Modeling Credit Risk and Credit Derivatives

Vincent Leijdekker

Modeling Credit Risk and Credit Derivatives

Vincent Leijdekker

Modeling Credit Risk and Credit Derivatives

Modeling Credit Risk and Credit Derivatives
Het Modelleren van Kredietrisico en Kredietderivaten
Vincent Johannes Gerardus Leijdekker
Proefschrift Universiteit van Amsterdam
ISBN: 978-90-9025339-8

Dit proefschrift werd mede mogelijk gemaakt door

Thomas Stieltjes Institute
for Mathematics

Modeling Credit Risk and Credit Derivatives

ACADEMISCH PROEFSCHRIFT

ter verkrijging van de graad van doctor
aan de Universiteit van Amsterdam
op gezag van de Rector Magnificus prof. dr. D.C. van den Boom
ten overstaan van een door het college voor promoties ingestelde
commissie, in het openbaar te verdedigen in de Agnietenkapel
op dinsdag 25 mei 2010, te 14:00 uur

door

Vincent Johannes Gerardus Leijdekker

geboren te Bussum

Promotiecommissie

Promotores:	prof. dr. M.R.H. Mandjes dr. P.J.C. Spreij
Overige leden:	prof. dr. J.J.A.G. Driessen prof. dr. C.A.J. Klaassen prof. dr. A. Lucas prof. dr. M.H. Vellekoop prof. dr. A.C.F. Vorst dr. R.G. de Vilder

Faculteit der Natuurwetenschappen, Wiskunde en Informatica

Dankwoord

Na ruim vier jaar aan mijn onderzoek en proefschrift te hebben gewerkt, ben ik nu dan eindelijk aanbeland bij het laatste, en ongetwijfeld belangrijkste, gedeelte: Het bedanken van alle personen zonder wie ik dit proefschrift nooit had kunnen voltooien.

Ten eerste wil ik mijn begeleider en copromotor Peter Spreij bedanken. Op jouw initiatief hebben we het project tussen de Universiteit van Amsterdam en de ABN Amro opgezet, wat mij in staat heeft gesteld om te gaan promoveren. Met mijn vragen kon ik altijd bij jou terecht, waarna jij mij weer verder op weg kon helpen, vaak met behulp van een van de boeken uit jouw eigen bibliotheek. Zonder jouw hulp was het mij waarschijnlijk nooit gelukt om het proefschrift in iets meer dan vier jaar af te ronden.

Natuurlijk wil ik ook mijn promotor Michel Mandjes bedanken. Gedurende mijn promotietijd hebben we steeds vaker contact gehad en zijn we elkaar steeds beter leren kennen. Zeker in de laatste twee jaar van mijn onderzoek hebben jouw ideeën een grote bijdrage geleverd aan mijn proefschrift. Door regelmatig de voortgang te bespreken, heb je ervoor gezorgd dat mijn onderzoek en mijn proefschrift op tijd af zijn gekomen.

De commissieleden wil ik graag bedanken voor de tijd en moeite die zij hebben genomen om mijn proefschrift door te lezen en voor het plaatsnemen in de promotiecommissie.

Zonder hulp van de ABN Amro, in het bijzonder van Ton Vorst en Martijn van der Voort, had ik dit onderzoek nooit kunnen uitvoeren. Ton, dankzij jouw bereidheid met mij en de Universiteit van Amsterdam in zee te gaan heb ik niet alleen mijn promotieonderzoek kunnen doen, maar ook heb ik veel werkervaring op kunnen doen

buiten het onderzoek. Martijn, ik heb veel van je kunnen leren en onze samenwerking is mij altijd goed bevallen. Ons gezamenlijke empirische onderzoek heeft ook een grote bijdrage geleverd aan dit proefschrift. Ook wil ik Bert-Jan Nauta bedanken dat hij mij binnenboord heeft gehouden na de overname van de ABN Amro.

Mijn ouders ben ik vanzelfsprekend zeer dankbaar voor de steun die ze me altijd hebben gegeven. Zonder jullie had ik mijn studie nooit zonder zorgen af kunnen ronden en had ik nooit kunnen beginnen met promoveren. Uiteraard wil ik mijn vriendin Jessica hier ook bedanken. Jij hebt steeds weer als eerste mijn geklaag aan moeten horen als ik de moed weer eens begon te verliezen. Ik ben je dan ook heel dankbaar dat mij iedere keer hebt weten te motiveren om toch weer verder te gaan.

Mijn broers Martijn en Sander wil ik bedanken dat ze mij willen bijstaan tijdens de verdediging van mijn proefschrift. Ik vind het erg leuk dat jullie mijn paranimfen willen zijn.

Contents

CHAPTER 1

Introduction

This thesis is about managing risk, which is an important task of banks and other financial institutions. These companies have large portfolios with different risk-bearing objects. We will contribute to the theory of the quantification of risk, and in particular credit risk. The first part of this thesis focuses on measuring the credit risk in large portfolios, whereas the second part deals with the valuation of credit derivatives.

1.1 Risk

We start by presenting different types of risk that a bank, or other financial institution, has to deal with. We first decompose the general notion of risk into three risk types, which are thereafter discussed more detail. We mainly focus on the credit risk, as this is the main topic of this thesis.

1.1.1 Risk Types

In its day-to-day business, a bank or other financial institution is always exposed to risk. This risk can be split up into a number of components, which we discuss in this section. Before we focus on the three main types of risk, we consider some examples of risks that a bank could encounter.

One of the core functions of a bank is lending out money to its customers, either private or corporate. The main risk that a bank faces when lending out money, is that the customer fails to make interest payments or that he fails to repay the loan. In such a case a the bank suffers financially.

Banks also face the risk that their investments drop in value. For example, banks usually own large numbers of stocks in several companies. In case these stocks drop in value, a bank incurs a loss. Also changes in foreign exchange rates could lead to losses for a bank. When, for example, a bank expects to receive a certain amount in a foreign currency one month from now, which is worth one million Euro today, could be worth less at the time the money is actually received, leading to loss for the bank.

Further, a bank also faces risks that are not directly related to loans or to their investments. These risks are related to the functioning of the bank. Banks for example suffer from fraud, either by employees or by clients, but also computer network problems could lead to serious problems and loss of clients.

In all these cases the bank suffers financially, and thus the occurrence of any of the events above has negative consequences for the bank. Therefore it is important that the bank measures and manages risk. In a first step to do so, we split up the risk into a number of components. In general we can identify three different types of risk.

- Credit risk, the risk that a debtor, or obligor, does not honor its payment obligations.
- Market risk, the risk that an investment decreases in value due to moves in market factors, such as the stock values or interest rates .
- Operational risk, the risk arising from the bank's business function.

Looking at the examples above, we can classify the first example as credit risk, the second as market risk and the last example as operational risk. The three risk types can be decomposed into more detailed risk types. Following [Sch03a], credit risk consists of the components

- Arrival risk, the risk that a default occurs.
- Timing risk, when the default occurs.
- Recovery risk, in case that a default occurs, which fraction of the investment is lost.
- Default correlation risk, one default could indicate a bad state of the economy, which could yield more defaults.

The arrival risk and timing risk are very much related. Clearly, when one knows the timing of the default, it is directly clear that the default event has occurred. In Section 1.1.2 we consider credit risk in more detail.

Market risk can be decomposed into the, self-evident, market factors

- equity risk, the risk of changes in stock prices,

- interest rate risk, the risk that interest rates change,
- currency risk, the risk that foreign exchange rates change,
- commodity risk, the risk that prices of commodities change.

Clearly, in most transactions that a bank enters into, it faces at least one of these risks.

Operational risk can be decomposed into many different types of risk, which cannot be seen as either market or credit risk. Event types that fall under operational risk are fraud, employment practice and workplace safety, clients product and business practice, damage to physical assets, business disruption and system failures, and execution, delivery and process management.

1.1.2 Credit Risk

In the previous section we have discussed a number of risk types to which banks, and other financial institutions are exposed. In this section we consider one specific type of risk, credit risk, which is commonly referred to as default risk as well. Above we have introduced the notion of credit risk, and we have identified four components of credit risk. All components rely in some way on the occurrence of the default event. In the previous section we have already seen some examples of default events. In this section we discuss two ways banks manage credit risk. A bank can reserve a certain amount of money to cover potential losses due to default events. Further, a bank can transfer the credit risk to other market parties, by entering a credit derivatives contract.

To determine the amount of money the bank has to reserve, it has to determine the total amount of money that could be at risk, and next to this it has to determine the probability that the default events occur. For this purpose the timing and size of the defaults have to modeled. When modeling these two quantities one has to take into account that

- default events are rare events,
- defaults occur unexpectedly,
- default events involve large losses,
- the size of the loss is not known until the default occurs,
- defaults are correlated.

The last point in this list in only relevant, and very important, in case one models more than one default time. In Section 1.2 we discuss the modeling of the size and timing of the losses, where the above points are discussed.

Next to reserving cash to cover losses, a bank can transfer default risk by entering into a credit derivatives contract. Such contract provide protection against losses due to defaults. If for example a bank has invested a large sum into a company, a credit default swap ensures that in case this company defaults, the incurred loss is repaid. Section 1.3 describes a number of credit derivatives and how they can be used to transfer default risk.

1.2 Credit Risk Modeling

Above we have seen that credit risk can be split up into several types of risk. When we model the credit risk, we have to focus on two main components, namely the timing of the default and the amount of the loss. The definition of the default event is not relevant in the modeling, as we are only interested in the timing of the default and the size of the resulting loss.

It is common to model the default time and the loss size separately. There are two main streams of modeling the default time, structural models and reduced form models, which are discussed in Sections 1.2.1 and 1.2.2, respectively. The size of the loss is modeled as the recovery rate, which is the fraction of the loan, or other type of debt, that is repaid, and hence not lost. In Section 1.2.3 some models for the recovery rate are considered. The section is concluded with an overview of techniques that can be used to measure portfolio credit risk.

1.2.1 Structural Models

In structural models, one models the ability of a company to repay its debt. A (continuous) process V_t is considered that represents the value of the company at time $t > 0$. Further a (continuous) threshold, or barrier, B_t, is considered. This barrier can be interpreted as the level of the (total) debt of the company. We say that the company defaults the first time the value of the firm is below the barrier. This means that the default time τ is defined as

$$\tau = \inf\{t > 0 | \, V_t < B_t\}.$$

Clearly, the possible choices for the value process V_t and the barrier B_t are endless. Historically, models based on Brownian motion, such as geometric Brownian motion with drift, have become popular, since such models, under suitable conditions, allow one to derive explicit expressions for e.g. default probabilities.

In this section we consider two types of structural models in more detail. First we consider Merton's model, which is one of the simplest structural models. Thereafter we discuss a simple first passage time model, which allows us to calculate the distribution of τ explicitly.

Merton's Model

The simplest form of structural models is due to Merton [Mer74]. Debt maturing at a fixed time horizon, T, is considered. The company only defaults in case its value is below the level of the debt at T. The evolution of the company's value process and its debt before this time T is not considered. The value process, $(V_t)_{t\geq 0}$, is modeled as geometric Brownian motion

$$\mathrm{d}V_t = \mu V_t \mathrm{d}t + \sigma V_t \mathrm{d}W_t, \tag{1.1}$$

where W_t is a standard Brownian motion, μ is the drift parameter and σ is a volatility parameter. Clearly we have $V_0 > 0$ and hence $V_t > 0$ for all t. It is assumed that the level of the debt is constant over time. As we assume that defaults can only occur at T, we set $B_t = K1_{\{t=T\}}$. The default time is thus defined as

$$\tau = \begin{cases} T & \text{if } V_T < K \\ \infty & \text{else.} \end{cases} \tag{1.2}$$

The probability, at time zero and under the physical measure $\mathbb{P}$, that the company defaults at time T can easily be calculated, since V_T is lognormally distributed. This yields

$$\mathbb{P}(\tau \leq T) = \mathbb{P}(V_T < K) = \Phi\left(\frac{\log\left(\frac{K}{V_0}\right) - \left(\mu - \frac{1}{2}\sigma^2\right)T}{\sigma\sqrt{T}}\right),$$

where V_0 is the initial value of the company, and Φ is the cumulative distribution function of the standard normal distribution.

First Passage Time Model

A clear disadvantage of Merton's model is that only a fixed point in time is considered, and hence the evolution of the company's value and the company's debt before this time is ignored. In the first passage time that we consider in this section, we deal with this disadvantage with the introduction of a constant barrier D. This corresponds to setting $B_t = D$. We allow for defaults in the whole interval $[0, T]$. The value of the company is again modeled as geometric Brownian motion, (1.1), and the default time can be defined as

$$\tau = \inf\{t \leq T : \ V_t < D\}. \tag{1.3}$$

We consider a slightly more general model by combining both Merton's model and the simple first passage model. We define the default time as

$$\tau = \min\{\tau_1, \tau_2\},$$

where τ_1 is given by (1.2) and τ_2 by (1.3).

Under these settings it is possible to calculate the probability, at time zero, that the default occurs before time T.

$$\mathbb{P}(\tau \leq T) = \Phi\left(\frac{\log(K/V_0) - mT}{\sigma\sqrt{T}}\right) + \left(\frac{D}{V_0}\right)^{2m/\sigma^2} \Phi\left(\frac{\log(D^2/(KV_0)) + mT}{\sigma\sqrt{T}}\right),$$

where $m = \mu - \frac{1}{2}\sigma^2$. The derivation of this expression relies on several properties of Brownian motion, such as the reflection principle. The distribution of τ for $t < T$, can easily be calculated from this expression.

The first passage model that we consider here is actually a special case of the first passage model considered by Black and Cox, [BC76], who consider the time dependent barrier

$$B_t = Ke^{-\gamma(T-t)}.$$

It is clear that this agrees with the model we consider here in case $\gamma = 0$.

Many other authors have considered geometric Brownian motion as a basis for a structural default model. Among others, Longstaff and Schwartz, [LS95], and Kim et al., [KRS93], both consider stochastic interest rates instead, and the company defaults if the firm value is below some constant barrier. Other variations to the model by Black and Cox, involve jump-diffusion models for the firm value process, e.g. [Sch96] and [Zho96]. A partial information approach, where one assumes that the firm value is not, or only partially, observed, is considered by Giesecke [Gie04]. Bielecki and Rutkowski [BR02] and [Sch03a] provide a more elaborate discussion of structural models.

Disadvantages of Structural Models

Most of the structural models suffer from a number of disadvantages. First of all, default times are predictable when they are modeled using structural models. It is possible to construct a sequence of increasing stopping times that converge to the default time. This is not a realistic property, since defaults in general cannot be anticipated. By the introduction of random jumps in the firm value process, one can partially solve this problem, since the company can 'jump to default'. On the other hand the company value can still pass the barrier as a continuous process.

In practice, credit spreads, which are the difference between the risk free rate and the rate at which a company can borrow, is relatively high for loans with a short maturity. When modeling the default time using structural models, and thus predictable default times, it is not possible to reproduce such spread, as the default probability is very small over a short period.

Another drawback of the structural models is that in general the value process V_t and the level of the debt are not, or only partially, observable. Due to this lack of information, there is uncertainty about the occurrence of the default event. Without this information it is not possible to accurately estimate or calibrate the model parameters. The partial information approach deals with this drawback explicitly.

1.2.2 Reduced Form Models

The second type of default time models that we consider concerns reduced form models. In these models the distribution of the default time is exogenously given, in contrast to the structural models, where the (distribution of the) default time corresponds directly to the value of the company. As the default time is based on an exogenous specification of the distribution, there is not a formal definition of the default time. Instead the default time is modeled as the first jump of a Cox Process, or doubly stochastic Poisson Process. The Cox process is an extension of the inhomogeneous Poisson process. Recall that for an inhomogeneous Poisson process $N = (N_t)_{t\geq 0}$, one has for $0 < s < t$,

$$\mathbb{P}\left(N_t - N_s = k\right) = \frac{\left(\int_s^t \lambda(u)\mathrm{d}u\right)^k}{k!} \exp\left(-\int_s^t \lambda(u)\mathrm{d}u\right), \tag{1.4}$$

where $\lambda(t) \geq 0$ is the deterministic (default) intensity, for $t \geq 0$. In case of Cox processes, stochastic intensities are considered instead, and (1.4) becomes a conditional probability. The application of Cox processes to the modeling of default times has been considered by e.g. [Lan98].

The default time of a company is modeled as the first jump time of N. This means that we can use (a variation to) (1.4) to calculate the distribution of the default time τ. In the remainder of this section we consider deterministic and stochastic intensities separately.

Deterministic Intensity

We first consider a deterministic intensity, $(\lambda(t))_{t\geq 0}$, and thus we model the default time as the first jump of a inhomogeneous Poisson process. Based on (1.4) we can easily determine the distribution of the default time. The probability that the company defaults before time $t > 0$ is given by

$$\mathbb{P}\left(\tau \leq t\right) = \mathbb{P}\left(N_t > 0\right) = 1 - \mathbb{P}\left(N_t = 0\right) = 1 - e^{-\int_0^t \lambda(u)\mathrm{d}u}.$$

In practical applications the deterministic default intensity is given a simple specification, typically (piecewise) constant or (piecewise) linear. Such specifications allow one the match the default probabilities to certain quantities observed in practice. The application of such intensities to the valuation of credit default swaps (introduced in Section 1.3.1) is considered in e.g. [Luo05].

Stochastic Intensity

When we use a stochastic, or random, intensity, the corresponding default process $(N_t)_{t\geq 0}$, is called a Cox process. Conditioned on the evolution of the intensity it is distributed as an inhomogeneous Poisson process. Thus, the distribution of the default time can easily be obtained by conditioning on the realization of the intensity. For $t > 0$ we obtain,

$$\mathbb{P}(\tau \leq t) = 1 - \mathbb{E}\left[1_{\{N_t > 0\}}\right] = 1 - \mathbb{E}\left[e^{-\int_0^t \lambda(u)\mathrm{d}u}\right].$$

Here we observe a close connection to the modeling of stochastic interest rates. Therefore many researchers have been inspired by well known interest rate models, such as the Vasicek model [Vas77], the Ho-Lee model [HL86] and the Cox-Ingersoll-Ross model [CIR85], for the modeling of the intensity. Further [BM06] gives an overview interest rate models, that could also be used to model the default intensity.

Advantages of Reduced Form Modeling

In Section 1.2.1 we have discussed a number of disadvantages of structural models. The two most important disadvantages are the predictability of the default time and the lack of information about a company's value and debt, which are required to estimate model parameters. In this section we show, without going into too much detail, that reduced form models do not suffer from these disadvantages.

To show that default times in reduced form models are not predictable, we observe that $(N_t)_{t \geq 0}$ is a nondecreasing process, and hence it is a sub-martingale with respect to its natural filtration if $\mathbb{E}[N_t] < \infty$ for all t. The Doob-Meyer decomposition, see [LS78], (under suitable conditions) yields that there exists a increasing process A such that $N - A$ is a martingale. This process A is called the compensator. Further, we known from p. 243 in [LS78] that the compensator is continuous if and only if the stopping time, or default time, τ is totally inaccessible. This means that for all predictable stopping times σ we have $\mathbb{P}(\tau = \sigma) = 0$. In the reduced form models we actually directly model the compensator process, which defines the Poisson or Cox process. Clearly, the integrated intensity is a continuous process, which means that the default time τ is not predictable.

As the reduced form models are based on an exogenous definition of the default time, we do not need information about the company's value and debt to estimate model parameters. Instead, we can estimate, or calibrate, the parameters of reduced form models directly from market prices of relevant instruments, such as defaultable bonds or credit default swaps.

1.2.3 Recovery Modeling

In the previous two sections we have considered the modeling of the default time. As we have mentioned above, we also have to model the loss incurred after a default. After a company has defaulted, one could still receive a fraction of the loan or bond notional. In this section we discuss recovery rate models, which model the fraction that is repaid in case of default. We focus on the modeling of the value of a defaultable bond after a default has occurred. However, the results can be applied to loans and other credit derivatives as well.

Deterministic Recovery

When one models the recovery rate as a deterministic quantity there are two different approaches, the fractional recovery of par value and the fractional recovery of market value. The former approach models the loss given default as a fraction of the face value of the bond. The valuation of the bond is then straightforward, as one can easily distinguish the default and no default cases. The bond's payoff at maturity T can be written as

$$D_T = \left(1_{\{\tau > T\}} + 1_{\{\tau \le T\}} R\right) N,$$

where D_T denotes the payoff, $0 \le R \le 1$ the recovery rate and N the face value of the bond. Here we have assumed that in case of default the recovered value of the bond is repaid at maturity. If we assume that the recovered value is paid at the time of default, we can accrue this amount at the risk free interest rate. This yields the payoff at maturity

$$D_T = \left(1_{\{\tau > T\}} + 1_{\{\tau \le T\}} R / B(\tau, T)\right) N,$$

where $B(t, T)$ is a (stochastic) discount factor, i.e. the value at t of receiving 1 at T.

In the fractional recovery of market value, it is assumed that at the time of default a fraction of the pre-default value of the bond is paid. The payoff of the bond at maturity can be written as

$$D_T = \left(1_{\{\tau > T\}} + 1_{\{\tau \le T\}} R D_{\tau^-} / B(\tau, T)\right) N,$$

where D_{τ^-} is the pre-default value of the bond.

These two approaches can easily be extended, for example by using a combination of the two approaches, or by making the recovery rate time dependent. More detailed discussions on the modeling of the deterministic recovery rate can be found in [Sch03a] and [BR02].

Stochastic Recovery

We have seen that when we model the recovery rate deterministically, bond values can easily be evaluated, as one can decompose the payoff into a default and no default scenario. When we model the recovery rate as a random quantity this property does not hold in general. Clearly, we could assume that the recovery rate R is a random variable independent of τ, with distribution F. In such a case the bond can be valued similar to the deterministic models, for example by conditioning on the outcomes of the recovery rate.

Alternatively, one could model the recovery rate as a random quantity, with the first passage time models from Section 1.2.2. In case the value of the company drops below the value of the debt, the company defaults. The recovery value in such a case is $V_\tau < D_\tau$.

For a more detailed discussion of stochastic recovery rates we refer to [Sch03a, Chapter 6]. There, a more general and formal framework is introduced, in which various recovery rate models are considered.

1.2.4 Measuring Portfolio Credit Risk

In general credit portfolios are very large, consisting of thousands of obligors. The measurement of the credit risk in such portfolios focusses on the tails of the credit loss distribution of these portfolios. Often banks are interested in the 95%, 99% or even more extreme percentiles of this distribution. In the previous section we have encountered two approaches to the modeling of the default time of a single company. In some cases it is straightforward to extend this to the modeling of the default times of several companies in a portfolio, which could in theory be used to measure portfolio credit risk. When the size of the portfolio is very large, which is usually the case for banks, such an approach could be very time consuming and therefore less interesting.

In this section we discuss a number of methods that can be used to measure the credit risk in large portfolios. First we briefly mention some popular risk measures, which can be evaluated using techniques that are discussed in the remainder of this section.

Risk Measures

One of the most popular risk measure is the value-at-risk (VaR) measure. The α percent VaR, for some portfolio, states the amount of money that can be lost with α percent certainty. In general VaR can be calculated for any kind of portfolio and any kind of risk type. In all cases one focusses on the loss on a single point in time, usually one year into the future. The VaR measure is discussed in more detail in e.g. [Jor06]. Further, other risk measures have been proposed to measure the risk in a portfolio. In any case one has to model the loss distribution of the portfolio in order to determine certain characteristics, such as the α percent quantile. Many different approaches can be considered for this purpose, and we list a number of these below.

Asymptotic Approximations

A common approach to modeling the loss in a very large credit portfolio is to consider a limiting portfolio, where the size of the portfolio tends to infinity. Under suitable assumptions, such as homogeneity and identical distributed underlyings, one can determine the distribution of the infinite portfolio, and use this as an approximation for the distribution of the finite portfolio.

A good example is the large pool model, where the portfolio is assumed to be homogeneous and of infinite size. The correlation between the different obligors is modeled using the one-factor Gaussian copula, which is discussed in more detail in Section 1.4. The advantage of this model is that one can calculate the loss distribution explicitly in terms of the normal distribution, which in turn could be used to easily determine quantiles of the loss distribution.

Another approach to the modeling of extreme losses are large deviation approximations of the tails of the loss distributions. Large deviations theory, through the large deviation principle, provides asymptotic upper and lower bounds for the logarithm of a family of distributions, such as the probability that the average loss in a portfolio of size n is above a certain level, as $n \to \infty$. Existing large deviation results, see e.g. Dembo and Zeitouni [DZ98], can easily be used to model to tails of the loss distribution at a single point in time. In Chapter 3 we give a brief introduction into the theory of large deviations and we present a large deviation principle for the whole path of the loss process.

Historical Data

A technique that is often used to construct the loss distribution is by using historical data. Here the loss distribution is constructed using a time series of historical data, for example as the empirical distribution function. When one considers the losses due to changes in asset prices, many data are available, and one can build a loss distribution. For the modeling of the credit loss distribution little data are available, as defaults in general are rare. Therefore such data might not suitable to construct the loss distribution. On the other hand, one can use daily market data on market quotes for credit derivatives, which are discussed in detail in Section 1.3, to approximate loss distribution. Such derivatives are used to transfer credit risk, and the associated premia give an indication of the credit risk involved.

Filtering

As an extension of the reduced form models, one can model the loss process as a Cox Process. Each jump of the Cox process represents a default of one of the companies in the portfolio. As is clear from the examples in Section 1.2.2 for the models for the (stochastic) intensities, it is common to model the intensity as

$$\mathrm{d}\lambda_t = \mu(\lambda_t, t)\mathrm{d}t + \sigma(\lambda_t, t)\mathrm{d}W_t, \tag{1.5}$$

for some functions μ, σ, and Brownian motion W_t with respect to some filtration $\mathcal{F}_t$. Following Bielecki and Rutkowski [BR02], one can write $\mathcal{F}_t = \mathcal{F}_t^N \vee \mathcal{G}_t$, where $\mathcal{F}_t^N$ is the filtration generated by the default counting process N, and $\mathcal{G}_t$ is another filtration.

In general, models assume that λ_t (or W_t) can be observed, as λ_t should be $\mathcal{F}_t$ measurable. In practice, one can usually only observe very little information about the a company, and therefore one could question the assumption that λ_t is observable. When we would assume that one can only observe the number of defaults in the portfolio, where the intensity λ_t is given by (1.5), we can use the theory of filtering for point process observations, which is described in e.g. Brémaud [Bré81], to compute relevant quantities with respect to the loss distribution. In Chapter 2 we show that we can derive the conditional moment generating function when we model the intensity as a CIR model. In that chapter, one also finds a brief introduction to filtering with point process observations.

1.3 Credit Derivatives

In this section we describe a number of credit derivatives, which can be used to transfer the default risk from one party to another. The payoff of these derivatives is contingent on the occurrence of a default event. Depending on the definitions of the contract of the credit derivative, the default event could be an unpaid coupon on a specific bond, the bankruptcy of the underlying company or a restructuring of the company.

In a credit derivative contract there are usually three parties involved, the protection buyer, the protection seller and the reference entity. The protection buyer has to pay a premium to the protection seller until maturity or until the default event with respect to the reference entity occurs. In case a default occurs before maturity, the protection seller has to compensate the loss incurred by the protection buyer. This can be done with physical settlement, i.e. the protection buyer delivers defaulted bonds to the protection seller, and receives the notional amount of the bond. Alternatively, the derivative can be settled in cash. In such a case, the loss is typically determined by asking several banks to provide prices for the defaulted bond. In both cases it is clear that the buyer of protection transfers the default risk to the protection seller. The only default risk that the protection buyer faces, is that the protection seller might not be able to cover the loss in case of default. This risk is known as counterparty risk, and it is present in almost any financial contract.

The reference entity can either be a (bond of a) single company, or it can be a basket of companies. In the former case we speak of single name credit derivatives, and in the latter case we speak of multi name credit derivatives. In this section we discuss the most popular credit derivatives traded.

1.3.1 Credit Default Swaps

The credit default swap (CDS) is the most liquid, and simplest credit derivative. It can be seen as an insurance against the default of the referenced company. An investor facing the default risk of a certain party, can enter into a CDS contract, by which he effectively eliminates the default risk with respect to this party. We illustrate the process behind a CDS in a simple example.

Example 1.1. Consider an investor that owns a bond with a notional value of 1000 Euro issued by ING bank that matures in two years, and say that the investor wants to eliminate the default risk with respect to ING bank. He can then enter into a CDS contract maturing after two years with, for example, ABN Amro bank, where the default event is defined with respect to this specific bond. The investor becomes the protection buyer and ABN Amro becomes the protection seller.

During the life of the CDS contract the investor has to make regular, e.g. quarterly, premium payments. These payments are defined as a fraction of the notional of the CDS, say 0.8% or 80 basis points per year. In case ING does not default with respect to this bond, the only payments made are those by the investor. Thus, in this case, eight payments of two Euro are made.

In case ING does default, ABN Amro has to pay the amount lost by the investor. Depending on the specifications of the contract there are two ways to make this payment, either physically or by cash settlement. Say that ING defaults after one year and two months. First assume that the CDS is physically settled. In the first year four payments of 2 Euro are paid by the investor. After the default the investor hands over the bond to ABN Amro and in return he receives 1000 Euro. Further, the investor has to pay a premium of 1.33 Euro, for the two month of protection during the fifth quarter of the contract.

Alternatively, the CDS contract could been cash settled. In such a case a number of dealers is requested to give prices for the defaulted bond. Say that the eventual price is 378 Euro, then ABN Amro has to pay the investor an amount of $1000 - 378 = 622$ Euro. $\diamondsuit$

The market for credit default swaps has become very liquid over the last two decades. This has led to a standardization of this market. Therefore market quotes are available for a wide range of maturities. For many companies, quotes with maturities, or tenors, of three and six months, and one, two, three, four, five, seven, ten, twelve and fifteen years can be found in the market. Further, due to standardization the market for credit default swaps, the time until the contract matures is usually not exactly equal to the periods above. The contracts last for at least this period, but they are extended until the first of four standard maturities, which are the twentieth of March, June, September or December. If, for example, two parties enter into a CDS contract with a tenor of 2 years on September 28, 2009, this contract will end on December 20, 2011, unless a default occurs before this date, which causes an earlier termination of the contract. Naturally, an investor can always try to negotiate different contract specifications, where the maturity can differ from the standard maturities.

1.3.2 Index Credit Default Swap

When an investor is exposed to the default risk of more than one company, he can choose to enter into CDS contracts for all companies to whose default risk it is exposed. If one of the companies defaults, the respective CDS covers the resulting loss, and the investor continues to make premium payments for the remaining CDS contracts. Alternatively the investor could enter into an index credit default swap, or index swap. This swap is similar to the CDS, but instead of a single underlying, a basket of companies is referenced. In case the CDS contracts are cash settlement, and the weights of the companies in the basket are the same as for the separate CDS contracts, the index swap makes the same default payments. The premium payments are different however. The premium is paid with respect to the outstanding, or remaining, notional of the basket. At the start of the contract the premium is agreed which has to be paid. After a default, less premium is paid, since the notional of the basket is reduced. The contract ends at its maturity or when all underlying companies have defaulted. Similar to the regular credit default swaps, the more liquid index credit default swaps are standardized, such that the underlying index consists of the same underlyings and the contracts expire at the standard maturities introduced in the previous section.

In the following section we discuss another credit derivative that can be used to transfer the default risk of a basket of reference entities.

1.3.3 Collateralized Debt Obligations

In the previous section we have seen a credit derivative that can be used to deal with the default risk in a basket of reference entities. The advantage of this is that a protection buyer does not incur losses in case any of these companies defaults. On the other hand, one has to pay premium with respect to the whole basket. Further, an investor could reserve a certain amount of money to cover the first losses in this basket himself, and he wants to receive protection against a certain fraction of the losses. The remaining default risk should then be transferred. The collateralized debt obligation (CDO) can be used to do this.

First, we consider a basket $\mathcal{B}$ of N reference entities. With respect to this basket we define the loss process

$$L_t^{\mathcal{B}} = \sum_{i=1}^{N} 1_{\{\tau_i \leq t\}} (1 - R_i) N_i, \tag{1.6}$$

where τ_i is the default time of reference entity i, R_i is the (fixed) recovery rate of reference entity i and N_i is the notional amount of the reference entity. A CDO tranche provides protection on a part of the losses in this basket. The level where the protection starts, the attachment point, is given as a percentage of the notional of the basket. When the losses in the basket exceed this level the protection buyer receives default payments. When the losses exceed the detachment point of the CDO tranche, and the contract has not matured, the protection buyer no longer receives default payments. One says that the tranche has been exhausted. The protection buyer has to make regular premium payments with respect to the notional amount of the tranche, which is equal to $(d - a)\%$ of the notional of the basket at the start of the contract, where d denotes the detachment point and a the attachment point. The tranche loss, $L_{a,d}^{\mathcal{B}}(t)$, as a fraction of the tranche notional is given by,

$$\begin{aligned} L_{a,d}^{\mathcal{B}}(t) &= \frac{d \cdot L_d^{\mathcal{B}}(t) - a \cdot L_a^{\mathcal{B}}(t)}{d - a}, \\ L_x^{\mathcal{B}}(t) &= \frac{1}{x} \min\left(L_t^{\mathcal{B}}, x\right), \text{ for } x = a, d. \end{aligned} \tag{1.7}$$

In general the basket is split up into a number of consecutive tranches (a_i, d_i), such that $a_i = d_{i-1}$, and $a_1 = 0\%$. The tranche $(a_1, d_1) = (0\%, d_1)$, clearly is the most risky tranche, since the first default of one of the reference entities leads to losses on this tranche. This tranche is usually called the equity tranche. The subsequent tranches are referred to as mezzanine tranches. The last tranche is called the super senior tranche.

Example 1.2. Consider a basket of 100 reference entities, each with a notional of 1000 Euro. Then the total notional of the basket is 100,000 Euro. Suppose that an investor is exposed to the default risk of (a part of) these reference entities, and that he wants to cover losses between 5% and 15% in the basket, for a period of five years. This means that when the loss process exceeds 5,000 Euro the investor starts to receive default payments, until the loss process exceeds 15,000 Euro. Further assume that the premium for this protection is 2% or 200 basis points, and say that the premium is paid quarterly.

When no defaults occur during the first three month, the investor has to pay 50 Euro premium, as the notional amount of the tranche is Euro 10,000. Suppose that in the following three months, four reference entities default, with recovery rates of 10%, 25%, 40% and 60%. The loss process then grows to 2,650 Euro. As this is still below 5,000 Euro, the investor has to pay 50 Euro premium after six months. Assume that until the end of the first year no further defaults occur. Then the investor has to make two more premium payments of 50 Euro.

When three more reference entities default in the first quarter of the next year, with an average recovery rate of 25%, the loss process after the fifteenth month equals Euro 4,900, yielding another premium payment of 50 Euro. If another company defaults, say after one year and four months, with a recovery rate of 50%, the loss process, now at 5,400 Euro, exceeds the attachment point of the CDO tranche, corresponding to 5,000 Euro. The investor thus receives 400 Euro. In return it has to make a premium payment of 48.33 Euro, as the notional of the tranche has been reduced after the default. (Note that over the first month of this quarter still full premium has to be paid).

The premium and default payments are made until the contract matures, or until the loss process exceeds the detachment point of 15%, which corresponds to 15,000 Euro. $\diamondsuit$

Originally, the basket underlying the CDO is constructed of several bonds, loans or other cash instruments. The main problem with this construction is that each bond can have a different maturity, and loans might be repaid early. This requires a manager that replaces bonds and loans if necessary. As the market for CDS contracts has become more liquid, it was realized that the basket of bonds and loans could be replaced by a basket of CDSs, resulting in synthetic CDO tranches. The advantages of using CDSs, is that these all have the same maturity, and that the basket becomes much easier to manage.

The introduction of synthetic CDOs has lead to a large increase in liquidity for CDO tranches, and to the introduction of CDO tranches with respect to standard baskets, or indices. Two of the most popular referenced indices are the iTraxx index, consisting of CDSs on 125 investment grade European companies, and the CDX index, consisting of 125 North-American companies. In the Chapters 4 and 5 these indices and CDO tranches in general are discussed in more detail.

1.3.4 Other Credit Derivatives

In the previous sections we have described some of the most popular credit derivatives. Next to these several other credit derivatives have emerged. Similar to the regular CDS, the constant maturity CDS has been introduced, where the premium payments depend on the market quote of a regular CDS at the beginning of a coupon period.

Another single name credit derivative, is the credit default swaption, which gives the holder the right to enter into a CDS contract at a certain future time against a certain premium. The details of the option specifies if one enters as protection buyer or seller.

Next to the CDO, the N-th to default swap is also a popular multi name credit derivative. In such a contract one is protected against the N-th default that occurs in the underlying basket. Usually this is the first or second default.

A popular variation of the CDO, is the forward starting CDO, where the protection starts at a future time T. Companies that default before this time are excluded from the basket or one can assume that the loss process increases, without triggering default payments.

1.4 Modeling Credit Derivatives

In this section we discuss models for the valuation of credit derivatives. We mainly focus on the credit default swap (CDS), the index swap and the collateralized debt obligation. These are the most liquid instruments. We first focus on the modeling of a single default time, which allows us to value a CDS contract. In addition we can value an index swap contract as well, as this does not depend on the dependence between the different companies in the underlying basket. Thereafter we discuss models that can be used to value (synthetic) CDO tranches.

1.4.1 Valuation of Credit Default Swaps

In Section 1.2 we have encountered two different approaches to the modeling of default times, the structural and reduced form approaches. As both approaches model a single default time, they can easily be used to determine the value of a CDS contract. Assume that the contract has a maturity of T years, and that premium has to be paid at times $T_1 < \cdots < T_N$. Further assume that the percentage of premium, or spread, that has to be paid is denoted by s. Then we can determine the value of the contract as the difference between the expected present value of future premium payment and the expected present value of a possible payment in case of default. If, for simplicity , we assume that the notional amount of the CDS contract equals 1, then the expected value of the premium payments, or premium leg (PL), is given by

$$\begin{aligned} \mathrm{PL} &= \mathbb{E}\left[\sum_{i=1}^{N} \delta_i D\left(T_i\right) 1_{\{\tau > T_i\}}\right] \\ &= \sum_{i=1}^{N} \delta_i D\left(T_i\right) \mathbb{P}\left(\tau > T_i\right), \end{aligned} \tag{1.8}$$

where $D(t)$ is the discount factor, i.e. the present value of receiving 1 unit of currency at time t, τ is the default time of the company underlying the CDS, and δ_i is the year fraction for the period from T_{i-1} to T_i. For simplicity we have ignored the partial premium payment in case of default. The expected amount of premium that will be paid, can be obtained by multiplying by the spread s.

The expected present value of the possible default payment, or default leg (DL), is given by

$$\begin{aligned} \mathrm{DL} &= \mathbb{E}\left[D(\tau) 1_{\{\tau \leq T\}} (1 - R)\right] \\ &\approx \sum_{i=1}^{M} D(t_i^{\star}) \left(\mathbb{P}\left(\tau \leq t_i\right) - \mathbb{P}\left(\tau \leq t_{i-1}\right)\right) \left(1 - R_{\text{fix}}\right). \end{aligned} \tag{1.9}$$

Here R is the recovery rate, the time points t_i form a discretization of the interval $[0, T]$, and $t_i^{\star} = (t_{i-1} + t_i)/2$. It is a common approach to assume a fixed, deterministic recovery rate R_{fix}. Further one usually approximates the expectation by assuming that the default can only occur on a certain time grid $t_1 < \cdots < t_M$, i.e. one computes the integral numerically. In case the recovery rate is modeled as a stochastic quantity (independent of τ) one can calculate the expectation by conditioning on the realization of R.

The value of the CDS contract is obtained as the difference between the two legs. The value for the protection buyer is thus given as DL - $s\cdot$ PL, and the value for the protection seller as $s\cdot$PL - DL. Usually the spread s is chosen such that the contract initially has zero value. Clearly this allows one to calibrate model parameters from market quotes for CDS contracts with different maturities. Especially the levels of the piecewise constant intensity that was discussed in Section 1.2.2, can easily be determined. One can choose the first level of the intensity such that the first CDS quote is matched. Next, one can match subsequent CDS quotes by adjusting the following levels of the intensity. In such a way the intensity is constant between the maturities of the CDS contracts used. This process is often referred to as bootstrapping of the intensity.

From Equations (1.8) and (1.9) it is clear that, in case a deterministic recovery rate is assumed, one only needs to know the distribution of the default time to value a CDS contract. This implies that we can (easily) use all models from Sections 1.2.1 and 1.2.2 to value CDSs.

For the valuation of an index CDS, it is sufficient to add the respective premium and default legs for CDSs on the companies underlying the index CDS, to obtain the premium and default leg of the index CDS.

1.4.2 Models for Collateralized Debt Obligations

Above we have seen that the valuation of CDS contracts is straightforward once the distribution of the default time can be calculated. In this section we look at the valuation of the more complex collateralized debt obligation (CDO). First we look at general pricing formulas. Thereafter we introduce the base correlation framework, which is the most popular model to value CDO tranches. The section is concluded by considering other models that have been proposed for the valuation of CDO tranches.

Pricing Formulas

To value a CDO tranche one has to determine the value of the protection and premium legs. These values depend directly on the cumulative loss process $L_t^{\mathcal{B}}$, as in Equation (1.6). We write a for the attachment point of the tranche and d for the detachment point, and write T for the maturity of the contract. Further we assume that premium payments are made at times $T_1 < \cdots < T_N$. The value of the premium leg, where the tranche loss in calculated as fraction of the tranche notional, is given by

$$\mathrm{PL}_{a,d}^{\mathcal{B},T} = \mathbb{E}\left[\sum_{i=1}^{N} \delta_i D\left(T_i\right) \frac{1}{T_i - T_{i-1}} \int_{T_{i-1}}^{T_i} \left(1 - L_{a,d}^{\mathcal{B}}(t)\right) \mathrm{d}t\right],$$

where the tranche loss is denoted by $L_{a,d}^{\mathcal{B}}(t)$, as defined in (1.7). Instead of calculating the integral above, one usually uses the average of the loss process at the start and the end of the period as an approximation. This yields

$$\mathrm{PL}_{a,d}^{\mathcal{B},T} \approx \sum_{i=1}^{N} \delta_i D\left(T_i\right)\left(1 - \frac{1}{2}\left(\mathbb{E}\left[L_{a,d}^{\mathcal{B}}\left(T_{i-1}\right)\right] + \mathbb{E}\left[L_{a,d}^{\mathcal{B}}\left(T_i\right)\right]\right)\right).$$

The default leg, i.e. the expected present value of the payments in case defaults occur, is given by

$$\begin{aligned}
\mathrm{DL}_{a,d}^{\mathcal{B},T} &= \mathbb{E}\left[\int_0^T D(t)\mathrm{d}L_{a,d}^{\mathcal{B}}(t)\right] \\
&\approx \sum_{i=1}^{K} D\left(t_i^{\star}\right)\left(\mathbb{E}\left[L_{a,d}^{\mathcal{B}}\left(t_i\right)\right] - \mathbb{E}\left[L_{a,d}^{\mathcal{B}}\left(t_{i-1}\right)\right]\right).
\end{aligned}$$

where the latter summation expression is an approximation of the integral.

From the formulas above it is clear that we need to calculate the expected tranche loss at certain points in time in order to determine the value of a CDO tranche. Further one can observe from Equation (1.7) that the tranche loss can be written as the difference between two equity tranches. Therefore we only need to compute expected equity tranche losses, i.e. $\mathbb{E}\left[L_d^{\mathcal{B}}(t)\right]$, for various d and t. In the remainder of this section we consider some popular models for the valuation of CDO tranches.

Copulas

The first type of models we consider are so-called copulas. Such models allow one to model the dependence between random variables, in such a way that the marginal distributions are not affected. If we write $F_1, \ldots, F_N$, for the marginal distributions of a set of random variables $X_1, \ldots, X_N$, then we can calculate the joint distribution of $X_1, \ldots, X_N$ as

$$\mathbb{P}\left(X_1 \leq x_1, \ldots, X_N \leq x_N\right) = C\left(F_1(x_1), \ldots, F_N(x_N)\right),$$

where $C : [0,1]^N \to [0,1]$ such that there exists a collection of uniform variables $U_1, \ldots, U_N$ for which the joint distribution function is given by the function C, and in addition $C(1, \ldots, 1, F_i(x_i), 1, \ldots) = F_i(x_i)$ for all i. The function C is called a copula function. A good introduction to copulas is by [Nel99]. Some well known copulas are the product copula, which results when independence is assumed, and the Gaussian copula, which is based on the multivariate normal distribution.

We consider a special, and intuitively clear, way of creating copulas, through the factor copula approach. We define the set of random variables $X_1, \ldots, X_N$ as

$$X_i = a_i \cdot Y + \sqrt{1 - ||a_i||^2} \cdot \varepsilon_i, \ \ i \leq N \tag{1.10}$$

where $Y \in \mathbb{R}^M$, Y and ε_i are independent, and $||a_i|| \leq 1$ for all $i \leq N$. The factor Y is usually referred to as the common factor, and the ε_i as the idiosyncratic factor. One can interpret the role of Y as a general state of the economy, and the ε_i as the state of company i.

We write F_i for the marginal distribution of company i, which is obtained from CDS market data for this company. In the factor copula setup we assume that the company defaults when for the first time $X_i \leq \chi_i(t)$, for some nondecreasing function $\chi_i(t)$, which is chosen such that for all t

$$F_i(t) = \mathbb{P}\left(\tau_i \leq t\right) = \mathbb{P}\left(X_i \leq \chi_i(t)\right).$$

In this way the marginal distributions are matched.

