On defaultable claims and credit derivatives.

Brian Huge
Department of Statistics and Operations Research
University of Copenhagen
Universitetsparken 5
DK-2100 Copenhagen Ø
Denmark

Ph.D. thesis
Thesis advisor : David Lando

December 19, 2001

Acknowledgments
First of all I would like to thank my adviser David Lando for inspiring discussions and contributions to this thesis. Further thanks are due to my colleagues at the Department of Statistics and Operations Research for a pleasant working environment. Furthermore, thanks to the University of Copenhagen for giving me this opportunity with a stipend.

In the fall 1999 I had the pleasure of staying in New York. I am very grateful to The Department of Statistics at Columbia University for sharing their facilities with me.

Last but not least, I am grateful to Pernille for coping with me during this period and for taking a second chance with a geek.

Contents

Chapter 1

Introduction

This Ph.D. thesis is on the pricing of different securities involving credit risk. In this chapter we will give a short introduction to credit risk by presenting a few contributions in this area. It is not an attempt to review all contributions in this area so a lot of important work has not been included.

Credit risk models can be divided into two main categories. The structural models which based on accounting information models the default probability. This type of modeling is theoretically solid and gives a good intuition for which variables are important for the default risk. However, pricing in these models tend to be somewhat difficult since they are usually based on asset value which can be hard to estimate. The other type of models is the reduced form/intensity based models which model default by an intensity. Here, it is not as clear how the default event is related to the firms structure even though the intensity may depend on firm specific variables. As we will see the pricing of defaultable claims becomes much more simple in this setting since, loosely speaking, the intensity plays the role of an extra discount factor. Therefore, pricing becomes very similar to pricing interest rate derivatives. Since this thesis focuses on the pricing of (complicated) defaultable claims we have chosen to work with a reduced form setting.

In Section 1.1 we will present four structural models which show most of the important features of this type of modeling. This will be a useful reference for example when we touch upon the KMV data in Chapter 2. First we will present Merton[49] which is a simple and very intuitive model. Next we will present Black and Cox[6] which handles some of the deficiencies in Merton[49]. Then we present Leland[44] which is a much more complicated model where the owners decision making also influence the default probability. Finally, we present Duffie and Lando[22] which is a link between the structural models and the intensity based models. Other models could have been presented but these models serve our purposes. Section 1.2 presents an intensity based model. Furthermore, we give a pricing formula for defaultable claims and show that in this setting different types of default settlements can be handled and also tax can be handled in this setting.

1.1 Structural models

It was not until the seminal paper by Black and Scholes [7] where they give an option pricing formula in a general equilibrium model that the pricing of corporate debt could be analyzed in a more quantitative framework. In this paper and in Merton[48] it is recognized that the option pricing formula presented in this paper can be used to price corporate securities. The argument is that the equity is a call option on the firm value. This is the basis for the first structural models.

In the structural models default is modeled as a function of the firm value. Therefore we can view corporate debt as a contingent claim on the firm. In the models we will consider in the next sections the firm value, V, is given by

$$dV_t = \mu(t, V_t)dt + \sigma V_t dW_t$$

There is also a constant riskless interest rate r. For most of the applications the calculations would be almost the same for a stochastic interest rate. It is also assumed that there exits an equivalent martingale measure under which any claim on the firms value can be priced. Hedging these claims, however, might be difficult since firm value is not a traded asset.

1.1.1 Merton

In Merton[49] the first structural model for pricing corporate debt is set up. He assumes that there are only two classes of claims on the firm value, one homogeneous class of debt and a residual class of equity. The firm can not issue new senior claims nor can it pay out dividends. The debt is D and it is issued as zero-coupon bonds with maturity, T. If the firm value is greater than D at time T the debt is paid and the equity holders receive the residual. If the firm can not meet the payment D at time T the bondholders liquidate the firm and receive the total firm value. In that case equity holders receive nothing.

In Figure 1.1 the debt at time T is plotted as a function of firm value. The debt at maturity can be evaluated as

$$B_T = \min(V_T, D) = V_T - \max(V_T - D, 0)$$

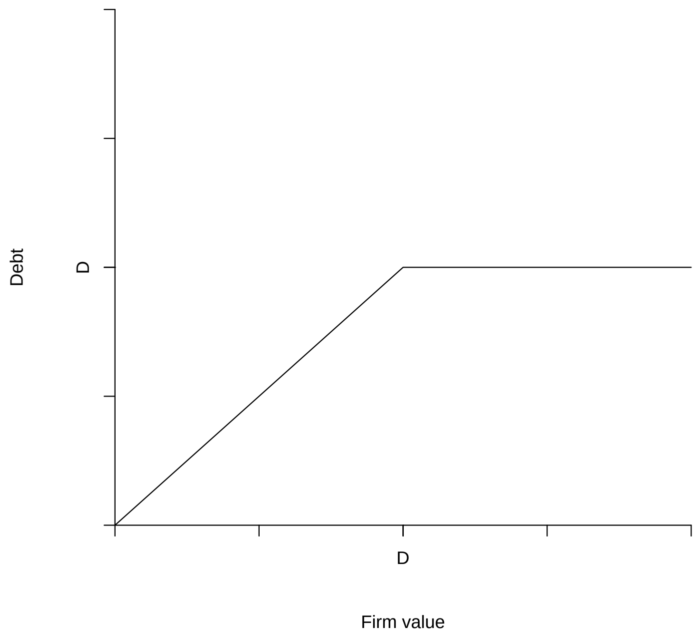


Figure 1.1: The debt payed at time T as a function of firm value.

which is firm value minus the value of the equity, E_T. As we can see the equity part can be valued as a European call option on the firm value and a strike price of D. Hence, the equity is valued using the Black-Scholes formula

$$E_t = BS(V_t, D, \sigma, r, T-t) \equiv V_t\Phi(d_1) - e^{-r(T-t)}D\Phi(d_2) \tag{1.1}$$
$$d_1 = \frac{\log \frac{V_t}{D} + (r + \frac{1}{2}\sigma^2)(T-t)}{\sigma\sqrt{T-t}}$$
$$d_2 = d_1 - \sigma\sqrt{T-t}$$

where Φ is the cumulative standard normal distribution function. From (1.1) we find that

$$\begin{aligned} B_t &= V_t - E_t \\ &= e^{-r(T-t)}D\left(\Phi(d_2) + \frac{V_t}{e^{-r(T-t)}D}\Phi(-d_1)\right) \end{aligned}$$

Define the leverage ratio as

$$l_t \equiv \frac{e^{-r(T-t)}D}{V_t}$$

and rewriting d_1 as

$$d_1 = \frac{-\log l_t + \frac{1}{2}\sigma^2(T-t)}{\sigma\sqrt{T-t}}$$

then the yield spread defined as

$$\begin{aligned} S_{t,T} &\equiv -\frac{1}{T-t}\log\frac{B_t}{D} - r \\ &= -\frac{1}{T-t}\log\left(\Phi(d_2) + \frac{1}{l_t}\Phi(-d_1)\right) \end{aligned} \qquad (1.2)$$

only depends on time to maturity, volatility, and leverage ratio. Notice, that the yield spread only depends on the value of the firm through the leverage ratio.

The Merton model is a good first attempt to price corporate debt. The idea to think of equity as a call option on firm value has been widely used since then. Merton also considers coupon bearing debt. In that case he can only achieve closed form solutions with an infinite horizon. Still, the debt structure in the Merton model has some deficiencies. The class of bondholders might not be homogeneous, e.g. senior and junior debt should not be priced equally. Also, bondholders will intervene if the firm value decreases to a certain level. They are not willing to sit and watch the equity holders ruining the firm. These two deficiencies have been studied by Black and Cox [6] in a similar setup.

1.1.2 Black and Cox

In Black and Cox [6] they consider the effect of safety covenants. Specifically, they introduce a lower boundary, H_t, at which bondholders will liquidate the firm and receive H_t. This boundary takes an exponential form

$$H_t \equiv Ce^{-\gamma(T-t)}$$

in which case they are still able to obtain a closed form solution for corporate debt in the form of zero coupon bonds. The idea is illustrated in Figure 1.2.

They also introduce prioritized debt so the pool of bondholders is no longer homogeneous. The pool of debt is divided into senior debt, SD, and junior debt, JD

$$D = SD + JD$$

Let the value of junior debt be JB_t and the value of senior debt SB_t then as shown in Figure 1.3

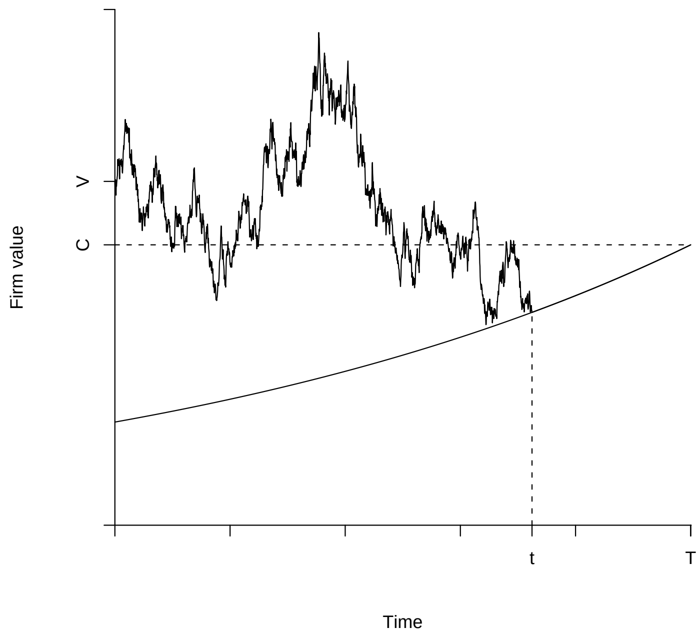


Figure 1.2: The firm is liquidated at time t and bondholders receive $Ce^{-\gamma(T-t)}$.

$$JB_T = [V_T - SD]^+ - [V_T - SD - JD]^+$$

Hence, JB_t is the difference between two call options, so if we can calculate the value of equity we can deduce the value of each class of debt.

Conditioning on no liquidation before maturity the equity is still a call option on the firm value. In case of liquidation equity holders get nothing. This is exactly a down-and-out call with an exponential barrier. This option has the value[1]

$$\begin{aligned} E_t &= BS\left(V_t, D, \sigma, r, T-t\right) \\ &\quad - \left(\frac{Ce^{-\gamma(T-t)}}{V_t}\right)^{2\frac{r-\gamma}{\sigma^2}-1} BS\left(\frac{\left(Ce^{-\gamma(T-t)}\right)^2}{V_t}, D, \sigma, r, T-t\right) \end{aligned}$$

Introducing a covenant increases the value of debt from the Merton model

[1] This value can be calculated using the distribution of $(V_t, \min_{0\leq s\leq t} V_s)$ which is a transformation of $(W_t, \max_{0\leq s\leq t} W_s)$ whose distribution can be found in Harrison [31]

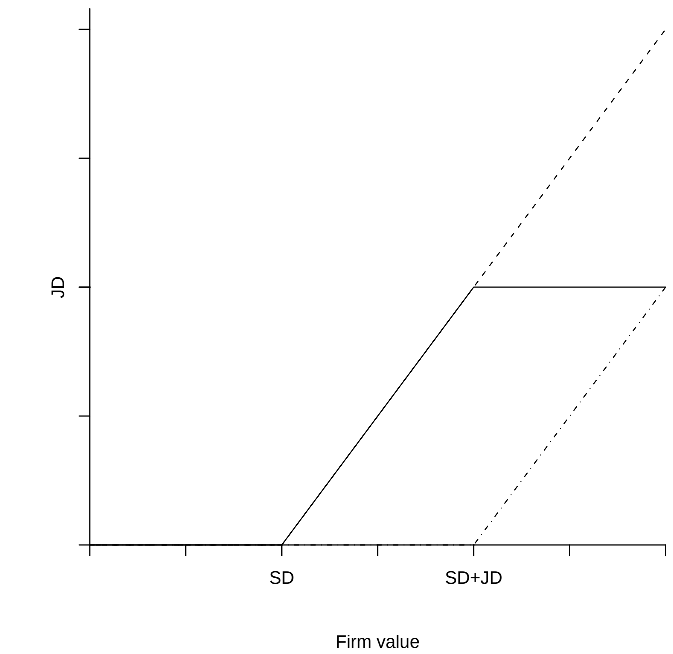


Figure 1.3: The debt payed at time T as a function of firm value.

with

$$\left(\frac{Ce^{-\gamma(T-t)}}{V_t}\right)^{2\frac{r-\gamma}{\sigma^2}-1} BS\left(\frac{\left(Ce^{-\gamma(T-t)}\right)^2}{V_t}, D, \sigma, r, T-t\right) > 0$$

In Figure 1.4 we have plotted the yield spreads for both the Merton model (1.2) and the Black and Cox model. We can see that the spreads for the Black and Cox model are smaller which is due to the introduction of the covenant which increases the value of debt.

They also have a discussion on debt in the form of coupon bonds. In that case they only obtain closed form solutions for the value of equity in a case with an infinite horizon, just as Merton.

In the Black and Cox model the liquidation boundary is chosen exogenously. In some cases an endogenously chosen boundary would be more appropriate. Black and Cox propose to choose such a boundary by maximizing firm value. This is done in Leland [44] where the liquidation boundary is found within the model by optimizing equity value.

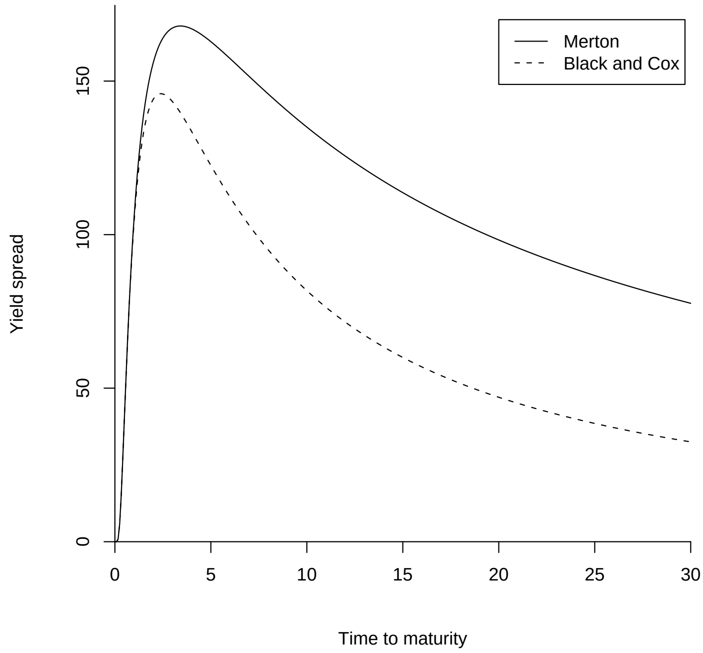


Figure 1.4: Yield spread (in bps) as a function of maturity. The parameters of the models are $V = 150, D = 100, C = 80, r = 5\%, \sigma = 30\%, \gamma = r$.

