferences.

When the first edition of this book appeared the CDO and – more general – ABS markets were rapidly evolving. The development came to a stop during the financial crisis starting in May 2007. Markets dried up completely, liquidity basically vanished and many structured credit instruments defaulted or at least dropped down tremendously w.r.t. their mark-to-market. In the course of the crisis there appeared many discussion papers and press articles about the role of the models in the crisis. People asked themselves whether models were in general capable of dealing with market scenarios like the current one. Some people also blamed the credit quants and analysts for not having warned executives and for working with overly optimistic model assumptions.

We can not explore the roots of the crisis and the model pros and cons in an introductory text like this book but instead refer to [26] for an extensive survey on *rethinking credit risk modeling and lessons from the crisis*. However, we want to mention a few facts which fit nicely in this chapter because the field of CDO modeling was a main motor for model developments in credit risk during the last years.

First we want to claim that the model universe is sufficiently rich to deal with crisis scenarios like the most recent one. As mentioned in the introduction already, it mainly is not the model itself which might need adjustments, it is the parameterization which might need a rethinking. And on top one can say that markets would have had much benefit from using models at all in the past. The fact is that large parts of the ABS and CDO market were managed basically free of stochastic modeling. Let us focus on that statement for a moment. Roughly summarizing, one can say that there are three stages of CDO and ABS valuation.

- **Believe in agencies**: market participants at this stage used no models at all but simply relied on ratings from Moody's, S&P and Fitch. It is shocking how often this has been the approach of investors in the past. A bitter wake-up for such investors was that agency ratings did not hold during the crisis and that large amounts of invested capital came under pressure or got lost.

- **Cash flow models**: market participants at this stage looked at ratings but in addition also operated tailor-made or off-the-shelf commercial cash flow models to understand the structures. This is already a better stage than the "believe in agencies" approach. However, cash flow models without an underlying model for the uncertainties in cash flows based on the stochastics of underlying credit assets are not sufficient for coming to meaningful conclusions. Very often such cash flow models allow for certain scenarios as an assumed model input but this can not be substitute for a full distribution of cash flows and the corresponding distribution of risks and returns. So in the end, information for such investors was missing and the picture of the transaction investment remained incomplete.

- **Cash flow model with underlying portfolio model**: sophisticated market players typically followed a combined approach of stochastic modeling of underlying credit risks and resulting (random) cash flows at the liability side of the structure; see Figure 8.6. This is the approach we consider as a "must-have" for CDO and ABS investors; unfortunately, it was the minority of investors who followed this approach.

So the truth is that many market participants did their investments basically "model-free" by just relying on the opinion of rating agencies and investor reports. A positive rating on a CDO tranche made such investors feel that their money was wisely invested. But in the course of the crisis it turned out that the situation can change very quickly and that agency ratings did not fully reflect the range of possible outcomes of a market heavily driven by stochastic influences.

Now, at the third afore-mentioned stage one still finds huge differences in the way market players developed models and applied them to their investments. One finds the whole range from fully-fledged Monte Carlo simulations to very much reduced and simplified analytic approximations. And here one also finds some valid arguments for criticism of model approaches because models were sometimes calibrated in overly optimistic ways. We come back to this point later in Section 8.6.

Another difficult part of the discussion is communication of model results to executive managers. Already years ago, credit risk modelers worked with models with fatter tails of the loss distribution based on tail dependencies which have not been used before. However, when it then comes to the point where a three times higher than usual credit risk quantity needs to be communicated to senior management the excitement about a tail dependent model and or a more defensive parameterization very soon comes to an end. A fatter tail of a loss distribution means lower return on capital and this is something which is hard to argue for just based on a re-parameterized model without corresponding clouds visible over the markets. We believe that due to such difficulties in communication and acceptance some very useful insights never made it into the briefing kit for decision making executive managers. Especially in a euphoric market environment as we had right before the crisis it is very difficult to adopt the role of a “trouble smeller.”

So altogether we find a superposition of effects which one can consider when thinking about the role of models in the most recent crisis. We stop the discussion here but revisit it in Section 8.6.

In a recent book [24] on *structured credit portfolio analysis, baskets and CDOs* two from the author team of this book already dedicated a whole textbook to questions of CDO modeling. It does not make sense to repeat part of the content from [24] here in this book. Everything said in the first edition of this book remained true and valid over the last years[1]. Therefore, we decided to keep most of the material from the first edition unchanged but extended our discussion of multi-period modeling which is a major building block of modern CDO modeling and added a section on recent developments in the field of structured credit. In the latter mentioned section we also integrated our hints for further reading and included a few remarks on model risk and the challenge of model parameterization.

So let us start with an introduction to CDOs in a "story-telling" style. Some references for further reading are given in the very last section of this chapter.

8.1 Introduction to Collateralized Debt Obligations

Figure 8.1 shows a segmentation of the CDO market. There are basically two types of debt on which CDOs are based, namely *bonds* and *loans*, constituting

- *Collateralized bond obligations* (CBO):
  In this case, the collateral pool contains credit risky bonds. Many of the CBOs we currently find in the market are motivated by *arbitrage spread* opportunities, see Section 8.2.1.

- *Collateralized loan obligations* (CLO):
  Here the collateral pool consists of loans. *Regulatory capital relief*, *cheaper funding*, and, more general, *regulatory arbitrage* combined with *economic risk transfer* are the major reasons for the *origination* of CLOs by banks all over the world, see Section 8.2.1.

Besides these two, CSOs (see the introductory remarks) are *market standard* today. Their advantage is the reduction of funding costs, because instead of *funded* instruments like loans or bonds, the cash flows from credit derivatives are structured in order to generate an attractive arbitrage spread. A second advantage of CSOs is the fact

[1]The only exception to that rule is the use of the binomial expansion method (see Section 8.5) by Moody's. This method is outdated and today rating agencies use more sophisticated models.

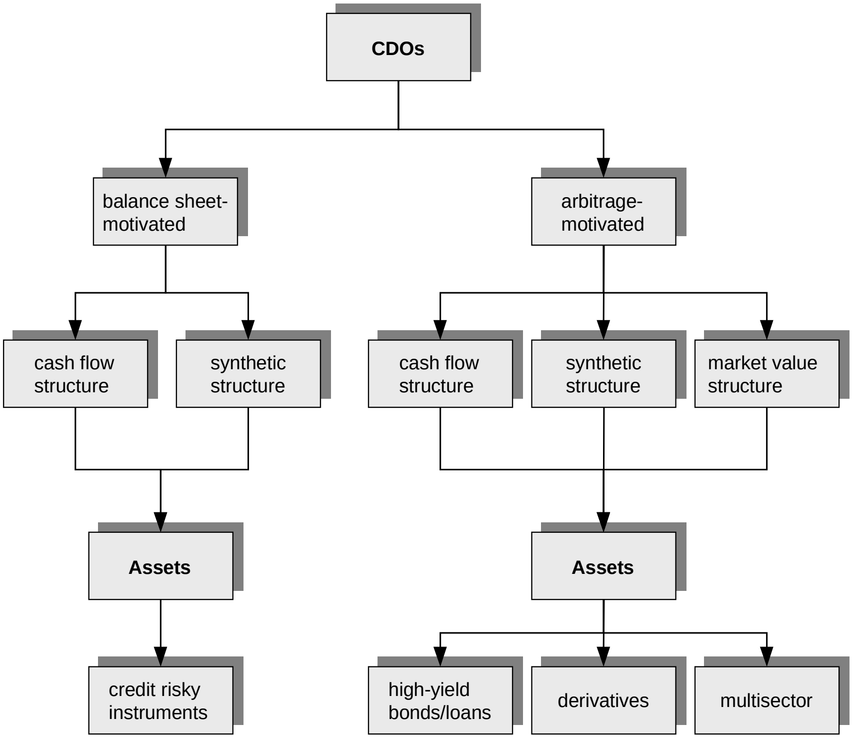


FIGURE 8.1: Classification of CDOs.

that credit derivatives are *actively traded* instruments, such that, based on the *fair market spread* of the collateral instruments, a *fair price* of the issued securities can be determined, for example, by means of a *risk-neutral valuation* approach.

In Section 8.6 we will see that CDO tranches based on derivatives constitute a very much standardized market instrument today.

Another class of CDOs gaining much attention are *multisector CDOs*. In this case, the collateral pool is a *mixture* of different ABS bonds, high-yield bonds or loans, CDO pieces, mortgage-backed securities, and other assets. Multisector CDOs are more difficult to analyze, mainly due to *cross-collateralization* effects, essentially meaning that bonds issued by a distressed company could be contained in more than one instrument in the collateral pool. For example, "fallen angels" (like Enron quite some years ago) typically cause performance difficulties simultaneously to all CDOs containing this particular risk. Cross-collateralization can only be treated by looking at the union of all collateral pools of all instruments in the multisector pool simultaneously, in order to get an idea about the *aggregated* risk of the combined portfolios. Then, based on every aggregated scenario in a Monte Carlo simulation, the cash flows of the different instruments have to be collected and combined in order to investigate the structured cash flows of the multisector CDO.

The credit risk modeling techniques explained in this book can be used for modeling (multisector) CDOs. Of course, a sound factor model, like the one explained in Section 1.2.3, is a necessary prerequisite for modeling CDOs by taking industry and country diversification effects into account. Moreover, in many cases one additionally has to incorporate an *interest rate term structure model* in order to capture interest rate risk in case of floating rate notes.

In general, *Market value* CDOs are more difficult to treat from a modeling point of view. These structures are more comparable to *hedge funds* than to traditional ABS structures. In a market value CDO, the *portfolio manager* has the right to freely trade the collateral. As a consequence, the portfolio constituents of a market value CDO today have sometimes very little overlap with the portfolio constituents a few months later. The performance of market value CDOs completely relies on the portfolio manager's expertise to trade the collateral in a way meeting the principal and interest obligations of the structure. Therefore, investors will mainly focus on the manager's *deal track record* and experience when deciding about an investment in the structure. The difficulties on the modeling side arise from the unknown *trading strategy* of the portfolio manager and the need of the manager to react to a *volatile economic environment*. Such subjective aspects are difficult if not impossible to model and will not be treated here.

8.1.1 Typical Cash Flow CDO Structure

In this section, we explain a typical cash flow CDO transaction; see Figure 8.2. For this purpose we focus on some of the main aspects without going too much into details.

- At the beginning there will always be some pool of credit risky assets. Admittedly, it will not always be the case that the pool intended to be securitized was existent at the originating bank's

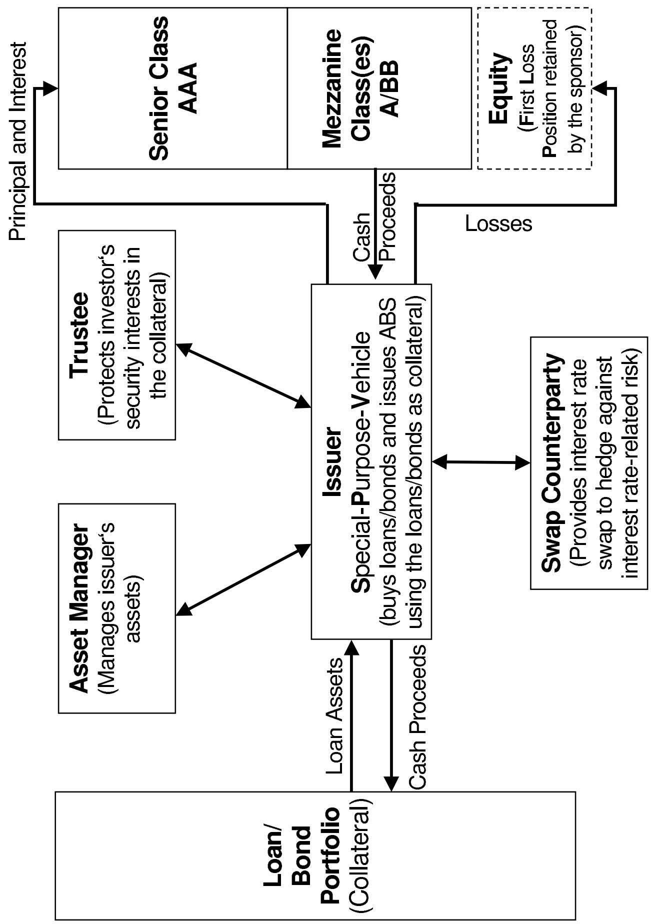


FIGURE 8.2: Example of a cash flow CDO transaction.

balance sheet for a long time; instead, there are many cases where banks purchased parts of the pool just a few months before launching the transaction. Such purchases are typically done in a so-called *ramp-up* period.

- In the next step, the assets are transferred to an SPV, which is a company set-up especially for the purpose of the transaction. This explains the notion *special purpose* vehicle. An important condition hereby is the *bankruptcy remoteness* of the SPV, essentially meaning that the SPV's own bankruptcy risk is minimized and that the SPV will not default on its obligations because of bankruptcy or insolvency of the originator. This is achieved by a strict legal separation between the SPV and the originator, implying legal and economic independence. Additionally, an SPV's obligations typically involve various structural features supporting the bankruptcy remoteness of the SPV.

  In case of cash flow structures, a "true-sale" of the assets from the originator to the SPV completely removes the securitized assets from the originator's balance sheet. However, most often the administration of the asset pool remains the originator's responsibility. The originator receives the principal balance of the pool as cash proceeds, such that from the originator's point of view the funding of the asset pool is completed.

- After the true sale, the assets are property of the SPV. Therefore, the SPV is the owner of all of the cash flows arising from the asset pool. This can be used to establish a funding source for the SPV's purchase of assets from the originator. Note that as a special purpose company, the SPV itself has no money for paying the principal balance of the asset pool to the originating institution. A way out is the issuance of securities or *structured notes* backed[2] by the cash flow of the asset pool. In other words, the SPV now issues notes to investors, such that the total notional of notes reflects the principal balance of the pool. Interest and principal for the notes are paid from interest and principal proceeds from the asset pool. This mechanism changes the meaning of the asset pool towards a *collateral* pool. From the issuance of notes, the

[2]This perfectly explains the name *asset backed* securities.

SPV receives cash proceeds from the investors, refinancing the original purchase of assets from the originating institution.

Because investor's proceeds (principal and interest) are paid from cash flows generated by the collateral pool, investors are taking the performance risk of the collateral pool. Because different investors have different risk appetite, the notes issued by an SPV are typically *tranched* into different *risk classes*. The *first loss piece* (FLP), often also called the *equity* piece[3] is the most *subordinated* tranche, receiving interest and principal payments only if all other notes investors received their contractually promised payments.

The FLP is followed by *junior*, *mezzanine*, and *senior* tranches, receiving interest and principal proceeds in the order of their seniority: Most senior note holders receive their payments first, more junior note investors receive payments only if more prioritized payments are in line with the documentation of the structure. Therefore, the most senior tranche always is the safest investment, carrying the lowest coupon. The more junior a tranche is, the higher the promised coupon, compensating investors for the taken risk.

An exception is the equity tranche, which typically carries no promised coupon. Instead, equity investors receive the *excess spread* of the structure in every payment period, where the excess spread is the cash left over after paying all fees of the structure and all payments to notes investors senior to the equity piece.

From the discussion above, it follows that *subordination* is kind of a *structural credit enhancement*. For example, in a structure with only one equity, mezzanine, and senior tranche, the senior note holders are protected by a cushion consisting of the equity and mezzanine capital, and the mezzanine tranche is protected by the equity tranche.

Figure 8.3 describes the interest and principal proceeds "waterfalls" in a typical cash flow CDO. The figure also indicates the *deleveraging* mechanism inherent in CDO structures, realized by *overcollateralization* (O/C) and *interest coverage* (I/C) tests, which brings us to our last topic in this section.

But before continuing we should mention that there are additional parties involved in a CDO transaction, including

[3]The equity tranche is sometimes kept by the originating institution, therefore constituting equity.

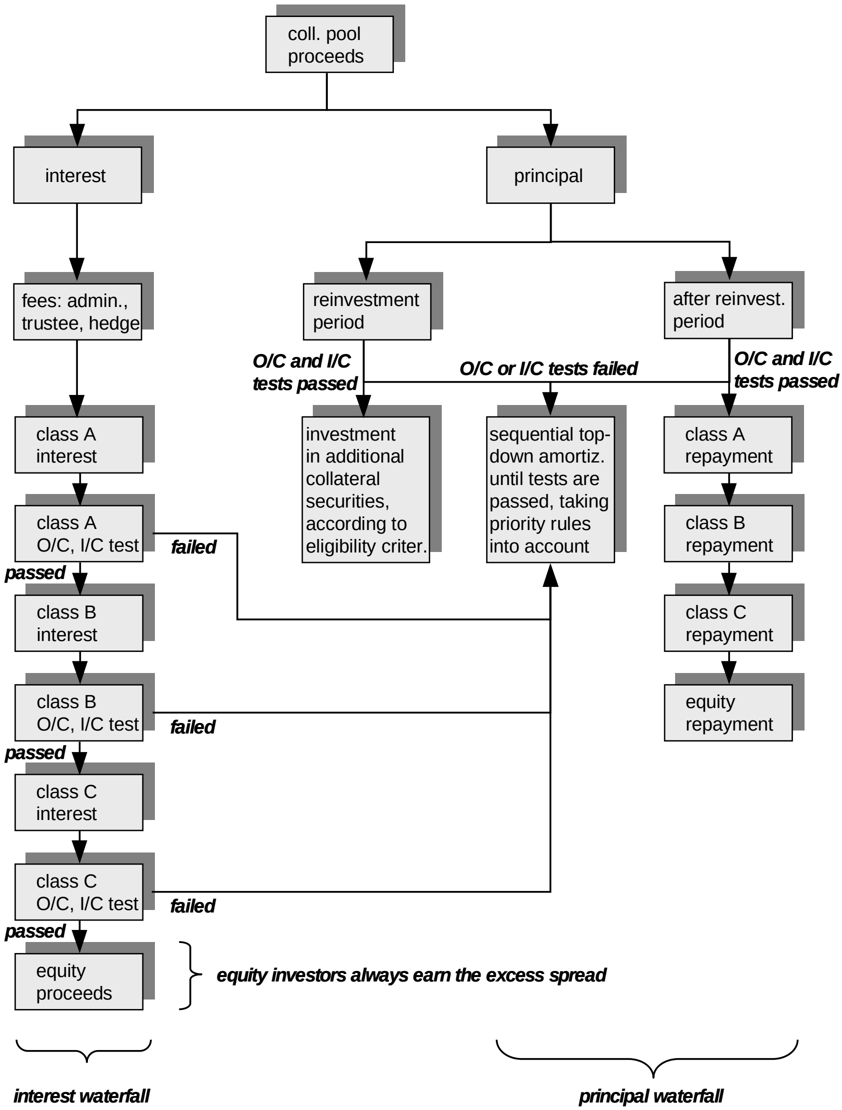


FIGURE 8.3: Example of cash flow waterfalls in a cash flow CDO.

TABLE 8.1: CDO example (illustrative only).

Class	Notional in USD	Tranche Size	Moody's Rating	Spread ov. LIBOR (bps)	minimum O/C ratio	minimum I/C ratio
A	225,000,000	75%	Aa2	100	120%	140%
B	30,000,000	10%	Baa3	250	110%	125%
C	15,000,000	5%	Ba1	550	105%	110%
Equity	30,000,000	10%	Not Rated	-	-	-

- rating agencies, which assign ratings to the issued notes,
- a trustee, which takes care that the legal documentation is honored and receives monthly trustee reports regarding the current performance of the structure,
- some swap counterparties in case interest or currency risk has to be hedged, and
- lawyers, structuring experts, and underwriters at the beginning of the transaction, where the latter mentioned are hired from another investment bank or from inhouse business units.