To calculate the expected tranche losses we assume that recovery rates are deterministic, but possibly different for each underlying. This fixes the size of losses, and what remains is the calculation of the probability that the loss ends up at a certain level, i.e. $\mathbb{P}\left(L^{\mathcal{B}}(t) = x\right)$. The advantage of the factor copula approach is that the variables X_i are independent when we condition on the realization of the common factor Y. This allows us to easily build the conditional distribution $\mathbb{P}\left(L^{\mathcal{B}}(t) = x | Y = y\right)$, as in explained in Chapter 4. The unconditional distribution can then be obtained by integration.

Base Correlation

The most popular factor copula model is the one-factor Gaussian Copula, which was introduced to credit risk modeling by Li [Li00], where both Y and ε_i are standard normally distributed, and one sets $a_i = \rho$ for all i, where $0 \leq \rho \leq 1$ denotes the correlation. Very accurate and fast approximations to the normal distribution allow for a fast valuation of CDO tranches.

The problem with the one-factor Gaussian copula is that one cannot calibrate to market prices with a single correlation value. In a first attempt to deal with this issue, the implied, or compound, correlation was introduced. Here a different correlation is used for each CDO tranche such that the tranche is repriced. This approach is analogous to the calculation of implied volatilities for call and put options in other markets, such as the equity and the foreign exchange markets. The disadvantage of the implied correlation approach is that the correlation depends on the attachment as well as the detachment point, which makes it difficult to interpolate the correlation in order to value a tranche with nonstandard attachment or detachment points, especially since market quotes are available for a small number of tranches.

Above we have observed that both legs of a CDO tranche contract can be valued as the difference of two equity tranches. This observation has led to the introduction of the base correlation framework. In general, market quotes are available for a number of consecutive tranches, and one can iteratively solve for a correlation for equity tranches.

After all base correlations have been calculated on can use interpolation to value any tranche (a, d), using the two correlations ρ_a and ρ_d. A big advantage over the implied correlation is that we now have to interpolate a function of a single variable. As we shall see in Chapter 4, the base correlation curve is increasing, whereas the implied correlation curve is usually not. In Chapter 4, we present a more detailed description of the base correlation framework, and we consider three different ways in which one can interpolate the base correlations, and we discuss the consequences of each interpolation method.

Other Models for CDOs

Due to the disadvantages of the base correlation framework, many different models for the valuation of CDO tranches have been introduced. In this section we briefly discuss a number of these models.

As an extension to the one-factor Gaussian copula, several authors have considered one-factor copulas. Different kinds of distributions are proposed for the common and idiosyncratic factors, or the correlation parameter is assumed to be a random variable. Each of these factor copulas attempts at matching a range of CDO quotes with a single set of parameters. Below, we briefly consider a number of these alternative one-factor copulas, which are discussed in more detail in Chapter 5.

Instead of assuming that the factors Y and ε_i, as in (1.10), have a standard normal distribution, Hull [HW04] considers the Student t-distribution for both factors, to obtain the *Double-*t *factor copula*. The Student t-distribution has much fatter tails compared to the normal distribution.

Similar to the double-t factor copula, Kalemanova et al. [KSW05] have considered the normal inverse Gaussian (NIG) distribution for both factors, which resulted in the *NIG factor copula*. This distribution also shows fatter tails than the normal distribution. Further, in contrast to the normal and student t-distribution, this distribution is not symmetric, which allows for fat left (or right) tails, and thin right (or left) tails.

A different kind of extension has been consider by, amongst others, Burtschell et al. [BGL05]. Here the distribution of the factors Y and ε_i is still the standard normal distribution, but now the correlation parameter ρ is assumed to be a random variable. In particular, it is assumed that the correlation can take on a finite number of values ρ_i, with probabilities p_i. The resulting factor copula model is referred to as the *mixture factor copula*.

The idea of a random correlation has also been adopted by Andersen and Sidenius [AS04], which let the correlation depend on the value of the factor Y, i.e. the state of the economy. The interpretation behind this assumption is that correlations tend to increase when the state of the economy deteriorates. In its simplest form, the correlation can take on two different values, which depends on Y being above or below a certain level. In Chapter 5 we consider a somewhat more advanced version of the *random factor loadings factor copula*.

Next to these four alternative factor copula models, which are described in more detail in Chapter 5, several other versions have been proposed. Some of these are briefly described in this chapter as well.

Besides the factor copula framework, many authors have attempted to match CDO market data with different kinds of models. In the remainder of this section we provide a brief, and far from complete, overview of such models.

The factor copula models that we have considered above model the loss distribution by modeling the individual default processes, and the dependence is introduced through a copula. This approach, where the individual default processes are modeled, are often referred to as bottom-up.

Next to the bottom-up modeling, some authors have modeled the loss process directly, often ignoring the individual characteristics. Such models are usually referred to as top-down models. Giesecke et al. have used this approach in [GGD05] and [EGG09], where the general setup of top-down modeling is introduced, together with more detailed examples. The top-down approach has also been followed by Brigo et al. [BPT06] and [BPT07], where a sum of Poisson models is used to model the loss process from the top down.

In an attempt to introduce a full dynamical model for the valuation of more advanced credit derivatives, such as options on CDO tranches, Schönbucher [Sch06] and Sidenius et al. [SPA06] have considered dynamical models for the evaluation of the loss process, which are inspired on the Heath-Jarrow-Morton modeling approach [HJM92] used in interest rates.

1.5 Credit Crisis

After the market for credit derivatives had grown explosively for almost two decades, this market collapsed when the credit crisis started halfway 2007. Many house owners in the United States could not pay their mortgages, which led to large losses on mortgage portfolios owned by banks or insurance companies. Banks had sold these mortgages in CDO-like, and more complex, structures to other banks or investors. Such structures were initially seen as very save investments, as in general high ratings were given to these structures. Soon after house owners defaulted on their mortgages, it turned out that the structures were in fact bearing high risks. This led to large losses and a lot of uncertainty. Therefore banks had to write off huge amounts of money on these and other types of credit derivatives off their balance sheets, incurring losses of many billions. Following these events, stock markets world wide collapsed in 2008, leading to a financial crisis. Several banks, such as Lehmann Brothers and Bear Stearns defaulted, and the trust between banks almost completely disappeared, making it almost impossible for banks to borrow money from other banks. Many banks received enormous amounts of money from governments in order to survive.

Many people have blamed the mathematical models, and especially the Gaussian copula, for the credit crisis. The Gaussian copula had been widely used to value all sorts of credit derivatives, despite its limitations. When this model was first introduced by Li [Li00], he already stressed that one should be careful applying this model in practice, as the normal distribution is not always a suitable distribution. In practice such warnings were ignored on a large scale, and practitioners invented ways to deal with the flaws of the Gaussian copula, e.g. by the base correlation approach [MBAW04]. Base correlations, however, can only be obtained for a limited number of underlying baskets, which meant that many credit derivatives were valued based on 'the wrong data' obtained with 'the wrong model'. Despite many people being aware of such limitations, the Gaussian copula still became the most popular model because of its simplicity and as it is easy to implement.

The credit crisis has shown that it is very important to correctly model loss process and dependence structures for (large) portfolios which bear credit risk. Furthermore, it became clear that many banks did not reserve enough money to account for credit losses. In this thesis we do not explicitly focus on the credit crisis, as many of the issues related to this crisis were outside the scope of the research underlying this thesis. However, in Chapters 2 and 3 we discuss two alternative ways of modeling the loss process, and analyzing some of its properties, first by modeling the loss process under the assumption of incomplete information, where the process driving the loss process is not observed. Secondly we focus on the asymptotic behavior of the loss process as the size of the portfolio increases. We derive characteristics of the path of the loss process, both as a large deviation principle and, for a special case, as the exact asymptotic behavior. The two modeling approaches address some of the shortcomings of current models, where full information is assumed or where the loss process is only considered at a fixed point in time. In Chapters 4 and 5 we present empirical analysis of pre-crisis CDO data, where the base correlation and other factor copula models are investigated. We show that the way base correlations are interpolated, can have a large influence, and in many cases substantial differences can be observed between the modeled and true values. Further we show that alternative factor copula models can do a reasonable task at fitting to pre-crisis data.

1.6 Outline

The remaining chapters in this thesis can be divided in two parts. The first part, Chapters 2 and 3, deals with two mainly mathematical topics, applied to credit risk. In the second part, Chapters 4 and 5, more practical results are discussed.

In Chapter 2 we model the number of defaults in a large credit portfolio as a Cox process where the intensity follows the square-root, or a Cox-Ingersoll-Ross process. We assume here that we cannot observe the Brownian motion that is driving the intensity process. Using filtering theory for point process observations we are able to derive explicit expressions for the conditional moment generating function of the loss counting process. We first derive partial differential equations for the evolution of the moment generating function between default times, which can be solved explicitly. Further we find the expression for this function at default times. By combining these results, and by assuming that the intensity initially has a gamma distribution, we are able to derive a recursive expression for the moment generating function.

In Chapter 3 we derive a number of asymptotic results for the loss process of a large credit portfolio. We consider a portfolio of n independent and identically distributed obligors, and we model their default time and loss given default separately. Under mild restrictions we derive a large deviation principle for the path of the loss process on a finite time grid. This large deviation principle allows us to calculate exponential bounds on the probability that the path of the loss process is in certain sets. Thereafter we present an (easy-to-check) condition under which the result holds. In the second part of the chapter, we derive the exact asymptotic behavior of the probability of the loss, or the increment of the loss, ever (on a countable time grid) exceeding a certain (time dependent) threshold. It turns out that the asymptotic behavior is completely determined by the 'most likely' time point. At the end of this chapter extensions of the presented results are discussed, where we consider a more diversified portfolio or dependence between the different obligors. Further we discuss how our results change if we consider a continuous time interval or a countable time grid, and the associated complications.

In Chapter 4 we analyze a large data set of market prices for CDO tranches, running from December 2004 up to November 2006, on iTraxx and CDX indices. Next to quotes for CDO tranches with five and ten year maturities, this data set contains CDS data for the underlying companies of the CDO tranches, and interest rates that are used for discounting payments. First a regression analysis is performed on the CDO quotes, to see up to which degree the levels of CDO tranche quotes can be explained by quotes for the index CDS on the same basket. Thereafter we calculate base correlations for all CDO tranche quotes, and we compute the correlation between the base correlations for all tranches, where it is observed that these correlations are in general very high. In the last part of the chapter we compare three different interpolation methodologies for the base correlation skew. We focus on the ability to price nonstandard CDO tranches, by pricing CDO tranches with five year maturities with the ten year skew, and vice versa. Further we price CDO tranches on the iTraxx index with the skew from the CDX index, and vice versa. The test results show that the most advanced interpolation method performs best, but the differences with the actual prices can still be significant.

In Chapter 5 we consider some of the alternative one-factor copula models from Section 1.4.2. We calibrate these models to market quotes from the same set of market data as in Chapter 4. Our main objective is to see how well the models are able to reproduce the market quotes. Further, we compare the performance of the models against each other. It turns out that the performance of the alternative factor copula models is of the same order. We conclude the chapter by investigating the models individually, where they are not only judged by there ability to match market quotes, but also by parameter stability, where unstable parameters can be an indication of over fitting to the data.

CHAPTER 2

A Filtering Problem with Applications to Credit Risk

In the first chapter of this thesis we have given a brief introduction into the modeling of credit risk and credit derivatives. In this chapter, that is based on [LS09], we present a model for the loss process of a credit risky portfolio. We model this loss process as a Cox process, where the default intensity is of the Cox-Ingersoll-Ross type. It is assumed that the Brownian motion driving the intensity process is not observed. With filtering using point process observations, we derive a recursive solution for the conditional moment generating function of the default counting process, given that we only observe the defaults.

2.1 Introduction

The main results of this chapter are explicit, closed form expressions for the solution of a filtering problem with counting process observations, where the unobserved intensity process is a solution to a square root stochastic differential equation. As a matter of fact the explicit solution we provide, is split into a part that concerns the update of the filter at jump times and a part that solves the problem between jump times. This of course reflects the usual strategy for filtering problems with counting observations. The evolution of the filter between jump times is commonly expressed by a partial differential equation for the conditional moment generating function. In general an explicit solution to this PDE is impossible to get, if only were it for the fact, that it can be viewed as an infinite dimensional problem, reflecting that the filter itself is in general infinite dimensional. However for the specific choice of the state process that we have made, we are able to provide explicit solutions. The choice for our specification of the state process is made upon two considerations. First it is known that for conditional Poisson process where the intensity is a random variable having a Gamma distribution, the filtered intensity at a time t also has a Gamma distribution, with parameters depending on t and the value N_t of the counting process. However, in the case that we analyze, the random intensity is also evolving in time, it solves a stochastic differential equation of the Cox-Ingersoll-Ross type, which admits a stationary solution for which all marginal distributions also belong to the Gamma family. By choosing the initial distribution of the intensity properly we are able to come up with explicit expressions for the conditional moment generating function, and we also show that the filtered intensities have distributions that are mixtures of Gamma distributions.

Another reason to study the chosen state process is that it comes up naturally in a simple model for credit risk, which has become a major field of interest in financial mathematics. Indeed, in [Sch03b] the author considers this model for the intensity. Further, [Duf05] considers more general affine models for credit risk. The filtering problem in this set up has previously been studied in [FPR07], where the focus was more on the update for the filter on jump times, whereas we also treat the evolution between jump times in great detail. The filtering problem as such has already been mentioned in [BB80], where the state process was assumed to follow an Ornstein-Uhlenbeck process and the intensity of the counting process was taken to be the square of the state process, which is easily shown to be a CIR process. Although in [BB80] attention has been paid to the evolution of the filter between jump times, an explicit formula for the solution of the resulting PDE has not been given. We obtain this part of the solution analytically by providing a closed form solution to a partial differential equation. Furthermore, we follow a different approach to obtain the recursive solution at jump times as compared to [FPR07]. By combining these solutions, we obtain a solution for all $t > 0$. It is further observed that the resulting conditional moment generating function at time t corresponds to a mixture of $N_t + 1$ Gamma distributions according to some discrete distribution.

Let us give some background for credit risk modeling and explain why filtering is a natural tool in this field of research. The main goal in credit risk is the modeling of the default time of a company or default times of several companies. The default times are often modeled using the so-called intensity-based approach, as opposed to the firm value approach. Here, the default time of a company is modeled as the first jump time of a Cox process, of which the intensity is driven by some stochastic process, e.g. Brownian motion, or, in case of more than one company, as consecutive jump times of this Cox process. This approach enables one to calculate survival probabilities, and to price financial derivatives depending on the default of one or more companies, such as defaultable bonds and credit default swaps. We refrain from a further presentation of these issues as it is beyond the objectives of this chapter and refer to [RMR07] and [Sch03a] for detailed expositions. Overviews of the intensity-based modeling approach can be found in [Lan98], [Gie04] and [Eli05]. In this approach, it is a common assumption that the driving process can be observed, i.e. the observed filtration is generated by the Cox process, which can be seen as the default counting process, and by the driving process.

In this chapter it is assumed that the driving process is *not* observed, and thus only a point process N_t is observed, which introduces a stochastic filtering problem for point processes. In particular the intensity is assumed to follow the Cox-Ingersoll-Ross (CIR) model, where the driving Brownian motion is not observed.

Our results are obtained under the assumption that the CIR process follows a SDE with constant parameters. We briefly discuss what happens if we let the parameters also depend on time. Such a model is more attractive from a practical point, since it allows for more flexible modeling. In general we will then loose the attractive feature of obtaining closed form solutions. But if one restricts the model by taking parameters which are piecewise constant functions, closed form solutions still exist. In practice these piecewise constant models have become popular in credit risk as its flexibility doesn't destroy calibration procedures, see [Luo05].

The chapter is organized as follows: In Section 2.2 the Cox-Ingersoll-Ross model is discussed and some results for the case of full information are discussed. Next, in Section 2.3, the filtering problem is introduced and some background is given for filtering of point process observations. First, the filtering formulas from [Bré81] are given, and the equations for the conditional intensity and conditional moment generating function are derived. Then, in the second part of Section 2.3, we introduce *filtering by the method of the probability of reference*, and the filtering equations are transformed using the ideas introduced in [BB80]. Section 2.4 deals with the filtering problem between the jump times of the point process, given the initial distribution of the intensity at jump times. In Section 2.5, the filtering problem is solved at jump times, and an explicit, recursive solution is obtained, which combines the solutions between and at jumps. Further the resulting conditional moment generating function is analyzed and it is observed that this function agrees with the moment generating function of a mixture of Gamma distributions. The section concludes with an illustration of the mixing probabilities. Finally, in Section 2.6 we discuss possible extensions of the problem under consideration, where the parameters of the SDE for the state process are allowed to be time varying or where the state process is more dimensional.

2.2 Model and Background

The main goal of the chapter is to derive explicit closed form expressions for a filtering problem with counting process observations. Filtering problems with such observations have been studied already some 30 years ago, see e.g. [SDK75], [Sch77] and [BJ77] and the later appearing book [Bré81]. Recently this kind of problem has gained renewed interest in the field of credit risk modeling, see also [FPR07], as we will outline below. One of the main goals in this field is the modeling of the default time of a company or the default times of several companies. Over the years two approaches have become popular, the *structural approach* and the *intensity-based* *approach*. In the structural approach the company value is modeled, for example as a (jump-)diffusion, and the company defaults when its value drops below a certain level. This approach is discussed in more detail in e.g. [Gie04], [BR02] and [Eli05]. In the intensity-based approach the default time is modeled as the first jump of a point process, e.g. a Poisson process or, more general, a Cox process, which is an inhomogeneous Poisson process conditional on the realization of its intensity. In case one considers more than one company, one can model the default times as consecutive jump times of the Cox Process. In [Lan98], [Gie04] and [Eli05] this modeling approach is discussed in more detail, and [Sch02] provides a detailed application. In this chapter we focus on the intensity-based approach, where the intensity λ_t of the Cox process, a nonnegative process, has an affine structure, similar to interest term structure models [DK96]. This means that the intensity process λ_t follows a stochastic differential equation (SDE) of the form:

$$\mathrm{d}\lambda_t = (a + b\lambda_t)\mathrm{d}t + \sqrt{c + d\lambda_t}\,\mathrm{d}W_t, \tag{2.1}$$

for a Brownian Motion W_t, with $d > 0$. In particular, the focus is on the Cox-Ingersoll-Ross square root (CIR) model with mean reversion, [CIR85], for the intensity, where the intensity λ_t satisfies

$$\mathrm{d}\lambda_t = -\alpha(\lambda_t - \mu_0)\mathrm{d}t + \beta\sqrt{\lambda_t}\,\mathrm{d}W_t. \tag{2.2}$$

In [LL96, Section 6.2.2.] one finds parameter restrictions for this model which guarantee positivity of λ_t. Naturally one should start with a positive initial value λ_0, and if $\alpha\mu_0 \geq \beta^2/2$, then λ_t remains strictly positive with probability one. Note that using the transformation $X_t = \lambda_t + c/d$ and by a reparametrization, X_t satisfies the general SDE (2.1), and λ_t satisfies (2.2). This implies that the general form (2.1) and the CIR intensity (2.2) are in fact equivalent. Therefore the CIR intensity will be considered in most of the remainder of this chapter.
A big advantage of the affine setup that we choose in this chapter, is that many relevant quantities in credit risk can be calculated explicitly. Using the formulas from [LL96, Section 6.2.2.] one can, for example, easily calculate the survival probability $\mathbb{P}(\tau > t|\mathcal{F}_s)$, with $t > s$ and $\mathcal{F}_t = \mathcal{F}_t^N \vee \mathcal{F}_t^Y$, where the former filtration is generated by the point process N_t and the latter by some process Y_t driving the intensity process.

Example 2.1. Consider, on the filtered probability space $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\geq 0}, \mathbb{P})$, a random time $\tau > 0$ as the first jump time of a Cox process N_t, which intensity follows the CIR model (2.2). Further assume that $\mathcal{F}_t = \mathcal{F}_t^N \vee \mathcal{F}_t^W$, where $\mathcal{F}_t^W$ is the filtration generated by the Brownian motion that drives the intensity process. Then one can calculate the survival probability for $t > s$ as

$$\mathbb{P}(\tau > t|\mathcal{F}_s) = 1_{\{\tau > s\}}\mathbb{E}\left[e^{-\int_s^t \lambda_u du} \middle| \mathcal{F}_s^W \right], \tag{2.3}$$

which follows from formulas in [BR02, Chapter 6]. Since λ_t is a Markov process, one can condition on λ_s instead of $\mathcal{F}_s^W$. An application of Proposition 6.2.4. from [LL96] to (2.3) yields

$$\mathbb{P}(\tau > t|\mathcal{F}_s) = 1_{\{\tau>s\}} \exp\left(-\alpha\mu_0\varphi(t-s) - \lambda_s\psi(t-s)\right), \tag{2.4}$$

where

$$\varphi(t) = -\frac{2}{\beta^2}\log\left(\frac{2\gamma e^{t(\gamma+\alpha)/2}}{\gamma-\alpha+e^{t\gamma}(\gamma+\alpha)}\right)$$
$$\psi(t) = \frac{2\left(e^{\gamma t}-1\right)}{\gamma-\alpha+e^{t\gamma}(\gamma+\alpha)}$$
$$\gamma = \sqrt{\alpha^2+2\beta^2}.\diamondsuit$$

Other relevant quantities, such as the price of a defaultable bond, can also be calculated analytically, under some restrictions on the interest rate, e.g. by posing that the interest rate evolves deterministically. In [FPR07] some of these quantities are considered in more detail.

It is a common assumption, which is also followed above, that the filtration $\mathcal{F}_t$ is built up using two filtrations, $\mathcal{F}_t^Y$ and $\mathcal{F}_t^N$, where the first filtration represents the information about the process driving the intensity and the second filtration contains information about past defaults. In this chapter it is assumed that the factor Y is not observed which results in a filtering problem of a point process.

In the following sections the problem is introduced formally and solved for the case where the intensity follows the CIR model.

2.3 The Filtering Problem

In filtering theory one deals with the problem of partial observations. Suppose that a process Z_t on the probability space $(\Omega, \mathcal{F}, \mathbb{P})$ is adapted to the filtration $\mathcal{F}_t$. Furthermore let the process Y_t be observed, where Y_t is measurable with respect to a smaller filtration $\mathcal{F}_t^Y \subsetneq \mathcal{F}_t$. One is then interested in conditional expectations of the form $\widehat{Z}_t = \mathbb{E}\left[Z_t|\mathcal{F}_t^Y\right]$, and one tries to find the dynamics of the process $\widehat{Z}_t$, for instance by showing that it is the solution of a stochastic differential equation.

In this section the filtering problem is considered in the case a point process is observed. First some general theory about filtering with point process observations is discussed, and Example 2.1 is continued within the filtering setup. The calculation of the survival probability depends on the conditional moment generating function, for which an SDE is derived. In the second part of this section this equation is transformed in such a way that the filtering problem allows an explicit solution.

Filtering Using Point Process Observations

In the case of point process observations the observed process Y_t is equal to the point process N_t, with jump times T_n, and with $\mathcal{F}_t$-intensity λ_t. The process Z_t is assumed to follow the SDE

$$\mathrm{d}Z_t = a_t \mathrm{d}t + \mathrm{d}M_t, \tag{2.5}$$

for an $\mathcal{F}_t$-progressive measurable a_t, with $\int_0^t |a_s|\mathrm{d}s < \infty$, and an $\mathcal{F}_t$-local martingale M_t. The filtering problem is often cast as the calculation of the conditional expectation $\mathbb{E}[Z_t|\mathcal{F}_t^N] =: \widehat{Z}_t$. Using the filtering formulas from [Bré81, Chapter IV], a representation of the solution to this filtering problem can be found. In case the (local) martingale M_t and the observed point process have no jumps in common, one has:

$$\mathrm{d}\widehat{Z}_t = \widehat{a}_t \mathrm{d}t + \left(\frac{\widehat{Z\lambda}_{t-}}{\widehat{\lambda}_{t-}} - \widehat{Z}_{t-} \right) \left(\mathrm{d}N_t - \widehat{\lambda}_t \mathrm{d}t \right), \tag{2.6}$$

with $\widehat{a}_t := \mathbb{E}[a_t|\mathcal{F}_t^N]$, and $X_{t-} := \lim_{s\uparrow t} X_s$.

Example 2.2 (Example 2.1 continued)**.** When one wants to calculate the survival probability given $\mathcal{F}_t^N$, one has $Z_t = 1_{\{\tau > t\}}$. Combining this with the survival probability in the case of full information, one can calculate the survival probability $\mathbb{P}(\tau > t|\mathcal{F}_s^N)$.

$$\begin{aligned} \mathbb{P}\left(\tau > t|\mathcal{F}_s^N\right) &= \mathbb{E}\left[\mathbb{P}\left(\tau > t|\mathcal{F}_s^N \vee \mathcal{F}_s^W\right) \middle| \mathcal{F}_s^N \right] \\ &= 1_{\{\tau > s\}} \exp\left(-\alpha\mu_0\varphi(t-s)\right) \mathbb{E}\left[\exp(-\psi(t-s)\lambda_s) \middle| \mathcal{F}_s^N \right], \end{aligned}$$

which can be calculated if an expression for the conditional moment generation function $\widehat{f}(s,t) := \mathbb{E}\left[e^{s\lambda_t} \middle| \mathcal{F}_t^N \right]$ is available. ♢

The above example illustrates that one can calculate the survival probability if the conditional moment generating function $\widehat{f}(s,t)$ is known. As a first step in the determination of this function, the SDEs of $\widehat{\lambda}_t := \mathbb{E}\left[\lambda_t \middle| \mathcal{F}_t^N \right]$ and $\widehat{f}(s,t)$ are determined. First Itô's formula is used to obtain the SDE for $e^{s\lambda_t}$, where λ_t satisfies (2.2)

$$\mathrm{d}e^{s\lambda_t} = \left[\left(-\alpha s + \frac{1}{2}s^2\beta^2 \right) \frac{\partial}{\partial s} e^{s\lambda_t} + s\alpha\mu_0 e^{s\lambda_t} \right] \mathrm{d}t + \beta\sqrt{\lambda_t} e^{s\lambda_t} \mathrm{d}W_t. \tag{2.7}$$

The filtered versions are obtained by applying formula (2.6). One obtains for $\widehat{\lambda}_t$

$$\mathrm{d}\widehat{\lambda}_t = -\alpha(\widehat{\lambda}_t - \mu_0)\mathrm{d}t + \left(\frac{\widehat{\lambda^2_{t-}}}{\widehat{\lambda}_{t-}} - \widehat{\lambda}_{t-} \right) \left(\mathrm{d}N_t - \widehat{\lambda}_t \mathrm{d}t \right), \tag{2.8}$$

and for $\widehat{f}(s,t)$ one finds

$$\begin{aligned} \mathrm{d}\widehat{f}(s,t) = &\left[\left(-\alpha s + \frac{1}{2}s^2\beta^2 \right) \frac{\partial}{\partial s} \widehat{f}(s,t) + s\alpha\mu_0 \widehat{f}(s,t) \right] \mathrm{d}t \\ &+ \left(\frac{\frac{\partial}{\partial s}\widehat{f}(s,t-)}{\widehat{\lambda}_{t-}} - \widehat{f}(s,t-) \right) \left(\mathrm{d}N_t - \widehat{\lambda}_t \mathrm{d}t \right). \end{aligned} \tag{2.9}$$

In general, filtering equations are very difficult, if possible at all, to solve explicitly, since the first equation involves terms with $\widehat{\lambda_t^2}$ and the second equation involves combinations of $\widehat{\lambda}_t$ and $\widehat{f}(s,t)$. In order to solve these equations one should also have equations for $\widehat{\lambda_t^2}$, but this involves $\widehat{\lambda_t^3}$ and so on, assuming that they exist. So instead of trying to solve these equations directly, a different approach is considered in order to find an expression for $\widehat{f}(s,t)$.

Filtering by the Method of Probability of Reference

In order to solve the problem introduced above, the *filtering by the method of probability of reference* is considered, see [Bré81, chapter VI] or [BB80, Section 2]. In this approach a second probability measure $\mathbb{P}_0$ and intensity process λ_t^0 are introduced, such that $N_t - \int_0^t \lambda_s^0 ds$ is a martingale with respect to $\mathcal{F}_t$ under $\mathbb{P}_0$. Corresponding to this change of measure one has the likelihood ratio, or density process Λ, given by

$$\Lambda_t := \mathbb{E}\left[\left.\frac{\mathrm{d}\mathbb{P}}{\mathrm{d}\mathbb{P}_0}\right| \mathcal{F}_t\right] = 1 + \int_0^t \Lambda_{s-}\frac{\lambda_{s-} - \lambda_{s-}^0}{\lambda_{s-}^0}\left(\mathrm{d}N_s - \lambda_s^0 \mathrm{d}s\right). \tag{2.10}$$

This likelihood ratio turns out to be a useful tool to solve the filtering problem for $\widehat{f}(s,t)$. It is known, see e.g. [Bré81] for the case $\lambda_t^0 \equiv 1$, that the filtered version of this likelihood ratio, $\widehat{\Lambda}_t := \mathbb{E}\left[\Lambda_t \middle| \mathcal{F}_t^N\right]$ follows an equation similar to (2.10). One has

$$\widehat{\Lambda}_t = 1 + \int_0^t \widehat{\Lambda}_{s-}\frac{\widehat{\lambda}_{s-} - \widehat{\lambda}_{s-}^0}{\widehat{\lambda}_{s-}^0}\left(\mathrm{d}N_s - \widehat{\lambda}_s^0 \mathrm{d}s\right)$$

To solve the filtering problem for $\widehat{f}(s,t)$ an auxiliary function $g(s,t)$ is introduced. It is defined by

$$g(s,t) := \widehat{f}(s,t)\widehat{\Lambda}_t \exp\left(-\int_0^t \widehat{\lambda}_u^0 \mathrm{d}u\right). \tag{2.11}$$

The exponent is used in order to obtain a simpler SDE of $g(s,t)$. After a solution to this equation has been found, one can obtain $\widehat{f}(s,t)$ by

$$\widehat{f}(s,t) = \frac{g(s,t)}{g(0,t)}. \tag{2.12}$$

It is directly clear that the first and third component of $g(s,t)$ are positive, and from (2.14) follows that also the second component is positive, and thus the division in (2.12) is well defined. The solution to the filtering problem is obtained as soon as an expression for $g(s,t)$ is found. In Proposition 2.3 an SDE is derived for $g(s,t)$ for the intensity following the CIR model.

Proposition 2.3. *Let $g(s,t)$ be given by (2.11), then one has, for $t \geq 0$*

$$\begin{aligned} \mathrm{d}g(s,t) = &\left[s\mu_0\alpha g(s,t) + \left(\frac{1}{2}s^2\beta^2 - s\alpha - 1\right)\frac{\partial}{\partial s}g(s,t)\right]\mathrm{d}t \\ &+ \left[\left(\widehat{\lambda}^0_{t-}\right)^{-1}\frac{\partial}{\partial s}g(s,t-) - g(s,t-)\right]\mathrm{d}N_t. \end{aligned} \tag{2.13}$$

Proof. As a first step in proving (2.13), one can rewrite the function $g(s,t)$. Denoting the jump times of N_t as T_n, an alternative expression for $\widehat{\Lambda}_t$ is given by

$$\widehat{\Lambda}_t = \prod_{T_n \leq t}\left(\frac{\widehat{\lambda}_{T_n-}}{\widehat{\lambda}^0_{T_n-}}\right)\exp\left(-\int_0^t\left(\widehat{\lambda}_u - \widehat{\lambda}^0_u\right)\mathrm{d}u\right), \tag{2.14}$$

which can be checked by a direct calculation. From this it is easy to see that

$$\begin{aligned} g(s,t) &\overset{(2.11)}{=} \widehat{f}(s,t)\widehat{\Lambda}_t\exp\left(-\int_0^t\widehat{\lambda}^0_u du\right) \\ &= \widehat{f}(s,t)\prod_{T_n\leq t}\left(\frac{\widehat{\lambda}_{T_n-}}{\widehat{\lambda}^0_{T_n-}}\right)\exp\left(-\int_0^t\widehat{\lambda}_u\mathrm{d}u\right) =: \widehat{f}(s,t)\widehat{L}_t. \end{aligned}$$

For $\widehat{L}_t$ one finds the SDE

$$\mathrm{d}\widehat{L}_t = \frac{\widehat{L}_{t-}\widehat{\lambda}_{t-}}{\widehat{\lambda}^0_{t-}}\left(\mathrm{d}N_t - \widehat{\lambda}^0_t\mathrm{d}t\right) - \widehat{L}_{t-}\mathrm{d}N_t.$$

The SDE in (2.13) follows from the product rule

$$\begin{aligned} \mathrm{d}g(s,t) &= \widehat{f}(s,t\text{-})\mathrm{d}\widehat{L}_t + \widehat{L}_{t-}\mathrm{d}\widehat{f}(s,t) + \Delta\widehat{f}(s,t)\Delta\widehat{L}_t \\ &= \widehat{f}(s,t-)\left(\frac{\widehat{L}_{t-}\widehat{\lambda}_{t-}}{\widehat{\lambda}^0_{t-}}\left(\mathrm{d}N_t - \widehat{\lambda}^0_t\mathrm{d}t\right) - \widehat{L}_{t-}\mathrm{d}N_t\right) \\ &\quad + \widehat{L}_{t-}\left(\left[\left(-\alpha s + \frac{1}{2}s^2\beta^2\right)\frac{\partial}{\partial s}\widehat{f}(s,t) + s\alpha\mu_0\widehat{f}(s,t)\right]\mathrm{d}t\right. \\ &\quad + \left.\left(\frac{\frac{\partial}{\partial s}\widehat{f}(s,t-)}{\widehat{\lambda}t-} - \widehat{f}(s,t-)\right)\left(\mathrm{d}N_t - \widehat{\lambda}_t\mathrm{d}t\right)\right) \\ &\quad + \left(\frac{\frac{\partial}{\partial s}\widehat{f}(s,t-)}{\widehat{\lambda}_{t-}} - \widehat{f}(s,t-)\right)\left(\frac{\widehat{L}_{t-}\widehat{\lambda}_{t-}}{\widehat{\lambda}^0_{t-}} - \widehat{L}_{t-}\right)\mathrm{d}N_t. \end{aligned}$$

Collecting the terms before $\mathrm{d}t$ and $\mathrm{d}N_t$, one obtains the equation

$$\mathrm{d}g(s,t) = \Bigg(-\widehat{\lambda}_t \widehat{f}(s,t)\widehat{L}_t + \left(-\alpha s + \frac{1}{2}s^2\beta^2\right)\frac{\partial}{\partial s}\widehat{f}(s,t)\widehat{L}_t + s\alpha\mu_0 \widehat{f}(s,t)\widehat{L}_t - \frac{\partial}{\partial s}\widehat{f}(s,t)\widehat{L}_t + \widehat{f}(s,t)\widehat{L}_t\widehat{\lambda}_t\Bigg)\mathrm{d}t$$
$$+ \Bigg(\frac{\widehat{f}(s,t-)\widehat{L}_{t-}\widehat{\lambda}_{t-}}{\widehat{\lambda}^0_{t-}} - \widehat{f}(s,t-)\widehat{L}_{t-} + \frac{\widehat{L}_{t-}\frac{\partial}{\partial s}\widehat{f}(s,t-)}{\widehat{\lambda}_{t-}} - \widehat{L}_{t-}\widehat{f}(s,t-) + \frac{\frac{\partial}{\partial s}\widehat{f}(s,t-)\widehat{L}_{t-}}{\widehat{\lambda}^0_{t-}} - \frac{\widehat{f}(s,t-)\widehat{L}_{t-}\widehat{\lambda}_{t-}}{\widehat{\lambda}^0_{t-}} - \frac{\frac{\partial}{\partial s}\widehat{f}(s,t-)\widehat{L}_{t-}}{\widehat{\lambda}_{t-}} + \widehat{f}(s,t-)\widehat{L}_{t-}\Bigg)\mathrm{d}N_t.$$

The result follows by simplifying the last equation. □

The right hand side of (2.13) depends only on $g(s,t)$ and its partial derivative with respect to s. In the next section this equation is solved between jumps, and in section 2.5 the equation is solved at jump times of the process N_t.

2.4 Filtering Between Jumps

In the previous sections the filtering problem for point processes has been defined in general terms, and the problem has further been considered for an intensity following the Cox-Ingersoll-Ross model. To solve the filtering problem, one has to solve equation (2.13). This equation can be split up into a partial differential equation between jumps of the process N_t and an equation at jumps. In this section the equation between jumps is solved for a general initial condition at time $T > 0$. Later on T will be considered as a jump time of N_t. Note that an initial condition for $g(s,t)$ is given as

$$g(s,T) = \widehat{f}(s,T)\widehat{\Lambda}_T \exp\left(-\int_0^T \widehat{\lambda}^0_u \mathrm{d}u\right).$$

For $T = 0$ it follows that

$$g(s,0) = \widehat{f}(s,0) = \mathbb{E}\left[e^{s\lambda_0} \,\middle|\, \mathcal{F}_0^N\right] = \mathbb{E}\left[e^{s\lambda_0}\right],$$

which is the moment generating function of the intensity at time $t = 0$, since the filtration at time $t = 0$, $\mathcal{F}_0^N$, consists of $\{\emptyset, \Omega\}$.

Before the solution to (2.13) is found, the specific case is considered, in which all the parameters in the CIR model are set to zero. Albeit a simple example, the analysis of it sheds some light on the approach that will be followed for the general case.

Example 2.4. Consider the CIR model in which all the parameters are set to zero. This results in a constant intensity, and thus $d\lambda_t = 0$. The filter equations (2.8) and (2.9) reduce to

$$\begin{aligned} \mathrm{d}\widehat{\lambda}_t &= \left(\frac{\widehat{\lambda^2_{t-}}}{\widehat{\lambda}_{t-}} - \widehat{\lambda}_{t-}\right)\left(\mathrm{d}N_t - \widehat{\lambda}_t \mathrm{d}t\right) \\ \mathrm{d}\widehat{f}_t &= \left(\frac{\widehat{\lambda_{t-} f_{t-}}}{\widehat{\lambda}_{t-}} - \widehat{f}_{t-}\right)\left(\mathrm{d}N_t - \widehat{\lambda}_t \mathrm{d}t\right). \end{aligned}$$

The partial differential equation for $g(s,t)$ between jumps reduces to:

$$\frac{\partial}{\partial t} g(s,t) = -\frac{\partial}{\partial s} g(s,t).$$

With an initial condition $g(s,T) = w(s)$, one easily finds that the solution to this equation is

$$g(s,t) = w(s - t + T).$$

◊

In the next section this example is considered once more, where the filter at jump times is considered. We proceed with the case of an intensity following the CIR model.

Proposition 2.5. *Let λ_t follow the Cox-Ingersoll-Ross model (2.2), and let $g(s,t)$ be given by (2.11), with an initial condition at time T, $g(s,T) = w(s)$. Then, for $T \le t < T_n$, with T_n the first jump time of N_t after T, $g(s,t)$ solves the partial differential equation*

$$\frac{\partial}{\partial t} g(s,t) = s\mu_0 \alpha g(s,t) + \frac{1}{2\rho}(\rho s - \alpha + \tau)(\rho s - \alpha - \tau)\frac{\partial}{\partial s} g(s,t), \tag{2.15}$$

where $\rho := \beta^2$ and $\tau := \sqrt{\alpha^2 + 2\beta^2}$. The unique solution to this equation is given by

$$\begin{aligned} g(s,t) = {} & e^{\theta(\alpha-\tau)(t-T)} \left(\frac{2\tau}{\rho s(e^{-\tau(t-T)} - 1) + (\tau - \alpha)e^{-\tau(t-T)} + \tau + \alpha}\right)^{2\theta} \\ & \times w\left(\frac{s\left((\alpha+\tau)e^{-\tau(t-T)} + \tau - \alpha\right) + 2e^{-\tau(t-T)} - 2}{\rho s(e^{-\tau(t-T)} - 1) + (\tau - \alpha)e^{-\tau(t-T)} + \tau + \alpha}\right), \end{aligned} \tag{2.16}$$

where $\theta := \frac{\mu_0 \alpha}{\rho}$.

Proof. The partial differential equation (2.16) for $g(s,t)$ follows directly from Proposition 2.3, since the jump part of this equation can be discarded.
To obtain a solution to this equation a candidate solution is derived by making a

number of transformations of the independent variables, until a simple PDE is found, which can be solved explicitly using known techniques. This candidate solution can then be checked to be the solution by calculating its partial derivatives, and inserting these into (2.15).
The first transformation is given by

$$(s,t) \longrightarrow \left(\frac{\rho s-\alpha+\tau}{\rho s-\alpha-\tau}, t\right) =: (u,t). \tag{2.17}$$

Instead of $g(s,t)$ one writes $f_1(u,s)$, in terms of the new variable u. Using this transformation and the PDE for $g(s,t)$, one can derive a PDE for $f_1(u,t)$, by expressing s in terms of u, and expressing the partial derivatives of $g(s,t)$ as partial derivatives of $f_1(u,t)$. The resulting PDE for $f_1(u,t)$ is

$$\frac{\partial}{\partial t} f_1(u,t) = \mu_0\alpha\left(\frac{\alpha}{\rho} + \frac{\tau(u+1)}{\rho(u+1)}\right) f_1(u,t) - \tau u \frac{\partial}{\partial u} f_1(u,t).$$

The second transformation that is used is given by

$$(u,t) \longrightarrow \left(\frac{\log(u)}{\tau}, t\right) =: (v,t),$$

where, for the time being, u is tacitly understood to be positive. Instead of the function $f_1(s,t)$, one considers the function $f_2(v,t) := f_1(u,t)$, in terms of the new variable v. This transformation results in a partial differential equation for $f_2(v,t)$,

$$\frac{\partial}{\partial t} f_2(v,t) = \mu_0\alpha\left(\frac{\alpha}{\rho} + \frac{\tau(e^{\tau v}+1)}{\rho(e^{\tau v}-1)}\right) f_2(v,t) - \frac{\partial}{\partial v} f_2(v,t).$$

The final transformation is given by

$$f_3(v,t) := \log(f_2(v,t)),$$

which results in the PDE for $f_3(v,t)$:

$$\frac{\partial}{\partial t} f_3(v,t) + \frac{\partial}{\partial v} f_3(v,t) = \mu_0\alpha\left(\frac{\alpha}{\rho} + \frac{\tau(e^{\tau v}+1)}{\rho(e^{\tau v}-1)}\right). \tag{2.18}$$

This equation can be solved using the method of characteristics, which is explained in chapter 1 and 8 of [Che71], for example. Using this technique the partial differential equation is transformed in an ordinary differential equation by introducing new variables $\xi(v,t)$ and $\zeta(v,t)$. The former is used to replace both v and t, and the latter is used to parameterize the initial curve. To be able to solve the PDE an initial condition is required for $f_3(v,t)$. By applying all the previous transformations to the initial condition $g(s,T) = w(s)$, with $t \geq T$, one obtains the initial condition

$$f_3(v,T) = \log\left(w\left(\frac{e^{\tau v}(\tau+\alpha)+\tau-\alpha}{\rho\left(e^{\tau v}-1\right)}\right)\right) =: G(v).$$

Next one has to solve the differential equations

$$\frac{\partial}{\partial \xi} t(\xi,\zeta) = 1, \qquad\qquad \frac{\partial}{\partial \xi} v(\xi,\zeta) = 1,$$

with the initial conditions $t(0,\zeta) = T$ and $v(0,\zeta) = \zeta$. The unique solution to these equations is trivially given by

$$t(\xi,\zeta) = \xi + T, \qquad\qquad v(\xi,\zeta) = \xi + \zeta.$$

Inverting these expressions, yields

$$\xi(v,t) = t - T, \qquad\qquad \zeta(v,t) = v - t + T.$$

Using these transformations, the partial differential equation (2.18) can be transformed into the ordinary differential equation (ODE)

$$\begin{aligned}\frac{\partial}{\partial \xi} f_3(\xi,\zeta) &= \mu_0\alpha\left(\frac{\alpha}{\rho} + \frac{\tau(e^{\tau(\xi+\zeta)}+1)}{\rho(e^{\tau(\xi+\zeta)}-1)}\right)\\ &= \frac{\mu_0\alpha(\alpha+\tau)}{\rho} + \frac{2\tau\mu_0\alpha}{\rho(e^{\tau(\xi+\zeta)}-1)} = \theta(\alpha+\tau) + \frac{2\tau\theta}{e^{\tau(\xi+\zeta)}-1},\end{aligned} \tag{2.19}$$

where $\theta = \frac{\mu_0\alpha}{\rho}$. This ordinary differential equation can be solved for the given initial condition $f_3(v,T) = G(v)$. To derive the solution one can start with a candidate solution

$$f_3(\xi,\zeta) = C_1 \log\left(e^{\tau(\xi+\zeta)} - 1\right) + C_2\xi + C_3.$$

For $\xi = 0$, one has $f_3(0,\zeta) = C_1 \log\left(e^{\tau\zeta} - 1\right) + C_3$, and f_3 has partial derivative with respect to ξ:

$$\frac{\partial}{\partial \xi} f_3(\xi,\zeta) = \tau C_1 + \frac{C_1\tau}{e^{\tau(\xi+\zeta)}-1} + C_2.$$

Using the initial condition $f_3(0,\zeta) = G(\zeta)$, together with the ODE (2.19), one can find the values of $C_1,\ C_2$ and C_3:

$$\begin{aligned}C_1 &= 2\theta,\\ C_2 &= \theta(\alpha-\tau),\\ C_3 &= G(\zeta) - 2\theta\log\left(e^{\tau\zeta}-1\right).\end{aligned}$$

This leads to the unique solution

$$f_3(\xi,\zeta) = \theta(\alpha-\tau)\xi + 2\theta\log\left(e^{\tau(\xi+\zeta)}-1\right) + G(\zeta) - 2\theta\log\left(e^{\tau\zeta}-1\right). \tag{2.20}$$

The proof of the uniqueness of this solution is postponed to the end of this proof.