1.1.3 Leland

Leland [44] introduces tax and bankruptcy costs to the problem of valuing corporate debt. He considers a model with infinite horizon such that he can obtain a closed form solution for the value of debt. In this model the firm promises to pay a continuous coupon of C. Now, any claim on the firm, $F(V)$, with an infinite horizon must be a solution to

$$\frac{1}{2}\sigma^2 V^2 F_{VV}(V) + rVF_V(V) - rF(V) + C = 0$$

The general solution is

$$\begin{aligned} F(V) &= A_0 + A_1 V + A_2 V^{-X} \qquad (1.3) \\ X &= \frac{2r}{\sigma^2} \end{aligned}$$

To find an optimal capital structure tax and bankruptcy costs are introduced. Tax deductible coupons benefits high leverage. High leverage, however, increases the chance of bankruptcy which is costly[2], therefore we are able to find an optimal capital structure.

The bankruptcy cost can be found the following way. At the bankruptcy level (which is not yet determined) V_B the bankruptcy cost, $BC(V)$, is assumed to be a fraction of the firms asset value

$$\begin{aligned} BC(V) &= \alpha V_B \quad \text{for } V = V_B \\ BC(V) &= 0 \quad \text{for } V = \infty \end{aligned}$$

At very high values of V the probability of bankruptcy is 0, hence $BC(V) = 0$. The solution of (1.3) with these boundary conditions is

$$BC(V) = \alpha V_B \left(\frac{V_B}{V}\right)^X$$

Let the tax rate be τ then we find the value of the tax shield, $TB(V)$, by a set of boundary conditions. At V_B the firm is liquidated so there are no more tax benefits. Hence,

$$\begin{aligned} TB(V) &= 0 \quad \text{for } V = V_B \\ TB(V) &= \frac{\tau C}{r} \quad \text{for } V = \infty \end{aligned}$$

and we find

$$TB(V) = \frac{\tau C}{r}\left(1 - \left(\frac{V_B}{V}\right)^X\right)$$

The total firm value, $v(V)$, can be found as the asset value plus the value of the tax shield less the bankruptcy cost

$$v(V) = V + TB(V) - BC(V)$$

The debt value $D(V)$ is found using the boundary conditions

$$\begin{aligned} D(V) &= (1-\alpha)V_B \quad \text{for } V = V_B \\ D(V) &= \frac{C}{r} \quad \text{for } V = \infty \end{aligned}$$

so a fraction α is lost to other parties in case of bankruptcy and for high asset value the debt is riskless. The solution is

$$D(V) = \frac{C}{r} + \left((1-\alpha)\, V_B - \frac{C}{r}\right)\left(\frac{V_B}{V}\right)^X$$

[2]The total cost of bankruptcy is both the bankruptcy costs but also the lost tax benefits. Brennan and Schwartz [9] show that even when bankruptcy costs are 0 the optimal leverage ratio is less than 100% with a positive tax rate.

If the bankruptcy level V_B is set by the equity owners they wish to maximize their value $E(V) \equiv v(V) - D(V)$. Setting $\frac{\partial E}{\partial V_B} = 0$ leaves

$$V_B = \frac{(1-\tau)\,C}{r + \frac{1}{2}\sigma^2}$$

which can be shown to maximize E. Notice that the bankruptcy level is proportional to the coupon, C, and does not depend on the bankruptcy cost, α.

In Figure 1.5 we have plotted both the total firm value and the debt value as a function of the coupon. By maximizing the total firm value we can find the optimal leverage ratio to be 64.3% for a coupon of $C = 7.8$. For debt

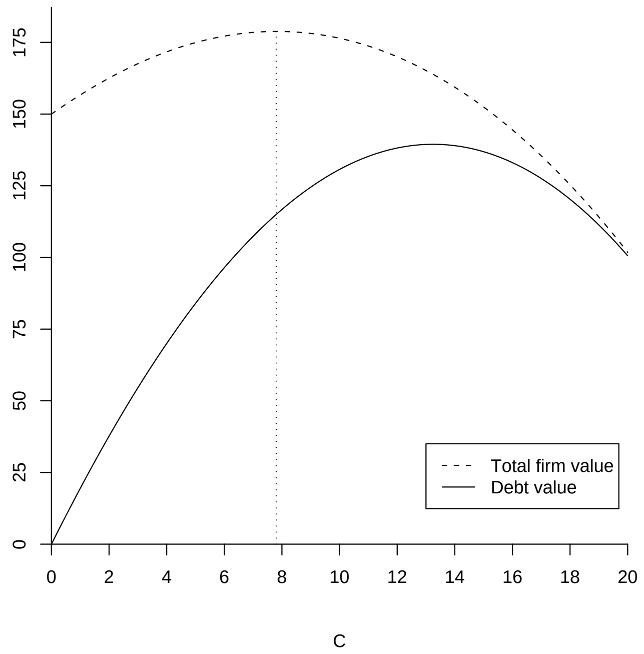


Figure 1.5: Total firm value and debt value as a function of the coupon for the parameters $V = 150, r = 5\%, \sigma = 30\%, \tau = 35\%, \alpha = 50\%$. The total firm value is maximized for $C = 7.8$ and the corresponding leverage ratio is 64.3%.

with an infinite horizon the yield spread is defined as

$$YS = C/D(V) - r$$

In Figure 1.6 we have plotted the yield spread as a function of the coupon. Using the optimal capital structure with $C = 7.8$ we find a yield spread of 179.1 bps.

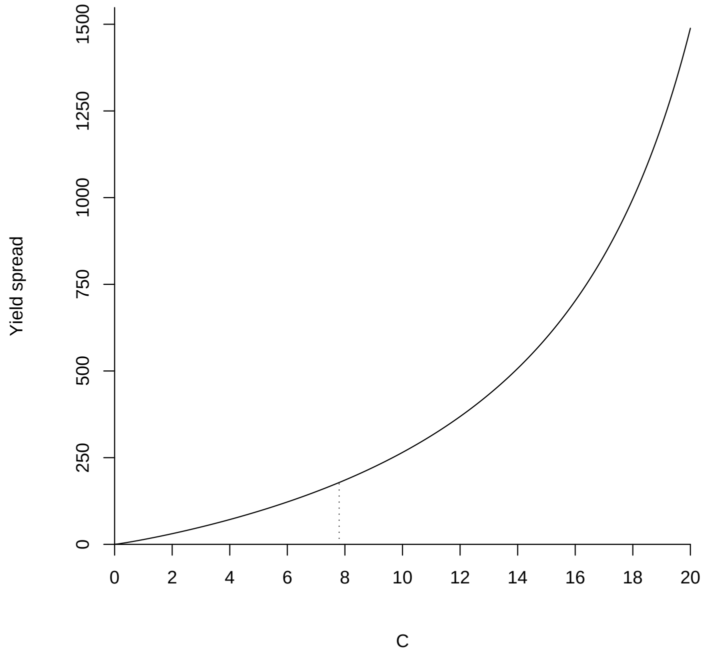


Figure 1.6: The yield spread as a function of the coupon with the parameters $V = 150, r = 5\%, \sigma = 30\%, \tau = 35\%, \alpha = 50\%$. For $C = 7.8$ the yield spread is 179.1 bps.

1.1.4 Duffie and Lando

One problem with most structural models including the three models presented in the previous sections is that the short yield spread is 0 which is different from observed short spreads. To see this consider the yield spread defined in (1.2), and let the short spread be defined as

$$S_t = S_{t,t} = \lim_{\Delta t \to 0} S_{t,t+\Delta t}$$

Assume $V_t > D$ then we have in the Merton model for small Δt

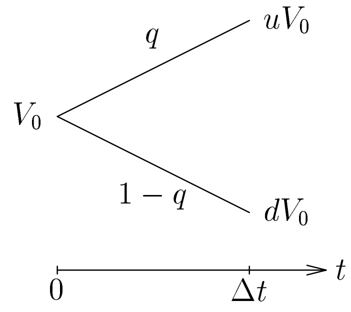
$$\begin{aligned} S_{t,t+\Delta t} &= -\frac{1}{\Delta t}\log\left(\Phi\left(d_2(\Delta t)\right) + \frac{1}{l_t}\Phi\left(-d_1(\Delta t)\right)\right) \\ &\simeq -\frac{1}{\Delta t}\log\Phi\left(d_2(\Delta t)\right) \\ &\simeq -\frac{\Phi\left(d_2(\Delta t)\right)-1}{\Delta t} \\ &= \frac{1-P\left(V_{t+\Delta t} > D\right)}{\Delta t} \\ &= \frac{P\left(V_{t+\Delta t} \le D\right)}{\Delta t} \end{aligned}$$

since $\Phi(d_2)$ is exactly the survival probability. This is the reason that we get a short spread of 0 in structural models since

$$\lim_{\Delta t \to 0} \frac{P\left(\text{default before time } t+\Delta t | \text{no default at time } t\right)}{\Delta t} = 0$$

for a predictable process.

This can also be seen in the binomial model of Cox, Ross, and Rubinstein [13]



Let

$$\begin{aligned} u &= e^{\sigma\sqrt{\Delta t}} \\ d &= \frac{1}{u} \\ q &= \frac{e^{r\Delta t}-d}{u-d} \end{aligned}$$

then we will get $V_t = V_0 e^{(r-\frac{1}{2}\sigma^2)t+\sigma W_t}$ as the limit when $\Delta t \to 0$.

Assume that the firm defaults if the value $V \le D$. We wish to calculate the probability of default at time Δt given that $V_0 > D$. Define $X_t = \log V_t$

then

$$\begin{aligned} & P\left(V_{\Delta t} \leq D \middle| V_0 > D\right) \\ = & P(\text{“down”}) P\left(V_0 e^{-\sigma\sqrt{\Delta t}} \leq D \middle| V_0 > D\right) \\ = & (1-q) P\left(\log V_0 \leq \log D + \sigma\sqrt{\Delta t} \middle| \log V_0 > \log D\right) \\ = & (1-q) P\left(X_0 \leq \log D + \sigma\sqrt{\Delta t} \middle| X_0 > \log D\right) \end{aligned}$$

If X_0 is known at time 0 this is 0 from a certain step, hence

$$\lim_{\Delta t \to 0} \frac{P\left(V_{\Delta t} \leq D \middle| V_0 > D\right)}{\Delta t} = 0$$

This is the case in the structural models of Merton, Black and Cox, and Leland[3]. In Duffie and Lando[22] they assume that the value of the firm can not be observed accurately. We can only observe if the firm has defaulted or not. In this case they find

$$\begin{aligned} & \lim_{\Delta t \to 0} \frac{P\left(V_{\Delta t} \leq D \middle| V_0 > D\right)}{\Delta t} \\ = & \lim_{\Delta t \to 0} (1-q) \lim_{\Delta t \to 0} \frac{P\left(X_0 \leq \log D + \sigma\sqrt{\Delta t} \middle| X_0 > \log D\right)}{\Delta t} \\ = & \frac{1}{2}\sigma^2 \lim_{\Delta t \to 0} \frac{P\left(X_0 \leq \log D + \sigma\sqrt{\Delta t} \middle| X_0 > \log D\right)}{\sigma^2 \Delta t} \\ = & \frac{1}{2}\sigma^2 f'(\log D) \end{aligned} \tag{1.4}$$

where f is the density function for $X| X > \log D$. If this density has a derivative different from 0 at the boundary the short spread is no longer 0.

1.2 Intensity Based Models

In intensity based models default is described by a stopping time which admits an intensity. As we will see in Section 1.2.1 the result of Duffie and Lando[22] justifies the use of an intensity even when working with structural models if there is uncertainty about the asset value of the firm.

[3]Leland's model as presented here only has debt with an infinite horizon. Therefore, it seems irrelevant to define short spreads for that model. However, the model could also be presented with a finite horizon in which case the short spread would be 0. With finite maturity there are no closed form solutions and numerical methods are called for.

The advantage of intensity based models is that they create a pricing framework very similar to interest rate theory. The main result is that we can discount promised payments with a default adjusted rate instead of discounting the realized payments with the interest rate. The adjustment is exactly the default intensity.

1.2.1 Model Setup

We will consider a filtered probability space $(\Omega, \mathbb{F}, (\mathbb{F}_t)_{t\geq 0}, P)$. We will assume the existence of an equivalent martingale measure Q under which all pricing is done. We will use the natural filtration generated by an n-dimensional state variable X_t defined by the SDE

$$dX_t = \mu(t, X_t)\, dt + \sigma(t, X_t)\, dW_t$$

where $\mu : [0,\infty) \times \mathbb{R}^n \to \mathbb{R}^n, \sigma : [0,\infty) \times \mathbb{R}^n \to \mathbb{R}^{n\times n}$ and W_t is an n-dimensional Brownian motion under P. We will also assume the existence of a (locally) risk less interest rate which will be a function of the state variables $r_t = R(X_t)$.

In intensity based models default is described by a stopping time τ which admits an intensity process λ_t. An intensity for τ is a non-negative, predictable process with

$$\int_0^t \lambda_s\, ds < \infty \text{ for every } t \text{ a.s.}$$

for which

$$1_{\{\tau\leq t\}} - \int_0^t \lambda_s 1_{\{\tau>s\}}\, ds \tag{1.5}$$

is a martingale. For more details see Brémaud[8]. We will assume that the intensity is a function of the state variables $\lambda : \mathbb{R}^n \to [0,\infty)$.

Using the martingale property (1.5) we find that

$$\begin{aligned} E\left[1_{\{\tau\leq T\}} - \int_0^T \lambda_s 1_{\{\tau>s\}}\, ds \middle| \mathbb{F}_t \right] &= 1_{\{\tau\leq t\}} - \int_0^t \lambda_s 1_{\{\tau>s\}}\, ds \\ \Leftrightarrow \quad P\left(\tau \leq T \middle| \mathbb{F}_t\right) &= 1_{\{\tau\leq t\}} + \int_t^T E\left[\lambda_s 1_{\{\tau>s\}} \middle| \mathbb{F}_t \right] ds \end{aligned}$$

where we have assumed that we can interchange the integrals. Given no default before time t and then differentiating with respect to T and taking the limit $T \to t$ we find

$$\frac{\partial P\left(\tau \leq t \middle| \mathbb{F}_t\right)}{\partial t} = \lambda_t$$

In other words the default intensity is the default probability within the next time interval multiplied by the length of the interval. Therefore, in intensity based models the short spread is typically positive.

