Pricing rating-dependent credit derivatives

Master's Thesis

Roger Lord

Rabobank International
Central Market Risk

Tilburg University
Faculty of Economics and Business Administration
Department of Econometrics and Operations Research

Rabobank International

Tilburg University

Preface

In front of you lies a thesis, which is the result of an internship of six months at Rabobank International in Utrecht and London. The research was carried out within the Central Market Risk department, which monitors the trading positions of the global Rabobank offices. The first three months of this internship served as a practical internship for my Master's degree in Applied Mathematics at the Eindhoven University of Technology in the Netherlands. The thesis in its entirety is my Master's thesis for my degree in Econometrics at Tilburg University, also in the Netherlands.

During my second year of Applied Mathematics in Eindhoven my interests gradually started focussing on financial applications, which is why I decided to take up Econometrics. For this final thesis I decided to pursue an international internship. Rabobank International offered me the opportunity to carry out interesting research on the topic of credit derivatives, which would be interesting both from a theoretical and a practical point of view, in their London branch.

My gratitude goes out to Diemer Salomé, who pointed me toward the right people within Rabobank International. I thank Sacha van Weeren for making this great opportunity happen. Moreover, I am very grateful to my supervisor Freddy van Dijk, who it was a pleasure to work with, and who always had constructive comments and answers to my questions. I would also like to thank my supervisor Kees van den Berg who, despite his busy schedule, always found time to discuss my work and gave insightful comments and new ideas. I am indebted to Patrick Houweling, a PhD student on the topic of credit risk. He always found time to answer my questions; several papers of his were helpful and spawned some of the research carried out in this thesis.

I would like to thank my supervisors at both universities: Feico Drost from Tilburg University for letting me carry out my research independently and for reading my manuscripts carefully, and Nino Mushkudiani from the Eindhoven University of Technology for her useful comments and advice. Although the topic of pricing credit derivatives is definitely not a common topic for a thesis in Eindhoven, she was interested in it and took the time to understand it.

Of course I have to thank my colleagues in Utrecht and London for the pleasant working environment they created, Richard Dagg in particular, with whom it was a pleasure to work in London. I hope the people in the kitchen have lost some of their obsession with leaves!

To my parents I am very grateful for supporting me throughout my study. Last, but not least, I would like to thank Mark, Olivier and Walter for a great six years. Onward and upward we go.

Roger Lord, MSc
Tilburg, July 2001

Contents

1. Introduction

Credit derivatives are the fastest growing sector of the global derivatives market. According to an article in the Financial Times by Smith [2000], the estimated value of notional amounts of credit derivative contracts outstanding at the end of 2000 was expected to be $800 billion. This thesis deals with the pricing of rating-dependent credit derivatives, in particular the recently issued telecom bonds with rating-dependent coupons. In Credit, Conroy [2000] states that over € 20 billion of Eurobonds issued in 2000 have rating-dependent coupons. In order to be able to price rating-dependent credit derivatives, we must first understand the kind of models that can be used to price such derivatives: rating-based models. The goal of this thesis can therefore be formulated as follows:

1. Investigate the properties of rating-based models;
2. Price rating-dependent credit derivatives with the aforementioned models.

After this chapter we first investigate the credit-related literature and give a simple introduction to such notions as credit derivatives, credit ratings, credit rating transitions and recovery rates. Furthermore, we give a brief overview of models for the pricing of credit derivatives. In order to accurately price the embedded credit derivative component of these bond issues, we need to use a pricing model which explicitly incorporates the credit rating of the underlying issue. Such models are called rating-based pricing models. The seminal rating-based model is that of Jarrow, Lando and Turnbull [1997]. In chapter 3 we review and discuss several properties of a broader class of models, which comprises the Jarrow, Lando and Turnbull assumption as a special case. In particular we thoroughly investigate the risk premiums needed in order to be able to implement this type of models.
In chapter 4 we describe and implement an alternative model, which also finds it basis in the Jarrow, Lando and Turnbull model. The model is calibrated on a data set of straight Euro-denominated bonds and the performance of this model is compared to several well-established models of the term structure. In the subsequent chapter 5 the focus lies on the pricing of the bond issues with rating-dependent coupons. We show how these bonds can be priced in the previously mentioned models, including empirical results. Chapter 6 reviews an article by Bielecki and Rutkowski [2000a], in which the defaultable term structure is modelled using an HJM-like approach with rating transitions. Chapter 7 concludes with a summary and some possible ideas for further research.

2. Introduction to credit-related literature

The main purpose of this chapter is to introduce some terminology and concepts widely used in the literature on credit derivatives. The list of references is far from complete. A more complete overview of the literature can be obtained by consulting the references in the papers that we refer to. We start out by reviewing some of the most common instruments traded in the credit market. The second paragraph deals with the notion of credit ratings, which give an indication of the creditworthiness of an obligor. When the issuer of a bond defaults, a bondholder usually receives some recovery amount. The third paragraph mentions some empirical facts about recoveries and discusses assumptions that are made about the recovery process in pricing models. The chapter is concluded with a brief overview of the various approaches to pricing credit derivatives, which clarifies why our focus lies on a particular class of models: the rating-based approach to pricing credit derivatives.

2.1. Overview of instruments

We first review the most actively traded types of credit-related contracts. For an extensive overview and for reasons why these products should be used, we refer you to Das [2000] or Tavakoli [1998]. The reference asset to which we refer in the description of the products can be almost any asset, index or basket of assets.

Corporate bonds and medium-term notes

A corporate bond is a bond issued by a corporation. They offer a higher yield than government bonds because they carry a higher default risk. The total size of the corporate bond market at the end of 1999 was € 7.9 trillion according to Merrill Lynch [2000], about 26% of the size of the world bond market. Instruments similar to corporate bonds are medium-term notes (MTNs); one of the main differences between the two is the way in which they are sold. MTNs are distributed on a best-efforts basis, and are typically offered in small amounts on a continuous basis, while corporate bonds are sold in relatively large offerings. In future, when we talk about corporate bonds we will mean either corporate bonds or MTNs.

In this thesis we will only focus on straight (bullet, fixed rate or plain vanilla) bonds. Such bonds have no special features, such as a callable or putable feature. A callable bond gives the issuer the option to redeem a bond at agreed specific prices and dates before the maturity date. Similarly, a putable bond gives the holder the option to sell the bond to the issuer at agreed prices and dates before the maturity date. Since bonds with such features will be priced differently to straight bonds, we will exclude them from our analysis.

Default swap

With a default swap, one party receives a fixed stream of payments or a lump-sum payment in return for which he pays nothing unless the reference asset defaults. The party receiving these payments provides insurance protecting the holder of the underlying from the consequences of default. In case of default and physical settlement, the reference asset is delivered to the insurer against repayment at par. In case of default and cash settlement, the notional value minus the value of the underlying is paid to the insurer. Default swaps are often issued in combination with the underlying bond, and are then referred to as credit linked notes.

Binary default swap

In case of a binary or digital default swap, the seller receives a fixed stream of payments or a lump-sum payment. In return he is required to pay 1 if the reference asset defaults. To clarify the difference between a default swap and a binary default swap we have set out their payoff schedules in the subsequent graph. If a payment is fixed, it is depicted by a straight arrow. If it is floating, i.e. dependent on some random variable, the arrow is wobbly. The letter τ indicates the time of default.

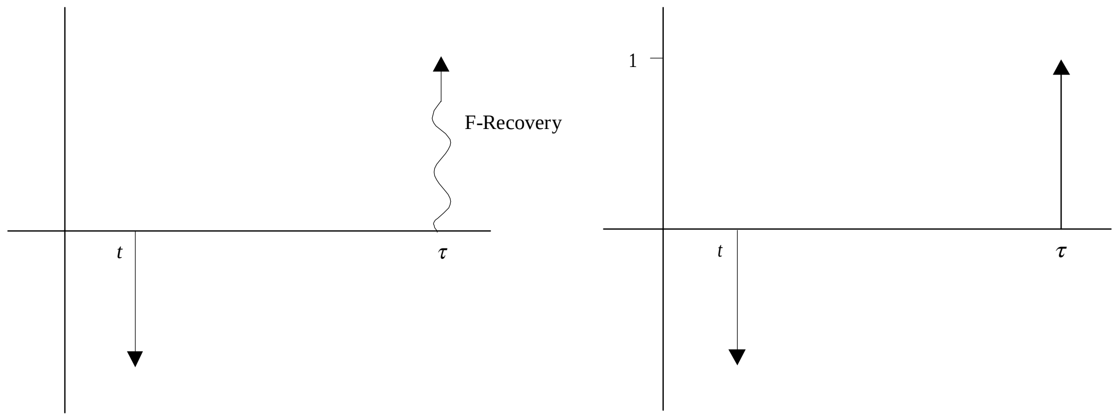


***Figure 2.1**: Payoff schedule of a default swap (left) and a binary default swap (right)*

Total return swap

In a total return swap party A makes a set of payments (either fixed or linked to an interest rate such as LIBOR) in return for a variable stream of payments from party B equal to the total return (coupons plus capital appreciation or depreciation) of the reference asset. At default party A agrees to repay the par value of the bond, in return for the recovery value of the bond. This is clarified in the following payoff schedule:

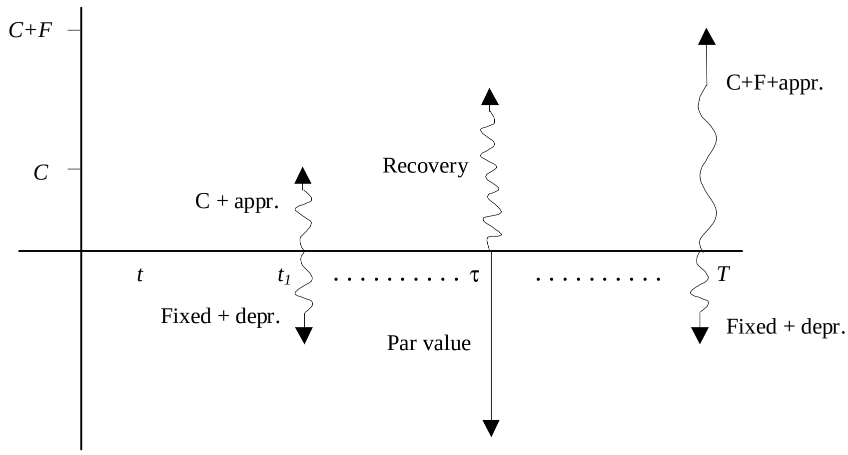


***Figure 2.2**: Payoff schedule of a total return swap (for party A)*

Credit spread options

Credit spread options are options that in some way or other depend on the credit spread of the reference asset. The credit spread is the difference between the yield of the reference asset and the default-free bond. For example a credit spread call option would pay off a certain value at maturity if the credit spread at that date exceeds a certain threshold value.

2.2. Credit ratings and transitions

2.2.1. Credit ratings

A credit rating is an assessment made by a rating agency of the creditworthiness of an obligor with respect to a certain obligation. It takes into consideration the creditworthiness of guarantors, insurers, or other forms of credit enhancement on the obligation and takes into account the currency in which the obligation is denominated. Ratings are supplied for individual issuers and bond issues. The two largest U.S. rating agencies, Moody's and Standard and Poor's (S&P) regularly publish ratings for a great deal of corporates and sovereigns. The following symbols are used to indicate the creditworthiness of an issuer (in decreasing order of creditworthiness):

Rating agency	*Ratings*									
Moody's	*Aaa*	*Aa*	*A*	*Baa*	*Ba*	*B*	*Caa*	*Ca*	*C*	*D*
S&P's	*AAA*	*AA*	*A*	*BBB*	*BB*	*B*	*CCC*	*CC*	*C*	*D*

Table 2.1: *Credit ratings*

Moody's and Standard & Poor's have both divided each rating from Aa/AA to Caa/CCC into three sub-classes. S&P's use a plus sign to signify the most creditworthy subrating and a minus sign for the least creditworthy subrating. Moody's uses different modifiers, namely '1' for the best subrating, '2' for the middle subrating and '3' for the worst subrating. Issuers or bond issues that are assigned a rating of Baa3/BBB- or higher are referred to as investment-grade. Issuers or issues below this rating are referred to as speculative-grade. The D rating category is used when payments on an obligation are not made on the date due even if the applicable grace period has not expired, unless the rating agency believes that such payments will be made during such grace period. The D rating will also be used upon the filing of a bankruptcy petition or the taking of a similar action if payments on an obligation are jeopardised.

The rating used in our analysis will be the Bloomberg composite rating. This rating is a combination of Moody's and S&P's rating. If Moody's and S&P's have a split rating, the composite rating is equivalent to the lower rating. If they are split two levels, the composite is equivalent to the middle rating. If only one of the rating agencies rates the issue, the composite rating is equivalent to this rating. The symbols used by Bloomberg are equal to the symbols used by S&P's, except for the modifiers, which are equal to the ones used by Moody's.

Both rating agencies also publish yearly reports with research on default rates, rating transition probabilities and recovery rates, see e.g. Moody's [2000] and Standard & Poor's [2001a]. We will mainly use the published rating transition probabilities and information on recovery rates in this thesis. With regard to the transition matrices, Moody's only publishes one-year transition matrices, whereas Standard & Poor's publishes multi-year transition matrices, from 1 up to 15 years.

2.2.2. Empirical facts on credit rating transitions

We briefly mentioned that the rating agencies publish transition probability matrices between the various credit ratings. Since we will often be discussing various properties of these matrices in our thesis, it is worthwhile to show the empirical transition probability matrix used in this thesis, namely the one-year transition matrix published in Standard & Poor's [2001a]. Rating transition probabilities for ratings lower than CCC are not published. Furthermore, we only depict unmodified rating transitions; we grouped the probabilities for the modified ratings using equal weights for each modified rating.

	AAA	AA	A	BBB	BB	B	CCC	D
AAA	91.9%	7.5%	0.5%	0.1%	0.0%	0.0%	0.0%	0.0%
AA	1.0%	92.5%	5.7%	0.6%	0.0%	0.1%	0.0%	0.0%
A	0.1%	2.3%	91.1%	5.6%	0.6%	0.3%	0.0%	0.0%
BBB	0.0%	0.3%	5.4%	87.5%	5.3%	1.1%	0.2%	0.2%
BB	0.1%	0.1%	0.7%	9.4%	80.7%	7.0%	1.1%	1.0%
B	0.0%	0.1%	0.3%	0.5%	5.0%	80.3%	6.1%	7.7%
CCC	0.2%	0.0%	0.4%	0.8%	2.4%	12.1%	60.5%	23.7%

Table 2.2: *Average one-year transition probability matrix*
Source: *Standard & Poor's [2001a]*

What these rating transition matrices do not show, is how large rating changes are per rating change. If e.g. instant rating downgrades of three rating classes or more do not occur, this imposes structure on our transition probability matrix, which we can use later on to impose certain restrictions when estimating our pricing models. Information about this can be obtained from Standard & Poor's [2001b]. Among other things, we can find all 983 rating changes that have taken place during the course of 2000 in this document. For the unmodified ratings the histogram for the size of rating changes is as follows[1]:

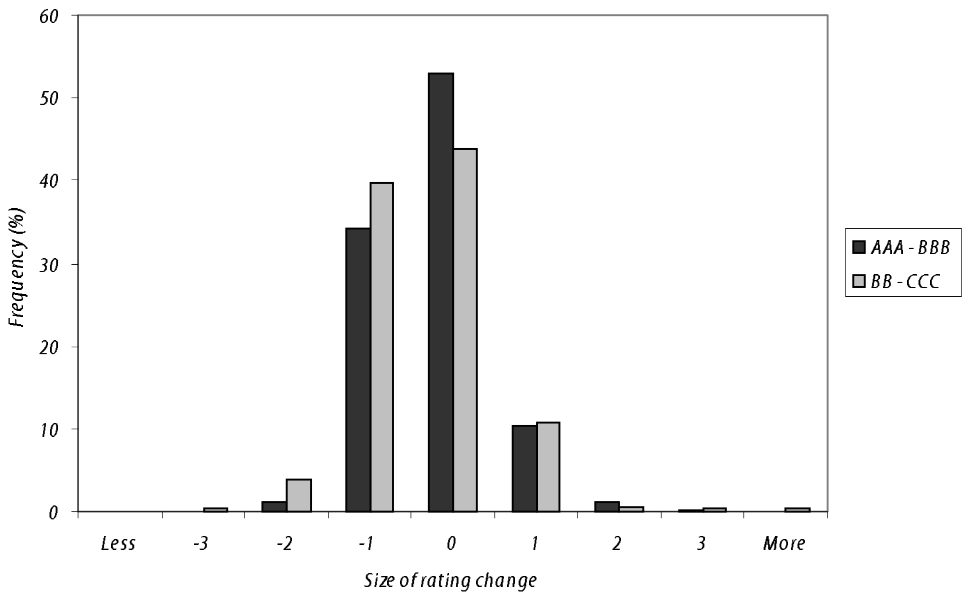


Figure 2.3: *Histogram of size of S&P's rating changes*

We notice that a rating change of more than 2 ratings is quite uncommon. Furthermore, just over one percent of the rating changes for investment grade bonds is of size 2 or larger, and for speculative grade bonds only a downgrade of size 2 is larger than one percent. If we were to impose that instant rating transitions of these sizes are impossible in our pricing model, this would introduce a large number of restrictions on our parameters, rendering the estimation process for our pricing model easier and our estimates more reliable.

[1] Rating changes of size zero occur since in this case the rating change was between two subratings of the same unmodified rating.

2.3. Recovery rates

2.3.1. Empirical facts on recovery rates

A critical aspect of corporate bond defaults is the severity of the loss that is incurred. In most cases, after default a bondholder is usually provided some amount of recovery, which may take the form of cash or other assets. Ideally recovery rates would be measured by tracking all payments made on a defaulted instrument and discounting them to one common date (e.g. the default date or the maturity date). One reason why this is hard to accomplish is that one must make assumptions concerning the value of certain payments. In case the bondholder receives various equity or derivative instruments or even physical assets as a recovery, it is hard to ascertain the exact value of such payments.
Both Moody's and Standard & Poor's therefore define the recovery rate as the percentage of face value returned to the bondholder. It is measured as the trading price of the defaulted instrument shortly after default, divided by the face value of the bond. Important determinants of the recovery rate of a bond are the seniority and the security of the bond, see e.g. Moody's [2000] and Standard & Poor's [2001a]. The seniority - senior, subordinated or junior - of a claim roughly determines in which order outstanding debt is repaid. Secured bonds are backed by a legal claim on some specific asset, whereas unsecured bonds are backed only by the promise of the issuer to pay. In the following graph the first, second (median) and third quartile of the recovery rate distribution are depicted for the combinations of seniority and security that can be found in the market.

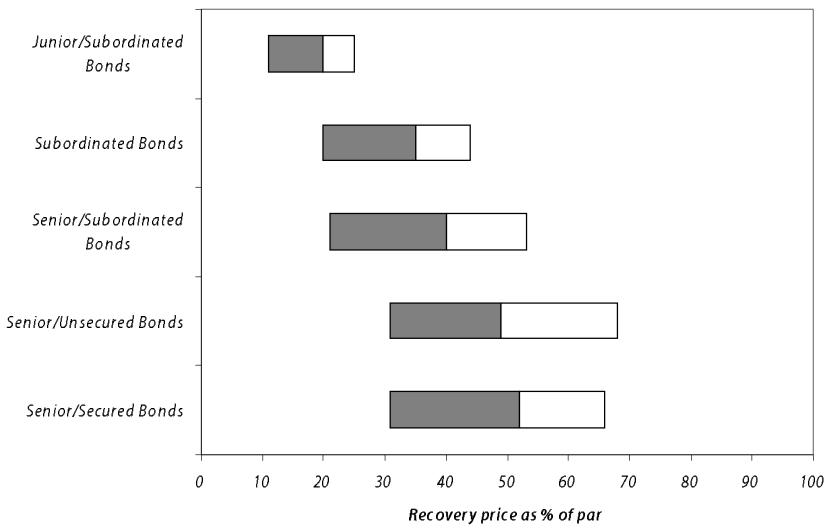


Figure 2.4: *Three quartiles of the recovery rate distribution*
Source: *Moody's [2000]*

We observe that there is a wide variation in recovery rates, and that the median recovery rate also depends strongly on the security/seniority of a bond.
Now we will turn to the treatment of recoveries in the literature on pricing credit derivatives.

2.3.2. Recovery rates in pricing models

Four different assumptions can be found in the literature concerning the recovery amount that is received by a bondholder. The bondholder could receive a fraction of:

(R1) the face value of the bond;
(R2) an equivalent default-free bond;
(R3) the value of the bond an instant before default;
(R4) the implicit accrued interest and the face value of the bond.

Furthermore, there is another assumption to be made, concerning the date on which the recovery amount is paid:

(a) immediately upon default;
(b) at the maturity of the original bond.

In the following, we will denote the recovery amount paid at a time t by $\varphi(t)$, as is common in some of the literature. Assumptions (R1) - (R4) are made on the level of a coupon bond. We will however start out with the value of a zero-coupon bond. If we assume no arbitrage, and therefore the existence of an equivalent martingale measure $\mathbb{Q}$ (see Harrison and Pliska [1981]), the price of a defaultable zero-coupon bond under assumption (a) can be written as follows:

$$\overline{P}(t,T) = B_t \cdot \mathbb{E}_t^{\mathbb{Q}}\left[\mathbb{1}_{[\tau>T]} \cdot B_T^{-1} + \varphi(\tau) \cdot \mathbb{1}_{[\tau\leq T]} \cdot B_\tau^{-1}\right]$$

where B_t is the value of the money market account at time t. For the basics of derivative pricing we refer you to appendix A - Mathematical Finance.
Under assumption (b) about the time of payment of the recovery amount we obtain:

$$\overline{P}(t,T) = B_t \cdot \mathbb{E}_t^{\mathbb{Q}}\left[\mathbb{1}_{[\tau>T]} \cdot B_T^{-1} + \varphi(T) \cdot \mathbb{1}_{[\tau\leq T]} \cdot B_T^{-1}\right]$$

We will now discuss the most common assumptions regarding the recovery process and relate them to the 8 possible assumptions we showed above.
According to the classification made by Duffie and Singleton [1999], assumption (R1a) is known as the *recovery of face value* (RFV) assumption, used in the article by e.g. Duffee [1998]. Moody's and S&P's use this assumption when measuring recovery rates, as we discussed in the previous subparagraph.
Assumptions (R1b), (R2a) and (R2b) can be seen to be equivalent for a zero-coupon bond, apart from the assumption whether the received fraction is known at the default time, or at the maturity date. Hereafter we denote the received fraction by the bondholders by δ, which in its most general form will be allowed to be a predictable stochastic process. Furthermore, we assume δ depends only on the information up to and including the time of default τ. This yields the following for the recovery under each of these assumptions:

(R1b) $\varphi(T) \equiv \delta(\tau)$
(R2a) $\varphi(\tau) \equiv \delta(\tau) \cdot P(\tau,T)$
(R2b) $\varphi(T) \equiv \delta(\tau)$

Assumption (R2a) is obviously equivalent to the other two assumptions, since the default-free zero-coupon bond equals 1 with certainty at the maturity date *T*.
Note that these assumptions are equivalent from a mathematical point of view. In practice there may however be some differences: the liquidity in the different situations is not necessarily equal, which may have consequences for the price of the product. These issues will be ignored throughout this thesis.
We will refer to assumptions (R2a) and (R2b) as the *recovery of treasury value* assumption (RT), used in the articles by e.g. Jarrow and Turnbull [1995] and Jarrow, Lando and Turnbull [1997]. As stated, assumption (R1b) is equivalent for a zero-coupon bond; when we are dealing with a coupon bond however, assumption (R1b) implies that the claim of the bondholder is limited to the face value of the coupon bond, not including the individual coupon payments. This assumption is employed by Bielecki and Rutkowski [1999, 2000a, 2000b].
Another common assumption made is (R3a), which is referred to as the *recovery of market value* assumption (RMV), used e.g. in the article by Duffie and Singleton [1999]. In this case the value of the defaultable bond at the time of default is equal to a fraction of the value just before default. Mathematically speaking:

$$\varphi(\tau) \equiv \delta(\tau) \cdot \overline{P}(\tau\text{-}, T) = \delta(\tau) \cdot \sup\left\{ \overline{P}(t,T) \;\middle|\; t < \tau \right\}$$

The RMV assumption is inspired by the default procedures in swap contracts. Standard agreements typically call for settlement upon default, based on an obligation represented by an otherwise equivalent, non-defaulted swap. It is also a quite natural assumption, as Schönbucher [1998] points out. The old claimants have to give up some of their claims in order to allow for rescue capital to be invested in the defaulted firm. They are not paid out in cash (this would drain the defaulted firm of valuable liquidity), but in new defaultable bonds of the same liquidity. In their article, Duffie and Singleton compare the RMV assumption to the RFV assumption, in a stylised example. It suffices to mention they assume the fractional recovery $\delta(t)$ to be constant, and that they specify a model for the hazard rate and the risk-free term structure. Under these assumptions, it turns out that the RMV assumption and the RFV assumption yield quite similar results for the bond prices. In cases where the bonds have a significant premium or discount, or where the term structures of interest rates are steeply upward or downward sloping, however, the RMV and RFV assumptions may differ more significantly.
Finally, Jarrow and Turnbull [2000] argue that the so-called *legal claim approach*, here represented by (R4a), may be the best assumption. This approach is easier to explain on a coupon bond level. The strip corresponding to the next coupon payment pays off C_i at time t_i. The claim of the bondholder is limited to accrued interest plus face value, i.e.:

$$\varphi(\tau) \equiv \delta(\tau) \cdot \left(C_i \cdot \alpha(t_{i-1,}\tau) + F \right)$$

where $\alpha(t,s)$ is the appropriate daycount fraction between dates *t* and *s*.

Implications for the prices of coupon bonds

In a risk-free setting, a coupon bond simply equals a portfolio of zero-coupon bonds. This desirable property is referred to as *value additivity*. It implies that the standard stripping procedures used to determine the implied zero-coupon bonds are appropriate.

Under the RT assumption, it is assumed that the claim of the bondholders at the time of default equals the present value of the not yet paid coupons, in addition to the principal of the coupon bond. It can easily be seen that in this case, a coupon bond is also equal to a portfolio of zero-coupon bonds. In the RMV case it can also be shown that this is true.

Under the remaining two assumptions we discussed before, the RFV assumption or the legal-claim approach, this unfortunately is not the case. Jarrow and Turnbull [2000] show this for the legal-claim approach; we will demonstrate this fact for the RFV assumption.

Suppose we have a coupon bond with coupons C_i at dates t_i, i = 1,...,n, and face value F. The value of a zero-coupon bond maturing at time t_i equals:

$$\overline{P}(t,t_i) = B_t \cdot \mathbb{E}_t^{\mathbb{Q}}\left[\mathbb{1}_{[\tau>t_i]} \cdot B_{t_i}^{-1} + \delta(\tau) \cdot \mathbb{1}_{[\tau\le t_i]} \cdot B_\tau^{-1}\right]$$

The value of the coupon bond as described before, can, under RFV, be seen to equal:

$$\sum_{i=1}^{n} C_i B_t \cdot \mathbb{E}_t^{\mathbb{Q}}\left[\mathbb{1}_{[\tau>t_i]} \cdot B_{t_i}^{-1}\right] + FB_t \cdot \mathbb{E}_t^{\mathbb{Q}}\left[\mathbb{1}_{[\tau>T]} \cdot B_T^{-1} + \delta(\tau) \cdot \mathbb{1}_{[\tau\le T]} \cdot B_\tau^{-1}\right]$$

In general, this cannot be expressed as a linear combination of various zero-coupon bonds, and therefore the value additivity property does not hold. If we can incorporate the stripping procedure within estimating the pricing model, this will not be a grave problem.

2.4. Pricing models for credit derivatives

Models for pricing credit derivatives can be split up in two groups: structural and reduced form models. The first type refers to models that describe the internal structure of the issuer of debt, so that default is a consequence of some internal event. These models are also referred to as firm's value models. Two further distinctions can be made within this class of models. In the first approach the firm's liabilities are modelled as claims issued against the firm's underlying assets. Bankruptcy is then determined via the evolution of the firm's assets in conjunction with the various debt covenants, see e.g. Merton [1974]. In the second approach risky debt is viewed as paying off an exogenously given fraction of each promised unit of cash in the event of bankruptcy. Bankruptcy is here determined when the value of the firm's underlying assets hits some prespecified boundary, see e.g. Longstaff and Schwartz [1995]. Both approaches are hard to implement, one reason being that they require all the firm's assets to be traded and observable. This is obviously not the case. Secondly, and this is important for our current research, these approaches do not utilise credit rating information, which makes them unable to price rating-dependent credit derivatives.