Replacing ξ by $t-T$ and ζ by $v-t+T$, results in

$$f_3(v,t) = \theta(\alpha-\tau)(t-T) + 2\theta\log\left(\frac{e^{\tau v}-1}{e^{\tau(v-t+T)}-1}\right) + \log\left(w\left(\frac{e^{\tau(v-t+T)}(\tau+\alpha)+\tau-\alpha}{\rho\left(e^{\tau(v-t+T)}-1\right)}\right)\right). \tag{2.21}$$

Next, one obtains a candidate solution for $g(s,t)$, by reversing all the transformations. This gives

$$f_2(s,t) = e^{\theta(\alpha-\tau)(t-T)}\left(\frac{e^{\tau v}-1}{e^{\tau(v-t+T)}-1}\right)^{2\theta} w\left(\frac{e^{\tau(v-t+T)}(\tau+\alpha)+\tau-\alpha}{\rho\left(e^{\tau(v-t+T)}-1\right)}\right),$$

$$f_1(s,t) = e^{\theta(\alpha-\tau)(t-T)}\left(\frac{u-1}{ue^{-\tau(t-T)}-1}\right)^{2\theta} w\left(\frac{ue^{-\tau(t-T)}(\tau+\alpha)+\tau-\alpha}{\rho\left(ue^{-\tau(t-T)}-1\right)}\right).$$

By performing the last substitution, (2.17), an expression for $g(s,t)$ is obtained. One has

$$\begin{aligned} g(s,t) &= e^{\theta(\alpha-\tau)(t-T)}\left(\frac{\frac{\rho s-\alpha+\tau}{\rho s-\alpha-\tau}-1}{\frac{\rho s-\alpha+\tau}{\rho s-\alpha-\tau}e^{-\tau(t-T)}-1}\right)^{2\theta} \\ &\quad\times w\left(\frac{\frac{\rho s-\alpha+\tau}{\rho s-\alpha-\tau}e^{-\tau(t-T)}(\tau+\alpha)+\tau-\alpha}{\rho\left(\frac{\rho s-\alpha+\tau}{\rho s-\alpha-\tau}e^{-\tau(t-T)}-1\right)}\right) \\ &= e^{\theta(\alpha-\tau)(t-T)}\left(\frac{2\tau}{\rho s(e^{-\tau(t-T)}-1)+(\tau-\alpha)e^{-\tau(t-T)}+\tau+\alpha}\right)^{2\theta} \\ &\quad\times w\left(\frac{s\left((\alpha+\tau)e^{-\tau(t-T)}+\tau-\alpha\right)+2e^{-\tau(t-T)}-2}{\rho s(e^{-\tau(t-T)}-1)+(\tau-\alpha)e^{-\tau(t-T)}+\tau+\alpha}\right), \end{aligned}$$

where it was used that $(\alpha+\tau)(\tau-\alpha) = 2\rho$. By inserting this candidate into equation (2.15), one can check that it indeed is the solution.

The last thing to prove is the uniqueness of the solution to equation (2.15). As all the transformations are clearly one-to-one, the uniqueness of this solution should follow from the uniqueness of the solution to equation (2.19). It is easy to see that the solution to this equation is unique, as the difference of two possible solutions, with the same initial condition, has zero derivative, which implies that the two solutions are in fact equal. □

The result of Proposition 2.5 tells us that one can calculate $g(s,t)$, for $T \le t < T_n$, where T_n is the first jump time of N_t after T. In order to completely solve the filtering problem, one further has to solve the equation (2.13) at jump times. This is the topic of the next section, where a recursive solution will be obtained for the case in which λ_0 has a Gamma distribution.

2.5 Filtering at Jump Times and a General Solution

In the previous section the filtering problem has been solved between jumps, for an arbitrary initial condition $w(s)$ for $g(s,t)$, at time $T > 0$. In this section the filtering problem is solved at jump times, first for Example 2.4, and after that for the case where the intensity follows the CIR model.

Example 2.6 (Example 2.4 (continued))**.** At jumps one obtains from equation (2.13)

$$\Delta g(s,t) = \left(\frac{\frac{\partial}{\partial s} g(s,t-)}{\widehat{\lambda}^0_{t-}} - g(s,t-) \right) \Delta N_t.$$

From this identity it easily follows that at a jump time $T > 0$:

$$g(s,T) = \left(\widehat{\lambda}^0_{T-}\right)^{-1} \frac{\partial}{\partial s} g(s,T-). \tag{2.22}$$

Combining the results between jumps and at jumps, one can obtain the solution to equation

$$\mathrm{d}g(s,t) = -\frac{\partial}{\partial s} g(s,t)\mathrm{d}t + \left(\frac{\frac{\partial}{\partial s} g(s,t-)}{\widehat{\lambda}^0_{t-}} - g(s,t-) \right) \mathrm{d}N_t.$$

At each jump time T_n, one has to take the derivative of the function $g(s,t)$, and divide by $\lambda^0_{T_n-}$; the resulting function can then be used as initial condition for the interval $[T_n, T_{n+1})$. Using an initial condition $g(s,0) = w(s)$, one obtains the solution

$$g(s,t) = w^{(N_t)}(s-t) \prod_{n=1}^{N_t} \left(\widehat{\lambda}^0_{T_n-}\right)^{-1},$$

where $w^{(n)}(s)$ denotes the n-th derivative of $w(s)$. The conditional moment generating function is found from (2.12), and is given by

$$\widehat{f}(s,t) = \frac{g(s,t)}{g(0,t)} = \frac{w^{(N_t)}(s-t)}{w^{(N_t)}(-t)}.$$

If one assumes that $\lambda_0 \sim \Gamma(\alpha,\beta)$, one has

$$f(s,0) = \widehat{f}(s,0) = \left(\frac{\beta}{\beta - s}\right)^{\alpha}, \qquad \widehat{f}(s,t) = \left(\frac{\beta+t}{\beta+t-s}\right)^{\alpha+N_t}. \tag{2.23}$$

From this follows that at time $t > 0$, λ_t given $\mathcal{F}^N_t$ is distributed as $\Gamma(\alpha + N_t, \beta + t)$. Further $\widehat{\lambda}_t$ can easily be derived by a differentiation with respect to s:

$$\widehat{\lambda}_t = \frac{\partial}{\partial s} \widehat{f}(s,t)\bigg|_{s=0} = \frac{\alpha + N_t}{\beta + t}.$$

◊

The solution in this example was easy to find, which could be expected, since λ_t is constant over time in this case. The general Cox-Ingersoll-Ross model for the intensity is more complicated, but in the remainder of this section, also this problem is solved. At jumps one has the same equation as in Example 2.6, which is already solved in (2.22). In Theorem 2.7 the solution for $g(s,t)$ for the CIR model is given. Before this theorem is stated some notation is introduced.
Let $x, y \in \mathbb{R}$, $t \geq 0$ and put

$$A(x,t,y) := x\left((\tau-\alpha)e^{-\tau t} + \tau + \alpha\right) + 2y\left(1 - e^{-\tau t}\right) \tag{2.24}$$
$$B(s,t) := \rho s\left(e^{-\tau t} - 1\right) + (\tau-\alpha)e^{-\tau t} + \tau + \alpha \tag{2.25}$$
$$C(x,t,y) := y\left((\alpha+\tau)e^{-\tau t} + \tau - \alpha\right) + \rho x\left(1 - e^{-\tau t}\right). \tag{2.26}$$

This notation allows us to write the general solution between jumps, (2.16), as

$$g(s,t) = e^{\theta(\alpha-\tau)(t-T)}\left(\frac{2\tau}{B(s,t-T)}\right)^{2\theta} w\left(\frac{C\left(-\frac{2}{\rho}, t-T, s\right)}{B(s,t-T)}\right) \tag{2.27}$$

Next let $T_1, T_2, \ldots$ denote the jump times, and let $T_0 = 0$. Then introduce the following notation:

$$\mathscr{A}(t,T_0) := A(\varphi,t,1) \qquad \text{for } 0 \leq t < T_1, \tag{2.28}$$
$$\mathscr{A}(t,T_n) := A(\mathscr{A}(T_n,T_{n-1}), t-T_n, \mathscr{C}(T_n,T_{n-1})) \qquad \text{for } T_n \leq t < T_{n+1}, \tag{2.29}$$
$$\mathscr{C}(t,T_0) := C(\varphi,t,1) \qquad \text{for } 0 \leq t < T_1, \tag{2.30}$$
$$\mathscr{C}(t,T_n) := C(\mathscr{A}(T_n,T_{n-1}), t-T_n, \mathscr{C}(T_n,T_{n-1})) \qquad \text{for } T_n \leq t < T_{n+1}. \tag{2.31}$$

With this notation, the main result of this chapter can be stated. A recursive solution to the filtering problem is obtained, for the case where λ_0 has a Gamma distribution.

Theorem 2.7. *Let $\lambda_0 \sim \Gamma(2\theta, \varphi)$, for $\varphi > 0$ and $\theta = \frac{\mu_0 \alpha}{\rho} > 0$. Then one has*

$$\widehat{f}_0(s) = g(s,0) = \left(\frac{\varphi}{\varphi - s}\right)^{2\theta},$$

which is the moment generating function of the $\Gamma(2\theta,\varphi)$ distribution. With the notation introduced in (2.24)-(2.26) and (2.28)-(2.31) one further has, for $T_n \leq t < T_{n+1}$,

$$g(s,t) = K(t)p_n(s,t)\left(\frac{1}{\mathscr{A}(t,T_n) - s\mathscr{C}(t,T_n)}\right)^{2\theta+n}, \tag{2.32}$$

where $p_0(s,t) \equiv 1$, and for $n \geq 1$, $p_n(s,t)$ is a polynomial of degree n in s, that satisfies the recursion,

$$\begin{aligned} p_n(s,t) = B^n(s,t-T_n)\Bigg[&p_{n-1}\left(\frac{C\left(-\frac{2}{\rho}, t-T_n, s\right)}{B(s,t-T_n)}, T_n\right)(2\theta+n-1)\mathscr{C}(T_n,T_{n-1}) \\ &+ \partial_1\left(p_{n-1}\left(\frac{C\left(-\frac{2}{\rho}, t-T_n, s\right)}{B(s,t-T_n)}, T_n\right)\right) \\ &\times\left(\mathscr{A}(T_n,T_{n-1}) - \frac{C\left(-\frac{2}{\rho}, t-T_n, s\right)}{B(s,t-T_n)}\mathscr{C}(T_n,T_{n-1})\right)\Bigg], \end{aligned} \tag{2.33}$$

where ∂_1 denotes the derivative with respect to the first argument of p_n, and

$$K(t) = e^{\theta(\alpha-\tau)t}(2\tau\varphi)^{2\theta} \prod_{m\geq 1,\ T_m \leq t} \left(\frac{(2\tau)^{2\theta}}{\widehat{\lambda}^0_{T_m-}}\right). \tag{2.34}$$

In the proof of this theorem the following lemma is used.

Lemma 2.8. *With the notation from (2.24)-(2.26) and (2.28)-(2.31), the following relations hold for $n \geq 1$ and $x, y \in \mathbb{R}$:*

(i) $\mathscr{A}(T_n,T_n) = 2\tau\mathscr{A}(T_n,T_{n-1})$

(ii) $\mathscr{C}(T_n,T_n) = 2\tau\mathscr{C}(T_n,T_{n-1})$

(iii) $xB(s,t) - yC\left(-\frac{2}{\rho}, t, s\right) = A(x,t,y) - sC(x,t,y)$.

Proof. (i) From equations (2.29) and (2.24) follows that

$$\begin{aligned} \mathscr{A}(T_n,T_n) &= A(\mathscr{A}(T_n,T_{n-1}), 0, \mathscr{C}(T_n,T_{n-1})) \\ &= \mathscr{A}(T_n,T_{n-1})\left((\tau-\alpha)e^0 + \tau + \alpha\right) + \mathscr{C}(T_n,T_{n-1})\left(1-e^0\right) \\ &= 2\tau\mathscr{A}(T_n,T_{n-1}). \end{aligned}$$

(ii) This follows along the same lines as in (i), using equations (2.31) and (2.26).

(iii) Using Equations (2.25) and (2.26) one finds:

$$\begin{aligned} xB(s,t) - yC\left(-\frac{2}{\rho}, t, s\right) &= x\left(\rho s\left(e^{-\tau t}-1\right) + (\tau-\alpha)e^{-\tau t} + \tau + \alpha\right) \\ &\quad - y\left(s\left((\alpha+\tau)e^{-\tau t} + \tau - \alpha\right) + 2\left(1-e^{-\tau t}\right)\right). \end{aligned}$$

Rearranging terms, and using Equations (2.24) and (2.26), we find

$$
\begin{aligned}
xB(s,t) - yC\left(-\frac{2}{\rho}, t, s\right) &= x\left((\tau-\alpha)e^{-\tau t} + \tau + \alpha\right) + 2y\left(1 - e^{-\tau t}\right) \\
&\quad - s\left(y\left((\alpha+\tau)e^{-\tau t} + \tau - \alpha\right) + x\rho\left(1 - e^{-\tau t}\right)\right) \\
&= A(x,t,y) - sC(x,t,y),
\end{aligned}
$$

which establishes the result.

□

Now, Theorem 2.7 can be proved.

Proof of Theorem 2.7. For each n it has to be shown that (2.32) holds at T_n, and between T_n and T_{n+1}. First this is shown for $n = 0$. Then the induction step is proved for $n \geq 1$.
$\mathbf{n = 0}$: For $t = T_0 = 0$ one has by assumption:

$$
g(s,0) = \left(\frac{\varphi}{\varphi - s}\right)^{2\theta}.
$$

From (2.32) one finds:

$$
\begin{aligned}
g(s,0) &= K(0)p_0(s,0)\left(\frac{1}{\mathscr{A}(0,0) - s\mathscr{C}(0,0)}\right)^{2\theta} \\
&= e^0(2\tau\varphi)^{2\theta}\left(\frac{1}{A(\varphi,0,1) - sC(\varphi,0,1)}\right)^{2\theta} \\
&= \left(\frac{2\tau\varphi}{2\tau\varphi - 2\tau s}\right)^{2\theta} = \left(\frac{\varphi}{\varphi - s}\right)^{2\theta}.
\end{aligned}
$$

Next the interval up to the first jump time, $0 < t < T_1$, is considered. From (2.27) and the expression for $w(s) = g(s,0)$, one finds:

$$
\begin{aligned}
g(s,t) &= e^{\theta(\alpha-\tau)t}\left(\frac{2\tau}{B(s,t)}\right)^{2\theta}\left(\frac{\varphi}{\varphi - \frac{C\left(-\frac{2}{\rho},t,s\right)}{B(s,t)}}\right)^{2\theta} \\
&= e^{\theta(\alpha-\tau)t}(2\tau\varphi)^{2\theta}\left(\frac{1}{B(s,t)\varphi - C\left(-\frac{2}{\rho},t,s\right)}\right)^{2\theta} \\
&= K(t)p_0(s,t)\left(\frac{1}{\mathscr{A}(t,0) - s\mathscr{C}(t,0)}\right)^{2\theta},
\end{aligned}
$$

which is the same expression as in (2.32) for $n = 0$. The final step in the derivation above follows from Lemma 2.8 (iii), with $x = \varphi$ and $y = 1$, together with the definition of $K(t)$ in (2.34).

$\mathbf{n \geq 1}$: Now it remains to prove the induction step. Therefore one can assume that equation (2.32) holds for $n-1$. It then remains to show that the equation holds for n, at T_n and between T_n and T_{n+1}. First the jump is considered. Thus one has to calculate the derivative of $g(s,t)$ with respect to s, and take the left limit in $t = T_n$, further the derivative is divided by $\widehat{\lambda}^0_{T_n-}$. By (2.22) one has

$$\begin{aligned} g(s,T_n) &= \left(\widehat{\lambda}^0_{T_n-}\right)^{-1} \frac{\partial}{\partial s} g(s,T_n-) \\ &= \left(\widehat{\lambda}^0_{T_n-}\right)^{-1} \frac{\partial}{\partial s}\Bigg(K(T_n-) p_{n-1}(s,T_n-) \\ &\quad \times \left(\frac{1}{\mathscr{A}(T_n,T_{n-1}) - s\mathscr{C}(T_n,T_{n-1})}\right)^{2\theta+n-1}\Bigg). \end{aligned} \tag{2.35}$$

Calculating the derivative with respect to s, leads to

$$\begin{aligned} g(s,t) = \left(\widehat{\lambda}^0_{T_n-}\right)^{-1} K(T_n-)\Bigg[& p_{n-1}(s,T_n)(2\theta+n-1)\mathscr{C}(T_n,T_{n-1}) \\ &+ \frac{\partial}{\partial s} p_{n-1}(s,T_n)\left(\mathscr{A}(T_n,T_{n-1}) - s\mathscr{C}(T_n,T_{n-1})\right)\Bigg] \\ &\times \left(\frac{1}{\mathscr{A}(T_n,T_{n-1}) - s\mathscr{C}(T_n,T_{n-1})}\right)^{2\theta+n}. \end{aligned} \tag{2.36}$$

From Lemma 2.8 (i) and (ii) follows that for the denominator in (2.36) one has

$$\mathscr{A}(T_n,T_{n-1}) - s\mathscr{C}(T_n,T_{n-1}) = (2\tau)^{-1}\left(\mathscr{A}(T_n,T_n) - s\mathscr{C}(T_n,T_n)\right).$$

Hence (2.36) can be written as

$$\begin{aligned} g(s,T_n) = \left(\widehat{\lambda}^0_{T_n-}\right)^{-1} K(T_n-)(2\tau)^{2\theta}(2\tau)^n \Bigg[& p_{n-1}(s,T_n)(2\theta+n-1)\mathscr{C}(T_n,T_{n-1}) \\ &+ \frac{\partial}{\partial s} p_{n-1}(s,T_n)\left(\mathscr{A}(T_n,T_{n-1}) - s\mathscr{C}(T_n,T_{n-1})\right)\Bigg] \\ &\times \left(\frac{1}{\mathscr{A}(T_n,T_n) - s\mathscr{C}(T_n,T_n)}\right)^{2\theta+n}. \end{aligned} \tag{2.37}$$

From (2.34) it is easy to see that $K(T_n) = K(T_n-)\left(\widehat{\lambda}^0_{T_n-}\right)^{-1}(2\tau)^{2\theta}$, and further one has $2\tau = B(s,0) = B(s,T_n - T_n)$. From this follows that (2.37) can be written as

$$
\begin{aligned}
g(s,T_n) = K(T_n)B^n(s,T_n - T_n)\Bigg[&p_{n-1}(s,T_n)(2\theta+n-1)\mathscr{C}(T_n,T_{n-1}) \\
&+ \frac{\partial}{\partial s}p_{n-1}(s,T_n)\left(\mathscr{A}(T_n,T_{n-1}) - s\mathscr{C}(T_n,T_{n-1})\right)\Bigg] \\
&\times\left(\frac{1}{\mathscr{A}(T_n,T_n) - s\mathscr{C}(T_n,T_n)}\right)^{2\theta+n}.
\end{aligned}
$$

This can be simplified further using the definition of $p_n(s,t)$ as given in (2.33), together with the identity $C\left(-\frac{2}{\rho},0,s\right) = \tau s$. This results in

$$
g(s,T_n) = K(T_n)p_n(s,T_n)\left(\frac{1}{\mathscr{A}(T_n,T_n) - s\mathscr{C}(T_n,T_n)}\right)^{2\theta+n},
$$

which is the required result at $t = T_n$. Finally one has to check that (2.32) holds for $T_n < t < T_{n+1}$. For this one can use the general solution (2.27) with initial condition $w(s) = g(s,T_n)$. One finds

$$
\begin{aligned}
g(s,t) = e^{\theta(\alpha-\tau)(t-T_n)}&\left(\frac{2\tau}{B(s,t-T_n)}\right)^{2\theta}e^{\theta(\alpha-\tau)T_n}(2\tau\varphi)^{2\theta} \\
&\times\prod_{m\geq 1,\ T_m\leq T_n}\left(\frac{(2\tau)^{2\theta}}{\widehat{\lambda}^0_{T_m-}}\right)p_n\left(\frac{C\left(-\frac{2}{\rho},t-T_n,s\right)}{B(s,t-T_n)},T_n\right) \\
&\times\left(\frac{1}{\mathscr{A}(T_n,T_n) - \frac{C\left(-\frac{2}{\rho},t-T_n,s\right)}{B(s,t-T_n)}\mathscr{C}(T_n,T_n)}\right)^{2\theta+n}.
\end{aligned}
$$

Simplifying this expression yields:

$$
\begin{aligned}
g(s,t) = e^{\theta(\alpha-\tau)t}(2\tau\varphi)^{2\theta}&\left(\prod_{m=1}^{n}\left(\frac{(2\tau)^{2\theta}}{\widehat{\lambda}^0_{T_m-}}\right)\right)(2\tau)^{2\theta} \\
&\times B^n(s,t-T_n)p_n\left(\frac{C\left(-\frac{2}{\rho},t-T_n,s\right)}{B(s,t-T_n)},T_n\right) \\
&\times\left(\frac{1}{2\tau B(s,t-T_n)\mathscr{A}(T_n,T_{n-1}) - 2\tau C\left(-\frac{2}{\rho},t-T_n,s\right)\mathscr{C}(T_n,T_{n-1})}\right)^{2\theta+n}.
\end{aligned}
$$

An application of Lemma 2.8, with $x = \mathscr{A}(T_n, T_{n-1})$ and $y = \mathscr{C}(T_n, T_{n-1})$, and the definitions of $\mathscr{A}(t, T_n)$ and $\mathscr{C}(t, T_n)$ in (2.29) and (2.31), together with the definition of $K(t)$ results in

$$g(s,t) = K(t)\frac{1}{(2\tau)^n}B^n(s,t-T_n)p_n\left(\frac{C\left(-\frac{2}{\rho},t-T_n,s\right)}{B(s,t-T_n)},T_n\right) \times \left(\frac{1}{\mathscr{A}(t,T_n) - s\mathscr{C}(t,T_n)}\right)^{2\theta+n}.$$

Next, with the definition of $p_n(s, T_n)$ from (2.33), evaluated in $t = T_n$, together with $C(x, 0, y) = 2\tau y$ and $B(s, 0) = 2\tau$ one rewrites this to

$$\begin{aligned} g(s,t) = K(t)&\left(\frac{1}{\mathscr{A}(t,T_n) - s\mathscr{C}(t,T_n)}\right)^{2\theta+n}\frac{1}{(2\tau)^n}B^n(s,t-T_n) \\ &\times (2\tau)^n\Bigg[p_{n-1}\left(\frac{C\left(-\frac{2}{\rho},t-T_n,s\right)}{B(s,t-T_n)},T_n\right)(2\theta+n-1)\mathscr{C}(T_n,T_{n-1}) \\ &+ \partial_1\left(p_{n-1}\left(\frac{C\left(-\frac{2}{\rho},t-T_n,s\right)}{B(s,t-T_n)},T_n\right)\right) \\ &\times\left(\mathscr{A}(T_n,T_{n-1}) - \frac{C\left(-\frac{2}{\rho},t-T_n,s\right)}{B(s,t-T_n)}\mathscr{C}(T_n,T_{n-1})\right)\Bigg] \\ = K(t)&p_n(s,t)\left(\frac{1}{\mathscr{A}(t,T_n) - s\mathscr{C}(t,T_n)}\right)^{2\theta+n}. \end{aligned}$$

In the final step the definition of $p_n(s, t)$ is used, this time evaluated in t, which concludes the proof of (2.32). From the definition of $B(s, t)$ and $C(x, t, y)$, with $y = s$, which are both linear in s, follows that $p_n(s, t)$ is a polynomial of degree n in s. □

This theorem provides a recursive solution to equation (2.13), in case λ_0 is distributed according to a $\Gamma(2\theta, \varphi)$ distribution. From (2.12) it already known that the conditional moment generating function can easily be obtained from an expression for $g(s, t)$. Now this has been found, the conditional moment generating function $\widehat{f}(s, t)$ can be obtained easily.

Corollary 2.9. *Under the assumptions of Theorem 2.7 the conditional moment generating function $\widehat{f}(s, t)$, for $T_n \le t < T_{n+1}$, can be expressed as:*

$$\widehat{f}(s,t) = q_n(s,t)\left(\frac{Q(t,T_n)}{Q(t,T_n) - s}\right)^{2\theta+n}, \tag{2.38}$$

where

$$q_n(s,t) = \frac{p_n(s,t)}{p_n(0,t)} \text{ and } Q(t,T_n) = \frac{\mathscr{A}(t,T_n)}{\mathscr{C}(t,T_n)}.$$

Here $q_n(s,t)$ is a polynomial of degree n in s.

Proof. The result follows directly from equation (2.12), Theorem 2.7 and the definitions of q_n and Q. □

With the derivation of the conditional moment generating function the filtering problem has been solved, and one is able to calculate conditional default probabilities using the results in Example 2.2. To conclude this section it is observed that the conditional moment generating function in (2.38) corresponds to a mixture of Gamma distributions.

Remark 2.10. Corollary 2.9 provides an expression for $\widehat{f}(s,t)$ that involves the polynomial $q_n(\cdot,t)$. Deriving an explicit expression for $q_n(s,t) = p_n(s,t)/p_n(0,t)$ for any $n \geq 0$ is quite complicated, but we can write

$$q_n(s,t) = \sum_{i=0}^{n} R_i^n(t) s^i,$$

where the coefficients, $R_i^n(t)$, of the polynomial follow directly from the coefficients of the polynomial in s, $p_n(s,t)$, which in turn can be obtained using the recursion (2.33).
Next, one can consider $n+1$ independent random variables Γ_i, where $\Gamma_i \sim \Gamma(2\theta + n - i, Q(t,T_n))$, for $i = 0,1,\ldots,n$. Further, consider the discrete random variable M^n, independent of the Γ_i, which assumes the values $0,1,\ldots,n$, with probabilities $\pi_i^n(t)$, and define the random variable

$$X_t^n = \sum_{i=0}^{n} 1_{\{M=i\}} \Gamma_i.$$

The moment generating function of X_t^n can easily be found, as Γ_i and M^n are independent, hence

$$\begin{aligned}\mathbb{E}\left[e^{sX_t^n}\right] &= \sum_{i=0}^{n} \mathbb{E}\left[e^{s\Gamma_i} 1_{\{M=i\}}\right] \\ &= \sum_{i=0}^{n} \pi_i^n(t) \mathbb{E}\left[e^{s\Gamma_i}\right] = \sum_{i=0}^{n} \pi_i^n(t) \left(\frac{Q(t,T_n)}{Q(t,T_n)-s}\right)^{2\theta+n-i}. \end{aligned} \tag{2.39}$$

The goal is to show that, by choosing the probabilities correctly, the moment generating function of X_t^n equals the conditional moment generation function $\widehat{f}(s,t)$.

Therefore (2.39) is first rewritten as

$$\mathbb{E}\left[e^{sX_t^n}\right] = \left(\frac{Q(t,T_n)}{Q(t,T_n)-s}\right)^{2\theta+n} \sum_{i=0}^{n} \pi_i^n(t) \left(\frac{Q(t,T_n)-s}{Q(t,T_n)}\right)^i .$$

To have that both moment generating functions $\widehat{f}(s,t)$ and (2.39) are equal, it is required that

$$q_n(s,t) = \sum_{i=0}^{n} R_i^n(t) s^i = \sum_{i=0}^{n} \pi_i^n(t) \left(\frac{Q(t,T_n)-s}{Q(t,T_n)}\right)^i .$$

The right hand side of this equation can be written as

$$\sum_{i=0}^{n} \pi_i^n(t) Q(t,T_n)^{-i} \sum_{j=0}^{i} \left(\begin{array}{c} i \\ j \end{array}\right) Q(t,T_n)^{i-j} s^j (-1)^j .$$

This equation can be turned into a polynomial in s, by interchanging the summations, which leads to

$$\begin{aligned} &\sum_{j=0}^{n}\sum_{i=j}^{n} \left(\begin{array}{c} i \\ j \end{array}\right) \pi_i^n(t) Q(t,T_n)^{-j} s^j (-1)^j \\ &= \sum_{j=0}^{n} s^j \left((-1)^j Q(t,T_n)^{-j} \sum_{i=j}^{n} \left(\begin{array}{c} i \\ j \end{array}\right) \pi_i^n(t) \right) . \end{aligned}$$

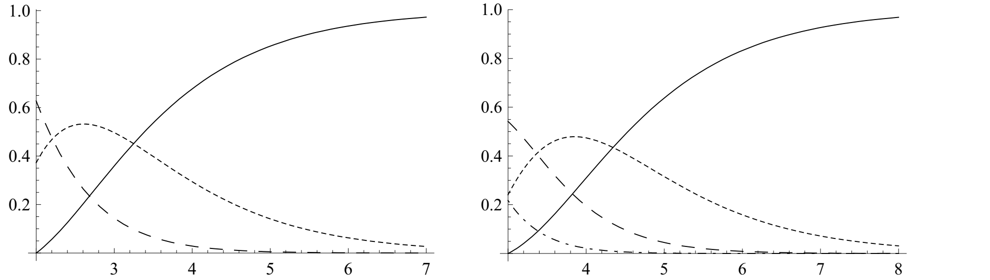
The moment generating functions are equal when

$$R_j^n(t) = (-1)^j Q(t,T_n)^{-j} \sum_{i=j}^{n} \left(\begin{array}{c} i \\ j \end{array}\right) \pi_i^n(t),$$

for $j = 0, 1, \ldots, n$. This can be solved iteratively, starting from $j = n$, which results in the probabilities

$$\pi_j^n(t) = (-1)^j R_j^n(t) Q(t,T_n)^j - \sum_{i=j+1}^{n} \pi_i^n(t) \left(\begin{array}{c} i \\ j \end{array}\right) . \tag{2.40}$$

It is not immediately clear from (2.40) that the $\pi_j^n(t)$ are all non-negative and sum to one. It turns out however that this is indeed the case for $T_n \leq t < T_{n+1}$, which means that the $\pi_j^n(t)$ can be interpreted as probabilities. It is however far from trivial to provide a general proof for all $n \geq 0$. We confine ourselves to illustrate this fact by some examples. In Figure 2.1, two graphs are given in which the probabilities are plotted.



(a) π_0^2 (short dashed line), π_1^2 (long dashed line) and π_2^2 (solid line)

(b) π_0^3 (dot dashed line), π_1^3 (short dashed line), π_2^3 (long dashed line) and π_3^3 (solid line)

Figure 2.1: Graphs of the mixing probabilities after two jumps of the process N_t, (a), and after three jumps, (b). The values of the previous jump times, T_1 and T_2 in case (a), and T_1, T_2 and T_3 in case (b), are taken as $T_i = i$, such that one is able to calculate the $\pi_j^n(t)$. The model parameters are chosen to be $\alpha = 0.5$, $\beta = 0.5$, $\mu_0 = 0.4$ and $\varphi = 4.0$.

2.6 Remarks on Model Extensions

We briefly discuss two ways of extending the model that we have considered in this chapter. First we consider time varying parameters for the intensity. This is a more realistic assumption from a practical point of view. Next, we look at a simple multi-factor specification of the intensity, where we see that the calculations for the one-dimensional case do not carry over to this multi-factor case.

Time Varying Parameters

In this section we briefly outline the consequences for our results when we replace the constant parameters in (2.2) with time varying ones. Clearly this introduces more flexibility of the model. So, we have $\alpha(t)$ instead of α, $\mu_0(t)$ instead of μ_0 etc. Many results in Sections 2 and 3 remain valid upon substitution of the constants by their time varying counterparts. In particular Equation (2.15) will change into

$$\frac{\partial}{\partial t}g(s,t) = s\mu_0(t)(\rho(t)s - \alpha(t) + \tau(t))(\rho(t)s - \alpha(t) - \tau(t))\frac{\partial}{\partial s}g(s,t). \tag{2.41}$$

But an explicit closed form solution for $g(s,t)$ that we were able to give for the constant parameter case by Equation (2.16) is in general impossible to obtain. The main reason for this is that transformation as given in (2.17) now introduces additional dependence on t and a simple PDE for $f_1(u,t)$ cannot be given. This complication carries over to similar ones for the functions $f_2(u,t)$ and $f_3(u,t)$.

If one uses piecewise constant functions for the parameters (as an approximation if needed), closed form solutions are still possible, although they will be given by complex expressions. We briefly outline how to get these. Suppose that $0 < t_1, t_2, \ldots$ (with

$t_i \to \infty$) denote the time instants where the parameters possibly change value. Consider a realization of the jump times $T_1, T_2, \ldots$. On each interval $[T_{n-1}, T_n)$ ($n \geq 1$) we relabel the t_i that fall in this interval by $\{t_1^n, \ldots, t_{k_n}^n\}$, which could be an empty set, in which case we can simply use (2.16) with the prevailing parameter values. Suppose now that this set is non-empty. On the subinterval $[T_{n-1}, t_1^n)$, we can compute the solution $g(s,t)$ to (2.41) again according to (2.16), eventually yielding $g(s, t_1^n-)$. Then we consider the PDE (2.41) on the interval $[t_1^n, t_2^n)$ with initial condition at t_1^n (instead of T) $w(s) = g(s, t_1^n-)$ and the values of the parameters on this interval. With the appropriate modifications, formula (2.16) can be used again. One then proceeds in this way until the final interval $[t_{k_n}^n, T_n)$ is reached, which eventually produces $g(s, T_n-)$.

We conclude by stating that more flexibility of the model by introducing time varying, but piecewise constant parameter functions also leads to closed form expressions, although they are more cumbersome to write down.

Multi-Factor Intensity

A second extension of the model that we have considered, is to assume that the intensity is driven by more than one Brownian motion, or factor. To illustrate the difficulties that emerge in such an extension, we look at a very simple two-factor model for the intensity,

$$\begin{aligned} \lambda_t &= \lambda_{1,t} + \lambda_{2,t} \\ \mathrm{d}\lambda_{i,t} &= -\alpha_i \left(\lambda_{i,t} - \mu_i\right) \mathrm{d}t + \beta_i \sqrt{\lambda_{i,t}}\, \mathrm{d}W_{i,t}, \ \text{for } i = 1, 2, \end{aligned}$$

where W_1 and W_2 are independent Brownian motions, and $\lambda_{1,t}$ and $\lambda_{2,t}$ both follow the CIR model with suitable parameter restrictions.

When we apply the filtering formulas (2.6) to λ_t we find

$$\mathrm{d}\widehat{\lambda}_t = \left(-\alpha_1 \left(\widehat{\lambda}_{1,t} - \mu_1\right) - \alpha_2 \left(\widehat{\lambda}_{2,t} - \mu_2\right)\right) \mathrm{d}t + \left(\frac{\widehat{\lambda^2_{t-}}}{\widehat{\lambda}_{t-}} - \widehat{\lambda}_{t-}\right) \left(\mathrm{d}N_t - \widehat{\lambda}_t \mathrm{d}t\right).$$

Just as in the one-dimensional case, this involves the term $\widehat{\lambda^2_t}$, and thus we again consider the conditional moment generating function $\widehat{f}(s,t) = \mathbb{E}\left[e^{s\lambda_t} \middle| \mathcal{F}_t^N\right]$. Therefore we have to determine the dynamics of $e^{s\lambda_t}$. An application of Itô's formula yields

$$\begin{aligned} \mathrm{d}e^{s\lambda_t} = &\left[\left(-\alpha_1 s + \frac{1}{2} s^2 \beta_1^2\right) \lambda_{1,t} e^{s\lambda_t} + \left(-\alpha_2 s + \frac{1}{2} s^2 \beta_2^2\right) \lambda_{2,t} e^{s\lambda_t} \right. \\ &\left. + s\left(\alpha_1 \mu_1 + \alpha_2 \mu_2\right)\right] \mathrm{d}t + s e^{s\lambda_t} \left(\beta_1 \sqrt{\lambda_{1,t}}\, \mathrm{d}W_{1,t} + \beta_2 \sqrt{\lambda_{2,t}}\, \mathrm{d}W_{2,t}\right). \end{aligned} \tag{2.42}$$

Comparing the terms in the square brackets above with those in (2.7), we directly observe that we have lost an important feature. In the one-dimensional case we could write the term $\lambda_t e^{s\lambda_t}$ as $\partial e^{s\lambda_t}/\partial s$, eventually resulting in the PDE that we could solve explicitly in Proposition 2.5. In the two factor model, the derivative of $e^{s\lambda_t}$ with respect to s, results in $(\lambda_{1,t} + \lambda_{2,t})\, e^{s\lambda_t}$. The terms $\lambda_{i,t} e^{s\lambda_t}$ in (2.42) thus cannot be written as $\partial e^{s\lambda_t}/\partial s$. This shows that a solution, similar to that of Proposition 2.5, cannot be obtained.

Alternatively, one could consider the conditional moment generating function of the pair $(\lambda_{1,t}, \lambda_{2,t})$, given by $h(s_1, s_2, t) = \mathbb{E}\left[e^{s_1\lambda_{1,t}+s_2\lambda_{2,t}} \,\middle|\, \mathcal{F}_t^N\right]$, and derive its dynamics. As we introduce an additional variable, the eventual PDE will be of a higher dimension and thus more complex. Obtaining an explicit closed form solution, if it exists, will be a substantially harder task and is beyond the scope of the present chapter. We conclude that it is far from straightforward to extend the explicit solution that we have obtained to models with more than one-factor.

CHAPTER 3

Sample Path Large Deviations in Credit Risk

In Chapter 2 we presented a model for the loss in a credit risky portfolio, where only partial information, observing the defaults only, was assumed. In the current chapter we again focus on the loss process of a credit risky portfolio. In the results we present, that are based on [LMS09], we assume that the constituents of the portfolio are independent and identically distributed. Under this assumption we derive a sample-path large deviation principle for the loss process. Furthermore, we derive the exact asymptotics for the probability that the loss process ever exceeds a certain threshold. The chapter is concluded with a discussion about extending the model to the case where the constituents can have different distributions or the constituents are correlated.

3.1 Introduction

For financial institutions, such as banks and insurance companies, it is of crucial importance to accurately assess the risk of their portfolios. These portfolios typically consist of a large number of obligors, such as mortgages, loans or insurance policies, and therefore it is computationally infeasible to treat each individual object in the portfolio separately. As a result, attention has shifted to measures that characterize the risk of the portfolio as a whole, see e.g. [DS03] for general principles concerning managing credit risk. The best known metric is the so-called *value at risk*, see [Jor06], measuring the minimum amount of money that can be lost with α percent certainty over some given period. Several other measures have been proposed, such as economic capital, the risk adjusted return on capital (RAROC), or expected shortfall, which is a coherent risk measure [ADEH99]. Each of these measure are applicable to market risk as well as credit risk. Measures such as loss given default (LGD) and exposure at default (EAD) are measures that purely apply to credit risk. These and other measures are discussed in detail in e.g. [MFE05].

The currently existing methods mainly focus on the distribution of the portfolio loss *up to a given point in time* (for instance one year into the future). It can be argued, however, that in many situations it makes more sense to use probabilities that involve the (cumulative) loss *process*, say $\{L(t) : t \geq 0\}$. Highly relevant, for instance, is the event that $L(\cdot)$ ever exceeds a given function $\zeta(\cdot)$ (within a certain time window, for instance between now and one year ahead), i.e., an event of the type

$$\{\exists t \leq T : L(t) \geq \zeta(t)\}. \tag{3.1}$$

It is clear that measures of the latter type are intrinsically harder to analyze, as it does not suffice anymore to have knowledge of the marginal distribution of the loss process at a given point in time; for instance the event (3.1) actually corresponds to the union of events $\{L(t) \geq \zeta(t)\}$, for $t \leq T$, and its probability will depend on the law of $L(\cdot)$ as a process on $[0, T]$.

In line with the remarks we made above, earlier papers on applications of large-deviation theory to credit risk, mainly address the (asymptotics of the) distribution of the loss process at a single point in time, see e.g. [DDD04] and [GKS07]. The former paper considers, in addition, also the probability that the *increments* of the loss process exceed a certain level. Other approaches to quantifying the tail distribution of the losses have been taken by [LKSS01], who use extreme-value theory (see [EKM97] for a background), [Gor02] and [MTB01], where the authors consider saddle point approximations to the tails of the loss distribution. Numerical and simulation techniques for credit risk can be found in e.g. [Gla04]. The first contribution of our work concerns a so-called sample-path large deviation principle (LDP) for the average cumulative losses for large portfolios. Loosely speaking, such an LDP means that, with $L_n(\cdot)$ denoting the loss process when n obligors are involved, we can compute the logarithmic asymptotics (for n large) of the normalized loss process $L_n(\cdot)/n$ being in a set of trajectories A:

$$\lim_{n\to\infty} \frac{1}{n} \log \mathbb{P}\left(\frac{1}{n} L_n(\cdot) \in A\right); \tag{3.2}$$

we could for instance pick a set A that corresponds to the event (3.1). Most of the sample-path LDPs that have been developed up to now involve stochastic processes with independent or nearly-independent increments, see for instance the results by Mogul'skiĭ for random walks [Mog76], de Acosta for Lévy processes [Aco94], and Chang [Cha95] for weakly-correlated processes; results for processes with a stronger correlation structure are restricted to special classes of processes, such as Gaussian processes, see e.g. [Aze80]. It is observed that our loss process is not covered by these results, and therefore new theory had to be developed. The proof of our LDP relies on 'classical' large-deviation results (such as Cramér's theorem, Sanov's theorem, Mogul'skiĭ's theorem), but in addition the concept of *epi-convergence* [Kal86] is relied upon.

Our second main result focuses specifically on the event (3.1) of ever (before some time horizon T) exceeding a given barrier function $\zeta(\cdot)$. Whereas we so far considered, inherently imprecise, logarithmic asymptotics of the type displayed in (3.2), we can now compute so-called *exact asymptotics*: we identify an explicit function $f(n)$ such that $f(n)/p_n \to 1$ as $n \to \infty$, where p_n is the probability of our interest. As is known from the literature, it is in general substantially harder to find exact asymptotics than logarithmic asymptotics. The proof uses the fact that, after discretizing time, the contribution of just a single time epoch dominates, in the sense that there is a $t^\star$ such that

$$\mathbb{P}\left(\frac{1}{n}\,L_n(t^\star) \geq \zeta(t^\star)\right)\Big/\, p_n \to 1, \quad \text{with } p_n := \mathbb{P}\left(\exists t : \frac{1}{n}\,L_n(t) \geq \zeta(t)\right). \tag{3.3}$$

This $t^\star$ can be interpreted as the most likely epoch of exceeding $\zeta(\cdot)$.

Turning back to the setting of credit risk, both of the results we present are derived in a setup where all obligors in the portfolio are i.i.d., in the sense that they behave independently and stochastically identically. A third contribution of our work concerns a discussion on how to extend our results to cases where the obligors are *dependent* (meaning that they, in the terminology of [DDD04], react to the same 'macro-environmental' variable, conditional upon which they are independent again). We also treat the case of obligor-heterogeneity: we show how to extend the results to the situation of multiple classes of obligors.

The chapter is structured as follows. In Section 3.2 we introduce the loss process and we describe the scaling under which we work. We also recapitulate a couple of relevant large-deviation results. Our first main result, the sample-path LDP for the cumulative loss process, is stated and proved in Section 3.3. Special attention is paid to, easily-checkable, sufficient conditions under which this result holds. As argued above, the LDP is a generally applicable result, as it yields an expression for the decay rate of any probability that depends on the entire sample path. Then, in Section 3.4, we derive the exact asymptotic behavior of the probability that, at some point in time, the loss exceeds a certain threshold, i.e., the asymptotics of p_n, as defined in (3.3). After this we derive a similar result for the *increments* of the loss process. Eventually, in Section 3.5, we discuss a number of possible extensions to the results we have presented. Special attention is given to allowing dependence between obligors, and to different classes of obligors each having its own specific distributional properties.

3.2 Notation and Definitions

The portfolios of banks and insurance companies are typically very large: they may consist of several thousands of assets. It is therefore computationally impossible to estimate the risks for each element, or obligor, in a portfolio. This explains why one attempts to assess the aggregated losses resulting from defaults, e.g. bankruptcies, failure to repay loans or insurance claims, for the portfolio as a whole. The risk in the portfolio is then measured through this (aggregate) loss process. In the following sections we introduce the loss process and the portfolio constituents more formally.

3.2.1 Loss Process

Let $(\Omega, \mathcal{F}, \mathbb{P})$ be the probability space on which all random variables below are defined. We assume that the portfolio consists of n obligors and we denote the default time of obligor i by τ_i. Further we write U_i for the loss incurred on a default of obligor i. We then define the cumulative loss process L_n as

$$L_n(t) := \sum_{i=1}^{n} U_i Z_i(t), \tag{3.4}$$

where $Z_i(t) = 1_{\{\tau_i \le t\}}$ is the default indicator of obligor i. We assume that the loss amounts $U_i \ge 0$ are i.i.d., and that the default times $\tau_i \ge 0$ are i.i.d. as well. In addition we assume that the loss amounts and the default times are mutually independent. In the remainder of this chapter U and $Z(t)$ denote generic random variables with the same distribution as the U_i and $Z_i(t)$, respectively.