Now, in order to explain the O/C and I/C mechanisms in a cash flow CDO, let us consider a simple illustrative example. Let us assume we are given a structure like the one outlined in Table[4] 8.1. Further we assume that

- the collateral pool contains 100 corporate bonds with an average default probability

  $$\text{PD} = 3\% ,$$

  and a weighted[5] average coupon (WAC) of

  $$\text{WAC} = 10.4\% ,$$

  reflecting the risk inherent in the collateral securities;
- spreads and default probabilities are annualized values. The following discussion is independent of the maturity of the structure.

[4]In the table, LIBOR refers to the 3-month *London Interbank Offered Rate*, which is a widely used benchmark or reference rate for short term interest rates.

[5]For reasons of simplicity, assuming that the bonds trade at par, the weighting is done with respect to the principal values of the bonds.

Now we are ready for explaining the O/C and I/C mechanisms. Basically these *coverage tests* are intended as an early warning (automatically redirecting cash flows) that interest or principal proceeds are running short for covering the notes coupons and/or repayments. In case of a broken coverage test, principal and interest proceeds are used for paying back the outstandings on notes sequentially (senior tranches first, mezzanine and junior tranches later) until all tests are passed again. This *deleveraging* mechanism of the structure reduces the exposure at risk for tranches in order of their seniority. So one can think of coverage tests as some kind of *credit enhancement* for protecting notes investors (according to prioritization rules) from suffering a loss (a missed coupon[6] or a repayment below par).

8.1.1.1 Overcollateralization Tests

In these tests, which are done for every single tranche except equity, the principal coverage of collateral securities compared to the required amount for paying back the notional of the considered tranche *and* the tranches senior to the considered tranche is tested. In the structure according to Table 8.1, three O/C tests have to be done:

O/C test for class A notes: Denote the par value of the pool by PV_{Pool} and the par value of class A notes by PV_A, where par values are taken w.r.t. the considered payment period in which the test is done. (Synonymously to "par value" we could also say "outstandings" on notes.) Define

$$(\text{O/C})_A \;=\; \frac{\text{PV}_{Pool}}{\text{PV}_A}\;.$$

The O/C test for class A notes is passed if

$$(\text{O/C})_A \;\geq\; (\text{O/C})_A^{\min} \;=\; 120\%$$

reflecting the minimum O/C ratio for class A as in Table 8.1.

O/C test for class B notes: Define

$$(\text{O/C})_B \;=\; \frac{\text{PV}_{Pool}}{\text{PV}_A + \text{PV}_B}\;.$$

[6]For mezzanine investors often a *deferred interest* is possible: If the cash flow from the collateral securities is not sufficient for passing the coverage tests, mezzanine investor's coupon payments are deferred to a later payment period, where all tests are in line again. Deferred interest is paid on an accrued basis.

TABLE 8.2: O/C ratios example (illustrative only).

Portfolio Par	O/C_A	O/C_B	O/C_C
300,000,000	133%	118%	111%
275,000,000	122%	108%	102%

The O/C test for class B is passed if

$$(\mathrm{O/C})_B \;\geq\; (\mathrm{O/C})_B^{\min} \;=\; 110\% \;.$$

Note that the O/C test for class B takes into account that class A notes have to be paid back before.

O/C test for class C notes: Set

$$(\mathrm{O/C})_C \;=\; \frac{\mathrm{PV}_{Pool}}{\mathrm{PV}_A + \mathrm{PV}_B + \mathrm{PV}_C} \;.$$

The O/C test for class C investors is passed if

$$(\mathrm{O/C})_C \;\geq\; (\mathrm{O/C})_C^{\min} \;=\; 105\% \;.$$

Note that the O/C test for class C takes into account that the outstandings of classes A and B have to be paid back before class C investors get their invested money back; see also Figure 8.3.

To give an example, assume we are in a payment period where the pool volume due to losses of 25,000,000 USD melted down to 275,000,000 USD from the previous to the current period. Table 8.2 shows the O/C ratios for the previous and the current period. One can see that the coverage is still sufficient for class A to pass the test, but insufficient for classes B and C. Their O/C tests are broken. This will cause a deleveraging of the CDO until all tests are in line again.

8.1.1.2 Interest Coverage Tests

An I/C test for a tranche basically measures if the interest proceeds from the collateral pool are sufficient for paying the fees and coupons of the structure. In our particular example there are three tests:

I/C test for class A notes: For the considered payment period, denote the par value of the pool by PV_{Pool}, the par value of class

A by PV_A, the amount of required annual fees by FEES, the weighted average coupon of the pool by WAC, and the coupon[7] on class A notes by C_A. Define

$$(\text{I/C})_A = \frac{\text{PV}_{Pool} \times \text{WAC} \times 0.5 - \text{FEES} \times 0.5}{\text{PV}_A \times \text{C}_A \times 0.5} .$$

Here, the factor 0.5 reflects that interest is calculated w.r.t. a semiannual horizon, covering two (quarterly) payment periods. Of course, the concrete calculation method for I/C and O/C ratios always has to be looked up in the documentation of the structure. The I/C test for class A notes is passed if

$$(\text{I/C})_A \geq (\text{I/C})_A^{\min} = 140\%$$

reflecting the minimum required I/C ratio for class A notes according to Table 8.1.

I/C test for class B notes: Define

$$(\text{I/C})_B = \frac{\text{PV}_{Pool} \times \text{WAC} \times 0.5 - \text{FEES} \times 0.5}{(\text{PV}_A \times \text{C}_A + \text{PV}_B \times \text{C}_B) \times 0.5} .$$

The I/C test for class B is passed if

$$(\text{I/C})_B \geq (\text{I/C})_B^{\min} = 125\% .$$

Analogous to the O/C tests, the calculation reflects that class A notes have priority before class B notes regarding coupon payments.

I/C test for class C notes: Setting

$$(\text{I/C})_C = \frac{\text{PV}_{Pool} \times \text{WAC} \times 0.5 - \text{FEES} \times 0.5}{(\text{PV}_A \times \text{C}_A + \text{PV}_B \times \text{C}_B + \text{PV}_C \times \text{C}_C) \times 0.5} ,$$

class C interest coverage requires classes A and B to be covered, before C-notes investors receive coupon payments. The test is passed if

$$(\text{I/C})_C \geq (\text{I/C})_C^{\min} = 110\% .$$

The interest waterfall as illustrated in Figure 8.3 is clearly reflected by these calculations.

[7] In our example we are dealing with floating-rate notes. Here, the coupon on notes is always defined as LIBOR+Spread.

TABLE 8.3: I/C ratios example (illustrative only).

Portfolio Par	I/C_A	I/C_B	I/C_C
300,000,000	273%	233%	210%

Total Fees	450,000	USD p.a.

Table 8.3 gives an example for the value of the three I/C ratios right at the beginning of the transaction. For calculating the I/C ratios we assumed the current 3-month LIBOR to be equal to 4%.

8.1.1.3 Other Tests

Typically, there are some more tests whose outcomes have to be reported to the trustee and to the investors. The collection of tests and criteria varies from deal to deal, and not all tests included in the monthly reports automatically have immediate consequences on the cash flow side of the structure. Some tests one frequently finds in deal documentations are

- an *average rating floor* test, reporting whether the weighted average rating of the collateral pool is above a critical threshold; a typical threshold for cash flow CDOs is Moody's B-rating;

- *industry* and *diversity score*[8] tests, alarming investors in case the industry diversification of the collateral pool decreased more than expected; a common range for the highest admissible industry concentration is 8-12%;

- an *obligor concentration* test, measuring the highest exposure concentration in the collateral pool, often restricted to concentrations below 3%,

and possibly some more tests helping the investors to identify and quantify potential sources of risk and financial distress of the structure.

This concludes our description of cash flow CDOs for now.

[8]Diversity scores are a measure of the industry diversification of a portfolio; diversity scores are due to Moody's and will be explained in Section 8.5.

8.1.2 Typical Synthetic CLO Structure

In contrast to cash flow CDOs, synthetic CDOs do not rely on the cash flows of the collateral pool. Instead, credit derivatives, e.g., credit-linked notes, are used to link the performance of securities issued by an SPV to the performance, e.g., the losses, of some *reference pool*. In other words, synthetic CDOs do not include a true sale, as we just discussed it in case of cash flow deals, but rather leave the reference assets on the originator's balance sheet. Figure 8.4 shows a typical synthetic CLO as we find it in the market. It could work as follows:

- The originator buys protection for *super senior* and *junior*[9] pieces of the reference portfolio by entering into two credit default swaps with some swap counterparties (protection sellers, see Chapter 7). The volume referring to the two swaps is called the *unfunded part*, because there is no sale requiring certain sources of funding.

- An SPV, which has to be bankruptcy remote for regulatory reasons, enters into a swap with the originator for the volume of the reference portfolio which is not covered by the senior and junior swaps.

- In order to guarantee the contingent payments to the originator in case of credit events in the reference pool, the SPV has to invest some money in collateral securities. Then, in case of a credit event, the contingent payments from the SPV (protection seller) to the originator (protection buyer) can be funded by selling collateral securities in an amount matching the realized losses in the reference portfolio.

- For purchasing collateral securities, the SPV needs some source of funding. In the same way as we already saw it for cash flow deals, the SPV issues credit-linked notes in the capital market, linked to the performance of the reference pool. The outstandings of the issued notes match the volume of the reference pool reduced by the size of the junior and senior swap tranches. The SPV invests the cash proceeds from issuing the notes in low-risk (AAA-quality) or riskless (treasury) notes.

[9]Selling the FLP of a synthetic transaction to an investor often involves a so-called *interest sub-participation*, essentially meaning that part of the reference pool's interest will be made available to the FLP-investor in case of losses.

- The spreads on the notes the SPV has to pay to notes investors match the premium the SPV receives from the originator who bought protection from the SPV for the funded part of the reference portfolio.

- If a credit event in the reference pool occurs, who has to pay for it depends on the already cumulated losses. Losses below the upper limit of the junior tranche are carried by the junior swap counterparty. Losses exceeding the junior piece but below the super senior tranche are (additionally to the contingent payment made by the junior swap protection seller) carried by the SPV from the originator's point of view and carried by the investor's from the SPV's point of view. Indeed, because collateral securities will be sold for funding the contingent payment the SPV has to make to the originator, investors will not get the complete face value of their invested money back at the final maturity of the structure. The more junior the notes, the more likely it is that investors will not be fully repaid. Super senior losses, which refer to loss events far out in the tail of the reference portfolio's loss distribution, are taken by the super senior swap counterparty. If super senior swap counterparties have to pay for losses, all subordinated investors already had to make their contingent payments on the swap agreements.

Please note that in the market one finds all kinds of variations of the illustrative synthetic CLO we just described. For example, instead of credit default swaps some form of *financial guarantee* could be used. In some cases there will be a *non-cash settlement* in that the protection buyer sells the defaulted loan to the protection seller for par right after a contractually specified credit event occurred.

Additionally, most synthetic structures involve *triggers* based on, e.g., rating distributions, diversity scores, collateral values, losses, defaults, etc. For example, a loss trigger could be defined by saying that in case losses exceed a critical threshold ("trigger event"), some structural features of the transaction change in a contractually pre-specified manner. In this way, triggers are structural elements providing protection to note holders, comparable to the coverage tests discussed above.

In the last years, many new innovative structures offered interesting investment opportunities. Due to inefficient markets and regulatory arbitrage (see our discussion in Section 8.2), this trend can be expected to continue.

Because in our example there is a funded *and* an unfunded part of the transaction, such a synthetic CLO is called *partially funded* accordingly. Again we should remark that all variations are possible and existent in the market: Fully funded, partially funded, and totally unfunded.

Agreements regarding the definition of credit events and the settlement after the occurrence of credit events can be made based on ISDA master agreements, see also Chapter 7.

The *tranching* of the reference portfolio into junior, funded, and super senior parts follows analogous rules, as we just saw in the case of cash flow CDOs. The more junior a note is, the higher the premium paid for buying protection for the considered tranche. The more senior a tranche is, the safer investors can invest, but the lower the premium they earn on the investment. Note that this is in line with the risk-adjusted pricing of swap contracts, see Chapter 7.

8.2 Different Roles of Banks in the CDO Market

There are in general many roles of banks in the ABS market. In many cases, a bank will play the role of the *originator* or the role of the *investor*. But there are certainly other roles, which will not be discussed in this book. For example, banks also *provide liquidity*, *guarantee* for promised cash flows, offer different types of *credit enhancement*, and sell their services for *structuring* or *underwriting* ABS transactions. Of course, different roles require different models, so that in general one can say that parallel to the ABS market a whole range of models is needed to measure the different risks the bank is exposed to when participating in the ABS market. In the following section we discuss origination, and in a short subsequent section we make some remarks on ABS investments.

8.2.1 The Originator's Point of View

This section discusses *securitization*. The original meaning of the word "securitization" is funding by means of issuing (structured) se-

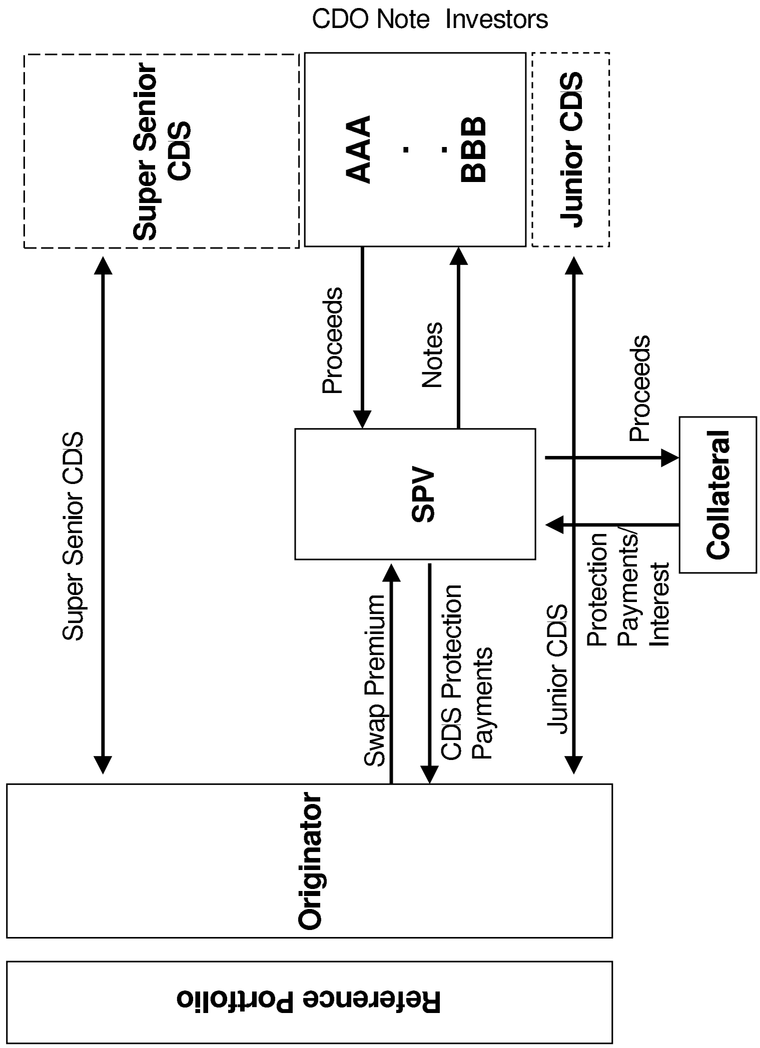


FIGURE 8.4: Example of a synthetic CDO transaction.

curities. Today, banks mainly do securitizations for several reasons, including

- *transferring risk*,
- *arbitrage spread* opportunities.
- *funding* at better (cheaper) conditions, and
- exploitation of *regulatory* or *tax arbitrage*,

From a portfolio modeling point of view, there is a fundamental difference between the first three and the last securitization benefits: Risk transfer, arbitrage spread opportunities, and (to some extent) better funding are *correlation-driven* effects, whereas regulatory capital relief and tax arbitrage are *correlation-free* effects.

8.2.1.1 Regulatory Arbitrage and Capital Relief

The keyword *regulatory arbitrage* refers to opportunities in the markets due to inappropriate regulation by the regulatory authorities. For example, as indicated in Section 1.3, regulatory capital for bank loans in line with the Basel II accord is not suitable for the treatment of correlation products like CDOs. This results in pricing distortions, due to the fact that the capital costs of a loan are not fully aligned with the credit quality of the borrower.

In this section we restrict our exposition to a few remarks only. *The part of the Basel II capital accord treating securitizations is expected to be modified in the near future because the framework is just not appropriate* as already indicated in Section 1.3. Therefore, it does not make sense to spend much time with a discussion on the current framework.

A recent trend in regulatory arbitrage is based on the so-called *supervisory formula* which is part of the *securitization framework* of the Basel II capital accord. In Section 1.3 we showed an example for a *corporate risk weight function* which then leads to the risk-weighted assets RWA of a corporate loan. Let us now assume a bank securitizes a subportfolio of its loan book. Then, looking at the securitization a bit simplified, the regulatory capital requirement of the subportfolio *before* securitization is given by

$$\text{Reg.Cap}_{before} \;=\; 8\% \times \sum_{i=1}^{m} \text{RWA}_i$$

where m is the number of securitized loans. A securitization has a positive capital effect if

$$\text{Reg.Cap}_{before} \; - \; \text{Reg.Cap}_{after} \; > \; 0.$$

Obviously, this is not enough to make a securitization motivated by capital relief a success because one has to *balance reduced capital cost against the cost of securitizing the assets*.

The capital requirement *after* securitization depends on the particular type of securitization. For instance, if the tranches are rated (except equity) then the so-called *ratings-based approach* (RBA) which is part of the Basel II securitization framework determines risk weights for tranches. The bank must hold capital for any tranche the bank is still holding after securitization. If the securitization is based on unrated tranches then the afore-mentioned *supervisory formula* is used for the determination of RWA of tranches. It is an open secret that banks exploit the fact that one can free-up capital comparably cheap by a replacement of the usual Basel II corporate risk weights by risk weights coming from a securitization treated by means of the supervisory formula. The optimization problem of such a *regulatory arbitrage opportunity* is to determine an optimal attachement point (lower boundary of the tranche) and detachment point (upper boundary of the tranche) such that for the given portfolio a *single-tranche securitization* gives the optimal capital relief when balancing capital relief against securitization cost.

8.2.1.2 Economic Risk Transfer

A securitization should also reduce the credit risk of the portfolio where credit risk can be quantified in various ways as we have seen in Chapter 1. For instance one can consider the EL or EC before and after securitization. Risk reduction then is in line with the *insurance* paradigm which originally motivated the notion of EL many years ago. After securitizing the portfolio there is no longer the need to have an insurance against the full loss potential of the portfolio. Instead, the bank is only exposed to the risk coming from kept tranches.

Moreover, the same argument conceptually also holds for the economic capital (EC; see Section 1.2.1) of the securitized portfolio. But because the EC involves correlations (and therefore incorporates *diversification* effects), a securitization not only impacts the securitized pool, but also influences the EC of the *source portfolio*, from which the securitized pool has been separated. In more mathematical terms we have the following situation

Denote by $I = \{1, ..., m\}$ an index set referring to the loans in the source portfolio, and let us assume that a subportfolio indexed by $S = \{i_1, ..., i_q\} \subset I$ has been selected for securitization by means of a CLO. The bank now wants to quantify the securitization impact on the source portfolio. For this purpose, the portfolio's EL and EC have to be re-calculated after the portfolio shrinking $I \to I \backslash S$.