In the model of Duffie and Lando[22] as we saw in (1.4) there exists a default intensity. In their model the default intensity is an intensity with respect to a filtration only containing information about the default time and the inaccurate observation of the firm value. The default intensity, λ_t, is exactly

$$\lambda_t = \frac{1}{2}\sigma^2 f'(\log D)$$

1.2.2 Pricing in intensity based models

In this section we wish to price defaultable claims. We will consider two types of payments. The first is a payment conditioning on no default prior to the payment date. This type of defaultable payment is considered in Duffie and Singleton[26] and Lando[43]. The other type is a payment paid upon default and this type of payment is also considered in Lando[43].

First, consider a payment due at time T, $Z_T 1_{\{\tau > T\}}$. In Duffie and Singleton[26] they work with a pre-default market value V_t which is the value of a claim given that there has been no default. Assume that in case of default a fraction, δ_t, of the pre-default value is recovered. In this case the discounted gain process is defined by

$$G_t = e^{-\int_0^t r_s\,ds} V_t(1-N_t) + \int_0^t e^{-\int_0^s r_u\,du}\delta_s V_{s-} dN_s$$

where $N_t = 1_{\{\tau \le t\}}$. From Duffie[17] we know that the discounted gain process is a martingale under Q. Using Itô's lemma on G_t we find

$$\begin{aligned} & dG_t \\ = \; & -r_t e^{-\int_0^t r_s\,ds} V_t(1-N_t)\,dt + e^{-\int_0^t r_s\,ds}(1-N_t)dV_t \\ & \quad - e^{-\int_0^t r_s\,ds} V_t dN_t + e^{-\int_0^t r_s\,ds}\delta_t V_t dN_t \\ = \; & -e^{-\int_0^t r_s\,ds}(1-N_t)\left(r_t + \lambda_t(1-\delta_t)\right) V_t\,dt + e^{-\int_0^t r_s\,ds}(1-N_t)dV_t + dM_t \end{aligned}$$

where M_t is a martingale. From this we find that

$$dV_t = \left(r_t + \lambda_t(1-\delta_t)\right) V_t\,dt + dM_t^* \tag{1.6}$$

where M_t^* is a martingale. Using the condition $V_T = Z_T$ we find the solution of (1.6)

$$V_t = E_t\left[e^{-\int_t^T r_s + \lambda_s(1-\delta_s)\,ds} Z_T\right] \tag{1.7}$$

Next, consider a payment $Y_\tau 1_{\{\tau\leq T\}}$ at the time of default and expiring at time T. This can be viewed as a different type of settlement than the fractional recovery of market value used previously. Again, define the discounted gain process

$$G_t = e^{-\int_0^t r_s\,ds} V_t(1-N_t) + \int_0^t e^{-\int_0^s r_u\,du} Y_s\,dN_s$$

and use Itô's lemma to find

$$\begin{aligned} & dG_t \\ = \;& -r_t e^{-\int_0^t r_s\,ds} V_t(1-N_t)\,dt + e^{-\int_0^t r_s\,ds}(1-N_t)dV_t \\ & \quad -e^{-\int_0^t r_s\,ds} V_t dN_t + e^{-\int_0^t r_s\,ds} Y_t dN_t \\ = \;& -e^{-\int_0^t r_s\,ds}(1-N_t)\,(r_t+\lambda_t)\,V_t\,dt + e^{-\int_0^t r_s\,ds}(1-N_t)dV_t \\ & \quad +e^{-\int_0^t r_s\,ds} Y_t(1-N_t)\lambda_t\,dt + dM_t \end{aligned}$$

where M_t is a martingale. From this we find that

$$dV_t = (r_t+\lambda_t)\,V_t - \lambda_t Y_t\,dt + dM_t^* \tag{1.8}$$

where M_t^* is a martingale. Using the condition $V_T = 0$ we find the solution of (1.8)

$$V_t = E_t\left[\int_t^T e^{-\int_t^s r_u+\lambda_u\,du}\lambda_s Y_s\,ds\right] \tag{1.9}$$

1.2.3 Default settlements

In general defaultable claims can be priced using (1.7) with $\delta = 0$ and then pricing the recovery separately. In Duffie and Singleton[26] they use a fraction of pre-default market value as recovery. This leads to the very nice pricing formula in (1.7) where both recovery and the regular payment can be priced at once.

In some cases, however, other settlements might be more appropriate. In Jarrow, Lando, and Turnbull[33] they assume that recovery is a fraction of the promised payment, $\bar{\delta}_T$. This type of settlement can be captured as

$$\bar{\delta}_T Z_T + (1-\bar{\delta}_T) Z_T 1_{\{\tau>T\}}$$

which can be priced using (1.7) with $\delta = 0$.

Any other settlement at the time of default can be priced using (1.9). For example a recovery of par assumption could be priced by $Y_t = \tilde{\delta}_t$ where $\tilde{\delta}_t$ is the fraction of par.

1.2.4 Tax in intensity based models

In Elton et. al.[28] they recognize the issue of different taxation of treasuries and corporate bonds in the U.S. They show that tax on corporate bonds will lead to a higher spread to treasury bonds which are not exposed to taxation. They use a discrete time model to show this. Here we will introduce tax in an intensity based model.

In Duffie and Singleton[26] and Lando[43] they find that instead of discounting the defaultable payments in corporate bonds with the riskless interest rate we might as well discount the promised payments with a higher risk adjusted rate. Let $T_1, T_2, \ldots, T_n$ be payment dates with a promised payment of X_i at time T_i. Let the intensity for default be λ_t with a recovery rate δ. I.e. this contract, V, can be evaluated as

$$V(t) = E_t\left[\sum_{i=1}^{n} e^{-\int_t^{t_i} r_s+\lambda_s(1-\delta)\,ds} X_i\right] \tag{1.10}$$

Let τ be the tax and define $X_i = Y_i + Z_i$ where Y_i is exposed to the tax rate τ and Z_i is not exposed to tax. Now, the real payment that this contract promises is

$$\tilde{X}_i = (1-\tau)Y_i + Z_i$$

Also, in the case of a default the loss will be deductible such that the real recovery is

$$\delta V(t) + (1-\delta)\tau V(t) = (\delta + (1-\delta)\tau)\, V(t) \equiv \tilde{\delta} V(t)$$

Now, instead of (1.10) we have

$$\begin{aligned} V(t) &= E_t\left[\sum_{i=1}^{n} e^{-\int_t^{t_i} r_s+\lambda_s(1-\tilde{\delta})\,ds} \tilde{X}_i\right] \\ &= E_t\left[\sum_{i=1}^{n} e^{-\int_t^{t_i} r_s+\lambda_s(1-\delta)(1-\tau)\,ds} \left((1-\tau)Y_i + Z_i\right)\right] \end{aligned}$$

so it is actually possible to include tax in the type of model from Duffie and Singleton(1997) and Lando(1997).

Consider a corporate bond where the coupons are taxed. Define the constant $\Delta t = T_i - T_{i-1}$ for every $i = 1, \ldots, N$ and let the coupon be c paid at time $T_1, \ldots, T_N$ then the value of this bond is

$$\begin{aligned} & V(t) \\ = \;& E_t\left[\sum_{i=1}^{n} e^{-\int_t^{T_i} r_s+\lambda_s(1-\delta)(1-\tau)\,ds} \Delta t(1-\tau)c + e^{-\int_t^{T_n} r_s+\lambda_s(1-\delta)(1-\tau)\,ds}\right] \end{aligned}$$

For $n = 1, c = 0.06, \lambda_s = 0.005, \delta = 0.5, T_1 = \Delta t = 1, \tau = 0.04$ we find the yield spread $S(0)$ to be

$$S(0) = 0.47\%$$

whereas with no tax the spread is 0.25%. I.e. even for a very small tax the spread is in this case increased by almost 100%. $\lambda_s = 0.005$ is similar to a BBB rating. It is mostly for the higher ratings that the tax effect is significant. For lower rated bonds the tax deduction of the loss in case of default almost cancels the taxation of coupons. E.g. for $\lambda_s = 0.25$ which is similar to a CCC rating the spread is actually smaller with tax (12.23%) than without tax (12.5%).

Chapter 2

On Corporate Bonds with Step-Up Provisions.

Acknowledgments
This chapter is based on a joint work with David Lando with the same title. We are very grateful to Søren Kyhl and Jacob Gyntelberg for useful discussions and for providing us with data.

2.1 Introduction

The corporate bond market in the European telecommunication sector is currently seeing a dramatic increase in the number and volume of bond issues with embedded step-up and step-down covenants, i.e. provisions which link the coupon payments of the bonds to the ratings of the issuing firms. The major companies, most of which are former state monopolies, have been forced by liberalization of the market into more high risk/high yield activities most notably in the mobile phone market. In particular, the major players have needed to finance their participation in auctions for so-called UMTS licenses, which are needed to operate the next 'third generation' mobile phone technology.

The huge expenses to acquire these licenses (some figures estimate that the auctions are expected to produce a revenue of 125 billion dollars in Europe), the large expenses in actually developing the new technology and the high uncertainty in estimating the cash flows which the new technology will produce, have given strong negative reactions in the equity markets. These negative reactions have been reinforced by the general downturn in the technology sector and have made equity financing less attractive as seen from the companies.[1] Figure 2.1 shows the evolution of the equity price of Deutsche Telekom - a company used for illustrative purposes in this paper. With internal cash flows far from sufficient to finance the new investments

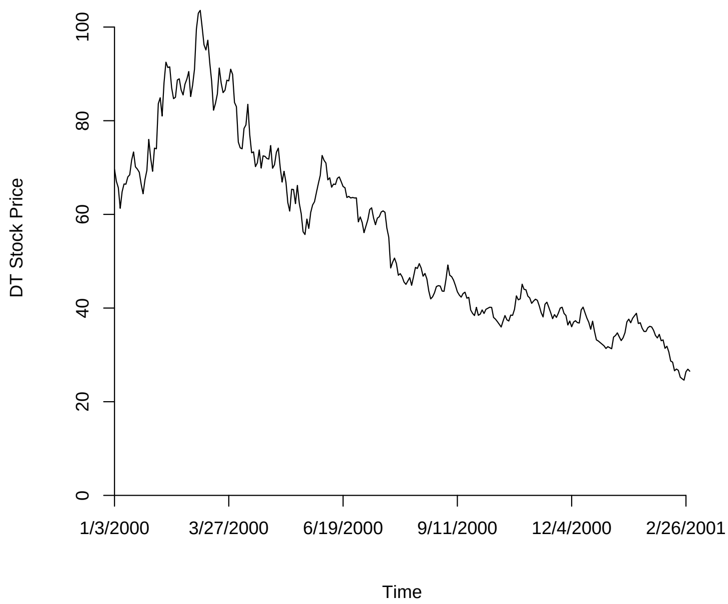


Figure 2.1: Daily observations of stock prices for Deutsche Telekom.

the companies have instead sought financing through large, corporate bond issues.

Almost all new issues by the major players in the telecom sector have step-up and/or step-down provisions.[2]

One can think of many reasons for these type of provisions to be seen as useful. From a theoretical point of view, it is consistent with an attempt by

[1]For example, the initial public offering by Orange, A France Telecom spinoff, in early 2001 only generated a revenue of 10 Euros per share, far from the 18 Euros expected in the beginning of the year, cf. Koo[39].

[2]Examples include the March 2001 issue by France Telecom with a dollar equivalent of 16,4 billion, the December 2000 and January 2001 issues by British Telecom with a total dollar equivalent of 18.8 billion and the June 2000 issue by Deutsch Telecom whose dollar equivalent value was 14.4 billion dollars. New issues by smaller players, such as TDC and Sonera, have not included step-up language.

management of the companies to signal that they will and can comply with the target leverage ratios required by the rating agencies to retain ratings which are above the step-up limit. In practice, many investors view the step-up clauses as providing a natural hedge against price changes due to downgradings. Investors may fear that the market will react strongly to downgrades possibly because fund managers will have to abandon positions as they come closer to having a speculative grade rating. At the same time, investors may be convinced that the companies offer a limited risk of actual default simply because the big companies still have a large fraction of shares owned by the government and because the companies are so vital in providing infrastructure that they are in a sense 'too big to fail'.

The precise language of the step-up provisions varies from issue to issue. Typically the rating level defining the clauses depend on changes in the rating by the two major agencies, Standard and Poor's and Moody's. The relevant rating is that of senior unsecured debt. In some cases, as with the Deutsche Telekom Euro issues used for illustrative purposes in this paper, the stipulated 50 bps step-up of the coupon requires a downgrade by both agencies below the comparable levels A3 and A- in the Moody's and S&P systems, respectively. In other cases, as with several British Telecom issues, there are step-up triggers in place for actions of each rating agency. Here, a one notch downgrade by one agency will trigger a 25 bps increase in the coupon, even if the rating of the other agency remains fixed.

Provisions also vary with respect to step-down provisions which, as the name suggests, trigger a lowering of the coupon if the company regains its original rating after a downgrade. For example, the KPN June 2003 5.75 % issue, has only a single step-up of 30 bps triggered by a downgrade either below Baa2 (Moody's) or below BBB+ (S&P) (a trigger which has become effective), but has no step-down if the company regains its original rating. A more common construction in the telecommunications sector, used for example in later issues by KPN and in the Deutsche Telekom issues studied in this paper, stipulates that coupons are stepped down if the rating rises above the trigger level again. In the case of two rating agencies, if a step-up has been triggered by a unanimous downgrade, then it requires a unanimous upgrade by both agencies to step down the coupon again.

Finally, the provisions may differ with respect to what happens with further downgrades. The Deutsche Telekom issues has only two possible levels of coupons, but the KPN 2006 and 2008 issues have continual step-ups for every further downgrade.

In all the constructions we are aware of, upgrades above the when-issued rating level do not trigger reductions in the coupon, so the step-up provisions are not completely like a floating rate security. Typically, the coupon change takes place on the first coupon date after the rating change has been announced.

The idea of linking credit quality to coupon payments is not new. A similar idea has been used in Floating Rate Notes, Credit Linked Notes and in credit triggers in swaps. What is remarkable, however, is the volume in which bonds with step-up clauses have been issued - making the secondary market for these bonds highly liquid.

The high liquidity makes the data well suited for testing and implementing the approach to arbitrage pricing of credit risky securities with rating dependence introduced in Jarrow, Lando and Turnbull [33] and Lando (1994,1998) and carried forward for example in the works by Das and Tufano [15], Arvanitis, Gregory and Laurent[4] and Kijima [36]. It also provides a natural opportunity to give an overview of the methods developed so far for pricing rating-sensitive instruments.