Throughout this chapter we assume that the defaults only occur on the time grid $\mathbb{N}$; in Section 3.5 we discuss how to deal with the default epochs taking continuous values. In some cases we explicitly consider a finite time grid, say $\{1, 2, \ldots, N\}$. The extension of the results we derive to a more general grid $\{0 < t_1 < t_2 < \ldots < t_N\}$ is completely trivial. The distribution of the default times, for each j, is denoted by

$$p_j := \mathbb{P}\left(\tau = j\right), \tag{3.5}$$

$$F_j := \mathbb{P}\left(\tau \le j\right) = \sum_{i=1}^{j} p_i. \tag{3.6}$$

Given the distribution of the loss amounts U_i and the default times τ_i, our goal is to investigate the loss process. Many of the techniques that have been developed so far, first fix a time T (typically one year), and then stochastic properties of the cumulative loss at time T, i.e., $L_n(T)$, are studied. Measures such as value at risk and economic capital are examples of these 'one-dimensional' characteristics. Many interesting measures, however, involve properties of the entire *path* of the loss process rather than those of just one time epoch, examples being the probability that $L_n(\cdot)$ exceeds some barrier function $\zeta(\cdot)$ for some t smaller than the horizon T, or the probability that (during a certain period) the loss always stays above a certain level. The event corresponding to the former probability might require the bank to attract more capital, or worse, it might lead to the bankruptcy of this bank. The event corresponding to the latter event might also lead to the bankruptcy of the bank, as a long period of stress may have substantial negative implications. We conclude that having a handle on these probabilities is therefore a useful instrument when assessing the risks involved in the bank's portfolios.

As mentioned above, the number of obligors n in a portfolio is typically very large, thus prohibiting analyses based on the specific properties of the individual obligors. Instead, it is more natural to study the asymptotical behavior of the loss process as $n \to \infty$. One could rely on a central-limit-theorem based approach, but in this chapter we focus on rare events, by using the theory of large deviations.

In the following subsection we provide some background of large-deviation theory, and we define a number of quantities that are used in the remainder of this chapter.

3.2.2 Large Deviation Principle

In this section we give a short introduction to the theory of large deviations. Here, in an abstract setting, the limiting behavior of a family of probability measures $\{\mu_n\}$ on the Borel sets $\mathcal{B}$ of a metric space $(\mathcal{X}, d)$ is studied, as $n \to \infty$. This behavior is referred to as the Large Deviation Principle (LDP), and it is characterized in terms of a *rate function*. The LDP states lower and upper exponential bounds for the value that the measures μ_n assign to sets in a topological space $\mathcal{X}$. Below we state the definition of the rate function that has been taken from [DZ98].

Definition 3.1. *A rate function is a lower semicontinuous mapping $I : \mathcal{X} \to [0, \infty]$, for all $\alpha \in [0, \infty)$ the level set $\Psi_I(\alpha) := \{x | \, I(x) \leq \alpha\}$ is a closed subset of $\mathcal{X}$. A good rate function is a rate function for which all the level sets are compact subsets of $\mathcal{X}$.*

With the definition of the rate function in mind we state the large deviation principle for the sequence of measure $\{\mu_n\}$.

Definition 3.2. *We say that $\{\mu_n\}$ satisfies the large deviation principle with a rate function $I(\cdot)$ if*

(i) (Upper bound) for any closed set $F \subseteq \mathcal{X}$

$$\limsup_{n\to\infty} \frac{1}{n} \log \mu_n(F) \leq - \inf_{x\in F} I(x). \tag{3.7}$$

(ii) (Lower bound) for any open set $G \subseteq \mathcal{X}$

$$\liminf_{n\to\infty} \frac{1}{n} \log \mu_n(G) \geq - \inf_{x\in G} I(x). \tag{3.8}$$

We say that a family of random variables $X = \{X_n\}$, with values in $\mathcal{X}$, satisfies an LDP with rate function $I_X(\cdot)$ iff the laws $\{\mu_n^X\}$ satisfy an LDP with rate function I_X, where μ_n^X is the law of X_n.

The so-called Fenchel-Legendre transform plays an important role in expressions for the rate function. Let for an arbitrary random variable X, the logarithmic moment generating function, sometimes referred to as cumulant generating function, be given by

$$\Lambda_X(\theta) := \log M_X(\theta) = \log \mathbb{E}\left[e^{\theta X}\right] \leq \infty, \tag{3.9}$$

for $\theta \in \mathbb{R}$. The Fenchel-Legendre transform $\Lambda_X^\star$ of Λ_X is then defined by

$$\Lambda_X^\star(x) := \sup_\theta \left(\theta x - \Lambda_X(\theta)\right). \tag{3.10}$$

We sometimes say that $\Lambda_X^\star$ is the Fenchel-Legendre transform of X.

The LDP from Definition 3.2 provides upper and lower bounds for the log-asymptotic behavior of measures μ_n. In case of the loss process (3.4), *fixed* at some time t, we can easily establish an LDP by an application of Cramér's theorem (Theorem 3.17). This theorem yields that the rate function is given by $\Lambda_{UZ(t)}^\star(\cdot)$, where $\Lambda_{UZ(t)}^\star(\cdot)$ is the Fenchel-Legendre transform of the random variable $UZ(t)$.

The results we present in this chapter involve either $\Lambda_U^\star(\cdot)$ (Section 3.3), which corresponds to i.i.d. loss amounts U_i only, or $\Lambda_{UZ(t)}^\star(\cdot)$ (Section 3.4), which corresponds to those loss amounts up to time t. In the following section we derive an LDP for the whole path of the loss process, which can be considered as an extension of Cramér's theorem.

3.3 A Sample-Path Large Deviation Result

In the previous section we have introduced the large deviation principle. In this section we derive a sample-path LDP for the cumulative loss process (3.4). We consider the exponential decay of the probability that the path of the loss process $L_n(\cdot)$ is in some set A, as the size n of the portfolio tends to infinity.

3.3.1 Assumptions

In order to state a sample-path LDP we need to define the topology that we work on. To this end we define the space $\mathscr{S}$ of all nonnegative and nondecreasing functions on $T_N = \{1, 2, \ldots, N\}$,

$$\mathscr{S} := \{f : T_N \to \mathbb{R}_0^+ \mid 0 \leq f_i \leq f_{i+1} \text{ for } i < N\}.$$

This set is identified with the space $\mathbb{R}^N_\leq := \{x \in \mathbb{R}^N \mid 0 \leq x_i \leq x_{i+1} \text{ for } i < N\}$. The topology on this space is the one induced by the supremum norm

$$||f||_\infty = \max_{i=1,\ldots,N} |f_i|.$$

As we work on a finite-dimensional space, the choice of the norm is not important, as any other norm on $\mathscr{S}$ would result in the same topology. We use the supremum norm as this is convenient in some of the proofs in this section.

We identify the space of all probability measures on T_N with the simplex Φ:

$$\Phi := \left\{ \varphi \in \mathbb{R}^N \;\middle|\; \sum_{i=1}^{N} \varphi_i = 1,\ \varphi_i \geq 0 \text{ for } i \leq N \right\}. \tag{3.11}$$

For a given $\varphi \in \Phi$ we denote the cumulative distribution function by ψ, i.e.,

$$\psi_i = \sum_{j=1}^{i} \varphi_j, \quad \text{for } i \leq N; \tag{3.12}$$

note that $\psi \in \mathscr{S}$ and $\psi_N = 1$.

Furthermore, we consider the loss amounts U_i as introduced in Section 3.2.1, a $\varphi \in \Phi$ with cdf ψ, and a sequence of $\varphi^n \in \Phi$, each with cdf ψ^n, such that $\varphi^n \to \varphi$ as $n \to \infty$, meaning that $\varphi_i^n \to \varphi_i$ for all $i \leq N$. We define two families of measures (μ_n) and (ν_n),

$$\mu_n(A) := \mathbb{P}\left(\left(\frac{1}{n} \sum_{j=1}^{[n\psi_i]} U_j \right)_{i=1}^{N} \in A \right), \tag{3.13}$$

$$\nu_n(A) := \mathbb{P}\left(\left(\frac{1}{n} \sum_{j=1}^{[n\psi_i^n]} U_j \right)_{i=1}^{N} \in A \right), \tag{3.14}$$

where $A \in \mathcal{B} := \mathcal{B}(\mathbb{R}^N)$ and $[x] := \sup\{k \in N | k \leq x\}$. Below we state an assumption under which the main result in this section holds. This assumption refers to the definition of *exponential equivalence*, which can be found in Definition 3.18.

Assumption 3.3. *Assume that $\varphi^n \to \varphi$ and moreover that the measures μ_n and ν_n as defined in* (3.13) *and* (3.14), *respectively, are exponentially equivalent.*

From Assumption 3.3 we learn that the differences between the two measures μ_n and ν_n go to zero at a ‘superexponential’ rate. In the next section, in Lemma 3.6, we provide a sufficient condition, that is easy to check, under which this assumption holds.

3.3.2 Main Result

The assumptions and definitions in the previous sections allow us to state the main result of this section. We show that the average loss process satisfies a large deviation principle as in (3.2). This principle allows us to approximate a large variety of probabilities related to the average loss process, such as the probability that the loss process stays above a certain time dependent level or the probability that the loss process exceeds a certain level before some given point in time.

Theorem 3.4. *With Φ as in (3.11) and under Assumption 3.3, the average loss process, $L_n(\cdot)/n$ satisfies an LDP with rate function $I_{U,p}$. Here, for $x \in \mathbb{R}^N_{\leq}$, $I_{U,p}$ is given by*

$$I_{U,p}(x) := \inf_{\varphi \in \Phi} \sum_{i=1}^{N} \varphi_i \left(\log\left(\frac{\varphi_i}{p_i}\right) + \Lambda_U^\star \left(\frac{\Delta x_i}{\varphi_i}\right)\right), \tag{3.15}$$

with $\Delta x_i := x_i - x_{i-1}$ and $x_0 := 0$.

Observing the rate function for this sample path LDP, we see that the effects of the default times τ_i and the loss amounts U_i are nicely decomposed into the two terms in the rate function, one involving the distribution of the default epoch τ (the 'Sanov term', cf. [DZ98, Thm. 6.2.10]), the other one involving the incurred loss size U (the 'Cramér term', cf. [DZ98, Thm. 2.2.3]. Observe that we recover Cramér's theorem by considering a time grid *consisting of a single time point*, which means that Theorem 3.4 extends Cramér's result. We also remark that, informally speaking, the optimizing $\varphi \in \Phi$ in (3.15) can be interpreted as the 'most likely' distribution of the loss epoch, given that the path of $L_n(\cdot)/n$ is close to x.

As a sanity check we calculate the value of the rate function $I_{U,p}(x)$ for the 'average path' of $L_n(\cdot)/n$, given by $x_j^\star = \mathbb{E}[U] F_j$ for $j \leq N$, where F_j is the cumulative distribution of the default times as given in (3.6); this path should give a rate function equal to 0. To see this we first remark that clearly $I_{U,p}(x) \geq 0$ for all x, since both the Sanov term and the Cramér term are non-negative. This yields the following chain of inequalities:

$$\begin{aligned}
0 \leq I_{U,p}(x^\star) &= \inf_{\varphi \in \Phi} \sum_{i=1}^{N} \varphi_i \left(\log\left(\frac{\varphi_i}{p_i}\right) + \Lambda_U^\star \left(\frac{\mathbb{E}[U]\, p_i}{\varphi_i}\right)\right) \\
&\overset{\varphi=p}{\leq} \sum_{i=1}^{N} p_i \left(\log\left(\frac{p_i}{p_i}\right) + \Lambda_U^\star \left(\frac{\mathbb{E}[U]\, p_i}{p_i}\right)\right) \\
&= \sum_{i=1}^{N} p_i \Lambda_U^\star (\mathbb{E}[U]) = \Lambda_U^\star(\mathbb{E}[U]) = 0,
\end{aligned}$$

where we have used that for $\mathbb{E}[U] < \infty$, it always holds that $\Lambda_U^\star(\mathbb{E}[U]) = 0$ cf. [DZ98, Lemma 2.2.5]. The inequalities above thus show that, if the 'average path' $x^\star$ lies in the set of interest, then the corresponding decay rate is 0, meaning that the probability of interest decays subexponentially.

In the proof of Theorem 3.4 we use the following lemma, which is related to the concept of *epi-convergence*, extensively discussed in [Kal86].

Lemma 3.5. *Let $f_n, f : D \to \mathbb{R}$, with $D \subset \mathbb{R}^m$ compact. Assume that for all $x \in D$ and for all $x_n \to x$ in D we have*

$$\limsup_{n\to\infty} f_n(x_n) \leq f(x). \tag{3.16}$$

Then we have

$$\limsup_{n\to\infty} \sup_{x\in D} f_n(x) \leq \sup_{x\in D} f(x).$$

Proof. Let $f_n^\star = \sup_{x\in D} f_n(x)$, $f^\star = \sup_{x\in D} f(x)$. Consider a subsequence $f_{n_k}^\star$ converging to $\limsup_{n\to\infty} f_n^\star$. Let $\varepsilon > 0$ and choose x_{n_k} such that $f_{n_k}^\star < f_{n_k}(x_{n_k}) + \varepsilon$ for all k. By the compactness of D, there exists a limit point x such that along a subsequence $x_{n_{k_j}} \to x$. By the hypothesis (3.16) we then have

$$\limsup_{n\to\infty} f_n^\star \leq \lim f_{n_{k_j}}(x_{n_{k_j}}) + \varepsilon = f(x) + \varepsilon \leq f^\star + \varepsilon.$$

Let $\varepsilon \downarrow 0$ to obtain the result. □

Proof of Theorem 3.4. We start by establishing an identity from which we show both bounds. We need to calculate the probability

$$\mathbb{P}\left(\frac{1}{n}L_n(\cdot) \in A\right) = \mathbb{P}\left(\left(\frac{1}{n}L_n(1), \ldots, \frac{1}{n}L_n(N)\right) \in A\right),$$

for certain $A \in \mathcal{B}$. For each point j on the time grid T_N we record by the 'default counter' $K_{n,j} \in \{0, \ldots, n\}$ the number of defaults at time j:

$$K_{n,j} := \#\{i \in \{1, \ldots, n\} \mid \tau_i = j\}.$$

These counters allow us to rewrite the probability to

$$\begin{aligned}\mathbb{P}\left(\frac{1}{n}L_n(\cdot) \in A\right) &= \mathbb{E}\left[\mathbb{P}\left(\left(\frac{1}{n}\sum_{j=1}^{K_{n,1}} U_{(j)}, \ldots, \frac{1}{n}\sum_{j=1}^{K_{n,1}+\ldots+K_{n,N}} U_{(j)}\right) \in A \,\middle|\, K_n\right)\right]\\ &= \sum_{k_1+\ldots+k_N=n} \mathbb{P}\left(K_{n,i} = k_i \text{ for } i \leq N\right) \times \mathbb{P}\left(\left(\frac{1}{n}\sum_{j=1}^{m_i} U_{(j)}\right)_{i=1}^{N} \in A\right),\end{aligned} \tag{3.17}$$

where $m_i := \sum_{j=1}^{i} k_j$ and the loss amounts U_j have been ordered, such that the first $U_{(j)}$ correspond to the losses at time 1, etc.

Upper bound. Starting from Equality (3.17), let us first establish the upper bound of the LDP. To this end, let F be a closed set and consider the decay rate

$$\limsup_{n\to\infty} \frac{1}{n} \log \mathbb{P}\left(\frac{1}{n}L_n(\cdot) \in F\right). \tag{3.18}$$

An application of Lemma 3.19 together with (3.17), implies that (3.18) equals

$$\limsup_{n\to\infty} \frac{1}{n} \log \mathbb{P}\left(\frac{1}{n} L_n(\cdot) \in F\right)$$
$$= \limsup_{n\to\infty} \max_{\sum k_j = n} \frac{1}{n} \left[\log \mathbb{P}\left(\frac{K_{n,i}}{n} = \frac{k_i}{n},\ i \le N\right) + \log \mathbb{P}\left(\left(\left(\frac{1}{n}\sum_{j=1}^{m_i} U_{(j)}\right)_{i=1}^{N} \in F\right)\right)\right], \tag{3.19}$$

Next we replace the dependence on n in the maximization by maximizing over the set Φ as in (3.11). In addition, we replace the k_i in (3.19) by

$$\widehat{\varphi}_{n,i} := \frac{[n\psi_i] - [n\psi_{i-1}]}{n}, \tag{3.20}$$

where the ψ_i have been defined in (3.12). As a result, (3.18) reads

$$\limsup_{n\to\infty} \sup_{\varphi\in\Phi} \frac{1}{n} \left[\log \mathbb{P}\left(\frac{K_{n,i}}{n} = \widehat{\varphi}_{n,i},\ i \le N\right) + \log \mathbb{P}\left(\left(\left(\frac{1}{n}\sum_{j=1}^{[n\psi_i]} U_{(j)}\right)_{i=1}^{N} \in F\right)\right)\right]. \tag{3.21}$$

Note that (3.18) equals (3.21), since for each n and vector $(k_1, \dots, k_N) \in \mathbb{N}^N$, with $\sum_{i=1}^N k_i = n$, there is a $\varphi \in \Phi$ with $\varphi_i = k_i/n$. On the other hand, we only cover outcomes of this form by rounding off the φ_i.

We can bound the first term in this expression from above using Lemma 3.21, which implies that the decay rate (3.18) is majorized by

$$\limsup_{n\to\infty} \sup_{\varphi\in\Phi} \left[-\sum_{i=1}^{N} \widehat{\varphi}_{n,i} \log\left(\frac{\widehat{\varphi}_{n,i}}{p_i}\right) + \frac{1}{n}\log \mathbb{P}\left(\left(\left(\frac{1}{n}\sum_{j=1}^{[n\psi_i]} U_{(j)}\right)_{i=1}^{N} \in F\right)\right)\right].$$

Now note that calculating the limsup in the previous expression is not straightforward due to the supremum over Φ. The idea is therefore to interchange the supremum and the limsup, by using Lemma 3.5. To apply this lemma we first introduce

$$f_n(\varphi) := -\sum_{i=1}^{N} \widehat{\varphi}_{n,i} \log\left(\frac{\widehat{\varphi}_{n,i}}{p_i}\right) + \frac{1}{n}\log \mathbb{P}\left(\left(\left(\frac{1}{n}\sum_{j=1}^{[n\psi_i]} U_{(j)}\right)_{i=1}^{N} \in F\right)\right),$$
$$f(\varphi) := -\inf_{x\in F} \sum_{i=1}^{N} \varphi_i \left(\log\left(\frac{\varphi_i}{p_i}\right) + \Lambda_U^\star\left(\frac{\Delta x_i}{\varphi_i}\right)\right),$$

and note that Φ is a compact subset of $\mathbb{R}^n$. We have to show that for any sequence $\varphi^n \to \varphi$

$$\limsup_{n\to\infty} f_n(\varphi^n) \leq f(\varphi), \tag{3.22}$$

such that the conditions of Lemma 3.5 are satisfied. We observe, with ψ_i^n as in (3.12) and $\widehat{\varphi}_i^n$ as in (3.20) with φ replaced by φ^n, that

$$\begin{aligned}\limsup_{n\to\infty} f_n(\varphi^n) \leq \limsup_{n\to\infty} &\left(-\sum_{i=1}^{N} \widehat{\varphi}_{n,i}^n \log\left(\frac{\widehat{\varphi}_{n,i}^n}{p_i}\right)\right) \\ &+ \limsup_{n\to\infty} \frac{1}{n} \log \mathbb{P}\left(\left(\frac{1}{n}\sum_{j=1}^{[n\psi_i^n]} U_{(j)}\right)_{i=1}^{N} \in F\right).\end{aligned}$$

Since $\varphi^n \to \varphi$ and since $\widehat{\varphi}_{n,i}^n$ differs at most by $1/n$ from φ_i^n, it immediately follows that $\widehat{\varphi}_{n,i}^n \to \varphi_i$. For an arbitrary continuous function g we thus have $g\left(\widehat{\varphi}_{n,i}^n\right) \to g(\varphi_i)$. This implies that

$$\limsup_{n\to\infty} \left(-\sum_{i=1}^{N} \widehat{\varphi}_{n,i} \log\left(\frac{\widehat{\varphi}_{n,i}^n}{p_i}\right)\right) = -\sum_{i=1}^{N} \varphi_i \log\left(\frac{\varphi_i}{p_i}\right). \tag{3.23}$$

Inequality (3.22) is established once we have shown that

$$\limsup_{n\to\infty} \frac{1}{n} \log \mathbb{P}\left(\left(\frac{1}{n}\sum_{j=1}^{[n\psi_i^n]} U_{(j)}\right)_{i=1}^{N} \in F\right) \leq -\inf_{x\in F} \sum_{i=1}^{N} \varphi_i \Lambda_U^\star\left(\frac{\Delta x_i}{\varphi_i}\right). \tag{3.24}$$

By Assumption 3.3, we can exploit the exponential equivalence together with Theorem 3.23, to see that (3.24) holds as soon as we have that

$$\limsup_{n\to\infty} \frac{1}{n} \log \mathbb{P}\left(\left(\frac{1}{n}\sum_{j=1}^{[n\psi_i]} U_{(j)}\right)_{i=1}^{N} \in F\right) \leq -\inf_{x\in F} \sum_{i=1}^{N} \varphi_i \Lambda_U^\star\left(\frac{\Delta x_i}{\varphi_i}\right).$$

But this inequality is a direct consequence of Lemma 3.22, and we conclude that (3.24) holds. Combining (3.23) with (3.24) yields

$$\begin{aligned}\limsup_{n\to\infty} f_n(\varphi^n) &\leq -\sum_{i=1}^{N} \varphi_i \log\left(\frac{\varphi_i}{p_i}\right) - \inf_{x\in F} \sum_{i=1}^{N} \varphi_i \Lambda_U^\star\left(\frac{\Delta x_i}{\varphi_i}\right) \\ &= -\inf_{x\in F} \sum_{i=1}^{N} \varphi_i \left(\log\left(\frac{\varphi_i}{p_i}\right) + \Lambda_U^\star\left(\frac{\Delta x_i}{\varphi_i}\right)\right) = f(\varphi),\end{aligned}$$

so that indeed the conditions of Lemma 3.5 are satisfied, and therefore

$$\limsup_{n\to\infty}\sup_{\varphi\in\Phi} f_n(\varphi) \le \sup_{\varphi\in\Phi} f(\varphi) = \sup_{\varphi\in\Phi}\left(-\inf_{x\in F}\sum_{i=1}^{N}\varphi_i\left(\log\left(\frac{\varphi_i}{p_i}\right)+\Lambda_U^\star\left(\frac{\Delta x_i}{\varphi_i}\right)\right)\right)$$

$$= -\inf_{x\in F}\inf_{\varphi\in\Phi}\sum_{i=1}^{N}\varphi_i\left(\log\left(\frac{\varphi_i}{p_i}\right)+\Lambda_U^\star\left(\frac{\Delta x_i}{\varphi_i}\right)\right) = -\inf_{x\in F} I_{U,p}(x).$$

This establishes the upper bound of the LDP.

Lower bound. To complete the proof, we need to establish the corresponding lower bound. Let G be an open set and consider

$$\liminf_{n\to\infty}\frac{1}{n}\log\mathbb{P}\left(\frac{1}{n}L_n(\cdot)\in G\right). \tag{3.25}$$

We apply Equality (3.17) to this liminf, with A replaced by G, and we observe that this sum is larger than the largest term in the sum, which shows that (where we directly switch to the enlarged space Φ) the decay rate (3.25) majorizes

$$\liminf_{n\to\infty}\sup_{\varphi\in\Phi}\frac{1}{n}\left(\log\mathbb{P}\left(\frac{1}{n}K_{n,i}=\widehat{\varphi}_{n,i},\ i\le N\right)+\log\mathbb{P}\left(\left(\frac{1}{n}\sum_{j=1}^{[n\psi_i]}U_{(j)}\right)_{i=1}^{N}\in G\right)\right).$$

Observe that for any sequence of functions $h_n(\cdot)$ it holds that $\liminf_n \sup_x h_n(x) \ge \liminf_n h_n(\tilde{x})$ for all $\tilde{x}$, so that we obtain the evident inequality

$$\liminf_{n\to\infty}\sup_x h_n(x) \ge \sup_x \liminf_{n\to\infty} h_n(x).$$

This observation yields that the decay rate of interest (3.25) is not smaller than

$$\sup_{\varphi\in\Phi}\left(\liminf_{n\to\infty}\frac{1}{n}\log\mathbb{P}\left(\frac{1}{n}K_{n,i}=\widehat{\varphi}_{n,i},\ i\le N\right)+\right.$$
$$\left.\liminf_{n\to\infty}\frac{1}{n}\log\mathbb{P}\left(\left(\frac{1}{n}\sum_{j=1}^{[n\psi_i]}U_{(j)}\right)_{i=1}^{N}\in G\right)\right) \tag{3.26}$$

where we have used that $\liminf_n (x_n + y_n) \geq \liminf_n x_n + \liminf_n y_n$. We apply Lemma 3.21 to the first liminf in (3.26), leading to

$$\begin{aligned}
&\liminf_{n\to\infty} \frac{1}{n} \log \mathbb{P}\left(\frac{1}{n} K_{n,i} = \widehat{\varphi}_{n,i},\ i \leq N\right) \\
&\geq \liminf_{n\to\infty} \left(-\sum_{i=1}^{N} \widehat{\varphi}_{n,i} \log\left(\frac{\widehat{\varphi}_{n,i}}{p_i}\right) - \frac{N}{n}\log(n+1)\right) \\
&= -\sum_{i=1}^{N} \varphi_i \log\left(\frac{\varphi_i}{p_i}\right), \qquad (3.27)
\end{aligned}$$

since $\log(n+1)/n \to 0$ as $n \to \infty$. The second liminf in (3.26) can be bounded from below by an application of Lemma 3.22. Since G is an open set, this lemma yields

$$\liminf_{n\to\infty} \frac{1}{n} \log \mathbb{P}\left(\left(\frac{1}{n}\sum_{j=1}^{[n\psi_i]} U_{(j)}\right)_{i=1}^{N} \in G\right) \geq -\inf_{x\in G} \sum_{i=1}^{N} \varphi_i \Lambda_U^{\star}\left(\frac{x_i - x_{i-1}}{\varphi_i}\right).$$

Upon combining these two results, we see that we have established the lower bound:

$$\begin{aligned}
&\liminf_{n\to\infty} \frac{1}{n} \log \mathbb{P}\left(\frac{1}{n} L_n(\cdot) \in G\right) \\
&\geq -\inf_{\varphi\in\Phi} \inf_{x\in A} \left(\sum_{i=1}^{N} \varphi_i \left(\log\left(\frac{\varphi_i}{pi}\right) + \Lambda_U^{\star}\left(\frac{x_i - x_{i-1}}{\varphi_i}\right)\right)\right) = -\inf_{x\in G} I_{U,p}(x).
\end{aligned}$$

This completes the proof of the theorem. □

In order to apply Theorem 3.4, one needs to check that Assumption 3.3 holds. In general this could be a quite cumbersome exercise. In Lemma 3.6 below we provide a sufficient, easy-to-check condition under which this assumption holds.

Lemma 3.6. *Assume that for all $\theta \in \mathbb{R} : \Lambda_U(\theta) < \infty$. Then Assumption 3.3 holds.*

Remark 3.7. The assumption we make in Lemma 3.6, i.e., that the logarithmic moment generating function is finite everywhere, is a common assumption in large deviations theory. We remark that for instance Mogul'skiĭ's theorem [DZ98, Thm. 5.1.2] also relies on this assumption; this theorem is a sample-path LDP for

$$Y_n(t) := \frac{1}{n} \sum_{i=1}^{[nt]} X_i,$$

on the interval $[0, 1]$. In Mogul'skiĭ's result the X_i are assumed to be i.i.d; in our model we have that $L_n(t) = \sum_{i=1}^{n} U_i Z_i(t)/n$, so that our sample-path result clearly does not fit into the setup of Mogul'skiĭ's theorem. ♢

Remark 3.8. In Lemma 3.6 it was assumed that $\Lambda_U(\theta) < \infty$, for all $\theta \in \mathbb{R}$, but an equivalent condition is

$$\lim_{x\to\infty} \frac{\Lambda_U^\star(x)}{x} = \infty. \tag{3.28}$$

In other words: this alternative condition can be used instead of the condition stated in Lemma 3.6. To see that both requirements are equivalent, make the following observations. In Lemma 3.20 it is shown that (3.28) is implied by the assumption in Lemma 3.6. In order to prove the converse, assume that (3.28) holds, and that there is a $0 < \theta_0 < \infty$ for which $\Lambda_U(\theta) = \infty$. Without loss of generality we can assume that $\Lambda_U(\theta)$ is finite for $\theta < \theta_0$ and infinite for $\theta \geq \theta_0$. For $x > \mathbb{E}[U]$, the Fenchel-Legendre transform is then given by

$$\Lambda_U^\star(x) = \sup_{0<\theta<\theta_0} (\theta x - \Lambda_U(\theta)).$$

Since $U \geq 0$ and $\Lambda_U(0) = 0$, we know that $\Lambda_U(\theta) \geq 0$ for $0 < \theta < \theta_0$, and hence

$$\frac{\Lambda_U^\star(x)}{x} \leq \theta_0,$$

which contradicts the assumption that this ratio tends to infinity as $x \to \infty$, and thus establishing the equivalence. ♢

Proof of Lemma 3.6. Let $\varphi^n \to \varphi$ for some sequence of $\varphi^n \in \Phi$ and $\varphi \in \Phi$. We introduce two families of random variables $\{Y_n\}$ and $\{Z_n\}$,

$$Y_n := \left(\frac{1}{n}\sum_{j=1}^{[n\psi_i]} U_j\right)_{i=1}^N, \quad Z_n := \left(\frac{1}{n}\sum_{j=1}^{[n\psi_i^n]} U_j\right)_{i=1}^N,$$

which have laws μ_n and ν_n, respectively, as in (3.13)–(3.14). Since $\varphi^n \to \varphi$ we know that for any $\varepsilon > 0$ there exists an M_ε such that for all $n > M_\varepsilon$ we have that the largest difference $\max_i \mid \varphi_i^n - \varphi_i| < \varepsilon/N$, and thus $|\psi_i^n - \psi_i| < \varepsilon$.

We have to show that for any $\delta > 0$

$$\limsup_{n\to\infty} \frac{1}{n} \log \mathbb{P}(||Y_n - Z_n||_\infty > \delta) = -\infty.$$

For $i \leq N$ consider the absolute difference between $Y_{n,i}$ and $Z_{n,i}$, i.e.,

$$|Y_{n,i} - Z_{n,i}| = \left|\frac{1}{n}\sum_{j=1}^{[n\psi_i]} U_j - \frac{1}{n}\sum_{j=1}^{[n\psi_i^n]} U_j\right|. \tag{3.29}$$

Next we have that for any $n > M_\varepsilon$ it holds that $|n\psi_i^n - n\psi_i| < n\varepsilon$, which yields the upper bound,

$$|[n\psi_i^n] - [n\psi_i]| < [n\varepsilon] + 2,$$

since the rounded numbers differ at most by 1 from their real counterparts. This means that the difference of the two sums in (3.29) can be bounded by at most $[n\varepsilon]+2$ elements of the U_j, which are for convenience denoted by $U_j^\star$. Recalling that the U_j are nonnegative, we obtain

$$\left| \frac{1}{n}\sum_{j=1}^{[n\psi_i]} U_j - \frac{1}{n}\sum_{j=1}^{[n\psi_i^n]} U_j \right| \leq \frac{1}{n}\sum_{j=1}^{[n\varepsilon]+2} U_j^\star.$$

Next we bound the probability that the difference exceeds δ, by using the above inequality:

$$\mathbb{P}\left(||Y_n - Z_n|| > \delta\right) \;\leq\; \mathbb{P}\left(\frac{1}{n}\sum_{j=1}^{[n\varepsilon]+2} U_j^\star > \delta\right) \;\leq\; \mathbb{E}\left[\exp\left(\theta U_1\right)\right]^{[n\varepsilon]+2} e^{-n\delta\theta},$$

where the last inequality follows from the Chernoff bound [DZ98, Eqn. (2.2.12)] for $\theta > 0$. Taking the log of this probability, dividing by n, and taking the limsup on both sides results in

$$\limsup_{n\to\infty} \frac{1}{n} \log \mathbb{P}\left(||Y_n - Z_n||_\infty > \delta\right) \leq \varepsilon\Lambda_U(\theta) - \delta\theta.$$

By the assumption, $\Lambda_U(\theta) < \infty$ for all θ. Thus, $\varepsilon \to 0$ yields

$$\limsup_{n\to\infty} \frac{1}{n} \log \mathbb{P}\left(||Y_n - Z_n||_\infty > \delta\right) \leq -\delta\theta.$$

As θ was arbitrary, the exponential equivalence follows by letting $\theta \to \infty$. □

We conclude this section with some examples.

Example 3.9. Assume that the loss amounts have finite support, say on the interval $[0, u]$. Then we clearly have

$$\Lambda_U(\theta) = \log \mathbb{E}\left[e^{\theta U}\right] \leq \theta u < \infty.$$

So for any distribution with finite support, the assumption for Lemma 3.6 is satisfied, and thus Theorem 3.4 holds. Here, the i.i.d. default times, τ_i, can have an arbitrary discrete distribution on the time grid $\{1, \ldots, N\}$.

In practical applications one (always) chooses a distribution with finite support for the loss amounts, since the exposure to every obligor is finite. Theorem 3.4 thus clearly holds for any (realistic) model of the loss given default.

An explicit expression for the rate function (3.15), or even the Fenchel-Legendre transform, is usually not available. On the other hand one can use numerical optimization techniques to calculate these quantities. ♢

We next present an example to which Lemma 3.6 applies.

Example 3.10. Assume that the loss amount U is measured in a certain unit, and takes on the values $u, 2u, \ldots$ for some $u > 0$. Assume that it has a distribution of Poisson type with parameter $\lambda > 0$, in the sense that for $i = 1, 2, \ldots$

$$\mathbb{P}\left(U = (i+1)u\right) = e^{-\lambda}\frac{\lambda^i}{i!}.$$

It is then easy to check that $\Lambda_U(\theta) = \theta u + \lambda\left(e^{\theta u} - 1\right)$, being finite for all θ. Further calculations yield

$$\Lambda_U^\star(x) = \left(\frac{x}{u} - 1\right)\log\left(\frac{1}{\lambda}\left(\frac{x}{u} - 1\right)\right) - \left(\frac{x}{u} - 1\right) + \lambda$$

for all $x > u$, and ∞ otherwise. Dividing this expression by x and letting $x \to \infty$, we observe that the resulting ratio goes to ∞. As a consequence, Remark 3.8 now entails that Theorem 3.4 applies. It can also be argued that for any distribution U with tail behavior comparable to that of a Poisson distribution, Theorem 3.4 applies as well. ♢

3.4 Exact Asymptotic Results

In the previous section we have established a sample-path large deviation principle on a finite time grid; this LDP provides us with logarithmic asymptotics of the probability that the sample path of $L_n(\cdot)/n$ is contained in a given set, say A. The results presented in this section are different in several ways. In the first place, we derive exact asymptotics (rather than logarithmic asymptotics). In the second place, our time domain is not assumed to be finite; instead we consider all integer numbers, $\mathbb{N}$. The price to be paid is that we restrict ourselves to special sets A, viz. those corresponding to the loss process (or the increment of the loss process) exceeding a given function. We work under the setup that we introduced in Section 3.2.1.

3.4.1 Crossing a Barrier

In this section we consider the asymptotic behavior of the probability that the loss process at some point in time is above a time-dependent level ζ. More precisely, we consider the set

$$A := \left\{f : T \to \mathbb{R}_0^+ \,|\, \exists t \in T:\; f(t) \geq \zeta(t)\right\}, \tag{3.30}$$

for some function $\zeta(t)$ satisfying

$$\zeta(t) > \mathbb{E}\left[UZ(t)\right] = \mathbb{E}\left[U\right]F_t \quad \text{for all } t \in T, \tag{3.31}$$

with F_t as in (3.6). If we would consider a function ζ that does not satisfy 3.31), we are not in a large-deviations setting, in the sense that the probability of the event $\{L_n(\cdot)/n \in A\}$ converges to 1 by the law of large numbers. In order to obtain a more interesting result, we thus limit ourselves to levels that satisfy (3.31). For such levels we state the first main result of this section.

Theorem 3.11. *Assume that*

$$\text{there is a unique } t^\star \in T \text{ such that } I_{UZ}(t^\star) = \min_{t\in T} I_{UZ}(t), \tag{3.32}$$

and that

$$\liminf_{t\to\infty} \frac{I_{UZ}(t)}{\log t} > 0, \tag{3.33}$$

where $I_{UZ}(t) = \sup_\theta \left\{\theta\zeta(t) - \Lambda_{UZ(t)}(\theta)\right\} = \Lambda^\star_{UZ(t)}(\zeta(t))$. *Then*

$$\mathbb{P}\left(\frac{1}{n}L_n(\cdot) \in A\right) = \frac{e^{-nI_{UZ}(t^\star)}C^\star}{\sqrt{n}}\left(1 + O\left(\frac{1}{n}\right)\right), \tag{3.34}$$

for A *as in* (3.30) *and* $\sigma^\star$ *is such that* $\Lambda'_{UZ(t^\star)}(\sigma^\star) = \zeta(t^\star)$. *The constant* $C^\star$ *follows from the Bahadur-Rao theorem (Theorem 3.24), with* $C^\star = C_{UZ(t^\star),\zeta(t^\star)}$.

Before proving our result, which will rely on arguments similar to those in [LM99], we first discuss the meaning and implications of Theorem 3.11. In addition we reflect on the role played by the assumptions. We do so by a sequence of remarks.

Remark 3.12. Comparing Theorem 3.11 to the Bahadur-Rao theorem (Theorem 3.24), we observe that the probability of a sample mean exceeding a rare value has the same type of decay as the probability of our interest (i.e., the probability that the normalized loss process $L_n(\cdot)/n$ ever exceeds some function ζ). This decay looks like $Ce^{-nI}/\sqrt{n}$ for positive constants C and I. This similarity can be explained as follows.

First observe that the probability of our interest is actually the probability of a *union* events. Evidently, this probability is larger than the probability of any of the events in this union, and hence also larger than the largest among these:

$$\mathbb{P}\left(\frac{1}{n}L_n(\cdot) \in A\right) \geq \sup_{t\in T} \mathbb{P}\left(\frac{1}{n}L_n(t) \geq \zeta(t)\right). \tag{3.35}$$

Theorem 3.11 indicates that the inequality in (3.35) is actually tight (under the conditions stated). Informally, this means that the contribution of the maximizing t in the right-hand side of (3.35), say $t^\star$, dominates the contributions of the other time epochs as n grows large. This essentially says that *given* that the rare event under consideration occurs, with overwhelming probability it happens at time $t^\star$. ♢

As is clear from the statement of Theorem 3.11, two assumptions are needed to prove the claim; we now briefly comment on the role played by these.

Remark 3.13. Assumption (3.32) is needed to make sure that there is not a time epoch $\bar{t}$, different from $t^\star$, having a contribution of the same order as $t^\star$. It can be verified from our proof that if the uniqueness assumption is not met, the probability under consideration remains asymptotically proportional to $e^{-nI}/\sqrt{n}$, but we lack a clean expression for the proportionality constant.

Assumption (3.33) has to be imposed to make sure that the contribution of the 'upper tail', that is, time epochs $t \in \{t^\star+1, t^\star+2, \ldots\}$, can be neglected; more formally, we should have

$$\mathbb{P}\left(\exists t \in \{t^\star+1, t^\star+2, \ldots\} : \frac{1}{n} L_n(t) \geq \zeta(t)\right) = o\left(\mathbb{P}\left(\frac{1}{n} L_n(\cdot) \in A\right)\right).$$

In order to achieve this, the probability that the normalized loss process exceeds ζ for large t should be sufficiently small. ♢

Remark 3.14. We now comment on what Assumption (3.33) means. Clearly,

$$\Lambda_{UZ(t)}(\theta) = \log\left(\mathbb{P}\left(\tau \leq t\right) \mathbb{E}\left[e^{\theta U}\right] + \mathbb{P}\left(\tau > t\right)\right) \leq \log \mathbb{E}\left[e^{\theta U}\right],$$

as t grows, $\theta \geq 0$; the limiting value is actually $\log \mathbb{E}\left[e^{\theta U}\right]$ if τ is non-defective. This entails that

$$I_{UZ}(t) = \Lambda^\star_{UZ(t^\star)}(\zeta(t)) \geq \Lambda^\star_U(\zeta(t)) = \sup_\theta \left(\theta\zeta(t) - \log \mathbb{E}\left[e^{\theta U}\right]\right).$$

We observe that Assumption (3.33) is fulfilled if $\liminf_{t\to\infty} \Lambda^\star_U(\zeta(t))/\log t > 0$, which turns out to be valid under extremely mild conditions. Indeed, relying on Lemma 3.20, we have that in great generality it holds that $\Lambda^\star_U(x)/x \to \infty$ as $x \to \infty$. Then clearly any $\zeta(t)$, for which $\liminf_t \zeta(t)/\log t > 0$, satisfies Assumption (3.33), since

$$\liminf_{t\to\infty} \frac{\Lambda^\star_U(\zeta(t))}{\log t} = \liminf_{t\to\infty} \frac{\Lambda^\star_U(\zeta(t))}{\zeta(t)} \frac{\zeta(t)}{\log t}.$$

Alternatively, if U is chosen distributed exponentially with mean λ (which does not satisfy the conditions of Lemma 3.20), then $\Lambda^\star_U(t) = \lambda t - 1 - \log(\lambda t)$, such that we have that

$$\liminf_{t\to\infty} \frac{I_U(\log t)}{\log t} = \lambda > 0.$$

Barrier functions ζ that grow at a rate slower than $\log t$, such as $\log\log t$, are in this setting clearly not allowed. ♢

Proof of Theorem 3.11. We start by rewriting the probability of interest as

$$\mathbb{P}\left(\frac{1}{n} L_n(\cdot) \in A\right) = \mathbb{P}\left(\exists t \in T : \frac{L_n(t)}{n} \geq \zeta(t)\right).$$

For an arbitrary point k in T we have

$$\mathbb{P}\left(\exists t \in T: \frac{L_n(t)}{n} \geq \zeta(t)\right) \leq \mathbb{P}\left(\exists t \leq k: \frac{L_n(t)}{n} \geq \zeta(t)\right) + \mathbb{P}\left(\exists t > k: \frac{L_n(t)}{n} \geq \zeta(t)\right). \tag{3.36}$$

We first focus on the second part in (3.36). We can bound this by

$$\begin{aligned}\mathbb{P}\left(\exists t > k: \frac{L_n(t)}{n} \geq \zeta(t)\right) &\leq \sum_{i=k+1}^{\infty} \mathbb{P}\left(\frac{L_n(i)}{n} \geq \zeta(i)\right) \\ &\leq \sum_{i=k+1}^{\infty} \inf_{\theta>0} \mathbb{E}\left[\exp\left(\theta \sum_{j=1}^{n} U_j Z_j(i)\right)\right] e^{-n\zeta(i)\theta},\end{aligned}$$

where the second inequality is due to the Chernoff bound [DZ98, Eqn. (2.2.12)]. The independence between the U_i and $Z_i(t)$, together with the assumption that the U_i are i.i.d. and the $Z_i(t)$ are i.i.d., yields

$$\begin{aligned}&\sum_{i=k+1}^{\infty} \inf_{\theta>0} \mathbb{E}\left[\exp\left(\theta \sum_{j=1}^{n} U_j Z_j(i)\right)\right] e^{-n\zeta(i)\theta} \\ &= \sum_{i=k+1}^{\infty} \inf_{\theta>0} \prod_{j=1}^{n} \mathbb{E}\left[\exp\left(\theta U_j Z_j(i)\right)\right] e^{-n\zeta(i)\theta} = \sum_{i=k+1}^{\infty} \exp\left(-n \sup_{\theta>0}\left(\zeta(i)\theta - \Lambda_{UZ(i)}(\theta)\right)\right) \\ &= \sum_{i=k+1}^{\infty} \exp\left(-n I_{UZ}(\zeta(i))\right).\end{aligned}$$

By (3.33) we have that

$$\liminf_{t\to\infty} \frac{I_{UZ}(t)}{\log t} = \beta,$$

for some $\beta > 0$ (possibly ∞). Hence there exists an m such that for all $i > m$

$$I_{UZ}(i) > \alpha \log i > I_{UZ}(t^\star), \tag{3.37}$$

where $\alpha = \beta/2$ (in case $\beta = \infty$, any $0 < \alpha < \infty$ suffices) and $t^\star$ defined in (3.32). Choosing $k = m$, we obtain by using the first inequality in (3.37) for $n > 1/\alpha$

$$\sum_{i=m+1}^{\infty} \exp\left(-n I_{UZ}(\zeta(i))\right) \leq \sum_{i=m+1}^{\infty} \exp\left(-n\alpha \log i\right) \leq \frac{1}{n\alpha - 1} \exp\left(\left(-n\alpha + 1\right) \log m\right),$$

where the last inequality trivially follows by bounding the summation (from above) by an appropriate integral. Next we multiply and divide this by $\mathbb{P}\left(L_n(t^\star)/n > \zeta(t^\star)\right)$

and we apply the Bahadur-Rao theorem, which results in

$$\frac{1}{n\alpha-1}e^{(-n\alpha+1)\log m} = \frac{1}{n\alpha-1}e^{(-n\alpha+1)\log m}\,\frac{\mathbb{P}\left(L_n(t^\star)/n > \zeta(t^\star)\right)}{\mathbb{P}\left(L_n(t^\star)/n > \zeta(t^\star)\right)}$$
$$= \mathbb{P}\left(\frac{1}{n}L_n(t^\star) > \zeta(t^\star)\right)\,\frac{m\sqrt{n}\,C^\star}{n\alpha-1}\,\left(1+O\left(\frac{1}{n}\right)\right)\,e^{-n(\alpha\log m - I_{UZ}(t^\star))}.$$

The second inequality in (3.37) yields $\alpha\log m - I_{UZ}(t^\star) > \delta$, for some $\delta > 0$. Applying this inequality, we see that this bounds the second term in (3.36), in the sense that as $n \to \infty$,

$$\mathbb{P}\left(\exists t > k:\ \frac{L_n(t)}{n} \geq \zeta(t)\right)\Big/\,\mathbb{P}\left(\frac{1}{n}L_n(t^\star) > \zeta(t^\star)\right) \to 0.$$

To complete the proof we need to bound the first term of (3.36), where we use that $k = m$. For this we again use the Bahadur-Rao theorem. Next to this theorem we use the uniqueness of $t^\star$, which implies that for $i \leq m$ and $i \neq t^\star$ there exists an $\varepsilon^\star > 0$, such that

$$I_{UZ}(t^\star) + \varepsilon^\star \leq I_{UZ}(i).$$

This observation yields, with σ_i such that $\Lambda'_{UZ(i)}(\sigma_i) = \zeta(i)$,

$$\mathbb{P}\left(\exists t \leq m:\ \frac{L_n(t)}{n} \geq \zeta(t)\right) \leq \sum_{i=1}^{m}\mathbb{P}\left(\frac{L_n(i)}{n} \geq \zeta(i)\right)$$
$$\leq \mathbb{P}\left(\frac{1}{n}L_n(t^\star) > \zeta(t^\star)\right)\,\left(1+O\left(\frac{1}{n}\right)\right)\,\left(\sum_{i=1}^{m}\frac{C^\star}{C_{UZ(i),\zeta(i)}}\,\frac{e^{-nI_{UZ}(t_i)}}{e^{-nI_{UZ}(t^\star)}}\right)$$
$$\leq \mathbb{P}\left(\frac{1}{n}L_n(t^\star) > \zeta(t^\star)\right)\,\left(1+O\left(\frac{1}{n}\right)\right)\,\left(1 + m\times\max_{i=1,\ldots,m}\left(\frac{C^\star}{C_{UZ(i),\zeta(i)}}\right)e^{-n\varepsilon^\star}\right)$$
$$= \mathbb{P}\left(\frac{1}{n}L_n(t^\star) > \zeta(t^\star)\right)\,\left(1+O\left(\frac{1}{n}\right)\right)\,\left(1+O\left(e^{-n\varepsilon^\star}\right)\right)$$

Combining the above findings, we observe

$$\mathbb{P}\left(\exists t \in T:\ \frac{L_n(t)}{n} \geq \zeta(t)\right) \leq \mathbb{P}\left(\frac{L_n(t^\star)}{n} \geq \zeta(t^\star)\right)\left(1+O\left(\frac{1}{n}\right)\right).$$

Together with the trivial bound

$$\mathbb{P}\left(\exists t \in T:\ \frac{L_n(t)}{n} \geq \zeta(t)\right) \geq \mathbb{P}\left(\frac{L_n(t^\star)}{n} \geq \zeta(t^\star)\right),$$

this yields

$$\mathbb{P}\left(\exists t \in T:\ \frac{L_n(t)}{n} \geq \zeta(t)\right) = \mathbb{P}\left(\frac{L_n(t^\star)}{n} > \zeta(t^\star)\right)\left(1+O\left(\frac{1}{n}\right)\right).$$

Applying the Bahadur-Rao theorem to the right-hand side of the previous display yields the desired result. □

3.4.2 Large Increments of the Loss Process

In the previous section we identified the asymptotic behavior of the probability that at some point in time the normalized loss process $L_n(\cdot)/n$ exceeds a certain level. We can carry out a similar procedure to obtain insight in the large deviations of the *increments* of the loss process. Here we consider times where the increment of the loss between time s and t exceeds a threshold $\xi(s,t)$. More precisely, we consider the event

$$A := \left\{ f : T \times T \to \mathbb{R}_0^+ \,|\, \exists s < t \in T :\ f(s,t) \geq \xi(s,t) \right\}. \tag{3.38}$$

Being able to deal with events of this type, we can for instance analyze the likelihood of the occurrence of a large loss during a short period; we remark that with the event (3.30) from the previous subsection, one cannot distinguish the cases where the loss is zero for all times before t and $x > \zeta(t)$ at time t, and the case where the loss is just below the level ζ for all times before time t and then ends up at x at time t. Clearly, events of the (3.38) make it possible to distinguish between such paths.