Now, based on Monte Carlo simulation techniques, the securitization impact is not difficult to calculate. Let us assume the bank would manage to sell all tranches of the CLO except the equity piece, which is then held by the bank. For reasons of simplicity we consider the one-year period from the launch of the deal until one year later. The size of the equity piece, FLP, is a random variable due to the uncertainty regarding the performance of the collateral securities. Denoting the loss statistics of the whole portfolio I by $(L_1, ..., L_m)$, see Chapter 2, the gross portfolio loss *before* the securitization transaction equals

$$L \;=\; \sum_{i=1}^{m} L_i \;,$$

hereby assuming an LGD of 100% and exposures equal to 1 for reasons of simplicity. The portfolio's gross loss *after* securitization obviously is given by

$$L_{Sec} \;=\; \sum_{i \in I \backslash S} L_i \;+\; \min\Big(\sum_{k=1}^{q} L_{i_k}, \mathrm{FLP} \Big) \;, \tag{8.1}$$

because the securitized portfolio S is protected against losses exceeding FLP. But the variables L and L_{Sec} can be easily simulated by use of the Monte Carlo engine of the bank. After simulation, we have a loss distribution of the portfolio *before* the transaction and a loss distribution of the portfolio *after* the securitization. The *expected loss gain*, respectively *economic capital gain*, of the transaction is given by

$$\Delta\mathrm{EL} \;=\; \mathbb{E}[L] - \mathbb{E}[L_{Sec}]$$

and

$$\Delta\mathrm{EC}_\alpha \;=\; \mathrm{EC}_\alpha(L) - \mathrm{EC}_\alpha(L_{Sec}) \;=$$

$$= (q_\alpha(L) - \mathbb{E}[L]) - (q_\alpha(L_{Sec}) - \mathbb{E}[L_{Sec}]) = \Delta q_\alpha - \Delta\mathrm{EL} ,$$

where q_α denotes the α-quantile of the respective loss distribution (before and after securitization, respectively), and

$$\Delta q_\alpha = q_\alpha(L) - q_\alpha(L_{Sec}) .$$

These calculations are sufficient for capturing the securitization impact on the source portfolio.

"Risk transfer" refers to the possibility to reduce the required capital cushion against losses of a portfolio by means of a securitization. Economic risk transfer happens, if the risk transfer can be measured in terms of the EL and EC, such that ΔEL and Δq_α are positive.

Now assume that the securitized pool S belongs to some business unit of the bank with its own profit center. Then, the securitization impact additionally has to be measured from that profit center's point of view, so we additionally need to quantify EL and EC benefits for the securitized pool only. Keeping the notation from above, the pool loss before and after securitization is

$$L_{Pool} = \sum_{k=1}^{q} L_{i_k} \qquad \text{and} \qquad L_{Pool}^{Sec} = \min\Big(\sum_{k=1}^{q} L_{i_k}, \mathrm{FLP}\Big) .$$

The pool's EL benefit of the securitization therefore is

$$\Delta\mathrm{EL}_{Pool} = \mathbb{E}[L_{Pool}] - \mathbb{E}[L_{Pool}^{Sec}] .$$

Obviously, $\Delta\mathrm{EL}_{Pool}$ is positive if and only if there is at least one cumulative loss path[10] for which the cap at FLP turns out to be effective.

Regarding EC, we now have to consider the *contributory economic capital* (CEC) of the securitized pool w.r.t. the source portfolio; see also Section 5.2. The gain in CEC of the securitization is given by

$$\Delta\mathrm{CEC}_\alpha = \mathrm{CEC}_{Pool} - \mathrm{CEC}_{Pool}^{Sec} ,$$

where CEC_{Pool} and $\mathrm{CEC}_{Pool}^{Sec}$ denotes the CEC of the securitized pool before and after securitization.

[10]Thinking in terms of a Monte Carlo simulation.

We conclude our discussion by briefly mentioning a common performance measure capturing the effects of economic risk transfer, namely *risk-adjusted return on capital* (RAROC). There are various definitions of RAROC measures in the literature, but here we use it just for illustrative purposes and therefore keep things as simple[11] as possible. More or less, RAROC always is defined as the risk-adjusted return of an instrument or portfolio divided by the corresponding EC. To illustrate the effect of securitization to RAROC benchmarks, let us assume that the CEC of the pool before securitization was $\text{CEC}_{Pool} = 5\%$. Let us further assume that after securitization the CEC of the pool melted down to $\text{CEC}^{Sec}_{Pool} = 150$ bps. The EL of the pool is assumed to be reduced from $\mathbb{E}[L_{Pool}] = 50$ bps to $\mathbb{E}[L^{Sec}_{Pool}] = 40$ bps, due to the securitization. This yields

$$\text{RAROC} \;=\; \frac{\text{NM} - \mathbb{E}[L_{Pool}]}{\text{CEC}_{Pool}} \;=\; \frac{130 - 50}{500} \;=\; 16\%$$

for the portfolio before the securitization transaction, and

$$\text{RAROC}_{Sec} \;=\; \frac{\text{NM} - \mathbb{E}[L^{Sec}_{Pool}] - \text{COST}}{\text{CEC}^{Sec}_{Pool}} \;=\; \frac{130 - 40 - 30}{150} \;=\; 40\%$$

after securitizing the portfolio. So the securitization improves the RAROC of the portfolio by a factor of 2.5, just due to the protection limit of 1.5%.

Note that the discussion above was based on a one-period view, e.g., based on an average lifetime consideration. For measuring economic risk transfer and securitization effects on RAROC more accurately, much more work and modeling efforts are required, very often accompanied by strong assumptions, e.g., regarding the evolution of the reference pool.

8.2.1.3 Funding at Better Conditions

Funding is an important issue for banks. Because every loan needs to be backed by regulatory capital, the capital costs associated with a loan to a customer can be too high for making the lending business profitable. But if loans are pooled into portfolios for securitization,

[11] For example, we do not take, as is often done, the *capital benefit* arising from risk free interest earned on the EC into account.

funding a loan can get significantly cheaper. The reasons why a securitization makes funding cheaper, are basically given in the two sections above: Because regulatory capital is relieved, equity costs of the securitized portfolio are much lower than they used to be. Moreover, if an economic risk transfer is achieved, EL costs and EC costs will be reduced to an extent reflecting the amount of risk transferred to the capital market. Both effects, and additional tax and other benefits can help a bank to refinance a loan portfolio at much better conditions than was the case before the securitization.

8.2.1.4 Arbitrage Spread Opportunities

Arbitrage spread opportunities are created in the following way. The assets in the collateral pool (in our example of a cash flow CDO we are talking about high-yield bonds) are priced on a *single* asset basis, such that every bond coupon in the portfolio reflects the risk of the bond. Of course, in general, the coupon of a bond *and* its price provide the "full" information about the risks inherent in the bond. However, assuming a bond trades at par allows for taking the coupon of the bond as a proxy for its riskiness. So the WAC of the collateral pool really is a weighted sum of *single* asset risks, ignoring the potential for diversification effects typically inherent in a portfolio.

In contrast, on the CDO side, it is the *portfolio* risk which endangers the performance of the structure. Recalling our discussion on cash flow CDOs, we see that the tranching of notes really is a *tranching of the loss distribution* of the collateral pool, taking all possible diversification effects into account. But diversification decreases the risk of a portfolio, so that the price of the portfolio risk must be lower than the price obtained by just summing up exposure-weighted single risks. This is reflected by the spreads on notes as given in Table 8.1: The spreads paid to notes investors are much lower than the spreads earned on the bonds in the collateral pool. Due to the risk tranching of notes, the spreads on senior notes is even lower, due to the credit enhancement by subordination provided from notes with lower seniority.

It is exactly the mismatch between the single asset based WAC of the portfolio and the much lower weighted average coupon on the notes of the CDO, which creates an arbitrage spread. This mismatch is in one part due to diversification effects, and in another part based on structural elements like subordination or other credit enhancement mechanisms. Calling special attention to the diversification point we already said earlier that *CDOs are "correlation products."*

An example regarding arbitrage spread is given in the next section in the context of CDO investments. Conceptually, any originator of an arbitrage cash flow CDO keeping the CDO's first loss piece automatically takes on the role of the *equity investor*, earning the *excess spread* of the structure in its own pockets. Therefore, we can postpone the arbitrage spread example to the next section.

8.2.2 The Investor's Point of View

Very often banks are on the investment side of a CDO. In many cases, ABS bonds offer interesting and attractive investment opportunities, but require (due to their complexity) careful analytic valuation methods for calculating the risks and benefits coming with an ABS investment into the bank's portfolio. This will be made explicit by means of the following example.

Recall the sample cash flow CDO from Table 8.1. In this example we assumed WAC = 10.4% and PD = 3%. Assuming an LGD of 80% on the collateral securities, we obtain the portfolio's expected loss, EL = 3% × 80% = 2.4%.

Considering the CDO from an *expected return* point of view, what would an equity investor expect to earn on an investment in the equity tranche? A typical "back-of-the-envelope" calculation reads as follows: From Table 8.1 we obtain the weighted average coupon WAC_{Notes} of the structure as

$$\text{WAC}_{Notes} \;=\; 75\% \times 5\% + 10\% \times 6.5\% + 5\% \times 9.5\% \;=\; 4.875\% \;,$$

again assuming the average 3-month LIBOR to be equal to 4%. Because cash flow CDOs completely rely on the cash flows from the collateral pool, the 10.4% of the pool's par value are the complete income of the structure. From this income, all expenses of the structure have to be paid. Paying[12] coupons to notes investors yields a *gross arbitrage spread* (gross excess spread) of

$$[\text{Pool Income}] - [\text{Notes Spreads}] \;=\; 10.4\% - 4.875\% \;=\; 5.525\% \;.$$

[12]Referring to an average scenario.

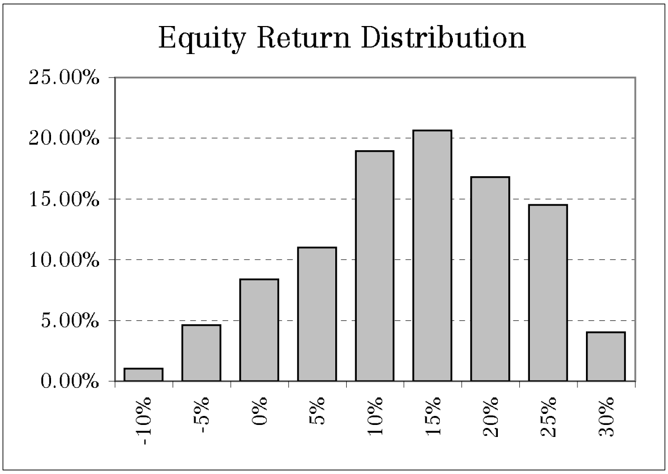


FIGURE 8.5: Equity return distribution of a CDO.

The *expected net excess spread* is then defined as

$$[\text{Gross Arbitrage Spread}] - \text{EL} - \text{COSTS} =$$

$$= 5.525\% - 2.4\% - \frac{450,000}{300,000,000} = 2.975\% .$$

The *equity return* is then given by

$$[\text{Exp. Net Excess Spread}] \times \frac{\text{Pool Volume}}{\text{Equity Volume}} = 29.75\% .$$

So the back-of-the-envelope calculation promises a very attractive equity return of almost 30%.

Now let us look at this seemingly attractive investment from a portfolio modeling point of view. For this purpose we calculated the *equity return distribution* of the CDO by means of a *correlated default times* approach as outlined later on in this chapter; see also Chapter 7. From a Monte Carlo simulation we obtained[13] Figure 8.5. Hereby we essentially followed the CDO modeling scheme as illustrated in Figure 8.3, adapted to a default times approach according to Figure 8.10.

[13]Under certain assumptions regarding the maturity of the bonds and the structure.

TABLE 8.4: Return statistics for class-A notes investors

Return Statistics for Class-A Notes	
Return Range	**Relative Frequency**
Return=5%	94.17%
4%<=Return<5%	2.78%
3%<=Return<4%	1.19%
2%<=Return<3%	0.59%
1%<=Return<2%	0.61%
0%<=Return<1%	0.39%
Return<0%	0.27%

Looking at the equity return distribution in Figure 8.5, it turns out that, in contrast to the above shown back-of-the-envelope calculation, the Monte Carlo simulation yields an *average equity return* of only 15.92%. Additionally, the *volatility of equity returns* turns out to be 9.05%, so by just one standard deviation move, the equity return can vary between 6.87% and 24.98%. This reflects the fact that equity investments are rather volatile and therefore very risky. Moreover, due to *tail events* of the collateral pool's loss distribution, it can happen that the *downside risks* of equity investments dominate the *upside chances*.

We continue our example by looking at the return distribution for class-A notes investors. Table 8.4 shows that in 94.17% of the cases the promised coupon of 5% has been paid to A-investors. However, in 5.83% of the cases, either not a full coupon payment or not a full repayment resulted in a loss. Here, loss means that at least one contractually promised dollar has not been paid. So the 5.83% are indeed the (cumulative) *default probability* of the senior tranche of the CDO. For an Aa2-rating, this is a very high chance for default. Additionally, the simulation yields an expected loss of the Aa2-tranche of 50bps, which again is very high compared to Aa2-rated bonds. Defining the loss given default of the tranche by

$$\mathrm{LGD}(T_{Aa2}) \;=\; \frac{\mathrm{EL}(T_{Aa2})}{\mathrm{PD}(T_{Aa2})} \;=\; \frac{50}{583} \;=\; 8.6\% \;,$$

shows that on the other side the LGD of the tranche is very low. This is also due to the large volume (thickness) of the tranche. In Section 8.5 we will discuss rating agency models, and it will turn out that agency

TABLE 8.5: Weighted average life of tranches

Tranche	Weighted Average Life
A	5.03
B	9.84
C	10.00
Equity	10.00

ratings of senior tranches typically underestimate the tranche's "true" risk. This is due to the fact that rating agency models often neglect the *fat tail* of credit portfolio loss distributions. In our example we can clearly see that the Aa2-rating does not really reflect the "true" risk of the Aa2-tranche.

Table 8.5 shows the weighted average life (WAL) of the four tranches. For the simulation, we assumed that the CDO matures in 10 years. The WAL for class-A notes is quite low, in part due to the amortization structure of the collateral pool, but to some extent also due to broken coverage tests leading to a *deleveraging* of the outstandings of the notes. Because of the waterfall structure illustrated in Figure 8.3, the most senior class has to be repaid before lower classes receive repayments. This yields the low WAL for class A.

We conclude this section by a brief summary. In the discussion above, our calculations showed that it is very dangerous to rely on "average value" considerations like our back-of-the-envelope calculation. Only a full Monte Carlo simulation, based on portfolio models as introduced in this book, or alternative techniques as presented in the last section of Chapter 2 will unveil the downside risks and upside chances of an investment in a CDO.

8.3 CDOs from the Modeling Point of View

In this section, a general framework for CDO modeling is presented. Not all structures require all elements mentioned in the sequel. In some cases, shortcuts, approximations, or working assumptions (e.g.

a fixed[14], possibly stress-tested, LIBOR) can be used for evaluating a CDO quicker than by means of implementing a simulation model where all random elements are also drawn at random, hereby increasing the complexity of the model.

In our presentation, we will keep a somewhat abstract level, because going into modeling details or presenting a fully worked-out case study is beyond the introductory scope of this chapter. However, we want to encourage readers[15] involved in ABS transactions to start modeling their deals by means of appropriate mathematical models instead of just following the common practice to evaluate deals by stress tests and the assumption of fixed loss rates. At the end of Chapter 2 we introduced some techniques which proved to be very valuable in CDO modeling. The example in the previous section demonstrates how dangerous shortcut models can be. Credit risk modelers should avoid shortcuts and instead make sure all aspects of a deal are captured by the applied model.

The evaluation of CDO transactions involves three major steps:

1. Step: *Constructing a model for the underlying portfolio*
   Underlying the structure is always an asset pool, for example a reference portfolio or a collateral pool. The structural elements of the considered deal are always linked to the performance of the underlying asset pool, so it is natural to start with a portfolio model similar to those presented in Chapters 1-4. Additionally, such a model should include

   - multi-year horizons due to maturities longer than one year,
   - a sound factor model for measuring industry and country diversification in an appropriate manner, and
   - a model for short term interest rates for capturing the interest rate risk of floating rate securities and notes.

   This first step is the only part involving probability theory. The second and third step are much more elementary.

[14] For example, if in the documentation of a structure one finds that fluctuations of LIBOR are limited by a predefined cap and floor, then one can think of stress testing the impact of LIBOR variations by just looking at the two extreme scenarios.

[15] As far as we know, most major banks use, in addition to the cash flow approaches and rating agency models, CDO models based on Monte Carlo simulation or techniques as explained at the end of Chapter 2.

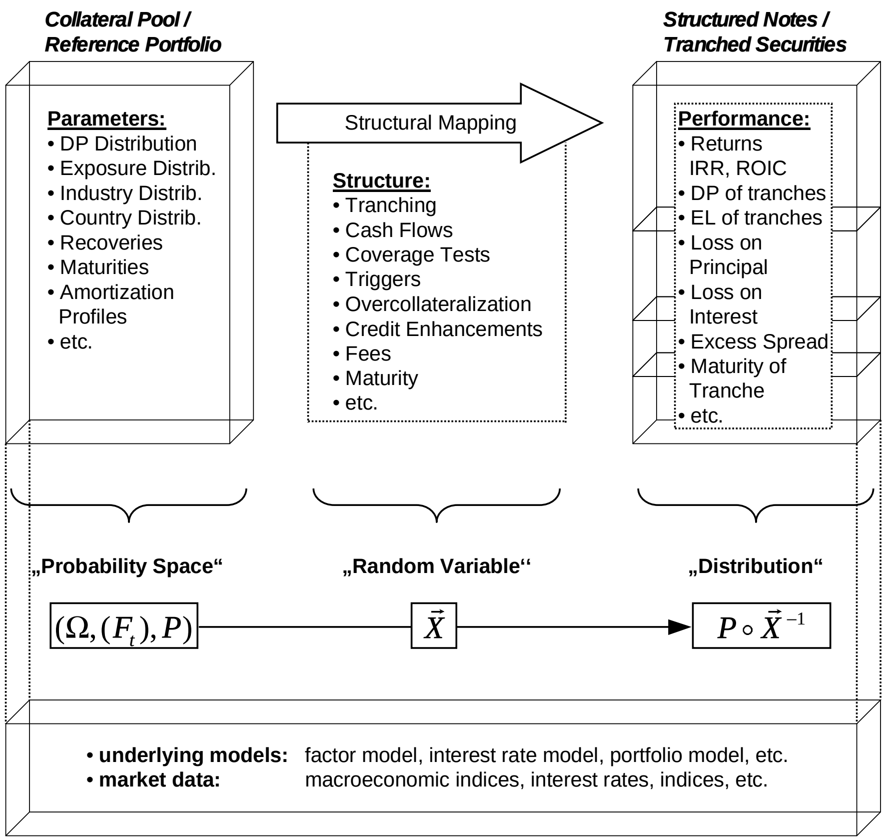


FIGURE 8.6: CDO modeling scheme.