This paper summarizes and compares various approaches one may take to pricing rating sensitive debt and estimates some of the models using data for Deutsche Telekom. A key trick to calibration in these models is to avoid implying out all parameters in the risk-neutral transition probabilities from scratch. The goal in this paper is to give an overview of the trade-offs that one faces in ratings-based models between mathematical and statistical tractability on one side and realism on the other. Ultimately, one may hope that a good model can shed more light on the value of step-up provisions. For example, despite the large liquidity of the bonds, a general feeling among many market participants that we have talked to, is that the step-up provisions of the bonds in the telecommunication sector are typically under-priced. The question now is whether the rating-based methodology to pricing rating sensitive instruments can be used in assessing the risk premia involved in pricing the step-up provisions.

The overview of the paper is as follows: In Section 2.2 we describe the data used throughout the paper. Section 2.3 sets up the framework in which we will be working. In Section 2.4 we define the rating based model and give the PDE which is used to price rating dependent claims. Also we give a setup in which pricing can be done in closed form. In Section 2.5 we analyze different rating models to get an idea on which variables can be implied out from the prices and for which parameters can we use the empirically estimated values. Section 2.6 gives a computationally tractable alternative to the rating based model when we need more state variables than can be handled in a rating based model. In Section 2.7 we analyze the threshold model presented in Section 2.6 using KMV data. Finally, in Section 2.8 we give an approach to estimate the parameters when we are observing coupon bonds.

2.2 Data

We have daily observations of the three major Euro denominated corporate bonds from Deutsche Telekom. The first bond is issued on June 28 2000 with an annual coupon of 6.125% and is maturing July 6 2005. The second bond has an annual coupon of 6.625% and is maturing July 6 2010. It is also issued on June 28 2000. Both bonds have a step-up feature where the coupon is increased with 50 bps if Deutsche Telekom is rated below A3 by Moody's and below A- by Standard and Poor's. If both the ratings rise above the trigger level again after a downgrade the coupon is stepped down again. We also have observations of a bond without a step-up provision. This bond is maturing May 20 2008 and has an annual 5.25% coupon. We have 176 observations of the two bonds with step-up provision and 303 observations of the regular corporate bond. The prices are shown in Figure 2.2. Deutsche Telekom also

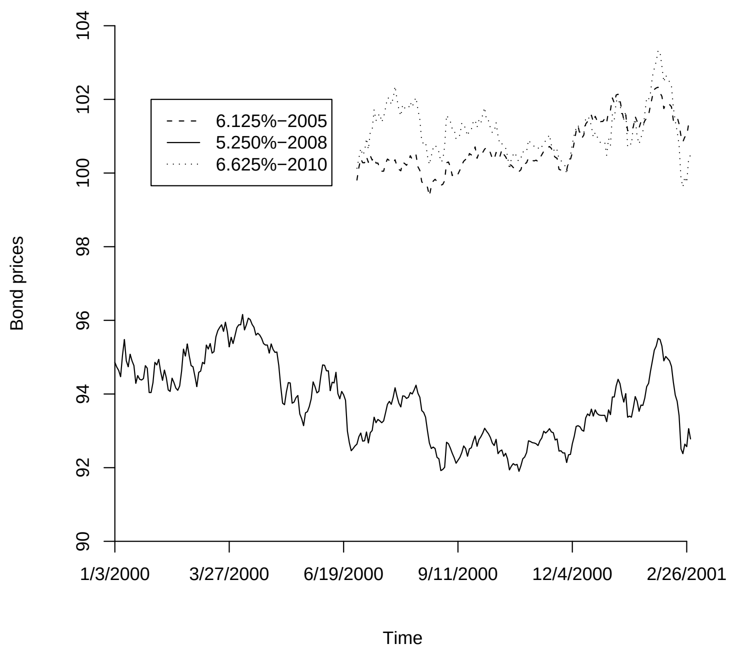


Figure 2.2: Daily observations of bond prices for Deutsche Telekom.

have issues in Dollar and Yen. We will disregard these issues since we are not including a full modeling of the capital structure of Deutsche Telekom and also since we can not be sure that these bonds are comparable with the

Euro issues. Since we are interested in models that can price comparable corporate bond issues the three Euro issues will be sufficient for this paper.

In the same time period we also have the yields on German treasuries for all maturities. The yields for the 5, 8, and 10 years treasuries are shown in Figure 2.3. We will be using the German treasury as discounting since this

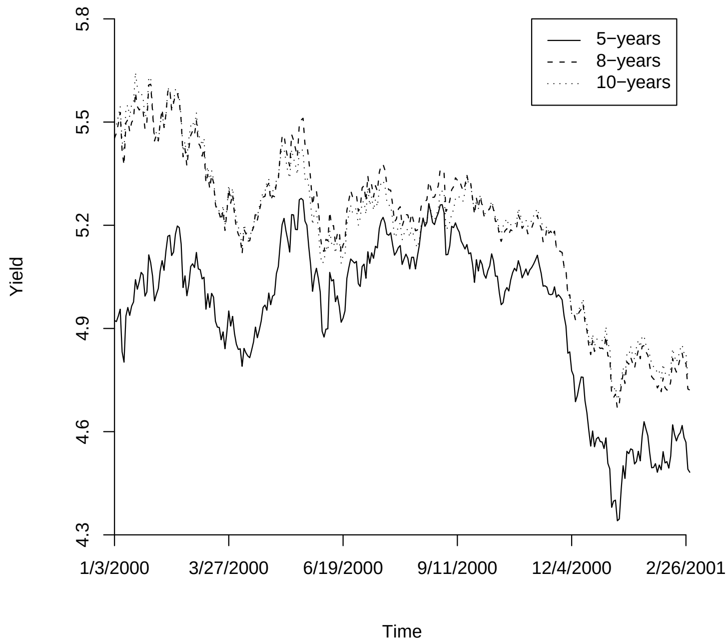


Figure 2.3: Daily observations of German treasury yields.

is the common benchmark for Euro denominated bonds.

Furthermore, we have yield spreads for AAA, AA, A, and BBB bonds in the same time period. These yields are shown in Figure 2.4 as referred in the Merrill Lynch index of 7-10 years European corporates. We are interested in the stochastics of a single firm but maybe the index is revealing something about the stochastics in the surroundings. We have chosen the 7-10 years index since this closely matches the Deutsche Telekom corporate bond maturities.

In Figure 2.4 we have also included the yield spread for the 5.25%-2008 Deutsche Telekom bond. Moody's rated Deutsche Telekom Aa2 with a negative outlook on January 1 2000. On April 10 the negative outlook was removed, but it was invoked again on June 22 2000. Finally, on October 5 it

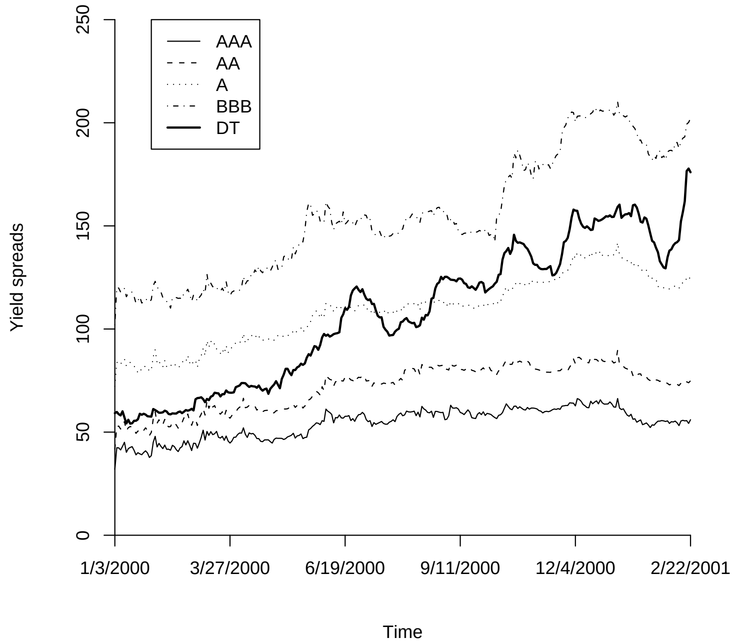


Figure 2.4: Yield spreads for the Merrill Lynch indexes of AAA,AA,A,BBB 7-10 years corporate bonds. We have included the yield spread for the Deutsche Telekom 2008 for comparison.

was downgraded by three notches to a A2 rating. Standard and Poor's rated Deutsche Telekom AA- in the beginning of year 2000 and gave it a negative outlook on April 28. On October 6 it was downgraded to A- and finally in the end of the sample on February 27 2001 Deutsche Telekom received a negative outlook. As we can see from the graph Deutsche Telekom's spread to the AA index started increasing in April 2000 and already in mid June 2000 it looked more like a A rated firm than a AA rated firm even though it was not downgraded till October 2000.

Finally, we have monthly estimates from KMV of the asset value, default point, volatility, and default probability for Deutsche Telekom. In Figure 2.5 we have plotted the asset value and the default point as a function of time. In Figure 2.6 we have plotted the volatility and default probability.

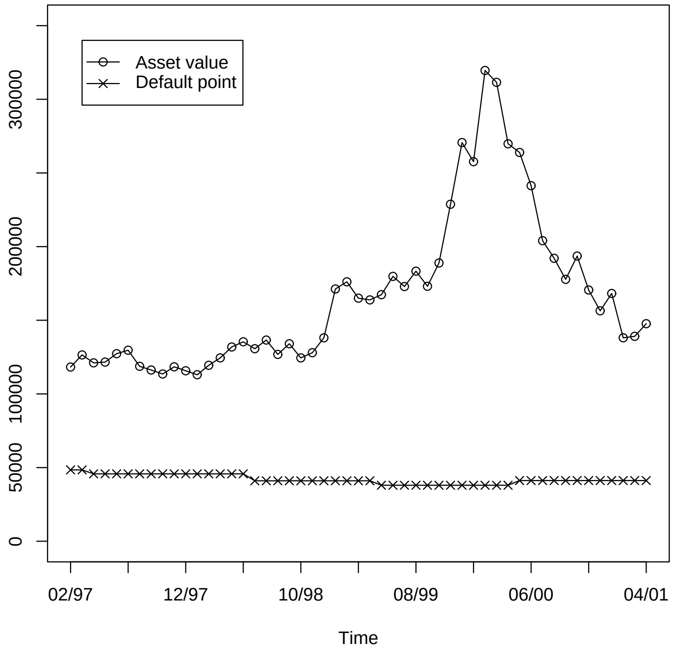


Figure 2.5: KMV estimates of the asset value and default point for Deutsche Telekom.

2.3 Model setup

We will consider a filtered probability space $\left(\Omega, \mathbb{F}, (\mathbb{F}_t)_{t\geq 0}, P\right)$. We will assume the existence of an equivalent martingale measure, Q, under which all pricing is done. We will be using the natural filtration generated by the state variables defined below.

In general we will work with an n-dimensional vector of state variables, X_t, defined by the SDE

$$dX_t = \mu(t, X_t)\,dt + \sigma(t, X_t)\,dW_t$$

where $\mu : [0,\infty) \times \mathbb{R}^n \to \mathbb{R}^n, \sigma : [0,\infty) \times \mathbb{R}^n \to \mathbb{R}^{n\times n}$ and W_t is a n-dimensional Brownian motion under P. We will denote the (locally) risk less interest rate by $r_t = R(X_t)$ which will be some function of the state variables.

For the remainder of this paper we will assume that the state variables have an affine specification i.e. $\mu(x), \sigma(x)\sigma(x)^T$ are both affine in x and follow

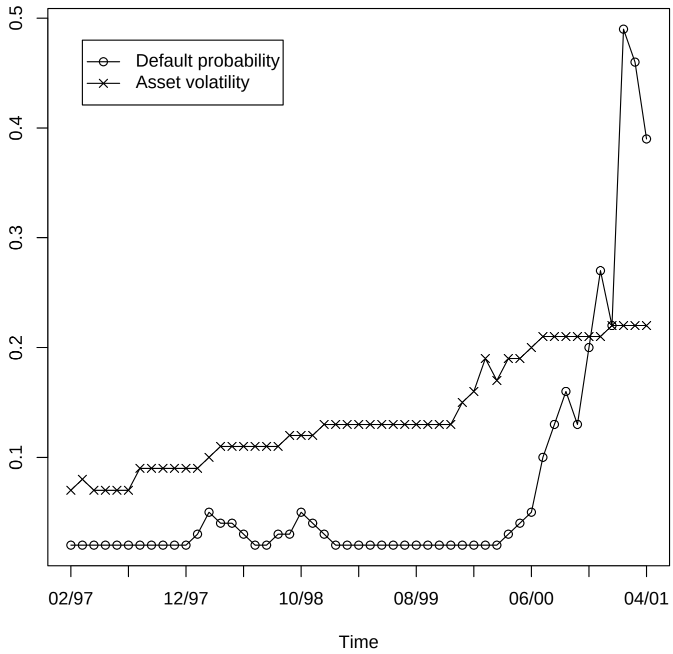


Figure 2.6: KMV estimates of volatility and default probability of Deutsche Telekom.

the pricing method of Duffie, Pan, and Singleton[23].[3] We will only consider a general CIR model

$$dX_t = K_0 + K_1 X_t \, dt + \Sigma \sqrt{X_t} \, dW_t^P$$

where $K_0 \in \mathbb{R}^n, K_1 \in \mathbb{R}^{n \times n}, \Sigma \in \mathbb{R}^{n \times n}$ and

$$\sqrt{x} = \begin{bmatrix} \sqrt{x_1} & 0 & \cdots & 0 \\ 0 & \sqrt{x_2} & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & \sqrt{x_n} \end{bmatrix}$$

With this setup

$$E_t \left[e^{-\int_t^T \rho_0 + \rho_1 \cdot X_s \, ds} e^{a \cdot X_t} \right] = e^{\alpha(T-t) + \beta(T-t) \cdot X_t} \tag{2.1}$$

[3]In Duffie, Pan, and Singleton[23] they introduce a fast pricing method for state variables with an affine specification. They also allow a jump process in the specification of the state variables.

where $\rho_0 \in \mathbb{R}, \rho_1, a \in \mathbb{R}^n$ and $\alpha : [0,T] \to \mathbb{R}, \beta : [0,T] \to \mathbb{R}^n$, and where α, β can be found as solutions to the set of ODE's

$$\beta'(s) = -\rho_1 + K_1^T \beta(s) + \frac{1}{2}\left(\Sigma^T \beta(s)\right)^2, \quad \beta(0) = a \tag{2.2}$$

$$\alpha'(s) = -\rho_0 + K_0^T \beta(s), \quad \alpha(0) = 0 \tag{2.3}$$

where $x^2 = (x_i^2)_{i=1,\dots,n}$. We will denote the dependence on a in the solutions by $\beta_a(t), \alpha_a(t)$. The derivation of these ODE's is an application of Itô's lemma. For more on this see Duffie, Pan, and Singleton[23].