In order to avoid trivial results, we impose a condition similar to (3.31), namely

$$\xi(s,t) > \mathbb{E}\left[U\right]\ (F_t - F_s)\,, \tag{3.39}$$

for all $s < t$. The law of large numbers entails that for functions ξ that do not satisfy this condition, the probability under consideration does not correspond to a rare event.

A similar probability has been considered in [DDD04], where the authors derive the logarithmic asymptotic behavior of the probability that the increment of the loss, for some $s < t$, in a bounded interval exceeds a thresholds that depends only on $t - s$. In contrast, our approach uses a more flexible threshold, which depends on both times s and t, and in addition we derive the *exact* asymptotic behavior of this probability.

Theorem 3.15. *Assume that*

$$\text{there is a unique } s^\star < t^\star \in T \text{ such that } I_{UZ}(s^\star, t^\star) = \min_{s<t} I_{UZ}(s,t), \tag{3.40}$$

and that

$$\inf_{s\in T} \liminf_{t\to\infty} \frac{I_{UZ}(s,t)}{\log t} > 0, \tag{3.41}$$

where $I_{UZ}(s,t) = \sup_\theta \left(\theta\xi(s,t) - \Lambda_{U(Z(t)-Z(s))}(\theta)\right) = \Lambda^\star_{U(Z(t)-Z(s))}(\xi(s,t))$. *Then*

$$\mathbb{P}\left(\frac{1}{n}L_n(\cdot) \in A\right) = \frac{e^{-nI_{UZ}(s^\star,t^\star)}C^\star}{\sqrt{n}}\left(1 + O\left(\frac{1}{n}\right)\right), \tag{3.42}$$

for A as in (3.39) and $\sigma^\star$ is such that $\Lambda'_{U(Z(t^\star)-Z(s^\star))}(\sigma^\star) = \xi(s^\star, t^\star)$. The constant $C^\star$ follows from the Bahadur-Rao theorem (Theorem 3.24), with $C^\star = C_{U(Z(t^\star)-Z(s^\star)),\ \xi(s^\star,t^\star)}$.

Remark 3.16. A first glance at Theorem 3.15 tells us the obtained result is very similar to the result of Theorem 3.11. The second condition, i.e., Inequality (3.41), however, seems to be more restrictive than the corresponding condition, i.e., Inequality (3.33), due to the infimum over s. This assumption has to make sure that the 'upper tail' is negligible for any s. In the previous subsection we have seen that, under mild restrictions, the upper tail can be safely ignored when the barrier function grows at a rate of at least $\log t$. We can extend this claim to our new setting of large increments, as follows.

First note that

$$\inf_{s\in T} \liminf_{t\to\infty} \frac{I_{UZ}(s,t)}{\log t} \geq \inf_{s\in T} \liminf_{t\to\infty} \frac{\Lambda_U^\star(\xi(s,t))}{\log t}.$$

Then consider thresholds that, next to condition (3.39), satisfy that for all s

$$\liminf_{t\to\infty} \frac{\xi(s,t)}{\log t} > 0. \tag{3.43}$$

Then, under the conditions of Lemma 3.20, we have that

$$\liminf_{t\to\infty} \frac{\Lambda_U^\star(\xi(s,t))}{\log t} = \liminf_{t\to\infty} \frac{\Lambda_U^\star(\xi(s,t))}{\xi(s,t)} \frac{\xi(s,t)}{\log t} = \infty, \tag{3.44}$$

since the second term remains positive by (3.43) and the first term tends to infinity by Lemma 3.20. Having established (3.44) for all s, it is clear that (3.41) is satisfied.

The sufficient condition (3.43) shows that the range of admissible barrier functions is quite substantial, and, importantly, imposing (3.41) is not as restrictive as it seems at first glance. ♢

Proof of Theorem 3.15. The proof of this theorem is very similar to that of Theorem 3.11. Therefore we only sketch the proof here. As before, the probability of interest is split up into a 'front part' and 'tail part'. The tail part can be bounded using Assumption (3.41); this is done analogously to the way Assumption (3.33) was used in the proof of Theorem 3.11. The uniqueness assumption (3.40) then shows that the probability of interest is asymptotically equal to the probability that the increment between time $s^\star$ and $t^\star$ exceeds $\xi(s^\star, t^\star)$; this is an application of the Bahadur-Rao theorem. Another application of the Bahadur-Rao theorem to the probability that the increment between time $s^\star$ and $t^\star$ exceeds $\xi(s^\star, t^\star)$ yields the result. □

3.5 Discussion and Concluding Remarks

In this chapter we have established a number of results with respect to the asymptotic behavior of the distribution of the loss process. In this section we discuss some of the assumptions in more detail and we consider extensions of the results that we have derived.

3.5.1 Extensions of the Sample-Path LDP

The first part of our work, Section 3.3, was devoted to establishing a sample-path large deviation principle on a finite time grid. Here we modeled the loss process as the sum of i.i.d. loss amounts multiplied by i.i.d. default indicators. From a practical point of view one can argue that the assumptions underlying our model are not always realistic. In particular, the random properties of the obligors cannot always be assumed independent. In addition, the assumption that all obligors behave in an i.i.d. fashion will not necessarily hold in practice. Both shortcomings can be dealt with, however, by adapting the model slightly. A common way to introduce dependence, taken from [DDD04], is by supposing that there is a 'macro-environmental' variable Y to which all obligors react, but conditional on which the loss epochs and loss amounts are independent. First observe that our results are then valid for any specific realization y of Y. Denoting the exponential decay rate by r_y, i.e.,

$$\lim_{n\to\infty} \frac{1}{n} \log \mathbb{P}\left(\frac{1}{n} L_n(\cdot) \in A \,\middle|\, Y = y \right) = r_y,$$

the *unconditional* decay rate is just the maximum over the r_y; this is trivial to prove if Y can attain values in a finite set only. A detailed treatment of this is beyond the scope of this chapter.

The assumption that all obligors have the same distribution can be relaxed to the case where we assume that there are m different classes of obligors (for instance determined by their default ratings). We further assume that each class i makes up a fraction a_i of the entire portfolio. Then we can extend the LDP of Theorem 3.3.2 to a more general one, by splitting up the loss process into m loss processes, each corresponding to a class. Conditioning on the realizations of these processes, we can derive the following rate function:

$$I_{U,p,m}(x) := \inf_{\varphi\in\Phi^m} \inf_{v\in V_x} \sum_{j=1}^{m} \sum_{i=1}^{N} a_i \varphi_i^j \left(\log\left(\frac{\varphi_i^j}{p_i^j} \right) + \Lambda_U^\star \left(\frac{v_i^j}{a_i \varphi_i^j} \right) \right), \tag{3.45}$$

where $V_x = \{ v \in \mathbb{R}_+^{m\times N} \mid \sum_{j=1}^m v_i^j = \Delta x_i \text{ for all } i \leq N \}$, and Φ^m is the Cartesian product $\Phi \times \ldots \times \Phi$ (m times), with Φ as in (3.11). The optimization over the set V_x follows directly from conditioning on the realizations of the per-class loss processes. We leave out the formal derivation of this result; this multi-class case is notationally considerably more involved than the single-class case, but essentially all steps carry over.

In our sample-path LDP we assumed that defaults can only occur on a finite grid. While this assumption is justifiable from a practical point of view, an interesting mathematical question is whether it can be relaxed. In self-evident notation, one would expect that the rate function

$$I_{U,p,\infty}(x) := \inf_{\varphi\in\Phi_\infty} \sum_{i=1}^{\infty} \varphi_i \left(\log\left(\frac{\varphi_i}{p_i} \right) + \Lambda_U^\star \left(\frac{\Delta x_i}{\varphi_i} \right) \right).$$

It can be checked, however, that the argumentation used in the proof of Theorem 3.4 does not work; in particular, the choice of a suitable topology plays an important role.

If losses can occur on a continuous entire interval, i.e., $[0, N]$, we expect, for a nondecreasing and differentiable path x, the rate function

$$I_{U,p,[0,N]}(x) := \inf_{\varphi \in \mathscr{M}} \int_0^N \varphi(t) \left(\log \left(\frac{\varphi(t)}{p(t)} \right) + \Lambda_U^\star \left(\frac{x'(t)}{\varphi(t)} \right) \right) \mathrm{d}t, \tag{3.46}$$

where $\mathscr{M}$ is the space of all densities on $[0, N]$ and p the density of the default time τ. One can easily guess the validity of (3.46) from (3.15) by using Riemann sums to approximate the integral. A formal proof, however, requires techniques that are essentially different from the ones used to establish Theorem 3.4, and therefore we leave this for future research.

3.5.2 Extensions of the Exact Asymptotics

In the second part of the chapter, i.e., Section 3.4, we have derived the exact asymptotic behavior for two special events. First we showed that, under certain conditions, the probability that the loss process exceeds a certain time-dependent level, is asymptotically equal to the probability that the process exceeds this level at the 'most likely' time $t^\star$. The exact asymptotics of this probability are obtained by applying the Bahadur-Rao theorem. A similar result has been obtained for an event related to the *increment* of the loss process. One could think of refining the logarithmic asymptotics, as developed in Section 3.3, to exact asymptotics. Note, however, that this is far from straightforward, as for general sets these asymptotics do not necessarily coincide with those of a univariate random variable, cf. [MMNvU06].

3.A Background Results

In this section we state a number of definitions and results, taken from [DZ98], which are used in the proofs in this chapter.

Theorem 3.17 (Cramér)**.** *Let X_i be i.i.d. real valued random variables with all exponential moments finite and let μ_n be the law of the average $S_n = \sum_{i=1}^n X_i/n$. Then the sequence $\{\mu_n\}$ satisfies an LDP with rate function $\Lambda^\star(\cdot)$, where $\Lambda^\star$ is the Fenchel-Legendre transform of the X_i.*

Proof. See for example [DZ98, Thm. 2.2.3]. □

Definition 3.18. *We say that two families of measures $\{\mu_n\}$ and $\{\nu_n\}$ on a metric space $(\mathcal{X}, d)$ are exponentially equivalent if there exist two families of $\mathcal{X}$-valued random variables $\{Y_n\}$ and $\{Z_n\}$ with marginal distributions $\{\mu_n\}$ and $\{\nu_n\}$, respectively, such that for all $\delta > 0$*

$$\limsup_{n\to\infty} \frac{1}{n} \log \mathbb{P}\left(d\left(X_n, Y_n\right) \geq \delta\right) = -\infty.$$

Lemma 3.19. *For every triangular array $a_n^i \geq 0$, $n \geq 1$, $1 \leq i \leq n$,*

$$\limsup_{n\to\infty} \frac{1}{n} \log \sum_{i=1}^{n} a_n^i = \limsup_{n\to\infty} \max_{i=1,\ldots,n} \frac{1}{n} \log a_n^i.$$

Proof. Elementary, but also a direct consequence of [DZ98, Lemma 1.2.15]. □

Lemma 3.20. *Let $\Lambda(\theta) < \infty$ for all $\theta \in \mathbb{R}$, then*

$$\lim_{|x|\to\infty} \frac{\Lambda^\star(x)}{|x|} = \infty.$$

Proof. This result is a part of [DZ98, Lemma 2.2.20]. □

Lemma 3.21. *Let $K_{n,i}$ be defined as $K_{n,j} := \#\{i \in \{1,\ldots,n\} \mid \tau_i = j\}$. Then for any vector $k \in \mathbb{N}^N$, such that $\sum_{i=1}^N k_i = n$, we have that*

$$(n+1)^{-N} \exp\left(-nH(k \mid p)\right) \leq \mathbb{P}\left(K_n = k\right) \leq \exp\left(-nH(k \mid p)\right),$$

where

$$H(k \mid p) = \sum_{i=1}^{N} \frac{k_i}{n} \log\left(\frac{k_i}{np_i}\right),$$

and p_i as defined in (3.5).

Proof. See [DZ98, Lemma 2.1.9]. □

Lemma 3.22. *Define*

$$Z_n(t) := \frac{1}{n} \sum_{i=1}^{[nt]} X_i,\ 0 \leq t \leq 1,$$

for an i.i.d. sequence of $\mathbb{R}^d$ valued random variables X_i. Let μ_n denote the law of $Z_n(\cdot)$ in $L_\infty([0,1])$. For any discretization $J = \{0 < t_1 < \ldots < t_{|J|} \leq 1\}$ and any $f : [0,1] \to \mathbb{R}^d$, let $p_J(f)$ denote the vector $(f(t_i))_{i=1}^{|J|} \in (\mathbb{R}^d)^{|J|}$. Then the sequence of laws $\{\mu_n \circ p_J^{-1}\}$ satisfies the LDP in $(\mathbb{R}^d)^{|j|}$ with the good rate function

$$I_J(z) = \sum_{i=1}^{|J|} (t_i - t_{i-1}) \Lambda^\star\left(\frac{z_i - z_{i-1}}{t_i - t_{i-1}}\right),$$

where $\Lambda^\star$ is the Fenchel-Legendre transform of X_1.

Proof. See [DZ98, Lemma 5.1.8]. This lemma is one of the key steps in proving Mogul'skiĭ's theorem, which provides a sample-path LDP for $Z_n(\cdot)$ on a bounded interval. □

Theorem 3.23. *If an LDP with a good rate function $I(\cdot)$ holds for the probability measures $\{\mu_n\}$, which are exponentially equivalent to $\{\nu_n\}$, then the same LDP holds for $\{\nu_n\}$.*

Proof. See [DZ98, Thm. 4.2.13]. □

Theorem 3.24 (Bahadur-Rao)**.** *Let X_i be an i.i.d. real-valued sequence random variables. Then we have*

$$\mathbb{P}\left(\frac{1}{n}\sum_{i=1}^{n} X_i \geq q\right) = \frac{e^{-n\Lambda_X^\star(q)}C_{X,q}}{\sqrt{n}}\left(1 + O\left(\frac{1}{n}\right)\right).$$

The constant $C_{X,q}$ depends on the type of distribution of X_1, as specified by the following two cases.

(i) The law of X_1 is lattice, i.e. for some x_0, d, the random variable $(X_1 - x_0)/d$ is (a.s.) an integer number, and d is the largest number with this property. Under the additional condition $0 < \mathbb{P}(X_1 = q) < 1$, the constant $C_{X,q}$ is given by

$$C_{X,q} = \frac{d}{(1 - e^{-\sigma d})\sqrt{2\pi\Lambda_X''(\sigma)}},$$

where σ satisfies $\Lambda_X'(\sigma) = q$.

(ii) If the law of X_1 is non-lattice, the constant $C_{X,q}$ is given by

$$C_{X,q} = \frac{1}{\sigma\sqrt{2\pi\Lambda_X''(\sigma)}},$$

with σ as in case (i).

Proof. We refer to [BR60] or [DZ98, Thm. 3.7.4] for the proof of this result. □

CHAPTER 4

An Empirical Analysis of CDO Data

In Chapters 2 and 3 we have presented two models for the loss process of a credit risky portfolio, where the focus was mainly on the modeling itself. This, and the next, chapter are of a more applied nature. In the current chapter, which is based on [LdVV08], we perform an empirical analysis on a large data set of CDO quotes. By means of a regression analysis, we filter out the credit risk that is present in each CDO tranche, to investigate the effects of correlation. Thereafter we consider three different ways in which one can use a set of base correlations, obtained from market quotes for CDO tranches, to value non-standard CDO tranches.

4.1 Introduction

Over the last decade, the growth in the credit derivatives market has been enormous. First, the market for credit default swaps (CDS) on single names has become very liquid. This was followed by the rapid development of the synthetic CDO market over the last couple of years. Following this large growth in liquidity there has been a need for models which can be calibrated to market quotes. Amongst practitioners, the one-factor Gaussian factor copula has become standard. In order to allow for calibration an ad-hoc method has become popular, known as the base correlation approach, presented in [MBAW04]. This method can not be regarded as a model explaining the observed market quotes, but rather as a fix of the standard one-factor Gaussian copula model, much like implied volatilities in the equity derivatives markets. When considering non-standard tranches, i.e. tranches with non-standard attachment and detachment levels, tranches with non-standard maturities or tranches on non-standard indices, some mapping choices have to be made. There are multiple ways to map the observed base correlation skew to different maturities or different baskets and three of them are considered in this paper.

In this chapter we perform an empirical analysis of market prices for CDO tranches, using a data set ranging from December 2004 up to November 2006. The data set consists of three different sets of quotes. Quotes of standard interest rate products are available and these are used to determine risk free discount factors. For each reference entity entire CDS term structures are available, which are used to determine marginal distributions of the default times. Finally quotes for synthetic CDO tranches are available and these are used to determine the market implied default dependency structure. Using only the data for synthetic CDO tranche quotes we first investigate whether correlation effects can be discovered from these quotes directly. Therefore the influence of the credit risk of the underlying index is filtered out by means of regression analysis. The resulting residuals should then capture correlation effects.

The second part of the analysis uses the base correlation method to imply the default dependency structure from the market. First the base correlation method is used to investigate the behavior of market implied credit correlation over time. Secondly, three different mapping methods are considered and their performance is investigated by means of out-of-sample tests. An out-of-sample test is considered where tranche quotes with five year maturity on a certain index are mapped to tranches with either seven or ten year maturities on the same index. In addition out-of-sample tests are performed where mapping is done using quotes for five year tranches on one index and mapping the resulting base correlations to five, seven and ten year tranches on the other index.

The outline of this chapter is as follows. In the next section the synthetic CDO market is discussed. The mechanics of the products are discussed and the quoting convention is explained. In addition it is shown that pricing of these products can be reduced to determining the expected loss on equity tranches. The third section focuses on modeling the default dependency by means of copulas and the one-factor Gaussian copula in particular. In addition the base correlation method is described, which has become the industry standard approach for visualizing default correlation implied in the market. Section 4.4 describes the data which has been used in the empirical analysis and some summary statistics are presented. In addition the increase in liquidity in the synthetic CDO market is briefly touched upon. Section 4.5 presents the results of the empirical analysis. First, the synthetic CDO tranche quotes are considered directly without using any default correlation model. The second part of this section applies the base correlation method to the available data. First the behavior of the different base correlations with a five year maturity is considered. Second the base correlation method is used in combination with different mapping methods in order to investigate the performance of three different mapping methods by means of out-of-sample tests. Finally, Section 4.6 concludes this chapter.

4.2 Synthetic CDO Tranches

In a Collateralized Debt Obligation, CDO, a pool of assets functions as collateral against which notes are issued. These notes represent tranches of the structure and strictly regulate the seniority of the notes. In Figure 4.1, the structure of a CDO contract is graphically represented.

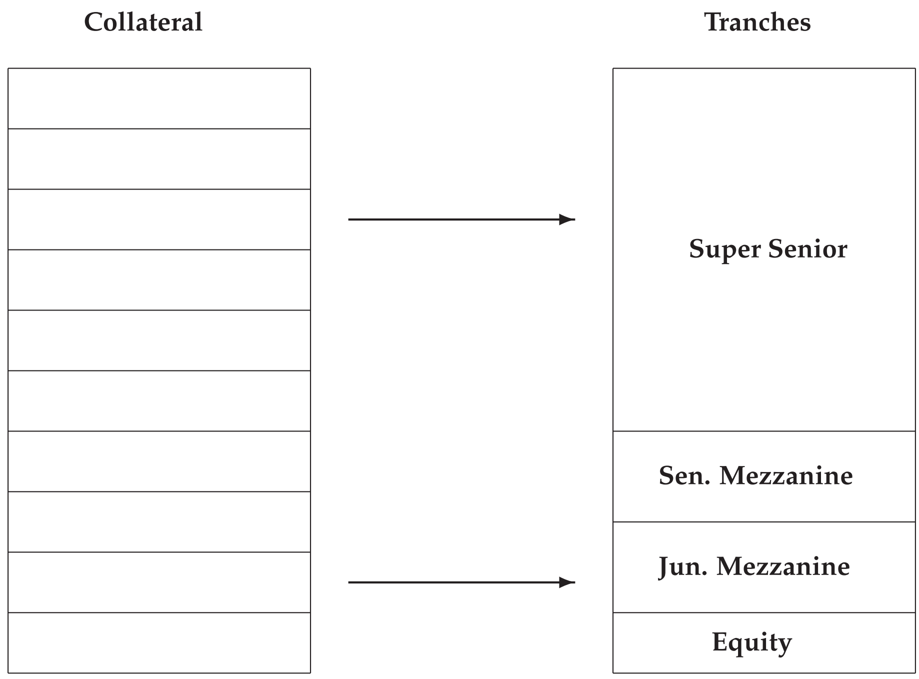


Figure 4.1: CDO structure with the pool of collateral assets on the left-hand-side and four different tranches on the right hand side The risky equity tranche is shown at the bottom of the CDO structure and is followed by a junior mezzanine tranche, a senior mezzanine tranche and a super senior tranche.

Here on the left hand side one can see the pool of collateral, while on the right hand side the CDO structure is represented where four different tranches are issued, the equity, the junior mezzanine, the senior mezzanine and the super senior tranche. In case of a default event in the pool of collateral the first losses are absorbed by the equity tranche investor. Once the cumulative losses increase beyond the size of the equity tranche, the junior mezzanine tranche is affected and so forth. Clearly the equity tranche investor is much more exposed to losses in the pool of collateral and therefore this investor is compensated by receiving a larger coupon. Rating agencies can be involved to assign credit ratings to the different tranches. Usually the equity tranche is not rated, due to the large risk involved. As the name suggests these tranches are similar to investing in equity, while the more senior notes can be compared to investing in bonds.

Traditionally these CDO tranches are backed by a pool of cash assets, such as bonds, student loans or credit card receivables. Due to issues such as different maturities and prepayment of assets such structures usually require a manager to manage the cash flows and invest in new assets. It was realized that one could also create the pool of collateral synthetically by means of credit default swaps. The clear advantage here is that one can select CDSs with equal maturity leading to easy to predict cash flows which are only affected by a default event of one of the underlying reference entities.

More information about different types of CDOs can be found in [Tav03].

4.2.1 Standard Tranches and Quoting Conventions

In recent years the market for synthetic CDO tranches has become increasingly liquid. Two reference indices have been defined, known as CDX NA IG and iTraxx Europe, which contain 125 North American reference entities and 125 European reference entities, respectively. All these names are high grade issuers and on 20 March and 20 September of each year the indices are rolled, leading to new series. At a roll date some names for which credit worthiness has deteriorated are possibly replaced by names with better credit quality. In addition to a possible change of index composition, at a roll date the maturity of quoted synthetic CDO tranches also changes with six months. Standard maturities are at 20 June and 20 December of each year. For instance, after an index roll at 20 March, the new on the run tranches reference the new index and have maturities at 20 June. Thus initially the five year quoted CDO tranches have a tenor of five years and three months. Just before the next roll date the tenor has reduced to four years and nine months.

Five tranches on the iTraxx index are quoted in the market with attachment and detachment levels: 0-3%, 3-6%, 6-9%, 9-12% and 12-22%. For the tranches on the CDX index detachment levels are slightly different at: 0-3%, 3-7%, 7-10%, 10-15% and 15-30%. All tranches have quarterly payments and the Actual/360 daycount convention is used to determine premium payments. For the equity tranche the quote is given as an percentage of the notional to be paid upfront by the protection buyer, in addition to premium payments of 500 basis points (bp) on an annual basis, paid quarterly. Typically the five tranches are quoted in combination with a notional delta exchange. This means that a protection buyer can buy protection for a certain tranche in return for the quoted premium and in addition will sell protection on the index credit default swap with notional equal to the tranche notional times the notional delta exchange. The reason for this delta exchange is that both parties are likely to hedge their credit delta risk. Trading the delta hedge with each other, in combination with the tranche, is thus simple and avoids extra costs due to the bid/offer spread. The quotes also consist of the index spread against which the notional delta exchange part of the trade is executed and this is known as the reference spread. Both reference spread and delta exchanges are agreed periodically by a dealer poll, where the reference spread follows the index spread closely.

Table 4.1 gives an example of a set of synthetic CDO quotes on the iTraxx index with maturity of five years.

iTraxx	Tranche	Bid	Offer	Delta
Ref. 24 BP	0 – 3	11.625	11.875	25.5
	3 – 6	54.5	54.75	5
	6 – 9	13.25	14.25	1.5
	9 – 12	5	6	0.5
	12 – 22	2	3	0.3

Table 4.1: Market quotes for the five standard synthetic CDO tranches on iTraxx, with a tenor of five years. For each tranche the bid and offer quotes are shown as well as the notional delta exchange. In addition a reference spread is shown. Quotes are from the 1^{st} of November 2006 and thus maturity for the quoted tranches is December 20^{th}, 2011.

Thus for instance an investor can sell credit protection on the 3-6% tranche for 1 million notional, receiving 54.5 bp. Simultaneously the investor buys protection on the index credit default swap for 5 million notional against a spread of 24 bp.

The synthetic CDO market on these indices started with quoting for a five year tenor only. In recent years the range of maturities has expanded to include seven and ten year maturities as well. However the mayor part of the analysis presented in this chapter focuses on tranches with a five year maturity as these are most liquid.

4.2.2 Pricing of Synthetic CDO Tranches

In order to price a synthetic CDO tranche the focus will be on the distribution of the cumulative loss of a certain basket, consisting of a number of CDSs, at future points in time. Let $L^{\mathcal{B}}(t)$ denote the percentage loss on basket $\mathcal{B}$ at time t. The percentage loss on an equity tranche, i.e. a tranche with zero attachment, and detachment level d is given by:

$$L_d^{\mathcal{B}}(t) = \frac{1}{d} \min\left(L^{\mathcal{B}}(t), d\right).$$

Next, a general tranche is considered, having both a non-zero attachment as well as a detachment level. The percentage loss of a tranche with attachment a and detachment

d is given by:

$$\begin{aligned} L^{\mathcal{B}}_{a,d}(t) &= \frac{\min\left(\max\left(L^{\mathcal{B}}(t)-a,0\right),d-a\right)}{d-a} \\ &= \frac{\min\left(L^{\mathcal{B}}(t),d\right)-\min\left(L^{\mathcal{B}}(t),a\right)}{d-a} \\ &= \frac{d\cdot L^{\mathcal{B}}_{d}(t)-a\cdot L^{\mathcal{B}}_{a}(t)}{d-a}. \end{aligned}$$

Thus the loss on any tranche can always be written in terms of the loss on equity tranches. the modeling of the tranche loss will focus on the modeling of the cumulative loss of the basket, $L^{\mathcal{B}}$. Once a certain model is chosen, the focus is on the calculation of the expected loss on a certain equity tranche at future times, i.e. $\mathbb{E}\left[L^{\mathcal{B}}_{d}(t)\right]$. This expected loss is calculated under the risk neutral measure as the model for the cumulative loss will be consistent with the market observed values for CDSs. The expected loss needs to be evaluated for different times in the future due to discounting and due to premium payments which are made over outstanding tranche notional, which is expected to decrease over time.

First, the value of the protection leg of an equity tranche is considered using a discrete approximation over time. Let $t_0,\ t_1,\ldots,T_M=T$ denote a partition of the time interval, where t_0 denotes the valuation date and t_M the maturity. The value, $\mathrm{DL}^{\mathcal{B},T}_{d}$, of the protection leg of an equity tranche with detachment d on basket $\mathcal{B}$ with maturity T can be approximated using the following expression:

$$\begin{aligned} \mathrm{DL}^{\mathcal{B},T}_{d} &\approx \mathbb{E}\left[\sum_{i=1}^{M} D\left(t_i^{\star}\right)\left(L^{\mathcal{B}}_{d}(t_i)-L^{\mathcal{B}}_{d}(t_{i-1})\right)\right] \\ &= \sum_{i=1}^{M} D\left(t_i^{\star}\right)\left(\mathbb{E}\left[L^{\mathcal{B}}_{d}(t_i)\right]-\mathbb{E}\left[L^{\mathcal{B}}_{d}(t_{i-1})\right]\right). \end{aligned} \tag{4.1}$$

Here $D(t)$ denotes the value at the valuation date of a risk free payment of one unit made at time t. In the expression, we assume that all protection payments during the period t_{i-1} up to t_i are made at the middle of the period. $t_i^{\star}$. Similarly one can derive the value, $\mathrm{PL}^{\mathcal{B},T}_{d}$, of the premium leg. Here we consider the case where premium payments are made over the outstanding notional of the tranche, which is the most commonly traded variation of synthetic CDO tranche swaps. The premium payment dates, $T_i, i=1,\ldots,N$, can be used for discretization. Furthermore we let T_0 denote the valuation date. The value of the protection leg is then given by

$$\begin{aligned}
\mathrm{PL}_d^{\mathcal{B},T} &\approx \mathbb{E}\left[\sum_{i=1}^{N} \delta_i D\left(T_i\right) \tfrac{1}{2}\left(1 - L_d^{\mathcal{B}}(T_i) + 1 - L_d^{\mathcal{B}}(T_{i-1})\right)\right] \\
&= \sum_{i=1}^{N} \delta_i D\left(T_i\right)\left(1 - \tfrac{1}{2}\left(\mathbb{E}\left[L_d^{\mathcal{B}}(T_i)\right] + \mathbb{E}\left[L_d^{\mathcal{B}}(T_{i-1})\right]\right)\right). \qquad (4.2)
\end{aligned}$$

Here the term δ_i denotes the day count fraction which applies to the premium payment to be made at time t_i. There are some important differences when comparing Equations (4.1) and (4.2). First, the discretization for the protection leg is general, while for the premium leg the payment dates have been used. Second, discounting for the protection leg is done in the middle of a period, while for the premium leg the end of the period is used, which corresponds to the payment date.

Due to the simple linear relationships it is straightforward to derive the following relationships for general tranches:

$$\begin{aligned}
\mathrm{DL}_{a,d}^{\mathcal{B},T} &= \frac{d \cdot \mathrm{DL}_d^{\mathcal{B},T} - a \cdot \mathrm{DL}_a^{\mathcal{B},T}}{d-a} \\
\mathrm{PL}_{a,d}^{\mathcal{B},T} &= \frac{d \cdot \mathrm{PL}_d^{\mathcal{B},T} - a \cdot \mathrm{PL}_a^{\mathcal{B},T}}{d-a}. \qquad (4.3)
\end{aligned}$$

As the value of the protection leg is defined for one unit of tranche notional, one has to weigh using the attachment and detachment levels of the contract. This relation shows that one can write both protection and premium leg values of any tranche in terms of two equity tranches. This relationship is the main observation used in the base correlation model discussed in Section 4.3.2. Given the values of the protection leg and premium leg, and the spread s, one could easily determine the value, $V_{a,d}^{\mathcal{B},T}$, of a CDO with contracted premium s^c, or one could determine the fair premium s^f of that tranche:

$$\begin{aligned}
V_{a,d}^{\mathcal{B},T} &= \mathrm{DL}_{a,d}^{\mathcal{B},T} - s^c \cdot \mathrm{PL}_{a,d}^{\mathcal{B},T} \\
s^f &= \frac{\mathrm{DL}_{a,d}^{\mathcal{B},T}}{\mathrm{PL}_{a,d}^{\mathcal{B},T}}. \qquad (4.4)
\end{aligned}$$

Here the value is given from the point of view of the protection buyer. In case an upfront premium payment is specified, one could easily determine the value by subtracting this upfront payment. Note that this is only relevant at initiation of a contract.

4.3 Modeling Default Dependency

Default dependency is usually modeled using copula methods, see [Nel99] for more details on copulas. The main advantage of copulas is that it allows one to split modeling of marginal distributions from the modeling of the dependency structure. The one-factor Gaussian copula is a special case of the more general Gaussian copula, presented in [Li00], where some additional structure is enforced which simplifies the model.

4.3.1 One-Factor Gaussian Copula

The market standard model has become the one-factor Gaussian copula. Default events are modeled by means of latent variables and default dependency is modeled by correlating the latent variables in some specific way. For the one-factor Gaussian copula model, the latent variables are modeled as follows:

$$X_i \equiv \rho_i \cdot Y + \sqrt{1-\rho_i^2} \cdot \xi_i \tag{4.5}$$
$$\tau_i \leq t \Leftrightarrow X_i \leq \chi_i(t).$$

Here τ_i denotes the default time for reference entity i. X_i denotes the latent variable for name i and a default event has occurred before time t, when the latent variable is smaller than some time dependent threshold function $\chi_i(t)$. The latent variables of the different names are correlated through the common factor, Y, which is assumed to have a standard normal distribution. The second term depends on the idiosyncratic random variable ξ_i, which also has a standard normal distribution. All the random normals, i.e. Y and ξ_i are independent. The latent variables X_i have a standard normal distribution as well. In order to make sure that the distribution of default times are calibrated to the market, one must use the distribution of these latent variables to determine the thresholds $\chi_i(t)$:

$$\chi_i(t) = \Phi^{-1}\left(p_i(t)\right). \tag{4.6}$$

Here $p_i(t)$ is the marginal default probability for name i for the period up to time t and these are obtained by bootstrapping to CDS quotes from the market. Φ denotes the cumulative distribution function for the standard normal distribution.

One can observe that the latent variables X_i and X_j, for $i \neq j$, as defined in Equation (4.5) have correlation $\rho_i \cdot \rho_j$. In practice, due to the large number of names to model, the parameters ρ_i are often chosen to be constant over all names, i.e. $\rho_i = \rho$ for all i. In this case the default dependency structure is driven by a single correlation parameter.

As discussed in Section 4.2.2 the valuation of synthetic CDO tranches comes down to modeling $\mathbb{E}\left[L_u^{\mathcal{B}}(t)\right]$. Under the class of one-factor copula models this can be done relatively fast using the observation that, conditional on the realization of the common factor, all remaining sources of risk are independent. The obvious equality

$$\mathbb{E}\left[L_u^{\mathcal{B}}(t)\right] = \int_{-\infty}^{\infty} \mathbb{E}\left[L_u^{\mathcal{B}}(t)\,|Y=y\right]\varphi(y)dy, \tag{4.7}$$

where φ denotes the probability density function for the standard normal distribution, tells us that we can focus on the modeling of the conditional loss process, given by $\mathbb{E}\left[L_u^{\mathcal{B}}(t)\,|Y=y\right]$, where we condition on the realization of the common factor Y.

One can use a simplified version of the methods described in [ASB03], [HW04] and [LG03] and we will illustrate this method using an example where the probability of exactly n defaults at some future time t will be determined. The approach can be seen as a recombining binomial tree which has N steps, each step corresponding to one name in the basket. At horizonal step i the probability of an up-move will be the default probability of name i conditional on the realization of the common factor $Y=y$: $p_i\,(t|y)$. Thus the conditional probability of a down move will be $1-p_i\,(t|y)$. Every up-move corresponds to a default event, while a down-move corresponds to survival. Working from the root of the tree to the end values, one obtains the probability of exactly n defaults, for $n=0,\ldots,N$. To illustrate this, let $\delta_{i,j}$ denote the path probability at step i and height j. We assume that i runs from $0,\ldots,N$, and $j=0,\ldots,i$. Thus we start the algorithm by setting $\delta_{0,0}=1$, after which the algorithm continues as follows:

$$\delta_{i,j}=\begin{cases}(1-p_i)\cdot\delta_{i-1,j} & \text{if } j=0\\ p_i\cdot\delta_{i-1,j-1}+(1-p_i)\cdot\delta_{i-1,j} & \text{if } 0<j<i\\ p_i\cdot\delta_{i-1,j-1} & \text{if } j=i\end{cases} \tag{4.8}$$

It is important to note that this algorithm can only be applied with independent default probabilities. For brevity we have used $p_i=p_i\,(t|y)$. If one is done working through the tree, the values $\delta_{N,j}=\delta_{N,j}\,(t|y)$ give the conditional probability of exactly j defaults in the basket of names at time t. Integrating these terms out over the distribution of the common factor will result in the probability of exactly j defaults, $\pi\,(j,t)$ for $j=0,\ldots,N$. Thus we get:

$$\pi\,(j,t)=\int_{-\infty}^{\infty}\delta_{N,j}\,(t|y)\,\varphi(y)dy \tag{4.9}$$

To calculate this integral one has to resort to numerical integration techniques, for instance Gauss-Legendre quadratures as described in [AS72]. During numerical integration one has to build the entire tree as shown in Equation (4.8) for each chosen value for the common factor realization, y,

From this one can easily determine the expected loss on a certain tranche. Alternatively one could determine the expected loss on a tranche conditional on the realization of the common factor and integrate out subsequently. This will lead to the same result as first integrating over probabilities. As can be seen from the pricing formulas (4.1) and (4.2) this procedure has to be repeated to determine the expected loss on a tranche for different future times.

4.3.2 Base Correlation

In recent years the market for Tranche CDSs has become very liquid leading to a need for a marking-to-market approach. Similar to implied volatilities, the notion of implied correlation can be introduced for these products using the standard one-factor Gaussian copula. Assuming that each pair of reference entities has a constant correlation parameter, one can derive a correlation such that the model price exactly matches the price observed in the market. This results in a tranche specific correlation parameter known as the implied, or compound correlation. One drawback of such an approach is that one is left with correlations which are a function of both attachment as well as detachment level. The base correlation method, due to [MBAW04], overcomes this issue by switching to equity tranches, or base tranches. In Section 4.2.2 it was shown that the expected loss of a certain tranche can always be written as the difference of expected losses for two equity tranches, see Equation (4.3). The idea of the base correlation method is that all tranches are reduced to base tranches and correlation depends on the detachment level of the tranche. One can iteratively solve for the correlations belonging to each detachment level for which a quote is available. The base correlation method has become the industry standard for valuing CDO tranches, see [Rey04] or [FR04] for instance.

Let $Q_{l,u}$ denote the quote of a tranche with attachment percentage a and detachment percentage d. Further, let $U_{a,d}$ denote the upfront value in percentages of tranche notional. A set of base correlations is calculated per basket and per maturity for which quotes are available. In the derivations below the basket and maturity are fixed and for brevity their dependency is not shown in the protection and premium values. We want to determine ρ_d recursively such that the model implied fair spread equals the spread quotes in the market. These correlations have to satisfy the equality

$$Q_{d,a} = \frac{d \cdot \mathrm{DL}_d(\rho_d) - a \cdot \mathrm{DL}_a(\rho_a) - (d-a) \cdot U_{a,d}}{d \cdot \mathrm{PL}_d(\rho_d) - a \cdot \mathrm{PL}_a(\rho_a)} \tag{4.10}$$

For the quoted equity tranche the attachment percentage is 0 and thus both DL_a and PL_a are zero. One can use a numerical root-finding routine, see for instance [PTVF02], to obtain the value of ρ_d such that the quote is exactly matched by the model. Once this has been determined, one can move to the next tranche and its attachment level, a, equals the detachment level of the previous tranche under consideration. Thus ρ_a is already known, as this is ρ_d solved in the previous step. This can be used to determine the next base correlation:

$$\begin{aligned} Q_{a,d}\left(d \cdot \mathrm{PL}_d(\rho_d) - a \cdot \mathrm{PL}_a(\rho_a)\right) &= d \cdot \mathrm{DL}_d(\rho_d) - a \cdot \mathrm{DL}_a(\rho_a) - (d-a) \cdot U_{a,d} \\ Q_{a,d} &= \frac{d \cdot \mathrm{DL}_d(\rho_d) - (d-a) \cdot \widetilde{U}_{a,d}}{\mathrm{PL}_d(\rho_d)} \end{aligned} \tag{4.11}$$

In the last line the terms already known, i.e. those depending on the attachment level l have been grouped in the adjusted upfront value:

$$\widetilde{U}_{a,d} \equiv U_{a,d} + \frac{a \cdot \mathrm{DL}_a(\rho_a) - a \cdot \mathrm{PL}_a(\rho_a) \cdot Q_{a,d}}{d-a}. \tag{4.12}$$

From Equation (4.11) one can see that the problem of finding ρ_d is reduced to finding the implied correlation of an equity tranche with detachment level d. Moreover as the value of an equity tranche is monotonically decreasing in default correlation, this relationship shows that the base correlation is unique, something which is not necessarily the case for implied correlations for general tranches. Moreover the monotonicity makes numerical root-finding a straightforward exercise.

Note that this algorithm can be applied only when quotes are available for successive tranches, where the detachment level of one tranche equals the attachment of another. Furthermore a quote for the equity tranche is required as well.

Figure 4.2 shows both compound correlations as well as base correlations using market data of the 1^{st} of November 2006, for tranches on the CDX index with maturity of five years. Both curves are plotted as a function of the detachment level.

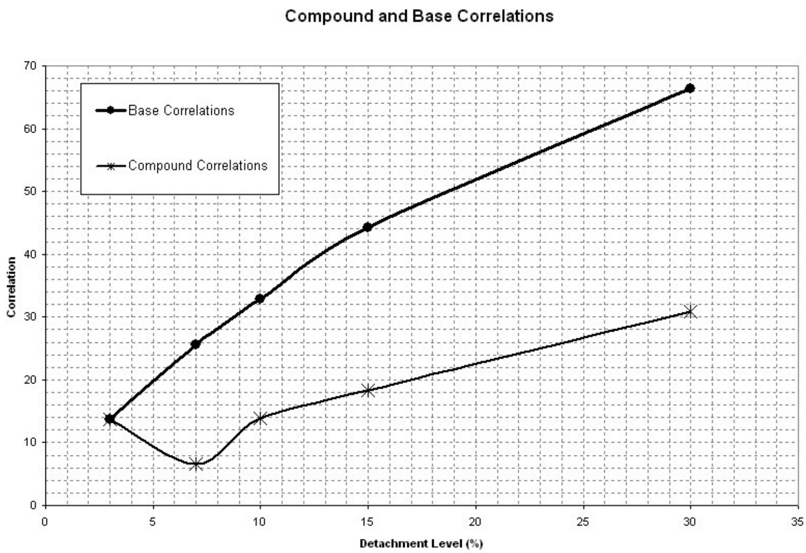


Figure 4.2: Compound and Base correlations plotted against the detachment level of the tranches. Quotes on standard tranches on the DJ CDX basket are used from the 1^{st} of November 2006.