2. Step: *Modeling the cash flows of the structure*
   Based on Step 1, the cash flows of the structure *conditioned* on the simulated scenario from the portfolio model representing the performance of the collateral securities should be modeled by taking all cash flow elements of the structure, including

   - subordination structure,
   - fees and hedge premiums,
   - principal and interest waterfalls,
   - coverage tests (O/C and I/C),
   - credit enhancements (e.g. overcollateralization),
   - triggers (e.g. early amortization, call options), etc.,

   into account. From a programming point of view, Step 2 consists of implementing an algorithm for "distributing money" (e.g., in a cash flow CDO the cash income from the collateral securities) into accounts (some specified variables reflecting, e.g., principal and interest accounts) defined by the contract or documentation of the deal. Such an algorithm should exactly reflect the cash flow mechanisms specified in the documentation, because leaving out just a single element can already significantly distort the simulation results towards wrong impressions regarding the performance of the structure. In addition to a cash flow model, a discounting method (e.g., a risk-neutral valuation model in case the risks, e.g., the default probabilities, of the collateral securities are determined according to a risk-neutral approach) should be in place in order to calculate present values of future cash flows

3. Step: *Interpreting the outcome of the simulation engine*
   After the simulation, the outcome has to be evaluated and interpreted. Because the performance of the structure is subject to *random fluctuations* based on the randomness of the behavior of the collateral securities, the basic outcome of the simulation will always consist of *distributions* (e.g., return distributions, loss distributions, etc.); see Figure 8.5 and the discussion there.

Figure 8.6 illustrates and summarizes the three steps by means of a modeling scheme.

In the more formal language of mathematical modeling the three major valuation steps discussed above can be described as follows:

Step 1 defines a *filtered probability space* $(\Omega, (\mathcal{F}_t), \mathbb{P})$, where:

- Ω consists of the whole universe of possible scenarios regarding the collateral pool and the interest rate model. More precisely, every scenario $\omega \in \Omega$ is a *vector* whose components are defined by the possible outcomes of the portfolio model, including a default/migration indicator realization for every collateral security, a realization of LIBOR, etc.

- $(\mathcal{F}_t)_{t=1,...,T}$ is a filtration of σ-algebras containing the measurable events up to the payment period t. Any σ-algebra F_t can be interpreted as the collection of events reflecting information known up to payment period t. For example, F_t contains the event that up to time t the portfolio loss already crossed a certain limit, etc. Here, T represents the final maturity of the structure.

- The probability measure $\mathbb{P}$ assigns probabilities to the events in the σ-algebras $\mathcal{F}_t$, $t = 1, ..., T$. For example, the probability that up to time t more than 20% of the collateral securities defaulted is given by $\mathbb{P}(F)$, where $F \in \mathcal{F}_t$ is the corresponding measurable event.

Step 2 defines a *random variable* $\vec{X}$, because as soon as a scenario $\omega \in \Omega$ is fixed by the simulation engine, the distribution of cash flows conditional on ω follows a *deterministic* workflow defined by the documentation of the structure. The variable $\vec{X}$ is a vector whose components contain the quantities relevant for the performance of the structure, e.g., realized returns for notes investors, the amount of realized repayments, the coupon payments made to notes investors, etc. The *distribution* $\mathbb{P} \circ \vec{X}^{-1}$ of the "performance vector" $\vec{X}$ then is the final output, which has to be analyzed and interpreted in Step 3. For example, the *relative frequency* of scenarios in which at least one promised dollar to a mezzanine investor has not been paid, constitutes the *default probability* of that mezzanine tranche.

The filtration $(\mathcal{F}_t)_{t=1,...,T}$ defines a *dynamic information flow* during the simulated lifetime of the deal. For example, the simulation step from time t to time $t+1$ will always be *conditioned* on the already realized *path* (the history up to time t). This very much reflects the approach an investor would follow during the term of a structure: At time t she or he will take all available information up to time t into account for making an analysis regarding the future performance of the structure.

8.4 Multi-Period Credit Models

An important aspect of CDO modeling is the treatment of multi-year horizons. Various multi-period credit models have been put forward during the last years. They differ quite substantially in their setup and hence it comes not as a surprise that their properties and outcomes show quite some disparity even when calibrated to the same input data. Some of the models are mere extensions of single period models, and thus time-discrete by construction; other models are continuous in time. But eventually all models have to be transformed to a discrete time grid as in practical application all happens on consecutive payment dates. FINGER [62] already compared some approaches to multi-period credit modeling and we will follow his presentation to some extent.

8.4.1 Migration Model

Multi-step models are natural extensions of single-period portfolio models, like the models we discussed in previous chapters. Essentially, a multi-step model can be thought of as many "intertemporally connected" single-period models successively simulated. So, one of the simplest extensions is to repeat the one-period default-only model of Chapter 2.4.1 for successive periods. Suppose we have a cumulative default term structure $F_i(t)$ for obligor i, then multiple default thresholds are given as

$$c_{i,t_{k+1}} = N^{-1}\left[\frac{F_i(t_{k+1}) - F_i(t_k)}{1 - F_i(t_k)}\right],$$

$0 = t_0 < t_1 < t_2 < t_3 < \ldots < t_n$. n independent standard normal random variables are drawn for obligor i, $r_{i,t_k} \sim N(0,1)$, and the obligor i defaults in period $]t_{k-1}, t_k]$ if $r_{i,t_l} \geq c_{i,t_l}$ for $l = 1, \ldots, k-1$ and $r_{i,t_k} < c_{i,t_k}$. Note that by construction there is no inter-period correlation of defaults, only within a period the defaults of the obligors are coupled by the Normal copula with a correlation matrix Σ, $(r_{i,t_k})_{i=1,\ldots,n} \sim N(0,\Sigma)$.

Building on the material on credit migrations in Section 6.3.3 this framework can easily be extended to several rating classes. A Markov state, or rating, $Y \in \{1, \dots, K\}$ is assigned to each single credit risky entity i, where the default state K is an absorbing state. A default probability term-structure $F_i(t)$ exists for each initial credit state i together with a sequence of migration matrices M_{t_k} that is adapted to meet the term-structure[16] [17]. These transition matrices are chained together and create a discrete credit migration process for each credit entity, $Y^i_{t_k}$, on a time grid $0 = t_0 < t_1 < t_2 < t_3 < \dots < t_n$. In the multi-firm context we add a dependency structure between different credit entities, i.e. credit migrations are coupled through a Gaussian copula function with correlation matrix Σ in each step. From each migration matrix we can now calculate migration thresholds that separate the transition buckets, see Figure 8.7.

For some period t_i the thresholds c_{kl,t_i} are obtained from

$$\begin{aligned}
c_{kl,t_i} &= N^{-1}\left(\sum_{j=l}^{K} M_{kj,t_i}\right), \quad \text{for}\, k,l = 1,\dots,K, \text{ with } \sum_{j=l}^{K} M_{kj,t_i} \neq 0,1 \\
c_{kl,t_i} &= -\infty, \qquad \text{for} \quad k,l = 1,\dots,K, \quad \text{with} \quad \sum_{j=l}^{K} M_{kj,t_i} = 0 \\
c_{kl,t_i} &= +\infty, \qquad \text{for} \quad k,l = 1,\dots,K, \quad \text{with} \quad \sum_{j=l}^{K} M_{kj,t_i} = 1.
\end{aligned} \tag{8.2}$$

For each period $]t_{j-1}, t_j]$ correlated normal random variables are sampled, $\left(r_{i,t_j}\right)_{i=1,\dots,n} \sim N(0,\Sigma)$, and credit i migrates from the initial state l to the final state k if

$$c_{lk-1,t_j} \leq r_{i,t_j} < c_{lk,t_j}.$$

[16]The migration matrix M_{t_n} defines a natural discretization of Y_t, but we can subdivide or refine the discretization arbitrarily through the introduction of a matrix square root $M_t^{1/2} = M_{t/2}$ or a generator matrix Q, $M_t = \exp(tQ)$.

[17]The discrete Markov process Y_t with time-homogeneous migration matrix does not necessarily meet a given PD-term structure i.e., $\left(M_0^k\right)_{iK} \neq F_i(t_k), k = 1,2,3,\dots,$ (with K as default state). This can easily be rectified by adapting the transition matrices recursively, i.e. the default column of the first matrix is set to the term structure and the remaining entries are renormalized. With some linear algebra the next matrix can be adjusted accordingly and so on.

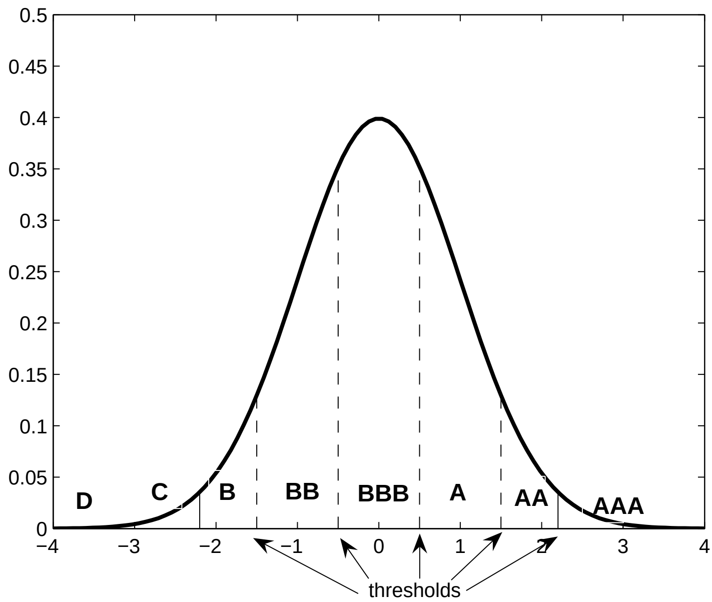


FIGURE 8.7: Migration thresholds.

There is no explicit interdependence between the steps apart from the autocorrelation generated by the migrations, hence it is essentially an extension of the CreditMetrics model. We also remark that this type of correlated credit migration model is also the basis of the credit component in some of the agency models.

But note one problem: The correlation structure of the model is not invariant under the refinement of the time discretization. Figure 8.8 shows the tail probability $\mathbb{P}[L > x]$ for a sample portfolio with non-vanishing correlation at the one year horizon under yearly, half-yearly, and quarterly discretization. For this, we have simpled calculated appropriate square-roots of the migration matrices. The fatness in the tail of the loss distribution is significantly reduced for smaller migration intervals. As soon as we introduce correlation to the rating transitions a link between global correlation and discretization is generated. By this we mean that choosing the same local correlation parameter ρ for each time step, the joint arrival probability in the states m, n of two

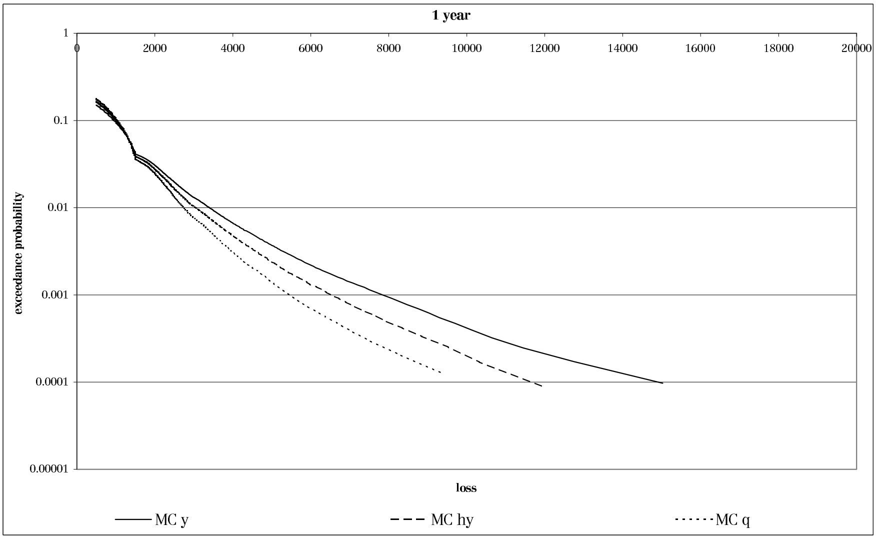


FIGURE 8.8: Refining time-discretization, migration model.

entities at time t, given they start at time 0 in states k, l

$$\mathbb{P}[Y_t^i = m, Y_t^j = n | Y_0^i = k, Y_0^j = l]$$

is a function of how fine we discretize the process, while keeping the local correlation constant. Smaller step-sizes de-correlate the processes Y_t^i and Y_t^j. This can easily be seen by the fact that for smaller step sizes the migration probabilities to the default state get smaller, but since the Gaussian copula has no tail dependence the correlation converges asymptotically to zero as we move the step size to zero. Obviously, the dispersion of the transition matrix also plays a role in this link.

In order to reconstitute the original correlation over a fixed time interval while halving the time step we have to adapt, i.e. increase the correlation. Suppose

$$\mathbb{P}[Y_1^i = K, Y_1^j = K | Y_0^i = k, Y_0^j = l]$$

is the joint default probability for one large step. Cutting the discretization in halves, the joint default probability is now

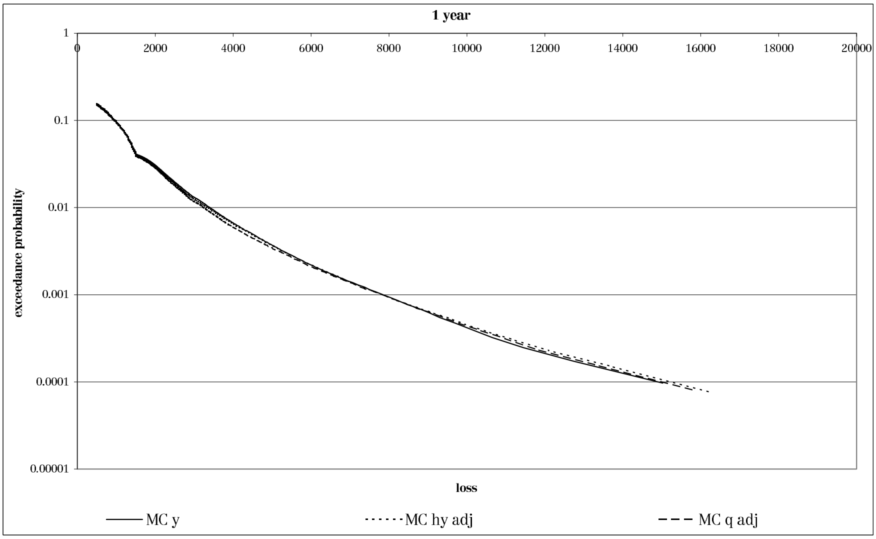


FIGURE 8.9: Refining time-discretization, migration model with adapted correlation.

$$
\begin{aligned}
&\mathbb{P}[Y_1^i = K, Y_1^j = K | Y_0^i = k, Y_0^j = l] = \\
&= \sum_{p,q} \mathbb{P}[Y_1^i = K, Y_1^j = K | Y_{1/2}^i = p, Y_{1/2}^j = q] \times \\
&\qquad \times \mathbb{P}[Y_{1/2}^i = p, Y_{1/2}^j = q | Y_0^i = k, Y_0^j = l]. \qquad (8.3)
\end{aligned}
$$

Instead of trying to adjust the correlation for all pairs i, j we confine ourselves to a homogeneous state in the sense of a large pool approximation. We obtain one adjustment factor and apply it to all names in the portfolio. For further discretization we simply nest the approach. Figure 8.9 shows the effect of the adjustment. We use an inhomogeneous portfolio of 100 positions with exposures distributed uniformly in $[500, 1500]$, 1-year default probabilities in $[10bp, 100bp]$, and correlation between $[10\%, 30\%]$. As can be seen from the graph both loss distribution are now commensurable. From a risk perspective this degree of similarity seems sufficient, particularely if risk measures like expected shortfall are used. Further improvement can be achieved by computing adjustment factors for each rating state and for each matrix in the sequence of transition matrices (if they are different).

8.4.2 Correlated Default Time Models

Another "best practice" approach is to generate *correlated default times* of the collateral securities. We already discussed this approach in Section 7.3. The correlated default times approach calibrates default times compatible to a given one-year horizon asset value model by means of credit curves, assigned to the default probability of the collateral securities, and some copula function, generating a multivariate dependency structure for the single default times. It is not by chance that this approach already has been used for the valuation of *default baskets*: Focusing only on defaults and not on rating migrations, the collateral pool (or reference portfolio) of a CDO can be interpreted as a somewhat large default basket. The only difference is the cash flow model on top of the basket.

From a simulation point of view, the default times approach involves much less random draws than a multi-step approach. For example, a multi-step model w.r.t. a collateral pool consisting of 100 bonds, would for quarterly payments over 10 years require $100 \times 10 \times 4$ simulated random draws in every scenario. The same situation by means of a default times approach would only require to simulate 100 random draws in a scenario, namely realizations of 100 default times for 100 bonds. This saves computation time, but has the disadvantage that *rating distributions* (e.g., for modeling rating triggers) can not be incorporated in a straightforward manner as it is in the case of multi-step models.

Time-consuming calculations in the default times approach could be expected in the part of the algorithm inverting the credit curves $F_i(t)$ in order to calculate default times according to the formula $\tau_i = F_i^{-1}(N[r_i])$; see Section 7.3. Fortunately, for CDO models the exact time when a default occurs is not relevant. Instead, the only relevant information is if an instrument defaults between two consecutive payment dates. Therefore, the copula function approach for default times can be easily discretized by calculating thresholds at each payment date $t_1 < t_2 < t_3 < \ldots < t_n$ according to

$$c_{i,t_k} \;=\; N^{-1}[F_i(t_k)] \;,$$

where F_i denotes the credit curve for some credit i, and $N[\cdot]$ denotes the cumulative standard normal distribution function. Clearly one has

$$c_{i,t_1} \;<\; c_{i,t_2} \;<\; \ldots \;<\; c_{i,t_n} \;.$$

Setting $c_{i,t_0} = -\infty$, asset i defaults in period $]t_{k-1}, t_k]$ if and only if

$$c_{i,t_{k-1}} \; < \; r_i \; \leq c_{i,t_k} \; ,$$

where $(r_1, ..., r_m) \sim N(0, \Sigma)$ denotes the random vector of standardized asset value log-returns with asset correlation matrix Σ. This reduces the computational efforts substantially, since the thresholds have to be calculated only once and can then be stored in a look-up table *before* the actual random events are simulated.

Semi-analytic methods like the saddle-point or the recursive technique (see the calculation techniques section at the end of Chapter 2) are also commonly used for solving the CDO pricing problem with correlated default times.

Note further that the correlated-default-times approach with Gaussian-copula is a rather static model. For this, we write the conditional joint default probability at different time horizons in a one-factor setting as

$$\begin{aligned}
&\mathbb{P}[\tau_1 < s, \tau_2 < t | Y = y] = \\
&= \mathbb{P}[r_1 < N^{-1}(F_1(s)), r_2 < N^{-1}(F_2(t)) | Y = y] \\
&= \mathbb{P}[Z_1 < \frac{N^{-1}(F_1(s)) - \sqrt{\varrho} y}{\sqrt{1-\varrho}}, Z_2 < \frac{N^{-1}(F_2(t)) - \sqrt{\varrho} y}{\sqrt{1-\varrho}}] \\
&= N\left[\frac{N^{-1}(F_1(s)) - \sqrt{\varrho} y}{\sqrt{1-\varrho}}\right] N\left[\frac{N^{-1}(F_2(t)) - \sqrt{\varrho} y}{\sqrt{1-\varrho}}\right]. \qquad (8.4)
\end{aligned}$$

The sample of the common factor Y is static for all time horizons; there is no dynamics through time, see also SCHÖNBUCHER [164].