Reduced form models do not try and explain why a default event has occurred, but instead model the probability of a credit event (either default or a rating transition) directly. In the literature two further distinctions are made. Firstly we have the intensity-based models, which usually model the default time as the time of the first jump of some generalisation of a Poisson process. The article which inspired many of these models was Jarrow and Turnbull [1995]. Further generalisations are made in e.g. Schönbucher [1998] and Duffie and Singleton [1999]. Intensity-based models are in fact a subclass of the rating-based models, which directly model the transitions between the various credit ratings, and include the default state as a rating. This class of models will be the focus of interest in this thesis: explicitly modelling the rating transition (or credit migration) process will enable us to price rating-dependent credit derivatives, which is our ultimate goal. The seminal model within this class of models is the Jarrow, Lando and Turnbull [1997] model, which will be the main starting point for the models discussed in this thesis. Further extensions to this model can be found in e.g. Kijima and Komoribayashi [1998], Das and Tufano [1996] and Lando [1998]. A different rating-based approach is employed in the models of e.g. Schönbucher [2000] and Bielecki and Rutkowski [1999, 2000a, 2000b]. In these articles a Heath-Jarrow-Morton model is proposed for the evolution of default-free forward rates and credit spreads for all rating classes. Rating jumps are modelled as realisations of Poisson processes. In both these models imposing no-arbitrage leads directly to the behaviour of the rating transition process under the risk-neutral measure, allowing credit derivatives to be priced easily.

Having described many important issues which are necessary to know when pricing credit derivatives, we can now start by describing the starting point for pricing rating-dependent credit derivatives: the Jarrow, Lando and Turnbull model. This will be the topic of the next chapter.

3. Jarrow, Lando and Turnbull-related models

Jarrow, Lando and Turnbull [1997] (hereafter JLT) have developed a model for valuing risky debt that explicitly incorporates a firm's credit rating as an indicator of the likelihood of default. In this model the rating migration process follows a Markov chain. This model will be our starting point for valuing rating dependent credit derivatives. We first discuss the assumptions underlying the model, and then discuss the discrete-time formulation of the model. A broader class of rating-based models, of which the JLT model is a subclass, is the topic of investigation in paragraph 3.3. Alternative models are also briefly reviewed. The final topic of this chapter is the completeness of the JLT model.

3.1. Assumptions underlying the model

We first remark that we will use a slightly different notation than in the original article; this is because we aim to use a consistent notation throughout this thesis. The first assumption made by the authors, is the following:

(1) *There exists a unique equivalent martingale measure $\mathbb{Q}$ making all the default-free and risky zero-coupon bond prices martingales, after normalisation by the money market account.*

This assumption is equivalent to the assumption that the markets for default-free and risky zero-coupon bonds are arbitrage-free and complete, see e.g. the article by Harrison and Pliska [1981] for the equivalence of these two assumptions. The assumption that both markets are arbitrage-free is quite conventional in models for valuing credit risk derivatives; completeness is however not always assumed. The latter property is not strictly needed to arrive at a price for the derivatives; it is merely required if we want to be able to replicate the derivatives perfectly.

(2) *The recovery of treasury assumption (RT) holds. Furthermore, the fraction $\delta(\tau)$ of the otherwise equivalent default-free bond which is repaid to the bondholder, is a constant $0 \leq \delta \leq 1$.*

As we saw in the paragraph about recoveries in the previous chapter, the recovery rate will vary quite considerably per issue. Later on in this paragraph we will show how this assumption can be relaxed somewhat, to still be consistent with the model of Jarrow, Lando and Turnbull. A final assumption which is made, irrespective of the choice of a discrete-time or continuous-time framework, is:

(3) *The stochastic process for the default-free spot rate r_t and the bankruptcy process, as specified by the default time τ, are independent under $\mathbb{Q}$.*

This assumption is imposed for simplicity of implementation. In the real world both processes are obviously related. If the default-free spot rate rises, borrowing money will become more expensive for firms, indirectly making the probability that a company will default on its outstanding loans higher. Whether this relation also holds under the equivalent martingale measure, is a different matter. Without this assumption it is still possible to arrive at the same model for valuing risky debt as in

the JLT model, which we shall demonstrate later on in this paragraph. First we derive the pricing equation for defaultable bonds in the JLT model.
Assumption (2) implies the following expression for the value at maturity:

$$\overline{P}(T,T) = \mathbb{1}_{[\tau > T]} + \delta \cdot \mathbb{1}_{[\tau \leq T]}$$

Using assumption (1) and normalising by the money market account, yields:

$$\begin{aligned}
\overline{P}(t,T) &= \mathbb{E}_t^{\mathbb{Q}}\left[\tfrac{B(t)}{B(T)} \cdot (\mathbb{1}_{[\tau > T]} + \delta \cdot \mathbb{1}_{[\tau \leq T]})\right] \\
&= \mathbb{E}_t^{\mathbb{Q}}\left[\tfrac{B(t)}{B(T)}\right] \cdot \mathbb{E}_t^{\mathbb{Q}}\left[\mathbb{1}_{[\tau > T]} + \delta \cdot \mathbb{1}_{[\tau \leq T]}\right] \\
&= P(t,T) \cdot \left(\delta + (1-\delta) \cdot \mathbb{Q}_t(\tau > T)\right)
\end{aligned} \tag{3.1}$$

where we used the third assumption in the second step, and the definition of the default-free zero-coupon bond in the final step. This is the pricing equation for defaultable zero-coupon bonds as presented in the article by JLT.

Using the forward measure
Instead of normalising by the money market account, we could normalise by the default-free zero-coupon bond maturing at time *T*. The assumption that the market is arbitrage-free implies that there exists an equivalent probability measure under which all traded assets, normalised by this zero-coupon bond, are martingales. In the literature this measure is referred to as the forward measure, and denoted by $\mathbb{P}^T$. We assume the same properties hold for this measure as did for $\mathbb{Q}$, see assumption 1. Furthermore, we combine the second and third assumption into the following:

(2') *The recovery of treasury assumption (RT) holds. We assume the fraction $\delta(\tau)$ of the otherwise equivalent default-free bond which is repaid to the bondholder, satisfies the following restrictions:*

(i) $\mathbb{E}_t^{\mathbb{P}^T}\left[\delta(\tau)\right] = \delta$ for all $0 \leq t \leq T$.

(ii) $Cov_t^{\mathbb{P}^T}\left(\delta(\tau),\, \mathbb{1}_{[\tau \leq T]}\right) = 0$

We will now see that by just using assumptions (1) and (2'), we arrive at a pricing equation of the same form as (3.1). The expression for the time *t* value of the risky zero-coupon bond can be written as:

$$\begin{aligned}
\overline{P}(t,T) &= \mathbb{E}_t^{\mathbb{P}^T}\left[\tfrac{P(t,T)}{P(T,T)} \cdot (\mathbb{1}_{[\tau > T]} + \delta(\tau) \cdot \mathbb{1}_{[\tau \leq T]})\right] \\
&= P(t,T) \cdot \left(\mathbb{P}_t^T(\tau > T) + \mathbb{E}_t^{\mathbb{P}^T}\left[\delta(\tau)\right] \cdot \mathbb{P}_t^T(\tau \leq T) + Cov_t^{\mathbb{P}^T}(\delta(\tau), \mathbb{1}_{[\tau \leq T]})\right) \\
&= P(t,T) \cdot \left(\delta + (1-\delta) \cdot \mathbb{P}_t^T(\tau > T)\right)
\end{aligned} \tag{3.2}$$

Using less restrictive assumptions still yields the same model as the JLT model. A problem with the forward measure is that it depends on the maturity of the underlying product. Therefore, if we are trying to infer the behaviour of the pricing measure from the quoted prices of traded credit derivative instruments, we can, without further assumptions, only use products of the same maturity. This will turn out to be quite a problem; to solve this, we will introduce the relaxed JLT framework, which is a more general version of assumptions (1) - (3).

Relaxed JLT framework

Hereafter we will price credit derivatives under the measure $\mathbb{Q}$. Completeness is not required to arrive at the same equation as (3.1). Also, assumptions (2) and (3) are again not strictly necessary. Adding one extra condition to (2') is however necessary:

(2") *The recovery of treasury assumption (RT) holds. We assume the variables satisfy the following restrictions:*

(i) $\mathbb{E}_t^{\mathbb{Q}}[\delta(\tau)] = \delta$ for all $0 \leq t \leq T$.

(ii) $Cov_t^{\mathbb{Q}}\left(\frac{1}{B(T)}, \mathbb{1}_{[\tau \leq T]}\right) = 0$

(iii) $Cov_t^{\mathbb{Q}}\left(\delta(\tau), \frac{1}{B(T)} \cdot \mathbb{1}_{[\tau \leq T]}\right) = 0$

A relaxed version of condition (3) is now incorporated into condition (2''). In this assumption we have allowed the recovery rate to be stochastic, as is the case in practice. Furthermore, the fact that the covariance of two processes is zero, does not necessarily mean that the two processes are independent. Therefore, under less restrictive and slightly more realistic assumptions, we arrive at the same pricing equation as (3.1).

3.2. Rating transition process: the discrete-time case

We consider an economy in the time interval $[0, T^*]$, where T^* is some horizon date. Trades, price changes and rating transitions may take place on $0, \triangle, \ldots, T^* - \triangle, T^*$. The rating transition process is modelled as a Markov chain[2] on the state space $S = \{1, \ldots, K\}$. This state space represents the set of possible credit ratings; state 1 is the highest possible rating, state 2 the second highest, etc. The final state represents bankruptcy and is assumed to be absorbing, i.e. once a firm enters bankruptcy, it cannot recover. Restructuring is therefore not possible in this model. We denote the credit rating at time t by X_t. The default time can now be written as $\tau = \inf\{t \geq 0 : X_t = K\}$.

The behaviour of the Markov chains will be described via transition matrices. The matrix $A(t,s)$ with elements $a_{ij}(t,s)$ specifies the transition probabilities from any state to another over the time interval $[t,s]$. The transition matrices under the equivalent pricing measure will be denoted as $\widetilde{A}(t,s)$ with elements $\widetilde{a}_{ij}(t,s)$. The entries of the transition matrix under $\mathbb{P}$ must satisfy (for all $t \leq s$):

$$
\begin{aligned}
0 \leq a_{ij}(t,s) &\leq 1 \\
\sum_{j=1}^{K} a_{ij}(t,s) &= 1 \qquad i = 1, \ldots, K \\
a_{Ki}(t,s) &= 0 \qquad i = 1, \ldots, K-1 \\
a_{KK}(t,s) &= 1
\end{aligned}
$$

[2] We refer to Kijima [1997] for more information about Markov chains.

In Jarrow, Lando and Turnbull, the rating transition process is modelled as a time-homogenous Markov chain. However, specifying the transition process under the real-world probability $\mathbb{P}$ is not required per se. For the purpose of pricing derivatives it is a well-known fact that we merely require the behaviour of all processes under some equivalent probability measure, which turns all traded assets into martingales once they are discounted by a numéraire asset. Modelling the behaviour of the default-free zero-coupon bonds under the pricing measure is a fairly well known subject in the literature; we therefore only require the process of rating transitions under the pricing measure $\mathbb{Q}$. To facilitate empirical estimation, most articles on this topic impose some restriction on the transition probabilities under the pricing measure. JLT are no exception, as they impose a restriction that specifies a dependency between the transition probabilities under $\mathbb{P}$ and the probabilities under the pricing measure. In this paragraph we will show why such a restriction is imposed. The next paragraph will deal with a number of methods proposed for obtaining the transition probabilities under $\mathbb{Q}$, including the original JLT assumption.

Transition probabilities under $\mathbb{Q}$

In the discrete-time framework, under the pricing measure $\mathbb{Q}$, the one-step transition matrix from time t to time $t + \triangle$, $\tilde{A}(t, t+\triangle)$, has entries $\tilde{a}_{ij}(t, t+\triangle)$ that in general satisfy the following equations:

$$\tilde{a}_{ij}(t, t+\triangle) = \pi_{ij}(t) \cdot a_{ij}(t, t+\triangle) \qquad i, j \in S, i \neq K$$

$$\sum_{j=1}^{K} \pi_{ij}(t) \cdot a_{ij}(t, t+\triangle) = 1 \qquad i \in S, i \neq K$$

$$\pi_{ij}(t) > 0 \Leftrightarrow a_{ij} > 0 \qquad i, j \in S$$

The $\pi_{ij}(t)$ can be interpreted as risk premiums, and are $\mathcal{F}_t$- measurable, if we do not impose any further restrictions. An economic interpretation for these risk premiums can not immediately be given. However, let us consider the probability that a firm with rating i defaults from time t to time T under $\mathbb{Q}$. Using elementary Markov chain theory, we know that:

$$\tilde{A}(t, T) = \tilde{A}(t, T-\triangle) \cdot \tilde{A}(T-\triangle, T)$$

hence:

$$\mathbb{Q}_t^i(\tau \leq T) = \tilde{a}_{iK}(t, T) = \sum_{j=1}^{K} \tilde{a}_{ij}(t, T-\triangle) \cdot \tilde{a}_{jK}(T-\triangle, T)$$

It is immediately clear that higher risk premiums cause higher default probabilities under $\mathbb{Q}$. Higher default probabilities imply lower prices for defaultable zero-coupon bonds, through (3.1).

To simplify notation we assume that all transition probabilities are deterministic, i.e. future transition probabilities do not depend on any as of yet unknown random variable. From the pricing equation (3.1) we see that we can obtain cumulative default probabilities from the data:

$$\mathbb{Q}_t^i(\tau \le T) = \tilde{a}_{iK}(t,T) = \frac{P(t,T) - \overline{P}^i(t,T)}{(1-\delta)\cdot P(t,T)} \tag{3.3}$$

where we denote the price of a defaultable bond maturing at time T, currently in rating class i, by $\overline{P}^i(t,T)$.

Turning to the amount of transition probabilities that need to be estimated per one-step transition matrix under $\mathbb{Q}$, we see that there are $(K–1)^2$ unknowns to be estimated[3] from the data, for which we require at least $(K–1)^2$ bonds per maturity date. If we have more (or less, which will typically be the case) than this amount, we will have to resort to some kind of least squares minimisation procedure. We will return to this type of procedure in chapter 4, since in practice this will turn out to be quite useful.

In their article, JLT assume that we have exactly one bond per rating class and maturity, and that bonds of all maturities are traded. Of course, this is not the case in practice, so that prior to calibrating the model we will need to estimate the corporate term structures, using e.g. some of the techniques in appendix B.

At any rate, a restrictive assumption about the structure of the risk premiums is necessary to be able to estimate the transition probabilities. The JLT restriction is a special case of a broader class of restrictions that can be imposed on the risk premiums. We will discuss these so-called column-independent risk premiums in the next paragraph.

3.3. Column-independent risk premiums

As demonstrated in the previous paragraph, we require the knowledge of the behaviour of the rating transition process under the equivalent probability measure $\mathbb{Q}$ in order to price credit derivatives. This paragraph will deal with a particular class of risk premiums: column-independent risk premiums[4]:

$$\pi_{ij}(t) = \pi_i(t) \qquad\qquad j \ne \nu(i) \tag{3.4}$$

where $\nu(i)$ is an entry in row i that is used for normalisation. We will discuss this in paragraph 3.3.1. A justification for column-independence is the following: as stated before, per one-step transition matrix there are $(K-1)^2$ unknowns to be estimated from the data. With the assumptions we have made, we only have $K-1$ bonds per maturity. Therefore, an assumption about the structure of the risk premiums is needed. From (3.3) it can be seen that at each time t we only observe the default column of each multi-step transition matrix from the data. A column-independent risk premium therefore seems a natural choice. The first assumption concerns the nature of the underlying Markov chain:

- **Time-homogeneous:** we observe the one-step transition matrix A; the forward and cumulative transition matrices are calculated as follows:

  $$A(t, t+\triangle) = A \qquad\qquad A(0, n\triangle) = A^n$$

[3] The transition probabilities from state K to any other state are known beforehand, since state K is assumed to be absorbing. Furthermore, the sum of each row has to equal 1.

[4] The classification of these risk premiums is due to Houweling [2000]. We dubbed this particular class of risk premiums the *column-independent risk premiums*.

- **Time-heterogeneous:** we observe the multi-step transition matrices A_t; the forward and cumulative transition matrices are calculated as follows:

$$A(t, t+\triangle) = A_t^{-1} \cdot A_{t+\triangle} \qquad A(0,t) = A_t$$

The one-year or multi-year transition matrices can readily be obtained from rating agencies, such as Moody's or Standard & Poor's. Secondly, we can choose to adjust either the forward or the cumulative transition matrices under $\mathbb{P}$ to obtain the transition probabilities under $\mathbb{Q}$. The two methods are:

- **Forward method:** adjust the forward matrix $A(t,t+\triangle)$ to obtain $\widetilde{A}(t,t+\triangle)$:

$$\widetilde{a}_{ij}(t,t+\triangle) = \pi_i(t,t+\triangle) \cdot a_{ij}(t,t+\triangle) \qquad j \neq \nu(i)$$

- **Cumulative method:** adjust the cumulative matrix $A(0,t)$ to obtain $\widetilde{A}(0,t)$:

$$\widetilde{a}_{ij}(0,t) = \pi_i(0,t) \cdot a_{ij}(0,t) \qquad j \neq \nu(i)$$

A choice for either method will have an impact on the numerical stability of the risk premiums, as we will see in this chapter. Finally, a choice has to be made about which entry in each row is used to ensure that the probabilities from each row sum to one. In the literature, two different approaches are proposed; the nomenclature will be explained in 3.3.1:

- **Ratio of default probabilities:** this method was initially used by Jarrow, Lando and Turnbull [1997];

- **Ratio of survival probabilities:** this is the alternative proposed by Kijima and Komoribayashi [1998].

This yields a total of eight combinations to calculate risk premiums. The next subparagraph deals with the choice of normalisation.

3.3.1. Normalisation

Irrespective of whether we assume a time-homogeneous or time-heterogeneous process under $\mathbb{P}$, or whether we use forward or cumulative risk premiums, assume we have an empirical transition matrix A and we want to obtain the transition matrix under $\mathbb{Q}$, $\widetilde{A}$, through the use of risk premiums. The entries of the default column of this matrix are available.
The risk premiums for row K are not interesting; due to the equivalence of $\mathbb{P}$ and $\mathbb{Q}$ the default state must also be absorbing under $\mathbb{Q}$, which fully specifies the transition probabilities for that row. From here onward we will not consider row K.
Suppose that for row $i < K$ we choose to normalise in column $\nu(i)$. This implies:

$$\widetilde{a}_{ij} = \begin{cases} \pi_i \cdot a_{ij} & j \neq \nu(i) \\ 1 - \pi_i \cdot (1 - a_{i\nu}) & j = \nu(i) \end{cases} \tag{3.5}$$

The problem we are faced with is that of the choice of $\nu(i)$. Since the default column of the multi-step transition matrices can be inferred from market data through the use of equation (3.3), the risk premium will be calculated using this information. From (3.5) we can deduce the default probabilities under $\mathbb{Q}$ from rating i; using this information it is possible to calculate two different risk premiums:

$$\pi_i = \begin{cases} \dfrac{\tilde{a}_{iK}}{a_{iK}} & \nu(i) < K \\ \dfrac{1-\tilde{a}_{iK}}{1-a_{iK}} & \nu(i) = K \end{cases} \tag{3.6}$$

This explains the earlier nomenclature: in the first case the risk premium is the ratio of default probabilities, in the second case the risk premium is the ratio of survival probabilities. If we use the ratio of default probabilities, we still have to choose which column $\nu(i) < K$ will be used for normalisation. Let us examine the lower and upper bounds for the risk premium if we normalise in column $\nu(i) < K$. In general each risk premium must be larger than or equal to zero. If in row i of matrix A there is an element $a_{ij} > 0$ for $j \neq \nu(i)$, the risk premium must be larger than 0 due to the assumed equivalence of both probability distributions. In order to derive an upper bound we must examine (3.5) closer. The elements $\tilde{a}_{ij}$ already sum to one, but they must also lie in the interval [0,1] to ensure that they are probabilities. This yields the following bounds[5] for the risk premium for row i:

$$0 \leq \pi_i(j) \leq u_i(\nu(i)) = \min\left\{\frac{1}{1-a_{i\nu(i)}}, \min_{j\neq\nu(i),\, a_{ij}>0} \frac{1}{a_{ij}}\right\} \tag{3.7}$$

We have here also assumed that $a_{i\nu(i)} < 1$ for all $\nu(i)$, which will in practice be the case. It seems wise to choose $\nu(i)$ so that we have the largest upper bound in (3.7). In this way we are ensured that the possibility of the risk premium being out of bounds is as small as possible. As we witnessed in table 2.2, the diagonal entry is usually the largest entry in each row of a historical transition matrix. Additionally we assume:

$$0 < a_{ii_{\min}} = \min_{j,\, a_{ij}>0} a_{ij} \leq a_{ii_{\max}} = \max_{j\neq i} a_{ij} \leq a_{ii} = \max_{j} a_{ij} < 1 \tag{3.8}$$

The historical transition matrices we have encountered in practice all satisfy this property. Given (3.7) and (3.8), the upper bounds satisfy:

$$u_i(\nu(i)) = \begin{cases} \min\left\{\dfrac{1}{1-a_{i\nu(i)}}, \dfrac{1}{a_{ii}}\right\} = \dfrac{1}{1-a_{i\nu(i)}} & \nu(i) \neq i \\ \min\left\{\dfrac{1}{1-a_{ii}}, \dfrac{1}{a_{ii_{\max}}}\right\} = \dfrac{1}{1-a_{ii}} & \nu(i) = i \end{cases} \tag{3.9}$$

This can be proven with (3.8) in mind. Take e.g. the equation for $\nu(i) = i$:

$$a_{ii_{\max}}^{-1} \leq (1-a_{ii})^{-1} \quad \Leftrightarrow \quad a_{ii_{\max}} \geq 1-a_{ii} \quad \Rightarrow \quad a_{ii_{\max}} + a_{ii} \geq 1$$

[5] Note that these bounds also hold for the case where $\nu(i) = K$.

The latter is impossible due to the restriction we imposed in (3.8). This leaves us with the result in (3.9). Since the diagonal entry was assumed to be the largest, we see that using ν = *i* for normalisation yields the largest possible upper bound for the risk premium. In future we will therefore only consider normalising on the diagonal when using risk premiums defined as the ratio of default probabilities. We now continue with the forward risk premiums.

3.3.2. Forward risk premiums

If we use the forward method, the forward transition probabilities are based on the observed forward probabilities:

$$\tilde{a}_{ij}(t,t+\triangle) = \pi_i(t,t+\triangle)\cdot a_{ij}(t,t+\triangle) \qquad j \neq \nu \tag{3.10}$$

From equation (3.3) it can be seen that we observe the cumulative default probabilities from the observed data on corporate bond prices. The forward probabilities can be related to the cumulative default probabilities as follows:

$$\tilde{A}(0,t+\triangle) = \tilde{A}(0,t)\cdot\tilde{A}(t,t+\triangle)$$

and hence (provided the transition matrix from 0 to *t* is invertible):

$$\tilde{A}(t,t+\triangle) = \tilde{A}^{-1}(0,t)\cdot\tilde{A}(0,t+\triangle)$$

Entry (*i*,*K*) of this forward matrix equals:

$$\tilde{a}_{iK}(t,t+\triangle) = \sum\nolimits_{j=1}^{K}\tilde{a}_{ij}^{-1}(0,t)\cdot\tilde{a}_{jK}(0,t+\triangle) \tag{3.11}$$

where $\tilde{a}_{ij}^{-1}(0,t)$ is entry (*i*,*j*) of $\tilde{A}^{-1}(0,t)$. Inserting the result from equation (3.3), which relates the cumulative default probabilities to the observable bond prices, yields:

$$\tilde{a}_{iK}(t,t+\triangle) = \sum\nolimits_{j=1}^{K-1}\tilde{a}_{ij}^{-1}(0,t)\cdot\frac{P(0,t+\triangle)-\overline{P}^{j}(0,t+\triangle)}{(1-\delta)\cdot P(0,t+\triangle)} + \tilde{a}_{iK}^{-1}(0,t) \tag{3.12}$$

where we used the fact that $\tilde{a}_{KK}(0,t+\triangle) = 1$.

We now turn to the choice of normalisation.

Ratio of default probabilities

We use the diagonal element of the transition probability matrix for normalisation. We can rewrite (3.10) as follows:

$$\tilde{a}_{ij}(t,t+\triangle) = \begin{cases}\pi_i(t,t+\triangle)\cdot a_{ij}(t,t+\triangle) & j \neq i \\ 1-\pi_i(t,t+\triangle)\cdot\left(1-a_{ii}(t,t+\triangle)\right) & j = i\end{cases}$$

The risk premium is the ratio of forward default probabilities:

$$\pi_i(t,t+\triangle) = \frac{\tilde{a}_{iK}(t,t+\triangle)}{a_{iK}(t,t+\triangle)} \tag{3.13}$$

Substituting (3.11) yields a more general version of the JLT risk premium:

$$\pi_i(t,t+\triangle) = \frac{1}{a_{iK}(t,t+\triangle)} \cdot \left[\sum_{j=1}^{K-1} \tilde{a}_{ij}^{-1}(0,t) \cdot \frac{P(0,t+\triangle) - \overline{P}^{j}(0,t+\triangle)}{(1-\delta) \cdot P(0,t+\triangle)} + \tilde{a}_{iK}^{-1}(0,t) \right] \tag{3.14}$$

The risk premium from the JLT article is obtained if we assume that we have a time-homogenous Markov chain under $\mathbb{P}$.

Ratio of survival probabilities
In this method, the final column is used for normalisation. Equation (3.10) becomes:

$$\tilde{a}_{ij}(t,t+\triangle) = \begin{cases} \pi_i(t,t+\triangle) \cdot a_{ij}(t,t+\triangle) & j \neq K \\ 1 - \pi_i(t,t+\triangle) \cdot \left(1 - a_{iK}(t,t+\triangle)\right) & j = K \end{cases}$$

From the equation for $j = K$, we see that the risk premium is the ratio of forward survival probabilities:

$$\pi_i(t,t+\triangle) = \frac{1 - \tilde{a}_{iK}(t,t+\triangle)}{1 - a_{iK}(t,t+\triangle)} \tag{3.15}$$

To yield an easy expression for the risk premium, in terms of the observable bond prices, we first observe the following property of the inverse transition matrix:

$$\begin{aligned} &\tilde{A}(0,t)^{-1} \cdot \tilde{A}(0,t) = I_K \iff \sum_{j=1}^{K} \tilde{a}_{ij}^{-1}(0,t) \cdot \tilde{a}_{jm}(0,t) = \mathbb{1}_{[m=i]} \\ &\sum_{j=1}^{K} \tilde{a}_{ij}^{-1}(0,t) = \sum_{j=1}^{K} \tilde{a}_{ij}^{-1}(0,t) \sum_{m=1}^{K} \tilde{a}_{jm}(0,t) \\ &\qquad = \sum_{m=1}^{K} \sum_{j=1}^{K} \tilde{a}_{ij}^{-1}(0,t) \cdot \tilde{a}_{jm}(0,t) = \sum_{m=1}^{K} \mathbb{1}_{[m=i]} = 1 \end{aligned} \tag{3.16}$$

Therefore, we can rewrite (3.11) as:

$$1 - \tilde{a}_{iK}(t,t+\triangle) = \sum_{j=1}^{K} \tilde{a}_{ij}^{-1}(0,t) \cdot \left(1 - \tilde{a}_{jK}(0,t+\triangle)\right)$$

Inserting the result from (3.3) yields:

$$1 - \tilde{a}_{iK}(t,t+\triangle) = \sum_{j=1}^{K-1} \tilde{a}_{ij}^{-1}(0,t) \cdot \frac{\overline{P}^{j}(0,t+\triangle) - \delta \cdot P(0,t+\triangle)}{(1-\delta) \cdot P(0,t+\triangle)}$$

where we used the fact that $1 - \tilde{a}_{KK}(0,t+\triangle) = 0$.