The compound correlation for the equity tranche is around 14%, after which it decreases to a level close to 6.5% for the junior mezzanine tranche. For more senior tranches it increases again from roughly 14% to 18% to reach a level of 31% for the most tranche. Further, one can observe a close to linear relationship between base correlations and detachment level. By construction the base correlation for the equity tranche equals the compound correlation. The base correlations increase to roughly 66% for the most senior tranche. The fact that the lines are not horizontal clearly illustrates the model's inability to fit the market with a single correlation parameter.

In order to apply the base correlation method to non standard tranches, for example tranches with non standard attachment and detachment levels, tranches with different maturities, or tranches on bespoke baskets, having a different set of underlyings as the standard baskets iTraxx and CDX, one has to choose the correlation that applies. Therefore the correlations obtained from quoted tranches on the standard baskets are 'mapped' such that they can be used to value non standard tranches. Ideally such an approach would take the credit risk in the underlying basket over the relevant period into account. For example a 0-3% tranche on a basket which has fair spread of 30 bp is very different from a 0-3% tranche on a basket having a fair spread at 100 bp. The following three different mapping methods have been described in literature:

- Direct *d*etachment-*d*etachment mapping, dd.
- *D*etachment as a *p*ercentage of expected basket loss, dp.
- Expected *l*oss *p*ercentage mapping, lp.

The first method states that the base correlation for any basket just depends on the detachment level only. It does not take into consideration differences in credit risk of different portfolios. The second method is described in [MBAW04] where it is proposed to divide the detachment level by the expected loss on the entire basket over the same time-span, resulting in a scaled detachment level. From this we define a functional form between scaled detachment and base correlation. In order to determine the correlation for a tranche with different maturity or on a different basket one first divides the detachment level by the expected loss of the underlying basket up to the relevant maturity and then determine the number using the functional form defined earlier. The third and final mapping method is the most advanced and is described in [RUK04]. For each combination of detachment level and corresponding base correlation one first determines the expected loss on the tranche, expressed as a percentage of expected loss on the total basket, in general leading to five combinations of expected loss percentage and base correlation, as typically quotes can be observed for five tranches. When mapping to a different maturity or different basket one first implies detachment levels on this basket such that the same combinations of expected loss percentage and base correlations are found. This requires a numerical root-finding technique. Finally the resulting detachment levels and corresponding base correlations can be used in combination with interpolation and extrapolation rules to determine the correlation to use for a certain tranche.

4.3.3 Other Default Dependency Models

The base correlation approach can not be seen as a proper model of defaults, but merely as a method to visualize market quotes in terms of correlations. It does not allow one to determine prices on bespoke baskets, without the use of mapping methods. Difficulties will increase even further when prices for more complex credit derivatives are required, such as a CDO of CDOs, or CDO^2. For these reasons many alternative models have been proposed in the literature, usually extending the one-factor Gaussian copula in one way or the other. For instance, [HW04] propose to model both the common factor and the idiosyncratic risk terms using a Student t distribution. Other extensions allow for randomness of the correlation parameter. A mixture copula for instance assumes that the correlation parameter can take on a limited number of values each with some probability. Another random correlation model proposed by [AS04] allows the correlation to be dependent on the common factor. In this case one can model behavior where a 'bad' state of the economy, reflected through the realization of the common factor, goes accompanied by large default correlation. In Chapter 5 we shall see that these and other one-factor copula models do a reasonable job in fitting to market quotes. A comparative analysis of a number of different one-factor copula models is given in [BGL05]. However, the base correlation method is still the industry standard approach for pricing tranches and to communicate about default correlation.

4.4 Data

In order to apply the base correlation method a large amount of input is required. For this study market data has been used which has been made available to us by ABN Amro. For the analysis we have used data from December 2004 up to November 2006. From the modeling description given in Section 4.3 it is clear that market data is needed for three parts of the model: risk free discounting, marginal default distributions and the default dependency. Each of these aspects is related to specific types of market instruments which are discussed below.

4.4.1 Default Free Discounting

In order to approximate risk free discounting, we have used quotes for money market accounts and plain vanilla interest rate swaps. From these quotes a risk free discounting curve is constructed using standard methods, see for instance [Hul05]. For each day of the data set a risk free discounting curve is constructed for both the US dollar as well as for the Euro.

4.4.2 CDS Term Structures

As our data set ranges from December 2004 up to November 2006, different series for both indices are involved due to the roll of the indices as discussed in Section 4.2.1.

For the iTraxx Europe index this means that series 2 up to series 6 are included. For CDX the data also involves five different compositions, namely series 3 up to series 7. On the Markit website: www.market.com, the composition of th iTraxx and CDX baskets can be found. It is worth noting that when rolling from series 4 to series 5 at September 20^{th} 2005, Ford and General Motors, were removed from the basket, amongst other reference entities. Due to the rolls the total number of different reference entities in the data set is thus larger than 250. For each of these reference entities a CDS term structure is available at each date in the period. A piecewise constant default intensity curve is used, which allows the model to be calibrated to the CDS quotes using a bootstrap method, see [Sch03a] or [BM06] for instance. Recovery rates are assumed to be 40% for all names. As discussed in [HV90] the choice of the recovery rate does not have a large influence on pricing, as long as one calibrates to observed CDS quotes. From the resulting marginal default distributions for all names, one can easily determine the fair spread for an index credit default swap on each of the two indices. This determines the total amount of default risk present in each of the baskets and is therefore a large determinant of the values of the different tranches as well. Figure 4.3 shows the evolution of the fair spread for an index CDS on iTraxx and CDX with five year maturity.

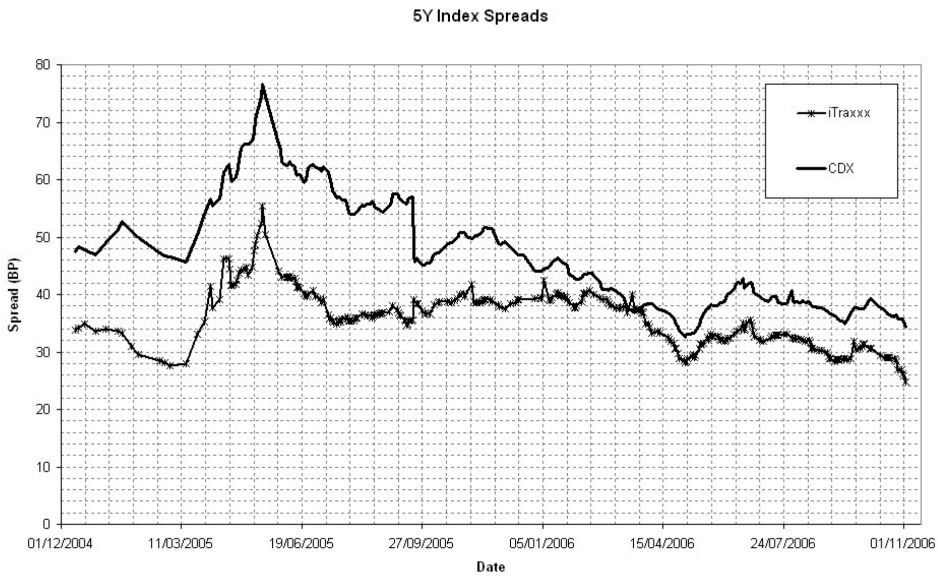


Figure 4.3: Fair spread for a five year index credit default swap on iTraxx and CDX from December 2004 up to November 2006.

From the figure one can observe the decrease of the fair spread for a five year CDS on the CDX index around the roll date of September 20^{th}, 2005, which is partly due to the removal of Ford and General Motors. Further one can observe that both indices show a similar behavior over time.

4.4.3 Synthetic CDO quotes

Once the marginal default distributions are calibrated to CDS quotes the remaining part is to include default dependency. Quotes for synthetic CDO tranches are available which, in combination with the base correlation method, determine the dependency structure. Quotes for tranches with a five year maturity are available from December 2004 up to November 2006, while quotes for tranches with seven and ten year maturity are available only over the period from July 2005 up to November 2006.

Table 4.2 presents a summary of statistics of both the quotes on iTraxx as well as those on CDX, all with a maturity of five years. In addition summary statistics are presented for the two reference spreads. For each series of tranche quotes or reference spread the table presents the average, the standard deviation, the maximum and minimum values. The last row presents the correlation of the quote under consideration with the reference spread.

iTraxx	0 – 3	3 – 6	6 – 9	9 – 12	12 – 22	Ref.
Avg.	24.7	86.3	26.4	13.5	7.4	35.15
Std. Dev.	6.3	31.6	10.6	6.6	4.9	5.9
Max.	49.6	191.0	63.0	35.0	25.0	57.0
Min.	10.4	43.8	13.0	5.0	2.1	23.0
Corr. Ref.	95%	82%	84%	88%	89%	100%
CDX	0 – 3	3 – 7	7 – 10	10 – 15	15 – 30	Ref.
Avg.	37.0	126.1	31.6	15.4	7.9	47.6
Std. Dev.	9.1	50.6	14.8	7.7	4.4	10.0
Max.	62.0	260.0	74.5	34.0	20.0	76.0
Min.	22.9	65.0	12.0	6.5	3.1	34.0
Corr. Ref.	98%	86%	75%	91%	94%	100%

Table 4.2: Summary statistics for quotes with a five year maturity on both the European iTraxx index as well as the North-American CDX index. For each tranche the average (Avg.), the standard deviation (Std. Dev.), the maximum (Max.), the minimum (Min.) and the correlation with the reference quote (Corr. Ref.) are shown.

One can observe that quotes have moved a lot over the past two years, as can be derived from the standard deviations as well as the large differences between minimum and maximum quotes of the different tranches. Further one can observe that for each of the ten different time series of tranche quotes there exists a large degree of correlation with the reference spread. Especially the equity tranches and the most senior tranches seem to be largely correlated with the value of the reference spread. This relationship is investigated in more detail in Section 4.5.1.

4.4.4 Liquidity

As discussed the synthetic CDO market has grown rapidly over the last few years. In order to investigate whether there is evidence of this growth in the available data set, the bid/offer spread of the quotes can be considered, which are available for each tranche over the entire period. A smaller bid/offer spread would indicate a larger degree of liquidity for the synthetic CDO tranches. In Table 4.3 the average bid offer spread of the first ten days is compared to the average bid/offer spread of the last ten days for each synthetic CDO tranche with a five year maturity.

iTraxx	Dec. 2004	Nov. 2006		CDX	Dec. 2004	Nov. 2006
0 – 3	0.67	0.24		0 – 3	1.33	0.24
3 – 6	3.30	0.60		3 – 7	4.35	1.70
6 – 9	3.30	0.95		7 – 10	5.40	1.40
9 – 12	3.75	0.95		10 – 15	5.75	1.23
12 – 22	2.00	0.76		15 – 30	2.58	0.95

Table 4.3: Bid/offer spread of the different tranches on the two indices. A ten day average at the beginning of the data set is compared to a ten day average at the end of the data set.

From the table one can observe that bid/offer spreads have decreased dramatically, roughly with a factor three. Although the spread of both underlying indices has decreased slightly as well this large decrease in bid/offer spread clearly indicates a large increase in synthetic CDO tranche liquidity over the last two years. A similar effect can be observed for the tranches on iTraxx and CDX with either seven year or ten year maturity. Relative decrease in bid/offer spreads were similar although one has to note that the period for which seven and ten year data is available only ranges from July 2005 up to November 2006.

4.5 Empirical Analysis

In the previous section we have discussed the market data that we shall use in this section to carry an empirical analysis. First we investigate the correlation between moves in the different tranche quotes. Thereafter we look at the correlation between the base correlations that we calculate for the available tranche quotes. This section is concluded by a comparison of the different mapping methodologies that were introduced in Section 4.3.2.

4.5.1 Correlation Moves in Quotes

In Section 4.4 it was observed that there exists a large correlation between quotes of tranches and the reference spread of the quote. Here we investigate this relationship in more detail using iTraxx and CDX tranche quotes with a five year maturity. Clearly, in a CDO structure, such as shown in Figure 4.1 any change in the risk of the pool of collateral assets must be translated to the different tranches of the CDO structure. When taking a closer look at modeling one finds that, apart from interest rates, the fair spread is a function of both the credit risk of the underlying index as well as default correlation, as can be seen from Equation 4.4. The effect of interest rate moves is typically small as it affects both legs in a similar manner and will therefore be ignored here. As before let Q denote the fair spread of a synthetic CDO tranche and S the fair spread of a credit default swap on the underlying index with the same maturity as the synthetic CDO tranche. Further let ρ denote a parameter for default correlation. Under these settings one has $Q = Q(S, \rho)$ and thus

$$\mathrm{d}Q = \frac{\partial Q}{\partial S}\mathrm{d}S + \frac{\partial Q}{\partial \rho}\mathrm{d}\rho.$$

This equation shows that after correcting for index spread movements, changes in default dependency remain the only driver of changes in quotes. We are interested in levels rather than changes in the quotes and therefore it is assumed that there exists a simple linear relationship between the tranche quote and the corresponding basket spread. In order to investigate correlation effects, one should thus adjust for the risk in the portfolio, which is reflected in the index spread, i.e. the spread of an index credit default swap on the index with same maturity as the CDO tranche. As a proxy for this index spread the reference spread is used, which is part of the quote. This can be done by means of an Ordinary Least Squares (OLS) regression where one regresses the observed quote on the corresponding reference spread. Let Q_i^k denote the mid quote for tranche k, for observation i. Assume that the iTraxx tranches are indexed $1, \ldots, 5$ and CDX tranches $6, \ldots, 10$ in order of seniority. Further let the corresponding reference spread be denoted by R_i^k, which is equal for all iTraxx quotes and is also equal for all CDX quotes. The following simple linear relationship is considered

$$Q_i^k = \alpha^k + \beta^k \cdot R_i^k + \varepsilon_i^k. \tag{4.13}$$

Here the disturbances ε_i^k are assumed to be i.i.d. normally distributed. The regression analysis has been performed for each of the available time series of quotes and regression results are shown in Appendix 4.A. Let a^k and b^k denote the estimate values for α^k and β^k, respectively. One can observe from the regression results that for each time series the estimated coefficient b^k is highly significant and that the coefficients of determination are reasonably large. Thus the reference spread explains a large part of the variation in the observed market quotes of synthetic CDO tranches. For each of the time series a second regression analysis has been performed where a quadratic form is considered by including the squared reference spread. However, for none of the series this improved explanatory power of the model significantly.

Using the estimated parameters as shown in Appendix 4.A, one can determine the residuals e_i^k as follows

$$e_i^k \equiv Q_i^k - a^k - b^k \cdot R_i^k, \qquad for\ k = 1, \ldots, 10. \tag{4.14}$$

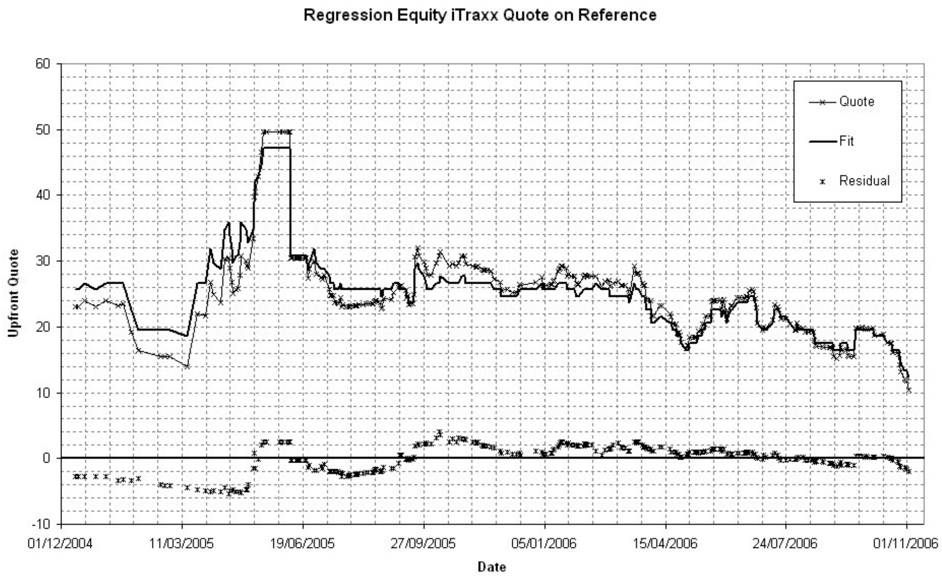


Figure 4.4: Regression of iTraxx equity quote on the reference spread. Both the original data as well as the fit and the residuals are presented.

As an example the series of iTraxx equity quotes is considered. As can be seen from the appendix the estimated coefficients are $a^1 = -11.1$ and $b^1 = 1.02$ and the coefficient of determination is close to 90%. In Figure 4.4 we present the time series of quotes along with the estimated values using the reference quote. In addition the estimated residuals are plotted.

As can be seen from the figure the simple linear relationship provides a good fit. Further one can observe that the estimation residuals do not seem to be independent but clearly follow a certain pattern over time. Furthermore one can clearly observe the increase in residuals at the beginning of May 2005. This increase means that, after correcting for moves in reference spread, the equity quote has moved upward. An increase in the value of protection for the equity tranche goes accompanied by a decrease in default correlation. This is exactly what was observed in the base correlations during the May 2005 events.

This simple OLS analysis can be applied to all different tranches which are quoted and one can observe clear patterns for the errors over time for each of these series. In Table 4.4 the correlation matrix of the ten different residuals series is presented.

From Table 4.4 one can observe some interesting results. First one can observe that for both indices the residuals for the equity tranche are negatively correlated with those for the other tranches. Note that this makes intuitive sense as when prices

Correlation of Errors		iTraxx 0-3	3-6	6-9	9-12	12-22	CDX 0-3	3-7	7-10	10-15	15-30
iTraxx	0-3	100%	-64%	-76%	-78%	-75%	80%	-54%	-73%	-75%	-49%
	3-6	-64%	100%	87%	74%	48%	-64%	93%	83%	70%	-1%
	6-9	-76%	87%	100%	83%	60%	-74%	82%	89%	87%	30%
	9-12	-78%	74%	83%	100%	85%	-80%	73%	80%	80%	44%
	12-22	-75%	48%	60%	85%	100%	-78%	47%	59%	63%	62%
CDX	0-3	80%	-64%	-74%	-80%	-78%	100%	-67%	-76%	-74%	-57%
	3-7	-54%	93%	82%	73%	47%	-67%	100%	84%	73%	1%
	7-10	-73%	83%	89%	80%	59%	-76%	84%	100%	90%	28%
	10-15	-75%	70%	87%	80%	63%	-74%	73%	90%	100%	50%
	15-30	-49%	-1%	30%	44%	62%	-57%	1%	28%	50%	100%

Table 4.4: Correlation of the regression residuals for the different series of synthetic CDO quotes. These residuals can be regarded as reference corrected quotes and are expected to capture correlation moves in the synthetic CDO market.

for the equity tranche are high relative to the reference spread, or average portfolio spread, it must be that the more senior tranches have lower prices relative to average spread. Another interesting observation is the larger correlation of the residuals for the different underlying baskets. For instance residuals for quotes on equity tranches have a correlation of 80%, while for the junior mezzanine tranche this is even 93%. Thus in Figure 4.3 it was shown that index CDS spreads for iTraxx and CDX moved closely and the table above suggests that the default dependency structure also moves closely between both indices.

4.5.2 Empirical Base Correlations

As discussed in Section 4.3.2 the base correlation method has become the industry standard approach which is used to visualize market implied correlations. This method has been applied to the data and in appendix 4.B the five base correlation series for iTraxx are shown, as well as the ones for the CDX all for a five year maturity. When considering the base correlation for the equity tranches for both iTraxx and CDX one can clearly observe the May 2005 event, where Ford and General Motors were downgraded. For both indices base correlations for the equity tranche dropped from around 20% to just below 10%. For the base correlations with higher detachment levels corresponding to the iTraxx one can observe similar drops. However, for the CDX base correlations corresponding to higher detachment levels one can see an increase.

As can be seen from the figures in Appendix 4.B the correlations move closely together. Table 4.5 presents the correlations of the different base correlations series.

These numbers confirm that base correlations for the different detachment levels move closely together and in addition show that the base correlations corresponding to the two different indices move closely together. One can not directly compare these numbers with the correlations of the regression residuals presented in Table 4.4.

Correlation of B.C.s		iTraxx					CDX				
		3	6	9	12	22	3	7	10	15	30
iTraxx	3	100%	89%	80%	75%	65%	90%	79%	70%	57%	58%
	6	89%	100%	98%	96%	89%	74%	92%	89%	81%	81%
	9	80%	98%	100%	99%	94%	64%	91%	91%	86%	86%
	12	75%	96%	99%	100%	96%	59%	89%	91%	87%	88%
	22	65%	89%	94%	96%	100%	46%	76%	80%	75%	80%
CDX	3	90%	74%	64%	59%	46%	100%	76%	62%	48%	49%
	7	79%	92%	91%	89%	76%	76%	100%	98%	93%	90%
	10	70%	89%	91%	91%	80%	62%	98%	100%	98%	95%
	15	57%	81%	86%	87%	75%	48%	93%	98%	100%	96%
	30	58%	81%	86%	88%	80%	49%	90%	95%	96%	100%

Table 4.5: Correlation of the different base correlations. For both iTraxx and CDX base correlations for five different detachment levels are determined for each date in the data set. This results in a time series for each base correlation and the table presents correlation for these series.

The reason is that the latter are based on the tranches seen in the market, which correspond to different parts of the CDO structure. When considering the base tranches, i.e. equity tranches with different detachment levels, there is a large amount of overlap. One could however compare the base correlations corresponding to the equity tranche, i.e. with detachment level of 3%, to the residuals determined in Section 4.5.1. For iTraxx the correlation for these two time series equals -94%, while for CDX the correlation is -83%. As discussed earlier, large residuals for OLS analysis for the equity tranches goes accompanied by low base correlation, explaining the negative sign of the two correlations between the time series of residuals and the time series of base correlations. These large correlations provide further evidence for the claim that the quotes, when adjusted for portfolio credit risk, indeed give a good indication of the default dependency implied in the market.

4.5.3 Mapping Methodologies

We now turn to the different mapping methodologies as described in Section 4.3.2. As discussed the base correlation method is not a proper model which becomes clear when one marks-to-market non-standard tranches, for instance tranches with non-standard attachment and detachment levels, tranches with non-standard maturities, or tranches on bespoke baskets. As only five base correlations are available, one has to resort to interpolation/extrapolation techniques in combination with mapping methods. Here we consider the three different mapping methodologies described earlier on.

1. Direct *d*etachment-*d*etachment mapping, dd.

2. *D*etachment as a *p*ercentage of expected basket loss, dp.

3. Expected *l*oss *p*ercentage mapping, lp.

In the test all different mapping methods are used in combination with cubic splines to determine a base correlation for each detachment level. Linear interpolation is used for detachment levels smaller than 3%. Note that the choice of interpolation technique is a complex problem as well, as one might introduce arbitrage opportunities, especially for short term CDO tranches. Here the base tranche loss as function of detachment point might not be increasing. of However this topic is not considered in this chapter.

For testing purposes we shall use the quotes on standard tranches with a five year maturity for building the base correlation skew, as these are by far the most liquid correlation instruments. After building the skew we apply one of the three different mapping methods in order to determine the model value of tranches with different maturities or on different underlying baskets. Comparing these model generated values with the actual values results in an error of the model compared to the market value.

Let $\mathcal{S}$ denote the set of tranches used for testing the mapping method. By definition when considering a tranche with upfront value and premium contracted as implied by the quotes, the correct value should be zero. Let V_i^j denote the model value for tranche i under mapping method j, when the upfront percentage and contracted premium correspond to the quoted numbers. As the market implied value of the tranche is zero, the error made by the mapping method is thus equal to V_i^j. These errors can be used to define the Root Mean Squared Error, for a mapping method j given a set of tranches $\mathcal{S}$. We refer to this as the value based Root Mean Squared Error, or ξ^V:

$$\xi^V\left(j, \mathcal{S}\right) \equiv \sqrt{\frac{1}{N_{\mathcal{S}}} \sum_{i \in \mathcal{S}} \left(V_i^j\right)^2}. \tag{4.15}$$

Here $N_{\mathcal{S}}$ denotes the number of tranches in the testing set. Instead of using the value of the tranche, one could alternative look at the fair upfront value or fair premium implied by the model, to determine the errors. The market quotes the upfront or premium, while a mapping method results in a model implied fair upfront value or fair premium, leading again to an error. Let Q_i denote the quote for tranche i and π_i^j the model implied upfront value of fair spread for the same tranche under mapping method j. In addition to the value based ξ we consider a quote based Root Mean Squared Error, ξ^Q:

$$\xi^Q\left(j, \mathcal{S}\right) \equiv \sqrt{\frac{1}{N_{\mathcal{S}}} \sum_{i \in \mathcal{S}} \omega_i \cdot \left(Q_i - \pi_i^j\right)^2}. \tag{4.16}$$

Here the terms ω_i allows one to give a certain weight to the different tranches. One could for instance set $\omega_i = 1/Q_i^2$ and consider the relative errors, a choice which is made throughout the remainder of this section.

When looking at the value based ξ one can expect that a large part of the result is caused by equity and junior mezzanine tranches. When using the quote based ξ with weights as above, the difficulty arises that quotes for the most senior tranches can be very low, making large relative errors likely and causing the ξ^Q to put a large weight on these senior tranches. Thus the ξ^V is likely to put more weight on the equity and junior tranches, while the ξ^Q is expected to put more weight on the more senior tranches. For this reason both of them are considered in the performance test for the different base correlation mapping methods.

Mapping in the Time Domain

First the performance of the different mapping methods is investigated when varying the maturity of the tranches. Using quotes on five year synthetic CDO tranches and one of the mapping methods, one can determine the model implied quote for a standard tranche with either seven or ten years maturity. As quotes are available for these products one can easily determine the error made due to the use of a mapping method. From these errors one can determine the different ξs as described above, and compare the performance of the different mapping methods among each other. In appendix 4.C the value based ξ is presented for each of the five tranches on both iTraxx and CDX, with either seven year or ten year maturity. In addition the ξ is aggregated over both indices and both maturities for each tranche separately. Further the ξ is aggregated over all tranches for a certain combination of index and maturity. Finally the ξ is determined over all tranches for each mapping method.

As can be seen from the results the mapping based on the expected loss percentage, *lp*, shows the best performance. With the exception of the equity tranche it shows the best performance for all tranches. A slightly worse performance is found for the scaled detachment method, *dp*, where the detachment is expressed as a fraction of expected loss of the total basket. The simple detachment based mapping, *dd*, shows the worst performance, even though for equity tranches it appears to generate the closest fit. A reason for this might be the chosen extrapolation method for the other two mapping methodologies, which is not an issue for the *dd* method.

Table 4.6 shows the ξ using all available tranches. Both the value based ξ is shown as well as the ξ based on relative errors for fair premia.

Mapping	ξ^V	ξ^Q
Detachment-Detachment, *dd*	3.78	0.46
Detachment Percentage, *dp*	3.38	0.28
Expected Loss Percentage, *lp*	2.54	0.18

Table 4.6: Total ξ under two approaches for all three different mapping methods. ξ^V is based on the value of the tranche when contracted premium equals fair premium. The ξ^Q is based on the quoted numbers and relative errors are used.

From this table one can observe that the expected loss percentage mapping, *lp*, shows the best performance under both measures. However the average relative error in terms of the quotes is still at a level of 18%.

Mapping to Bespoke Baskets

A similar test can be performed in order to test the different mapping methods when applied to bespoke baskets. Quotes for bespoke baskets are not available, but one could use quotes on iTraxx tranches and subsequently value CDX tranches. The model generated quotes can then be compared to the market quotes and the performance of the different mapping methods can be investigated. We have used the five year quotes on the iTraxx index to build a base correlation skew. Using one of the three mapping methods the standard tranches on CDX with maturities of five, seven and ten years, are priced and results are compared with the correct values. In Appendix 4.C one can find the value based ξ split over the different mapping methods, the different tranches and the different maturities. Again the results show that the expected loss percentage, *lp*, mapping performs best in most cases. Again one can observe that the simple detachment mapping, *dd*, performs best for equity tranches. It is interesting to see that this is only the case when mapping from CDX to iTraxx tranches. When mapping from iTraxx to CDX tranches the *dd* mapping method shows the worst performance.

Table 4.7 shows the ξ using all available tranches. Both the value based ξ is shown as well as the ξ based on relative errors for fair premia.

Mapping	ξ^V	ξ^Q
Detachment-Detachment, *dd*	3.21	0.42
Detachment Percentage, *dp*	3.48	0.32
Expected Loss Percentage, *lp*	2.75	0.23

Table 4.7: Total ξ under two approaches for all three different mapping methods. ξ^V is based on the value of the tranche when contracted premium equals fair premium. The ξ^Q is based on the quoted numbers and relative errors are used. iTraxx quotes are used to map to CDX tranches and vice versa.

From this table one can observe that the expected loss percentage mapping, *lp*, shows again the best performance under both measures. However the average relative error in terms of the quotes is still at a level of 23%. Further one can observe that the simple detachment mapping, *dd*, performs better than the scaled detachment level, *dp*, when considering the value based ξ. For the ξ based on relative error in fair premium, the results are the other way around. This is caused by the large relative errors for the simple detachment mapping method for senior tranches.

4.6 Conclusion

Over the last years the market for synthetic CDO tranches has undergone a rapid development together with the research efforts in order to value these products consistently with the observed market quotes. A widely used model for imposing default correlation is the one-factor Gaussian copula approach and calibration is done using the base correlation method. One of the main drawbacks is that the base correlation method is not a proper model and mapping methods are required when pricing tranches with non-standard attachment and detachment levels, non-standard maturities or tranches on bespoke baskets.

In this paper we have focussed on default correlation implied by market quotes. Using a large data set with quotes for synthetic CDO tranches between December 2004 and November 2006 we have considered different aspects of market implied default correlation. First the quotes have been used directly, using a simple linear regression to correct for the level of default risk of the reference index. The residuals show a clear pattern over time, indicating moves in market implied default dependency. One could clearly observe the May 2005 events in these residuals. Furthermore it was shown that the correlation of these residuals amongst each other has been large and of expected sign, where residuals for equity quotes where negatively correlated with those for the other tranches. In addition it was shown that there exists a large correlation between the time series of the residuals for equity tranches and the time series of base correlations corresponding to the 3% detachment level.

In addition the data has been used to investigate the performance of three different mapping methods. These methods determine the correlation for a non-standard tranche using the base correlations implied from market quotes. Two out-of-sample tests have been performed. First quotes for tranches with a five year maturity on a certain index were mapped through the base correlation approach to tranches with seven and ten year maturity on the same index. This allowed us to investigate the performance of the different mapping methods in the maturity domain. Secondly, quotes for five year tranches on a certain index where mapped to tranches with five, seven and ten year maturities referencing the other index. This allowed us to investigate the performance of the mapping methods when mapping to different baskets. It was found that the method based on the expected loss percentage outperforms the use of the method of scaled detachment levels and simple detachment only mapping. However in all cases the errors were still very large with relative errors in quotes around levels of 20%.

4.A Regression Results: Quotes Regressed on Reference Spread

iTraxx	**0 - 3**		
R^2	0.898		
Variable	**Coeff.**	**Std. Err.**	**t ratio**
Constant	-11.139	0.719	-15.502
Ref. Spread	1.022	0.020	50.617

CDX	**0 - 3**		
R^2	0.956		
Variable	**Coeff.**	**Std. Err.**	**t ratio**
Constant	-5.268	0.569	-9.259
Ref. Spread	0.888	0.012	75.886

iTraxx	**3 - 6**		
R^2	0.676		
Variable	**Coeff.**	**Std. Err.**	**t ratio**
Constant	-68.661	6.386	-10.752
Ref. Spread	4.417	0.180	24.603

CDX	**3 - 7**		
R^2	0.742		
Variable	**Coeff.**	**Std. Err.**	**t ratio**
Constant	-81.994	7.660	-10.704
Ref. Spread	4.370	0.157	27.754

iTraxx	**6 - 9**		
R^2	0.706		
Variable	**Coeff.**	**Std. Err.**	**t ratio**
Constant	-26.618	2.040	-13.051
Ref. Spread	1.512	0.057	26.363

CDX	**7 - 10**		
R^2	0.560		
Variable	**Coeff.**	**Std. Err.**	**t ratio**
Constant	-21.302	2.926	-7.280
Ref. Spread	1.111	0.060	18.475

iTraxx	**9 - 12**		
R^2	0.773		
Variable	**Coeff.**	**Std. Err.**	**t ratio**
Constant	-21.109	1.115	-18.928
Ref. Spread	0.986	0.031	31.445

CDX	**10 - 15**		
R^2	0.824		
Variable	**Coeff.**	**Std. Err.**	**t ratio**
Constant	-18.003	0.964	-18.671
Ref. Spread	0.702	0.020	35.435

iTraxx	**12 - 22**		
R^2	0.798		
Variable	**Coeff.**	**Std. Err.**	**t ratio**
Constant	-18.892	0.786	-24.023
Ref. Spread	0.749	0.022	33.876

CDX	**15 - 30**		
R^2	0.879		
Variable	**Coeff.**	**Std. Err.**	**t ratio**
Constant	-11.981	0.460	-26.043
Ref. Spread	0.417	0.009	44.140

Table 4.8: Results of regression analysis, where the quote for each tranche is regressed on the reference spread. On the left hand side results are shown for iTraxx, while on the right hand side results for CDX are shown. For each regression the coefficient of determination, R^2 is shown and for each parameter its estimate, standard error and t ratio.

4.B Base Correlations

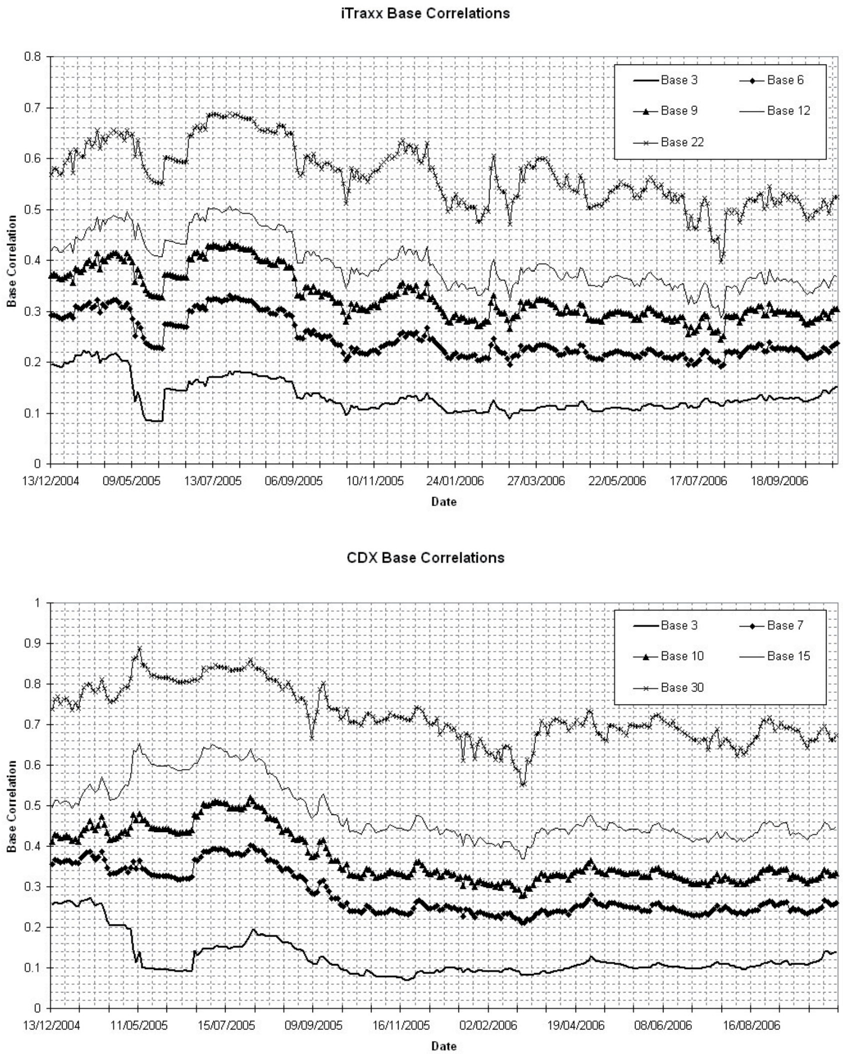


Figure 4.5: iTraxx and CDX base correlations from December 2004 up to November 2006.

4.C Performance of Mapping Methods

			Tranche 1	Tranche 2	Tranche 3	Tranche 4	Tranche 5	Total
iTraxx	7Y	*dd*	2.83	1.32	1.17	0.88	0.24	1.55
		dp	2.94	3.72	0.62	0.28	0.08	2.14
		lp	2.29	2.51	0.35	0.21	0.13	1.53
	10Y	*dd*	2.58	10.11	3.33	2.90	1.18	5.10
		dp	6.20	6.16	5.31	1.32	0.31	4.61
		lp	5.89	4.28	2.92	0.57	0.41	3.52
CDX	7Y	*dd*	2.79	1.36	1.15	0.93	0.26	1.54
		dp	2.58	2.89	0.61	0.19	0.14	1.76
		lp	2.31	1.93	0.42	0.18	0.19	1.36
	10Y	*dd*	2.04	10.68	3.00	2.81	1.23	5.23
		dp	3.94	5.51	6.00	0.99	0.16	4.07
		lp	3.99	3.97	3.58	0.95	0.39	3.02
All		*dd*	2.59	7.35	2.38	2.11	0.86	3.78
		dp	4.20	4.76	3.98	0.85	0.19	3.38
		lp	3.94	3.32	2.29	0.56	0.30	2.54

Table 4.9: Root Mean Square Error, ξ of three different mapping methods. Map '*dd*' simply maps based on the relevant detachment level. Map '*dp*' uses the method described in [MBAW04], where detachment is expressed as a percentage of the expected loss on the entire basket. Finally, map '*lp*' considers mapping based on the expected loss on a tranche as a percentage of total expected loss, as described in [RUK04]. Quotes with 5Y maturities on iTraxx and CDX are used as input and errors in value are determined for tranches with 7Y and 10Y maturities, after mapping. The final row considers both indices and both maturities for the tranche. The final column combines the five different tranches.

			Tranche 1	Tranche 2	Tranche 3	Tranche 4	Tranche 5	Total
iTraxx	5Y	*dd*	3.70	1.42	0.66	0.38	0.32	1.81
		dp	4.71	1.74	0.64	0.32	0.35	2.27
		lp	3.71	0.80	0.29	0.18	0.17	1.71
	7Y	*dd*	1.88	1.99	0.83	0.69	0.31	1.32
		dp	4.99	3.73	0.63	0.23	0.16	2.80
		lp	4.50	3.15	0.49	0.26	0.10	2.47
	10Y	*dd*	1.95	8.81	2.73	2.84	1.22	4.44
		dp	8.13	6.67	5.20	1.23	0.19	5.28
		lp	7.95	5.08	3.49	0.74	0.42	4.51
CDX	5Y	*dd*	4.03	1.36	0.59	0.26	0.16	1.92
		dp	4.21	1.92	0.33	0.13	0.11	2.08
		lp	3.83	1.08	0.26	0.10	0.19	1.79
	7Y	*dd*	5.17	1.72	2.03	1.04	0.26	2.64
		dp	2.89	4.96	1.24	0.22	0.18	2.63
		lp	2.56	3.08	0.41	0.18	0.25	1.81
	10Y	*dd*	3.73	10.67	4.19	3.11	1.11	5.59
		dp	3.13	7.35	7.74	1.74	0.29	5.04
		lp	3.17	5.57	4.51	0.62	0.35	3.52
All		*dd*	3.62	5.50	2.14	1.71	0.68	3.21
		dp	4.96	4.68	3.64	0.84	0.23	3.48
		lp	4.58	3.44	2.22	0.40	0.26	2.75

Table 4.10: Root Mean Square Error, ξ of three different mapping methods. Map '*dd*' simply maps based on the relevant detachment level. Map '*dp*' uses the method described in [MBAW04], where detachment is expressed as a percentage of the expected loss on the entire basket. Finally, map '*lp*' considers mapping based on the expected loss on a tranche as a percentage of total expected loss, as described in [RUK04]. Quotes with 5Y maturities on iTraxx are used as input and errors in value are determined for tranches on CDX with 5Y, 7Y and 10Y maturities, after mapping. This is repeated using CDX quotes as input and pricing iTraxx tranches. The final row considers both indices and both maturities for the tranche. The final column combines the five different tranches.

CHAPTER 5

Pricing CDO Tranches with Factor Copula Models

In the previous chapter we have investigated a large data set of CDO tranche prices. In particular, we have looked at the performance of the, market standard, base correlation framework and a number of interpolation, or mapping, schemes applied to this framework. As the base correlation framework uses different correlation values for each CDO tranche, one has to resort to interpolation, in order to value tranches with non standard attachment and detachment points. In this chapter we investigate a number of models as alternatives to the base correlation framework, which all belong to the class of one factor copula models, which has been discussed briefly in Chapter 1. Further, for ease of reading, we present some background of credit derivatives and the base correlation framework.

5.1 Introduction

In recent years, the market for credit derivatives has undergone an enormous growth. The market for credit default swaps has grown from 14 trillion USD in 2005, to over 500 trillion USD in 2007. Further, in 2004 CDO contracts were issued for a total notional of 157 billion USD, which grew to 552 billion USD in 2007. This development has led to the need for models which can be calibrated to the available market quotes. Due to its simplicity, the one factor Gaussian copula has become market standard. As this model is not able to calibrate to all market quotes with a single correlation parameter, the base correlation framework was introduced by McGinty, Beinstein, Ahluwalia, and Watts [MBAW04]. The ideas behind this framework are similar to the implied volatility in the equity derivatives market, as different correlations are used for different CDO tranches.

The main disadvantage of the base correlation framework is that it is not an actual model of default, it merely is an advanced interpolation scheme. Therefore, many authors have suggested different models for the valuation of CDO tranches, with the goal to calibrate to CDO quotes with a single set of parameters. Hull and White [HW04] consider the student t-distribution instead of the Gaussian distribution in their factor copula. Kalemanova et al. [KSW05] consider the normal inverse Gaussian distribution. Other types of extensions have been considered by e.g. Andersen and Sidenius [AS04], who stick to the Gaussian distribution, but they let the correlation depend on the state of the economy. This reflects that correlations tend to increase when the state of the economy deteriorates. Another approach, where the correlation is modeled stochastically, is the mixture factor copula, where a mixture of gaussian copulas is considered. This idea has been discusses by Burtschell et al. [BGL05].

Other models in the factor copula setup that have been proposed are by [Moo06], who considers the variance gamma distribution for the factors. This distribution has been introduced by [MS90] into financial mathematics, and it has been used to match the volatility smile of equity options by [MCC98]. The alpha-stable distributions are considered by [PS06], and [vdV07] considers an, additional, external factor to explain for external causes of default. A brief overview of several of the models above is given by [BGL05].

Next to the models in the factor copula framework, many other models have been considered. Whereas the factor copula models belong to the class of bottom-up models, as one first models the marginal distributions and then the dependence structure, other authors have considered a top-down approach, where one focusses directly on the loss process. Giesecke et al. [GGD05] and [EGG09] consider the general setup of such top-down models, as well as some detailed examples. Brigo et al. [BPT06] and [BPT07] consider the sum of Poisson processes to model the loss process.

In this chapter we investigate the performance of a number of the models described above, where we calibrate the models to a large set of (daily) CDO tranche prices, from December 2004 up to November 2006. We consider the calibration errors for each model and each date at which we calibrate to the market data. Further, we discuss the stability of the optimal model parameters, which could indicate if a model is 'over fitting' to the market data.

This chapter is structured as follows. In Section 5.2 we briefly define the two most popular credit derivatives, the credit default swap and the collateralized debt obligation, and in Section 5.3 we consider the valuation of such derivatives. Thereafter, in Section 5.4 we describe the one factor copula models that we investigate in this Chapter. Section 5.5 describes the approach we have taken to compare the models, and we describe the set of market data that has been used. The results from the tests that we have performed are presented in Section 5.6, where first the calibration errors are compared. Thereafter, we compare the model quotes to the market quotes, to see how well the models match the market data. In Section 5.7 we summarize the results and we conclude with some remarks about the obtained results.

5.2 Credit Derivatives

In this section we briefly describe two of the most liquid credit derivatives currently traded. These are the credit default swap (CDS) and the collateralized debt obligation (CDO).

5.2.1 Credit Default Swaps

The credit default swap is a contract between two parties providing protection against the default of a third party. The party that buys the protection is called the protection buyer, and he is compensated for the losses incurred on the default of the third party, which is called the reference credit. The compensation is paid by the protection seller who in turn receives (periodic) premium payments from the protection buyer. This premium is usually defined in basis points per year on the notional value of the contract. The level of the premium is set such that the contract has zero value at the start of the contract, and it is paid until the maturity of the contract or until the reference credit defaults.

The default event can be defined in several ways, such as bankruptcy, restructuring or failure to pay back a loan. In case of default the settlement of the CDS can be in physical delivery, which means that the protection buyer delivers bonds of the defaulted company and in return it receives the full nominal value of the bonds. Alternatively the settlement can be settled in cash, where the value of the defaulted bonds is determined, the recovery rate, and the protection buyer receives the nominal value of the bonds reduced by this recovery rate.