Figure 8.10 illustrates the workflow of a CDO model which is (in its portfolio scenario engine) based on default times.

8.4.3 First-Passage-Time Models

8.4.3.1 Discrete Barrier Model

FINGER [62] and HULL and WHITE [95] proposed a discrete multi-period barrier model on a time grid $t_0 < t_1 < \ldots < t_n$ based on correlated Brownian processes B_t^i where the default thresholds $c_i(t_k)$ are decreasing functions of time calibrated to satisfy the marginal term structure $F_i(t_k)$. Credit entity i defaults in period k if for the first time $B_{t_k}^i < c_i(t_k)$, i.e.

$$\tau_i = \min\left\{t_k \geq 0 : B_{t_k}^i < c_i(t_k), k = 0, \ldots, n\right\}.$$

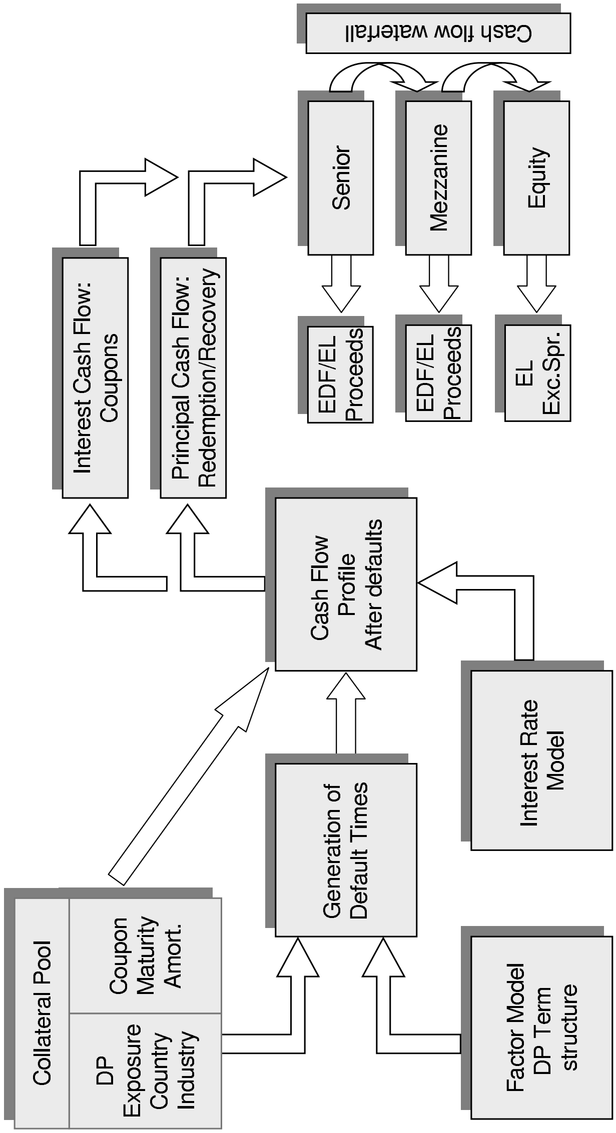


FIGURE 8.10: CDO modeling workflow based on default times.

The default barriers $c_i(t_k)$ are to be calibrated to match $F_i(t_k)$ such that

$$F_i(t_k) \stackrel{!}{=} \mathbb{P}[\tau_i < t_k].$$

Denote $\delta_k = t_k - t_{k-1}$, then from

$$\mathbb{P}[B^i_{t_1} < c_i(t_1)] = F_i(t_1)$$

follows that

$$c_i(t_1) = \sqrt{\delta_1} N^{-1}[F_i(t_1)].$$

The successive thresholds are then found by solving

$$\begin{aligned}
&F_i(t_k) - F_i(t_{k-1}) = \\
&= \mathbb{P}[B^i_{t_1} > c_i(t_1) \cap \cdots \cap B^i_{t_{k-1}} > c_i(t_{k-1}) \cap B^i_{t_k} < c_i(t_k)] \\
&= \int_{c_i(t_{k-1})}^{\infty} f_i(t_{k-1}, u) N\left[\frac{c_i(t_k) - u}{\sqrt{\delta_k}}\right] du,
\end{aligned}$$

where $f_i(t_k, x)$ is the density of $B^i_{t_k}$ given $B^i_{t_j} > c_i(t_j)$ for all $j < k$:

$$\begin{aligned}
f_i(t_1, x) &= \frac{1}{\sqrt{2\pi\delta_1}} \exp\left(-\frac{x^2}{2\delta_1}\right) \\
f_i(t_k, x) &= \int_{c_i(t_{k-1})}^{\infty} f_i(t_{k-1}, u) \frac{1}{\sqrt{2\pi\delta_k}} \exp\left(-\frac{(x-u)^2}{2\delta_k}\right) du.
\end{aligned}$$

Hence, the calibration of the default thresholds is an iterative process and requires the numerical evaluation of integrals with increasing dimension, which renders the model computationally very heavy. Another shortcoming of the model is that it is not invariant under the refinement of the time discretization [178]. Figure 8.11 shows the tail probability $\mathbb{P}[L > x]$ of a portfolio loss with different discretizations (yearly, half-yearly, quarterly) of the model. Obviously, the volatility and tail fatness of the loss distribution decreases with increasing refinement, and it is not clear where the limiting distribution is.

8.4.3.2 Continuous Time-Changed Barrier Model

The above mentioned discrete barrier model is drawn from a continuous version, i.e. correlated Brownian processes B^i_t with time-dependent barriers $c_i(t)$. The default time of credit i is then the first hitting time of the barrier $c_i(t)$ by the driving process B^i_t:

$$\tau_i = \inf\left\{t \geq 0 : B^i_t < c_i(t)\right\}.$$

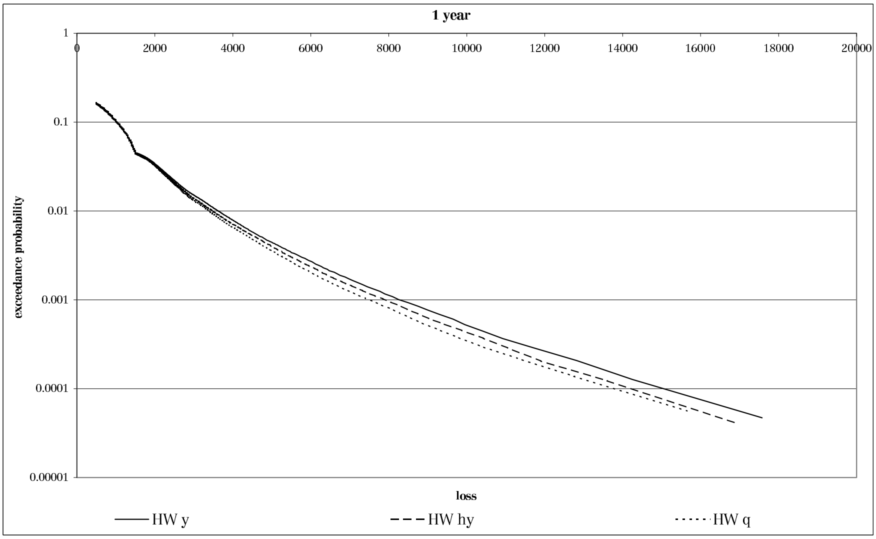


FIGURE 8.11: Refining time-discretization, discrete barrier model.

If $c_i(t)$ is absolutely continuous, we can write

$$c_i(t) = c_i(0) + \int_0^t \mu_s^i ds,$$

and the default time τ_i is the first hitting time of the constant barrier $c_i(0)$ by a Wiener process with drift.

$$\begin{aligned} Y_t^i &= B_t^i - \int_0^t \mu_s^i ds \\ \tau_i &= \inf \left\{ t \geq 0 : Y_t^i < c_i(0) \right\}. \end{aligned} \tag{8.5}$$

The problem now is to calibrate the model to the prescribed default term structure, $\mathbb{P}[\tau_i < t] = F_i(t)$. To this end Overbeck and Schmidt [154] put forward a barrier model based on Brownian processes B_t^i with suitably transformed time scales, (T_t^i), strictly increasing, $T_0^i = 0$. The first passage time to default τ_i of credit entity i is defined through the process

$$Y_t^i = B_{T_t^i}^i$$

and

$$\tau_i = \inf \left\{ s \geq 0 : Y_s^i < c_i \right\},$$

with a time independent barrier c_i. From the strong Markov property or the reflection principle of the Brownian motion follows (see, e.g., KARATZAS and SHREVE [112, 113]) that the first passage time of an untransformed Brownian motion with respect to a constant barrier c

$$\tilde{\tau} = \inf\{t \geq 0 : B_t < c\}$$

is distributed as

$$\mathbb{P}[\tilde{\tau} < t] = \mathbb{P}\left[\min_{0<s<t} B_s < c\right] = 2N\left[\frac{c}{\sqrt{t}}\right]. \tag{8.6}$$

As T_t^i is strictly increasing we find that

$$\begin{aligned}\mathbb{P}[\tau_i < t] &= \mathbb{P}\left[\min_{0<s<t} B^i_{T^i_t} < c_i\right] = \mathbb{P}\left[\min_{0<s<T^i_t} B^i_t < c_i\right] \\ &= 2N\left[\frac{c_i}{\sqrt{T^i_t}}\right] \end{aligned} \tag{8.7}$$

Hence, given a default term structure $F_i(t)$ the model is calibrated to the marginals via the time transformation

$$T^i_t = \left[\frac{c_i}{\mathrm{N}^{(-1)}\left(F_i(t)/2\right)}\right]^2. \tag{8.8}$$

Since $F(t)$ is strictly increasing this also follows for T_t. The constant default barrier c_i is then obtained by fixing a time t_0 with $T^i_{t_0} = t_0$ which implies

$$c_i = \mathrm{N}^{(-1)}\left(F_i(t_0)/2\right)\sqrt{t_0}. \tag{8.9}$$

An obvious, but not necessarily the only sensible choice is to take t_0 as the final maturity. Figure 8.12 depicts the time transformation for various credit qualities with $t_0 = 1$. Dependency between credits is introduced here through the (local) instantaneous correlation matrix Σ of the Brownian processes B^i_t. The joint default probabilities $\mathbb{P}[\tau_i < t, \tau_j < t]$ can be written in analytical, but rather technical form (we refer to [154]), which allows the calibration of the model to prescribed joint default probabilities. The derivation of the JDPs is based on an analytical representation via modified Bessel functions as described by ZHOU [193].

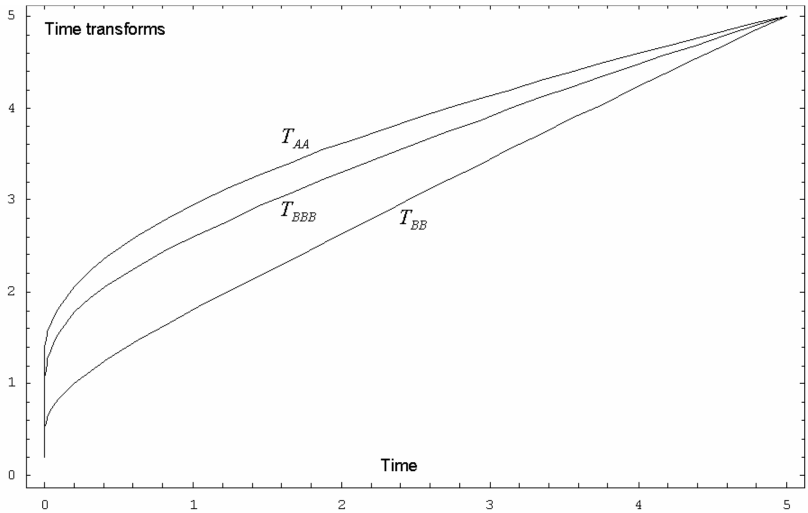


FIGURE 8.12: Time transformation according to Equation (8.8) for three different credit qualities with term structure $F_i(t) = 1 - e^{-\lambda t}$, $\lambda = 10\text{bp}$, 100bp and 500bp.

8.4.4 Stochastic Default Intensity Models

The stochastic intensity approach [47, 49] is a time continuous model and has already been presented in Section 2.4.4. DUFFIE and GÂRLEANU [47] studied a stochastic intensity approach to the valuation of CDOs by considering a basic affine process for the intensity λ, solving a stochastic differential equation of the form

$$d\lambda(t) \;=\; \kappa(\theta - \lambda(t))dt + \sigma\sqrt{\lambda(t)}dB(t) + \Delta J(t), \tag{8.10}$$

where B is a Wiener process and J is a pure-jump process, independent of B. In the course of their paper, they consider a simple subordinated structure, consisting of only three tranches: An equity piece, a mezzanine tranche, and a senior tranche. They experimented with different overcollateralization levels and different correlations and showed that *correlations significantly impact the market value of individual tranches.* For example, in cases where the senior tranche has only a small cushion of subordinated capital, the market value of the senior tranche decreases with decreasing correlation, whereas the market value of the equity piece increases with increasing correlation. Their calculations further show that this effect can be mitigated, but not removed, by assuming a higher level of overcollateralization. Regarding the behavior of the mezzanine tranche in dependence on a changing correlation, they find that the net effect of the impact of correlation changes on the market value of the senior and equity tranche is *absorbed by the mezzanine tranche*. This interestingly results in an ambiguous behavior of the mezzanine tranche: Increasing default correlation may raise or lower the mezzanine spreads.

For a practical implementation, the stochastic differential equation (2.47) has to be solved *numerically* by discretization methods, i.e., the intensity is integrated in appropriately small time steps. Unfortunately, this procedure can be quite *time-consuming* compared to other CDO modeling approaches.

8.4.5 Intertemporal Dependence and Autocorrelation

The above models are all calibrated to the same marginal default probabilities in order to ensure that

$$\mathbb{P}[\tau_i < t] = F_i(t).$$

Furthermore, for a meaningful comparison of single-period credit portfolio models another calibration requirement is to equate the joint default probabilities at a given horizon, i.e.

$$\mathbb{P}[\tau_i < T, \tau_j < T].$$

But in the multi-period context, from the last condition does not necessarily follow that the joint default probabilities at a horizon s different from the calibration horizon remain equal between different models,

$$\mathbb{P}[\tau_i < s, \tau_j < s], \quad s \neq T.$$

And even more, it is clear that the properties of joint default probabilities at two different maturities

$$\mathbb{P}[\tau_i < s, \tau_j < t],$$

that reflect the intertemporal dependencies, are in general different for different multi-period models. This is already a problem when we deal with products like plain vanilla CDOs under the subjective historical measure because the same CDO at a maturity that is not the calibration horizon leads to different results under different models. And as soon as we would like to model a whole book of CDOs with different vintage years and a decent overlap in single name risks, or equally, a time-dependent product like a Forward-Start-CDO, we should be aware that the intertemporal properties of our model play indeed an important role. We also refer to ANDERSEN [6] where useful insights to portfolio losses in factor models and their term structure and intertemporal loss dependence can be gained. Figures 8.13 and 8.14 serve to demonstrate these differences in the intertemporal behavior. We have taken a sample portfolio, as in Section 8.4.1, and depict the joint 2-year/4-year loss distribution in form of a two-dimensional heat map on a logarithmic scale, i.e. dark colors represent high probabilities. All four models are calibrated to the same marginal single credit name term structure through time and to the same joint default probabilities at the final 4-year horizon, i.e. the final loss distributions are congruent. There is no probability mass in the upper left cone of the graph as the portfolio loss is a monotonous function of time. High intertemporal dependence is reflected through large joint probabilities in the middle between the x-axis and the 45°-line. Comparing the graphs, it is no surprise that the correlated-default-time model (CDT) shows the highest degree of intertemporal dependence (Fig. 8.13, bottom) as it is essentially a static model; confer also Section 8.4.2. On the other end of the scale is the Markov-Chain-Migration model (MC) (Fig. 8.13, top), where we find high probabilities on the edges of the lower triangle. By construction there is no interperiod correlation of default, hence most of the loss volatility originates from cross correlation. The discrete barrier model (HW) and time-changed barrier model (OS) show similar behavior as their intertemporal dependence is governed by the autocorrelation of the driving Brownian motion. But there is still a visible difference, the joint probability distribution of the time-changed model (Fig. 8.14, bottom) has more a drop-like shape than the one from the discrete barrier model (Fig. 8.14, top). This means that, although all models produce the same final portfolio loss, the various models accumulate this loss differently over time.

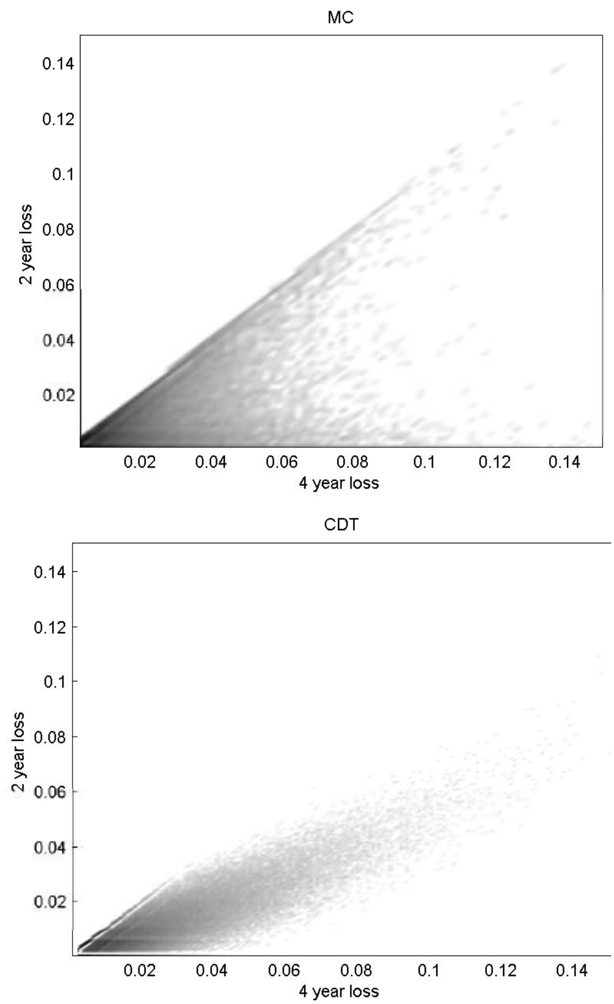


FIGURE 8.13: Joint loss distributions at the 2-year and 4-year horizon for different multiperiod models: (top) migrations model (Section 8.4.1), (bottom) correlated-default-time model (Section 8.4.2).