We substitute this in (3.15):

$$\pi_i(t,t+\triangle)=\frac{1}{1-a_{iK}(t,t+\triangle)}\cdot\left[\sum_{j=1}^{K-1}\tilde{a}_{ij}^{-1}(0,t)\cdot\frac{\overline{P}^{j}(0,t+\triangle)-\delta\cdot P(0,t+\triangle)}{(1-\delta)\cdot P(0,t+\triangle)}\right] \tag{3.17}$$

This is a general version of the risk premium proposed by Kijima and Komoribayashi [1998]. Comparing this to the JLT version, we find that the expression between brackets is here divided by the one-step survival probability in rating class *i*. In the JLT risk premium, we divided by the one-step default probability. This is notoriously difficult to estimate: for high rating classes it will be very close to zero. This would make the corresponding risk premium very large, which may in turn cause the transition matrix not to be a true transition probability matrix. In (3.17) this problem is circumvented. These kinds of numerical problems in calculating risk premiums shall be discussed in paragraph 3.3.4. We now turn to the cumulative risk premiums.

3.3.3. Cumulative risk premiums

When using the cumulative method we adjust the *t*-step transition matrix in order to obtain the transition matrix under $\mathbb{Q}$:

$$\tilde{a}_{ij}(0,t)=\pi_i(0,t)\cdot a_{ij}(0,t) \qquad j\neq \nu(i) \tag{3.18}$$

We again discuss the two normalisation rules for the risk premiums.

Ratio of default probabilities

Under this normalisation method, equation (3.18) can be rewritten as:

$$\tilde{a}_{ij}(0,t)=\begin{cases}\pi_i(0,t)\cdot a_{ij}(0,t) & j\neq i\\ 1-\pi_i(0,t)\cdot\left(1-a_{ii}(0,t)\right) & j=i\end{cases}$$

The expressions for the cumulative risk premiums are a lot easier than the forward risk premiums. We immediately see:

$$\pi_i(0,t)=\frac{1}{a_{iK}(0,t)}\cdot\frac{P(0,t)-\overline{P}^{i}(0,t)}{(1-\delta)\cdot P(0,t)} \tag{3.19}$$

Though (3.19) is a much more simple expression than the JLT risk premium, we still divide by a *t*-step default probability. Provided that *t* is large enough, this should not be a problem; however, it is more stable to divide by a survival probability, as will be done in the last risk premium which we discuss.

Ratio of survival probabilities

Under this normalisation method equation (3.18) becomes:

$$\tilde{a}_{ij}(0,t)=\begin{cases}\pi_i(0,t)\cdot a_{ij}(0,t) & j\neq K\\ 1-\pi_i(0,t)\cdot\left(1-a_{iK}(0,t)\right) & j=K\end{cases}$$

We immediately obtain the following for the risk premium:

$$\pi_i(0,t) = \frac{1}{1-a_{iK}(0,t)} \cdot \frac{\overline{P}^i(0,t) - \delta \cdot P(0,t)}{(1-\delta)\cdot P(0,t)} \tag{3.20}$$

The merits of (3.20) over (3.19) are immediately obvious. Since equation (3.20) also is less computationally intensive than the general version of the Kijima and Komoribayashi risk premium in (3.17), we at the moment expect that this type of risk premium will yield the most stable results.

3.3.4. Problems with column-independent risk premiums

We will now address various problems that can arise when calculating column-independent risk premiums. The following four problems are the most common:

1. *The arbitrariness of the assumption of column-independence.*

There is no real theoretical motivation for the column-independence of risk premiums. Furthermore, even if the column-independence assumption is true, we still do not know on which column we should normalise. That the choice of normalisation can have quite a significant impact can be seen from the following example. We took the transition matrices and prices of zero-coupon bonds for the various ratings from the Jarrow, Lando and Turnbull [1997] article. We calculated the Jarrow, Lando and Turnbull risk premiums (normalisation on the diagonal) as well as the Kijima and Komoribayashi (KK) risk premiums (normalisation on the default column). Using the transition matrices under $\mathbb{Q}$ we are now able to calculate the price of credit derivatives. Let us take a downgrade put, that pays € 1 in one year time if a bond, which is currently rated A, is downgraded below A. We assume zero recovery for this product, i.e. if the bond defaults, the downgrade put pays nothing. The price of this product using the JLT risk premiums equals € 0.14, whereas the price of this product using the KK risk premiums equals € 0.08. This is a very considerable difference.

2. *The resulting adjusted transition matrix is not a transition probability matrix.*

We may very well be able to fit the initial term-structure of defaultable bonds perfectly with the risk premiums, but the results are meaningless if the resulting transition matrices are not transition probability matrices. This is the case if the appropriate bounds as derived in (3.7) are violated. We shall demonstrate this potential problem for the case of cumulative risk premiums, defined as the ratio of default probabilities. The risk premiums satisfy:

$$\tilde{a}_{ij}(0,t) = \begin{cases} \pi_i(0,t)\cdot a_{ij}(0,t) & j \neq i \\ 1-\pi_i(0,t)\cdot\left(1-a_{ii}(0,t)\right) & j = i \end{cases}$$

From the results derived in section 3.3 we find that the following bounds hold:

$$0 \le \pi_i(0,t) \le \frac{1}{1-a_{ii}(0,t)}$$

Furthermore, if $a_{ii}(0,t) < 1$ for all $i < K$, we must have that $\pi_i(0,t) > 0$. If these constraints are violated for the risk premium as calculated by (3.19), the model obviously does not represent the market appropriately. For example, suppose that we have $\pi_i(0,t) > \frac{1}{a_{iK}(0,t)}$, which in turn implies a violation of the upper bound. From the JLT pricing equation we see that this means that $\overline{P}^i(0,t) < \delta \cdot P(0,t)$, which is something we would not expect given our model. This could either mean there is an arbitrage opportunity in the market or that our model is wrong.
One possible way around the problem would be to use a different normalisation method for that specific row; this however does not always overcome the problem, as can be seen in the following example.

Example
We suppose that $\pi_i(0,t) > \frac{1}{a_{iK}(0,t)}$ and denote the survival risk premium by $\ell_i(0,t)$. We could decide to switch to using survival risk premiums for row i. From (3.19) and (3.20) we find the following relation between both risk premiums:

$$\ell_i(0,t) = \frac{-a_{iK}(0,t)\cdot\pi_i(0,t)+1}{1-a_{iK}(0,t)} < \frac{-a_{iK}(0,t)\cdot\frac{1}{a_{iK}(0,t)}+1}{1-a_{iK}(0,t)} = 0$$

In this case the survival risk premium will violate its lower bound, and will also not yield a proper transition probability matrix. □

If we are willing to sacrifice the property that the initial term-structure will be perfectly fitted, we may replace the risk premium by its upper bound (if it is the upper bound which is violated), or by a very small positive number[6] (if the lower bound is violated). Since the price of a defaultable bond is monotonic in the risk premium, this will ensure that the price of that certain bond is matched as best as possible, given the model we are working in.
If we encounter such a problem when using forward risk premiums, the problem is more serious. The error in the adjusted transition matrix will most likely have repercussions for the transition matrices over longer periods.

3. Equivalence of $\mathbb{Q}$ *and* $\mathbb{P}$ *is not guaranteed.*

We initially assumed that the risk-neutral probability measure $\mathbb{Q}$ is equivalent to $\mathbb{P}$. If we denote the historical transition matrix (be it the forward or the cumulative transition matrix, dependent on which method we use) by A, this entails:

$$a_{ij} = 0 \quad \Leftrightarrow \quad \tilde{a}_{ij} = 0 \qquad\qquad a_{ij} > 0 \quad \Leftrightarrow \quad \tilde{a}_{ij} > 0$$

When using risk premiums, we may not obtain two equivalent probability distributions. In practice, the only case that arises is that of a_{iK} being equal to zero. As mentioned before, default probabilities are notoriously hard to estimate and their estimates may even equal zero. We will discuss the potential problems of a_{iK} being zero for both normalisation methods.

[6] A risk premium equal to zero would in general not yield two equivalent probability distributions.

- **Ratio of default probabilities**

The risk premium in this case is not well defined by (3.13) or (3.19). Any risk premium within the bounds we supplied earlier will yield a proper transition probability matrix. This unfortunately does not guarantee that the fitted price will be equal to the market price. When using cumulative risk premiums (this result holds for both normalisation methods), the model price of a defaultable bond must equal:

$$\overline{P}^{i}(0,t) = P(0,t) \cdot \left(\delta + (1-\delta) \cdot \left(1 - \tilde{a}_{iK}(0,t)\right)\right) = P(0,t)$$

due to the equivalence of both probability distributions. In general, the market price will not satisfy this equation.

- **Ratio of survival probabilities**

The risk premium in this case is well defined. However, for $\tilde{a}_{iK}$ to be equal to zero, the risk premium must be equal to one. For cumulative risk premiums this is again only satisfied if $\overline{P}^{i}(0,t) = P(0,t)$, which in general will not be the case.

A way to circumvent the aforementioned problems is to replace the zero entry in the default column by a small number and adjust the other probabilities accordingly, so that we still have a transition probability matrix. Naturally, this will distort the empirical transition matrix. Nevertheless, since the default probabilities are hard to estimate, the true default probability may very well be different than zero. Hereafter we will always replace zero entries in the default column by the smallest non-zero entry in the transition matrix[7]. To ensure that we still have rows that sum to one, we will subtract this number from the diagonal entry - typically the largest entry in each row - in that row.

4. *One-step transition matrices calculated from cumulative transition probability matrices do not necessarily have to be transition probability matrices.*

This problem arises in three instances: if we decide to use empirical multi-step transition matrices, as supplied by e.g. Standard & Poor's [2000], we can calculate the forward transition probability matrix from time t to time $t + \triangle$ as follows[8]:

$$\hat{A}(t, t+\triangle) = \hat{A}(0,t)^{-1} \cdot \hat{A}(0, t+\triangle) \qquad (3.21)$$

Secondly, when using forward risk premiums this kind of calculation is performed to obtain the one-step default probabilities, see e.g. (3.11). Finally, when calculating cumulative risk premiums these kind of calculations are needed in order to obtain the forward transition probability matrices, which are needed for the pricing of certain credit derivatives.

Does (3.21) always yield a proper transition probability matrix? For transition matrices on state spaces with more than two states the answer is unfortunately no. It is easy to verify that the product of two square matrices A and B with row sums equal to one always is again a matrix with row sums equal to one. From (3.16) we know that any inverse transition probability matrix also has rows which sum to one.

---

[7] This is a suggestion that Wei [2000] makes in his article. In this case the size of the adjustment will be in line with the significance of the other numbers in the matrix.

[8] We use a hat in order to show that these matrices are estimated from data.

Unfortunately the elements of the resulting matrix in (3.21) are not necessarily constrained to [0,1].
One can check that this problem definitely arises in the case where we use empirical multi-step transition matrices. Take for example the one-year transition matrix and the two-year transition matrix as supplied in Standard & Poor's [2000]. If we calculate the implied forward transition probability matrix from year 1 to year 2, we do not obtain a proper transition probability matrix.
In the following example we will demonstrate that this problem also can occur in a situation that is not uncommon in practice: when yield curves for two ratings cross.

Example
We take the state space $S = \{1,2,3\}$, and choose $\triangle$ to be equal to one. From the data on zero-coupon bonds we have determined the one-step transition probability matrix from $t = 0$ to $t = 1$, $\tilde{A}(0,1)$, and the two-step default probabilities $\tilde{a}_{\bullet 3}(0,2)$. We would now like to calculate the forward default probabilities from $t = 1$ to $t = 2$. Using (3.21) it is not hard to find the following analytical expression:

$$\tilde{a}_{\bullet 3}(1,2) = \tilde{A}(0,1)^{-1} \cdot \tilde{a}_{\bullet 3}(0,2) = \frac{1}{\det(\tilde{A}(0,1))} \cdot \begin{pmatrix} \tilde{a}_{22}(0,1) \cdot \triangle_{13}(1,2) - \tilde{a}_{12}(0,1) \cdot \triangle_{23}(1,2) \\ \tilde{a}_{11}(0,1) \cdot \triangle_{23}(1,2) - \tilde{a}_{21}(0,1) \cdot \triangle_{13}(1,2) \\ \det(\tilde{A}(0,1)) \end{pmatrix}$$

where $\det(\tilde{A}(0,1)) = \tilde{a}_{11}(0,1) \cdot \tilde{a}_{22}(0,1) - \tilde{a}_{12}(0,1) \cdot \tilde{a}_{21}(0,1)$ and for simplicity we introduce the following notation: $\triangle_{ij}(s,t) = \tilde{a}_{ij}(0,t) - \tilde{a}_{ij}(0,s)$. The diagonal entries of the transition probability matrix will usually be the largest entries in the matrix; the calculated determinant will therefore typically be positive. If this is the case, it is rather straightforward to derive conditions under which the resulting forward default probabilities are not probabilities, but are, say, negative. The condition under which e.g. $\tilde{a}_{23}(1,2)$ is negative, is the following:

$$\tilde{a}_{11}(0,1) \cdot \triangle_{23}(1,2) < \tilde{a}_{21}(0,1) \cdot \triangle_{13}(1,2) \tag{3.22}$$

As we stated before, we will usually have $\tilde{a}_{11}(0,1) \gg \tilde{a}_{21}(0,1)$. It seems that (3.22) will not happen very often. However, suppose that we have a situation where the yield curves for rating 1 and 2 cross[9], i.e. $\overline{P}^2(0,2) > \overline{P}^1(0,2)$, which from the JLT pricing equation immediately implies that $\tilde{a}_{23}(0,2) < \tilde{a}_{13}(0,2)$. Suppose that this did not happen for maturity 1, so that $\overline{P}^2(0,1) < \overline{P}^1(0,1)$ and thus $\tilde{a}_{23}(0,1) > \tilde{a}_{13}(0,1)$. If these effects are large enough to offset the fact that $\tilde{a}_{11}(0,1) \gg \tilde{a}_{21}(0,1)$, we will obviously get a negative implied forward default probability. That this is not a freak occurrence will be demonstrated in the following numerical example. Suppose we are given the following zero-coupon yields for all rating classes:

[9] This may very well happen in practice. For instance, a lower rating class could be more liquid than a higher one. We encountered this for several rating classes in our dataset.

Maturity	Default-free	Rating 1	Rating 2
1	5	6	6.5
2	5.5	6.5	6

***Table 3.1**: Zero-coupon yields (in %) for the various rating classes*

Using a recovery rate of 50%, we can easily check that these zero-coupon yields will lead to defaultable zero-coupon bond prices satisfying the aforementioned conditions. Furthermore suppose that $\tilde{a}_{11}(0,1) = 0.85$ and that $\tilde{a}_{22}(0,1) = 0.8$. Using all this information and the JLT pricing equation yields the following transition probabilities:

$$\tilde{A}(0,1) = \begin{pmatrix} 0.85 & 0.14 & 0.01 \\ 0.18 & 0.8 & 0.02 \\ 0 & 0 & 1 \end{pmatrix} \qquad \tilde{a}_{\bullet K}(0,2) = \begin{pmatrix} 0.04 \\ 0.02 \\ 1 \end{pmatrix}$$

From (3.21) it follows that $\tilde{a}_{\bullet K}(1,2) = (0.03, -0.01, 1)^T$. □

There is no immediate solution to this problem. Since for rating-dependent products we require proper forward transition probabilities, it seems wise not to use cumulative risk premiums. As we just demonstrated, the same problem occurs with forward risk premiums. It therefore seems wise to use constrained optimisation in order to determine the risk premiums. Since for every step we have to estimate K - 1 risk premiums, we might have to impose extra structure, e.g. that the risk premiums are constant over a longer period of time.

3.4. Alternative risk premium assumptions

Some other risk premium assumptions that have been suggested in the literature have been based on the generator matrix of the underlying Markov chain. For more information about generator matrices we refer you to appendix E – Generator matrices. Here we can suffice with the information that the generator matrix Λ for a transition probability matrix A is a matrix with row-sums equal to 0 and non-negative off-diagonal entries, such that $\exp(\Lambda) = A$. The exponent of a matrix is defined as the following Taylor series expansion:

$$\exp(t\Lambda) = \lim_{n\to\infty} \sum\nolimits_{k=0}^{n} (t\Lambda)^k / k! \tag{3.23}$$

where $\Lambda^0 = I$, the identity matrix. Various powers of A can be determined easily by using various values of t in (3.23).

JLT consider a particular way of modifying the generator matrix by risk premiums for the continuous-time version of their model. It suffices to say that the approach is very similar to the column-independent risk premiums, and will also be affected by similar kinds of numerical problems. Lando [1998] pursues another approach. He suggests to modify the eigenvalues of the transition probability matrix with risk premiums. Again, this method is faced with numerical problems.

3.5. Completeness in the JLT model

As a final topic we discuss the completeness of the discrete-time version of the JLT model. Here it does not matter whether we use the assumptions as in the JLT model or the relaxed JLT framework which we introduced in paragraph 3.1. As in JLT, we assume that for any specific firm bonds of all maturities are traded.

To show that the JLT model is indeed complete, first consider the forward rate for zero-coupon bonds. In discrete time these are defined as follows:

$$f(t,T) \equiv -\triangle^{-1} \ln\left(\frac{P(t,T+\triangle)}{P(t,T)} \right) \qquad f^i(t,T) \equiv -\triangle^{-1} \ln\left(\frac{\overline{P}^i(t,T+\triangle)}{\overline{P}^i(t,T)} \right)$$

Note that for $i = K$ we have $f^K(t,T) = f(t,T)$. Using (3.1), we obtain the following:

$$\begin{aligned} f^i(t,T) &= -\ln\left(\frac{P(t,T+\triangle)\cdot(\delta+(1-\delta)\cdot \mathbb{Q}_t^i(\tau > T+\triangle))}{P(t,T)\cdot(\delta+(1-\delta)\cdot \mathbb{Q}_t^i(\tau > T))} \right) \\ &= f(t,T) + \ln\left(\frac{\delta+(1-\delta)\cdot \mathbb{Q}_t^i(\tau > T)}{\delta+(1-\delta)\cdot \mathbb{Q}_t^i(\tau > T+\triangle)} \right) \end{aligned} \tag{3.24}$$

We define the second term in this equation as $s^i(t,T)$, the forward spread for rating i. Suppose the current rating of the defaultable bond under consideration is $X_t = i$ and that the rating at time $t + \triangle$ equals j. The change in the forward rate of this bond over $[t, t+\triangle]$ can be expressed as:

$$f^j(t+\triangle,T) - f^i(t,T) = \left(f(t+\triangle,T) - f(t,T)\right) + \left(s^j(t+\triangle,T) - s^i(t,T)\right)$$

The first component of the change is caused by shifts in the default-free forward rate and the shortening of the time to maturity. If we assume that the market for default-free debt is complete, this part can be hedged. We will now focus on the change in the spread, which is caused by the shortening of the time to maturity and a possible change in the credit rating. From the definition of the spread in (3.24), we see that:

$$s^j(t+\triangle,T) - s^i(t,T) = \ln\left(\frac{\delta+(1-\delta)\cdot \mathbb{Q}_{t+\triangle}^j(\tau > T)}{\delta+(1-\delta)\cdot \mathbb{Q}_{t+\triangle}^j(\tau > T+\triangle)} \right) - s^i(t,T) \tag{3.25}$$

Since the current spread is known at time t, the only unknowns seem to be the two non-default probabilities $\mathbb{Q}_{t+\triangle}^j(\tau > T)$ and $\mathbb{Q}_{t+\triangle}^j(\tau > T+\triangle)$. However, it can be shown that these too are known at time t. For any maturity S we obviously have:

$$\mathbb{Q}_{t+\triangle}^j(\tau > S) = 1 - \tilde{a}_{jK}(t+\triangle,S)$$

This last probability is part of the transition matrix $\tilde{A}(t+\triangle,S)$, which satisfies:

$$\begin{aligned} \tilde{A}(t,S) &= \tilde{A}(t,t+\triangle)\cdot \tilde{A}(t+\triangle,S) \\ \tilde{A}(t+\triangle,S) &= \tilde{A}^{-1}(t,t+\triangle)\cdot \tilde{A}(t,S) \end{aligned}$$

provided the one-step matrix is invertible. In the right-hand side of the second equation we have two matrices, which are both known at time t. Hence, the aforementioned non-default probability is also known at time t. It follows directly that the change in the spread term in equation (3.25) is a constant, provided that the future rating of the bond is known. The change in the forward spread therefore has at most K different outcomes. How can we hedge this change in the forward spread? Remember from before that we assumed that the market for default-free debt is complete. In conjunction with the earlier assumption that prices are only allowed to change on discrete dates, this implies that we have a finite number of possible outcomes for the price of a default-free bond. To see why this is important, we turn to the price of a defaultable zero-coupon bond.

Assuming that $T = t + n\triangle$, and using the definition of the forward rate, we have:

$$\begin{aligned}\overline{P}^i(t,t+n\triangle) &= \exp\left(-\sum_{m=0}^{n-1} f^i(t,t+m\triangle)\right)\\ &= P(t,t+n\triangle)\cdot\exp\left(-\sum_{m=0}^{n-1} s^i(t,t+m\triangle)\right)\end{aligned} \tag{3.26}$$

Now, let us try and hedge the change in the forward spread with a portfolio consisting of default-free and defaultable bonds (all of the same firm). Conditional upon the rating of the firm at time $t + \triangle$, at time $t + \triangle$ we obtain a portfolio which is equivalent to a portfolio of default-free bonds, as can be seen from (3.26). If each default-free bond has a finite number of possible outcomes, and default-free and defaultable bonds of all maturities are traded, we can hedge the change in the spread term. We will demonstrate this in the following simple example.

Example

Let us assume that $K = 2$. Furthermore, we assume a flat forward curve at all times and for all maturities. We also assume that the rating transition process under $\mathbb{Q}$ is governed by a time-homogenous Markov chain, so that we have a constant spread for each rating and time to maturity. In other words:

$$s^i(t,T) = s^i(T-t)$$

The prices of default-free zero coupon bonds at time $t + \triangle$ are fully determined by one factor, which can take on two values: u or d (up or down state). We will set up a portfolio V of default-free and defaultable bonds at time t, which will satisfy the following at time $t + \triangle$:

$$V(t+\triangle) = s^{X_{t+\triangle}}(t+\triangle,T) \quad \text{a.s.}$$

Here $X_{t+\triangle}$ equals the rating of the firm under consideration at time $t + \triangle$. The portfolio will equal the forward spread from time $t + \triangle$ to T. This is equivalent to hedging the forward spread, which we considered earlier. If we assume that the firm under consideration has rating 1 at time t, we have the following possible outcomes:

Default-free state	*Rating of firm*	*Forward spread of firm*
u	*1*	$s^1(t+\triangle,T)$
u	*2*	*0*
d	*1*	$s^1(t+\triangle,T)$
d	*2*	*0*

As the spread is constant, we could set up our replicating portfolio with only two instruments, say the default-free bond maturing at time $t + \triangle$ and a defaultable zero-coupon bond of that firm, also maturing at time $t + \triangle$. The first always pays 1, the second pays 1 if its rating is unequal to 2 (the default state), and δ if it has defaulted. The following portfolio does the trick:

$$V(t) = -\frac{\delta\, s^1(t+\triangle,T)}{1-\delta} \cdot P(t,t+\triangle) + \frac{s^1(t+\triangle,T)}{1-\delta} \cdot \overline{P}^1(t,t+\triangle)$$

Obviously if $\delta < 1$ this is a valid portfolio, which has the correct payoff at time $t+1$. □

We have shown that we can hedge the change in the forward spread, using default-free and defaultable securities. This in turn implies that the market for defaultable debt is complete in the JLT model.

3.6. Conclusions

In this chapter we reviewed a broad class of rating-based models based on the original model by Jarrow, Lando and Turnbull. These models use column-independent risk premiums to determine the transition probabilities under the equivalent pricing measure $\mathbb{Q}$ from the real-world probability measure $\mathbb{P}$. As we have shown, the choice of column-independence is rather arbitrary and has serious repercussions for the prices of certain credit derivatives. Apart from this matter, using column-independent risk premiums can cause several other problems. Below these problems are briefly mentioned, and we summarise which action has to be taken if any of these problems are encountered[10].

1. *The resulting adjusted transition matrix is not a transition probability matrix.*

We can work around this problem, if we sacrifice the property that the initial term structure of zero-coupon bonds is to be fitted perfectly. We specified lower and upper bounds for the risk premiums. If the upper bound is violated, the risk premium could be put equal to its upper bound. If it is the lower bound which is violated, we can replace the risk premium by a small positive number. Note that when using forward risk premiums, a wrongly chosen risk premium will affect the choice of the risk premium for all future steps.

2. *Equivalence of $\mathbb{Q}$ and $\mathbb{P}$ is not guaranteed.*

This problem tends to arise when entries in the default column of the empirical transition matrix are equal to zero. If we replace these by a small number, e.g. the smallest non-zero number in the transition matrix, this problem is 'solved'.

[10] Of course, there is also the option of not using this kind of risk premiums.

3. *One-step transition matrices calculated from cumulative transition probability matrices do not necessarily have to be transition probability matrices.*

This is a rather unfortunate problem, which at the moment has no clear solution. It could occur when we calculate forward transition probabilities from:

- Empirical multi-year transition matrices;
- Transition probability matrices under $\mathbb{Q}$, when using cumulative risk premiums;
- Transition probability matrices under $\mathbb{Q}$, when using forward risk premiums;

In the last problem an additional problem is that the bounds for the risk premium will be violated, which brings us back to problem 1. We demonstrated that this last problem could very well occur in practice, for example when the yield curves for two ratings cross.

Given all these problems, and the obvious arbitrariness of the assumption of column-independence, it seems that other methods of obtaining the transition probabilities under the pricing measure may be more appropriate. One of these methods, a constrained optimisation model proposed by e.g. Arvanitis et al. [1999] will be discussed and implemented in the next chapter.

4. A time-homogenous model under $\mathbb{Q}$

As discussed in the previous chapter, the arbitrariness of the column-independence assumption and various numerical problems which are associated with these risk premiums call for a different approach to inferring the transition probabilities under the equivalent pricing measure $\mathbb{Q}$. In this chapter we first review and then implement a model, based in the JLT framework, which assumes that the credit migration process under $\mathbb{Q}$ is time-homogenous. The transition matrix under the pricing measure will be inferred from the market prices of corporate bonds. The used data will be described in the second paragraph.

The goal of this thesis is to price rating-dependent credit derivatives, in particular the corporate bonds with step-up features which have recently been issued by several telecommunications companies. However, in the articles we have read on the rating transition approach to pricing credit derivatives, the ability of the models to price straight bonds has never been empirically investigated. Most articles just focus on the theory underlying the pricing model; some articles use a bond index for each of the various rating classes, and calibrate their pricing model to this index. Since straight bonds are basically the simplest class of credit derivatives, it is interesting to check how well our simple pricing models can replicate the market prices of these bonds. This will be accomplished in the third and fourth paragraphs. We will conclude with a description of several identification problems that are encountered when using rating-based models.

4.1. Description of the model

In the previous chapter we suggested that we could use constrained optimisation in order to avoid the problems associated with column-independent risk premiums. In the remainder of this paragraph we will discuss how this can be done in the particular case where we have a time-homogenous model under $\mathbb{Q}$. The model can easily be modified to allow for time-varying risk premiums. The approach we follow is inspired by one of the models considered in Arvanitis, Gregory and Laurent [1999]. The advantages of a constrained optimisation approach in this setting are:

- An arbitrary risk premium assumption is not required;
- The assumption made with the column-independent risk premiums, that there is exactly one bond price per maturity and rating class, will not be satisfied. Therefore this model will have to be estimated on a previously estimated term structure. In our model the term structure can directly be stripped during the estimation process, which is more efficient;
- Any recovery rate mechanism can be incorporated.

With respect to the last remark we refer the reader to appendix D, which contains the prices of coupon bonds under the most common recovery rate assumptions in the literature. We will use each of these recovery rate assumptions when implementing the model later in this chapter.