Similar to the CDS the index CDS provides protection against defaults. In case of the CDS there is only one reference credit but the index CDS has a basket of reference credits as underlyings, each with its own notional value, but usually all underlyings have the same notional value. The defaults of the reference credits results in losses which are paid by the protection seller. In return the protection buyer pays a periodic premium over the remaining notional value of the reference credits which have not defaulted.

5.2.2 Collateralized Debt Obligations

Another credit derivative that provides protection on a basket of reference credits is the collateralized debt obligation (CDO). Whereas the index CDS provides the protection against all defaults during the life of the contract, the CDO provides protection on a part, or tranche, of the basket, which is defined by the attachment and detachment points. As soon as the sum of the losses on the basket reaches the level of the attachment point, the protection buyer start to receive loss payments. He stops receiving loss payments, when the losses reach the detachment point or when the maturity of the contract is reached. In return for this possible loss payments the protection seller receives a periodic premium. In some cases a part of the premium is paid up front at the start of the contract. The CDO can be defined on several types of reference credits, such as loans, bonds or CDSs. In case the underlying basket consists of credit default swaps, we speak of a synthetic CDO. In the analysis that we perform in this chapter we only focus on this latter type of CDO structures.

5.2.3 iTraxx and CDX

In recent years the liquidity in credit derivatives have increased enormously. Currently CDSs on almost all large companies are traded, with maturities of three and six months, and one, two, three, four, five, seven, ten and twelve years. The contracts have also been standardized, such that the maturity always falls on the twentieth of March, June, September or December.

Also the market for CDOs has become much more liquid, which has led to standardization of contracts and to the existence of the iTraxx and CDX indices. These are baskets consisting of (CDSs on) 125 equally weighted European or North-American companies, respectively. On these baskets index CDSs are traded, for which daily quotes are available. Further CDO tranches are traded on these baskets. In case of the European iTraxx basket market quotes are available on five different tranches, where the first tranche, or equity tranche, covers the first 3% of the basket notional. A part of the premium on this tranche is paid upfront together with a periodic premium of 500 basis points per year. The next tranche, or junior mezzanine tranche, covers the next 3% of the basket notional and only a periodic premium is paid. The mezzanine tranche covers the 6% up to 9% region of the basket notional, the senior mezzanine tranche covers the 9% up to 12% region and the (super) senior tranche covers the 12% up to 22% region of the basket loss. Sometimes quotes are available for a sixth tranche, which covers the remaining part of the basket. These latter quotes have not been considered.

For the North-American CDX index, quotes for five year CDO tranches are given in the market. The equity tranche covers the first 3% of the basket notional and a part of the premium is paid upfront and there is a periodic premium of 500 basis points. The junior mezzanine tranche covers the 3% up to 7% region of the basket notional, the mezzanine tranche covers the 7% up to 10% of the losses and the senior mezzanine tranche covers the 10% up to 15% of the notional. The (super) senior tranche covers the 15% up to 30% of the basket notional. Again, quotes for a sixth tranche, covering the remaining part of the basket, can be given.

5.3 Pricing Credit Derivatives

In the previous section we have briefly explained two of the most traded credit derivatives. In this section we give some formulas to determine the value of a trade.

5.3.1 Pricing a Credit Default Swap

We start with the pricing of credit default swaps and we discuss the model that is used in the analysis in the present chapter. In order to price the CDS we need to value the two legs, which are the two payment streams of the swap. The first leg is the default or protection leg (DL) and the second leg is the premium leg (PL). Further we want to determine the fair spread s, which is the spread in basis points per year such that the contract has zero value at the start of the contract.

We write $\tau \geq 0$ for the default time of the reference credit, T for the maturity of the contract and we write $D(t)$ for the risk free discount factor, which we assume to be deterministic. The expected recovery rate is given by R and the notional amount of the contract is given by N. The expected present value of the default leg can be written as

$$\begin{aligned} \mathrm{DL} &= N\,\mathbb{E}\left[D(\tau)\,(1-R)\,1_{\{\tau \leq T\}}\right] \\ &\approx N \sum_{i=1}^{M} D\left(t_i^{\star}\right)\left(1-R_{\text{fix}}\right)\mathbb{P}\left(\tau \leq T\right). \end{aligned} \tag{5.1}$$

The premium is paid periodically, usually every three months at the end of the period. In case the reference credit defaults, an accrued premium payment has to be paid over the period from the previous premium payment date up to the default date. When we denote the premium payment dates by $T_1 < T_2 < \ldots < T_N = T$, and we write $T_0 = 0$, we can write the present value of the fixed leg as

$$\mathrm{PL} = N \sum_{i=1}^{N} \left(D\left(T_i\right) 1_{\{\tau > T_i\}}(t_j - t_{j-1}) + D(\tau) 1_{\{t_{j-1} < \tau \leq t_j\}}(\tau - t_{j-1})\right). \tag{5.2}$$

The value of both legs is obtained by taking the expectations of the present values in (5.1) and (5.2). We write F_τ for the distribution function of τ. The value of the two legs of the credit default swap are given by

$$\mathbb{E}\mathrm{DL} = (1-R)\,N \int_0^T D(t)\mathrm{d}F_\tau(t) \tag{5.3}$$

$$s\,\mathbb{E}\mathrm{PL} = s\,N \sum_{j=1}^{m} \left(D(t_j)(t_j - t_{j-1})(1 - F_\tau(t_j)) + \int_{t_{j-1}}^{t_j} D(t)(t - t_{j-1})\mathrm{d}F_\tau(t)\right). \tag{5.4}$$

We choose to model the default time τ as the first jump of a inhomogeneous Poisson process, with a deterministic piecewise constant intensity. The levels of the intensity allow us to match the market quotes for CDSs with different maturities, by bootstrapping the intensity one maturity at a time, starting with the lowest maturity. The distribution function of τ, or default probability function, is given by $\mathcal{F}_\tau(t) = p(t) = 1 - \exp(-\int_0^t \lambda(s)ds)$. In order to evaluate the integrals in (5.3) and (5.4) we consider a discretization of the interval $[0, T]$ into time points $0 = T_0 < T_1 < \ldots < T_n = T$, and we approximate $D(t)$ on $[T_i, T_{i+1})$ by $D\left(\frac{T_i+T_{i+1}}{2}\right)$ and by $D\left(\frac{t_j+t_{j+1}}{2}\right)$ on $[t_j, t_{j+1})$. This leads to the formulas

$$
\begin{aligned}
\mathbb{E}\mathrm{DL} &\approx (1-R)\,N \sum_{i=1}^{n} D\left(\frac{T_{i-1}+T_i}{2}\right)(p(T_i) - p(T_{i-1})) \\
s\,\mathbb{E}\mathrm{PL} &\approx s\,N \sum_{j=1}^{m} \Bigg(D(t_j)(t_j - t_{j-1})(1 - p(t_j)) \\
&\qquad + D\left(\frac{t_{j-1}+t_j}{2}\right)\frac{t_j - t_{j-1}}{2}(p(t_j) - p(t_{j-1}))\Bigg).
\end{aligned}
$$

The valuation of the index CDS proceeds along the same lines as the valuation of the CDS above. An index CDS can be viewed as a portfolio of CDSs where each CDS has the same spread. Using this observation one can simply obtain the fixed and credit leg of an index CDS by summing the fixed legs and premium legs of the CDSs, respectively.

5.3.2 Pricing a Collateralized Debt Obligation

In order to price a CDO tranche we again have to value the credit and fixed leg. We write $L(t)$ for the cumulative loss process of the underlying basket consisting of K reference credits:

$$
L(t) = \sum_{k=1}^{K} (1 - R_k)\,N_k\,1_{\{\tau_k \le t\}},
$$

where R_k is the recovery rate of reference credit k and N_k is its notional value and its default time is given by τ_k. When we denote the attachment point of the tranche by a and the detachment point by d then we define the tranche loss at time t as $L_{a,d}(t)$ and its value can be obtained from the loss process by

$$
L_{a,d}(t) = \min\left(\max(L(t) - a, 0), d - a\right).
$$

The present value of the two legs can be written in terms of the tranche loss process. This gives

$$\mathrm{DL}_{a,d} = \int_0^T D(t)\mathrm{d}L_{a,d}(t)$$

$$s\,\mathrm{PL}_{a,d} = s\sum_{j=1}^{m} D(t_j)\Bigg(((d-a) - L_{a,d}(t_j))(t_j - t_{j-1}) + \sum_{k=1}^{K} 1_{\{t_{j-1}<\tau_k\le t_j\}}(\tau_k - t_{j-1})\Delta L_{a,d}(\tau_k)\Bigg).$$

Here $\Delta L_{a,d}(\tau_k)$ is the change in the loss amount on the CDO tranche (a,d) due to a default of name k at time τ_k.

The values of the legs are obtained by taking the expectation of both expressions. The values can be approximated using the same discretization and assumptions as in the previous section. A straightforward calculation shows that

$$\mathbb{E}\mathrm{DL}_{a,d} \approx \sum_{i=1}^{n}\left(\mathbb{E}L_{a,d}(T_i) - \mathbb{E}L_{a,d}(T_{a,d})\right) D\left(\frac{T_i + T_{i-1}}{2}\right) \tag{5.5}$$

$$s\,\mathbb{E}\mathrm{PL}_{a,d} \approx s\sum_{j=1}^{m}(t_j - t_{j-1})\left(1 - \frac{\mathbb{E}L_{a,d}(t_j) + \mathbb{E}L_{a,d}(t_{j-1})}{2}\right) D(t_j) \tag{5.6}$$

From formulas (5.5) and (5.6) it is easy to see that in order to value both legs of a CDO tranche one only has to calculate expected tranche losses at the premium payment dates t_j and at the time points T_i. In the next sections we discuss techniques and models which can be used to determine these expected losses.

5.4 Factor Copula Models

In the previous section we have seen that to value CDO tranches, one has to calculate expected tranche losses at a number of points in time. In order to calculate these expected losses one has to model the dependence structure between the different reference credits in the basket. In the analysis in this chapter we consider several factor copula models which can exactly do this. The advantage of (factor) copula models is that they allow one to model the dependence structure and marginal distributions separately. They also allow one to condition on a common factor, by which default times become independent. We start with a general description of the factor copula models and we consider a technique that can be used to determine the loss distribution. In the remainder of the section we consider the standard model and a number of alternatives to this model.

5.4.1 General Setup

We define the factor copula models through a set of random variables, which determine the dependence structure. For each of the K companies in the basket underlying the CDO we define a random variable X_k by

$$X_k = Z + E_k, \quad k = 1, \ldots, K, \tag{5.7}$$

where Z and the $E_k's$ are independent and $Z \sim F_Z$ and $E_k \sim F_{E_k}$. Further, we write F_{X_k} for the distribution of X_k. This factor setup has been introduced to the pricing of CDOs by [Li00], where the normal distribution is considered for all factors. The distribution function F_{X_k} can be obtained from the distribution function of Z and E_k, by conditioning on the factor Z which results in

$$P(X_k \leq x) = \int_{\mathbb{R}} F_{E_k}(x - Z)\mathrm{d}F_Z(x).$$

The factor Z is used to model the global state of the economy. A low value corresponds to a bad state of the economy, and a high value corresponds to a good state of the economy. The factor E_k models the state of the individual company. Again, a low value represents a bad state and a high value represents a good state of the company.

In order to link the static factor copula structure to the marginal distributions, that were obtained in Section 5.3.1, we assume that the company k defaults as soon as a non-decreasing time-dependent barrier $\chi_k(t)$ reaches the level X_k. This results in default probabilities

$$P(\tau \leq t) = P(X_k \leq \chi_k(t)).$$

We, naturally, want both probabilities to be equal and this implies the value of the barrier. Using $p_k(t)$ to denote the marginal default probability of company k then we have

$$\chi_k(t) = F_{X_k}^{-1}\left(p_k(t)\right).$$

In order to calculate the expected (tranche) losses we need to know the loss distribution. Therefore we want to compute probabilities of the form $P(L(t) \leq \ell)$, for certain loss levels ℓ. We discretize the loss range by introducing δ_ℓ, which described the minimum loss step. This step is such that $(1 - R_k)N_k = n_k\delta_\ell$, for some integer n_k. With this discretization $L(t)$ has a discrete distribution on the points $n\,\delta_\ell$, with $n \in \mathbb{N}$.

We can build the loss distribution using ideas explored by [ASB03] and [HW04]. We first condition on the factor Z and then use the independence of the factors E_k. The loss distribution can be obtained using the following algorithm, where we add one company at a time to the distribution of $L(t)$.

1) Start with $P^{(0)}(L(t) = n\delta_\ell | Z) = 1_{\{n=0\}}$, for $n = 0, 1, 2, \ldots$.

2) Assume that we are at step k and we want to determine $P^{(k+1)}(L(t) = n\delta_\ell|Z)$ by adding company $k+1$ to the distribution. Then we have

$$P^{(k+1)}(L(t) = n\delta_\ell|Z) = P^{(k)}(L(t) = n\delta l|Y)(1 - F_{E_{k+1}}(\chi_{k+1}(t) - Z)) + P^{(k)}(L(t) = (n - n_{k+1})\delta_\ell|Y)F_{E_{k+1}}(\chi_{k+1}(t) - Z)).$$

Repeat this step for $k = 1, 2, \ldots, K$.

3) Obtain the conditional loss distribution by

$$P(L(t) = n\delta_\ell|Z) = P^{(K)}(L(t) = n\delta_\ell|Z).$$

The unconditional loss distribution is obtained by integrating out over factor Z,

$$P(L(t) = n\delta_\ell) = \int_{D_Z} P(L(t) = n\delta_\ell|Z = x)dF_Z(x)$$

With this loss distribution we can easily calculate expected (tranche) losses as

$$\mathbb{E}L(t) = \sum_{k=1}^{\lceil K/\delta_\ell \rceil} k\delta_\ell P(L(t) = k\delta_\ell)$$

$$\mathbb{E}L_{a,d}(t) = \sum_{k=1}^{\lceil K/\delta_\ell \rceil} \min(\max(k\delta_\ell - a, 0), d - a)P(L(t) = k\delta_\ell)$$

5.4.2 Standard Model

With the formulas to determine the loss distribution in place, we can focus on specific models which fall into the general setup of formula (5.7). In this section we consider the standard model where Z and the E_k's have a normal distribution. To be more precise, we let

$$Z \sim N(0, \rho^2), \;\; E_k \sim N(0, 1 - \rho^2),$$

for a correlation parameter $0 \leq \rho \leq 1$. We can then rewrite (5.7) to the equivalent formulation,

$$X_k = \rho Y + \sqrt{1 - \rho^2}\, \varepsilon_k,$$

where $Y, \varepsilon_k \sim N(0, 1)$ and hence $X_k \sim N(0, 1)$. The resulting factor copula model is often referred to as the standard model or the one factor Gaussian copula. The one factor Gaussian copula has first been suggested by [Li00].

The main drawback of this model is that it is not able to match the market quotes on the European iTraxx and North-American CDX indices. This is illustrated in Table 5.1

Tranche	Market Quote	Model Quote
(0%,3%)	23.00 %	23.00 %
(3%,6%)	131.0 bp	187.26 bp
(6%,9%)	43.5 bp	45.25 bp
(9%,12%)	27.25 bp	12.18 bp
(12%,22%)	14.25 bp	1.43 bp

Table 5.1: Market and Model Quotes for CDO tranches on the iTraxx index on valuation date 13 December 2004 with a maturity of five years. The correlation parameter has been chosen such that the upfront premium on the equity tranche is matched.

for a set of CDO tranches on the iTraxx index. In this table we can see that, if we choose the correlation such that the upfront premia is matched, the standard model overestimates the premium on the junior mezzanine tranche and it underestimates the premium on the more senior tranches.
A solution to this problem is given through the base correlation framework, where a CDO tranche (a, d) is split up into two tranches $(0, d)$ and $(0, a)$. Starting with the equity tranche, one can try to find a correlation value such that the upfront premium on this tranche is matched. Next one iterates over the remaining tranches by finding correlation values such that the annual premia are matched. At each step one uses the correlation from the previous step to price the tranche $(0, a)$ and the tranche $(0, d)$ is priced with the new value of the correlation parameter. The problem of this approach is that it is not a model of default times. Furthermore it is not clear what correlation parameters one should use to value a nonstandard CDO tranche, and one has to resort to interpolation, which could lead to arbitrage opportunities, as the expected loss on base tranches, as function of detachment point, might not be increasing. A more detailed description of the base correlation approach can be found in [AM04] and [ABM04].

Due to the deficiencies of the standard model and the base correlation several alternative models have been proposed. In the analysis in this chapter we restrict ourselves to models that fit the general setup of Equation (5.7). In the remainder of this section we describe the alternative models that we consider in this chapter.

5.4.3 Double *t*-Copula

The first alternative model that we consider is the double t-copula. This model is obtained by assuming a student t-distribution for the factors Z and E_k. The density of random variable V with a t_ν distribution, with $\nu > 0$ degrees of freedom, is given by

$$f_\nu(t) = \frac{\Gamma\left(\frac{\nu+1}{2}\right)}{\sqrt{\nu\,\pi}\,\Gamma\left(\frac{\nu}{2}\right)}\left(1+\frac{t^2}{\nu}\right)^{-(\nu+1)/2}$$

$$\Gamma(x) = \int_0^\infty t^{x-1}e^{-t}\mathrm{d}t$$

In case $\nu > 1$ the expectation of V is zero and, in case $\nu > 2$ the variance of V is $\nu/(\nu-2)$. For low values of ν The student t-distribution has much fatter tails than the normal distribution, which makes extreme outcomes more likely.

Using t-distributions for both factors has been considered by [HW04], where the following description is considered.

$$
\begin{aligned}
Z &= \sqrt{\frac{\nu_Y - 2}{\nu_Y}}\,\rho\, Y \\
E_k &= \sqrt{\frac{\nu_\varepsilon - 2}{\nu_\varepsilon}\,(1-\rho^2)}\,\varepsilon_k,
\end{aligned}
$$

where $Y \sim t_{\nu_Y}$ and $\varepsilon \sim t_{\nu_\varepsilon}$, with $\nu_Y, \nu_\varepsilon > 2$ and $0 \le \rho \le 1$. The factors are scaled by the inverse of their variance, such that X_k has zero expectation and a variance of one.

In the beginning of this section we have seen that for the Student t-distribution we have $\nu > 0$. The description given above only allows for $\nu > 2$, which excludes the use of other t-distributions. Therefore we consider a different description by removing the scaling factors in front of both factors, which results in the double-t factor copula.

$$
\begin{aligned}
Z &= \rho\, Y \\
E_k &= \sqrt{1-\rho^2}\,\varepsilon_k,
\end{aligned}
$$

where we still have $Y \sim t_{\nu_Y}$ and $\varepsilon \sim t_{\nu_\varepsilon}$, only in this case we have $\nu_Y, \nu_\varepsilon > 0$. This description does allow us to use a wider range of degrees of freedom, as the range from 0 up to 2 is available as well. In the analysis in this chapter we consider this latter description because of the wider parameter range. This wider range is important as tests show that in some cases one finds optimal parameter values where the number of degrees of freedom are below two.

5.4.4 Normal Inverse Gaussian Factor Copula

Similar to the student t-distribution we consider the normal inverse Gaussian (NIG) distribution, where we, compared to the standard model, also consider different distributions for the factors. The NIG distribution has the density

$$
\begin{aligned}
f_{\alpha,\beta,\mu,\delta}(x) &= \frac{\delta\alpha \exp(\delta\gamma + \beta(x-\mu))}{\pi\sqrt{\delta^2 + (x-\mu)^2}} K_1\left(\alpha\sqrt{\delta^2 + (x-\mu)^2}\right) \\
\gamma &= \sqrt{\alpha^2 - \beta^2} \\
K_1(\omega) &= \frac{1}{2}\int_0^\infty \exp\left(-\frac{1}{2}\omega\left(t + t^{-1}\right)\right) \mathrm{d}t,
\end{aligned}
$$

where $0 < |\beta| < \alpha$ and $\delta > 0$. The expectation and variance of a random variable V following an $NIG(\alpha, \beta, \mu, \delta)$ distribution are given by

$$\begin{aligned} \mathbb{E}V &= \frac{\mu\gamma + \beta\delta}{\gamma} \\ Var(V) &= \frac{\delta\alpha^2}{\gamma^3}. \end{aligned}$$

Further the NIG distribution satisfies the properties

$$V \sim NIG(\alpha, \beta, \mu, \delta) \Rightarrow cV \sim NIG\left(\frac{\alpha}{c}, \frac{\beta}{c}, c\mu, c\delta\right) \tag{5.8}$$

$$V_i \sim NIG(\alpha, \beta, \mu_i, \delta_i), \text{ for } i \leq M \Rightarrow \sum_{i=1}^{M} V_i \sim NIG\left(\alpha, \beta, \sum_{i=1}^{M} \mu_i, \sum_{i=1}^{M} \delta_i\right) \tag{5.9}$$

We obtain the normal inverse Gaussian factor copula from the general description (5.7) setting

$$\begin{aligned} Z &= \rho Y \\ E_k &= \sqrt{1-\rho^2}\,\varepsilon_k, \end{aligned}$$

where we assume that the common factor Y is distributed as

$$Y \sim NIG\left(\alpha, \beta, -\frac{\alpha\beta}{\sqrt{\alpha^2 - \beta^2}}, \alpha\right)$$

and the idiosyncratic factor ε as

$$\varepsilon_k \sim NIG\left(\frac{\sqrt{1-\rho^2}}{\rho}\alpha, \frac{\sqrt{1-\rho^2}}{\rho}\beta, -\frac{\sqrt{1-\rho^2}}{\rho}\frac{\alpha\beta}{\sqrt{\alpha^2-\beta^2}}, \frac{\sqrt{1-\rho^2}}{\rho}\alpha\right).$$

From Equation (5.8) and (5.9) it follows that

$$X_k \sim NIG\left(\frac{\alpha}{\rho}, \frac{\beta}{\rho}, -\frac{\alpha\beta}{\rho\sqrt{\alpha^2-\beta^2}}, \frac{\alpha}{\rho}\right).$$

The parameters are chosen such that X_k has zero expectation. Using the normal inverse Gaussian distribution in this form for the pricing of CDO tranches has been proposed by [KSW05]. In this paper the authors price CDO tranches in the large homogeneous pool setup, where it is assumed that the underlying basket is homogenous and its size is infinite. In this setup one ignores the different characteristics of the companies in the underlying basket, but it allows for (much) faster calculations.

Just as the student t-distribution, the NIG distribution has fatter tails than the normal distribution, and thus extreme events, such as a large number of losses, should be more likely.

5.4.5 Random Correlation

The previous two models changed the distribution of X_k by changing the distributions of the factors. An alternative approach is to make the correlation random and maintaining the normal distribution of both factors. We consider two different models where the correlation is random.

Mixture Factor Copula

The first model with random correlation that we consider is a mixture of one factor Gaussian copulas. We assume that the correlation can take on n different levels $\rho_1, \ldots, \rho_n$ with probability $p_1, \ldots, p_n$, with $\sum_{i=1}^n p_i = 1$. We arrive at the mixture factor copula from the general description (5.7) by setting

$$Z = \sum_{i=1}^n 1_{\{B=i\}} \rho_i Y$$
$$E_k = \sum_{i=1}^n 1_{\{B=i\}} \sqrt{1-\rho_i^2}\,\varepsilon_k,$$

where B is a discrete random variable that can take the values $i = 1, \ldots, n$ with probabilities p_i, and $Y, \varepsilon_k \sim N(0,1)$. In the tests in Section 5.6 we will use the mixture factor copula with $n = 3$.

This simple extension of the standard model represents the uncertainty about the correlation state. Typical values of the correlation will include a very low correlation, which can be associated with an economy that is in a good state, and a very high correlation value, which corresponds to a bad state of the economy. The remaining correlation parameters is expected to lie somewhere in the middle, representing a 'normal' state of the economy.

The mixture factor has been considered by, amongst others, [BGL07], where also the three state version is considered. Test results for a small number of valuation dates in this paper suggest that the mixture factor copula can be calibrated well to market data. Further they provide some figures with calibrated parameters values over a larger data set, without providing the quality of the fits. In the current chapter we investigate the calibration on a larger data set and we consider both the iTraxx and CDX indices, where [BGL07] only consider the iTraxx index on a large data set.

Random Factor Loadings Copula

Another way to introduce randomness in the correlation parameters is by making the correlation parameter dependent on the factor Y. As we already briefly mentioned above, correlations tend to by higher when the market is in a bad state and they tend to be low when the market is in a good state. We choose Z and E_k such that they incorporate this behavior, and such that X_k has zero expectation and a variance of one.

$$Z = \rho(Y)Y \tag{5.10}$$

$$E_k = \sqrt{1 - Var(\rho(Y)Y)}\,\varepsilon_k + \mathbb{E}\rho(Y)Y \tag{5.11}$$

$$\rho(y) = \alpha + \beta \frac{\exp(-\delta(y-\gamma)}{1+\exp(-\delta(y-\gamma))}. \tag{5.12}$$

Again we let $Y, \varepsilon_k \sim N(0,1)$ and further we let $\alpha, \beta \geq 0$, $\alpha + \beta \leq 1$, $\delta > 0$. The function to let the correlation depend on state of the economy is chosen such that it always lies between zero and one, and it is strictly decreasing. Further we can adjust the function easily through the four scaling and shifting parameters.

The general setup of the random factor loading factor copula is suggested by [AS04]. In their paper they use a much simpler description for the function $\rho(y)$, which can take on only two different values. With the formulation (5.12) we let the correlation depend on the state of the economy in a continuous way. In a paper by [BGL07] a similar model is considered, where the correlation also depends on the state of the economy. The idiosyncratic term is defined differently though, as the term in front of the ε_i varies with the correlation function as well, instead of in our description, where we use the variance of $\rho(Y)Y$ and its expectation. Further the authors in [BGL07] try to directly calibrate the shape of the function for the correlation as a function of Y, where we have parameterized this function. Some tests with both models, with the same description for the correlation function, have shown a better performance for the random factor loading copula as described in (5.10) - (5.12). Therefore we have chosen to consider this description in the analysis in this chapter.

5.5 Approach

In this chapter we compare the models discussed in Sections 5.4.2 up to 5.4.5. First we want to see if the extensions can outperform the standard model, which is highly likely as the standard model can be obtained from the alternative models. In case of the Double-t factor copula we can let the degrees of freedom go to infinity, such that the distribution of the factors equals the normal distribution. For the mixture factor copula we can simply let all correlation values be equal and for the random factor loadings factor copula we set $\beta = 0$. The standard model cannot be obtained from the NIG factor copula.

Next we want to see if we can reproduce market data by calibrating the models to the market data such that the data is matched as closely as possible. Further we are interested if the performance of the models changes over the different points in the data set.

In this section we will mainly focus on the different techniques that we use to find these optimal parameters. Further we discuss some (possible) issues one might encounter when trying to find these optimal parameters. We conclude this section by describing the market data that we use for the analysis. The results of this analysis are discussed in the next section.

5.5.1 The Sum of Squared Errors

The market data for one day includes CDS spreads for each company in the basket underlying the CDO tranches. Further we have quotes for the five CDO tranches, either on the European iTraxx basket or on the North-American CDX basket. The CDS spreads are matched by choosing the levels of the piecewise constant intensity, starting with the first period.

In order to match the CDO tranche data as closely as possible for a given day we need to define the distance between the market data and the fair upfront or the fair spread which results from the alternative model. Therefore we consider the sum of squared errors between the market quotes and the fair spreads obtained by an alternative model.

$$S(\mathbf{p}) := \sum_{i=1}^{5} w_i \left(Q_i - \bar{s}_i(\mathbf{p})\right)^2, \tag{5.13}$$

where Q_i represents the market quote for tranche i and $\bar{s}_i$ represents the fair spread for tranche i obtained with the set of parameters $\mathbf{p}$, which is different for each model. Further we have introduced weights w_i. In case $i = 1$ we consider the equity tranche which in our analysis always has attachment point 0% and detachment point 3%, and the market quote and the fair spread are in this case an upfront percentage of the tranche notional. Together with this upfront premium a spread of 500 basis point has to be paid per year on the outstanding notional. In case $i > 1$, Q_i and $\bar{s}_i$ represent the premium in basis points which has to be paid per year on the outstanding notional of the tranche.

Before we are able to compute to sum of squared error $S(\mathbf{p})$ we need to define the weights that will be used. We use two different kinds of weights. First we consider the errors relative to the market quotes, which means that we choose $w_i = Q_i^{-2}$. An alternative set of weights that we consider is given through $w_i = Q_i^{-1}$. The former set of weights favors market quotes with low values, which are observed for the most senior tranches. The latter set of weights, compared to the forme, puts more weight on market quotes with higher values, as one divides by a smaller number and hence the contribution to the SSE becomes larger. Another possible choice of w_i is $w_i \equiv 1$, which measures absolute errors. This choice puts high weights on tranches with high spreads, and very low weight on tranches with very low spreads. Therefore we do not consider this choice in the analysis in this chapter.

5.5.2 Finding the Minimum of the SSE

In order the fit the alternative models to the market data, we need to find the set of optimal parameters such that $S(\mathbf{p})$ reaches its minimum value. When finding this minimum one needs to keep parameters restrictions in mind. It is thus clear that we have a constrained minimization problem.

Before we try to solve this problem, we first transform the constrained minimization problem into a unconstrained minimization problem by defining transformations from $\mathbb{R}$ to the (possibly) bounded region for each parameter. The transformations are described in Section 5.5.2. After we have obtained the unconstrained problem we apply the Levenberg-Marquardt algorithm, which is designed for unconstrained minimization problems, where the objective function is of the form (5.13). In the appendix we briefly describe this algorithm. Using the Levenberg-Marquardt algorithm we are able to obtain real valued optimal parameters for the unconstrained problem. By applying the parameter transformations to these optimal values we obtain the optimal parameter values for the constrained problem.

Parameters Transformation

As we discussed in the previous section, the Levenberg-Marquardt algorithm only works on real valued parameters. In our analysis the parameters are usually restricted to some interval or to a part of the real line. Further, parameters may depend on the value of other parameters. We denote the set of valid parameters values $\mathbf{p}$ by $\Theta \subset \mathbb{R}^k$. In this section we consider some transformations to map Θ to $\mathbb{R}$ and vice versa. Some restriction on these transformation are that they should be bijective and continuous. The first transformation maps the interval (a, b), with $-\infty < a < b < \infty$ to $\mathbb{R}$. This transformation can be applied to the correlation which should be in the interval $(0, 1)$. Together with the transformation we also provide its inverse.

$$x \mapsto \log\left(\frac{x-a}{b-x}\right) : \ (a,b) \to \mathbb{R}$$

$$y \mapsto a + (b-a)\frac{\exp(y)}{1+\exp(y)} : \ \mathbb{R} \to (a,b).$$

These mappings are bijective and continuous. If we a have parameters which bounds are determined by the value of other parameters, e.g. $g(p_1) < p_2 < h(p_1)$ we can still apply this transformation as soon as the value of this parameter p_1 is known. We just have to set $a = g(p_1)$ and $b = h(p_1)$.

Another transformation that we need is a transformation that maps the interval (a, ∞) onto $\mathbb{R}$. This can easily be done by the transformations

$$x \mapsto \log(x-a) : \ (a,\infty) \to \mathbb{R}$$

$$y \mapsto a + \exp(y) : \ \mathbb{R} \to (a,\infty).$$

Again it is directly clear that the transformations are bijective and continuous.

We only consider transformation for open intervals. In case the parameters are restricted to a closed interval we can safely ignore the parameter values at the boundaries since the objective functions depend continuously on the parameters, and thus we can get arbitrarily close to the boundaries.

Choice of Initial Parameters and Numerical Issues

Above we have shown how we can apply the Levenberg-Marquardt algorithm to the constrained optimization problem (5.13). In order to start this algorithm an initial choice for the parameters has to be made. With the initial choice of parameters the algorithm starts looking for a new set of parameters where the sum of squared errors is smaller, by calculating derivatives with respect to the parameters in the starting point. As we do not have an analytical expression it is not possible to determine an analytical expression for the derivative with respect to the parameters. Therefore we have to resort to numerical derivatives by changing the parameter by a small amount, e.g. 10^{-5}.

In order to obtain an accurate approximation to the minimum of the SSE we need to make a good initial guess. If we choose the 'wrong' initial condition, the SSE might be very high, and the algorithm could end up in a local minimum that is significantly higher than the global minimum. We deal with this problem by randomly calculating initial parameters until the SSE is lower than a certain level. This level is obtained as a fraction of the global minimum obtained with the standard model. In this way we are guaranteed to end up in a minimum that is lower than the minimum obtained with the standard model. We thus avoid that the Levenberg-Marquardt algorithm chooses parameters such that the alternative model is equivalent to the standard model. On the other hand we still might not end up in the global minimum.

Another problem is that the global minimum could be in an almost flat region, by which we could end up very close to the minimum with parameters that are far from the optimal parameters. This problem can also be caused by numerical integration routines where we approximate integrals up to a certain accuracy. Due to small errors in this integration the flat region might not be smooth enough, which results in a many local minima in this flat region, all of them having approximately the same value.

5.5.3 Market Data

The analysis in this chapter is performed on a data set of CDO tranche quotes on iTraxx as well as CDX, which was provided by ABN AMRO bank. The data ranges from December 2004 up to November 2006. For a large number of days in this set we have CDS quotes from which we can obtain marginal distributions. Further we have CDO tranche quotes for maturities of five years to which we calibrate the alternative model. Further the data set includes discount curves which are used to determine the present value of future payments.

For some companies in the iTraxx or CDX indices we have some incomplete data, where CDS quotes are only available for a limited number of days. If we calibrate the models on such a day, we use the CDS data on the day that is closest to the valuation date. Further, for some days in the data set we have multiple quotes for each CDO tranche. In these cases we compare the quotes and pick the set that seems most reasonable compared to the days before and after this date.

As this data set ranges from December 2004 up to November 2006 we do not have data for credit derivatives during the credit crisis which started in July 2007. During this period quotes have increased significantly for CDSs as well as for CDO tranches, especially the more senior tranches. The current data set does contain data from May 2005, when Ford's and General Motors' credit rating were downgraded. This caused some shocks in the market and the structure of the quotes changed which resulted in a steeper base correlation curve, which indicates a different correlation structure. When calibrating the alternative models it is thus interesting to see if models fit better before or after May 2005.

In the next section we discuss the results from the calibration of the alternative models to the data described above.

5.6 Results

In the previous section we have discussed the numerical techniques and the data for the tests that we have performed. In this section we discuss the results of these tests and we look which model calibrates best to the market data. For this we use several measures such as the sum of squared errors and the actual differences between market and model quotes. Further we look at the stability of the parameter values and the stability of the sum of squared errors.

5.6.1 The Sum of Squared Errors

First we consider the sum of squared errors (SSE) which is the objective function of the calibration. In case the model fits perfectly to the market data this function has zero value. In any other case it has a positive value. As we discussed in Section 5.5.1 we consider two different sets of weights w_i in the sum of squared errors, that is defined in Equation (5.13). We calibrate the models with $w_i = Q_i^{-2}$ as well as $w_i = Q_i^{-1}$. For each valuation day in the data set we calibrate the models the market data, which results in a set of SSEs. In Table 5.2 we present the average, minimum, maximum and standard deviation of these SSEs corresponding to the weights $w_i = Q_i^{-2}$. The corresponding CDO tranches have the iTraxx index as underlying and they have a five year maturity.

From this table we can see that the alternative models all perform much better than the standard model. On the other hand we see that the models, except for the mixture factor copula are not able to provide a perfect fit to the market data. In the case of the mixture factor copula we observe the perfect fit for only one day in the data set. Further we see that the differences between the averages are not very large.

	Standard	Double-t	NIG	Mixture	RFL
avg	2.257	0.0672	0.0574	0.0457	0.0565
min	0.734	0.0043	0.0010	0.0000	0.0002
max	3.38	0.366	0.317	0.322	0.310
stdev	0.69	0.062	0.051	0.053	0.048

Table 5.2: The average, minimum, maximum and standard deviation of the SSE for the standard model, the double-t factor copula (Double-t), the normal inverse Gaussian factor copula (NIG), the mixture factor copula (Mixture) and the random factor loadings copula (RFL). The values are obtained by calibrating the models to CDO tranches on the iTraxx index with five year maturities. The weights in the SSE (5.13) are given by $w_i = Q_i^{-2}$.

The same analysis has been performed for CDO tranches on the North-American CDX index, again with maturities of five years. In Table 5.3 the results are presented corresponding to the weights $w_i = Q_i^{-2}$.

SSE	Standard	Double-t	NIG	Mixture	RFL
average	2.159	0.0642	0.0622	0.0747	0.0573
minimum	0.666	0.0009	0.0003	0.0002	0.0007
maximum	3.50	0.271	0.194	0.537	0.211
st.deviation	0.70	0.056	0.044	0.106	0.048

Table 5.3: The average, minimum, maximum and standard deviation of the SSE for the standard model, the double-t factor copula (Double-t), the normal inverse Gaussian factor copula (NIG), the mixture factor copula (Mixture) and the random factor loadings copula (RFL). The values are obtained by calibrating the models to CDO tranches on the CDX index with five year maturities. The weights in the SSE (5.13) are given by $w_i = Q_i^{-2}$.

From this table we can draw similar conclusions as from Table 5.2, although the performance of the mixture factor copula is slightly worse in this case. This is mainly due a small number of dates on which the mixture factor copula calibration performs poorly.

In Tables 5.4 and 5.5 below, the same statistics are presented, but in this case we use the weights $w_i = Q_i^{-1}$. With this choice the resulting SSEs are larger than in the previous case, since we have $Q_i^{-2} < Q_i^{-1}$ as long as $Q_i > 1$, which is (almost) always the case. Further we expect that we fit better to tranches with high spreads, and worse to tranches with low spreads.

In these two tables, we observe that the calibration errors are substantially larger that in Tables 5.2 and 5.3, which was to be expected, as the pricing errors are divided by the quote instead of the squared quote. Further we again observe that alternative models performs better than the standard model. The mixture factor copula again shows a number of large calibration errors, which results in the higher average.

Tables 5.2 - 5.5 only give a general view on the performance of the models. The statistics in these tables can be influenced by some outliers. It could be the case that

	Standard	Double-t	NIG	Mixture	RFL
avg	41.9	1.61	1.47	1.50	1.33
min	17.1	0.107	0.026	0.000	0.004
max	143	8.75	17.7	27.2	8.57
stdev	19.5	1.72	2.11	3.25	1.33

Table 5.4: The average, minimum, maximum and standard deviation of the SSEs, calculated on each valuation date for the standard model, the double-t factor copula (Double-t), the normal inverse Gaussian factor copula (NIG), the mixture factor copula (Mixture) and the random factor loadings copula (RFL). The values are obtained by calibrating the models to CDO tranches on the iTraxx index with five year maturities. The weights in the SSE (5.13) are given by $w_i = Q_i^{-1}$.

	standard	Double-t	NIG	mixture	RFL
avg	57.0	2.28	1.79	6.13	2.04
min	21.0	0.023	0.012	0.006	0.016
max	337.6	17.8	20.8	97.1	12.0
stdev	44.0	2.89	2.56	13.9	2.34

Table 5.5: The average, minimum, maximum and standard deviation of the SSE for the standard model, the double-t factor copula (Double-t), the normal inverse Gaussian factor copula (NIG), the mixture factor copula (Mixture) and the random factor loadings copula (RFL). The values are obtained by calibrating the models to CDO tranches on the CDX index with five year maturities. The weights in the SSE (5.13) are given by $w_i = Q_i^{-1}$.

a model can provide a good fit to market data, except for a small number of dates on which a poor fit is found, as is the case for the mixture factor copula. In this case the average and maximum might not be a good measure of the performance of the model. We therefore have added figures where the SSE are plotted for each date in the data set. The graphs of the SSEs for the four different cases can be found in Appendix 5.B in Figures 5.1 - 5.4. In these figures we see that in general the SSE are quite close for the four models, although for a small number of dates in the first half of the data set we can observe that the mixture factor copula performs worse than the other models, when calibrating the CDO quotes on the CDX index. In the second half of the data set however, we see that the mixture copula performs better than the other models. From the figures, it does not become clear whether the models calibrate better before or after May 2005.

5.6.2 Fair Premia vs. Market Quotes

In the previous section we have investigated the SSE, where we have looked at the average SSE for each of the alternative models, and compared these values with the standard model and with the other alternative models. Further, we have looked at some figures showing the SSE for each day in the data set. The SSE, even for a single day, only gives a global view on the calibration performance of the models. From this statistic it does not become clear how well, or how poor, the model fits to each

CDO tranche. In this section we look at the fair spreads, that are obtained from the model with the optimal parameters for each CDO tranche, and we compare these to the corresponding market quotes. This makes it easier to compare the performance of the models and it becomes easier to compare the calibration with the different sets of weights which is not possible by only looking at the SSE.

We first consider the weights Q_i^{-2}, where we calculate the average absolute differences as well as the relative absolute difference with respect to the corresponding market quote. In Tables 5.6 and 5.7 the averages are presented.

CDO tranche	differences	Standard	Double-t	NIG	Mixture	RFL
(0%,3%)	absolute	4.26	4.08	3.74	3.65	4.16
	relative	16.79%	16.52%	15.60%	14.44%	16.86%
(3%,6%)	absolute	67.97	7.74	5.02	5.93	4.45
	relative	76.73%	9.02%	5.61%	6.47%	4.90%
(6%,9%)	absolute	9.43	2.11	1.57	1.42	1.76
	relative	33.91%	8.28%	5.87%	5.01%	6.52%
(9%,12%)	absolute	8.62	0.81	1.07	0.46	0.81
	relative	67.13%	5.69%	7.47%	3.18%	6.39%
(12%,22%)	absolute	7.11	0.38	0.27	0.25	0.21
	relative	93.55%	4.08%	3.85%	2.57%	2.52%

Table 5.6: The averages of the absolute and absolute relative differences between the fair spread obtained with the alternative models and the market quotes. The CDO tranches have the iTraxx index as underlying and the weights Q_i^{-2} are used.

CDO tranche	differences	Standard	Double-t	NIG	Mixture	RFL
(0%,3%)	absolute	4.82	4.32	4.09	3.98	4.80
	relative	12.69%	11.48%	11.04%	10.56%	12.71%
(3%,7%)	absolute	100.08	15.22	13.12	20.82	11.61
	relative	78.42%	10.04%	10.12%	13.33%	7.25%
(7%,10%)	absolute	9.23	3.14	2.13	2.56	2.92
	relative	22.72%	8.73%	5.89%	6.86%	8.64%
(10%,15%)	absolute	9.28	1.03	1.23	0.62	1.28
	relative	62.94%	7.78%	9.23%	3.82%	8.54%
(15%,30%)	absolute	7.29	0.56	0.46	0.38	0.31
	relative	94.19%	6.29%	6.74%	3.48%	3.48%

Table 5.7: The averages of the absolute and absolute relative differences between the fair spread obtained with the alternative models and the market quotes. The CDO tranches have the CDX index as underlying and the weights Q_i^{-2} are used.

The results given in these tables give more insights into the performance of the models, since we have broken down the calibration error over the five different tranches. We again observe the superior behavior of the alternative models with respect to the standard model. Further we can see that the absolute as well as the relative errors are reasonably small for the most senior tranches. The errors on the equity and junior mezzanine tranches are quite large. This is what we expected from using these types of weights, as smaller quotes are given larger weights. Comparing the alternative models, we can observe that none of the models performs better than the other, as one model performs better on one tranche and another model performs better on a different tranche.

Next we consider the same statistics, only in this case we use the weights Q_i^{-1}. In Section 5.5.1 we have mentioned that the two sets of weights should provide a different fit to the market data, as the first set of weights, $w_i = Q_i^{-2}$, focusses more on the more senior tranches and the second set, $w_i = Q_i^{-1}$ focusses less on the senior tranches and more on the junior mezzanine tranches. We can thus expect that the errors on the senior tranches become larger and the errors on the junior mezzanine tranches become smaller.

Tables 5.8 and 5.9 present the results from calibrating with weights $w_i = Q_i^{-1}$.

CDO tranche	differences	Standard	Double-t	NIG	Mixture	RFL
(0%,3%)	absolute	6.95	3.99	3.82	3.62	4.27
	relative	27.95%	16.18%	15.93%	14.31%	17.30%
(3%,6%)	absolute	18.29	2.33	2.18	2.75	1.36
	relative	17.17%	2.77%	2.36%	2.84%	1.54%
(6%,9%)	absolute	18.40	2.29	1.48	1.61	1.35
	relative	74.16%	9.26%	5.57%	5.86%	4.94%
(9%,12%)	absolute	12.77	1.10	1.17	0.76	1.19
	relative	92.63%	8.40%	8.24%	5.62%	9.26%
(12%,22%)	absolute	7.63	1.03	0.79	0.68	0.45
	relative	98.85%	14.78%	10.70%	9.51%	6.05%

Table 5.8: The averages of the absolute and absolute relative differences between the fair spread obtained with the alternative models and the market quotes. The CDO tranches have the iTraxx index as underlying and the weights Q_i^{-1} are used.

Looking at the tables, we can indeed observe that the calibration errors on the senior tranches have increased, and that the errors on the junior mezzanine tranches have decreased. Again, we cannot clearly point out a model that performs better than the other models.