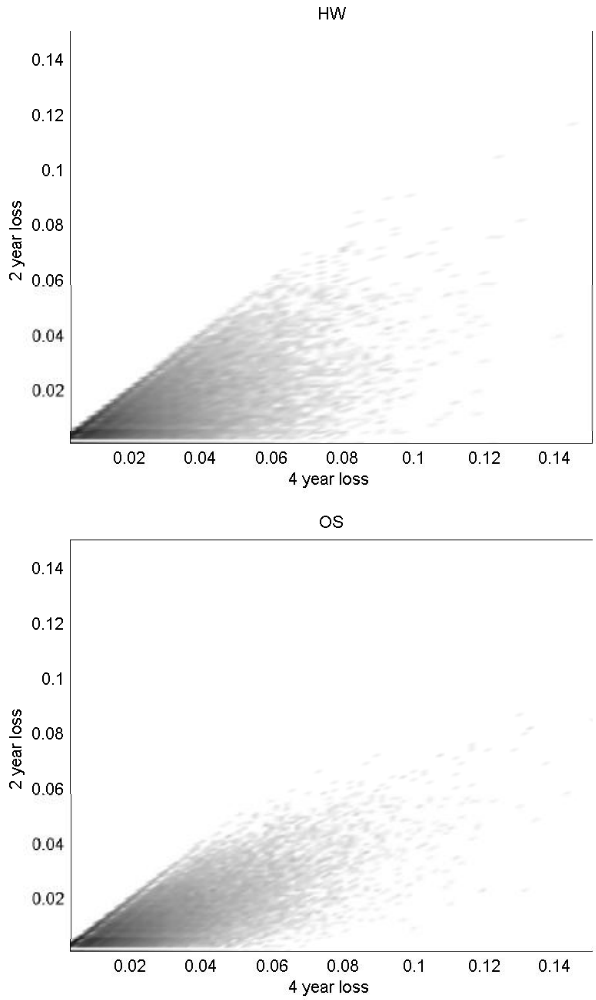


FIGURE 8.14: Joint loss distributions at the 2-year and 4-year horizon for different multiperiod models: (top) discrete barrier model (Section 8.4.3.1), (bottom) time-changed models (Section 8.4.3.2).

8.5 Former Rating Agency Model: Moody's BET

This section is a relict from the first edition of this book and we seriously considered removing it because Moody's no longer relies on their "BET method" in the rating process of CDOs. However, we decided to keep the section for two reasons. First, it shows a typical rating agency approach from some years ago. Second, there are still many reports in the market in which the so-called *diversity score* and Moody's BET are mentioned. For participants in the CDO market it is still good to know how this method works. So for the rest of this section we forget that the BET is outdated and follow the presentation of the first edition of this book.

Moody's rating analysis of CDOs some years ago was based on the following idea:

Instead of calculating the loss distribution of the original collateral portfolio of a CDO, Moody's constructs a *homogeneous comparison portfolio* satisfying the following conditions:

- All instruments have equal face values, summing up to the collateral pool's total par value.
- All instruments have equal default probability p, calibrated according to the *weighted average rating factor* (WARF), assigned to the portfolio by means of Moody's rating analysis.
- The instruments in the comparison portfolio are independent.

Moody's calibrates such a homogenous portfolio to any given pool of loans or bond, taking the rating distribution, exposure distribution, industry distribution, and the maturities of the assets into account. Then, according to the assumptions made, the portfolio loss of the homogeneous comparison portfolio follows a *binomial distribution*; see also Chapter 2.

The crucial parameter in this setting is the number n of instruments in the comparison portfolio. This parameter constitutes a key measure of diversification in the collateral pool developed by Moody's and is therefore called Moody's *diversity score* (DS) of the collateral portfolio. Regarding diversification, Moody's makes two additional assumptions:

TABLE 8.6: Moody's Diversity Score; see [140].

Number of Firms in Same Industry	Diversity Score
1	1.00
2	1.50
3	2.00
4	2.33
5	2.67
6	3.00
7	3.25
8	3.50
9	3.75
10	4.00
Source: ***Moody's Investors Services***	

- Every instrument in the comparison portfolio can be uniquely assigned to one industry group.
- Two intruments in the comparison portfolio have positive correlation if and only if they belong to the same industry group.

Based on this assumption, the only driver of diversification is the industry distribution of the collateral pool. Table 8.6 reports the diversity score for different industry groupings[18]. The diversity score of a portfolio is then calculated by summing up the diversity scores for the single industries represented in the collateral pool. For illustration purposes, let us calculate two sample constellations.

1. Consider 10 bonds from 10 different firms, distributed over 3 industries:
   2 firms in industry no. 1, yielding a diversity score of $\mathrm{DS}_1 = 1.50$
   3 firms in industry no. 2, yielding a diversity score of $\mathrm{DS}_2 = 2.00$
   5 firms in industry no. 3, yielding a diversity score of $\mathrm{DS}_3 = 2.67$
   The portfolio's total diversity score equals
   $\mathrm{DS} = \mathrm{DS}_1 + \mathrm{DS}_2 + \mathrm{DS}_3 = 6.17$.

2. Consider 10 bonds from 10 different firms, distributed over 10 industries:
   10 times one firm in one single industry means

[18] For more than 10 instruments in one industry group, the diversity score is determined by means of a case-by-case evaluation.

10 times a diversity score of 1, such that the portfolio's total diversity score sums up to DS = 10.

The industry distribution of Constellation 2 leads to an obviously better industry diversification, and therefore yields a higher diversity score. Altogether, Moody's distinguishes between 33 industry groups, yielding[19] a best possible diversity score of $132 = 33 \times 4$.

The loss distribution of the homogeneous comparison portfolio is assumed to be binomially distributed with parameters DS and WARF, $L \sim B(\mathrm{DS}, \mathrm{WARF})$, such that the probability of k defaults in the comparison portfolio equals

$$\mathbb{P}[L = k] \;=\; \frac{(\mathrm{DS})!}{k!(\mathrm{DS}-k)!}\,(\mathrm{WARF})^k\,(1-\mathrm{WARF})^{\mathrm{DS}-k}\;,$$

where k ranges from 0 to DS. Based on the so-obtained loss distribution, *cash flow scenarios* are evaluated in order to determine the *rating of a tranche*. Dependent on how many losses in the collateral pool a tranche can bear without suffering a loss due to the credit enhancement mechanisms of the structure, the tranche gets assigned a rating reflecting its "default remoteness." For example, senior notes have to pass much stronger stress scenarios without suffering a loss than junior or mezzanine notes.

From time to time CDO tranches are *down- or upgraded* by the rating agencies, because their default remoteness decreased or increased. For example, last and this year we saw many downgrades of CDO tranches, sometimes downgraded by more than one notch on the respective rating scale, due to the heavy recession in the global economy.

In a next step, we now want to consider the BET from a more mathematical point of view. For this purpose we consider a sample portfolio of m bonds, all bonds having the same default probability p and equal face values. Additionally we assume that the pairwise default correlation[20] of the bonds is uniform for the whole portfolio and given by r. Our modeling framework is a uniform Bernoulli mixture model, with asset values as latent variables, as introduced in Section 2.5.1. According to Equation 2.10 and Proposition 2.5.1, the corresponding uniform

[19]Ignoring deviations from Table 8.6 due to special case-by-case evaluations.

[20]In contrast to the rest of this book we here denote the default correlation by r.

asset correlation ϱ of the model can be calculated by solving

$$r \;=\; \frac{N_2\big[N^{-1}[p], N^{-1}[p]; \varrho\big] - p^2}{p(1-p)}$$

for ϱ. For example, for $r = 3\%$ and $p = 1\%$ we calculate $\varrho = 23.06\%$. Recall that the uniform Bernoulli mixture model is completely determined by specifying p and r (respectively ϱ).

In Proposition 2.5.7 we already discussed the two extreme cases regarding ϱ. In case of $\varrho = 0$, the distribution of the portfolio loss is binomial, $L \sim B(m, mp)$. In case of $\varrho = 1$, the loss distribution is of Bernoulli type, $L \sim B(1, p)$. Both extreme case distributions are *binomial* distributions with probability p. Looking at the respective first parameter of both distributions, we discover m bonds in the first case and 1 bond in the second case. The idea of the BET now is to introduce also the intermediate cases by establishing a relation between the assumed level of correlation and the number of bonds in a homogeneous comparison portfolio. More formally, for a given portfolio of m bonds, the BET establishes a functional relation

$$n : [0,1] \to \{0, 1, ..., m\}, \;\; r \mapsto n(r),$$

between the default correlation and the number of bonds in a homogeneous portfolio of *independent* bonds with binomial loss distribution.

The function n can be determined by a matching of first and second moments. The first moments of both portfolios must be equal to p. The second moment of the original portfolio can be calculated as

$$\mathbb{V}\big[L^{(m)}\big] \;=\; \frac{1}{m^2}\,\mathbb{V}\Big[\sum_{i=1}^{m} L_i\Big] \;=\; \frac{1}{m^2}\sum_{i,j=1}^{m}\mathrm{Cov}[L_i, L_j] \;=$$

$$=\; \frac{1}{m^2}\Big(mp(1-p) + \sum_{i\neq j} rp(1-p)\Big) \;=$$

$$=\; \frac{mp(1-p) + m(m-1)rp(1-p)}{m^2} \;=\; \frac{p(1-p)(1+(m-1)r)}{m}\,.$$

The variance of the homogeneous comparison portfolio, consisting of $n(r)$ independent bonds, equals

$$\mathbb{V}\big[L^{(n(r))}\big] \;=\; \frac{p(1-p)}{n(r)}\,.$$

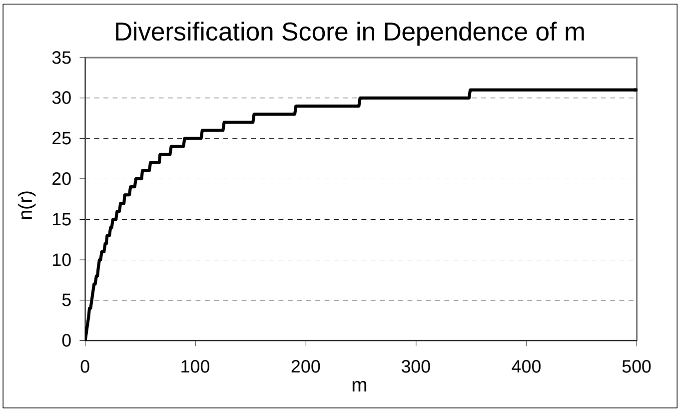


FIGURE 8.15: Diversification score as a function of m for $r = 3\%$.

Matching both second moments finally yields

$$n(r) \;=\; \frac{m}{1 + r(m-1)}\;. \tag{8.11}$$

This number $n(r)$ is not necessarily an integer value, so that we have to round it to the closest integer. The so-obtained number is comparable to Moody's diversity score. In order to distinguish[21] the two scores, we call $n(r)$ the *diversification score* of the original portfolio. Figure 8.15 shows the diversification score $n(r)$ in dependence of m for $r = 3\%$. Two facts illustrated by the plot follow from Equation 8.11:

1. The diversification score is independent of the credit quality of the pool, captured by the default probability p.

2. The diversification score is bounded from above by $1/r$. For $r = 3\%$, the maximum diversification score is DS $= 33$, which is attained for $m \geq 1,261$. The reason for an upper bound of the diversification score w.r.t. a fixed default correlation comes from the fact that only *specific risk* can be eliminated by diversification. *Systematic risk* remains in the portfolio, no matter by how many obligors we enlarge the portfolio.

We now want to compare the loss distributions of a fictitious sample portfolio and the homogeneous portfolio of independent bonds fitted to

[21] Note that Moody's diversity score purely relies on industry diversification.

the original portfolio by means of the BET. We assume that the original portfolio contains $m = 100$ bonds with uniform default probability $p = 1\%$ and uniform default correlation $r = 3\%$. As already mentioned, these assumptions yield an asset correlation of $\varrho = 23.06\%$.

According to Equations 2.8 and 2.49, the probability for k defaults in the original bond portfolio is given by

$$\mathbb{P}[L^{(100)} = k] \;=\; \binom{100}{k} \int_0^1 p(y)^k (1-p(y))^{100-k} dN(y) \;,$$

$$\text{where} \qquad p(y) = N\Big[\frac{N^{-1}(0.01) - \sqrt{0.2306}\,y}{\sqrt{1-0.2306}}\Big] \;.$$

Therefore, we can easily calculate the loss distribution of the original portfolio. Next, we calculate the diversification score of the original portfolio. According to Equation 8.11, we obtain

$$n(3\%) \;=\; \frac{100}{1 + 3\%(100-1)} \;=\; \frac{100}{1 + 2.97} \;=\; 25.19 \;,$$

such that the diversification score after rounding equals 25. Therefore, the loss distribution of the homogeneous comparison portfolio follows a binomial distribution, $L^{(25)} \sim B(25, 2.5)$. So here the BET claims that *25 independent bonds carry the same risk as 100 bonds with default correlation* $r = 3\%$.

Figure 8.16 compares the original loss distribution with the BET-fitted binomial distribution. The plot clearly shows that the BET-fit significantly underestimates the *tail probabilities* of the original loss distribution.

This does not come as much of a surprise, because due to the *central limit theorem* binomial distributions tend to be approximately normal for a large number of bonds, whereas typical credit portfolio loss distributions are skewed with fat tails. Moreover, it is generally true that moment matching procedures do not automatically also fit the tails of the considered distributions in an accurate manner.

Now we come to an important conclusion: Because the BET significantly underestimates the tail probabilities of the original portfolio, the *risk of senior notes will typically be underestimated by the BET approach.*

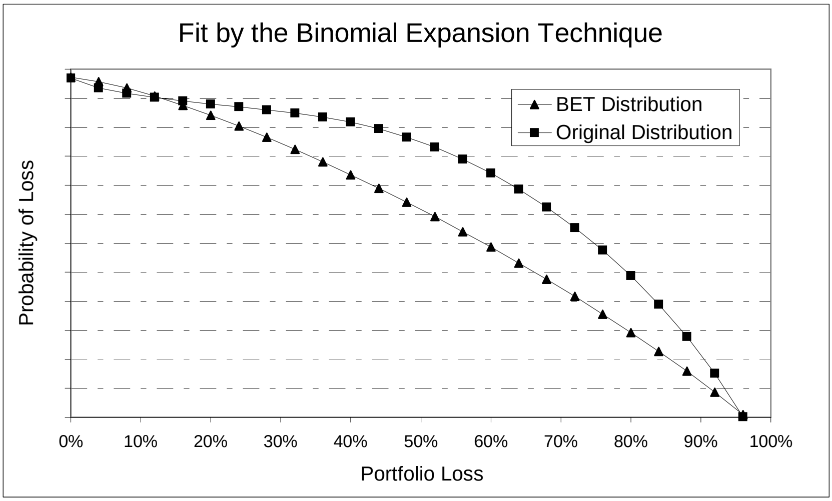


FIGURE 8.16: Fitting a loss distribution by means of the BET (original uniform portfolio: $p = 1\%$, $r = 3\%$, $m = 100$); note that the y-axis is logarithmically scaled.

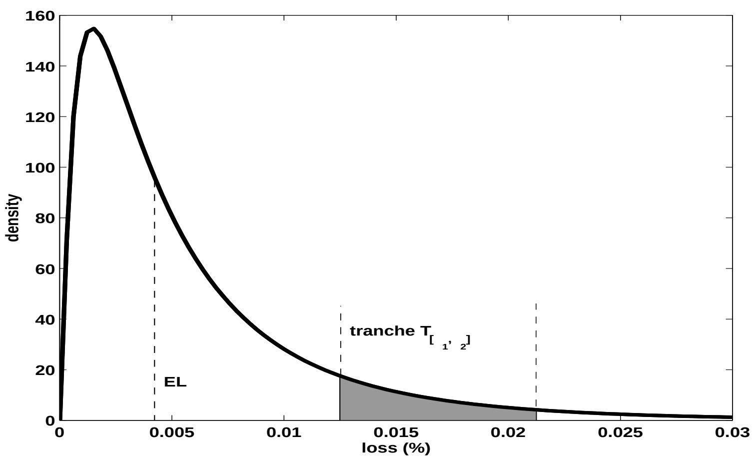


FIGURE 8.17: Tranching a loss distribution.

To make this more explicit, we consider a situation as illustrated in Figure 8.17. Assume that the plot shows the probability density of the distribution of the cumulative net losses L of some collateral pool, calculated over the whole term of the structure. Let us further assume that the bank invested in an upper mezzanine tranche $T_{[\alpha_1,\alpha_2]}$ with lower bound α_1 and upper bound α_2. Then, the default probability (PD) of this tranche and its expected loss (EL) can be calculated[22] as

$$\mathrm{PD}(T_{[\alpha_1,\alpha_2]}) \;=\; \mathbb{P}[L > \alpha_1]\;,$$

$$\mathrm{EL}(T_{[\alpha_1,\alpha_2]}) \;=\; \frac{1}{\alpha_2 - \alpha_1}\int \min\Big(\max[x-\alpha_1, 0], \alpha_2 - \alpha_1\Big) d\,\mathbb{P}_L(x)\;,$$

where $\mathbb{P}_L$ denotes the probability density of L. If we now would replace the loss distribution $\mathbb{P}_L$ by a binomial distribution fitted to $\mathbb{P}_L$ by means of the BET, we can expect that $\mathrm{PD}(T_{[\alpha_1,\alpha_2]})$ and $\mathrm{EL}(T_{[\alpha_1,\alpha_2]})$ will be significantly lower; see Figure 8.16. A moment-matched binomial distribution will not appropriately capture the risk of a tranche more outside in the tail, as it is the case for $T_{[\alpha_1,\alpha_2]}$.

Our discussion has far-reaching consequences. Whenever a bank intends to invest in a senior note, the model the bank uses for the evaluation of the investment should capture the *tail risk* of the collateral pool. But the tail risk of the collateral pool is driven by the *correlation* inherent in the collateral portfolio. The higher the overall correlation, the larger the tail probabilities and therefore the potential for losses in senior pieces of the structure. Because the bank wants to be compensated for taking this risk, it can not rely on the BET or other methods ignoring the *skewed fat-tailed* character of credit portfolio loss distributions. Only a full Monte Carlo simulation of an appropriate portfolio model, combined with a sound modeling of all relevant cash flow elements of the structure, will really show how much premium payment the bank needs to be compensated for the taken risk and to make some profit at the end.

The criticism of the BET just made is a major reason why rating agency models are more sophisticated today than they used to be in

[22] Of course, certain cash flow elements in a structure can distort the direct effect of losses on a particular tranche, as we claimed it here, but for reasons of simplicity we ignore this greater complexity for the moment. However, in synthetic CLOs, where the performance of notes is linked to the performance of a reference pool (e.g. by means of credit-linked notes) this simplified view is very close to the truth.

the early days of CDOs. But still, rating agencies make quite some simplifying assumptions in their approaches and the most recent crisis gave rise for a lot of scepticism w.r.t. the way rating agencies evaluate structured credit instruments. We neither positively nor negatively comment on that but leave the judgment to the reader.

8.6 Developments, Model Issues and Further Reading

Between the publication date of the first edition of this book and today the CDO market rapidly evolved. In parallel the literature on credit risk modeling as well as the focus of credit risk modeling turned in large part towards questions of CDO modeling. This is in contrast to the first edition of this book where we wrote that there is not much (academic) literature available on CDOs.

A comprehensive introductory book to CDOs as capital market instruments is the book by Choudry [32]. It has not a particular focus on modeling but describes common structures and mechanisms from different angles so that readers not familiar with structured credit get access to a nice presentation of recent developments. Another book in this area is the book [58] by Fabozzi. There is much more literature on the topic available and readers will find them quickly via booksellers and webshops. On the modeling side we refer to the afore-mentioned book [24] where readers find the necessary mathematical background for modeling recent CDO transactions as well as many examples of real-life transactions and deals, including an elaborated approach on how one could model them. Another source for CDO modeling topics are books with a primary focus on active credit portfolio management which can be seen as a major driver of the CDO market evolvement during the last years. For instance, the book by Felsenheimer et al. [60], the very recent book by Gregoriou and Hoppe [85] and the guide to active credit portfolio management with a focus on illiquid credit by Benvegnu et al. contain rich material in the area of CDO modeling and related fields like hedging and securitizations, quantitative evaluation of investment opportunities, and so on. During the last eight years many research papers on CDOs have been written. In the afore-mentioned books readers will find guidance regarding papers on different topics. We mention four out of many papers here, namely the one by Baxter [14], a paper on the comparison of CDO pricings models [30] by Burtschell et al., and two papers on modeling techniques [22, 21] which turned out to be useful for modeling baskets and CDOs.