As stated, we will consider a simple model where under $\mathbb{Q}$ the rating transition process follows a time-homogenous Markov chain. We therefore have:

$$\tilde{A}(t, t+\triangle) = \tilde{A} \qquad \text{for all } t \text{ in } [0, T^*].$$

We will start with calibrating our model to market prices of bonds. Suppose we have M_i bonds in credit class i, and that we denote the current market and model prices of bond m with rating i by $\overline{V}^{i}_{m,\text{model}}(\tilde{A})$ and $\overline{V}^{i}_{m,\text{market}}$, respectively. We explicitly state the dependence of the model price on the transition matrix. Having chosen for a constrained optimisation approach, we decide to determine $\tilde{A}$ via minimising:

$$\begin{aligned}
&\min_{\tilde{A}} \sum_{i=1}^{K} \sum_{m=1}^{M_i} w_{i,m} \cdot \left(\overline{V}^{i}_{m,\text{model}}(\tilde{A}) - \overline{V}^{i}_{m,\text{market}}\right)^2 \\
&\text{s.t.} \quad 0 \le \tilde{a}_{ij} \le 1 \\
&\qquad \sum_{j=1}^{K} \tilde{a}_{ij} = 1 \\
&\qquad \tilde{a}_{Kj} = a_{Kj}
\end{aligned} \tag{4.1}$$

This model setup is similar to that of Arvanitis et al. We have chosen to minimise the (weighted) sum of squared errors due to the statistical properties of non-linear least squares estimation, which can be found in appendix C. The reason for minimising price errors rather than yield to maturity (YTM) errors is that the calculations necessary to minimise YTM errors are significantly more time consuming in a constrained optimisation framework. It is nevertheless important that the model be capable of fitting the yield curves of the bonds used in the calibration process. To this extent, we must choose the weights with which each price error is weighted carefully.
First consider the case where the majority of the bonds belong to one rating. The optimisation routine will mostly focus on these bonds. In order to ensure that the ratings for which there are fewer bonds available are equally important in the goal function, each error is weighted by the inverse of the number of bonds in that class.
Secondly, consider the fact that long-maturity bonds are much more sensitive to changes in the interest rate than short-maturity bonds. Put differently, a large price change for a long-maturity bond may lead to an identical change in yield when compared to a much smaller change in price in a short-maturity bond. The optimisation technique that seeks to minimise price errors will try to overfit the long-maturity bond prices at the expense of the short-term prices, which in turn will lead to the overfitting of long-term yields at the expense of the short-term prices. In order to correct for this problem, each squared error is additionally weighted by the inverse of its duration. We will define duration in terms of YTM. The YTM y of a coupon bond with coupon dates t_1, …, $t_N = T$, coupon C and face value F is implicitly defined as:

$$\overline{V}(t,T) = \sum_{i=1}^{N} C \cdot e^{-y\cdot(t_i - t)^+} + F \cdot e^{-y\cdot(T-t)^+}$$

The duration of a bond D is now equal to:

$$D = -\frac{\partial \ln \overline{V}}{\partial y} = -\frac{1}{\overline{V}} \cdot \frac{\partial \overline{V}}{\partial y} = \sum_{i=1}^{N} (t_i - t)^+ \cdot \frac{C \cdot e^{-y\cdot(t_i - t)^+}}{\overline{V}} + (T-t)^+ \cdot \frac{F \cdot e^{-y\cdot(T-t)^+}}{\overline{V}}$$

We see that the duration of a bond is a weighted average of the times to maturity on each outstanding cashflow, where the weights are equal to the share of the outstanding cashflow in the total amount of outstanding cashflows. The definition of duration used here is equivalent to Macaulay's duration, except for the fact that we use continuously compounded yields instead of discrete compounded yields.

Apart from the transition matrix we could include another parameter in our model, the average recovery rate δ, which is constant through time. We will investigate two approaches:

- setting it equal to the average recovery rate as quoted by S&P's: 51.14%;
- inferring this parameter from our market prices by including it as an extra parameter, of course adding the condition that $0 \leq \delta \leq 1$.

The latter method may yield better results due to its inclusion of an extra parameter. Furthermore, it is a well-known fact that recoveries are hard to measure. In fact, the recovery rate in the JLT model is not strictly the same as the recoveries measured by e.g. Moody's and S&P's, see paragraph 2.3.1. Therefore it may be more sensible to infer the recovery rate from the data.

Estimation of the model

We have to discuss how to estimate (4.1). This is a non-linear least squares problem, subject to restrictions on the parameters. We could use a Lagrange multiplier approach, but this would increase the dimensionality of the optimisation problem quite considerably; here we have roughly as many parameters as restrictions. In this instance it is actually quite easy to transform (4.1) into an unconstrained problem, and estimate our parameters consistently using a non-linear least squares estimation technique. For these techniques and the statistical properties of non-linear least squares estimation we refer the reader to appendix C.

One way to get rid of a number of these restrictions is to use the underlying generator matrix instead of the one-step transition probability matrix. Generator matrices are described in appendix E and were first introduced in paragraph 3.4. The reason why we use a generator matrix here, is that we can:

a) get rid of the restriction that the row-sums of the transition matrix must equal one;
b) incorporate practical restrictions on the generator matrix.

To explain a), we look more closely at the transition matrix. Each probability in this matrix must lie in the interval [0,1], and furthermore each row-sum must equal one. Suppose we have a row containing n elements. We can set an element, say n, equal to one minus the sum of all other entries in this row, which ensures that the row-sum equals one. Furthermore we must impose constraints on the other n-1 entries to make sure that they are contained in [0,1]. We are still not sure that element n lies in [0,1], so that we must add additional constraints.

In a generator matrix, the row-sum must equal zero. Let us now consider the constraints for one row in a generator matrix. The diagonal element (which is smaller than or equal to zero) can be set equal to the negative sum of all other entries in that row. Furthermore we must impose constraints on the off-diagonal entries to ensure that they are larger than or equal to zero. We see that the diagonal element is now completely eliminated as a variable: we do not have to add additional constraints.

As for b), in paragraph 2.2.2 we examined all rating actions that S&P's took during the course of 2000. Using unmodified ratings, it turned out that for all ratings instantaneous upgrades of more than one rating are quite rare. For investment grade bonds, instantaneous downgrades of more than one rating, and for speculative grade bonds downgrades of more than two ratings, do not occur very often. We could decide to exclude the possibility of such large rating changes. We can conveniently incorporate these restrictions in our generator matrix, since this implies that the corresponding intensities are equal to zero. If we assume the aforementioned structure and use a rating system of 7 ratings, this would mean that we are only left with 12 parameters (as opposed to 36 parameters) in our transition matrix, which have to be estimated.

Transforming (4.1) into the form where we estimate the generator Λ of $\tilde{A}$ yields:

$$
\begin{aligned}
&\min_{\Lambda} \sum_{i=1}^{K} \sum_{m=1}^{M_i} w_{i,m} \left(\bar{V}_{m,\text{model}}^{i}(\Lambda) - \bar{V}_{m,\text{market}}^{i} \right)^2 \\
&\text{s.t.} \;\; \forall_{i \neq j} \; \lambda_{ij} \geq 0 \\
&\qquad \lambda_{ii} = -\sum_{j \neq i} \lambda_{ij} \\
&\qquad \lambda_{Kj} = 0
\end{aligned} \tag{4.2}
$$

The latter two restrictions can be directly substituted into the model, which leaves us with $(K\text{-}1)^2$ parameters to be estimated.

Now consider the following two types of restrictions:

- $b \geq 0$ (here imposed on the off-diagonal elements of the generator matrix);
- $0 \leq b \leq 1$ (here imposed on the recovery rate δ, if we include it in the parameters that have to be inferred from the market prices).

We can reparameterise e.g. $b = \exp(\beta)$ for the first type and $b = \frac{1}{1+\exp(\beta)}$ for the second type to get rid of these restrictions. The new parameter β is now unconstrained.
We have now transformed (4.1) into an unrestricted non-linear least squares problem, for which estimation techniques are readily available. We only require a starting solution. It would seem natural to take the empirical generator matrix[11] and the empirical average recovery rate as starting values for our parameters. We can also generate random starting values for our parameters.

4.2. Description of the used data

All data sets were obtained from the bond database maintained by *Rabobank International's Central Market Risk* department. This database contains data from *Reuters*, an international news and information organisation.
For the default-free zero-coupon prices we used stripped zero-coupon bond prices for Dutch treasury bonds, obtained from the *Central Market Risk* database. To obtain zero-coupon bond prices, *Central Market Risk* models the default-free curve as a B-spline, see Houweling, Hoek and Kleibergen [2001] or our appendix on term structure estimation for details. The remainder of our data set was selected on basis of the following criteria:

- Euro-denominated, straight bullet coupon bonds or mid-term notes;
- Senior secured issues;
- Each bond has to have at least 200 quotes available per year, on average; we require this in order to ensure that the bonds we examine are relatively liquid;
- Only rated bonds are considered;

[11] For numerical problems associated with this we refer the reader to appendix E - Generator matrices.

- National governments with an AAA-rating are excluded; these are usually regarded as default-free.

The estimation date was chosen to be equal to the week starting on 31/07/2000 and ending on 04/08/2000[12]. Quotes for all bonds were available in that week, we used the most recent midquote for each bond (i.e. bid and ask quote divided by two). In future, when we refer to the full data set, we will speak of the BOND data set. We present some summary statistics for this data set:

BOND *Composite rating*	*AAA*	*AA*	*A*	*BBB*	*BB*	*B*	*Total*
No. of bonds	*465*	*380*	*171*	*43*	*37*	*28*	*1124*
Average coupon rate (annual, %)	*5.9%*	*5.8%*	*5.9%*	*6.3%*	*9.0%*	*8.8%*	*6.1%*
Time to maturity (years)							
Average	*4.6*	*4.3*	*4.7*	*4.4*	*4.9*	*4.0*	*4.5*
Maximum	*27.6*	*31.5*	*27.9*	*9.6*	*16.8*	*16.9*	*31.5*
Relative bid-ask spread							
Average	*0.3%*	*0.4%*	*0.4%*	*0.6%*	*0.9%*	*0.9%*	*0.4%*
Standard deviation	*0.3%*	*0.3%*	*0.2%*	*0.4%*	*0.4%*	*0.4%*	*0.3%*

Table 4.1: *Summary statistics for BOND on 04/08/2000*

The relative bid-ask spread is here defined as the bid-ask spread as a percentage of the midquote[13].

4.3. Term structure estimation with traditional methods

Before estimating the model described in paragraph 4.1 on the BOND data set, we estimated the term structure using the B-spline and the Svensson model described in Appendix B - Term structure estimation. Both models will function as a benchmark for the performance of the credit migration models.

Which statistics will we report for the estimated models? Firstly, we will calculate an R^2-like statistic to assess the goodness of fit of the model. If we denote the average market price by $\overline{V}_{\text{market}}$, the R^2-statistic is here equal to:

$$R^2 = 1 - \frac{\sum_{i=1}^{K-1} \sum_{m=1}^{N_i} w_{i,m} \left(\overline{V}^{\,i}_{m,\text{model}} - \overline{V}^{\,i}_{m,\text{market}} \right)^2}{\sum_{i=1}^{K-1} \sum_{m=1}^{N_i} w_{i,m} \left(\overline{V}^{\,i}_{m,\text{market}} - \overline{V}_{\text{market}} \right)^2} \tag{4.3}$$

Unfortunately for estimation techniques other than ordinary least squares this statistic is no longer guaranteed to be larger than 0. This is not a real problem as larger R^2 values still correspond to a better fit. Also, for the instances we consider here, the variation between the bond prices is quite large, so that the denominator in (4.3) is quite large in comparison with the numerator. We should therefore not be misled by R^2 values close to 1: this definitely does not mean we are close to a perfect fit! Due to the fact that the various R^2 values of the models will lie close to each other, we will also report the sum of squared errors (SSE) that we minimise, the numerator in (4.3).

[12] Her Majesty Queen Elizabeth The Queen Mother's 100th birthday.

[13] Obviously, this can only be calculated for those issues that had bid and ask quotes available.

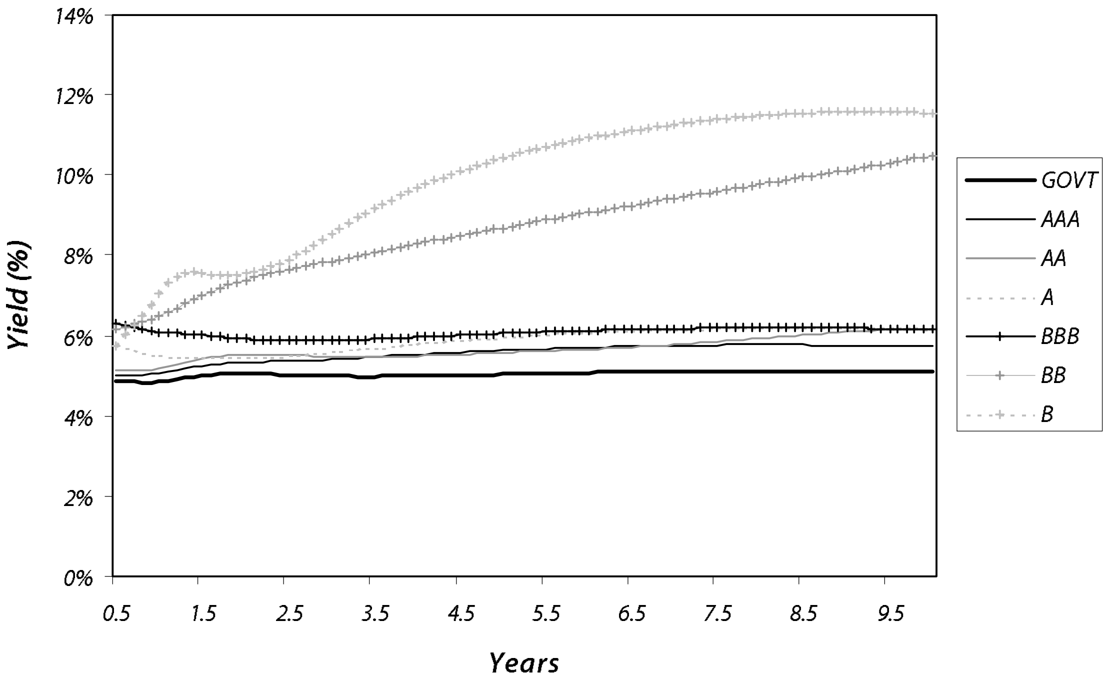
The second statistic we calculate is the absolute relative mispricing. Given model and market prices for bond i ($\overline{V}_{i,\text{model}}$ and $\overline{V}_{i,\text{market}}$ respectively) we define this as:

$$\frac{|\overline{V}_{i,\text{model}} - \overline{V}_{i,\text{market}}|}{\overline{V}_{i,\text{market}}}$$

This statistic is sensitive to large outliers, so that we will also report its median. Although our model does not minimise yield errors, we will also report the average difference between the yield to maturity of the model price and that of the market price, as well as their absolute average difference. This should give a more intuitive feeling about the performance of the model than the earlier mentioned R^2 statistic.
Finally, when we compare two yield curves to each other, it will be useful to have a quantitative measure to be able to compare the difference between both yield curves. Given yield curves R_1 and R_2, a step size $\triangle$ and an interval $[t, t + n\triangle]$ we will calculate:

$$\tfrac{1}{n+1}\sum_{i=0}^{n}\left|R_1(t,t+i\triangle) - R_2(t,t+i\triangle)\right|\cdot 10000$$

This will give us the average absolute difference of both yield curves in basis points[14].
In the B-spline model we estimated the government curve with a 3rd degree polynomial on each subinterval. The spread curves superimposed on the government curve were chosen to be a 2nd degree polynomial. The knots were carefully chosen. In the graph below the estimated B-spline yield curves are shown for maturities ranging from 0.5 to 10 years.



Graph 4.1: *B-spline yield curves*

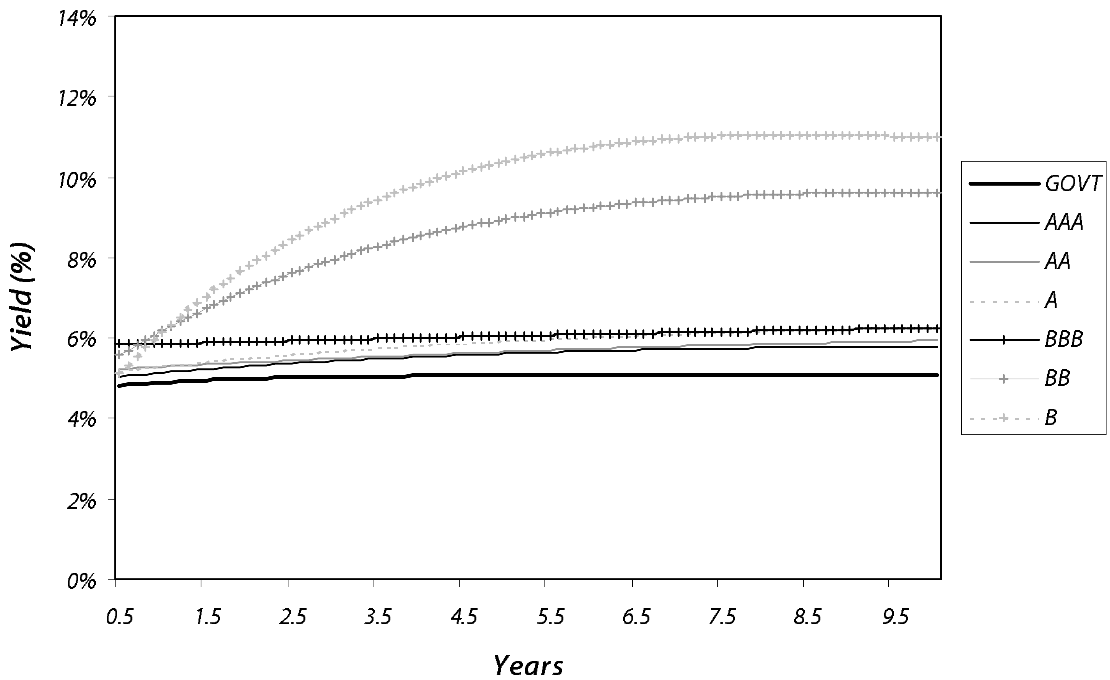
[14] One basis point (in future bp) is equal to one one-hundredth of a percent.

The calculated statistics for this model are presented in the following table:

Statistic	AAA	AA	A	BBB	BB	B	Overall
Absolute relative mispricing							
Average	0.83%	0.99%	0.88%	1.33%	3.83%	4.64%	1.10%
Median	0.20%	0.21%	0.41%	0.79%	2.39%	2.58%	0.28%
Standard deviation	2.79%	3.69%	1.77%	1.82%	3.88%	5.42%	3.20%
YTM difference (in bp)							
Average	-2.10	-0.28	-0.73	-0.65	-4.91	-8.70	-1.48
Absolute average	16.15	20.30	25.57	45.24	105.04	128.00	25.81
Sum of squared errors	22.52		R^2-statistic		0.9962		

Table 4.2: *Performance of the B-spline model*

The performance for the higher ratings is not too bad, although the average mispricing is still larger than the average bid-ask spread. This means that the estimated price of the bond will typically not lie within the bid-ask spread.
The estimated Svensson curves are shown at the bottom of this page. Comparing both graphs visually we can see that there are some differences between both estimated yield curves. Inferring the true term structure from market prices is a notoriously difficult problem. One thing we also observe in both graphs is that we obtain yield curves which at times cross each other, see e.g. the yield curve for the AA-rating in the B-spline graph. At times the AA-yield is lower than the AAA-yield. We explained in the previous chapter that this can happen in practice. It does however pose serious problems when calculating column-independent risk premiums, which is why we do not attempt to calculate these in this chapter.
Finally, we observe that the short-term behaviour of both yield curves is also questionable. A reason for this is that we have relatively few bonds in the short end of the maturity range. Houweling et al. [2001] suggest that the estimated yield curve can be steered in a certain direction by adding synthetic bonds, using e.g. fair market



Graph 4.2: *Svensson yield curves*

yields for the various ratings, as published by Bloomberg. At this moment in time we are however interested in the overall performance of the various techniques. Since we have few bonds in the short end of the maturity range, we expect that this will not affect our results, except for the average difference between the yield curves. We will account for this by excluding the short end of the curves from the statistic.
The calculated statistics for the Svensson model are:

Statistic	*AAA*	*AA*	*A*	*BBB*	*BB*	*B*	*Overall*
Absolute relative mispricing							
Average	0.77%	0.75%	0.84%	1.32%	4.29%	4.84%	1.01%
Median	0.18%	0.18%	0.41%	0.72%	2.50%	2.88%	0.24%
Standard deviation	2.58%	2.34%	1.77%	1.82%	4.04%	5.37%	2.68%
YTM difference (in bp)							
Average	5.81	2.12	-3.48	-1.45	-19.23	-5.28	1.77
Absolute average	17.80	19.68	22.87	43.53	117.97	134.09	26.39
Sum of squared errors	23.52		R^2-statistic		0.9960		

Table 4.3: *Performance of the Svensson model*

Finally, we calculate the average difference between the B-Spline and the Svensson yield curves. To avoid the aforementioned problems with the short end of the yield curves, we choose the interval over which we measure the statistic to include maturities between 1.5 and 10 years, with a step size of 0.1 years.

	AAA	*AA*	*A*	*BBB*	*BB*	*B*	*Overall*
Average difference (bp)	2	9	5	4	26	32	13

Table 4.4: *Average difference between the B-spline and Svensson yield curves*

4.4. Term structure estimation with a credit migration model

From the previous paragraph we now have a slight idea of what the 'true' term structure looks like, if there is such a thing to begin with. The Svensson and B-spline models can also act as a benchmark for the model we consider in this paragraph. We will now estimate the time-homogenous model we introduced in the first paragraph of this chapter.
We implemented the model with 7 unmodified ratings (AAA through B), as we only had data of 4 CCC-rated bond issues. Additional to the formulation in (4.2), we imposed the following restrictions:

- For numerical stability, off-diagonal entries of the generator matrix are bounded from above by 5[15];
- We incorporated the restrictions mentioned in 4.1, following from the discussion in paragraph 2.2.2: for investment-grade bonds, instantaneous up- and downgrades of more than one rating class are impossible. For speculative grade, the same holds, with the exception that instantaneous downgrades of two rating classes are possible.

[15] In the current model this would imply that a lower bound for the average stay in any rating is either 36 days (investment grade) or 24 days (speculative grade). According to historical transition matrices the average stay in any credit rating under the real-world probability measure equals at least 8 years (investment grade) or 4.5 years (speculative grade).

With this formulation we have 12 parameters to be estimated in the generator matrix and one extra parameter if we allow the recovery rate to be variable. We built a Visual Basic program that, given a set of bond issues, estimates our model with the Davidon-Fletcher-Powell Quasi-Newton algorithm. We opted for a Quasi-Newton algorithm in order to avoid costly calculations of the Hessian. For more details on the estimation algorithm we refer you to appendix C. The algorithm was terminated when the square of the L_2-norm of the gradient was smaller than 0.05.
In order to be more certain that an optimisation algorithm has arrived at a global minimum (as opposed to a local minimum), it is wise to use several sets of initial values for the parameters. We generated 9 random sets of parameter values. One additional set of parameter values represents the empirical transition matrix as published in Standard & Poors [2001a]. If the recovery rate was not included as a parameter, we set it equal to the average value reported in that report.

As for the recovery rate assumptions, we investigated the following:
1. Recovery of treasury assumption
2. Recovery of face value, payoff at maturity
3. Recovery of face value, payoff at default
4. Legal claim approach (face value + accrued interest)

For the prices of straight zero-coupon and coupon bonds under these recovery rate assumptions, we refer you to appendix D. These recovery rate assumptions were discussed to some extent in paragraph 2.3.
The first topic of investigation was whether the additional weighting of each price error with the inverse duration of each bond had any impact on the performance or the speed of the estimation. As noted before, we expected that the optimisation algorithm would try to overfit long-maturity bonds, as these are more sensitive to changes in the interest rates. This could slow down the estimation process.
In order to test this, we estimated our problem on the 10 parameters sets with the recovery of face value (payoff at maturity) assumption, where the recovery rate was included as a parameter. On a Pentium III 650 MHz computer, additionally weighting the price errors by the inverse duration of each bond speeded up the estimation process by a factor of ca. 3. In order to compare the performance of the weighted and unweighted estimates, we recalculated the sum of squared errors, now using the additional weighting with the duration of each bond. The weighted sum of squared errors equalled 37.06. It turns out that if we do additionally weight the price errors by the duration of each bond, we end up with a sum of squared errors equal to 37.17. That this is a very small price to pay for such a significant improvement in speed will be clear when we look at the sums of squared errors reported throughout this chapter.
Subsequently, we let our program estimate the parameters, using the four earlier mentioned recovery rate assumptions, letting the recovery rate be either fixed or a parameter, and making use of the 10 sets of starting values for the parameters. This means we end up with 80 sets of results. Per recovery rate assumption and the choice of fixed/variable recovery rate, we picked the best result over the 10 sets. To give an impression: roughly about 75% of the sets gave a value of the goal function that was within 5% of the best of the 10 sets. For the best two combinations (one per choice of fixed/variable recovery rate), we included the performance statistics, as well as the estimated one-year transition matrix and the recovery rate in the appendix to this chapter. In this text we merely depict the weighted sum of squared price errors for each of the 8 combinations (including the rank within its own column):

Recovery rate assumption	Fixed recovery rate	Inferred recovery rate
Recovery of treasury	50.41 (4)	47.62 (4)
Recovery of face value		
Payoff at maturity	45.16 (2)	37.17 (3)
Payoff at default	48.72 (3)	32.77 (1)
Legal claim approach	44.38 (1)	33.24 (2)

Table 4.5: *Weighted sum of squared errors for each of the 8 combinations*

It turns out that the recovery of face value assumption, where the recovery amount is paid at default, yields the lowest sum of squared errors. When the recovery rate is not included as a parameter, the legal claim approach seems to yield the best results. Of course, the inclusion of the recovery rate as an extra parameter improves the performance, as this increases the flexibility of the model.
We will now examine the results of the best model specification more closely. First let us examine the calculated performance statistics:

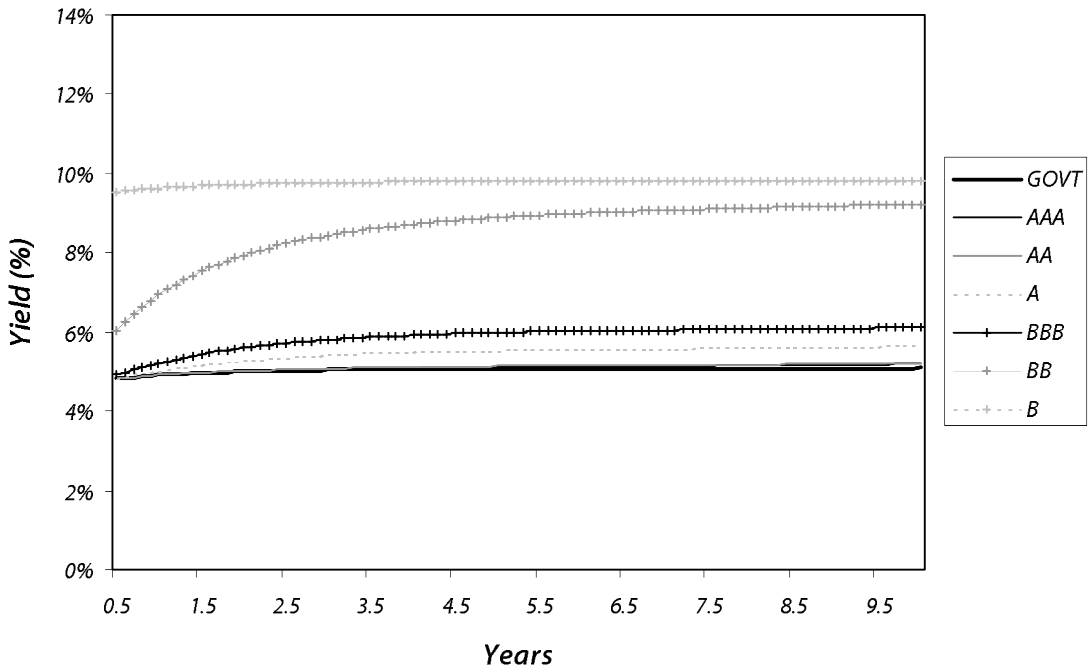
Statistic	AAA	AA	A	BBB	BB	B	Overall
Absolute relative mispricing							
Average	1.31%	1.33%	1.16%	1.64%	4.13%	3.85%	1.46%
Median	0.94%	0.94%	0.45%	1.14%	3.00%	2.59%	0.94%
Standard deviation	2.04%	2.15%	2.12%	1.46%	3.61%	4.39%	2.31%
YTM difference (in bp)							
Average	27.19	31.94	-16.72	-14.1	56.86	67.14	22.51
Absolute average	38.98	47.84	31.4	49.49	136.9	152.04	47.26
Sum of squared errors	32.77		R^2-statistic		0.9945		

Table 4.6: *Recovery of face value (payoff at default), variable recovery rate*

Comparing the results to both the B-spline and the Svensson results out of the previous paragraph, we notice that the average and median mispricings are definitely larger for investment grade bonds, although the standard deviations are lower, except for A-rated bonds. For speculative grade bonds, the performance of our model seems to be better. For example, B-rated bonds have an average mispricing which is approximately 1% lower than with the B-spline and Svensson models. Overall, it can be said that our model has a higher average and median mispricing than the traditional methods, although the standard deviation of the absolute relative mispricing is definitely lower in our model. The ytm differences seem to be larger with our model.
It should not be a big surprise that our model has a worse performance than the traditional methods. The Svensson model we used to estimate the term-structure used a total of 36 parameters, whereas the B-spline model even used 49. The model we implemented in this chapter used 12 (or 13, if the recovery rate was included) parameters. Although the performance is slightly worse, with our model we hope we will be able to price rating-dependent credit derivatives, something that is not possible with either the B-spline or the Svensson curves.
The estimated recovery rate in the appendix to this chapter is reported to be equal to 99.98%. In case of a default, this means we would only lose 0.02% of the face value of the bond. In practice this definitely is not the case, so that something else must be going on. If we look at the average ytm differences[16], we see that, except for A and

[16] Let it be noted that in the table the difference between the model ytm and the market ytm is reported.