The Tables 5.6 - 5.9 again provide a global view on the performance, since we consider averages over all valuation dates at the same time. In order to analyze the day-to-day performance we could look at figures where the market quotes are plotted together with the fair spreads calculated with the alternative models. We do not show such figures here, as this would require twenty different figures, which do not add much additional information. Instead, we confine ourselves with a few remarks about the results. First, it was observed that the fair premia from the alternative models, tend

CDO tranche	differences	Standard	Double-t	NIG	Mixture	RFL
(0%,3%)	absolute	6.96	4.28	3.47	3.91	4.85
	relative	18.00%	11.53%	9.59%	10.39%	13.03%
(3%,7%)	absolute	34.80	4.30	3.95	13.08	3.21
	relative	21.85%	2.79%	2.76%	8.42%	2.10%
(7%,10%)	absolute	22.40	3.12	2.31	3.81	2.32
	relative	76.73%	8.11%	6.65%	9.88%	6.75%
(10%,15%)	absolute	13.97	1.75	1.56	1.46	1.93
	relative	91.76%	11.33%	10.12%	7.63%	12.15%
(15%,30%)	absolute	7.68	1.58	1.36	1.74	1.08
	relative	98.89%	17.29%	18.17%	17.31%	12.03%

Table 5.9: The averages of the absolute and absolute relative differences between the fair spread obtained with the alternative models and the market quotes. The CDO tranches have the CDX index as underlying and the weights Q_i^{-1} are used.

to closely follow each other and to a lesser extend the market quotes. Second, in the second half of the data set, the fair premia for the alternative models are very close to the market quotes on the most senior tranches when using the relative weights. The fair premia were very close to the market quotes on the junior mezzanine tranche when using the weights $w_i = Q_i^{-1}$. Further similar conclusions to the ones from Tables 5.6 - 5.9 can be drawn.

5.6.3 Parameter Stability

In the previous two sections we have investigated the calibration errors of the alternative one-factor copula models. We did this first by directly looking at these errors, and secondly by looking at the differences between model and market prices. In this section we focus on a third important issue, namely the parameter stability. A model might be able to match market prices on a daily basis, but with different parameter values on each day. This could indicate that we are overfitting the market data, which is an undesirable feature of a model.

To test the parameter stability, we compute a number of statistics for the optimal parameters, such as averages, standard deviations and correlations between the different model parameters. As each model has different kinds of parameters, we present these statistics separately for each model. For the one-factor Gaussian copula we only have to deal with a single parameter. In Table 5.10 one can find the average correlations and its standard deviations.

From this table we can observe that the optimal correlation is in general small, and it is quite stable, as the standard deviation is not very large. However, the results for the one-factor Gaussian copula are not very interesting, as it cannot match the market quotes.

	$w_i = Q_i^{-2}$		$w_i = Q_i^{-1}$	
	iTraxx	CDX	iTraxx	CDX
avg. ρ	0.107	0.125	0.056	0.054
std.dev. ρ	0.048	0.057	0.042	0.053

Table 5.10: The average and standard deviations of the optimal values for the correlation parameter ρ in the one-factor Gaussian copula.

Next we consider the double-t factor copula, which matches the market quotes considerably better. There are three parameters that we need to consider in this case, which are the correlation ρ and the degrees of freedom for the common and idiosyncratic factor, ν_Y and ν_ε, respectively. In Table 5.11 one finds the average and standard deviation of the optimal choices of these parameters.

	$w_i = Q_i^{-2}$		$w_i = Q_i^{-1}$	
	iTraxx	CDX	iTraxx	CDX
avg. ρ	0.550	0.388	0.493	0.453
avg. ν_Y	4.89	2.53	4.43	19.99
avg. ν_ε	50.26	21.74	54.53	117.59
std.dev. ρ	0.312	0.265	0.337	0.332
std.dev. ν_Y	4.23	0.82	4.55	83.55
std.dev. ν_ε	120.72	62.86	120.43	205.71

Table 5.11: The average and standard deviations of the optimal parameter values in the double-t factor copula.

From this table we can observe that the standard deviation of ν_ε is very large, just as for ν_Y in case we look at the market quotes for CDO tranches on the CDX index, with the weights $w_i = Q_i^{-1}$. Furthermore, the corresponding averages of the degrees of freedom are also quite high. These high values are caused by very high optimal values for the idiosyncratic degrees of freedom on about 30% of the dates considered. In these cases the degrees of freedom are such that the corresponding factor is in fact very close to being normally distributed. Ignoring these dates yields much lower and more stable degrees of freedom, although the fraction of 30% does not justify ignoring this procedure. We therefore conclude that the optimal parameters are not very stable.

For the NIG factor copula, we have to consider the correlation ρ and the two parameters related to the NIG distribution, namely α and β, as introduced in Section 5.4. For the optimal parameter values we have again calculated the averages and standard deviations, which can be found in Table 5.12.

Considering the standard deviations, we observe these are lower than for the double-t factor copula, indicating more stable optimal parameters. Furthermore we see that the average correlation is about the same for the four different cases that are consid-

	$w_i = Q_i^{-2}$		$w_i = Q_i^{-1}$	
	iTraxx	CDX	iTraxx	CDX
avg. α	0.828	0.711	0.934	0.944
avg β	-0.552	-0.427	-0.699	-0.721
avg. ρ	0.335	0.366	0.338	0.373
std.dev. α	0.544	0.591	0.586	0.988
std.dev. β	0.729	0.705	0.773	1.104
std.dev. ρ	0.070	0.065	0.068	0.069

Table 5.12: The average and standard deviations of the optimal parameter values in the NIG factor copula.

ered. A point of concern for this model, or at least for the current implementation of the model and the calibration, is that there is a high negative correlation. For each of the cases this is less than -94%. On a large number of dates, this means that approximately $\alpha = -\beta$. As one divides by $\sqrt{\alpha^2 - \beta^2}$ in the specification of the model, this could lead to numerical difficulties.

Next, we consider the mixture factor copula, where we have to consider three correlation parameters, namely ρ_1, ρ_2, and ρ_3 and the corresponding probabilities p_1, p_2 and $p_3 = 1 - p_1 - p_2$. As the calibration algorithm does not necessarily keep the correlations ordered, we first order the optimal correlation parameters, and thereafter we calculate the averages and standard deviations, which are given in Table 5.13.

	$w_i = Q_i^{-2}$		$w_i = Q_i^{-1}$	
	iTraxx	CDX	iTraxx	CDX
avg. ρ_1	0.003	0.000	0.002	0.000
avg. ρ_2	0.222	0.253	0.232	0.306
avg. ρ_3	0.947	0.948	0.947	0.938
avg. p_1	0.637	0.628	0.646	0.641
avg. p_2	0.265	0.263	0.252	0.243
std.dev. ρ_1	0.009	0.002	0.008	0.002
std.dev. ρ_2	0.092	0.135	0.102	0.248
std.dev. ρ_3	0.117	0.101	0.123	0.106
std.dev. p_1	0.189	0.155	0.193	0.149
std.dev. p_2	0.153	0.160	0.153	0.153

Table 5.13: The average and standard deviations of the optimal parameter values in the mixture factor copula.

From this table we observe that the lowest correlation value is almost (always) equal to zero, as the average as well as the standard deviation are close to zero. Furthermore, the highest correlation value is close to one, again with a low standard deviation. The remaining correlation parameter, ρ_2, takes values around 0.25, where the standard deviation is largest when we calibrate the CDO tranches on the CDX basket, with the weights $w_i = Q_i^{-1}$. Similar to the correlations, also the probabilities show a quite stable behavior, as the average probabilities are roughly equal across the test cases, and the standard deviations are quite low. Comparing these results with those of the double-t and NIG factor copulas, we observe a higher parameter stability for the mixture factor copula.

We conclude this section with an investigation of the optimal parameter values of the random factor loadings factor copula, where we have to consider four parameters, namely α, β, γ and δ, where α and β determine the minimum and maximum correlation value, δ determines the speed at which the correlation decreases and γ determines the point where the correlations are exactly halfway between α and β. In Table 5.14 the averages and standard deviations of the optimal parameters are shown.

	$w_i = Q_i^{-2}$		$w_i = Q_i^{-1}$	
	iTraxx	CDX	iTraxx	CDX
avg. α	0.131	0.151	0.111	0.108
avg. β	0.794	0.825	0.817	0.846
avg. γ	-2.757	-2.690	-2.730	-2.582
avg. δ	17.609	3.583	14.214	4.575
std.dev. α	0.106	0.107	0.106	0.084
std.dev. β	0.202	0.141	0.186	0.147
std.dev. γ	0.344	0.255	0.335	0.300
std.dev. δ	98.224	5.703	78.435	7.914

Table 5.14: The average and standard deviations of the optimal parameter values in the random factor loadings factor copula.

From this table we observe that the fist three averages are roughly equal across the four test cases, the δ however can be quite large, and its standard deviation is very high, especially when we use the weights $w_i = Q_i^{-2}$. These high values are a direct result of a small number of very high values of δ. When we would ignore these results, which occur on less than 5% of the dates, we find much lower average deltas, between 2.1 and 2.7, with standard deviations between 1.1 and 2.1, which shows that, except for the 'outliers', the parameters for the random factor loadings factor copula are quite stable.

5.7 Concluding Remarks

In this chapter we have compared four different one-factor copula models. First of all we have looked at the ability of the models to fit to market data for CDO tranches on the iTraxx and CDX index. We have compared the calibration errors of these models, with those of the one-factor Gaussian copula, and with errors of the other models. To investigate if we overfit to the market data, we have carried out a stability analysis of the optimal parameters. Based on the test results that we have presented in this chapter, we summarize a number of general conclusions with respect to the models.

The first, and most straightforward, conclusion that can be drawn from the results is that each of the four models performs much better than the one-factor Gaussian copula. Comparing the calibration errors for the factor copula models, we can see that none of the models has the lowest error in each of the four test cases. On the other hand we have seen that the errors for the models, do not differ much.

With respect to the comparison between the fair premia and the market quotes we have seen that the two different types of weights used in the calibration function favor different tranches. Furthermore, it was observed that the fair premia of the models are quite close to each other. Based on the test results we cannot select a model that overall significantly outperforms the other models.

When looking at the optimal parameters of the models, a number of interesting conclusions has been drawn. As we have seen in Section 5.6.3, the optimal parameters for the double-t factor copula are not very stable, and for the NIG factor copula, the parameters α and β showed an undesirable correlation. The two random correlation models however, showed more stable parameters, although for a small number of cases in the random factor loadings factor copula the optimal values of δ can be very large.

5.A The Levenberg-Marquardt Algorithm

The Levenberg-Marquardt algorithm is designed for optimization problems where the parameters are all real valued and the objective function has the form

$$S(\mathbf{p}) := \sum_{i=1}^{N} w_i (f(x_i, \mathbf{p}) - y_i)^2, \tag{5.14}$$

where the parameters are given by $\mathbf{p} \in \mathbb{R}^M$, and the goal is to fit the function $f(x, \mathbf{p})$ to the data $\{(x_i, y_i) \,|\, i = 1, \dots, N\}$. In this section we briefly describe the Levenberg-Marquardt algorithm. For a more detailed description of the algorithm we refer to [PTVF02].

The Levenberg-Marquardt algorithm combines two techniques that can be used to find the minimum of a problem of the form (5.14). The first is the steepest descent method, where one determines the gradient in a given point and change the parameters by a constant times the gradient, such that the value of the objective function decreases. This method works well if we are far away from the minimum, but if we get closer to the minimum we might wander around the minimum, since the derivatives are close to zero.

The second approach uses a quadratic approximation of (5.14). It is easy to find the minimum of this quadratic approximation. This results in an iterative procedure to find the minimum of the nonlinear problem. This approach works well if the quadratic approximation is accurate, which is typically the case when we are reasonably close to the minimum. By combining both methods it is possible to construct an algorithm that should work well in both cases. In the remainder of this section we provide the Levenberg-Marquardt algorithm.

Start with an initial choice $\mathbf{p}^{(0)} \in \mathbb{R}^M$ and set $\lambda = \lambda_0$, where λ is used to smoothly switch between the two methods. The initial value λ_0 is chosen to be a modest number, e.g. 0.001.

At step m we want to find $\mathbf{p}^{(m+1)} \in \mathbb{R}^M$ based on $\mathbf{p}^{(m)}$. Therefore we need to compute the following quantities

- the gradient

$$\beta_k = 2\sum_{i=1}^{N} w_i(f(x_i, \mathbf{p}) - y_i)\frac{\partial f(x_i, \mathbf{p}^{(m)})}{\partial \mathbf{p}_k},$$

- an approximation to the Hessian

$$\alpha_{kl} = 2\sum_{i=1}^{N} w_i \frac{\partial f(x_i, \mathbf{p}^{(m)})}{\partial \mathbf{p}_k}\frac{\partial f(x_i, \mathbf{p}^{(m)})}{\partial \mathbf{p}_l}.$$

  Terms involving second derivatives are ignored, as these are typically small. For additional motivation of this approximation we refer to [PTVF02].

- the matrix $A_{kl} = \begin{cases} \alpha_{kl} & k \neq l \\ \alpha_{kl}(1+\lambda) & k = l \end{cases}.$

We obtain a candidate for $\mathbf{p}^{(m+1)}$ by solving (5.15) for $\delta\mathbf{p} = \mathbf{p}^{(m+1)} - \mathbf{p}^{(m)}$

$$A \cdot \delta\mathbf{p} = \beta. \tag{5.15}$$

When $\delta\mathbf{p}$ is found, we set $\tilde{\mathbf{p}} = \mathbf{p}^{(m)} + \delta\mathbf{p}$ and we calculate $S(\tilde{\mathbf{p}})$. Then we check if $\tilde{\mathbf{p}}$ results in an improvement.

- $S(\tilde{\mathbf{p}}) < S(\mathbf{p}^{(m)})$: Set $\mathbf{p}^{(m+1)} = \tilde{\mathbf{p}}$ and divide λ by λ_{scale}, where λ_{scale} is a substantial number, e.g. 10. By this division the algorithm behaves more like the quadratic approximation method.
- $S(\tilde{\mathbf{p}}) \geq S(\mathbf{p}^{(m)})$: Set $\mathbf{p}^{(m+1)} = \mathbf{p}^{(m)}$ as there was no improvement. Further we multiply λ by λ_{scale} such that the algorithm behaves more like the steepest descent method.

We stop the algorithm when the change in $S(\mathbf{p})$ is very small. The algorithm should not be stopped when there was no improvement, as λ might now have reached to proper level. In case there was an improvement we determine the absolute improvement $S(\mathbf{p}^{(m)}) - S(\mathbf{p}^{(m+1)})$ together with the relative improvement. $\frac{S(\mathbf{p}^{(m)}) - S(\mathbf{p}^{(m+1)})}{S(\mathbf{p}^{(m)})}$. We stop the algorithm if one or both of these errors are below some small level, e.g. 0.0001.

5.B Figures

In this section one can find the figures that are referred to in this paper.

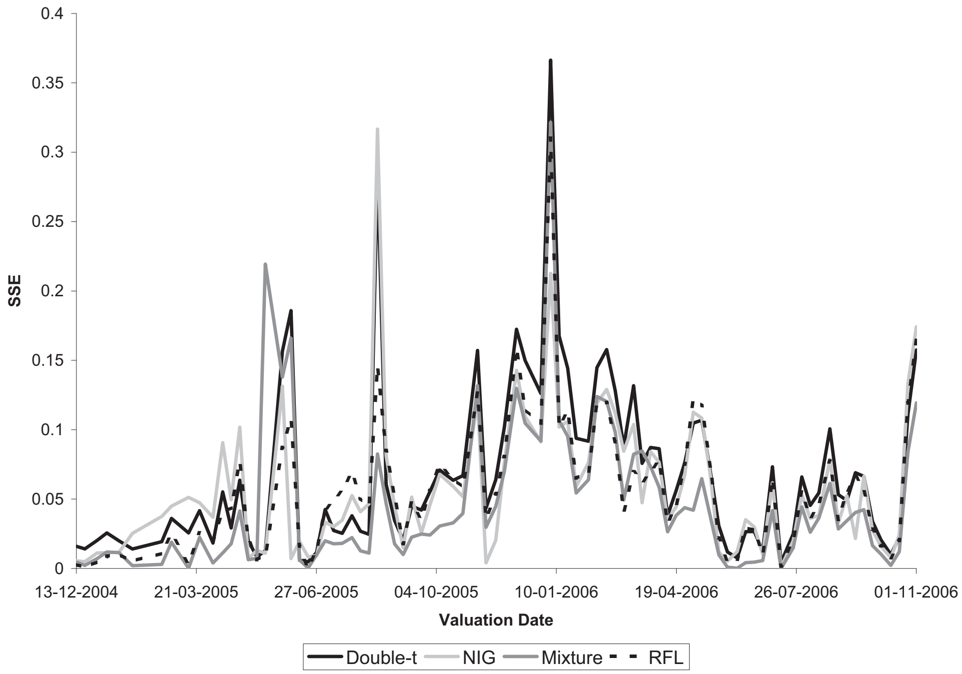


Figure 5.1: SSE's for the alternative models calibrated to CDO tranches on the iTraxx index with a five year maturity and using the weights Q_i^{-2}.

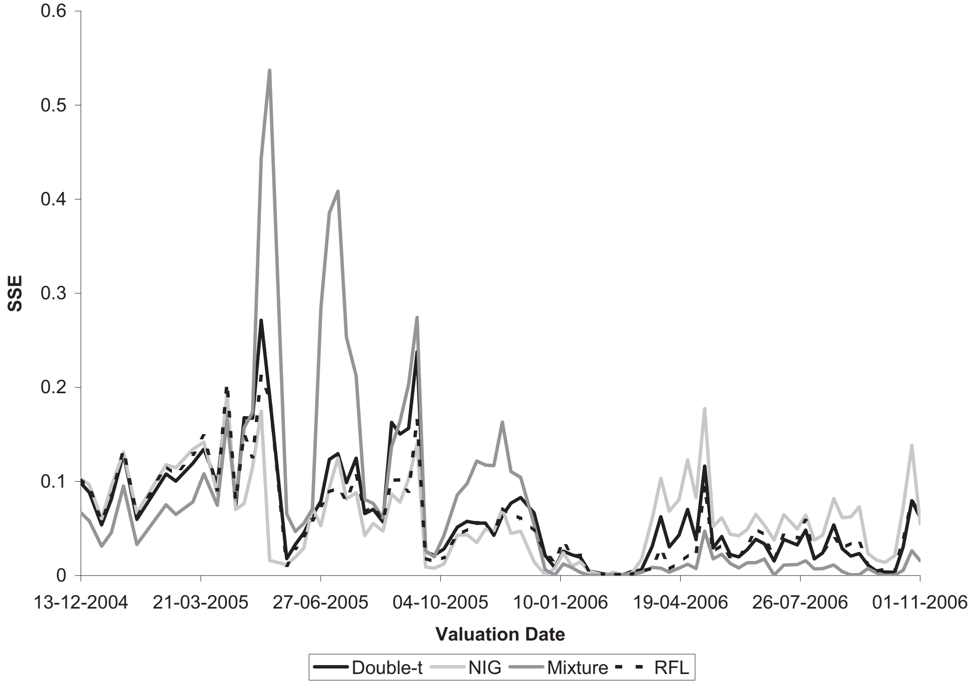


Figure 5.2: SSE's for the alternative models calibrated to CDO tranches on the CDX index with a five year maturity and using the weights Q_i^{-2}.

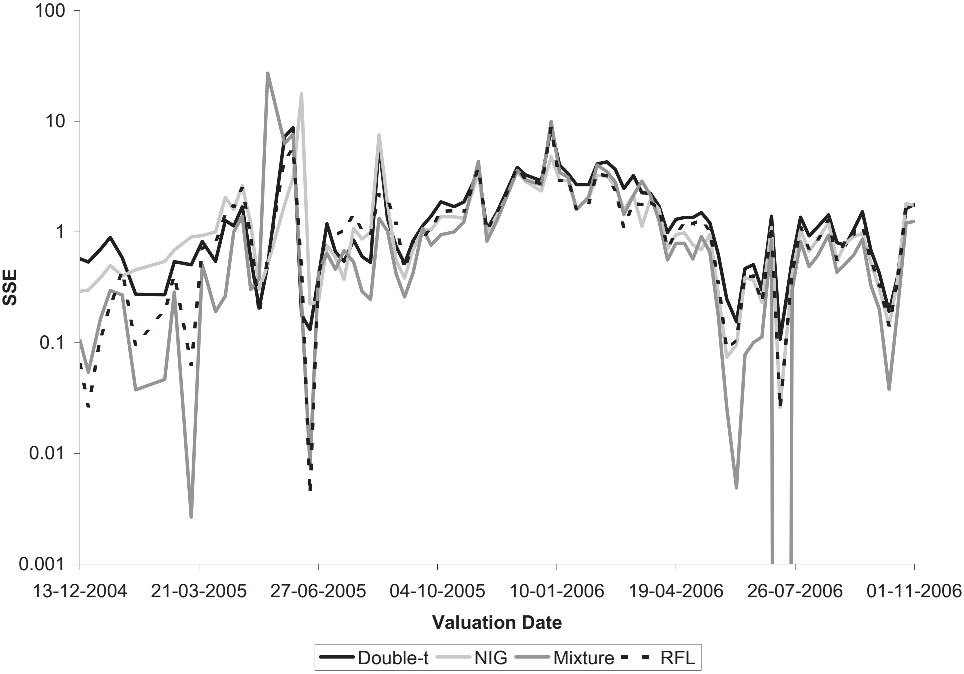


Figure 5.3: SSE's for the alternative models calibrated to CDO tranches on the iTraxx index with a five year maturity and using the weights Q_i^{-1}. The results are plotted on a logarithmic scale due to the wide range.

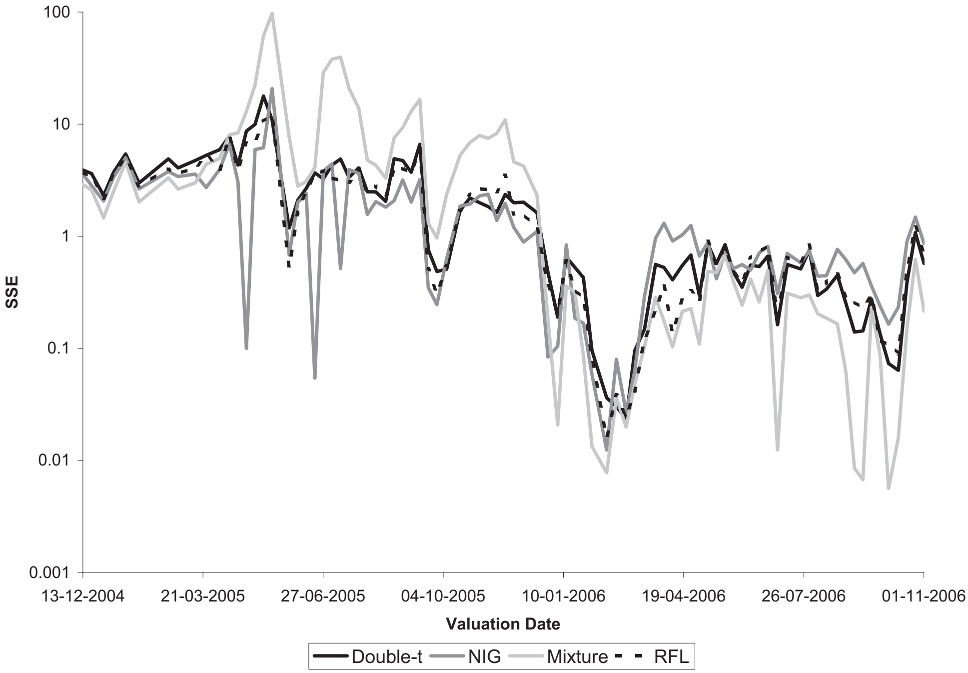


Figure 5.4: SSE's for the alternative models calibrated to CDO tranches on the CDX index with a five year maturity and using the weights Q_i^{-1}. The results are plotted on a logarithmic scale due to the wide range.

Summary

This thesis is on modeling credit risk and credit derivatives. Credit risk is the risk that a debtor, or obligor, does not honor its payment obligations. This can be the risk that a customer does not pay its bills, or that a loan is not, or only partially, paid. Using credit derivatives one can transfer credit risk to a third party. Two of the best known credit derivatives are the 'credit default swap' (CDS), which provides insurance against the loss due to the bankruptcy of a company, and the 'collateralized debt obligation' (CDO), which covers a predescribed part of the default losses in a portfolio.

The first chapter of this thesis gives a brief introduction to credit risk and credit derivatives. First, the notion of credit risk is formalized, and some well known modeling techniques are discussed. Second, this chapter discusses a number of credit derivatives, and models to value these derivatives.

Chapters 2 and 3 focus on the mathematical modeling of the evolution of the cumulative loss process in a large portfolio. In Chapter 2 this process is modeled as a point process, or, more specifically, as a *Cox process*. The intensity of such a process is stochastic, and we assume it is driven by Brownian motion. We further assume that the intensity evolves according to the *Cox-Ingersoll-Ross equation*, also known as the square root process. In addition, we assume that the intensity is *not* observed. Using the theory of filtering with point process observations, we derive equations for the conditional moment generating function between jumps as well as at jump times of the loss process. These are subsequently solved, and by combining the solutions, a recursive expression for the conditional moment generating function is obtained. The chapter concludes with a discussion about possible extensions of the model.

In Chapter 3 we investigate how the loss process behaves when the size of the portfolio increases. We assume that the contributions of each company to the cumulative loss process are independent, and we furthermore assume that these amounts are independent of the time at which the company defaults. Under these, and a number of other 'mild', assumptions we derive a 'large deviations principle' for the path of the cumulative loss process. This principle allows one to obtain upper and lower bounds on how certain probabilities with respect to the loss process behave when the size of the portfolio increases. For two special cases, we present the exact form of the asymptotic behavior.

Chapters 4 and 5 are of a more practical nature, and they mainly focus on the valuation of CDO tranches. In Chapter 4 the prices of CDO tranches are investigated using a large collection of market data. We investigate if correlation moves can be directly explained from tranche prizes, after these have been corrected for general credit risk. Further, we study the *base correlation model* for the valuation of CDOs. By using, for each CDO tranche, different correlation values in the one-factor Gaussian copula model, it is possible to exactly match market prices. The problem with this approach is that, to value a nonstandard CDO tranche, one has to interpolate between these correlations. We compare three different interpolation techniques, and it turns out that the routine that takes losses in the portfolio, as well as tranche losses into account, yields the best performance.

Finally, Chapter 5 discusses some models for the valuation of CDO tranches. Using the same data set as in Chapter 4, we study the performance of four alternative models to the one-factor Gaussian copula. We consider the *student t-distribution* as well as the *normal inverse Gaussian distribution* as alternative to the normal distribution. Furthermore, we consider two models with random correlation. Firstly we consider a mixture of one-factor copulas, and secondly we let the correlation depend on the state of the economy. For each date in the data set, we calibrate the models to CDO tranche prizes. It turns out that the alternative models provide a much better fit to the market data. Based on the calibration errors, however, it is not possible to point out a model that significantly outperforms the other models. We conclude the chapter with an investigation of the optimal model parameters, and we find that the parameters for the models with stochastic correlation are more stable.

Samenvatting

Dit proefschrift gaat over het modelleren van kredietrisico en kredietderivaten. Kredietrisico is het risico dat een schuldenaar, of debiteur, niet aan zijn betalingsplicht voldoet. Dit kan het risico zijn dat een klant zijn rekening niet betaalt, maar ook het risico dat een lening niet, of slechts gedeeltelijk, wordt terugbetaald. Met behulp van kredietderivaten is het mogelijk om kredietrisico geheel of gedeeltelijk over te dragen aan een derde partij. Twee van de meeste bekende kredietderivaten zijn de 'credit default swap' (CDS) en de 'collateralized debt obligation' (CDO). Een CDS is een 'verzekering' tegen het faillissement van een bedrijf. Met behulp van een CDO kan men zich 'verzekeren' tegen een gedeelte van het totale verlies binnen een portfolio.

Het eerste hoofdstuk van dit proefschrift geeft een beknopte inleiding in kredietrisico en kredietderivaten. Ten eerste wordt kredietrisico geformaliseerd, en wordt een aantal bekende modelleertechnieken behandeld. Verder behandelt dit hoofdstuk een aantal kredietderivaten, waaronder de CDS en CDO. Daarnaast wordt aandacht besteed aan bekende modellen om deze derivaten te waarderen.

Hoofdstukken 2 en 3 richten zich op het wiskundig modeleren van de ontwikkeling van het cumulatieve verliesproces in een grote portefeuille. In Hoofdstuk 2 wordt dit proces gemodelleerd als een puntproces, en in het bijzonder als een *Cox proces*. De intensiteit van dit proces is stochastisch, en de intensiteit wordt aangedreven door een Brownse beweging. We nemen aan dat de intensiteit zich ontwikkeld volgend de *Cox-Ingersoll-Ross vergelijking*. Het Cox proces staat ook bekend als het wortelproces. Verder nemen we aan dat de intensiteit *niet* wordt waargenomen. Door gebruik te maken van de theorie voor het filteren met observaties van puntprocessen, worden er vergelijkingen afgeleid voor de conditionele moment genererende functie, zowel tussen sprongen als op sprong tijdstippen van het verliesproces. Vervolgens hebben we deze vergelijkingen exact opgelost. Door de oplossingen tussen en op sprongen te combineren hebben we een recursieve oplossing gevonden voor de conditionele moment genererende functie. Het hoofdstuk besluit met een discussie over uitbreidingen van het model.

In Hoofdstuk 3 onderzoeken we hoe het verliesproces zich gedraagt als de portefeuille steeds groter wordt. We nemen aan dat de bijdragen aan het totale verlies van elk bedrijf in de portefeuille onafhankelijk zijn. Verder nemen we aan dat deze bijdragen onafhankelijk zijn van het tijdstip waarop een bedrijf failliet gaat. Onder deze, en een aantal andere milde, aannames hebben we een zogeheten *principe van de grote afwijkingen* voor het pad van het cumulatieve verliesproces af kunnen leiden. Hiermee is het mogelijk boven- en ondergrenzen te bepalen voor de manier waarin bepaalde kansen met betrekking tot het verliesproces zich gedragen naarmate de portefeuille groter wordt. Daarnaast hebben we voor twee specifieke gevallen exacte uitdrukkingen voor deze asymptotiek kunnen afleiden.

Hoofdstukken 4 en 5 zijn praktischer van aard, en richten zich op het waarderen van CDO tranches. In Hoofdstuk 4 worden aan de hand van een grote hoeveelheid marktdata, die een periode van twee jaar beslaat, de prijzen van CDO tranches onderzocht. We onderzoeken of correlatiebewegingen direct zijn af te leiden uit trancheprijzen, door deze prijzen eerst te corrigeren voor algemeen kredietrisico. Daarnaast onderzoeken we, met behulp van dezelfde verzameling data, het *base correlation model* voor het waarderen van CDO tranches. Door, voor elke CDO tranche, verschillende correlaties te gebruiken in de één-factor Gaussische copula is het mogelijk marktprijzen precies te reproduceren. Het probleem is echter dat, om een CDO tranche met afwijkende specificaties te waarderen, er geïnterpoleerd moet worden tussen deze correlaties. We vergelijken drie verschillende interpolatiemethodes. Het blijkt dat de meest gecompliceerde methode van de drie, waarbij rekening gehouden wordt met het verwachte verlies in de gehele portefeuille en het verlies met betrekking tot de CDO tranche, het beste resultaat oplevert.

Tot slot behandelt Hoofdstuk 5 alternatieve modellen voor het waarderen van CDO tranches. Met behulp van dezelfde verzameling data als in Hoofdstuk 4 bestuderen we de prestaties van vier alternatieve modellen voor de één-factor Gaussische copula. We bekijken de *student t-verdeling* en de *normale inverse Gaussische verdeling* als alternatief voor de normale verdeling. Daarnaast bekijken we twee modellen met een stochastische correlatieparameter. Ten eerste bekijken een mix van één-factor Gaussische copula's, en ten tweede laten we de correlation afhangen van de toestand van de economie. Voor elke dag in de verzameling data kalibreren we deze vier modellen naar CDO trancheprijzen. Het blijkt dat het met de alternatieve modellen veel beter mogelijk is om in de buurt van de marktprijzen te komen dan met de één-factor Gaussische copula. Op basis van de afwijkingen ten opzichte van de marktprijzen is het niet mogelijk om een model aan te wijzen dat significant het beste presteert. We besluiten het hoofdstuk met het bestuderen van de optimale modelparameters, waarbij het duidelijk wordt dat de parameters voor de modellen met stochastische correlatie stabieler zijn.

Bibliography

[ABM04] R. Ahluwalia, E. Beinstein, and L. McGinty. A relative value framework for credit correlation. JPMorgan, 2004.

[Aco94] A. de Acosta. Large deviations for vector-valued Lévy processes. *Stochastic Processes and their Applications*, 51:75–115, 1994.

[ADEH99] P. Artzner, F. Delbaen, J.-M. Eber, and D. Heath. Coherent measures of risk. *Mathematical Finance*, 9:203–228, 1999.

[AM04] R. Ahluwalia and L. McGinty. Credit correlation: A guide. JPMorgan, 2004.

[AS72] M. Abramowitz and I.A. Stegun. *Handbook of Mathematical Functions with Formulas, Graphs, and Mathematical Tables*. U.S. Department of Commerce, 1972.

[AS04] L. Andersen and J. Sidenius. Extensions to the Gaussian copula: Random recovery and random factor loadings. *The Journal of Credit Risk*, 1(1):29–70, 2004.

[ASB03] L. Andersen, J. Sidenius, and S. Basu. All your hedges in one basket. *RISK*, November:67–72, 2003.

[Aze80] R. Azencott. *Ecole D'Eté de Probabilités de Saint-Flour VIII-1978*, volume 774 of *Lecture Notes in Mathematics*, chapter Grandes Déviations et Applications, pages 1–176. Springer, Berlin, 1980.

[BB80] R.K. Boel and V.E. Beneš. Recursive nonlinear estimation of a diffusion acting as the rate of an observed Poisson process. *IEEE Transactions on Information Theory*, 26(5):561–575, 1980.

[BC76] F. Black and J.C. Cox. Valuing corporate securities: Some effects of bond indenture provisions. *Journal of Finance*, 31:351–367, 1976.

[BGL05] X. Burtschell, J. Gregory, and J.P. Laurent. A comparative analysis of CDO pricing models. Working paper, BNP Paribas, University of Lyon, 2005.

[BGL07] X. Burtschell, J. Gregory, and J.-P. Laurent. Beyond the Gaussian copula: Stochastic and local correlation. *Journal of Credit Risk*, 3(1):31–62, Spring 2007.

[BJ77] P. Brémaud and J. Jacod. Processus ponctuels et martingales: Résultats récents sur la modélisation et le filtrage. *Advances in Appl. Probability*, 9(2):362–416, 1977.

[BM06] D. Brigo and F. Mercurio. *Interest Rate Models - Theory and Practice. With Smile Inflation and Credit.* Springer Finance, 2^{nd} edition, 2006.

[BPT06] D. Brigo, A. Pallavicini, and R. Torresetti. Default correlation, cluster dynamics and single names: The GPCL dynamical loss model. 2006.

[BPT07] D. Brigo, A. Pallavicini, and R. Torresetti. Calibration of CDO tranches with the dynamical generalized-poisson loss model. 2007.

[BR60] R.R. Bahadur and R. Ranga Rao. On deviations of the sample mean. *Annals of Mathematical Statistics*, 31:1015–1027, 1960.

[BR02] T.R. Bielecki and M. Rutkowski. *Credit Risk: Modeling, Valuation and Hedging*. Springer-Verlag, 2002.

[Bré81] P. Brémaud. *Point Processes and Queues*. Springer-Verlag New York Inc, 1981.

[Cha95] C.-S. Chang. Sample path large deviations and intree networks. *Queueing Systems*, 20:7–36, 1995.

[Che71] C.R. Chester. *Techniques in Partial Differential Equations*. McGraw-Hill Book Company, 1971.

[CIR85] J.C. Cox, J.E. Ingersoll, and S.A. Ross. A theory of the term structure of interest rates. *Econometrica*, 53(2):385–407, 1985.

[DDD04] A. Dembo, J.-D. Deuschel, and D. Duffie. Large portfolio losses. *Finance and Stochastics*, 8:3–16, 2004.

[DK96] D. Duffie and R. Kan. A yield-factor model of interest rates. *Mathematical Finance*, 6:379–406, 1996.

[DS03] D. Duffie and K. Singleton. *Credit Risk: Pricing, Measurement, and Management*. Princeton University Press, 2003.

[Duf05] D. Duffie. Credit risk modeling with affine processes. *Journal of Banking and Finance*, 29:2751–2802, 2005.

[DZ98] A. Dembo and O. Zeitouni. *Large Deviations Techniques and Applications*. Springer Verlag, New York, 2nd edition, 1998.

[EGG09] E. Errais, K. Giesecke, and L.R. Goldberg. Affine point process and portfolio credit risk. September 2009.

[EKM97] P. Embrechts, C. Klüppelberg, and T. Mikosch. *Modelling extremal events*, volume 33 of *Applications of Mathematics*. Springer-Verlag, 1997.

[Eli05] A. Elizalde. Credit risk models II: Structural models. `http://abelelizalde.com/pdf/survey2%20-%20structural.pdf`, 2005.

[FPR07] R. Frey, C. Prosdocimi, and W.J. Runggaldier. Affine credit risk under incomplete information. In *Stochastic Processes and Applications to Mathematical Finance - Proceedings of the 6th Ritsumeikan International Symposium*, pages 97–113. World Scientific Publishing, 2007.

[FR04] A. Friend and E. Rogge. Correlation at first sight. *Forthcoming: Economic Notes: Review of Banking, Finance and Monetary Economics*, 2004.

[GGD05] K. Giesecke, L.R. Goldberg, and X. Ding. A top-down approach to multi-name credit. Februari 2005.

[Gie04] K. Giesecke. Credit risk modeling and valuation: An introduction. Working paper, Cornell University, 2004.

[GKS07] P. Glasserman, W. Kang, and P. Shahabuddin. Large deviations in multifactor portfolio credit risk. *Mathematical Finance*, 17:345–379, July 2007.

[Gla04] P. Glasserman. *Monte Carlo Methods in Financial Engineering*, volume 53 of *Applications of Mathematics*. Springer-Verlag, 2004.

[Gor02] M.B. Gordy. Saddlepoint approximation of CreditRisk+. *Journal of Banking and Finance*, 26:1337–1355, 2002.

[HJM92] D. Heath, R.A. 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.

[HL86] T.S.Y. Ho and S.B. Lee. Term structure movements and pricing interest rate contingent claims. *Journal of Finance*, 41, 1986.

[Hul05] J. Hull. *Options, Futures and Other Derivatives*. Prentice Hall, 6^{th} edition, 2005.

[HV90] P. Houweling and T. Vorst. Pricing default swaps: Emperical evidence. *Journal of International Money and Finance*, 24:1200–1225, 1990.

[HW04] J. Hull and A. White. Valuation of a CDO and an n-th to default CDS without monte carlo simulation. *Journal of Derivatives*, 2:8–23, 2004.

[Jor06] P. Jorion. *Value at Risk: The New Benchmark for Managing Financial Risk*. McGraw-Hill, 3rd edition, 2006.

[Kal86] P. Kall. Approximation to optimization problems: An elementary review. *Mathematics of Operations Research*, 11:9–18, February 1986.

[KRS93] I.J. Kim, K. Ramaswamy, and S. Sundaresan. The valuation of corporate fixed income securities. Technical report, Wharton School, University of Pennsylvania, 1993.

[KSW05] A. Kalemanova, B. Schmid, and R. Werner. The normal inverse Gaussian distribution for synthetic CDO. Working paper, risklab Germany, Algorithmics Incorporated, Allianz, 2005.

[Lan98] D. Lando. On Cox processes and credit risky securities. *Review of Derivatives research*, 2:99–120, 1998.

[LdVV08] V.J.G. Leijdekker, M.F.A. Van der Voort, and A.C.F. Vorst. *An Empirical Analysis of CDO Data*, chapter 23, pages 457–484. Chapman & Hall, 2008.

[LG03] J.-P. Laurent and J. Gregory. Basket default swaps, CDO's and factor copulas. Technical report, BNP Paribas and ISFA Actuarial School, University of Lyon, 2003.

[Li00] D. Li. On default correlation: A copula approach. *Journal of Fixed Income*, 9:43–54, 2000.

[LKSS01] A. Lucas, P. Klaassen, P. Spreij, and S. Straetmans. An analytical approach to credit risk in large corporate bond and loan portfolios. *Journal of Banking and Finance*, 25:1635–1664, 2001.

[LL96] D. Lamberton and B. Lapeyre. *Introduction to Stochastic Calculus Applied to Finance*. Chapman & Hall, 1996.

[LM99] N. Likhanov and R. Mazumdar. Cell loss asymptotics for buffers fed with a large number of independent stationary sources. *Journal of Applied Probability*, 36:86–96, 1999.

[LMS09] V.J.G. Leijdekker, M.R.H. Mandjes, and P.J.C. Spreij. Sample-path large deviations in credit risk. 2009.

[LS78] R.S. Liptser and A.N. Shiryayev. *Statistics of Random Processes II*. Springer-Verlag, 1978.

[LS95] F.A. Longstaff and E.S. Schwartz. A simple approach to valuing risky fixed and floating rate debt. *Journal of Finance*, 50:789–819, 1995.

[LS09] V.J.G. Leijdekker and P.J.C. Spreij. Explicit computations for a filtering problem with point process observations with applications to credit risk. 2009.

[Luo05] L.S.J. Luo. Bootstrapping default probability curves. *Journal of Credit Risk*, 1, 2005.

[MBAW04] L. McGinty, E. Beinstein, R. Ahluwalia, and M. Watts. Introducing base correlations. *Credit Derivatives Strategy, JP Morgan*, March 2004.

[MCC98] D.B. Madan, P.P. Carr, and E.C. Chang. The variance gamma process and option pricing. *European Finance Review*, 2:79–105, 1998.

[Mer74] R.C. Merton. On the pricing of corporate debt: The risk structure of interest rates. *Journal of Finance*, 29:449–470, 1974.

[MFE05] A. McNeil, R. Frey, and P. Embrechts. *Quantitative Risk Management: Concepts, Techniques and Tools*. Princeton University Press, 2005.

[MMNvU06] M. Mandjes, P. Mannersalo, I. Norros, and M. van Uitert. Large deviations of infinite intersections of events in Gaussian processes. *Stochastic Processes and their Applications*, 116:1269–1293, 2006.

[Mog76] A. Mogul'skiĭ. Large deviations for trajectories of multi dimensional random walks. *Theory of Probability and its Applications*, 21:300–315, 1976.

[Moo06] T. Moosbrucker. Pricing CDOs with correlation variance gamma distributions. working paper, University of Cologne, February 2006.

[MS90] D.B. Madan and E. Seneta. The variance gamma model for market share returns. *Journal of Business*, 63:511–524, 1990.

[MTB01] R. Martin, K. Thompson, and C. Browne. Taking to the saddle. *Risk*, 14:91–94, 2001.

[Nel99] R.B. Nelsen. *An Introduction to Copulas*. Number 139 in Lecture Notes in Statistics. Springer, Heidelberg, Berlin, New York, 1999.

[PS06] D. Prange and W. Scherer. Correlation smile matching with alpha-stable distributions and fitted Archimedan copula models. working paper, DrKW, March 2006.

[PTVF02] W.H. Press, S.A. Teukolsky, W.T. Vetterling, and B.P. Flannery. *Numerical Recipes in C++: The Art of Scientific Computing*. Cambridge University Press, 2^{nd} edition, 2002.

[Rey04] A. Reyfman. *Valuing and Hedging Synthetic CDO Tranches Using Base Correlations*. Bear Stearns & Co. Inc., May 2004.

[RMR07] A. Rajan, G. McDermott, and R. Roy. *The Structured Credit Handbook*. Wiley & Sons Inc, 2007.

[RUK04] A. Reyfman, K. Ushakova, and W. Kong. *Credit Derivatives Research: How to Value Bespoke Tranches Consistently with Standard Ones*. Bear Stearns & Co. Inc., September 2004.

[Sch77] J.H. Van Schuppen. Filtering, prediction and smoothing for counting process observations, a martingale approach. *SIAM J. Appl. Math.*, 32(3):552–570, 1977.

[Sch96] P.J. Schönbucher. Term structure of defaultable bond prices. Technical report, University of Bonn, 1996.

[Sch02] P.J. Schönbucher. A tree implementation of a credit spread model for credit derivatives. *Journal of Computational Finance*, 6(2):1–38, 2002.

[Sch03a] P.J. Schönbucher. *Credit Derivatives Pricing Models*. Wiley, 2003.

[Sch03b] P.J. Schönbucher. *Information-Driven Default Contagion*. working paper, ETH Zürich, 2003.

[Sch06] P.J. Schönbucher. Portfolio losses and the term structure of loss transition rates: A new methodology for the pricing of portfolio credit derivatives. 2006.

[SDK75] A. Segall, M.H.A. Davis, and T. Kailath. Nonlinear filtering with counting observations. *IEEE Trans. Information Theory*, IT-21:143–149, 1975.

[SPA06] J. Sidenius, V. Piterbarg, and L. Andersen. A new framework for dynamic credit portfolio loss modelling. 2006.

[Tav03] J.M. Tavakoli. *Collateralized Debt Obligations & Structured Finance, New Developments in Cash & Synthetic Securitization*. John Wiley & Sons, Inc., Hoboken, New Jersey, 2003.

[Vas77] O. Vasicek. An equilibrium characterisation of the term structure. *Journal of Financial Economics*, 5:177–188, 1977.

[vdV07] M. van der Voort. Factor copulas: External defaults. *The Journal of Derivatives*, 15(1):91–99, 2007.

[Zho96] C. Zhou. A jump-diffusion approach to modeling credit risk and valuing defaultable securities. Technical report, Federal Reserve Board, Washington, 1996.