For the sequel we recall some notation. Let us assume that a portfolio of m credit-risky assets is given. The assets can be everything bearing credit risk, for instance, loans, bonds, asset-backed securities (like, e.g., RMBS, other CDOs, etc.), credit derivatives, and so on. The portfolio can be tranched in various ways. The lower boundary of a tranche is called its *attachment point* and the upper boundary of a tranche is called its *detachment point*. The loss profile of a tranche, denoted by $T_{\alpha,\beta}$ with attachment point α and detachment point β is described by

$$L^{(t)}_{T_{\alpha,\beta}} \;=\; \min\bigl[\max[0, L^{(t)} - \alpha], \beta - \alpha\bigr] \tag{8.12}$$

where $L^{(t)}$ denotes the portfolio loss up to time t. Over the last years there evolved a whole market just dealing with the loss profile of single tranches which are accordingly called *single-tranche CDOs* (STCDOs). For an introduction to STCDOs and their underlyings we refer to [24], Section 3.4, and to the research papers by Amato and Gyntelberg [4], Andersen et al. [7] and Felsenheimer et al. [59].

Underlyings of STCDOs are just credit default swaps (CDS) so that such deals are completely *unfunded*; they are pure derivative transactions. Participants of the STCDO market invented a notion of *implied correlation* which could be used by *correlation desks* to trade systematic risk in tailor-made ways; see [24], Figures 3.37 and 3.38 and the discussion there. The concept of implied correlations is a nice example for the phenomenon that a market trend motivated a whole series of research papers with various approaches to model the relation between STCDO market prices and risk parameters like correlation in correspondence to a given curve of spread-implied default probabilities.

Besides STCDOs referenced to a standardized pool (say, index) of CDS one also finds a lot of so-called *bespoke* transactions where an issuer structures a portfolio and a tranche in a tailor-made way to satisfy the demand of a particular investor. These are bilateral transactions, often negotiated in a non-public way in the form of a private placement. One can say that the last years (before the crisis) revealed an almost unlimited range of possibilities and ways to structure credit. However, this all came to an abrupt end with the financial crisis which started in May 2007. We come to that in a moment.

Besides all kinds of unfunded synthetic transactions one also finds a lot of *partially funded* transactions, which is the classical case for securitizations with a focus on hedging and/or balance sheet management. Partially funded here means that the portfolio is tranched into some *funded* lower tranches (including equity) and an *unfunded* super senior tranche. Typically, the funded part has an attachment point at 0 and a detachment point at some $x\%$ of the portfolio where x often is between 5 and 15, depending on the considered portfolio, the purpose of the deal from originator view, and the current market environment (say, investor demand and risk appetite, pricing, liquidity, and so on).

Another trend which accelerated over the last years is *leverage*. The word "leverage" can mean many things and is used in many ways. What we mean by it can best be described by the sloppy slogan "invest less for more." Consider the following illustrative (!) example which can help to explain what that means. Assume that a portfolio pays 150bps spread per annum. Now we tranche the portfolio into two tranches, a junior tranche and a senior tranche:

$$T_{jun} \;=\; T_{0\%,5\%} \quad \text{and} \quad T_{sen} \;=\; T_{5\%,100\%}\,.$$

Because T_{jun} bears the *first loss risk* and is wiped out soon when losses occur (because its thickness is only 5%) we have to pay to investors in T_{jun} a much higher spread than the average 150bps spread which is earned on the whole portfolio. Without going too much into details and illustrative calculations let us assume that from the

$$150\text{bps} \times 100\% \text{ [portfolio face value]}$$

we pay 120bps to T_{jun}-investors. So we have a remaining

$$30\text{bps} \times 100\% \text{ [portfolio face value]}$$

left over for T_{sen}-investors. So investors in the senior tranche get much less than the average spread on the portfolio and the underlying rationale is that T_{jun} provides a 5%-*capital cushion* which protects T_{sen} against the first 5% of occurring losses. Our example is meant only in an illustrative sense but let us see what we get in terms of return from tranching the portfolio. For T_{jun} we get a return of

$$r_{jun} \;=\; \frac{\text{spread income}}{\text{invested capital}} \;=\; \frac{1.2}{5} \;=\; 24\%.$$

For the senior tranche T_{sen} we get

$$r_{sen} = \frac{\text{spread income}}{\text{invested capital}} = \frac{0.3}{95} \approx 32\text{bps.}$$

Assuming that the tranche investments are unfunded, this is a realistic situation where equity (T_{jun}) earns an attractive return at the cost of bearing the first loss risk and the senior tranche investor, which is capital buffered and at least for the first 5% remote from losses, gets a correspondingly low spread for a (hopefully[23]) low risk investment. Coming back to *leverage*, our little example illustrates how a portfolio can be leveraged in principal into a high risk investment with high return and a low risk investment with a correspondingly low return.

The tranching example provides some flavor of what *financial engineering* can do. It is no problem to structure portfolios in tailor-made ways to create new credit-risky assets with given *risk/return profile*. However, investment banks and portfolio arrangers did not stop after leveraging a portfolio *once*. They often leveraged again and again by collecting CDO tranches and ABS deals into new portfolios which were then tranched and structured again.

Over time such *multi-leverage* led to weird situations. For instance, one can find so-called *structured finance CDOs* (SFCDOs) with an *embedded circle reference* in the market. We do not disclose actual names of such deals here for obvious reasons but explain briefly what "embedded circle reference" means. Let us give the two SFCDOs names, say, one is SFCDO_A and the other is SFCDO_B. Both deals are *four-times leveraged*. This means that the reference pool of the two SFCDOs is a portfolio of SFCDOs which are referenced to SFCDOs which are referenced to multi-sector portfolios consisting of CDO tranches and asset-backed securities. To give it a name we can call this a *CDO of CDOs of CDOs of CDO&ABS*. The very strange situation in this particular example of deals in the market is that if one *looks through* to the lowest level of SFCDO_A then one finds a tranche from the SFCDO_B as one of the underlying assets and if one analyzes SFCDO_B then one finds a tranche of the SFCDO_A as the underlying asset of one of the underlying reference portfolios at the lowest level. This is what we mean by a *circle reference* embedded in two SCFDOs.

[23] The recent crisis showed that many assumed low risk tranches at senior level were not as much remote from losses as people originally thought.

This example demonstrates how an originally good idea, namely, leverage and financial engineering, finally was overplayed by market participants. In the crisis between May 2007 and December 2008 it turned out that SFCDOs and many other structured credit assets heavily got under pressure. Distress occurred in two ways. The first category were real (physical) defaults leading to a realized loss. For instance, SFCDOs and ABS referenced to RMBS with underlying US subprime mortgages belong into this category. The second category were mark-to-market losses based on pessimistic market opinion leading to high spreads. Mark-to-market volatility bounces back as soon as markets recover so for the second category one can expect future gains.

For a more comprehensive "lessons learned" analysis we refer to [26] but from what has been said so far one can at least draw the conclusion that multi-leverage can be dangerous and should be avoided. It is possible to model two times leveraged CDOs ("CDO-squared") via a look-through approach such that the transaction is appropriately analyzed, but it starts to be tremendously complex if not impossible to apply the same analysis and model standards to CDOs with a leverage of three and higher. The arising tree when one climbs down the structure from the very top CDO to the knots of the tree, say, the assets underlying at lowest level, gets really complex and data gathering becomes more than a challenge.

There are also interesting conclusions from a modeling perspective. First of all, we would like to repeat our remark from the beginning of this chapter, that large parts of the structured credit market existed more or less free from stochastic modeling. However, what we are looking at now is the part of the CDO market where models as explained in this book played a serious role. To mention a few examples, rating agencies and many sophisticated banks base their valuation of CDOs on quantitative models as well as on qualitative criteria. It is beyond the scope of this section to go into details regarding CDO model risk but we want at least to give a few hints about where the challenges are.

One can divide the CDO modeling world in two parts. One part which one can almost consider as routine because the estimation of parameters is standard and well understood, and another part where model choice and parameterization problems are a real challenge. Figure 8.6 depicts components of a stochastic CDO model as we would recommend it. The routine part includes the following:

- single-name risk quantities, e.g., PDs (physical or spread-implied);
- structural elements, for instance, the cash flow waterfall or the tranching (subordination) structure;
- interest rate modeling (if interest rates are not fully hedged).

The challenging part is hidden in the keywords

- factor model and
- portfolio model

at the bottom of Figure 8.6. In other words, the modeling of dependencies in the underlying reference portfolio is the truly challenging part in a CDO model. In [24] the impact of different copula choices on the risk/return profile of tranches is mathematically explained and by means of examples illustrated. The conclusion is that *the choice of the dependence model, say, the portfolio copula function, really makes a huge difference.* In our discussion of different copulas in Chapter 2 we already experienced why the copula choice really matters. The higher the dependencies the fatter will be the tail of the loss distribution and the higher will be the risk for senior tranches. The more uncoordinated risks in the portfolio behave, the lower will be the risk of senior tranches and the higher will be the risk for lower tranches. These and other rules of thumb can be made mathematically precise and have immediate implications in CDO modeling and valuation.

Again we can draw a conclusion from that in order to understand what happened in the most recent crisis. Most of the current market CDO models rely on the Gaussian copula function (see Chapter 2). As elaborated in [24], the Gaussian copula implicitely assumes that *in a heavy tail scenario single-name credit risks behave completely uncoordinated.* The mathematical term for this phenomenon is *asymptotic independence.* Unfortunately, the most recent crisis as well as past crisis scenarios clearly showed that the opposite is the case: in a heavy crisis scenario we find more coordinated behavior than in a normal market scenario. Therefore, the Gaussian copula is not the most suitable choice if markets are captured by a heavy crisis. Some people conjectured that the comprehensive use of the Gaussian copula across markets made analysts "blind" for possible scenarios in a tail event like the crisis we just experienced. One could continue to argue and debate forever when it comes to copulas. What one can say is that it is really necessary to re-think the dependence model part of CDO models in particular and, as already indicated earlier in this book, the modeling of economic capital and other risk measures in general. The mathematical tools are there, banks just have to use them. An overall guiding principle hereby is to introduce just as much complexity in models as necessary but sufficient complexity to avoid overlooking important aspects of the problem. Model risk is a reality banks have to deal with. It is important that senior credit risk modelers with access to decision makers in banks use their face time with executives to make this point clear whenever necessary. And, as already mentioned in the preface of the second edition of this book, looking at a problem from different angles via different models also is a good idea. Very often one particular model captures several particular aspects of a problem but another model might help to shed some light on some aspect the other model is overlooking. CDO modeling in the future will remain an important subdiscipline of credit risk modeling and we hope that the various products in the market will continue to drive model development.

References

[1] C. Acerbi. Spectral measures of risk: A coherent representation of subjective risk aversion. *Journal of Banking & Finance*, 26:1505–1518, 2002.

[2] C. Acerbi and D. Tasche. On the coherence of expected shortfall. *Journal of Banking & Finance*, 26:1487–1503, 2002.

[3] E. I. Altman, A. Resti, and A. Sironi. Analyzing and explaining default recovery rates. A report submitted to The International Swaps & Derivatives Association, December 2001.

[4] J. D. Amato and J. Gyntelberg. CDS index tranches and the pricing of credit risk correlations. *BIS Quarterly Review*, 2005.

[5] M. Ammann. *Credit Risk Valuation*. Springer, 2002.

[6] L. Andersen. Portfolio losses in factor models: Term structures and intertemporal loss dependence. `http://www.defaultrisk.com`, 2006.

[7] L. Andersen, D. Baum, and B. Kologlu. Single tranche cdos. tailored investment grade portfolio exposure. *Bank of America Securities*, 2003.

[8] A. Antonov, S. Mechkov, and T. Misirpashaev. Analytical techniques for synthetic CDOs and credit default risk measures. `http://www.defaultrisk.com`, 2005.

[9] M. Araten and L.Angbazo. Roots of transition matrices: Application to settlement risk. Chase Manhattan Bank, 1997. Practical Paper.

[10] P. Artzner, F. Delbaen, J. Eber, and D. Heath. Thinking coherently. *RISK*, 10(11):68–71, 1997.

[11] P. Artzner, F. Delbaen, J. Eber, and D. Heath. Coherent measures of risk. *Mathematical Finance*, 9(3):203–228, 1999.

[12] H. Bauer. *Probability Theory.* de Gruyter, 1996.

[13] H. Bauer. *Measure and Integration Theory.* de Gruyter, 2001.

[14] M. Baxter. Dynamic modelling of single-name credits and CDO tranches. Preprint, 2006.

[15] M. Baxter and A. Rennie. *Financial Calculus. An introduction to derivative pricing.* Cambridge University Press, 1996.

[16] S. Benvegnu, C. Bluhm, and C. Müller. *A Guide to Active Credit Portfolio Management.* RISK Books, 2008.

[17] N. H. Bingham and R. Kiesel. *Risk-Neutral Valuation. Pricing and Hedging of Financial Derivatives.* Springer Finance. Springer, 3rd edition, 2000.

[18] F. Black and J. C. Cox. Valuing corporate securities: Some effects of bond indenture provisions. *Journal of Finance*, 31:351–367, 1976.

[19] F. Black and M. Scholes. The pricing of options and corporate liabilities. *Journal of Political Economy*, (81):637–654, 1973.

[20] C. Bluhm and L. Overbeck. Estimating systematic risk in uniform credit portfolios. Working paper, Deutsche Bank Group, December 1999.

[21] C. Bluhm and L. Overbeck. Comonotonic default quote paths for basket evaluation. *RISK*, 18(8):67–71, 2005.

[22] C. Bluhm and L. Overbeck. Semi-analytic approaches to CDO modeling. *Economic Notes*, 33(2):233–255, 2005.

[23] C. Bluhm and L. Overbeck. Calibration of PD term structures: to be Markov or not to be. *RISK*, 20(11):98–103, 2007.

[24] C. Bluhm and L. Overbeck. *Structured Credit Portfolio Analysis, Baskets & CDOs.* Chapman & Hall/CRC, 2007.

[25] C. Bluhm, L. Overbeck, and C. Wagner. Irreconcilable differences. *RISK*, 14(10):S33–S37, October 2001.

[26] C. Bluhm and C. Wagner. Rethinking credit risk modeling. to appear, 2010.

[27] J. R. Bohn. A survey of contingent-claims approaches to risky debt valuation. Working paper, June 1999.

[28] N. L. Bowers, H. U. Gerber, J. C. Hickman, D. A. Jones, and C. J. Nesbitt. *Actuarial Mathematics.* Schaumberg, Illinois, 2nd edition, 1997. Society of Actuaries.

[29] H.-J. Brasch. A note on efficient pricing and risk calculation of credit basket products. `http://www.defaultrisk.com`, 2004.

[30] X. Burtschell, J. Gregory, and J.-P. Laurent. A comparative analysis of CDO pricing models, 2005.

[31] U. Cherubini, E. Luciano, and W. Vecciato. *Copula Methods in Finance.* Wiley, 2004.

[32] M. Choudry. *Structured Credit Products: Credit Derivatives and Synthetic Securitization.* Wiley, 2004.

[33] D. R. Cox and D. Oakes. *Analysis of Survival Data.* Chapman & Hall/CRC, 1984.

[34] J. C. Cox, J. E. Ingersoll, and S. A. Ross. A theory of term structure of interest rates. *Econometrica*, 53:385–407, 1985.

[35] Credit Suisse Financial Products. *CreditRisk$^+$ – A Credit Risk Management Framework*, 1997.

[36] P. Crosbie. Modeling default risk. KMV Corporation, `http://www.kmv.com`, 1999.

[37] M. Crouhy, D. Galai, and R. Mark. A comparative analysis of current credit risk models. *Journal of Banking & Finance*, 24:59–117, 2000.

[38] M. Crouhy, D. Galai, and R. Mark. *Risk Management.* McGraw-Hill, 2000.

[39] H. E. Daniels. Saddlepoint approximations in statistics. *Annals of Mathematical Statistics*, 25(25):631–650, 1954.

[40] H. E. Daniels. Tail probability approximation. *International Statistical Review*, 55(1):37–48, 1987.

[41] S. R. Das. Structured notes and derivatives embedded securities. Euromoney Publications PLC, 1996.

[42] A. Dembo and O. Zeitouni. *Large deviation techniques and applications (Applications of Mathematics).* Springer, 1998.

[43] M. Denault. Coherent allocation of risk capital. `http://www.risklab.ch/Papers.html`, 1999.

[44] Deutsche Bank Fixed Income Research. *Credit Derivatives and Structured Credit*, August 2000.

[45] D.v. Deventer, K. Imai, and M. Mesler. *Advanced Financial Risk Management: Tools & Techniques for Integrated Credit Risk and Interest Rate Risk Management*. Wiley, 2004.

[46] D. Duffie. *Dynamic Asset Pricing Theory*. Princeton University Press, 1992.

[47] D. Duffie and N. Gârleanu. Risk and valuation of collateralized debt obligations. Working Paper, 1999. Graduate School of Business, Stanford University.

[48] D. Duffie and K. Singleton. An econometric model of the term structure of interest-rate swap yields. *Journal of Finance*, 51(1):1287–1321, 1997.

[49] D. Duffie and K. Singleton. Simulating correlated defaults. Working Paper, 1998. Graduate School of Business, Stanford University.

[50] D. Duffie and K. Singleton. Modeling term structures of defaultable bonds. *Review of Financial Studies*, 12:687–720, 1999.

[51] D. Duffie and K. J. Singleton. *Credit Risk: Pricing, Measurement, and Management*. Princeton University Press, New Jersey, 2003.

[52] N. Dunford and J. Schwartz. *Linear Operators - Part I: General Theory*. Interscience Publishers, New York, 1958.

[53] E. Eberlein, R. Frey, M. Kalkbrener, and L. Overbeck. Mathematics in financial risk management. `http://www.defaultrisk.com`, 2007.

[54] E. Eberlein and U. Keller. Hyperbolic distributions in finance. *Bernoulli*, 1:281–299, 1995.

[55] D. Egloff, M. Leippold, S. Jöhri, and C. Dalbert. Optimal importance sampling for credit portfolios with stochastic approximation. `http://www.defaultrisk.com`, 2005.

[56] P. Embrechts, A. McNeil, and D. Straumann. Correlation and dependence in risk management: Properties and pitfalls. Preprint, July 1999. ETH Zürich.

[57] B. Engelmann and R. Rauhmeier. *The Basel II Risk Parameters: Estimation, Validation, and Stress Testing*. Springer, 2006.

[58] F. J. Fabozzi. *Investing in Asset-Backed Securities*. McGraw Hill, 2001.

[59] J. Felsenheimer, P. Gisdakis, and M. Zaiser. DJ iTraxx: Credit at its best! Credit Derivatives Special, HypoVereinsbank Corporates & Markets, 2004.

[60] J. Felsenheimer, P. Gisdakis, and M. Zaiser. *Active Credit Portfolio Management*. Wiley-VCH, 2006.

[61] C. C. Finger. Conditional approaches for creditmetrics portfolio distributions. *CreditMetrics*^{TM} *Monitor*, April 1999.