BBB-rated bonds, the average difference between the model and the market ytm is positive, indicating that the model prices of the bonds are on average lower than the quoted market prices. The optimisation algorithm will try to fit these bonds better by making the model price larger. One way of accomplishing this, is by raising the recovery rate. This is possibly what has happened here. A recovery rate of 99.98% could indicate that there is a misspecification in our model. We will return to this possibility in a short while.
As the value additivity property, i.e. coupon bonds are a portfolio of zero-coupon bonds, is not valid for this recovery rate assumption, it makes no sense to calculate the zero-coupon yield curves for this model. We would not really be able to compare them to the B-spline or Svensson yield curves. Therefore we depict the estimated yield curves for the best recovery of treasury specification. Under this assumption, the value additivity property is valid.



Graph 4.3: *Yield curves estimated with the recovery of treasury assumption*

Comparing this graph to both graphs 4.1 and 4.2, the most notable differences can be found in the AAA, AA and B-yield curves. The first two are almost equal to the default-free yield curve, which is not the case in graphs 4.1 and 4.2. The B-yield curve is flat in our model, whereas it is upward sloping for the B-spline and Svensson models. The average difference in yields between the traditional methods and the time-homogenous model are depicted below:

	AAA	AA	A	BBB	BB	B	Overall
Average difference (bp)							
Vs. B-spline model	51	59	43	12	52	132	58
Vs. Svensson model	51	56	43	13	37	104	51

Table 4.7: *Average difference in yields with the estimates from the traditional methods*

In appendix C we show how to calculate asymptotic 95% confidence intervals for the estimated parameters. Using the delta method, also treated there, we can also calculate 95% asymptotic confidence intervals for functions of the parameters, e.g. the estimated 1-year default probabilities under $\mathbb{Q}$. In the table below the confidence intervals for the 1-year default probabilities are printed, now using the best performing specification, the recovery of face value assumption (payoff at default), where the recovery rate is included as a parameter:

Rating	*Estimated 1-year default prob.*	*95% confidence interval*
AAA	*0.00*	*(0.00,0.01)*
AA	*0.00*	*(0.00,0.01)*
A	*0.01*	*(0.00,0.02)*
BBB	*0.18*	*(0.15,0.20)*
BB	*0.71*	*(0.67,0.76)*
B	*0.77*	*(0.72,0.82)*

Table 4.8: *Asymptotic 95% confidence intervals for the 1-year default probabilities under* $\mathbb{Q}$

We see that some of the confidence intervals are still quite large. We could use more weeks in the estimation process, thereby compensating for the lack of data, and hopefully improving the accuracy of the estimates. Of course, we must also not include too many weeks of data. By calibrating our model on more weeks of data, we assume that the time-homogenous model is valid over longer periods of time, with the same parameter values. This does not have to be the case.

At any rate, we tried this for the recovery of face value assumption (payoff at maturity)[17], using 8 weeks of data. The accuracy of the estimates was improved, and the performance of the algorithm was comparable to that when we just used one week of data. However, it was still the case that most of the other migration probabilities, which are of interest to us, were not significantly different from zero. There are a number of possible explanations for this:

- The asymptotic approximation is not appropriate here;
- Though we used 8 weeks of data, this still is not enough to be able to say more about the exact location of the parameters;
- The model is misspecified;
- The straight bonds and MTNs on which we calibrated our model only convey information about default probabilities and are relatively insensitive with respect to any other transition probabilities.

The last of these reasons will be the topic of investigation in the next paragraph. We will conclude this paragraph by briefly looking at the third reason. Is the model misspecified? The JLT framework, which we use for our model, assumes that the default-free interest rate, the bankruptcy process and the recovery rate are mutually uncorrelated. If this is not the case, then our model definitely is misspecified. However, as data on defaults and recovery rates are quite scarce and very expensive to obtain, at the moment it does not seem that much can be done about this. Das and Tufano [1996] have extended the original JLT model to include correlation between the recovery rate and the default-free interest rate. As stated, it is however difficult to measure this correlation.

[17] The estimation process in the case of the recovery of face value (payoff at default) specification took several hours for one set of starting values. If we were to include more weeks of data, the estimation time would become much larger.

Another possible misspecification may lie in the fact that we only looked at unmodified ratings. Since the Central Market Risk database contains modified ratings as well, we can investigate whether extending the model to include modified ratings as well would significantly improve the fit. To investigate this, we regressed the pricing errors on a constant and 6 dummy variables. In the equation $\overline{V}_m^{i,j}$ is the price of the m^{th} coupon bond with rating i and modifier j (which equals 1 if the bond has a + as modifier in S&P's terminology, 2 if the bond has no modifier, and 3 if the bond has a – as a modifier). The regression equation then becomes:

$$\overline{V}_{m,\text{model}}^{i,j} - \overline{V}_{m,\text{market}}^{i,j} = \beta_1 + \sum_{\ell=2}^{K-2} \beta_\ell \cdot \mathbb{1}_{[i=\ell]} + \sum_{\ell=1}^{2} \gamma_\ell \cdot \mathbb{1}_{[j=\ell]} + \varepsilon_{ijm}$$

We did not include bonds with an AAA-rating in the equation, as these do not have a modifier. We assumed the errors were mutually independent and normally distributed and estimated the equation with OLS. The results of the regression are as follows:

Variable	*Coefficient*	*Standard deviation*	*P-value*
β_1	0.10	0.57	0.86
β_2	-0.59	0.54	0.28
β_3	1.04	0.57	0.07
β_4	0.94	0.67	0.16
β_5	-0.27	0.70	0.71
γ_1	-0.48	0.28	0.08
γ_2	-0.35	0.27	0.21

Table 4.9: *Regressing the price errors on the modifiers*

At a confidence level of 95%, none of the modifiers seem to add information about the pricing error. We therefore do not expect that adding modified ratings to our model would improve the performance of our model much, and choose not to do this, as it would only complicate the estimation process.
In the next paragraph we will investigate whether the straight bonds and MTNs on which we calibrated the model, carry enough information about non-default probabilities.

4.5. Identifiability of the parameters

As we have seen in the previous paragraph, it turns out that, apart from the default probabilities, we are not really able to say much about the size of the migration probabilities. This could be caused by misspecification of the model, but it could also be the fact that these parameters simply cannot be identified. We will first show the problems that stem from using the recovery of treasury assumption. Secondly, we will show in a simple example that default probabilities are quite insensitive with respect to the other migration probabilities.

4.5.1. Problems with the recovery of treasury assumption

Using the RT assumption implies that every coupon bond is a portfolio of defaultable zero coupon bonds, which are priced according to equation (3.1). Writing this equation slightly differently, we find:

$$\overline{P}(t,T) = P(t,T) \cdot \left(1 - (1-\delta) \cdot \mathbb{Q}_t(\tau \le T)\right)$$

It is evident that only the product of the default probabilities and 1-δ are identifiable. This can be shown with the estimated one-year transition probability matrices. For this table we estimated the recovery of treasury model with recovery rates fixed at respectively 51.14% and 75%. The product of 1-δ and the risk-neutral one-year default probabilities per rating is depicted in the table below:

	Fixed recovery rate (51.14%)	*Fixed recovery rate (75%)*
AAA	*0.00%*	*0.00%*
AA	*0.01%*	*0.03%*
A	*0.23%*	*0.21%*
BBB	*0.58%*	*0.60%*
BB	*3.99%*	*3.92%*
B	*11.15%*	*11.43%*

Table 4.10: *Product of 1-δ and the default probability for the RT assumption*

As we see, these numbers are pretty much equal, whether the recovery rate is fixed at 51.14% (S&P's estimated average recovery rate) or 75%. This could definitely be a problem if we want to price certain rating-dependent credit derivatives. In certain elements of their price the rating transition probabilities may appear without 1-δ in front of them, in which case we cannot be certain of which price to quote.

4.5.2. Insensitivity of default probabilities to migration probabilities

Corporate bonds are products that are contingent on default. It is therefore not surprising that only risk neutral default probabilities enter their pricing formulas.
Of course, default probabilities definitely do contain information about the non-default migration probabilities, but it remains a question how sensitive they are with respect to changes in the non-default probabilities. In the following simple example we will see that they are relatively insensitive, which implies that it is hard to infer the non-default migration probabilities, which we need in order to price rating-dependent derivatives, from the aforementioned products.
Suppose the risk-neutral credit migration process follows a time-homogeneous Markov chain, with the following (constant) one-step transition matrix:

$$A = \begin{pmatrix} B & D \\ 0 & 1 \end{pmatrix}$$

where B is a K - 1 $\times$ K - 1 matrix and D is a K - 1 $\times$ 1 row vector, representing the default probabilities from each non-default state. It is easy to check that:

$$A^n = \begin{pmatrix} B^n & \sum_{m=0}^{n-1} B^m D \\ 0 & 1 \end{pmatrix} \qquad (4.4)$$

For this example, assume that two of the default probabilities are equal to each other, say $d_i = d_j$. In practice this could occur for two investment-grade ratings, where the default probabilities are very small. Even if they are not exactly equal, their difference will be almost negligible. Let us now perturb the one-step matrix by switching around columns i and j. Denote this matrix by $\tilde{A}$, and the perturbed version of B by $\tilde{B}$. From (4.4) it is clear that entry ℓ of the difference between their two-step default probabilities equals:

$$\sum_{m=0}^{n-1}(b_{\ell m}-\tilde{b}_{\ell m})\cdot d_m=\sum_{m\neq i,j}(b_{\ell m}-b_{\ell m})\cdot d_m+(b_{\ell i}-b_{\ell j})\cdot d_i+(b_{\ell j}-b_{\ell i})\cdot d_i=0$$

In other words, both matrices have equal one and two-step default probabilities. We could carry on and analyse more-step default probabilities, but the results are less clear. The following numerical example will bring over our results a lot more clearly.

Example:
We took matrix A to be equal to the estimated one-day transition matrix from the model with the recovery of face value assumption, where the recovery rate was included as a variable. We switched around the 1st and 2nd columns. Although the one-day default probability starting from the 1st and 2nd rating were not equal, they were comparable in size. As the differences in default probabilities are not representative of price differences, we recalculated the model prices using the perturbed version of the transition matrix. The relative differences between the original model price and the price resulting from the perturbed estimates are depicted in the following table:

Rating	*AAA*	*AA*	*A*	*BBB*	*BB*	*B*
Average relative difference	0.00%	0.12%	0.80%	0.73%	0.59%	0.36%

Table 4.11: Average relative difference between model price with estimated one-day transition matrix and the model price with a perturbed one-day transition matrix

On average the differences do not seem to be large at all. In fact, the performance statistics are virtually indistinguishable from table 4.6:

Statistic	*AAA*	*AA*	*A*	*BBB*	*BB*	*B*	*Overall*
Absolute relative mispricing							
Average	1.31%	1.35%	1.20%	1.63%	4.10%	3.86%	1.47%
Median	0.95%	0.93%	0.47%	1.15%	2.96%	2.64%	0.95%
Standard deviation	2.05%	2.18%	2.15%	1.42%	3.63%	4.46%	2.33%
YTM difference (in bp)							
Average	28.34	30.63	-17.03	-12.05	56.99	69.05	22.34
Absolute average	37.96	49.2	30.16	50.05	138.4	155.04	48.14
Sum of squared errors	33.23		*R^2-statistic*		0.9944		

Table 4.12: Recovery of face value (payoff at maturity), variable recovery rate
Columns 1 and 2 switched around in estimated one-step matrix

Switching the two columns has however had a significant impact on the price of a downgrade put, as can be seen from the following example. Consider a downgrade put on a AAA rated bond that pays € 1 in two years time if the rating of the underlying bond at that point in time is below AAA. With the original estimates, the price for this downgrade put is € 0.90. With the perturbed version of the estimates, the price for the put becomes € 0.01, a very considerable difference. It is clear that we need to calibrate our model on more complex products if we want to be able to price rating-dependent credit derivatives.

4.6. Conclusions

In this chapter we described and estimated a time-homogenous model under the pricing measure $\mathbb{Q}$. Regardless of whether we do or do not infer the recovery rate from the data, it turns out that the recovery of face value assumption, where the recovery amount is paid at maturity, yields the best results. Overall the performance is worse than that of the traditional methods, but in terms of pricing errors it seems to outperform the traditional methods slightly for bonds with ratings from BBB to B.

From an analysis of the estimates we learn that there is still a great deal of uncertainty about the exact location of the non-default probabilities, which we require in order to be able to price rating-dependent credit derivatives. Two main reasons for this are:

- Possible misspecification of the model;
- The straight bonds and MTNs on which we calibrated our model are insensitive to non-default migration probabilities.

To solve the second possible reason, we suggested to calibrate our model on more complex credit producs, which possibly do contain information about non-default migration probabilities. Which other credit products are traded in the market? We recall from paragraph 2.1 that the following are the most actively traded contracts in the credit market:

- Default swap;
- Binary default swap;
- Credit spread options;
- Total return swap.

The payoff of the default swap and the binary default swap only depends on default, so that the analysis we made earlier in this chapter is also applicable here.

In addition to default, a credit spread option is also dependent on the evolution of the credit spread. This in its turn is partially determined by the credit rating of the underlying. However, we expect that there are other determinants more important than the credit rating of the underlying. Not many different credit spread options are traded currently, so that we cannot investigate this presumption.

The payoff of a total return swap is dependent on the appreciation and depreciation of the underlying bond. This is loosely dependent on credit events. However, for pricing the total return swap, many more factors are important, such as liquidity. Therefore we do not expect that these products will help us further in our quest for better estimation of the non-default migration probabilities.

Our final hope, product-wise, is to include the rating-dependent credit derivatives in the calibration process. Whether there are enough rating-dependent issues in the market, and whether they are trading at sensible prices, will be investigated in the next chapter.

One last possible solution that has not been mentioned yet, is to extend structural models to incorporate credit ratings. This type of model would endogenously determine the risk premiums, which in turn means that we would not have to estimate them. The CreditMetrics model by Gupton, Finger and Bhatia [1997] can for example easily be extended for the pricing of credit derivatives. The same caveats which hold for all structural models however still remain: the asset process is neither observable nor tradable. Therefore we have not been successful in specifying such a model in a sensible fashion.

Appendix - Estimation results

For completeness we include the most important estimation results for the two best specifications (one for a fixed, and one for an inferred recovery rate). The results depicted here are the best over the 10 parameter sets and were estimated on the full BOND dataset. Note that since no CCC or lower rated bonds were available in our data set, the B-rating here comprises all ratings lower than and including the B-rating.

1. Legal claim approach (fixed recovery rate)

Statistic	*AAA*	*AA*	*A*	*BBB*	*BB*	*B*	*Overall*
Absolute relative mispricing							
Average	1.12%	1.62%	1.06%	1.21%	4.69%	5.79%	1.52%
Median	0.57%	0.60%	0.42%	0.65%	3.50%	4.22%	0.60%
Standard deviation	2.89%	5.28%	2.03%	1.60%	3.77%	5.03%	3.95%
YTM difference (in bp)							
Average	-17.83	11.86	-20.22	-11.98	52.33	111.88	-2.39
Absolute average	28.84	34.65	30.34	37.82	128.51	208.47	39.13
Sum of squared errors	44.38		R^2-statistic		0.9925		

Table 4.13: *Performance results, recovery of face vakue assumption (payoff at default)*

	AAA	*AA*	*A*	*BBB*	*BB*	*B*	*D*
AAA	79.7%	8.3%	6.1%	4.8%	0.4%	0.7%	0.0%
AA	38.3%	17.0%	18.3%	18.0%	1.8%	6.4%	0.2%
A	28.4%	18.3%	20.4%	20.5%	2.2%	9.8%	0.4%
BBB	22.1%	18.0%	20.6%	20.9%	2.3%	15.2%	0.8%
BB	3.1%	3.2%	3.8%	4.0%	0.7%	78.7%	6.6%
B	0.0%	0.0%	0.0%	0.0%	0.0%	90.7%	9.3%

Table 4.14: *Implied one-year transition matrix, recovery of face value assumption (payoff at default)*

Recovery rate: 51.14%.

2. Recovery of face value assumption (payoff at default, inferred recovery rate)

Statistic	AAA	AA	A	BBB	BB	B	Overall
Absolute relative mispricing							
Average	1.31%	1.33%	1.16%	1.64%	4.13%	3.85%	1.46%
Median	0.94%	0.94%	0.45%	1.14%	3.00%	2.59%	0.94%
Standard deviation	2.04%	2.15%	2.12%	1.46%	3.61%	4.39%	2.31%
YTM difference (in bp)							
Average	27.19	31.94	-16.72	-14.1	56.86	67.14	22.51
Absolute average	38.98	47.84	31.4	49.49	136.9	152.04	47.26
Sum of squared errors	32.77		R^2-statistic		0.9945		

Table 4.15: *Performance results, recovery of face vakue assumption (payoff at default)*

	AAA	AA	A	BBB	BB	B	D
AAA	0.8%	99.1%	0.2%	0.0%	0.0%	0.0%	0.0%
AA	0.1%	99.8%	0.2%	0.0%	0.0%	0.0%	0.0%
A	0.1%	98.1%	0.9%	0.1%	0.0%	0.2%	0.6%
BBB	0.1%	75.7%	2.2%	0.4%	0.1%	3.7%	17.8%
BB	0.0%	16.5%	0.7%	0.2%	0.0%	11.1%	71.4%
B	0.0%	0.1%	0.0%	0.0%	0.0%	23.0%	76.9%

Table 4.16: *Implied one-year transition matrix, recovery of face value assumption (payoff at default)*

Estimated recovery rate: 99.98%.

5. Valuing ratings-sensitive coupon language

Over € 20 billion of Eurobonds issued in 2000 have ratings-sensitive coupon language, for example the bonds of some of the largest European telecom issuers such as Vodafone, Olivetti/Tecnost, Telstra, Deutsche Telekom and KPN. In the July 2000 issue of Credit, Conroy [2000] states a couple of reasons as to why ratings-sensitive coupon language can be a valuable feature for investors. We name two of the most important ones:

- It limits the pain experienced by investors. In the event of a downgrade the probability of default for a firm will most likely rise, lowering the price of the bond. With a ratings-sensitive coupon, i.e. a coupon that rises when the bond is downgraded, a downgrade will raise the value of a coupon, thereby compensating the effect of the downgrade. The price of a security with such language will not drop as much as the price of one without it.
- It reinforces an issuer's commitment to its rating, by establishing an extra penalty for engaging in actions that may deteriorate credit quality (e.g. acquisition, share buy-back, reorganisation, etc.). With a lower rating, issuers will find that borrowing money will become increasingly expensive. However, if they have sold bonds with ratings-dependent coupons, they will incur an extra penalty by having to pay higher coupons. The downside to this is that coupon step-ups may also have a deteriorating effect on credit quality: they could prompt further downgrades.

Having discussed some of the pros and cons of ratings-sensitive coupon language, it is now time to discuss the valuation of the embedded credit derivatives in the recent issues of the European telecommunications companies. The framework we will be working in is the JLT one. In the July 2000 issue of Risk, Gregory-Costello discusses the properties of these issues in more detail. From this article it becomes clear that in the market we can discern two different types of downgrade puts:

- A downgrade put that remains in place for the life of the bonds as can be found in e.g. the KPN and the Deutsche Telekom (DT) issues. For the KPN issues, the following holds: if at the time of a coupon payment either Moody's rating for KPN equals Baa1 or lower, or Standard & Poor's rating equals BBB+ or lower, the coupon rises to a certain value. If however at the next coupon payment both ratings are above Baa1/BBB+, the bondholder will receive the original coupon. In future when we talk about *downgrade puts*, we will mean this type.
- A downgrade put that can only be triggered once. If the put is triggered, the coupon will permanently be fixed at a higher level. We can make a distinction concerning the time when the put can be triggered:
  - The rating is reviewed only at a certain point in time. For the Vodafone issue, this was the time directly after the Mannesmann acquisition. We will refer to these downgrade puts as *one-off down-and-in downgrade puts*.
  - The rating is reviewed during the whole duration of the bond. If at any time the rating drops below a certain prespecified rating, the coupon is permanently fixed at a higher level. We will refer to these downgrade puts as *continuously reviewed down-and-in downgrade puts*.

The valuation of these downgrade puts will be the main focus of this chapter. Note that in practice, whether the downgrade puts are triggered or not is determined by the

lower of either Moody's or Standard & Poor's rating. This is not entirely equivalent to Bloomberg's composite rating (see the part in this thesis about the various credit rating agencies), which is provided in the Central Market Risk bond database. This may distort the results somewhat.
The chapter will start of with an empirical investigation of the market prices of both the KPN and DT issues, to examine whether they are sensibly priced in relation to the straight bonds of the same obligor. If they are, we could possibly include them in the calibration process for our model, as was suggested in the previous chapter. Three paragraphs will follow, deriving the theoretical prices of respectively the downgrade put, the one-off down-and-in downgrade put and the continuously reviewed down-and-in downgrade put. Subsequently we will calculate theoretical prices for the KPN and DT issues, using the estimated models from the previous chapter, and compare these to their market prices. The chapter will finish with conclusions.

5.1. Data on embedded rating-dependent credit derivatives

Unfortunately prices of the embedded rating-dependent credit derivatives in the recent telecommunications issues are not directly available to us. Therefore, if we want to have an indication of the price of the embedded credit derivative, we will have to find a method of inferring it from the available quotes. In this paragraph we will look at the embedded credit derivatives in the Deutsche Telekom and the two KPN issues. The specifications of the Euro-denominated issues for which quotes are available to us, are the following:

Obligor	*Issue date*	*Maturity*	*Coupon*	*Rating-dependent component*
KPN	*June 2000*	*13/06/2003*	*5.75/ann.*	*+30bp if rated Baa1/BBB+ or below*
	April 2001[18]	*12/04/2006*	*7.25/ann.*	*+37.5bp if rated below Baa2/BBB*
Deutsche Telekom	*July 2000*	*06/07/2005*	*6.125/ann.*	*+50bp if rated Baa1/BBB+ or below*
	July 2000	*06/07/2010*	*6.625/ann.*	*+50bp if rated Baa1/BBB+ or below*

Table 5.1: *Specification of the rating-dependent step-up bonds by KPN and DT*

All of these issues contain a long position in the embedded credit derivatives. Therefore their price can be seen as the sum of a straight coupon bond and the embedded credit derivative. Using quotes from other bonds of the same issuer, it should be possible to obtain an estimate of the embedded credit derivatives in these issues. The next subparagraph will deal with the issue of stripping out the straight bond component of the special issues in table 5.1.

5.1.1. Estimating the yield curve of an obligor

To strip out the straight bond component of each special issue from table 5.1, we require a good estimate of both the KPN and DT yield curves, which captures the firm-specific characteristics of these yield curves. We could estimate these yield curves using the B-spline or Svensson methods. Due to the fact that we will have a very small amount of bonds available, smaller than the amount of parameters for these models, we will consider two alternate methods.
First however, we briefly discuss how each straight bond of an obligor is built. Suppose we have M bonds of the same obligor. Bond j has $N(j)$ outstanding coupons, with coupon dates t_{j1}, ..., $t_{j,N(j)}$ and corresponding coupons C_{j1}, ..., $C_{j,N(j)}$. It also has a

[18] This issue also contains a rating-dependent put option: If the company loses its investment-grade rating because of a change in structure, then KPN will repay the bonds.

face value, F_j. We assume the value-additivity property holds, and therefore the price of zero-coupon equals:

$$V_j(t) = \sum_{i=1}^{N(j)} C_{ji} \overline{P}(t, t_{ji}) + F_j \overline{P}(t, t_{j,N(j)})$$

where $\overline{P}(t,T)$ is the time t price of a zero-coupon defaultable bond with maturity date T of that obligor. We now discuss two methods we used to strip the yield curve.

1. Bootstrapping the yield curve of the obligor

With this method, we just use straight bonds of the same obligor and bootstrap the yield curve from the market prices of the bonds. Suppose that the bonds are sorted in order of maturity, and that each bond has a different maturity. With the bootstrapping method we assume the yield curve has the following form:

$$\begin{aligned}
&R(t,T) = R_1 && T \le t_{1,N(1)} \\
&R(t,T) = R_j && t_{j-1,N(j-1)} < T \le t_{j,N(j)}, \quad 1 < j \le M \\
&R(t,T) = R_M && t_{M,N(M)} \le T
\end{aligned}$$

The zero-coupon defaultable bonds satisfy: $\overline{P}(t,T) = \exp\bigl(-R(t,T)\cdot(T-t)\bigr)$. Using this information, it is straightforward to determine the yield curve from the market prices of available bonds.

2. Using a Svensson yield curve for obligors of the same rating

We can estimate a Svensson yield-curve for obligors of the same rating in the same fashion as we did before. After having estimated this curve, we assume that the yield for the obligor under investigation equals the yield for that specific rating plus a fixed spread s that is particular to that obligor. Now, using just bonds of the obligor under investigation, we estimated this fixed spread by minimising the sum of squared errors between the market prices and the model prices. The advantage of this method over the previous method is that we obtain a smooth yield curve for each obligor, which will capture the dynamics that are particular to that rating. A disadvantage is that typically the model prices will deviate slightly from the market prices of the available straight bonds, something that is not the case in the previous method.