If we assume that the market prices of risk are given by $\sqrt{X_t} b$ where $b \in \mathbb{R}^n$ then under the equivalent martingale measure Q we have

$$dX_t = K_0 + (K_1 + \Sigma D(b)) X_t \, dt + \Sigma \sqrt{X_t} \, dW_t^Q$$

where

$$D(x) = \begin{bmatrix} x_1 & 0 & \cdots & 0 \\ 0 & x_2 & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & x_n \end{bmatrix}$$

Hereby, we can preserve the affine structure under Q.

We will be considering defaultable claims and the time of default is τ. We will assume that this stopping time τ admits an intensity process which we will denote by λ_t. An intensity for τ is a non-negative, predictable process with

$$\int_0^t \lambda_s \, ds < \infty \text{ for every } t \text{ a.s.}$$

for which

$$1_{\{\tau \le t\}} - \int_0^t \lambda_s 1_{\{\tau > s\}} \, ds$$

is a martingale.[4] We will assume that the intensity is a function of the state variables $\lambda : \mathbb{R}^n \to [0, \infty)$.

If we assume that the recovery, δ_t, is a percentage of the pre-default market value of the claim, then pricing a defaultable claim can be done by discounting the *promised* payment with $r_t + \lambda_t(1 - \delta_t)$ instead of just the interest rate. This is shown in both Duffie & Singleton[26] and Lando[43]. Hence, the price of a defaultable claim with positive recovery is the same as one with 0-recovery but with a lower default intensity. Therefore, in some cases it might be easier to fit $\lambda_t(1 - \delta_t)$ instead of both fitting the default intensity and the recovery rate.

[4] For more details see Brémaud [8].

Furthermore, we will define K different (non-default) rating categories, $1, \ldots, K$, and denote the rating process by η_t. A rating dependent defaultable payment, $C(T, \eta_T)$, can be priced as

$$E^Q \left[e^{-\int_0^T r_t + \lambda_t \, dt} C(T, \eta_T) \right]$$

In this paper we will be considering payments of the form

$$C(t, \eta_t) = c + s 1_{\{\eta_t \geq i\}}$$

for some choice of rating i. In this case the coupon is stepped up whenever the rating is above i and it will be stepped down as soon as the rating gets below i again. This is exactly the case for two of the Deutsche Telekom issues we are studying in this paper. A different provision is

$$C(t, \eta_t) = c + s 1_{\{\max_{0 \leq u \leq t} \eta_u \geq i\}}$$

where the coupon cannot be stepped down again once the step-up provision is triggered. This is the case for the KPN June 2003 5.75% issue.

2.4 A rating based model

The step-up provision is typically linked to an explicit rating category given by a rating agency. In Jarrow, Lando, and Turnbull[33] they introduce a pricing framework which explicitly takes the ratings into account. In this model the rating transitions are governed by a continuous time Markov chain. The Markov chain is defined by a generator (under Q)

$$\Lambda = (\lambda_{ij})_{i,j=1,\ldots,K}$$

where λ_{ij} is the transition intensity for the Markov chain to jump from i to j if $i \neq j$ and $\lambda_{ii} = -\sum_{j \neq i} \lambda_{ij}$. Now, the price of a rating dependent payment, $C(T, \eta_T)$ can be calculated as

$$\begin{aligned} & E^Q \left[e^{-\int_0^T r_t \, dt} \right] \sum_{i=1}^{K} \left(e^{\Lambda T} \right)_{\eta_0 i} C(T, i) \\ = \quad & p(0, T) \sum_{i=1}^{K} \left(e^{\Lambda T} \right)_{\eta_0 i} C(T, i) \end{aligned}$$

where e^{Λ} is the matrix exponential of Λ and $p(0, T)$ is the price of a zero-coupon treasury with maturity T.

More generally, we will let the transition intensities depend on the state variables. Denote the matrix of transition intensities by $\Lambda_X(t)$ and the default intensity by $\lambda_X(\eta_t)$. We will work with two specifications of the generator. For some recoveries one specification work better than the other and for other recovery assumptions it is the other way around. The first possibility is to let default be a state "outside" the rating system. In this case the Markov Chain is a pre-default rating process and the generator for a system with two non-default states would be

$$\Lambda_X^1(t) = \begin{bmatrix} -\lambda_{12} & \lambda_{12} \\ \lambda_{21} & -\lambda_{21} \end{bmatrix}$$

Here the default intensity, $\lambda_X^1(\eta_t)$, would be modulated by the pre-default rating process. The other possibility is to let the Markov Chain include a specific state for default which typically will be the last state, K. This state is absorbing so for a system with two non-default states the generator would be

$$\Lambda_X^2(t) = \begin{bmatrix} -\lambda_{12} - \lambda_{13} & \lambda_{12} & \lambda_{13} \\ \lambda_{21} & -\lambda_{21} - \lambda_{23} & \lambda_{23} \\ 0 & 0 & 0 \end{bmatrix}$$

In this case default has already been taken care of so $\lambda_X^2(\eta_t) = 0$. If $\lambda_{13} = \lambda_X^1(1)$ and $\lambda_{23} = \lambda_X^1(2)$ these to systems are equivalent.

We wish to be able to handle both specifications, therefore we will work with a general generator $\Lambda_X(t)$ and an intensity for leaving the rating system $\lambda_X(\eta_t)$. In the following we will call the intensity for leaving the rating system a default intensity. This is a slight abuse of language. As we saw if the rating system already includes default $\lambda_X(\eta_t) = 0$ since there is no way to leave the rating system. The actual default intensity is included in $\Lambda_X(t)$.

From Duffie and Singleton[26] and Lando[43] we know that the value of a defaultable payment can be found by discounting the promised payment with $r(X_t) + \lambda_X(\eta_t)$. Define the martingale Y_t by

$$Y_t = E_t\left[e^{-\int_0^T r(X_s)+\lambda_X(\eta_s)\,ds} C(T, \eta_T)\right]$$

and the function $f(t, X_t, \eta_t)$ as

$$f(t, X_t, \eta_t) = E_t\left[e^{-\int_t^T r(X_s)+\lambda_X(\eta_s)\,ds} C(T, \eta_T)\right]$$

which is the price of a defaultable claim paying $C(T, \eta_T)$ at time T.

Using Itô's lemma

$$dY_t = -\left(r(X_t) + \lambda_X(\eta_t)\right) Y_t\,dt + e^{-\int_0^t r(X_s)+\lambda_X(\eta_s)\,ds}\,df$$

Since Y_t is a martingale the drift must be 0, hence the drift of f, $\mu_f(t)$, have to satisfy

$$\mu_f(t) = (r(X_t) + \lambda_X(\eta_t))\, f(t, X_t, \eta_t) \tag{2.4}$$

Define the differential operator by

$$\mathcal{D}f = \frac{\partial f}{\partial t} + \frac{\partial f}{\partial x}(X_t) \cdot (K_0 + K_1 X_t) + \frac{1}{2} tr\left(\Sigma D(X_t)\Sigma^T \frac{\partial^2 f}{\partial x^2}\right)$$

then using Itô's lemma for semimartingales with jumps (see e.g. Protter[52]) we find

$$\begin{aligned} & f(t, X_t, \eta_t) - f(0, X_0, \eta_0) \\ = & \int_0^t \mathcal{D}f(s, X_s, \eta_s)\, ds + \sum_{0<s\leq t} (f(s, X_s, \eta_s) - f(s, X_s, \eta_{s-})) + M_t \quad (2.5) \end{aligned}$$

where M_t is a local martingale.

Define $\lambda_i = -\left(\Lambda_X(t)\right)_{ii}$ which is the intensity for leaving state i. Then,

$$\begin{aligned} & \sum_{0<s\leq t} (f(s, X_s, \eta_s) - f(s, X_s, \eta_{s-})) \\ = & \int_0^t \lambda_{\eta_s} E_s\left[\Delta f(s, X_s, \eta_s) \middle| \Delta\eta_s \neq 0\right] ds + M_t^* \\ = & \int_0^t \lambda_{\eta_s} \sum_{k=1, k\neq\eta_s}^{K} \frac{\lambda_{\eta_s k}}{\lambda_{\eta_s}} (f(s, X_s, k) - f(s, X_s, \eta_s))\, ds + M_t^* \\ = & \int_0^t \sum_{k=1}^{K} \lambda_{\eta_s k} f(s, X_s, k)\, ds + M_t^* \end{aligned} \tag{2.6}$$

where M_t^* is a local martingale.

Denote the K-dimensional vector of any rating dependent function with a bar on top. For example

$$\bar{f}(t, x) = \begin{bmatrix} f(t, x, 1) \\ \vdots \\ f(t, x, K) \end{bmatrix}$$

then using (2.4), (2.5), and (2.6) we find a K dimensional PDE for $\bar{f}$

$$\mathcal{D}\bar{f}(t, x) + \Lambda_x(t)\bar{f}(t, x) - D\left(r(x) + \bar{\lambda}_x\right)\bar{f}(t, x) = 0 \tag{2.7}$$

with the terminal condition $\bar{f}(T, x) = \bar{C}(T)$.

2.4.1 Closed form solutions for the rating model

In general (2.7) has no closed form solution and numerical procedures are called for. However, if the number of state variables is greater than two numerical schemes as finite difference will be rather slow. When considering defaultable claims it is reasonable to assume that there is one state variable for the interest rate, one for the general default intensity in the economy, and one firm specific state variable. Therefore, we are interested in models that have closed form solutions. In Lando[43] closed form solutions are obtained if we assume that there exists K linearly independent eigenvectors of $\Lambda_X(t)$ all independent of X_t. We will show this result more generally.

Let $B \in \mathbb{R}^{K\times K\times n}$ and define

$$\Lambda_X(t) = BX_t$$

where the product is defined as $(Bx)_{ij} = \sum_k B_{ijk}x_k$. Also define $K \times K$ matrices $B_1, \ldots, B_n$ by

$$(B_i)_{jk} = B_{jki} \text{ for } i = 1, \ldots, n$$

Now, given the evolution of the state variables $\Lambda_X(t)$ is just a matrix of time-dependent intensities. Hence, we can define a conditional Markov chain given the state variables. Define the conditional transition probabilities by $P_X(s,t)$ where

$$\left(P_X(s,t)\right)_{ij} = P\left(\eta_t = j | \, \mathbb{F}_t \vee \sigma(\eta_s = i)\right)$$

and the unconditional transition probabilities as

$$P(s,t) = E\left[P_X(s,t)|\, \mathbb{F}_s\right]$$

The conditional Markov chain solves the forward Kolmogorov equation

$$\frac{\partial}{\partial t}P_X(s,t) = P_X(s,t)\Lambda_X(t) \tag{2.8}$$

and hereby,

$$\frac{\partial}{\partial t}P(s,t) = E\left[P_X(s,t)\Lambda_X(t)|\, \mathbb{F}_s\right]$$

Notice, that in general the unconditional transition probabilities P do not solve the forward Kolmogorov equation. In Appendix 2.9.1 the differential equation for P is given.

Equation (2.8) is solved by the product integral

$$P_X(s,t) = \prod_{s<u\leq t} I + \Lambda_X(u)\,du \equiv \lim_{m\to\infty}\prod_{i=1}^{m} I + \frac{t-s}{m}\Lambda_X(s + i\frac{t-s}{m})$$

which is defined more generally in Gill and Johansen[29] where they also show some properties of the product integral.

For example for a matrix $\Lambda_X(u) = UD_X(u)U^{-1}$

$$\prod_{s<u\leq t} I + \Lambda_X(u)\,du = U\left(\prod_{s<u\leq t} I + D_X(u)\,du\right)U^{-1}$$

This can be seen by observing that the right hand side solves (2.8).

In the case where $n = 1$ the product integral is just the exponential function and in the case of diagonal matrices it is the matrix exponential. In both cases

$$\prod_{s<u\leq t} I + \Lambda_X(u)\,du = e^{\int_s^t \Lambda_X(u)\,du}$$

We are interested in the transition probabilities and we find

$$P(s,t) = E_s\left[\prod_{s<u\leq t} I + \Lambda_X(u)\,du\right]$$

which is not easily solved in general. However, if we assume that

$$B_i = UD_iU^{-1} \text{ for every } i = 1,\ldots,n$$

where D_i is a diagonal matrix of the eigenvalues for B_i and U is a matrix of eigenvectors. I.e. we assume that $B_1,\ldots,B_n$ can all be diagonalized by the same set of eigenvectors. Let D be the collection of $D_1,\ldots,D_n$ such that $D \in \mathbb{R}^{K\times K\times n}$ then

$$\Lambda_X(t) = U(DX_t)U^{-1}$$

has the same set of eigenvectors and we can evaluate the transition probabilities as

$$\begin{aligned} P(s,t) &= E_s\left[\prod_{s<u\leq t} I + U(DX_u)U^{-1}\,du\right] \\ &= UE_s\left[\prod_{s<u\leq t} I + DX_u\,du\right]U^{-1} \\ &= UE_s\left[e^{\int_s^t DX_u\,du}\right]U^{-1} \end{aligned}$$

If X_t is an affine process we can solve this by the technique described previously and

$$P(s,t) = Ue^{\mathfrak{A}(s,t)+\mathfrak{B}(s,t)X_s}U^{-1}$$

where $\mathfrak{A}(s,t), \mathfrak{B}_1(s,t), \ldots, \mathfrak{B}_n(s,t)$ are $K \times K$ diagonal matrices where each element solve an ODE of the same type as (2.2) and (2.3) and $\mathfrak{B}(s,t)$ is the collection of $\mathfrak{B}_1(s,t), \ldots, \mathfrak{B}_n(s,t)$.

We are interested in finding the price of a rating dependent payment, $C(T, \eta_T)$. The price of such a payment at time t given that $\eta_t = i$ is

$$v(t,T,i) = E_t^Q \left[e^{-\int_t^T R(X_s)\,ds} \sum_{k=1}^{K} \left(P_X(t,T)\right)_{ik} C(T,k) \right]$$

Now, if $\bar{C}(T)$ is deterministic

$$\begin{aligned} \bar{v}(t,T) &= E_t^Q \left[e^{-\int_t^T R(X_s)\,ds} P_X(t,T) C_T \right] \\ &= U E_t^Q \left[e^{\int_t^T DX_s - R(X_s) I\,ds} \right] U^{-1} C_T \end{aligned}$$

where we can use (2.1) if $R(x) = \rho_0 + \rho_1 \cdot x$.