[62] C. C. Finger. A comparison of stochastic default rate models. *RiskMetrics Journal*, pages 49–73, November 2000.

[63] H. Föllmer. Calcul d'Itô sans Probabilités. volume 850 of *Lecture Notes in Mathematics*, pages 143–150. Springer, New York, 1981. Seminaire Probabilités, Strasbourg XV.

[64] H. Föllmer and A. Schied. Convex measures of risk and trading constraints. *Finance and Stochastics*, 6:429–447, 2002.

[65] H. Föllmer and A. Schied. *Stochastic Finance - An introduction in discrete time*. de Gruyter, Berlin, New York, 2004.

[66] J. Fons. Using default rates to model the term structure of credit risk. *Financial Analysts Journal*, pages 25–32, 1994.

[67] E. W. Frees and E. A. Valdez. Understanding relationships using copulas. 1997. Paper presented at the 32nd Actuarial Research Conference, held August 6-8, at the University of Calgary, Alberta, Canada.

[68] R. Frey and A. J. McNeil. Modelling dependent defaults. `http://www.defaultrisk.com`, March 2001.

[69] R. Frey and A. J. McNeil. Var and expected shortfall in portfolios of dependent credit risks: Conceptual and practical insights. `http://www.math.ethz.ch/~mcneil/ftp/rome.pdf`, 2002.

[70] R. Frey, A. J. McNeil, and M. Nyfeler. Copulas and credit models. *RISK*, 14(10):111–114, 2001.

[71] M. Frittelli and E. R. Gianin. Putting order in risk measures. *Journal of Banking & Finance*, 26:1473–1486, 2002.

[72] S. Fritz, M. Luxenburger, and T. Miehe. Implementation of an IRB compliant rating system. In M. Ong, editor, *The Basel Handbook - A Guide for Financial Practitioners*. RISK Books, 2004.

[73] H. Frydman and T. Schuermann. Credit rating dynamics and Markov mixture models. Working Paper, 2005.

[74] J. Frye. Collateral damage. *RISK*, 13(4):91–94, 2000.

[75] J. Frye. Depressing recoveries. *RISK*, 13(11):108–111, 2000.

[76] P. Georges, A-G. Maly, E. Nicolas, G. Quibel, and T. Roncalli. Multivariate survival modelling: a unified approach with copulas. `http://gro.creditlyonnais.fr`, 1996.

[77] K. Giesecke. Successive correlated defaults: Pricing trends and simulation. Cornell University, 2003.

[78] K. Giesecke. Default and information. *Journal of Economic Dynamics and Control*, 30:2281–2303, 2006.

[79] P. Glasserman. *Monte Carlo Methods in Financial Engineering*. Springer, Berlin, 2003.

[80] P. Glasserman and J. Li. Importance sampling for a mixed Poisson model of portfolio credit risk. In S. Chick, P. Sanchez, D. Ferrin, and D. Morrice, editors, *Proceedings of the 2003 Winter Simulation Conference*. IEEE, 2005.

[81] P. Glasserman and J. Li. Importance sampling for portfolio credit risk. *Management Science*, 51:1643–1656, 2005.

[82] M. Gordy. Saddlepoint approximation of CreditRisk+. *Journal of Banking & Finance*, 26(7):1335–1353, 2002.

[83] M. B. Gordy. A comparative anatomy of credit risk models. *Journal of Banking & Finance*, 24:119–149, 2000.

[84] M. B. Gordy. A risk-factor model foundation for ratings-based bank capital rules. Draft, February 2001.

[85] G. N. Gregoriou and C. Hoppe. *The Handbook of Credit Portfolio Management.* McGraw-Hill, 2009.

[86] G. R. Grimmet and D. Stirzaker. *Probability and Random Processes.* Oxford University Press, 2nd edition, 1992.

[87] M. Gundlach and F. Lehrbass, editors. *CreditRisk+ in the Banking Industry.* Springer, Berlin, 2004.

[88] G. M. Gupton, C. C. Finger, and M. Bhatia. CreditMetrics – Technical Document. Morgan Guaranty Trust Co., `http://www.creditmetrics.com`, 1997.

[89] J. M. Harrison and S. Pliska. Martingales and stochastic integrals in the theory of continuous trading. *Stochastic Processes and Their Applications*, 11:215–260, 1981.

[90] D. Heath, R. Jarrow, and A. Morton. Bond pricing and the term structure of interest rates: A new methodology for contingent claims valuation. *Econometrica*, 60:77–105, 1992.

[91] D. Heath and H. Kuhn. Pareto equilibria with coherent measures of risk. *Mathematical Finance*, 14:163–172, 2004.

[92] M. Hillebrand. Modeling and estimating dependent loss given default. Working Paper, Munich University of Technology, 2005.

[93] S. Höse, S. Huschens, and R. Wania. Rating migration. In W. Härdle, N. Hautsch, and L. Overbeck, editors, *Applied Quantitative Finance*, pages 105–123. Springer, 2009.

[94] J. C. Hull. *Option, Futures, and Other Derivatives.* Prentice Hall, 4th edition, 2000.

[95] J. C. Hull and A. White. Valuing credit default swaps II: Modeling default correlations. `http://www.defaultrisk.com`.

[96] J. C. Hull and A. White. The impact of default risk on options and other derivatives. *Journal of Banking & Finance*, 19(2):299–322, 1995.

[97] J. C. Hull and A. White. Valuing credit default swaps I: No counterparty default risk. *Journal of Derivatives*, 8(1):29–40, 2000.

[98] International Swap and Derivatives Association Inc. *Credit derivatives documentation*, 1998.

[99] R. Israel, J. Rosenthal, and J. Wei. Finding generators for Markov chains via empirical transition matrices with application to credit ratings. *Mathematical Finance*, 11(2):245–265, 2001.

[100] J. P. Morgan Securities. *The J. P. Morgan guide to credit derivatives*, 1999.

[101] P. Jäckel. *Monte Carlo Methods in Finance.* John Wiley, Berlin, 2002.

[102] J. James. Total return swaps. *The Electronic Journal of Financial Risk*, 1(2), 1998.

[103] R. A. Jarrow, D. Lando, and S. M. Turnbull. A Markov model for the term structure of credit risk spreads. *Review of Financial Studies*, 10:481–523, 1997.

[104] R. A. Jarrow and S. M. Turnbull. Pricing options on financial securities subject to default risk. *Journal of Finance*, 50:53–86, 1995.

[105] S. Jaschke and U. Küchler. Coherent risk measures, valuation bounds, and (μ, σ)-portfolio optimization. *Finance and Stochastics*, 5(2):181–200, 2000.

[106] J. Jensen. *Saddle Point Approximation.* Oxford University Press, 1995.

[107] H. Joe. *Multivariate Models and Dependence Concepts.* Chapman & Hall/CRC, 1997.

[108] A. Kadam and P. Lenk. Heterogeneity in ratings migration. Working Paper, 2005.

[109] M. Kalkbrener. An axiomatic approach to capital allocation. *Mathematical Finance*, 15(3):425–437, 2005.

[110] M. Kalkbrener, H. Lotter, and L. Overbeck. Sensible and efficient capital allocation for credit portfolios. *RISK*, 1:S19–S24, January 2004.

[111] O. Kallenberg. *Foundations of Modern Probability.* Probability and its Applications. Springer, 1997.

[112] I. Karatzas and S. E. Shreve. *Brownian Motion and Stochastic Calculus.* Springer, 2nd edition, 1991.

[113] I. Karatzas and S. E. Shreve. *Methods of Mathematical Finance.* Springer, 1998.

[114] S. Kealhofer, S. Kwonk, and W. Weng. Uses and abuses of bond default rates. KMV Corporation, 1998.

[115] H. U. Koyluoglu and A. Hickman. Reconcilable differences. *RISK*, 56, October 1998.

[116] A. Kreinin and M. Sidelnikova. Regularization algorithms for transition matrices. *Algo Research Quarterly*, 4(1/2):25–40, 2001.

[117] S. Kusuoka. On law invariant coherent risk measures. *Advances in Mathematical Economics*, 3:83–95, 2001.

[118] D. Lamberton and B. Lapeyre. *Introduction to Stochastic Calculus Applied to Finance.* Chapman & Hall/CRC, 1996.

[119] D. Lando. *Modelling Bonds and Derivatives with Default Risk.* Mathematics of Derivatives Securities, M. Dempster and S. Pliska, eds. Cambridge University Press, 1997.

[120] D. Lando. *Credit Risk Modelling: Theory and Applications.* Princeton University Press, 2004.

[121] H. E. Leland. Corporate debt value, bond convenants, and optimal capital structure. *Journal of Finance*, 49:1213–1252, 1994.

[122] H. E. Leland. A simple approach to valuing risky fixed and floating rate debt. *Journal of Finance*, 50:789–819, 1995.

[123] H. E. Leland and K. Toft. Optimal capital structure, endogeneous bankruptcy, and the term structure of credit spreads. *Journal of Finance*, 51:987–1019, 1994.

[124] D. X. Li. The valuation of basket credit derivatives. *CreditMetrics^{TM} Monitor*, pages 34–50, April 1999.

[125] D. X. Li. On default correlation: A copula function approach. *Journal of Fixed Income*, 6:43–54, March 2000.

[126] A. Lipton and A. Rennie. *Credit Correlation. Life after Copulas.* World Scientific, 2008.

[127] F. Longstaff and E. Schwartz. A simple approach to valuing risky fixed and floating rate debt. *Journal of Finance*, 50:789–819, 1995.

[128] F. Longstaff and E. Schwartz. Valuing credit derivatives. *Journal of Fixed Income*, pages 6–12, June 1995.

[129] P. Malliavin. *Integration and Probability.* Springer, 1995.

[130] R. Martin. *Credit Portfolio Modeling Handbook.* Credit Suisse First Boston, 2004.

[131] R. Martin and R. Ordovas. An indirect view from the saddle. *RISK*, 19(10):94–99, 2006.

[132] R. Martin, K. Thompson, and C. Browne. Taking to the saddle. *RISK*, 14(7):91–94, June 2001.

[133] R. Mashal and M. Naldi. Extreme vents and default baskets. *RISK*, 15(6):119–122, June 2002.

[134] B. Masters. Credit derivatives and the management of credit risk. *The Electronic Journal of Financial Risk*, 1(2), 1998.

[135] McKinsey & Company, German Office. *CreditPortfolioView 2.0*, June 2001. Technische Dokumentation.

[136] A. J. McNeil, R. Frey, and P. Embrechts. *Quantitative Risk Management. Concepts, Techniques and Tools.* Princeton University Press, 2005.

[137] R. Merton. On the pricing of corporate debt: The risk structure of interest rates. *The Journal of Finance*, 29:449–470, 1974.

[138] S. Meyer-Dautrich and C.Wagner. Minimum-entropy calibration of cdo-tranches. `http://www.defaultrisk.com`, 2007.

[139] T. Mikosch. *Elementary Stochastic Calculus – with Finance in View*, volume 6 of *Advanced Series on Statistical Science & Applied Probability.* World Scientific Publishing, 1998.

[140] Moody's Investors Service. *Rating Cash Flow Transactions Backed by Corporate Debt 1995 Update*, April 1995.

[141] Moody's Investors Service. *Default and Recovery Rates of Corporate Bond Issuers: 2000*, February 2001.

[142] D. Murphy. Keeping credit under control. *RISK*, 9, September 1996.

[143] S. R. Neal. Credit derivatives: New financial instruments for controlling credit risk. *Economic Review*, 1996.

[144] R. Nelsen. *An Introduction to Copulas.* Springer, New York, 1999.

[145] C. Nelson and A. Siegel. Parsimonious modeling of yield curves. *Journal of Business*, 60:473–489, 1987.

[146] P. Nickell, W. Perraudin, and S. Varotto. Ratings- versus equity-based credit risk modeling: an empirical analysis. `http://www.bankofengland.co.uk/workingpapers/`, 1999. Working paper.

[147] J. Norris. *Markov Chains.* Cambridge Series in Statistical and Probabilistic Mathematics. Cambridge University Press, 1998.

[148] Basel Committee on Banking Supervision. *The Internal Ratings-Based Approach.* Supporting Document to the New Basel Capital Accord, 2001.

[149] Basel Committee on Banking Supervision. International convergence of capital measurement and capital standards. `http///www.bis.org`, June 2004.

[150] M. K. Ong. *Internal Credit Risk Models.* RISK Books, 1999.

[151] M. K. Ong. *The Basel Handbook: A Guide for Financial Practitioners.* RISK Books, 2006.

[152] L. Overbeck. Allocation of economic capital in loan portfolios. In U. Franke, W. Härdle, and G. Stahl, editors, *Measuring Risk in Complex Stochastic Systems.* Springer, New York, 2000.

[153] L. Overbeck. Spectral capital allocation. In *Economic Capital: A practitioner guide.* RISK Books, London, 2004.

[154] L. Overbeck and W. Schmidt. Modeling default dependence with threshold models. *Journal of Derivatives*, 12(4):10–19, 2005.

[155] W. R. Pestman. *Mathematical Statistics.* de Gruyter, 1998.

[156] W. Press, S. Teukolsky, W. Vetterling, and B. Flannery. *Numerical Recipes: The Art of Scientific Computing.* Cambridge University Press, 3rd edition edition, 2007.

[157] PriceWaterhouseCoopers. *The PriceWaterhouseCoopers Credit Derivatives Primer*, 1999.

[158] J. Rank. *Copulas: From Theory to Application in Finance.* RISK Book, 2006.

[159] D. Revuz and M. Yor. *Continuous Martingales and Brownian Motion*. Springer, 1991. Chapter IV (3.13).

[160] J. A. Rice. *Mathematical Statistics and Data Analysis*. Duxbury Press, 2nd edition, 1995.

[161] B. Schmid. *Credit Risk Pricing Models: Theory and Practice*. Springer, 2004.

[162] W. Schmidt and I. Ward. Pricing default baskets. *RISK*, 15(1):111–114, January 2002.

[163] P. Schönbucher. Factor models for portfolio credit risk. *Preprint, University of Bonn, Germany*, 2001.

[164] P. Schönbucher. *Credit Derivatives Pricing Models*. John Wiley, 2003.

[165] T. Schuermann and Y. Jafry. Measurement and estimation of credit migration matrices. Wharton School Center for Financial Institutions, University of Pennsylvania, Working Paper, 2003.

[166] T. Schuermann and Y. Jafry. Metrics for comparing credit migration matrices. Wharton School Center for Financial Institutions, University of Pennsylvania, Working Paper, 2003.

[167] Moody's Investor Services. Corporate default and recovery rates, 1920–2008. *Moody's Global Credit Policy*, (February), 2009.

[168] A. Sklar. Fonction de repartition à n dimension et leur marges. *Publications de l'Insitute Statistique de l'Université de Paris*, 8:229–231, 1959.

[169] A. Sklar. Random variables, joint distribution functions and copulas. *Kybernetika*, 9:449–460, 1973.

[170] J. R. Sobehart and S. C. Keenan. An introduction to market-based credit analysis. *Moody's Risk Management Services*, November 1999.

[171] SPNew. 2008 annual global corporate default study and rating transitions. *RatingsDirect*, (April), 2009.

[172] Standard & Poor's. *Standard & Poor's Corporate Ratings Criteria 1998*.

[173] W. Stromquist. Roots of transition matrices. Daniel H. Wagner Associates, 1996. Practical Paper.

[174] D. Tasche. Risk contributions and performance measurement. `http://www.ma.tum.de/stat/`, 1999.

[175] D. Tasche. Calculating value-at-risk contributions in CreditRisk+. `http://arxiv.org/abs/cond-mat/0112045`, 2000.

[176] D. Tasche. Conditional expectation as quantile derivative. `http://www.ma.tum.de/stat/`, 2000.

[177] D. Tasche. Expected shortfall and beyond. *Journal of Banking & Finance*, 26:1519–1533, 2002.

[178] K. Thompson, A. McLeod, P. Teklos, and S. Gupta. Time for multi-period capital models. *RISK*, pages 74–78, October 2005.

[179] S. Trueck and E. Oezturkmen. Adjustment and application of transition matrices in credit risk models. University of Karlsruhe, Working Paper, 2003.

[180] S. Trueck and S. T. Rachev. *Rating Based Modeling of Credit Risk*. Academic Press, 2009.

[181] O. A. Vasicek. An equilibrium characterization of the term structure. *Journal of Financial Economics*, 5:177–188, 1977.

[182] O. A. Vasicek. Probability of loss on loan portfolio. KMV Corporation, 1987.

[183] K. Vogl. *Modellierung der Zeitstruktur von Ratingmodellen*. Shaker Verlag, 2008.

[184] N. Wagner. *Credit Risk: Models, Derivatives, and Management*. Chapman & Hall/CRC, 2008.

[185] S. S. Wang. Aggregation of correlated risk portfolios: Models & algorithms.

[186] S. Weber. Distribution invariant risk measures, entropy, and large deviations. Dept. of Mathematics, Humboldt Universität Berlin, preprint, 2004.

[187] T. Wilde. IRB approach explained. *RISK*, 14(5), 2001.

[188] T. Wilde and L. Jackson. Low-default portfolios without simulation. *RISK*, 19(8):60–63, 2006.

[189] T. Wilson. Portfolio credit risk, part i. *RISK*, pages 111–117, 1997.

[190] T. Wilson. Portfolio credit risk, part ii. *RISK*, pages 56–61, 1997.

[191] J. Xiao. Importance sampling for credit portfolio simulation. *RiskMetrics Group*, pages 24–30, January 2002.

[192] J. Yang, T. Hurd, and X. Zhang. Saddle point approximation method for pricing CDOs. *Journal of Computational Finance*, 10(1):1–20, 2006.

[193] C. Zhou. A jump-diffusion approach to modeling credit risk and valuing defaultable securities. `http://www.defaultrisk.com`, 1997.

[194] C. Zhou. An analysis of default correlations and multiple defaults. *The Review of Financial Studies*, 14:555–576, 2001.

[195] D. Zwillinger. *Handbook of Differential Equations*. Academic Press, Boston, 1995.

Index

Finance and Investing

Introduction to Credit Risk Modeling

Second Edition

The recent financial crisis has shown that credit risk in particular and finance in general remain important fields for the application of mathematical concepts to real-life situations. While continuing to focus on common mathematical approaches to model credit portfolios, **Introduction to Credit Risk Modeling, Second Edition** presents updates on model developments that have occurred since the publication of the best-selling first edition.

New to the Second Edition

- An expanded section on techniques for the generation of loss distributions
- Introductory sections on new topics, such as spectral risk measures, an axiomatic approach to capital allocation, and nonhomogeneous Markov chains
- Updated sections on the probability of default, exposure-at-default, loss-given-default, and regulatory capital
- A new section on multi-period models
- Recent developments in structured credit

The financial crisis illustrated the importance of effectively communicating model outcomes and ensuring that the variation in results is clearly understood by decision makers. The crisis also showed that more modeling and more analysis are superior to only one model. This accessible, self-contained book recommends using a variety of models to shed light on different aspects of the true nature of a credit risk problem, thereby allowing the problem to be viewed from different angles.

C8992

ISBN: 978-1-58488-992-2

90000

9 781584 889922

CRC Press
Taylor & Francis Group
an informa business
www.crcpress.com

6000 Broken Sound Parkway, NW
Suite 300, Boca Raton, FL 33487
270 Madison Avenue
New York, NY 10016
2 Park Square, Milton Park
Abingdon, Oxon OX14 4RN, UK