5.1.2. Empirical results

We will investigate if the two methods above yield sensible prices for the embedded credit derivatives by applying them to the market data. From Bloomberg we obtained quotes for the following straight Euro-denominated issues from KPN and DT:

Obligor	*Issue date*	*Maturity*	*Coupon*
KPN	*July 1996*	*03/07/2006*	*6.5/ann.*
	June 1998	*30/06/2004*	*4/ann.*
	November 1998	*05/11/2008*	*4.75/ann.*
	October 2000	*04/10/2005*	*6.25/ann.*
Deutsche Telekom	*May 1998*	*20/05/2008*	*5.25/ann.*
	May 2001	*14/05/2002*	*5/ann.*
	April 2001	*10/04/2002*	*4.8/ann.*

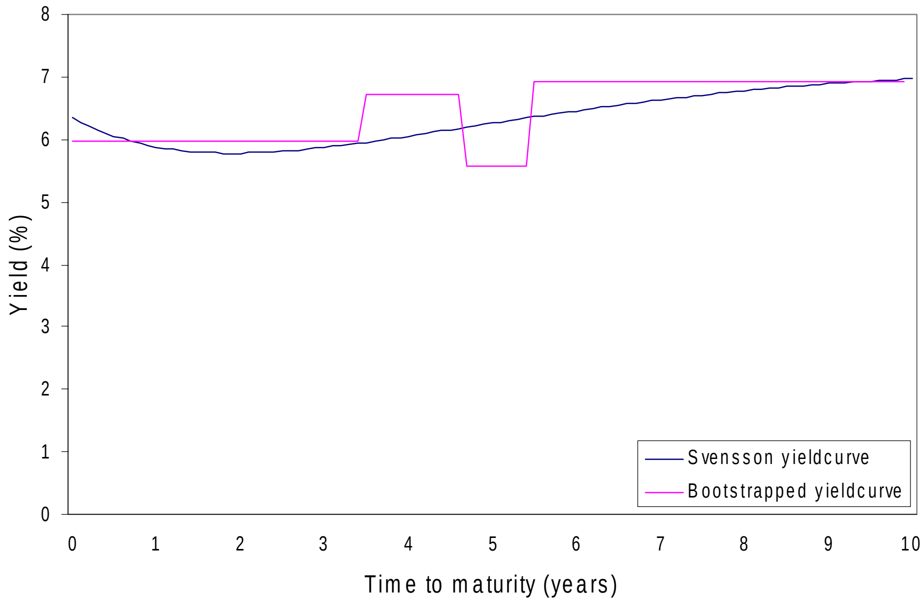
Table 5.2: *Straight Euro-denominated issues available in the market from KPN and DT*

Both obligors have other issues in the market. Unfortunately quotes for issues other than those in table 5.2 were not available due to the fact that these were not traded frequently. For all of the issues in table 5.2, we took the so-called Bloomberg Fair Values (BFV's). A BFV is defined as 'a bond's theoretical value, based on where similar bonds, as defined by credit quality and market sector, have traded'.
Unfortunately, our bond database only contains quotes for all straight Eurobonds starting from 09/03/1999 until 10/01/2001. Since most of the available Deutsche Telekom bonds were issued after that date, we will not be able to use the second method to obtain an estimate for the embedded credit derivative in the DT issues. We will therefore:

1. Compare both methods for KPN bonds for the last trading date in four consecutive months, namely the last four months in 2000.
2. Use the bootstrapping method on four weeks in May, and compare the prices we obtain for the DT and the KPN credit derivatives.

Comparison of both methods

We applied both methodologies to the available quotes of KPN bonds for the last trading dates in the four last months of 2000. To give an impression of the resulting yield curves for both methods, we depict both estimated yield curves for 29/12/2000 in the subsequent graph:



Graph 5.1: Yield curves estimated with the recovery of treasury assumption

The bootstrapping method does not produce a very stable yield curve, due to the fact that only four straight bonds are used in the process. The resulting Svensson yield curve is smoother; the average absolute relative error of the model price with respect to the market price was 0.5%. In the following table we present estimates for the price of the embedded credit derivative component (in €, assuming the face value of the bond is 100 €) in the KPN June 2000 issue:

	Bootstrapped estimate	Svensson estimate
29/09/2000	-0.52	-0.89
27/10/2000	0.31	0.03
24/11/2000	0.25	0.09
29/12/2000	0.30	0.02

Table 5.3: *Estimates for the embedded credit derivative in the KPN June 2000 issue*

As can be seen, both methods produce different results. Also we see that in 29/09/2000 the implied price of the credit derivative was negative. This implies that the yield on the embedded bond in the KPN June 2000 issue at that time is significantly higher than should be expected. Liquidity issues could cause this.

Comparison of DT and KPN embedded credit derivatives

Using the bootstrap method for both DT and KPN, we estimated the prices of the embedded credit derivatives within the aforementioned four special issues for the last four weeks in May.

	KPN 13/06/2003	KPN 12/04/2006	DT 06/07/2005	DT 06/07/2010
04/05/2001	1.26	0.11	2.69	0.45
11/05/2001	1.44	0.15	3.12	0.16
18/05/2001	1.40	0.13	2.48	-0.26
25/05/2001	1.51	0.18	2.82	-0.44

Table 5.4: *Estimates for the embedded credit derivatives in several KPN and DT issues*

The results can be seen in the table above. Again, we notice that some prices are negative, which is obviously not possible. From the table in which the specifications of the special issues are mentioned, it is clear that the embedded derivative in the DT 06/07/2010 issue contains the embedded derivative in the DT 06/07/2005 issue. Its price should therefore be at least equally high. This is not reflected in the results.
We find it safe to conclude that, if we only have a small number of traded straight bonds available, it will be almost impossible to obtain a fair market price for the embedded credit derivative. The reason why we wanted to do this in the first place was to use only the prices of the rating-dependent credit derivatives to calibrate our pricing model. As they are traded in combination with the underlying bond, they have a relatively small influence on the price of the complete package. Information about rating migrations will therefore typically not be picked up by an estimation algorithm. We will now look at the theoretical prices of several types of downgrade puts.

5.2. Downgrade put

We will be working on the zero-coupon level in order to value the downgrade put. Let us consider a downgrade put on a zero-coupon bond, maturing at time *T*, which raises the payoff by 1 in case the rating at the payoff date is larger[19] than rating *j*. We partition the state space *S* as follows:

$$S_u(j) = \{1,\dots,j\} \qquad S_d(j) = \{j+1,\dots,K-1\}$$
$$S = S_u(j) \cup S_d(j) \cup \{K\}$$

[19] Note: we are considering a downgrade from *j*, which in our model means that we go to rating $i > j$.

Throughout this chapter we will be working in the discrete-time JLT framework, irrespective of any further assumptions about the credit migration process. For the formulas presented in this chapter we assume the recovery of treasury assumption. To obtain the prices under the recovery of face value assumption, we merely need to replace the recovery rate δ by zero. For the legal claim approach the derivation becomes slightly different, as we have to take into account the accrued interest.
Under the recovery of treasury assumption, the value of the zero-coupon bond after default equals a fraction of the no-default value. This means that the payoff of the downgrade put at maturity equals:

$$\mathbb{1}_{[\tau>T,\, X_T \in S_d(j)]} + \delta \cdot \mathbb{1}_{[\tau\le T,\, X_{\tau-1} \in S_d(j)]}$$

The time t value of this downgrade put, which we will denote by $p_j(t,T)$, equals:

$$\begin{aligned} p_j(t,T) &= \mathbb{E}_t^{\mathbb{Q}}[\tfrac{B(t)}{B(T)} \cdot (\mathbb{1}_{[\tau>T,\, X_T \in S_d(j)]} + \delta \cdot \mathbb{1}_{[\tau\le T,\, X_{\tau-1} \in S_d(j)]})] \\ &= P(t,T) \cdot \mathbb{Q}_t\big(X_T \in S_d(j)\big) + P(t,T) \cdot \mathbb{E}_t^{\mathbb{Q}}\Big[\mathbb{E}_t^{\mathbb{Q}}[\delta \cdot \mathbb{1}_{[\tau\le T,\, X_{\tau-1} \in S_d(j)]} \mid \tau]\Big] \\ &= P(t,T) \cdot \mathbb{Q}_t\big(X_T \in S_d(j)\big) + \delta \cdot P(t,T) \cdot \sum_{m=1}^{n} \mathbb{Q}_t(\tau = t + m\triangle,\, X_{\tau-1} \in S_d) \end{aligned}$$

where we assumed that $T = t + n\triangle$. The $\mathbb{Q}$-probabilities appearing in this valuation formula can easily be expressed in terms of the transition matrices. Assuming that the rating at time t equals i:

$$\begin{aligned} &\mathbb{Q}_t\big(X_T \in S_d(j)\big) = \sum_{\ell \in S_d(j)} \tilde{a}_{i\ell}(t,T) \\ &\mathbb{Q}_t(\tau = t + m\triangle,\, X_{\tau-1} \in S_d) = \sum_{\ell \in S_d(j)} \tilde{a}_{i\ell}\big(t, t+(m-1)\triangle\big) \cdot \tilde{a}_{\ell K}\big(t+(m-1)\triangle, t+m\triangle\big) \end{aligned}$$

This completes the valuation of the downgrade put. Suppose we have a coupon bond, currently rated i, with coupon payments C at times t_1 through $t_n = T$ and face value F. This bond also has an embedded downgrade put: if at any coupon date t_ℓ the rating is larger than rating j, the coupon will be equal to $C + \triangle C$. The value of this portfolio at time t equals:

$$\sum_{\ell=1}^{n} \big(C \cdot \overline{P}^i(t,t_\ell) + \triangle C \cdot p_j(t,t_\ell)\big) + F \cdot \overline{P}^i(t,T)$$

5.3. One-off down-and-in downgrade put

Again we will be working on the zero-coupon level. A one-off down-and-in downgrade put on a zero-coupon bond, maturing at time T, raises the payoff by 1 if the rating at the review date $t_R = t + r\triangle$ was larger than rating j. In the JLT framework this implies the following for the payoff at maturity:

$$\mathbb{1}_{[\tau>T,\, X_{t_R} \in S_d(j)]} + \delta \cdot \mathbb{1}_{[t_R<\tau\le T,\, X_{t_R} \in S_d(j)]}$$

We here assume that if default occurs before the review date, the downgrade put will not be triggered. As before, we can easily deduce that the time t value equals:

$$p_j(t,T,t_R) = \mathbb{E}_t^{\mathbb{Q}}[\tfrac{B(t)}{B(T)} \cdot (\mathbb{1}_{[\tau>T,\, X_{t_R} \in S_d(j)]} + \delta \cdot \mathbb{1}_{[t_R<\tau\le T,\, X_{t_R} \in S_d(j)]})]$$
$$= P(t,T) \cdot \left(\mathbb{Q}_t(\tau > T, X_{t_R} \in S_d(j)) + \delta \sum_{m=1}^{n-r} \mathbb{Q}_t(\tau = t_R + m\triangle,\, X_{t_R} \in S_d)\right)$$

The $\mathbb{Q}$-probabilities can again be expressed in terms of the transition matrices, assuming that at time t we are in rating i:

$$\mathbb{Q}_t(\tau > T, X_{t_R} \in S_d(j)) = \sum_{\ell \in S_d(j)} \sum_{m \neq K} \tilde{a}_{i\ell}(t,t_R) \cdot \tilde{a}_{\ell m}(t_R,T)$$
$$= \sum_{\ell \in S_d(j)} \tilde{a}_{i\ell}(t,t_R) \cdot (1 - \tilde{a}_{\ell K}(t_R,T))$$
$$\mathbb{Q}_t(\tau = t_R + m\triangle,\, X_{t_R} \in S_d(j)) = \sum_{\ell \in S_d(j)} \tilde{a}_{i\ell}(t,t_R) \cdot \tilde{a}_{\ell K}(t_R, t_R + m\triangle)$$

Coupon bonds with embedded one-off down-and-in downgrade puts can now, as before, be expressed as a portfolio of defaultable zero-coupon bonds and one-off down-and-in downgrade puts. Note that a one-off down-and-in downgrade put with review date equal to $t_R = T$ is the previously treated downgrade put.

5.4. Continuously reviewed down-and-in downgrade put

A continuously reviewed down-and-in downgrade put on a zero-coupon bond, maturing at time T, raises the payoff by 1 if at any time during the life of the bond the rating has been lower than j. The payoff at maturity equals:

$$\mathbb{1}_{[\tau>T]} \cdot \left(1 - \prod_{k=1}^{n} \mathbb{1}_{[X_{t+k\triangle} \notin S_d(j)]}\right) + \delta \cdot \mathbb{1}_{[\tau=t+r\triangle\le T]} \cdot \left(1 - \prod_{k=1}^{r-1} \mathbb{1}_{[X_{t+k\triangle} \notin S_d(j)]}\right)$$

This payoff can be substantially simplified if we modify the state space slightly. We introduce the following modified state space:

$$S'(j) = S_u(j) \cup \{K\} \cup S^*(j) = S_u(j) \cup \{K\} \cup \{1^*(j), 2^*(j), \ldots, K^*(j)\}$$

where being in state $i^*(j)$ means that the original Markov process is in rating i and that it has been below rating j during the life of the bond. The payoff then is:

$$\mathbb{1}_{[\tau>T,\, X_T \in S^*(j)]} + \delta \cdot \mathbb{1}_{[\tau\le T,\, X_\tau = K^*(j)]}$$

The default time τ is here of course the first time at which either K or $K^*(j)$ is reached. We now consider the one-step transition matrices on the enlarged state space. Let us first partition the one-step transition matrix on the original state space:

$$\tilde{A}(t,t+\triangle) = \begin{pmatrix} \tilde{A}_{uu}(t,t+\triangle) & \tilde{A}_{ud}(t,t+\triangle) & \tilde{A}_{uK}(t,t+\triangle) \\ \tilde{A}_{du}(t,t+\triangle) & \tilde{A}_{dd}(t,t+\triangle) & \tilde{A}_{dK}(t,t+\triangle) \\ 0_{1\times j} & 0_{1\times(K-j)} & 1 \end{pmatrix}$$

where e.g. $\widetilde{A}_{ud}(t,t+\triangle)$ is the transition probability matrix from states $S_u(j)$ to $S_d(j)$, and $0_{m\times n}$ is an $m \times n$ matrix with zeros. What does the transition probability matrix for the enlarged state space look like? As with S we now partition $S^*(j)$ into $S_u^*(j)$, $S_d^*(j)$ and $\{K^*(j)\}$. The transition probability matrix now simply becomes:

$$\widetilde{A}'(t,t+\triangle)=\begin{pmatrix} \widetilde{A}_{uu}(t,t+\triangle) & \widetilde{A}_{uK}(t,t+\triangle) & 0_{j\times j} & \widetilde{A}_{ud}(t,t+\triangle) & 0_{j\times 1} \\ 0_{1\times j} & 1 & 0_{1\times j} & 0_{1\times(K-j)} & 0 \\ 0_{j\times j} & 0_{j\times(K-j)} & \widetilde{A}_{uu}(t,t+\triangle) & \widetilde{A}_{ud}(t,t+\triangle) & \widetilde{A}_{uK}(t,t+\triangle) \\ 0_{(K-j)\times j} & 0_{(K-j)\times(K-j)} & \widetilde{A}_{du}(t,t+\triangle) & \widetilde{A}_{dd}(t,t+\triangle) & \widetilde{A}_{dK}(t,t+\triangle) \\ 0_{1\times j} & 0_{1\times(K-j)} & 0_{1\times j} & 0_{1\times(K-j)} & 1 \end{pmatrix}$$

Note that we have not specified the dependency of the matrix on j to not further complicate notation. This matrix is easy to read if we keep in mind that:

$$S'(j)=S_u(j)\cup\{K\}\cup S_u^*(j)\cup S_d^*(j)\cup\{K^*(j)\}$$

Therefore e.g. element (4,3) (as depicted here, i.e. $\widetilde{A}_{du}(t,t+\triangle)$) is the transition probability matrix from $S_d^*(j)$ to $S_u^*(j)$. Multi-step transition probability matrices can be calculated in the same fashion as before. Having specified the transition matrices for the enlarged state space, it is now straightforward to calculate the time t value of the continuously reviewed down-and-in downgrade put:

$$\begin{aligned} p_j(t,T) &= \mathbb{E}_t^{\mathbb{Q}}\left[\tfrac{B(t)}{B(T)}\cdot\left(\mathbb{1}_{[\tau>T,\,X_T\in S^*(j)]}+\delta\cdot\mathbb{1}_{[\tau\le T,\,X_\tau=K^*(j)]}\right)\right] \\ &= P(t,T)\cdot\left(\mathbb{Q}_t\left(\tau>T,\,X_T\in S^*(j)\right)+\delta\cdot\mathbb{Q}_t\left(\tau\le T,X_\tau=K^*(j)\right)\right) \\ &= P(t,T)\cdot\sum_{\ell\in S^*(j)}\tilde{a}'_{i\ell}(t,T)+\delta\cdot P(t,T)\cdot\sum_{\ell=1}^{n}\tilde{a}'_{iK^*(j)}(t,t+\ell\triangle) \end{aligned}$$

where we assumed that $T = t + n\triangle$ and that the rating at time t equals $i \le j$. If the rating at time t equals $i > j$ (but not equal to K), the put has been triggered and its value is equal to $\overline{P}^i(t,T)$.

5.5. Numerical example

Using the constrained optimisation model which we estimated in the previous chapter, we will now check what price this model yields for the KPN June 2000 issue, which has embedded downgrade puts, as described in paragraph 5.2. The specification of this issue, which can also be found in table 5.1, is the following:

Issue date	*Maturity*	*Coupon*	*Rating-dependent component*
June 2000	*13/06/2003*	*5.75/ann.*	*+30bp if rated Baa1/BBB+ or below*

Table 5.5: *Specification of the KPN June 2000 issue*

The embedded downgrade puts are triggered at each coupon date, if the issue is rated Baa1/BBB+ or below on the coupon date. If this is the case, the coupon is raised by 30 basis points. At the estimation date, which we chose to be equal to 04/08/2000, the KPN issue was rated Aa2 by Moody's. The last midquote for the bond on that date (including accrued interest) equalled € 101.06 (assuming the face value was € 100).
In the subsequent table we give the model prices of an otherwise equal AA-rated bond and the price of the embedded downgrade puts for the eight recovery rate assumptions investigated in the previous chapter.

Recovery rate assumption	*Bond price*	*Downgrade puts*	*Estimated total price*
Fixed recovery rate			
Recovery of treasury	*102.29*	*0.35 (0.05,0.64)*	*102.64*
Recovery of face value			
Payoff at maturity	*102.25*	*0.14 (0.00,0.46)*	*102.39*
Payoff at default	*102.05*	*0.00 (0.00,0.16)*	*102.05*
Legal claim approach	*102.42*	*0.26 (0.00,0.64)*	*102.68*
Inferred recovery rate			
Recovery of treasury	*102.29*	*0.03 (0.00,0.43)*	*102.32*
Recovery of face value			
Payoff at maturity	*102.30*	*0.00 (0.00,0.25)*	*102.30*
Payoff at default	*101.75*	*0.00 (0.00,0.08)*	*101.75*
Legal claim approach	*101.86*	*0.01 (0.00,0.08)*	*101.87*

Table 5.6: *Prices of the KPN issue for the various estimates*

In the table we also included 95% asymptotic confidence intervals for the prices of the downgrade puts. The first thing we notice is that the model price of the straight bond component is already higher than the quoted market price of the full issue. KPN bonds tend to have higher yields than other bonds of the same rating, but this could also be caused by something which we noticed for the market price on 29/09/2000 in paragraph 5.1.2: the KPN June 2000 issue seems to be trading at a discount relative to other KPN bonds.
The impact of the embedded credit derivatives on the total price is quite small. The estimates for the straight bond component are quite close for all the recovery rate assumptions, the prices of the embedded downgrade puts however, are not (relatively speaking). Whether the recovery rate was fixed at 51.14% or whether it was inferred from the market prices of bonds also seems to have quite an impact. Finally we notice that the confidence intervals for the downgrade puts are very wide, indicating a large uncertainty about the exact price of the downgrade put. This is not surprising given the results of the previous chapter.

5.6. Conclusions

In this chapter we had a closer look at the pricing of bonds with ratings-sensitive coupons. Empirical data on these issues is available, but it is hard to strip out the prices of the embedded derivative components. The impact of the embedded rating-dependent components on the total price of the considered issues seems to be quite small. Therefore we conclude that including the few rating-dependent issues that are available in the calibration process of the model, as suggested in the previous chapter, will not improve the estimates of the non-default migration probabilities significantly.
Using the estimated model from the previous chapter to calculate the theoretical prices for the KPN June 2000 issue, we found that there is a great uncertainty concerning the exact price of the embedded credit derivatives. This is not only due to the uncertainty in the estimates of the parameters, but also due to the uncertainty about which recovery rate assumptions should be used.
Due to the somewhat disappointing results in this and the previous chapter, we will review a more general and elegant model in the next chapter, namely that of Bielecki and Rutkowski [2000a], which introduces an HJM-like model with credit ratings.

6. HJM with credit ratings

In a recent article, Bielecki and Rutkowski [2000a] have introduced a model which according to themselves can best be described as: ‘the credit-spreads-based HJM-type arbitrage-free term structure model with multiple ratings’. We will briefly show how their model can be derived and will later discuss how this model can possibly be implemented, as this is not indicated in the article. To be complete, the technical details of the proofs of the model can be found in Bielecki and Rutkowski [1999]. In the near future this article, including some minor extensions, will be published in Bielecki and Rutkowski [2001]. We are grateful to Prof. Bielecki for sending us these last two articles. Finally, a non-technical summary of this paper can also be found in Bielecki and Rutkowksi [2000b].

6.1. Description of the forward rates

First we introduce some notation (note that we use a somewhat different notation than the discussed paper). We assume the following set of rating classes: $S = \{1,\dots, K\}$, where class K corresponds to default. The default state is absorbing, i.e. there is no restructuring after default. The price of a default-free zero-coupon bond, maturing at time T, is defined as:

$$P(t,T) \equiv \exp\left(-\int_t^T f(t,s)\, ds\right) \tag{6.1}$$

Furthermore we introduce pseudo-bonds, each corresponding to a non-default rating:

$$P_i(t,T) \equiv \exp\left(-\int_t^T f_i(t,s)\, ds\right) \qquad i = 1,\dots, K-1$$

Note that these pseudo-bonds do not represent price processes of traded securities. The process $P_i(t,T)$ should be interpreted as the conditional price at time t of the defaultable zero-coupon bond, given that default has not yet occurred. Finally, we introduce the money market account:

$$dB_t = r_t\, B_t\, dt$$

where $r_t = f(t,t)$ is the spot rate. To simplify the derivation of the model, we will use the discounted processes $Z(t,T) = B_t^{-1} P(t,T)$ and likewise $Z_i(t,T) = B_t^{-1} P_i(t,T)$. Bielecki and Rutkowski now model the term structure of the zero-coupon bonds by modelling the default-free forward-rates and the forward rates for each rating class. They assume:

$$\begin{aligned} d\, f(t,T) &= \alpha(t,T)\, dt + \sigma(t,T)^T \cdot dW_t \\ d\, f_i(t,T) &= \alpha_i(t,T)\, dt + \sigma_i(t,T)^T \cdot dW_t \end{aligned} \tag{6.2}$$

where $\alpha(t,T)$ and $\sigma(t,T)$ (and correspondingly $\alpha_i(t,T)$ and $\sigma_i(t,T)$) are adapted stochastic processes in $\mathbb{R}$ and $\mathbb{R}^d$, respectively.
We assume that the forward rates do not cross:

$$f_{K-1}(t,T) > f_{K-2}(t,T) > \ldots > f_1(t,T) > f(t,T)$$

This is satisfied for example if we take the same volatility function for all ratings and ensure that the drift functions are ordered in the same way.
We will now derive the HJM model and show that the same results of HJM can be carried over to defaultable forward rates. From the description of the default-free forward rates in (6.2) and the definition in (6.1), we can derive the dynamics of the bond price in terms of the dynamics of the forward rates. Using Itô calculus:

$$\begin{aligned} d\left(-\int_t^T f(t,u)\,du\right) &= f(t,t)\,dt - \int_t^T d\,f(t,u)\,du \\ &= f(t,t)\,dt - \int_t^T \left(\alpha(t,u)\,dt + \sigma(t,u)^T\,dW_t\right)du \\ &= \left(f(t,t) - \alpha^*(t,T)\right)dt - \sigma^*(t,T)^T\,dW_t \end{aligned}$$

where we introduced the following notation:

$$\alpha^*(t,T) = \int_t^T a(t,u)\,du \qquad \sigma^*(t,T)^T = \int_t^T \sigma(t,u)^T\,du$$

For the dynamics of the bond price, this implies:

$$\frac{d\,P(t,T)}{P(t,T)} = \left(f(t,t) - \alpha^*(t,T) + \tfrac{1}{2}\left\|\sigma^*(t,T)\right\|^2\right)dt - \sigma^*(t,T)^T\,dW_t \tag{6.3}$$

where $\|\cdot\|$ denotes the L_2-norm of a vector. Now, suppose we change measure from $\mathbb{P}$ to $\mathbb{Q}$, the pricing measure, which has the money market account as the numéraire asset. We define a Brownian motion under $\mathbb{Q}$ by setting:

$$\tilde{W}_t = W_t - \int_0^t \gamma_u\,du$$

Using this, (6.3) becomes:

$$\frac{d\,P(t,T)}{P(t,T)} = \left(f(t,t) - \alpha^*(t,T) + \tfrac{1}{2}\left\|\sigma^*(t,T)\right\|^2 - \sigma^*(t,T)^T\cdot\gamma_t\right)dt - \sigma^*(t,T)^T\,d\tilde{W}_t$$

For the discounted process $Z(t,T)$ this implies:

$$\frac{d\,Z(t,T)}{Z(t,T)} = \left(-\alpha^*(t,T) + \tfrac{1}{2}\left\|\sigma^*(t,T)\right\|^2 - \sigma^*(t,T)^T\cdot\gamma_t\right)dt - \sigma^*(t,T)^T\,d\tilde{W}_t$$

If there is no arbitrage, under $\mathbb{Q}$ the process $Z(t,T)$ must be a martingale. This yields:

$$\alpha^*(t,T) = \tfrac{1}{2}\left\|\sigma^*(t,T)\right\|^2 - \sigma^*(t,T)^T\cdot\gamma_t$$

We could assume that γ_t is uniquely determined by this condition, so that the default-free market is complete. This is however not necessary. Following the same rationale, and introducing $\alpha_i^*(t,T)$ and $\sigma_i^*(t,T)$ in a similar fashion, we obtain the following dynamics for the pseudo-bonds under the pricing measure $\mathbb{Q}$:

$$\frac{d\,P_i(t,T)}{P_i(t,T)} = \left(f_i(t,t) - \alpha_i^*(t,T) + \tfrac{1}{2}\left\| \sigma_i^*(t,T) \right\|^2 - \sigma_i^*(t,T)^T \cdot \gamma_t \right) dt - \sigma_i^*(t,T)^T \; d\tilde{W}_t$$

The drift of $\dfrac{d\,Z_i(t,T)}{Z_i(t,T)}$ under $\mathbb{Q}$ now equals:

$$\ell_i(t,T) \equiv f_i(t,t) - f(t,t) - \alpha_i^*(t,T) + \tfrac{1}{2}\left\| \sigma_i^*(t,T) \right\|^2 - \sigma_i^*(t,T)^T \cdot \gamma_t \tag{6.4}$$

which Bielecki and Rutkowski assume not to depend on *T*, but this is not a necessary condition for the development of the model. This drift term does not necessarily equal zero, as the pseudo-bonds are not traded assets. A first difference with the traditional HJM modelling approach will soon become visible. First observe that:

$$\frac{\partial \sigma^*(t,T)}{\partial T} = \sigma(t,T) \qquad\qquad \frac{\partial \alpha^*(t,T)}{\partial T} = \alpha(t,T)$$

$$\frac{\partial \tfrac{1}{2}\left\| \sigma^*(t,T) \right\|^2}{\partial T} = \sigma(t,T)^T \sigma^*(t,T)$$

We can therefore rewrite the condition on γ_t as follows:

$$\alpha(t,T) = \sigma(t,T)^T \sigma^*(t,T) - \sigma(t,T)^T \cdot \gamma_t$$

Under $\mathbb{Q}$ we therefore have the following evolution of the default-free forward rates:

$$d\,f(t,T) = \sigma(t,T)^T \sigma^*(t,T)\,dt + \sigma(t,T)^T \cdot d\tilde{W}_t$$

The evolution of the forward rates under $\mathbb{Q}$ solely depends on the volatility function. This means that in order to obtain prices for interest-rate derivatives, we do not require to model the drift of the default-free forward rates. However, we also have forward rates for ratings $i = 1, ..., K\text{-}1$:

$$d\,f_i(t,T) = \left(\alpha_i(t,T) + \sigma_i(t,T)^T \gamma_t \right) dt + \sigma_i(t,T)^T \cdot d\tilde{W}_t$$

In this setting therefore, we definitely do need to specify the drift terms of both the default-free and the defaultable forward rates. This increases the difficulty of the calibration considerably.

We have finished the description of the forward rates under the pricing measure. We will now continue with the description of the credit migration process.