A problem with stochastic eigenvalues and constant eigenvectors is that we can get a positive probability that some of the intensities are negative. This problem is very similar to the problem with negative interest rates in the Vasicek model. Mathematically this is clearly incorrect but it can be viewed as an approximation to using the generator $\Lambda_X(t)^+$ defined by

$$\begin{aligned} \left(\Lambda_X(t)^+\right)_{ij} &= \max(\lambda_{ij}, 0) \quad i \neq j \\ \left(\Lambda_X(t)^+\right)_{ii} &= -\sum_{i \neq j} \left(\Lambda_X(t)^+\right)_{ij} \end{aligned}$$

One way to ensure the positivity of the intensities is to assume

$$\Lambda_X(t) = A g(t, X_t)$$

for a positive function g and $A \in \mathbb{R}^{K \times K}$. For example A could be the observed matrix of transition intensities and g could be a function of the risk adjustment and firm specific variables. This method is equivalent to having a stochastic maturity. To see this remember that for $g = 1$

$$P(t,T) = e^{A(T-t)}$$

and in general

$$P(t,T) = E\left[e^{A \int_t^T g(s,X_s)\,ds} \right]$$

so the maturity is $\int_t^T g(s, X_s)\,ds$ instead of $T - t$.

A problem with this model is that $\frac{\lambda_{ij}}{\lambda_{kl}}$ is constant for every $i,j,k,l = 1,\ldots,K$. However, notice that the fraction between the spreads observed on different firms see Figure 2.4 do not have to be independent of the state variables since we allow for firm specific variables in the function g.

2.5 Analysis of rating based models

For corporate indexes we have a lot of data available but for single firms with special features, e.g. highly volatile firms or counter cyclic firms, the data might be inadequate. However, we might still be able to use some of the general data, for example yield spreads of indexes or estimated generators. In the following we will analyze different rating models to get an idea about which variables can be implied out from the prices and for which parameters can we use the empirically estimated values.

2.5.1 Empirical transition intensities

Since the step-up provision is linked to the rating of Deutsche Telekom it is natural to use a rating based model for default. For a bond with step-up we are interested in the transition probabilities between rating classes, not just the default probabilities. The simplest model is to assume independence between the interest rate and the transition probabilities and assume no risk adjustment of the transition intensities. For this model we will also assume that there is 0-recovery. In this case the pricing formula for a corporate bond with rating dependent coupons is

$$\bar{v}(t,T) = \sum_{i=1}^{N} p(t,T_i)P(t,T_i)\bar{C}_{T_i}$$

where T_i are the payment dates for $i = 1, \ldots, N$.

On CreditMetrics's website we have downloaded an empirical estimated matrix[5] of 1-year transition probabilities. In Table 2.1 we have given these estimates for a rating system with 17 non-default rating categories. Using these transition probabilities and the yields on the treasuries shown in Figure 2.3 we find theoretical corporate bond prices of the three Deutsche Telekom bonds. The results are shown in Figure 2.7.

We can see that these prices are all too high compared to the observed prices even though we have assumed 0-recovery which is not a realistic assumption. This suggests that a risk adjustment on the transition probabilities is needed to get a higher discounting and hereby lower prices. This is of course not surprising. Furthermore, the differences between the theoretical prices of the three bonds are quite different from the observed differences. Another problem with this model is that bond prices would be the same for different firms if they are rated the same.

[5]The matrix of empirically estimated transition probabilities was downloaded on April 7 2001.

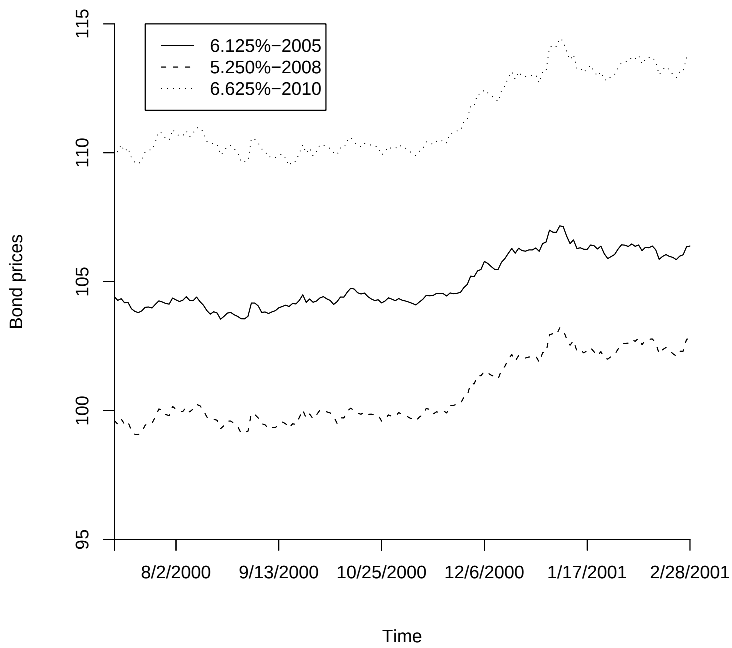


Figure 2.7: Theoretical bond prices using empirically estimated transition probabilities.

2.5.2 Stochastic recovery rate

In Das and Tufano [15] they propose a model with a stochastic recovery rate such that prices differ for firms with the same rating. Their model is simple to calibrate to the observed prices and works well for a single issue of bonds. However, in our case we have 3 different bond issues from the same firm with identical recovery rates.

Define the risk adjustment by $b \in \mathbb{R}$ and let the transition probabilities under Q be denoted by $Q(s,t)$. Define the one year risk adjusted transition probabilities by

$$Q(0,1) = (1-b)I + bP(0,1)$$

where $P(0,1)$ is the matrix of empirical transition probabilities given in Table 2.1. As we can see in Table 2.1 default is defined as the last category which is absorbing.

We will use a fraction, δ_t, of the promised payment as recovery. This is the same assumption used in Jarrow, Lando, and Turnbull[33] and in Das and Tufano[15]. Define

$$\bar{v}(t,T) = \sum_{i=1}^{N} Q(t,T_i) E_t^Q \left[e^{-\int_t^{T_i} r_s\, ds} \bar{C}_{T_i} \right]$$

where $\left(\bar{C}_{T_i}\right)_K = \delta_{T_i} \left(\bar{C}_{T_i}\right)_{K-1}$ and $\left(\bar{C}_{T_i}\right)_k$ is the payment at time T_i if the rating is $k = 1, \ldots, K-1$.

We wish to calibrate the recovery to the observed prices. In this setup all we are interested in is the expectation of the recovery rate under the forward measure, which we will denote $\delta(t,T_i)$. This is what we will calibrate, hence we only need a time dependent recovery rate. Now, the pricing equation is

$$\bar{v}(t,T) = \sum_{i=1}^{N} p(t,T_i) Q(t,T_i) \bar{C}(t,T_i)$$

where $\left(\bar{C}(t,T_i)\right)_K = \delta(t,T_i) \left(C(t,T_i)\right)_{K-1}$ and $\left(\bar{C}(t,T_i)\right)_k = \left(\bar{C}_{T_i}\right)_k$ for every $k = 1, \ldots, K-1$.

Since we only have a 5, 8, and 10 years bond we assume that $\delta(t,1) = \cdots = \delta(t,5)$, $\delta(t,6) = \cdots = \delta(t,8)$, and $\delta(t,9) = \delta(t,10)$. First, we find $\delta(t,1)$ by calibrating the 2005 bond, then we find $\delta(t,6)$ by calibrating the 2008 bond, and finally we find $\delta(t,9)$ by calibrating the 2010 bond. We have done this for the data on October 10 2000 for different choices of risk adjustments and the results are shown in Figure 2.8. As we can see we will need a rather large risk adjustment to get recovery rates between 0 and 1. Actually, we get negative probabilities in the diagonal which is because of the way we have chosen to change measure. This is similar to the method used in Das and Tufano[15]. They use a different risk adjustment for each row whereas we use a single adjustment for the entire matrix. Furthermore, the expected recovery rates changes with more than 0.2 for every risk adjustment. It is hard to find an economic reason for this large increase in expected recovery rates. Also if we choose other ways to change measure the conclusion will be the same, that the expected recovery rate has to vary a lot over time.

In both Duffie and Singleton[26] and Lando[43] it is shown that there is a close relationship between the default intensity and the recovery rate of pre-default market value. The risk adjustment for default with a positive recovery rate should be $\lambda_t(1-\delta_t)$ so not surprisingly a higher default intensity can be compensated by a higher recovery rate. So it might be hard to distinguish the two. Instead of making the recovery rate firm specific and stochastic we will in the following let the intensities be firm specific and stochastic.

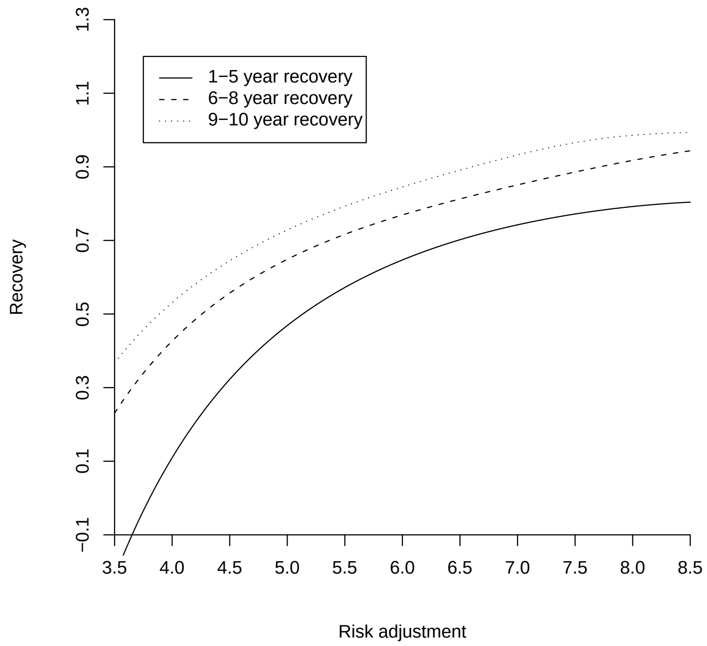


Figure 2.8: Recovery rates chosen such that the theoretical bond prices matches the observed prices on October 10 2000 for different values of the risk adjustment.

2.5.3 Q-dynamics for the Markov Chain

It is not easy to find a specific generator for Deutsche Telekom with 17 different non-default ratings. We are only interested in the behavior under Q so we might not need the empirical structure with 17 non-default ratings. We only need to distinguish between different payments. Therefore, we will collect all ratings giving the same coupon in one rating. Since both the Deutsche Telekom bonds with step-up provisions only have one trigger level we will consider a Markov chain with two non-default ratings, which we will try to fit to the observed data.

When the bonds with step-up provisions was issued Deutsche Telekom was rated Aa2 but as the yield spread in Figure 2.4 shows the market viewed Deutsche Telekom more as an A rated firm. Therefore, we have chosen to let the yield spread for corporate bonds with an A rating also seen in Figure

2.4 be a state variable. So the generator is a 3×3 matrix including default

$$\Lambda_X(t) = BX_t$$

where B is a matrix of coefficients which we will estimate using the observed prices on all the three Deutsche Telekom bonds.

We have estimated B by least squares. Let $v_j(t, T, k_t)$ be the theoretical bond price for bond $j = 1, 2, 3$ in this model where k_t is the rating of Deutsche Telekom at time t. Furthermore, let $w_j(t, T)$ be the observed bond price for $j = 1, 2, 3$. We find the estimated coefficient matrix as

$$\hat{B} = \arg \min_{\{B | B \text{ is a generator}\}} \sum_{i=1}^{I} \sum_{j=1}^{3} \left(v_j(t_i, T, k_{t_i}) - w_j(t_i, T)\right)^2$$

where I is the number of observations and t_i is an observation time.

For our sample we find

$$\hat{B} = \begin{bmatrix} -9.3662 & 9.3662 & 0.0000 \\ 8.2627 & -14.6603 & 6.3976 \\ 0.0000 & 0.0000 & 0.0000 \end{bmatrix}$$

Using $\hat{B}$ generates the prices shown in Figure 2.9. As we can see we are getting the right levels for the three bond prices but we do not get enough volatility. This is not surprising since the index has less volatility than a single bond issue. One way to get around this is to use a function of the index yield spread and then try to fit the volatility to the observed data. Another way to get the volatility right is to use a stochastic variable specific for Deutsche Telekom.

2.5.4 State variable for Deutsche Telekom

In this section we will model the matrix of transition intensities as

$$\Lambda_X(t) = Ag(t, X_t)$$

where A is the estimated intensities given in Table 2.2. We will assume that $Y_t = g(t, X_t)$ is a CIR process

$$dY_t = k_0 + k_1 Y_t \, dt + \sigma \sqrt{Y_t} \, dW_t$$

So Y_t can be a function of specific variables for Deutsche Telekom and possibly also the risk adjustment.

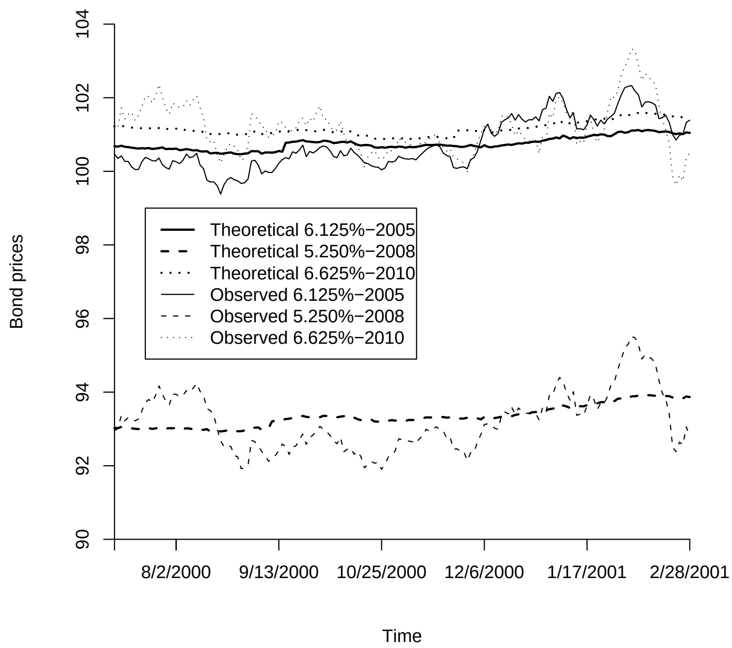


Figure 2.9: Theoretical bond prices using Q-dynamics for the generator.

A has 18 different eigenvalues, two of which are complex.[6] Let U be a matrix with the eigenvectors and D a diagonal matrix with the eigenvalues. We will assume that Y_t is independent of the treasury, then

$$\bar{v}(t,T) = \sum_{i=1}^{N} p(t,T_i) U E_t^Q \left[e^{D \int_t^{T_i} Y_s \, ds} \right] U^{-1} \bar{C}_{T_i}$$

For our calculations we have chosen the parameters

$$\begin{aligned} k_0 &= 6.5 \\ k_1 &= -2.75 \\ \sigma &= 2.45 \end{aligned}$$

Y_t is chosen to match the observed prices of the 2008 5.25% bond which is the bond with no step-up provision. Both the theoretical and observed prices are

[6] The complex eigenvalues do not create a problem. We just need to solve complex versions of (2.2) and (2.3). It can also be verified that we end up with real prices.

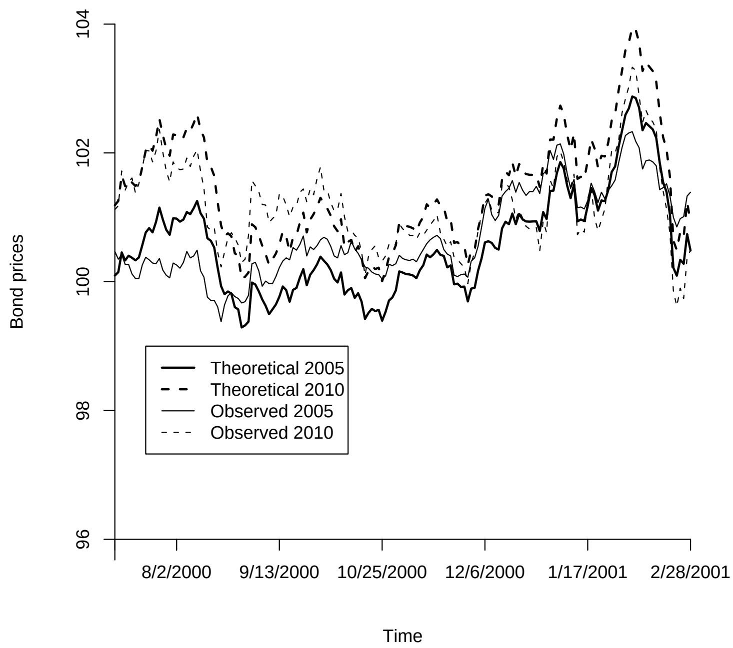


Figure 2.10: Theoretical bond prices using a specific stochastic variable for Deutsche Telekom.

shown in Figure 2.10. Since the parameters of Y_t might depend on the rating of Deutsche Telekom we have assumed that Deutsche Telekom has been rated A2 since the large issue of the two bonds with step-up provisions. As we saw in Figure 2.4 it looks as if the market already viewed Deutsche Telekom as an A rated firm at that time so this might be a close approximation. We have excluded 2008 5.25% bond since Y_t is found such that the theoretical price exactly matches the observed price. As we can see from the figure we get the right volatility and level using this model. The results might be even better for a different set of parameters. In Section 2.8 we will briefly give an estimation procedure for corporate bond observations in an affine setting.

2.6 Threshold model

The rating-based model seems like a natural choice for pricing corporate bonds with step-up provisions. One problem with the approach, however, is the need to specify the parameters for the transition intensities and the associated risk premia. The calibration methods developed to deal with this typically assume that we have a large universe of comparable bonds which behave fairly homogeneously within each rating category. For a sector undergoing very significant structural change, as is the case with the telecom sector, there is a large amount of idiosyncratic risk for each bond issuer which needs to be modeled. We can accommodate this in our rating based setting by having idiosyncratic risk factors affecting the transition intensities for each firm separately in addition to systematic terms influencing credit spreads in general. The associated model easily becomes hard to estimate. In this section we will present a computationally tractable model with fewer parameters.

Define a rating function by $\nu : \mathbb{R}^n \to [0, \infty)$ and $K - 1$ levels $0 < \gamma_1 < \gamma_2 < \cdots < \gamma_{K-1}$. For $i = 1, \ldots, K$ the rating process is defined to be

$$\eta_t = i \Leftrightarrow \nu(X_t) \in [\gamma_{i-1}, \gamma_i) \text{ and } t < \tau$$

where $\gamma_0 = 0, \gamma_K = \infty$. We also define an absorbing state D for default such that $\eta_t = D \Leftrightarrow \tau \leq t$. Now, the transition probabilities are

$$P(\eta_t = i) = E\left[e^{-\int_0^t \lambda(X_s)\, ds} 1_{\{\gamma_{i-1} \leq \nu(X_t) < \gamma_i\}}\right] \text{ for } i = 1, \ldots, K$$

and

$$P(\eta_t = D) = 1 - E\left[e^{-\int_0^t \lambda(X_s)\, ds}\right]$$

For example, if $\nu(X_t) = \lambda(X_t)$ then the rating jumps whenever the default intensity crosses a level. Another example would be $\nu(X_t) = \lambda(X_t) - \lambda^{\text{LIBOR}}(X_t)$ where $\lambda^{\text{LIBOR}}(X_t)$ is the default intensity for a refreshed AA. In this case it is the spread to the LIBOR rate which determines the rating category.

In the case where the state variable X_t represents firm value this is similar to having a classical model, as in Merton[49] where default is defined as the first hitting-time of this process to a certain boundary defined by the debt commitments of the firm. It is also similar to the CreditMetrics definition of rating transitions induced by changes in asset values. However, it is different in two important respects. First, it considers an intensity of default for all asset values, and does not link default to the first-hitting time of these assets. As shown in Duffie and Lando[22], it is consistent with a model of the Merton type - even in a diffusion based setting, to have an intensity of default as long as there is imperfect information of the firm's assets. Furthermore, contrary to, for example, the CreditMetrics approach, it allows for variables different from asset value to influence the rating movement. As shown in Sobehart et.

al.[56] there is considerable evidence that variables in addition to asset value have predictive power for default, even when asset value estimates are included as covariates. This is also consistent with the incomplete information explanation laid out in Duffie and Lando[22].

To calculate the transition probabilities we can use the method of Duffie, Pan, and Singleton[23] if we assume that both $\lambda(x), \nu(x)$ are affine in x. For $d \in \mathbb{R}^n, y \in \mathbb{R}$ define the function f by

$$f(y,t,d,X_t) = E\left[e^{-\int_t^T \rho_0+\rho_1\cdot X_s\,ds} e^{a\cdot X_T} 1_{\{d\cdot X_T \leq y\}}\right]$$

Now, the Fourier transform of f is

$$\begin{aligned}\phi(v) &= E\left[e^{-\int_t^T \rho_0+\rho_1\cdot X_s\,ds} e^{(a+ivd)\cdot X_T}\right] \\ &= e^{\alpha_{a+ivd}(T)+\beta_{a+ivd}(T)\cdot X_0}\end{aligned}$$

The Fourier transform is inverted to evaluate the function f. Assume that $\int_{\mathbb{R}} |\phi(v)|\,dv < \infty$ then for $y_1, y_2 \in \mathbb{R}$ with $y_1 < y_2$

$$f(y_2,t,d,X_t) - f(y_1,t,d,X_t) = \frac{1}{2\pi}\int_{\mathbb{R}} \frac{e^{-ivy_1} - e^{-ivy_2}}{iv}\phi(v)\,dv$$

Since ϕ is the Fourier transform of a real function $\phi(-v) = \phi(v)^*$ where x^* is the complex conjugate of x. Using this we find

$$\frac{1}{2\pi}\int_{\mathbb{R}} \frac{e^{-ivy_1} - e^{-ivy_2}}{iv}\phi(v)\,dv = \frac{1}{\pi}\int_0^\infty \frac{Im\left[\phi(v)\left(e^{-ivy_1} - e^{-ivy_2}\right)\right]}{v}\,dv$$

This technique can also be used to price the step-up provisions of the type

$$C(T,\eta_T) = c + s1_{\{\eta_T \geq i\}} = c + s1_{\{\nu(X_T) \geq \gamma_{i-1}\}}$$

if we assume that $R(x) + \lambda(x) = \rho_0 + \rho_1 \cdot x, \nu(x) = \nu_0 + \nu_1 \cdot x$. In that case the price is given as

$$E_t^Q\left[e^{-\int_t^T \rho_0+\rho_1\cdot X_u\,du}\left(c + s - s1_{\{\nu_1\cdot X_T \leq \gamma_{i-1}-\nu_0\}}\right)\right]$$

and the Fourier inversion technique can be used to evaluate this expression.

2.7 Analysis of the threshold model

KMV uses asset value and a default point to estimate a default probability. We will in the following give an example, using monthly KMV data, to find the value of the step-up feature for the Deutsche Telekom bonds. Define the asset value, X_t, as the solution to the SDE

$$dX_t = rX_t\,dt + \sigma(X_t)\,dW_t \tag{2.9}$$

and let the default point be a constant D. In Figure 2.11 we have plotted the logarithm of the default intensity $\log \lambda_t$ as a function of $\log(\frac{X_t}{D})$ for the time after the issue of the bonds with step-up provisions. As we can see this

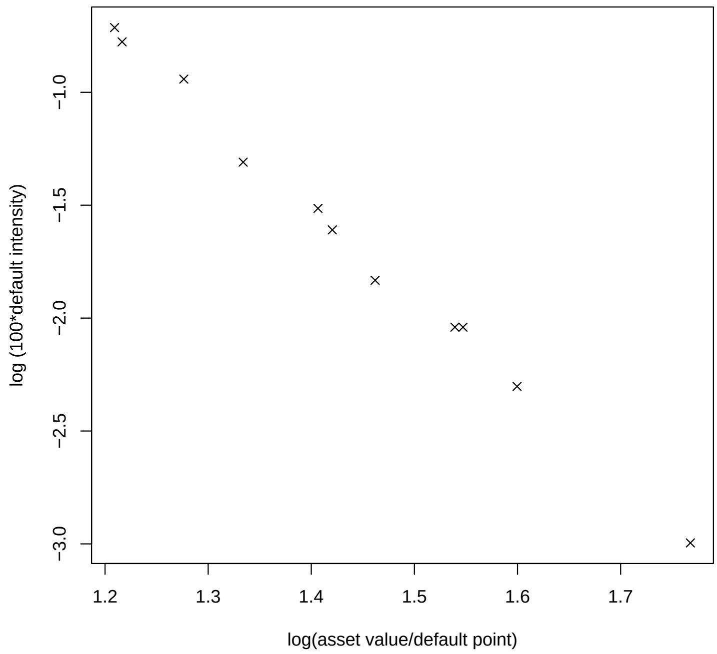


Figure 2.11: We have plotted $\log \lambda_t$ as a function of $\log(\frac{X_t}{D})$ for the time after the issue of the bonds with step-up provisions Notice the almost affine relationship.

function can be approximated by an affine function, hence, we approximate

$$\begin{aligned}\log(100\lambda_t) &= 4 - 4\log\frac{X_t}{D} \\ \Leftrightarrow \quad \lambda_t &= \frac{e^4}{100}\left(\frac{X_t}{D}\right)^{-4} \\ \Leftrightarrow \quad X_t &= D\left(\frac{100}{e^4}\lambda_t\right)^{-0.25} = D\frac{e}{\sqrt{10}}\lambda_t^{-0.25}\end{aligned} \tag{2.10}$$

Assume that λ_t is a CIR process

$$d\lambda_t = k_0 + k_1\lambda_t\,dt + \sigma\sqrt{\lambda_t}\,dW_t$$

We are interested in finding parameters for λ_t using the KMV data. Use Itô's lemma on (2.10)

$$\begin{aligned} & dX_t \\ = \; & -D\frac{e}{4\sqrt{10}}\lambda_t^{-1.25}\,d\lambda_t + D\frac{5e\sigma^2}{32\sqrt{10}}\lambda_t^{-1.25}\,dt \\ = \; & \frac{De}{4\sqrt{10}}\lambda_t^{-1.25}\left(\frac{5\sigma^2}{8} - k_0\right) - \frac{k_1}{4}X_t\,dt - \frac{5\sigma}{2e^2D^2}X_t^3\,dW_t \end{aligned}$$

and then compare this expression to (2.9). We find

$$\begin{aligned} k_0 &= \frac{5\sigma^2}{8} \\ k_1 &= -4r \\ \sigma &= -\frac{2e^2D^2}{5X_t^3}\sigma(X_t) \end{aligned}$$

We will assume that $r = 0.05$ and approximate σ by using the KMV estimate of the asset volatility, hence

$$\begin{aligned} \sigma(X_t) &= 0.2X_t \\ \Leftrightarrow \quad \sigma &\simeq 0.037 \end{aligned}$$

and we can find

$$\begin{aligned} k_0 &= 0.000853 \\ k_1 &= -0.2 \end{aligned}$$

where the long run mean is 0.0043.

In Figure 2.12 we have plotted both the default probabilities estimated by KMV and the rating given by Standard and Poor's. As we can see modeling ratings as different levels of the default intensity seems to be a good idea. Also, the increase in default probability is earlier than the rating change by Standard and Poor's. Just as the market seemed to price Deutsche Telekom as an A rated firm even before it was actually downgraded. From the figure we find a default probability of more than 0.3% is equivalent to a rating of BBB or lower.

We are now able to price the step-up provision. I.e. the possible 50 bps increase in coupon. The results are shown in Figure 2.13. Not surprisingly, the fluctuations are very similar for both bonds. Also, comparing to Figure 2.12 we can see that the value of the step-up provision increases as the default intensity increases.

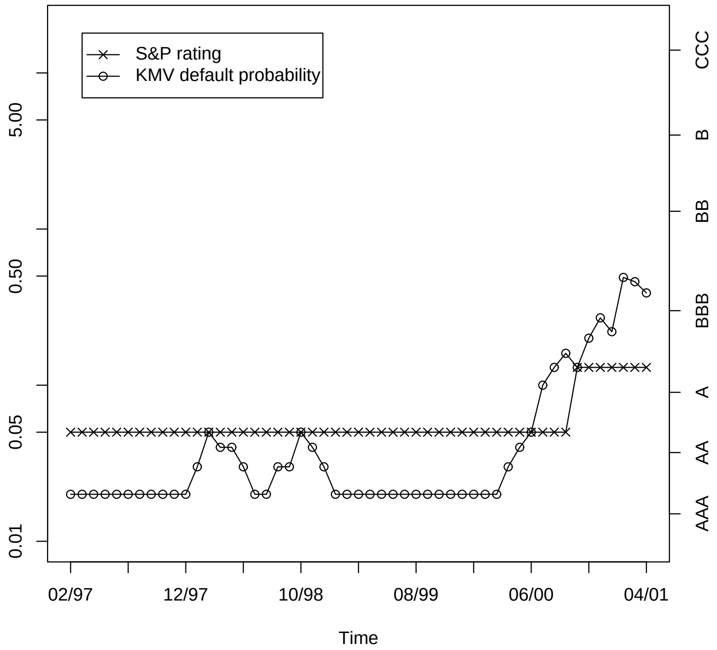


Figure 2.12: We have plotted both the rating of Deutsche Telekom given by Standard and Poor's and the estimated default probability given by KMV. Notice the log scale on the y-axis. Source Nykredit Markets and KMV.