6.2. Description of the credit migration process

We introduce a conditionally Markov chain, denoted by X_t, on the state space S. Formally, this has to be done by enlarging the probability space on which we were implicitly working earlier. To not complicate matters further, we will ignore this. The conditional generator under $\mathbb{Q}$ at time t, given the σ-field $\mathcal{F}_t$ is:

$$\Lambda_t = \left(\lambda_{ij}(t)\right)_{i,j=1,\dots,K}$$

where $\lambda_{Kj}(t) = 0$ for $j = 1,\dots, K$ and $\lambda_{ii}(t) = -\sum_{j\neq i} \lambda_{ij}(t)$.

The transition intensities are adapted, strictly positive processes. We now state the following property with respect to this Markov chain.

Lemma:

First define $H_i(t) \equiv \mathbb{1}_{[X_t = i]}$ for $i \in S$. Secondly, let $H_{ij}(t)$ represent the number of transitions from i to j over $[0, t]$. The following is a martingale for $i < K$ and $i \neq j$:

$$M_{ij}(t) \equiv H_{ij}(t) - \int_0^t \lambda_{ij}(u) H_i(u)\, du$$

Proof:

For a formal proof we refer you to Bielecki and Rutkowski [1999]. It is however intuitively clear that $M_{ij}(t)$ is a martingale. Using Itô's lemma that allows for jump processes, we obtain the following:

$$dM_{ij}(t) = dH_{ij}(t) - \lambda_{ij}(t) H_i(t)\, dt$$

If we show that the right-hand side of this equation has no drift, we have shown that $M_{ij}(t)$ is a martingale. Let us take expectations with respect to the information available up to and including time $s \leq t$:

$$\begin{aligned}
\mathbb{E}_s^{\mathbb{Q}}[dM_{ij}(t)] &= \mathbb{E}_s^{\mathbb{Q}}[dH_{ij}(t)] - \mathbb{E}_s^{\mathbb{Q}}[\lambda_{ij}(t) H_i(t)\, dt] \\
&= \mathbb{E}_s^{\mathbb{Q}}[\mathbb{E}_t^{\mathbb{Q}}[dH_{ij}(t)]] - \mathbb{E}_s^{\mathbb{Q}}[\lambda_{ij}(t) H_i(t)\, dt] \\
&= \mathbb{E}_s^{\mathbb{Q}}[H_i(t) \cdot (\lambda_{ij}(t)\, dt \cdot 1 + (1 - \lambda_{ij}(t)\, dt) \cdot 0)] - \mathbb{E}_s^{\mathbb{Q}}[\lambda_{ij}(t) H_i(t)\, dt] = 0
\end{aligned}$$

where we used the well-known fact that the probability of going from state i to j from time t to time $t + dt$ equals $\lambda_{ij}(t)\, dt$. □

6.3. The price process of a defaultable zero-coupon bond

Now, in order to derive the price process of a defaultable bond, Bielecki and Rutkowski assume that after default, the bondholder receives a fraction of an otherwise equivalent default-free bond (i.e. the recovery of treasury assumption), where the recovery rate is constant, but may depend on the rating the bond was in prior to default. To keep things as simple as possible, we will just assume a constant recovery rate δ here. We will now derive the price process of a defaultable bond, by guessing its form, and introducing some additional conditions on the conditional generator matrix of the credit migration process under which this price process is indeed a martingale. We guess the discounted process of a defaultable bond can be written as follows (as before, we use a bar to denote that the bond is defaultable):

$$\overline{Z}(t,T) = \mathbb{1}_{[X_t \neq K]} \cdot Z_{X_t}(t,T) + \delta \cdot \mathbb{1}_{[X_t = K]} \cdot Z(t,T) \tag{6.5}$$

First observe that we must have: $\sum_{i=1}^{K-1} H_{iK}(t) = \mathbb{1}_{[X_t = K]}$. Indeed, if the left-hand side of this equation equals one (it cannot become larger due to the absorbing nature of the default state), the current state must be K. Using this, we can write:

$$\overline{Z}(t,T) = \sum_{i=1}^{K-1} \left(H_i(t) \cdot Z_i(t,T) + \delta \cdot H_{iK}(t) \cdot Z(t,T) \right) \tag{6.6}$$

Also, observe that we have:

$$H_i(t) = H_i(0) + \sum_{\substack{j=1 \\ j \neq i}}^{K-1} \left(H_{ji}(t) - H_{ij}(t) \right) - H_{iK}(t)$$

Using the modified version of Itô's lemma that allows for jump processes yields:

$$dH_i(t) = \sum_{\substack{j=1 \\ j \neq i}}^{K-1} \left(dH_{ji}(t) - dH_{ij}(t) \right) - dH_{iK}(t) \tag{6.7}$$

Also, Itô's lemma on each term in (6.6) gives:

$$\begin{aligned} d\left(H_i(t) \cdot Z_i(t,T) \right) &= H_i(t)\ dZ_i(t,T) + Z_i(t,T)\ dH_i(t) \\ d\left(H_{iK}(t) \cdot Z(t,T) \right) &= H_{iK}(t)\ dZ(t,T) + Z(t,T)\ dH_{iK}(t) \end{aligned} \tag{6.8}$$

Using (6.7) and (6.8) in (6.6) gives us the following for the dynamics of the price of a discounted defaultable bond:

$$\begin{aligned} d\,\overline{Z}(t,T) &= \sum_{i=1}^{K-1} Z_i(t,T) \cdot \left(\sum_{j=1, j\neq i}^{K-1} \left(dH_{ji}(t) - dH_{ij}(t) \right) - dH_{iK}(t) \right) \\ &\quad + \sum_{i=1}^{K-1} \left(H_i(t)\, dZ_i(t) + \delta\, Z(t,T)\, dH_{iK}(t) + \delta\, H_{iK}(t)\, dZ(t,T) \right) \\ &= \sum_{i,j=1, j\neq i}^{K-1} \left(Z_j(t,T) - Z_i(t,T) \right) dH_{ij}(t) + \sum_{i=1}^{K-1} \left(\delta Z(t,T) - Z_i(t,T) \right) dH_{iK}(t) \\ &\quad + \sum_{i=1}^{K-1} \left(H_i(t)\, dZ_i(t) + \delta\, H_{iK}(t)\, dZ(t,T) \right) \end{aligned}$$

with the initial condition:

$$\overline{Z}(0,T) = \sum_{i=1}^{K-1} H_i(0) \cdot Z_i(t,T).$$

We can simplify this formula considerably using the lemma from paragraph 6.2 and the HJM results from paragraph 6.1. The first result is:

$$dM_{ij}(t) = dH_{ij}(t) - \lambda_{ij}(t)\, H_i(t)\, dt$$

The second, from the HJM model:

$$\frac{d\,Z(t,T)}{Z(t,T)} = -\sigma^*(t,T)^T \; d\widetilde{W}_t \qquad \frac{d\,Z_i(t,T)}{Z_i(t,T)} = \ell_i(t,T)\,dt - \sigma_i^*(t,T)^T \; d\widetilde{W}_t \tag{6.9}$$

We can now simplify the expression for the drift term of the discounted defaultable bond price process $d\overline{Z}(t,T)$:

$$\sum_{i,j=1,j\neq i}^{K-1} \big(Z_j(t,T) - Z_i(t,T)\big)\cdot \lambda_{ij}(t)\, H_i(t)\, dt +$$
$$\sum_{i=1}^{K-1} \big(\delta Z(t,T) - Z_i(t,T)\big)\cdot \lambda_{iK}(t)\, H_i(t)\, dt +$$
$$\sum_{i=1}^{K-1} H_i(t)\cdot \ell_i(t,T)\cdot Z_i(t,T)\, dt$$

This drift term must be 0 a.s. under $\mathbb{Q}$, as $\overline{Z}(t,T)$ must be a martingale under the pricing measure $\mathbb{Q}$. The following conditions for i = 1,... K-1, and for $t \in [0,T]$[20] must therefore be imposed:

$$\sum_{\substack{j=1\\ j\neq i}}^{K-1} \lambda_{ij}(t)\big(Z_j(t,T) - Z_i(t,T)\big) + \lambda_{iK}(t)\big(\delta\, Z(t,T) - Z_i(t,T)\big) + \ell_i(t,T)\cdot Z_i(t,T) = 0 \tag{6.10}$$

If the transition intensities satisfy (6.10), $\overline{Z}(t,T)$ is indeed a martingale and therefore the discounted price process of a traded asset. Note that condition (6.10) generally does not uniquely determine the transition intensities under $\mathbb{Q}$.
Finally, we have obtained the price of the defaultable zero-coupon bond at time t:

$$\overline{P}(t,T) = B_t \cdot \mathbb{E}_t^{\mathbb{Q}}[\delta\, B_T^{-1}\, \mathbb{1}_{[\tau\leq T]} + B_T^{-1}\, \mathbb{1}_{[\tau>T]}] \ ,$$

where $\tau = \inf\left\{t \in \mathbb{R}_+ \,\middle|\, X_t = K\right\}$ indicates the default time. This expression is similar to the one we obtained in the JLT framework, with the difference that here the interest rate process is in general not independent of the default process. As before, we can switch to the T-forward measure $\mathbb{P}^T$ (using the default-free T-bond as numéraire) to obtain an easier expression:

$$\overline{P}(t,T) = P(t,T)\cdot \mathbb{E}_t^{\mathbb{P}^T}[\delta\, \mathbb{1}_{[\tau\leq T]} + \mathbb{1}_{[\tau>T]}] = P(t,T)\cdot\big(\delta + (1-\delta)\cdot \mathbb{P}^T(\tau > T)\big)$$

The downside of this approach however is, that we have to rederive some of the dynamics to obtain a revised version of condition (6.10). In order to exclude arbitrage, we require that under $\mathbb{P}^T$ the money market account, discounted by the T-

[20] Note that this condition is incorrectly stated in Bielecki and Rutkowski [2000a] and the preprint of Bielecki and Rutkowski [2001] that we received.

bond, is a martingale. We can now derive that the market price of interest rate risk must satisfy:

$$\alpha^*(t,T)+\tfrac{1}{2}\left\|\sigma^*(t,T)\right\|^2+\sigma^*(t,T)^T\cdot\gamma_t=0$$

Since we are now using a different numéraire, we have to redefine:

$$Z(t,T)=P(t,T)^{-1}P(t,T)=1 \qquad Z_i(t,T)=P(t,T)^{-1}P_i(t,T)$$

The drift of $Z_i(t,T)$ under the T-forward measure must now equal:

$$\begin{aligned}\ell_i(t,T)\equiv{}& f_i(t,t)-f(t,t)-\alpha_i^*(t,T)+\tfrac{1}{2}\left\|\sigma_i^*(t,T)\right\|^2-\sigma_i^*(t,T)^T\cdot\gamma_t\\ &+\tfrac{1}{2}\left\|\sigma_i^*(t,T)\right\|^2+\tfrac{1}{2}\sigma_i^*(t,T)^T\sigma^*(t,T)\end{aligned}$$

Condition (6.10) now becomes:

$$\sum_{\substack{j=1\\ j\neq i}}^{K-1}\lambda_{ij}(t)\big(Z_j(t,T)-Z_i(t,T)\big)+\lambda_{iK}(t)\big(\delta-Z_i(t,T)\big)+\ell_i(t,T)\cdot Z_i(t,T)=0 \qquad (6.10')$$

Bielecki and Rutkowski remark that calibrating the model to market data is currently under investigation.

6.4. Pricing defaultable coupon bonds

Consider a defaultable coupon bond with face value F that matures at time T and has coupons C_i at times t_i (i = 1,..., N, $t_N = T$). Under the recovery of treasury assumption, the value additivity property holds, so that we have the following value of the bond:

$$V(t,T)=\sum_{i=1}^{N}C_i\cdot\overline{P}^{X_t}(t,t_i)+F\overline{P}^{X_t}(t,T)$$

Bielecki and Rutkowski in their article assume that for coupon bonds, the recovery of face value holds. As can be seen in condition (6.10), the rating transitions depend on the maturity of the bond under consideration and on the recovery rate. If we write this possible dependence as $X_t(t_i,\delta)$, then the value of the bond becomes:

$$V(t,T)=\sum_{i=1}^{N}C_i\cdot\overline{P}^{X_t(t_i,0)}(t,t_i)+F\overline{P}^{X_t(T,\delta)}(t,T)$$

However, according to Bielecki and Rutkowski this dependence on the recovery rate and the maturity of the bond under consideration, implies that each zero coupon component has its own ratings process. From a practical point of view, this does not seem to be sensible. A way to circumvent this is to consider a defaultable coupon bond as a nondivisible asset, and to introduce its own ratings process. Another justification for this is that in the market for defaultable debt, zero-coupon bonds are rarely traded. In our database with straight Eurobonds only 2% of the AAA-rated, 1.6% of the AA-rated issues and 0.3% of the A-rated issues were zero-coupon bonds.

6.5. Calibrating the model to market data

For notational reasons, we will assume from now on that we use the pricing measure obtained by using the money market account as the numéraire asset. The discussion however also holds for other pricing measures.
As we remarked before, a significant difference with the defaultable HJM approach presented by Bielecki and Rutkowski and the traditional HJM approach of modelling the term structure is that we here need to specify the drift of the default-free and defaultable forward rates under the real-world probability measure $\mathbb{P}$. Therefore, the first two steps for any calibration typically are:

1. Specify $\alpha(t,T)$, $\sigma(t,T)$, $\alpha_i(t,T)$ and $\sigma_i(t,T)$ for $i = 1, ..., K\text{-}1$.
2. Estimate the parameters of the previously specified functions using historical data, implied volatilities from quoted bond prices, or a combination of both.

Given the observed zero-coupon bond prices at time t, condition (6.10) imposes necessary conditions on the transition intensities. As we have stated before, this condition (6.10) does not in general uniquely specify the transition intensities. With regard to (6.10) we can also remark that if defaultable zero-coupon bonds with the same rating process are traded for more than K-1 different maturities, there will exist no transition intensities that satisfy (6.10) in general. This implies that there will be arbitrage possibilities in the market.
If equation (6.10) does not uniquely specify the transition probabilities, will the relationship between the equivalent probability measures $\mathbb{P}$ and $\mathbb{Q}$ tell us anything? Bielecki and Rutkowski show that we have the following relationship:

$$\lambda_{ij}^{\mathbb{P}}(t) = \phi_{ij}(t) \cdot \lambda_{ij}^{\mathbb{Q}}(t) \quad \text{for } i \neq j$$

where $\phi_{ij}(t)$ for any $i \neq j$ is an arbitrary nonnegative predictable process. Unfortunately, this does not give us any extra information. Therefore, in order to be able to practically implement this model, we might have to, as in the discrete-time models we studied earlier, make extra assumptions about the structure of the transition intensities. In the continuous version of the JLT model for example, the assumption is made that $\phi_{ij}(t) = \phi_i(t)$, again a column-independent risk premium.

We could also specify a stochastic process for the off-diagonal transition intensities that partially depends on the Brownian motions that determines the evolution of the default-free and the defaultable forward rates. We must ensure that this process stays positive. For an example of how this is done in a one rating setting, we refer the reader to Driessen [2001]. In this article the evolution of the default intensity is specified using a square-root process.
Suppose that, through extra structure, we are able to determine the behaviour of the transition intensities over time. We can then numerically solve the forward Kolmogorov equation to determine the rating transition matrices under the chosen pricing measure:

$$\frac{dA(t,s)}{ds} = A(t,s) \cdot \mathbb{E}_t^{\mathbb{Q}}[\Lambda_s]$$

This will enable us to price all types of credit derivatives.

6.6. Conclusions

In this chapter we have reviewed the defaultable HJM model with credit ratings presented by Bielecki and Rutkowski. Unfortunately the calibration of the model is much harder than in the original HJM model. The reasons for this are threefold. Firstly, the drift functions of both the default-free and the defaultable forward rates must be specified in this model.

Secondly, although the assumption of no arbitrage imposes certain conditions on the transition intensities, it does not uniquely specify them. Therefore, it is not clear how to model the transition intensities under the pricing measure.

Finally, even if we have modelled the transition intensities appropriately, any model will typically have to be calibrated on market prices. As the resulting pricing equations are very similar to the JLT model, we expect that the results from chapter 4 carry over. We expect that the products that are currently traded do not convey enough information about the transition intensities between non-default states in order to be able to calibrate the model for the pricing of rating-dependent products.

7. Conclusions and recommendations

In this thesis we reviewed the rating transition based pricing model of Jarrow, Lando and Turnbull. We investigated the risk premium assumption they make in their article and looked at possible extensions to this assumption, the so-called column-independent risk premiums. Due to several problems that can arise when calculating this particular type of risk premiums, we decided to turn to a simpler model. The model assumes that the rating transitions under the pricing measure follow a time-homogenous Markov chain model. Furthermore, due to some empirical facts on rating transitions that we investigated in the second chapter, we were able to impose a structure on the underlying generator matrix of the Markov chain, thereby reducing the amount of parameters considerably. This model is described in the fourth chapter.
Consequently we investigated the ability of this simple model to replicate market prices of bonds. Since straight bonds are the simplest type of credit derivatives, it is definitely worthwhile to check how well market prices are recovered. The model performed slightly worse than the traditional term structure estimation methods, such as the Svensson model and the B-spline model. We noticed that there was quite some uncertainty about the exact location of the parameters, in particular for the non-default probabilities, which are needed if we want to price rating-dependent credit derivatives. This could be caused by a misspecification of the model, but definitely also by the fact that corporate bonds just depend on default probabilities, which do not convey enough information about non-default transition probabilities.
To improve the estimates, we suggested two possible approaches:

- Using a structural model with rating transitions, hereby endogenously specifying the risk premiums;
- Including more complex products in the calibration process, that explicitly depend on rating transitions.

The first suggestion was not pursued further in this thesis. For the second suggestion, the best candidates seem to be the rating-dependent credit derivatives themselves. In the penultimate chapter however we found that it is quite hard to strip out the prices of the embedded credit derivatives from the bond issues in which they are sold. As their impact on the total price is quite small, including the price of full issue in the calibration process would not improve our estimates significantly.
In the penultimate chapter we turned to the goal of this thesis, that of pricing bond issues with rating-dependent coupons. In the framework in which we were working it is relatively easy to derive the price of the embedded credit derivatives in the bond issues with rating-dependent coupons. We considered three types of embedded downgrade puts:

- a regular downgrade put, which pays off 1 if the rating at the maturity date is lower than a prespecified rating;
- a one-off down-and-in downgrade put, which pays off 1 if the rating at a certain review date is lower than a prespecified rating;
- a continuously reviewed down-and-in downgrade put, which pays off 1 at the maturity date if the rating at any time during the life of the put has been below a prespecified rating.

Using linear combinations of these downgrade puts, the issues with rating-dependent coupons can easily be recreated.

In a numerical example we examined one of the rating-dependent KPN issues, which has a portfolio of embedded regular downgrade puts. Unfortunately, the estimated transition probabilities under the pricing measure seem to vary quite considerably, dependent on the recovery rate assumption. Also, due to the uncertainty about the location of the parameters, there is also a large uncertainty about the exact price of the rating-dependent credit derivatives.

Finally, the thesis concludes with an examination of a Heath, Jarrow and Morton-like model by Bielecki and Rutkowski, which includes the defaultable term structure and credit migration. Again, the great difficulty with this model seems to be the calibration. The resulting pricing equations are very similar to those from the Jarrow, Lando and Turnbull model, so that we do not expect that this model can be calibrated sufficiently either.

Concluding we can state that, given the current products in the credit market, it seems quite hard to price rating-dependent products with a rating-based model, such as the Jarrow, Lando and Turnbull model. If more rating-dependent products are introduced and are traded more frequently, we might be able to generate better results. For further research it may be interesting to look at structural models which include rating transitions. In this way the risk premiums will be specified endogenously, rendering the calibration process a lot easier.

Appendix A - Mathematical finance

In this appendix we briefly summarise some results from the area of mathematical finance. We merely mean to introduce some concepts so that we can use them throughout this thesis. For more information on mathematical finance and derivative pricing in particular we refer you to excellent references such as Baxter and Rennie [1996] and Musiela and Rutkowski [1998].

A.1. Derivative pricing

We shall treat the topics in this paragraph in a continuous setting. Their discrete analogues can easily be formulated. The market under consideration usually exists out of some traded stocks and a *money market account*, which usually satisfies the following stochastic differential equation:

$$dB(t) = r(t)B(t)\,dt$$

The money market account can be viewed as a bank account: one unit of currency invested in it will accumulate interest. Specifically, at time t the so-*called spot interest rate* is paid: $r(t)$.

We call a portfolio *self-financing* if and only if the change in its value only depends on the change of the asset prices. For example, suppose we have a portfolio V, which exists of an investment in a stock S, and the money market account B:

$$V(t) = \phi(t)\, S(t) + \psi(t)\, B(t)$$

Mathematically, the portfolio is self-financing if and only if:

$$dV(t) = \phi(t)\, dS(t) + \psi(t)\, dB(t)$$

An *arbitrage* possibility on a financial market is a self-financed portfolio V such that:

$$V(0) = 0$$

$$V(T) \geq 0 \quad \text{a.s.}$$

$$\mathbb{P}\big(V(T) > 0\big) > 0$$

The market is *arbitrage free* if there are no arbitrage possibilities. An arbitrage possibility is thus equivalent to the possibility of making a non-negative amount of money out of nothing with probability 1. We say that a claim at time T, $X \in \mathcal{F}_T$ (the information set up to time T), can be *replicated*, alternatively that it is *reachable* or *hedgeable*, if there exists a self-financing portfolio V such that:

$$V(T) = X \quad \text{a.s.}$$

In this case we say that the portfolio is a *hedge* against X. Alternatively, the portfolio is a *replicating* or *hedging* portfolio. If every claim is reachable we say the market is *complete*.

Suppose we have a market of securities and a money market account under a probability measure $\mathbb{P}$. An *equivalent martingale measure* (*EMM*), or *risk-neutral probability measure* is a probability measure $\mathbb{Q}$ which is equivalent (or absolutely continuous) with respect to $\mathbb{P}$ and has the property that all securities, divided by the money market account ($S(t)/B(t)$ in the previous notation), under this probability measure $\mathbb{Q}$ are martingales.

Arbitrage-free and completeness theorem (Harrison and Pliska [1981])
Suppose we have a market of securities and a numeraire bond. Then:

(1) the market is arbitrage-free if and only if there is at least one EMM;
(2) in which case, the market is complete if and only if there is one unique, EMM.

If there exists an EMM we have:

$$\frac{V(t)}{B(t)} = \mathbb{E}^{\mathbb{Q}}\left[\frac{V(T)}{B(T)} \mid \mathcal{F}_t\right] = \mathbb{E}_t^{\mathbb{Q}}\left[\frac{V(T)}{B(T)}\right]$$

From this the price of the claim at time t can be calculated.

A.2. Interest-rate related concepts

In the previous paragraph we introduced some concepts from mathematical finance that are used when calculating the price of a derivative. In this paragraph we briefly introduce some concepts related to the interest rate market.
Apart from the already introduced money market account, we also *have zero coupon bonds*. A default-free zero coupon bond with maturity date T, pays out one unit of currency for certain at time T. The time t price of this security is denoted by $P(t,T)$. The *term structure* of zero coupon bonds simply describes $P(t,T)$ for all possible values of T. The *forward rates* are implicitly defined through:

$$P(t,T) = \exp(-\int_t^T f(t,s)\,ds)$$

or in a discrete time setting:

$$P(t,t+n\triangle) = \exp(-\sum_{i=0}^{n-1} f(t,t+i\triangle))$$

The spot rate r(t) simply equals $f(t,t)$. Another concept we shall use is the yield of a zero coupon bond, which is implicitly defined via:

$$P(t,T) = \exp\left(-R(t,T)\cdot(T-t)\right)$$

In the debt market we usually find traded zero coupon bonds, but also coupon bonds. A default-free coupon bond is a security which pays off a coupon C at times t_1 to t_n and has face value F. It can be shown that the price of such a coupon bond equals:

$$V(t,T) = \sum_{i=1}^{n} C\cdot P(t,t_i) + F\cdot P(t,T)$$

Appendix B - Term structure estimation

A vast amount of coupon bonds are traded in the debt market. Suppose we are faced with the problem of determining the price of a zero coupon bond within a certain class, say the class of all Euro-denominated bonds with an AA-rating. How do we go about this? Provided we have exactly one bond issue for each possible maturity date and assuming that the value-additivity property of coupon bonds holds, it is easy to infer the price of a zero coupon bond. Unfortunately this is never the case, so that we have to resort to more advanced methods. We will discuss two possible methods of estimating default-free term structures and credit spreads in the following paragraphs. The first paragraph deals with a parametric model of the yield curve, the Svensson model. In the second paragraph we describe how to estimate the discount curve using B-splines.

B.1. The Svensson model

Nelson and Siegel [1987] originally developed a parametric model of the instantaneous forward rate curve. An extension to this model was developed by Svensson [1994] and can capture a wider range of forward curves. Our brief discussion of both models is largely based on a practical discussion of the models in Bolder and Stréliski [1999].

In the Svensson model, the instantaneous forward rate $f(t,T)$ equals:

$$f(t,T) = \beta_0 + \beta_1 \cdot \exp(-\tfrac{\tau}{\tau_1}) + \beta_2 \cdot \tfrac{\tau}{\tau_1} \cdot \exp(-\tfrac{\tau}{\tau_1}) + \beta_3 \cdot \tfrac{\tau}{\tau_2} \cdot \exp(-\tfrac{\tau}{\tau_2}) \qquad \text{(B.1)}$$

where $\tau = T - t$ and the parameters can be interpreted as follows:

- $\beta_0 > 0$: the asymptotic value of the forward rate curve;
- β_1: The vertical intercept of the forward rate curve is $\beta_0 + \beta_1$; furthermore, this parameter determines the speed with which the curve tends to its asymptote;
- $\tau_1 > 0$: This parameter specifies the position of the first hump or U-shape;
- β_2: If positive, a hump will occur at τ_1, otherwise a U-shaped form will occur;
- $\tau_2 > 0$: This parameter specifies the position of the second hump or U-shape;
- β_3: If positive, a hump will occur at τ_2, otherwise a U-shaped form will occur.

The Nelson-Siegel model for the forward rate curve is obtained by letting β_3 equal 0. Using (B.1) and the implicit definition of a forward rate given in the previous appendix, allows us to find an explicit form for a zero coupon bond price under the Svensson model. If we furthermore assume that the value additivity property holds, we have an analytical expression for the price of coupon bonds. Suppose now that we have N bonds (i = 1 ,..., N). In order to estimate the parameters in (B.1), we solve:

$$\min_{\beta_0,\beta_1,\beta_2,\beta_3,\tau_1,\tau_2} \sum_{i=1}^{N} w_i \cdot (V_{i,\text{market}} - V_{i,\text{Svensson}})^2$$

subject to the restrictions placed on the parameters. This is a typical non-linear least squares problem, which can be solved with methods in e.g. Greene [1997].

B.2. Modelling the discount function with B-splines

The model in the previous paragraph can be used for instance when the number of available bonds is small. A more flexible approximation, yet harder to calibrate, to the term structure can be found by using B-splines. One possible way of achieving this is discussed in Houweling, Kleibergen and Hoek [2001]. We briefly discuss their model here, for more details we refer to their paper. *Central Market Risk* currently uses this approach to estimate the default-free term structure and credit spreads. For more information on splines we refer you to Powell [1981].

A spline is a piecewise polynomial. The approximation interval $[a,b]$ (for our application this runs from 0 to the longest bond maturity in the sample) is divided into n subintervals $[\tau_{i-1}, \tau_i]$ for $i = 1, \ldots, n$, $\tau_0 = a$ and $\tau_0 = b$. The function which we are trying to approximate is modelled as a k^{th} degree polynomial in each subinterval. Furthermore, the polynomials are constrained by the condition that they must be k-1 times continuously differentiable. This imposes k constraints on the coefficients of two adjacent polynomials. Summing up, we have $n(k+1)$ coefficients and $(n-1)k$ constraints, which leaves us with $n+k$ degrees of freedom.

A way of representing splines is through basis functions. Any k^{th} degree spline function can be expressed as a linear combination of $n+k$ basis functions. We will now show how to do this for a particular type of splines, namely B-splines.

We will now give a brief description of B-splines. Given the $n+1$ knots we specified earlier, a k^{th} degree B-spline function is defined as:

$$B_s^k(t) = \sum_{\ell=s}^{s+k+1} \prod_{h=s,h\neq\ell}^{s+k+1} \frac{1}{\tau_h - \tau_\ell} \cdot \max(t-\tau_\ell, 0)^k$$

where s indicates that the B-spline is only non-zero if $t \in [\tau_s, \tau_{s+k+1}]$. In order to construct a basis, we need $n+k$ linearly independent B-splines. A k^{th} order B-spline is only non-zero in $k+1$ subintervals, so within the interval $[\tau_0, \tau_n]$ only $n-k$ B-splines are defined. One possible way to construct the additional $2k$ splines is by introducing extra knots as follows:

$$\begin{aligned} \tau_i &= \tau_0 + i\cdot(\tau_1 - \tau_0) && i = -k,\ldots,-1 \\ \tau_i &= \tau_n + (i-n)\cdot(\tau_n - \tau_{n-1}) && i = n+1,\ldots,n+k \end{aligned}$$

A basis of $n+k$ B-splines then consists of B_s^k for $s = -k, \ldots, n-1$.

In Houweling et al. the prices of zero coupon bonds are modelled as follows:

$$\begin{aligned} P(t,T) &= g_0(\tau)^T \beta_0 \\ \overline{P}^i(t,T) &= g_0(\tau)^T \beta_0 + g_i(\tau)^T \beta_i \qquad i = 1,\ldots,K-1 \end{aligned} \tag{B.2}$$

where $g_i(\tau)$ contains $n_i + k_i$ B-spline basis functions and $\tau = T - t$. If we assume that the value additivity property holds for coupon bonds, the value of a coupon bond can now be expressed in terms of (B.2), except for the unknowns β_i.

As in the previous paragraph we have market prices of bonds. In Houweling et al. the following linear model is postulated for the market price of bond *j* in rating class *i*:

$$V^{i}_{j,\text{market}} = V^{\text{i}}_{j,\text{B-spline}} + \varepsilon^{i}_{j} \tag{B.3}$$

where the error terms $\varepsilon^{i}_{j} \sim$ i.i.d. $(0,\sigma^{2}_{i})$. The error terms have different variances for each rating category, due to the empirical fact that lower rated bonds are noisier due to lower liquidity and a higher uncertainty about their creditworthiness.
Finally, one restriction must be imposed on the parameters in (B.2). The value of zero coupon bonds at maturity must equal one, hence:

$$\begin{aligned} P(t,t) &= g_0(0)^T \beta_0 = 1 \\ \overline{P}^{i}(t,t) &= g_0(0)^T \beta_0 + g_i(0)^T \beta_i = 1 \qquad i = 1,\ldots,K-1 \end{aligned}$$

These restrictions are linear in the parameters and can easily be incorporated when estimating (B.3). The model can now be estimated using restricted feasible generalised least squares (RFGLS), for which we refer to Greene [1997]. In a nutshell, RFGLS boils down to:

1. Estimate (B.3) using restricted ordinary least squares (ROLS); using the estimated error terms, we can obtain consistent estimates $\hat{\sigma}^{2}_{i}$ for the variances of the error terms within each rating class.
2. Using the estimated covariance matrix from step 1, we estimate (B.3) using restricted generalised least squares (RGLS).

Due to the fact that the default-free discount curve and the spread curves are jointly estimated, the reliability of the estimates is improved. However, working with splines is a tricky subject area. The knot and degree specifications must be chosen with great care. Furthermore, as always, great care must be taken when calculating zero coupon bond prices for maturity ranges that lie outside the sample width.
Note that weights can easily be incorporated in this estimation procedure. We then just need to estimate the model with weighted restricted feasible generalised least squares (WRFGLS).

Appendix C – Non-linear least squares

Assume we are considering the following regression model:

$$y_i = f(x_i, \beta) + \varepsilon_i \qquad\qquad i = 1,\ldots,n$$
$$\varepsilon_i \sim \text{iid}\,(0, \sigma^2)$$

where $y_i \in \mathbb{R}$, $x_i \in \mathbb{R}^m$ and $\beta \in \mathbb{R}^k$. The function *f* is a non-linear function of the parameter vector β. Analogous to the case when *f* is a linear function of the parameters, we will try to minimise the sum of least squares in order to estimate the parameter vector β:

$$S(\beta) = (y - f(x,\beta))^T (y - f(x,\beta))$$

In order to simplify the notation, we have dropped the observation subscript.

Computing the non-linear least squares estimator

There are many ways to solve non-linear optimisation problems. The most common algorithms are based on the following iterative algorithm:

- Initial estimate $\hat{\beta}_0$;
- If at iteration *t* $\hat{\beta}_t$ is not a local minimum for $S(\beta)$, compute the direction vector Δ_t and a step size λ_t. Update the estimate using: $\hat{\beta}_{t+1} = \hat{\beta}_t + \lambda_t \Delta_t$.

Gradient methods choose the direction vector as follows:

$$\Delta_t = -D_t \frac{\partial S(\hat{\beta}_t)}{\partial \hat{\beta}_t}$$

where D_t is a positive definite matrix. We then still have to specify how to choose this matrix D_t and the step size. Newton's method, which is based on a linear Taylor series approximation, uses:

$$D_t = \frac{\partial^2 S(\hat{\beta}_t)}{\partial \hat{\beta}_t} \qquad\qquad \lambda_t = 1$$

Newton's method finds the optimum in one iteration from any starting point if the function is quadratic. However, if it is not, the step size may not be the appropriate step size and furthermore, the computational burden of computing the Hessian may be excessive. A quite effective class of algorithms, that does not use second order derivatives are quasi-newton methods, which use the following updating scheme:

$$D_{t+1} = D_t + E_t$$

where E_t is a positive definite matrix. We will briefly discuss the Davidon-Fletcher-Powell (DFP) method, as this is the method we have used in this thesis.

The DFP algorithm uses the following updating scheme:

$$D_{t+1} = D_t + \frac{p_t p_t^T}{p_t^T q_t} - \frac{D_t q_t q_t^T D_t}{q_t^T D_t q_t}$$

where

$$p_t = \lambda_t \Delta_t = \hat{\beta}_{t+1} - \hat{\beta}_t$$
$$q_t = \frac{\partial S(\hat{\beta}_{t+1})}{\partial \hat{\beta}_{t+1}} - \frac{\partial S(\hat{\beta}_t)}{\partial \hat{\beta}_t}$$

The optimal step size λ_t is determined via a standard line minimisation. The algorithm is initialised with $D_0 = I$ and an initial estimate $\hat{\beta}_0$. As a termination criterion, the L_2-norm of the gradient vector can be used.
For a more detailed discussion of the DFP algorithm in particular and non-linear optimisation methods in general we refer you to Bazaraa, Sherali and Shetty [1993]. A good numerical discussion can be found in Press, Teukolsky, Vetterling and Flannery [1996].

Asymptotic properties of the non-linear least squares estimator

Let us linearise our previous regression model at the point β^0, which is the true parameter vector:

$$y = f(x,\beta^0) + \frac{\partial f(x,\beta^0)}{\partial \beta}^T (\beta - \beta^0) + O(\beta^2) + \varepsilon$$

Rearranging terms:

$$y - f(x,\beta^0) + \frac{\partial f(x,\beta^0)}{\partial \beta}\beta^0 \doteq \frac{\partial f(x,\beta^0)}{\partial \beta}^T \beta + \varepsilon$$

For the technical details, we refer you to Davidson and MacKinnon [1993]. As Greene [1997] remarks, it turns out that the asymptotic results are based on the linearised version of our regression model. We assume:

$$\frac{1}{n}\frac{\partial f(x,\beta^0)}{\partial \beta}^T \frac{\partial f(x,\beta^0)}{\partial \beta} \xrightarrow{\text{P}} Q(\beta^0) \qquad \text{(C.1)}$$

where $Q(\beta^0)$ is a positive definite matrix. Furthermore, provided that:

$$\frac{1}{n}\frac{\partial f(x,\beta^0)}{\partial \beta}^T \varepsilon \xrightarrow{\text{P}} 0, \qquad \text{(C.2)}$$

the non-linear least squares estimator is consistent.

Asymptotic normality can be established if:

$$\frac{1}{n}\frac{\partial f(x,\beta^0)}{\partial\beta}^T \varepsilon \xrightarrow{d} N\left(0,\sigma^2 Q(\beta^0)\right) \tag{C.3}$$

In this case, the limiting distribution of the non-linear least squares estimator is:

$$\hat{\beta} \xrightarrow{d} N\left(\beta^0, \frac{\sigma^2}{n} Q(\beta^0)^{-1}\right) \qquad\qquad n \to \infty$$

The sample asymptotic covariance matrix is:

$$\hat{\sigma}^2\left(\frac{\partial f(x,\hat{\beta})}{\partial\beta}^T \frac{\partial f(x,\hat{\beta})}{\partial\beta}\right)^{-1} \text{,where } \hat{\sigma}^2 = \tfrac{1}{n}S(\hat{\beta}).$$

For a function g of the parameters, we can derive the following asymptotic distribution, using the delta method:

$$g(\hat{\beta}) \xrightarrow{d} N\left(g(\beta^0), \frac{\sigma^2}{n}\frac{\partial g}{\partial\beta} Q(\beta^0)^{-1}\frac{\partial g}{\partial\beta}^T\right) \qquad n \to \infty$$

Another, maybe more appropriate way to derive confidence intervals of a function of the parameters in a small sample, can be obtained if the convergence of the parameters to a normal distribution is quite speedy. In this case, we can draw random samples from the distribution of the parameter vector and evaluate the functions at these points to obtain a sample from the distribution of g. This random sample can be used to estimate any confidence interval. An advantage of this method is that the proper bounds for the function will be satisfied, whereas the delta method may yield nonsensical results if the confidence interval is quite wide.
We will now briefly cover two topics, which can be treated similarly to the linear model. Firstly, we consider heteroskedasticity.

Heteroskedasticity
We consider the following specification for the errors:

$$\varepsilon_i \sim \text{iid}(0,\sigma_i^2)$$

Define $\sigma^2\Omega$ as the diagonal matrix containing the variances on the diagonal. If we impose the following condition:

$$\frac{1}{n}\frac{\partial f(x,\beta^0)}{\partial\beta}^T \Omega \frac{\partial f(x,\beta^0)}{\partial\beta} \xrightarrow{P} Q(\Omega,\beta^0) \tag{C.1'}$$

and both (C.1) and (C.2) hold, NLS is consistent and unbiased.

The limiting distribution now becomes:

$$\hat{\beta} \xrightarrow{d} N\left(\beta^0, \frac{\sigma^2}{n} Q(\beta^0)^{-1} Q(\Omega, \beta^0)\, Q(\beta^0)^{-1}\right) \qquad n \to \infty$$

A consistent estimator of the asymptotic covariance matrix under heteroskedasticity is the White estimator, which equals:

$$n \cdot \left(\frac{\partial f(x,\hat{\beta})}{\partial \beta}^T \frac{\partial f(x,\hat{\beta})}{\partial \beta} \right)^{-1} S_0 \left(\frac{\partial f(x,\hat{\beta})}{\partial \beta}^T \frac{\partial f(x,\hat{\beta})}{\partial \beta} \right)^{-1}$$

where $S_0 = \frac{1}{n} \frac{\partial f(x,\hat{\beta})}{\partial \beta}^T \operatorname{diag}\left((y - f(x,\beta))(y - f(x,\beta))^T \right) \frac{\partial f(x,\hat{\beta})}{\partial \beta}$.

Weighted non-linear least squares

Suppose we have heteroskedastic errors, as in the previous subparagraph, but that in addition we want to weight each observation with a weight w_i. Let W be the diagonal matrix containing the weights w_i on the diagonal. Now consider:

$$Wy = Wf(x,\beta) + W\varepsilon \qquad\qquad \varepsilon \sim (0, \sigma^2 \Omega)$$

or equivalently:

$$y^* = f^*(x,\beta) + \varepsilon^* \qquad\qquad \varepsilon^* \sim (0, \sigma^2 W^2 \Omega)$$

where the superscript star indicates that the original function has been premultiplied with the matrix W. The transformed problem can be treated in the same framework as the previous subparagraph, as we have heteroskedastic errors here as well.

Appendix D - Various recovery rate assumptions

The use of various recovery rate assumptions has various implications for the prices of defaultable coupon bonds. In this appendix we will present the prices of coupon bonds under the recovery rate assumptions that we have used within this thesis, merely for reference purposes. For a discussion about the various recovery rate assumptions we refer you to paragraph 2.3.2. All formulas are based on the JLT framework as discussed in paragraph 3.1. Assumptions about the credit migration are not necessary.

D.1. Recovery of treasury

Under the recovery of treasury assumption, the recovery amount is based on all outstanding cashflows. Each claim is replaced by a fraction δ of an otherwise equivalent default-free claim. The price of a defaultable zero-coupon bond under this assumption is equivalent to:

$$\overline{P}(t,T) = P(t,T) \cdot \left(\delta + (1-\delta) \cdot \mathbb{Q}_t(\tau > T)\right) \tag{D.1}$$

The price of a coupon bond with coupon dates $t_1, \ldots, t_N = T$, coupon C and face value F therefore equals:

$$\overline{V}(t,T) = \sum_{i=1}^{N} C \cdot \overline{P}(t,t_i) + F \cdot \overline{P}(t,T) \tag{D.2}$$

where we use the convention that $\overline{P}(t,s) = 0$ if $s < t$.

D.2. Recovery of face value

The recovery amount under this assumption is solely based on the face value of the coupon bond. The price of a zero-coupon therefore remains the same as under the previous assumption, however, the price of a defaultable coupon bond is in general no longer equal to a portfolio of defaultable zero-coupon bonds. The price also depends on the timing of the recovery payment.

D.2.1. Payoff at maturity

If the recovery amount is paid at maturity, the price of a coupon bond simply equals:

$$\overline{V}(t,T) = \sum_{i=1}^{N} C \cdot P(t,t_i) \cdot \mathbb{Q}_t(\tau > t_i) + F \cdot \overline{P}(t,T) \tag{D.3}$$

The coupon bond is therefore equal to a portfolio of zero-coupon bonds with recovery rate equal to zero, and a zero-coupon bond for the face value payment, which has a recovery rate equal to δ.

D.2.2. Payoff at default

If the recovery amount is paid at default, we have a slightly different situation. The price of a defaultable zero-coupon bond then becomes:

$$\begin{aligned}\overline{P}(t,T) &= \mathbb{E}_t^{\mathbb{Q}}\left[\tfrac{B(t)}{B(T)}\cdot \mathbb{1}_{[\tau>T]} + \tfrac{B(t)}{B(\tau)}\cdot\delta\cdot\mathbb{1}_{[\tau\le T]}\right] \\ &= P(t,T)\cdot\mathbb{Q}_t(\tau>T) + \delta\cdot\mathbb{E}_t^{\mathbb{Q}}\left[\tfrac{B(t)}{B(\tau)}\cdot\mathbb{1}_{[\tau\le T]}\right] \qquad \text{(D.4)} \\ &= P(t,T)\cdot\mathbb{Q}_t(\tau>T) + \delta\cdot\sum\nolimits_{i=1}^{n} P(t,t+i\triangle)\cdot\mathbb{Q}_t(\tau = t+i\triangle)\end{aligned}$$

where $T = t + n\triangle$. The price of a coupon bond is now as under (D.3).

D.3. Legal claim approach

This approach, advocated by Jarrow and Turnbull [2000], assumes that the claim is based on the face value plus accrued interest. The payoff of the recovery amount takes place at the default time.

As zero-coupon bonds have no accrued interest, the price of a zero-coupon bond remains equal to (D.1). However, the price of a defaultable coupon bond changes slightly in comparison with (D.3). We namely have to add a term that accounts for the accrued interest. Let us first assume the bond was issued at time t_0. At any time t we define the time of the previous coupon payment (or issue date if there was no previous coupon payment):

$$t_{\text{prev}}(t) = \min\{t_i \le t \mid i = 0,\ldots,N\}$$

The accrued interest at time t is now equal to:

$$\alpha(t_{\text{prev}}(t),t)\cdot C$$

where $\alpha(t,s)$ is the year fraction according to the proper daycount convention for that bond. We will not get into the details of the various daycount conventions here. For this we refer you to e.g. Fabozzi [1999]. At any rate, the amount that has to be added to (D.3) to obtain the correct bond price under this recovery rate assumption, equals:

$$\begin{aligned}\mathbb{E}_t^{\mathbb{Q}}[\tfrac{B(t)}{B(T)}\cdot\delta\cdot\mathbb{1}_{[\tau\le T]}\cdot\alpha(t_{\text{prev}}(\tau),\tau)] &= \delta\cdot\mathbb{E}_t^{\mathbb{Q}}[\tfrac{B(t)}{B(T)}\cdot\mathbb{E}_\tau^{\mathbb{Q}}[\mathbb{1}_{[\tau\le T]}\cdot\alpha(t_{\text{prev}}(\tau),\tau)]] \\ &= \delta\cdot\sum\nolimits_{i=1}^{n} P(t,t+i\triangle)\cdot\mathbb{Q}_t(\tau=t+i\triangle)\cdot\alpha(t_{\text{prev}}(t+i\triangle),t+i\triangle)\end{aligned}$$

Appendix E - Generator matrices

In this thesis we use the concept of a generator matrix of a Markov chain. We will group all the information we need about generator matrices in this appendix, so it is not scattered throughout the thesis.
We consider a time-homogenous Markov chain with a one-step transition matrix *A*. The generator matrix Λ for this Markov chain is a matrix with row-sums equal to 0 and non-negative off-diagonal entries, such that $\exp(\Lambda) = A$. The exponent of a matrix is defined as the following Taylor series expansion:

$$\exp(t\Lambda) = \lim_{n\to\infty} \sum_{k=0}^{n} (t\Lambda)^k \big/ k! \qquad \text{(E.1)}$$

Alternately, we can write the generator matrix as follows:

$$\Lambda = \lim_{n\to\infty} \sum_{k=1}^{n} -(I-A)^k \big/ k \qquad \text{(E.2)}$$

First we will show that if Λ is a matrix with row-sums equal to 0 and non-negative off-diagonal entries, $\exp(\Lambda)$ is a transition probability matrix.

Lemma E.1
If Λ *is a generator matrix, i.e. has rows that sum to zero and all off-diagonal entries are non-negative, then* $\exp(\Lambda)$ *is a transition probability matrix.*

Proof:
First we will sketch the proof that $\exp(\Lambda)$ yields a matrix with non-negative entries. We can easily see that for sufficiently small positive values of *s*, the following holds:

$$\exp(s\Lambda) \approx I + s\Lambda$$

Entry (i,j) of this matrix will be equal to $1_{[i=j]} + s\lambda_{ij}$. If $i \neq j$ this is obviously non-negative; when $i = j$, this entry is still non-negative, provided that *s* is sufficiently small. Suppose that we now have such a value *s*. For any $t < s$, $\exp(t\Lambda)$ will also only contain non-negative values, in particular for $t = 1 / n$, where $n \in \mathbb{N}$. The n^{th} power of $\exp(t\Lambda)$ simply equals $\exp(\Lambda)$. We can easily check that any integer power of a matrix with non-negative entries yields a matrix with non-negative entries, and therefore $\exp(\Lambda)$ also has non-negative entries.
For the exponent to be a transition probability matrix, we also require that the rows sum to one. Suppose we have K x K matrices *A* and *B* with rows that sum to zero. The sum of any row of their product *AB* will also equal zero:

$$\sum_{j=1}^{K} \sum_{m=1}^{K} a_{im} b_{mj} = \sum_{m=1}^{K} a_{im} \sum_{j=1}^{K} b_{mj} = 0$$

In particular any integer power of Λ will have rows that sum to zero, and therefore from the power expansion (E.1) we see that $\exp(\Lambda)$ will have rows that sum to one, due to the fact that $\Lambda^0 = I$ is included in the summation. □

So, if we have a generator matrix, there is one transition matrix that corresponds with it. However, we are not guaranteed that given a transition matrix, there is a generator matrix that corresponds with it. These problems are dealt with in Israel, Rosenthal and Wei [2000]. We can calculate (E.2), but we are not ensured that this expression will converge or will be a generator matrix. A sufficient condition for convergence is however mostly satisfied in empirical transition probability matrices: the diagonal probabilities must all be larger than 0.5. Therefore we are left with the problem that the resulting logarithm may not be a generator matrix. The row-sums of the resulting matrix in (E.2) will always be equal to zero, so that this can only be caused by possible negative off-diagonal elements.
In the continuous-time version of the Jarrow, Lando and Turnbull model, generator matrices are also used. JLT did not mention this problem, but worked their way around it. Since they worked with one-year transition matrices, they assumed that the probability of making more than one transition per year equals zero, thereby making it easy to determine a functional form of the one-year transition matrix in terms of the elements of the generator matrix. Let us name their resulting modified generator Λ_{JLT}. Israel et al. also considered this problem, and proposed two different algorithms. We will here discuss the simpler of the two, since the performance of both algorithms seems to be quite close. Their algorithm yields a generator matrix Λ' as follows:

$$\lambda'_{ij} = \max(\lambda_{ij}, 0) \qquad i \neq j$$
$$\lambda'_{ii} = \sum_{j \neq i} \lambda'_{ij}$$

For various historical transition matrices supplied by the rating agencies, Israel et al. find that $\left\| \tilde{A} - \exp(\Lambda') \right\|$ is much smaller than $\left\| \tilde{A} - \exp(\Lambda_{JLT}) \right\|$, implying that their algorithm yields a generator matrix which matches the observed probabilities better than the JLT algorithm. Here $\left\| \cdot \right\|$ is the L^1 norm, defined as the sum of the absolute values of the matrix.
This concludes our discussion of generator matrices.

Bibliography

ARVANITIS, A., J. GREGORY AND J-P. LAURENT (1999). "Building Models for Credit Spreads", *Journal of Derivatives* Spring, pp. 27-43.

BAXTER, M. AND A. RENNIE (1996). *Financial Calculus: An Introduction to Derivative Pricing*, Cambridge University Press.

BAZARAA, M.S., SHERALI, H.D. AND C.M. SHETTY (1993). *Nonlinear Programming: Theory and Algorithms* (Second Edition), John Wiley & Sons.

BIELECKI, T. AND M. RUTKOWSKI (1999). "Modelling of the Defaultable Term Structure: Conditionally Markov Approach", Working Paper.

BIELECKI, T. AND M. RUTKOWSKI (2000a). "Multiple ratings model of defaultable term structure", *Mathematical Finance*, vol. 10, no. 2, pp. 125-139.

BIELECKI, T. AND M. RUTKOWSKI (2000b). "HJM with multiples", *Risk*, vol. 13, no. 4, pp. 95-97.

BOLDER, D. AND D. STRÉLISKI (1999). "Yield Curve Modelling at the Bank of Canada", Bank of Canada, Technical Report No. 84.

CONROY, P. (2000). "Valuing ratings-sensitive coupon language", *Credit* July 2000, pp. 62-64.

GUPTON, G.M., C.C. FINGER AND M. BHATIA (1997). "CreditMetrics – Technical Document".

DAVIDSON, J. AND J. MACKINNON (1993). *Estimation and Inference in Econometrics*, Oxford University Press.

DAS, S. (2000). *Credit Derivatives and Credit Linked Notes* (2nd edition), John Wiley & Sons.

DAS, S. AND P. TUFANO (1996). "Pricing Credit Sensitive Debt when Interest Rates, Credit Ratings and Credit Spreads are Stochastic", *Journal of Financial Engineering* vol.5, no. 2, pp. 161-198.

DRIESSEN, J. (2001). "The Cross-Firm Behaviour of Credit Spreads", Working Paper, University of Amsterdam.

DUFFEE, G.R. (1998). "The Relation between Treasury Yields and Corporate Bond Yield Spreads", *Journal of Finance* 53, pp. 2225-2242.

DUFFIE, D. AND K.J. SINGLETON (1999). "Modeling Term Structures of Defaultable Bonds", *Review of Financial Studies*, vol. 12, no. 4, pp. 687-720.

FABOZZI, F.J. (1999). *Bond Markets: Analysis and Strategies* (4th edition), Prentice Hall.

GREENE, W.H. (1997). *Econometric Analysis* (3rd edition), Prentice Hall, New Jersey.

GREGORY-COSTELLO, M. (2000). "Credit limits - How credit protection is changing the bond markets", *Risk* July 2000, pp. 24-26,29.

HARRISON, M. AND S. PLISKA (1981). "Martingales and Stochastic Integrals in the Theory of Continuous Trading", *Stochastic Processes and their Applications* 11, pp. 215-260.

HOUWELING, P. (2000). "Risk Premia in Markov Chain Credit Risk Models", Working Paper, Rabobank International and Tinbergen Institute, Erasmus University Rotterdam.

HOUWELING, P., J. HOEK AND F. KLEIBERGEN (2001). "The Joint Estimation of Term Structures and Credit Spreads", *Journal of Empirical Finance*, vol. 8, no. 3, pp. 297-323.

ISRAEL, R.B., J.S. ROSENTHAL AND J.Z. WEI (2001). "Finding generators for Markov chains via empirical transition matrices, with applications to credit ratings", *Mathematical Finance*, v11: 2.

JARROW, R.A. AND S.M. TURNBULL (1995). "Pricing Derivatives on Financial Securities Subject to Credit Risk", *Journal of Finance*, vol. 50, no. 1, pp. 53-86.

JARROW, R.A., D. LANDO AND S.M. TURNBULL (1997). "A Markov Model for the Term Structure of Credit Risk Spreads", *Review of Financial Studies*, vol. 10, no. 2, pp. 481-523.

JARROW, R.A. AND S.M. TURNBULL (2000). "The intersection of market and credit risk", *Journal of Banking and Finance* 24, pp. 271-299.

KIJIMA, M. AND K. KOMORIBAYASHI (1998). "A Markov Chain Model for Valuing Credit Risk Derivatives", *Journal of Derivatives* Fall, pp. 97-108.

KIJIMA, M. (1997). *Markov Processes for Stochastic Modeling*, Chapman & Hall, London.

LANDO, D. (1998). "On Cox processes and credit risky securities", *Review of Derivatives Research*, vol. 2, no. 2 and 3, pp. 99-120.

LANDO, D. (2000). "Some Elements of Rating-Based Credit Risk Modeling", in: N. Jegadeesh and B. Tuckman (editors): *Advanced Fixed-Income Valuation Tools*, Wiley, pp. 193-215.

LONGSTAFF, F. AND E. SCHWARTZ (1995). "A Simple Approach to Valuing Risky Fixed and Floating Rate Debt", *Journal of Finance* 50, pp. 789-821.

MERRILL LYNCH (2000). "Size & Structure of the World Bond Market: 2000".

MERTON, R. (1974). "On The Pricing of Corporate Debt: The Risk Structure of Interest Rates", *Journal of Finance* 29, pp. 449-470.

MOODY'S (2000). "Historical Default Rates of Corporate Bond Issuers, 1920-1999".

MUSIELA, M. AND M. RUTKOWSKI (1998). *Martingale Methods in Financial Modelling*, Second Printing, Springer-Verlag.

NELSON, C.R. AND A.F. SIEGEL (1987). "Parsimonious Modeling of Yield Curves", *Journal of Business* 60, pp. 473-489.

POWELL, M.J.D. (1981). *Approximation Theory and Methods*, Cambridge University Press.

PRESS, W.H., TEUKOLSKY, S.A., VETTERLING, W.T. AND B.P. FLANNERY (1996). *Numerical Recipes in C* (Second Edition), Cambridge University Press.

SCHÖNBUCHER, P.J. (1998). "Term Structure Modelling of Defaultable Bonds", *Review of Derivatives Research*, vol. 2, no. 2 and 3, pp. 161-192.

SMITH, C. (2000). "Credit Derivatives: A route to emerging markets", *Financial Times*, 19/05/2000.

STANDARD & POOR'S (2001a). "Ratings Performance 2000".

STANDARD & POOR'S (2001b). "Ratings Roundup 2000".

SVENSSON, L.E. (1994). "Estimating and Interpreting Forward Interest Rates: Sweden 1992-1994", Working Paper, Centre for Economic Policy Research, Discussion Paper 1051.

TAVAKOLI, J.M. (1998). *Credit Derivatives: A Guide to Instruments and Applications*, Wiley & Sons.

WEI, J.Z. (2000). "A Multi-Factor, Markov Chain Model for Credit Migrations and Credit Spreads", Working Paper, Rotman School of Management, University of Toronto.