2.8 Estimation

For affine models we can very easily calculate the characteristic function for the state variables using (2.1) with $\rho_0 = \rho_1 = 0$ and $a = iu$. However, the state variables are usually not observed. We are observing a transformation of the state variables. Assume, that we have one state variable which is the default adjusted interest rate, $X_t = r_t + \lambda_t$ defined by the SDE

$$X_t = k_0 + k_1 X_t \, dt + \sigma \sqrt{X_t} \, dW_t$$

This is only for illustrative purposes so to simplify things a little we will assume that the risk adjustment is 0, hence, X_t has the same distribution under P and Q.

We are now observing corporate bond prices $Y_{t_1}, \ldots, Y_{t_I}$ paying c_j at time

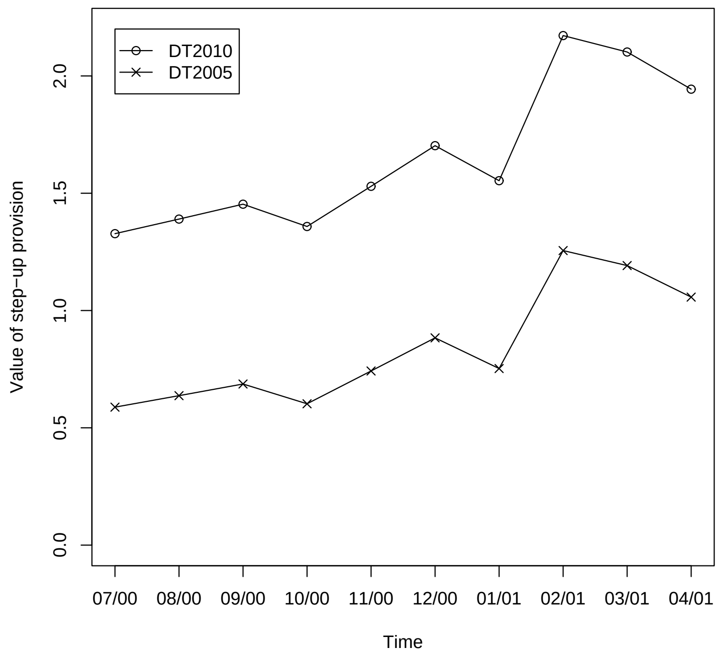


Figure 2.13: We have plotted the value of the step-up provision for both the Deutsche Telekom 2010 and 2005 bonds using the KMV data.

T_j for $j = 1, \ldots, N$. Y_{t_i} is a function of X_{t_i} since

$$Y_{t_i} = \sum_{j=1}^{N} c_j e^{\alpha_0(T_j - t_i) + \beta_0(T_j - t_i) X_{t_i}} \equiv f(t_i, X_{t_i}) \quad \text{for every } i = 1, \ldots, I$$

f is increasing in the second argument so the inverse exists and can be found numerically

$$X_{t_i} = f^{-1}(t_i, Y_{t_i})$$

To estimate the parameters of the model we will assume that the conditional distribution of Y_{t_i} given $Y_{t_{i-1}}$ is Gaussian. We can find the conditional mean as

$$
\begin{aligned}
& E\left[Y_{t_i} \middle| Y_{t_{i-1}} \right] \\
= \; & E\left[\sum_{j=1}^{N} c_j e^{\alpha_0(T_j - t_i) + \beta_0(T_j - t_i) X_{t_i}} \middle| X_{t_{i-1}} \right] \\
= \; & \sum_{j=1}^{N} c_j e^{\alpha_{\alpha_0(T_j - t_i)}(t_i - t_{i-1}) + \beta_{\beta_0(T_j - t_i)}(t_i - t_{i-1}) X_{t_{i-1}}} \\
= \; & \sum_{j=1}^{N} c_j e^{\alpha_{\alpha_0(T_j - t_i)}(t_i - t_{i-1}) + \beta_{\beta_0(T_j - t_i)}(t_i - t_{i-1}) f^{-1}(t_i, Y_{i-1})}
\end{aligned}
$$

and similarly we find the conditional variance. We now have an approximate conditional distribution of Y_{t_i} given $Y_{t_{i-1}}$ and we can use maximum likelihood estimation. In the case of Deutsche Telekom 2008 5.25% bond after the large issue of bonds with step-up provisions we find the parameters to be

$$
\begin{aligned}
k_0 &= 0.0231 \\
k_1 &= -0.329 \\
\sigma &= 0.0643
\end{aligned}
$$

Hence, the long term mean is 7.02%.

2.9 Appendix

2.9.1 Stochastic intensities

Let λ_{ij} be a stochastic variable defined as the intensity for a jump of the Markov chain from i to j. At time t the intensity for a jump to j is $\lambda_{\eta_t j}$ so a martingale is defined by

$$
1_{\{\eta_t = j\}} - \int_0^t \lambda_{\eta_u j} \, du
$$

hence

$$
\begin{aligned}
E\left[1_{\{\eta_t = j\}} - \int_0^t \lambda_{\eta_u j} \, du \middle| \mathbb{F}_s \right] &= 1_{\{\eta_s = j\}} - \int_0^s \lambda_{\eta_u j} \, du \\
\Leftrightarrow \quad P\left(\eta_t = j \middle| \mathbb{F}_s \right) &= 1_{\{\eta_s = j\}} + E\left[\int_s^t \lambda_{\eta_u j} \, du \middle| \mathbb{F}_s \right]
\end{aligned}
$$

Define

$$p_{ij}(s,t) = P\left(\eta_t = j \middle| \mathbb{F}_s\right)$$

if the Markov chain is in i at time s, i.e. $\eta_s = i$. Now,

$$\begin{aligned}
p_{ij}(s,t) &= 1_{\{i=j\}} + E\left[\int_s^t \lambda_{\eta_u j}\,du \middle| \mathbb{F}_s\right] \\
&= 1_{\{i=j\}} + \int_s^t E\left[\lambda_{\eta_u j} \middle| \mathbb{F}_s\right] du \\
&= 1_{\{i=j\}} + \int_s^t \sum_{k=1}^{K} E\left[\lambda_{\eta_u j} 1_{\{\eta_u = k\}} \middle| \mathbb{F}_s\right] du \\
&= 1_{\{i=j\}} + \int_s^t \sum_{k=1}^{K} P\left(\eta_u = k \middle| \mathbb{F}_s\right) E_{\{\eta_u = k\}}\left[\lambda_{kj} \middle| \mathbb{F}_s\right] du \\
&= 1_{\{i=j\}} + \int_s^t \sum_{k} p_{ik}(s,u) E_{\{\eta_s = i, \eta_u = k\}}\left[\lambda_{kj} \middle| \mathbb{F}_s\right] du
\end{aligned}$$

where $E_{\{\eta_t = k\}}[X]$ is the expectation of X on the set $\{\omega | \, \eta_t(\omega) = k\}$.

Differentiating with respect to t and using matrix notation

$$\frac{\partial P(s,t)}{\partial t} = P(s,t)\Gamma_s(t) \quad P(s,s) = I$$

where $\Gamma_s(t) \in \mathbb{R}^{K \times K \times K}$ with $(\Gamma_s(t))_{ikj} = E_{\{\eta_s = i, \eta_t = k\}}\left[\lambda_{kj} \middle| \mathbb{F}_s\right]$ and we define the product as $(AB)_{ij} = \sum_k A_{ik} B_{ikj}$.

	Aaa	Aa1	Aa2	Aa3	A1	A2	A3	Baa1	Baa2	Baa3	Ba1	Ba2	Ba3	B1	B2	B3	Caa	D
Aaa	.8865	.0680	.0287	.0062	.0064	.0028	.0011	.0000	.0000	.0000	.0004	.0000	.0000	.0000	.0000	.0000	.0000	0.0000
Aa1	.0295	.7950	.0815	.0652	.0238	.0019	.0000	.0021	.0000	.0000	.0010	.0000	.0000	.0000	.0000	.0000	.0000	0.0000
Aa2	.0079	.0250	.8093	.0927	.0423	.0113	.0079	.0018	.0005	.0000	.0000	.0000	.0006	.0006	.0000	.0000	.0000	0.0000
Aa3	.0012	.0039	.0321	.8058	.1010	.0375	.0090	.0021	.0026	.0020	.0000	.0009	.0011	.0000	.0000	.0000	.0000	0.0009
A1	.0005	.0009	.0067	.0483	.8134	.0771	.0299	.0081	.0033	.0016	.0046	.0034	.0007	.0015	.0000	.0000	.0000	0.0000
A2	.0003	.0007	.0022	.0065	.0575	.8069	.0740	.0329	.0081	.0044	.0029	.0012	.0013	.0003	.0006	.0000	.0003	0.0000
A3	.0005	.0010	.0002	.0022	.0150	.0892	.7554	.0676	.0395	.0144	.0063	.0019	.0024	.0040	.0004	.0000	.0000	0.0000
Baa1	.0005	.0000	.0013	.0013	.0017	.0315	.0849	.7393	.0764	.0357	.0116	.0043	.0041	.0059	.0011	.0000	.0000	0.0005
Baa2	.0000	.0011	.0015	.0014	.0015	.0091	.0353	.0773	.7521	.0752	.0193	.0054	.0072	.0053	.0048	.0027	.0000	0.0007
Baa3	.0004	.0000	.0000	.0006	.0023	.0063	.0050	.0386	.1001	.7052	.0697	.0304	.0222	.0086	.0031	.0014	.0016	0.0044
Ba1	.0010	.0000	.0000	.0000	.0021	.0011	.0067	.0090	.0333	.0736	.7376	.0505	.0396	.0100	.0134	.0109	.0039	0.0072
Ba2	.0000	.0000	.0000	.0004	.0000	.0016	.0017	.0042	.0057	.0256	.0853	.7286	.0619	.0137	.0427	.0158	.0058	0.0070
Ba3	.0000	.0003	.0002	.0000	.0000	.0020	.0018	.0014	.0025	.0087	.0258	.0544	.7590	.0287	.0579	.0244	.0071	0.0258
B1	.0003	.0000	.0003	.0000	.0006	.0005	.0020	.0009	.0034	.0038	.0039	.0251	.0687	.7619	.0275	.0496	.0100	0.0416
B2	.0000	.0000	.0008	.0000	.0014	.0000	.0006	.0014	.0011	.0017	.0038	.0172	.0372	.0576	.6789	.0818	.0280	0.0886
B3	.0000	.0000	.0006	.0000	.0000	.0000	.0004	.0015	.0016	.0022	.0020	.0043	.0152	.0502	.0347	.6977	.0522	0.1375
Caa	.0000	.0000	.0000	.0000	.0063	.0000	.0000	.0000	.0063	.0063	.0084	.0000	.0223	.0214	.0151	.0278	.6105	0.2756
D	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	1.0000

Table 2.1: A matrix of empiricaly estimated one year transition probabilities.

	Aa[a]	Aa1	Aa2	Aa3	A1	A2	A3	Baa1	Baa2	Baa3	Ba1	Ba2	Ba3	B1	B2	B3	Caa	D
Aa[a]	-.1224	.0806	.0298	.0024	.0055	.0027	.0010	.0000	.0000	.0000	.0004	.0000	.0000	.0000	.0000	.0000	.0000	.0000
Aa1	.0347	-.2359	.0997	.0752	.0224	.0000	.0000	.0025	.0000	.0000	.0013	.0000	.0000	.0000	.0000	.0000	.0000	.0000
Aa2	.0087	.0307	-.2165	.1125	.0444	.0089	.0083	.0014	.0001	.0000	.0000	.0000	.0007	.0007	.0000	.0000	.0000	.0000
Aa3	.0012	.0042	.0391	-.2231	.1227	.0403	.0071	.0008	.0026	.0022	.0000	.0008	.0012	.0000	.0000	.0000	.0000	.0010
A1	.0005	.0007	.0070	.0591	-.2144	.0923	.0332	.0069	.0024	.0009	.0053	.0040	.0003	.0017	.0000	.0000	.0000	.0000
A2	.0003	.0007	.0022	.0057	.0701	-.2243	.0919	.0381	.0058	.0036	.0026	.0009	.0013	.0000	.0007	.0000	.0004	.0000
A3	.0005	.0011	.0000	.0019	.0150	.1124	-.2928	.0859	.0469	.0147	.0063	.0014	.0020	.0046	.0000	.0000	.0000	.0000
Baa1	.0006	.0000	.0015	.0013	.0000	.0341	.1105	-.3153	.0972	.0429	.0119	.0039	.0035	.0069	.0006	.0000	.0000	.0003
Baa2	.0000	.0014	.0017	.0014	.0010	.0070	.0410	.0999	-.2988	.1003	.0202	.0040	.0068	.0054	.0059	.0029	.0000	.0000
Baa3	.0004	.0000	.0000	.0005	.0025	.0066	.0005	.0462	.1342	-.3634	.0928	.0380	.0254	.0098	.0009	.0000	.0018	.0039
Ba1	.0013	.0000	.0000	.0000	.0025	.0001	.0074	.0072	.0375	.0992	-.3145	.0650	.0481	.0103	.0144	.0124	.0043	.0049
Ba2	.0000	.0000	.0000	.0005	.0000	.0017	.0013	.0041	.0029	.0293	.1140	-.3254	.0781	.0134	.0561	.0163	.0061	.0016
Ba3	.0000	.0004	.0002	.0000	.0000	.0023	.0019	.0011	.0016	.0088	.0298	.0709	-.2836	.0332	.0773	.0270	.0071	.0218
B1	.0003	.0000	.0003	.0000	.0006	.0003	.0023	.0006	.0038	.0037	.0016	.0300	.0879	-.2774	.0322	.0644	.0109	.0385
B2	.0000	.0000	.0010	.0000	.0017	.0000	.0004	.0016	.0007	.0011	.0027	.0212	.0460	.0749	-.3949	.1151	.0381	.0903
B3	.0000	.0000	.0007	.0000	.0000	.0000	.0004	.0018	.0013	.0023	.0015	.0037	.0154	.0659	.0478	-.3675	.0787	.1480
Caa	.0000	.0000	.0000	.0000	.0089	.0000	.0000	.0000	.0084	.0084	.0115	.0000	.0301	.0286	.0206	.0399	-.5008	.3445
D	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000

Table 2.2: A matrix of transition intensities calculated from empiricaly estimated transition probabilities.

Chapter 3

Valuation of m-to-default contracts.

Acknowledgments
This chapter is based on a paper submitted to The Journal of Computational Finance. The paper has been presented at the PhD workshop in Ebeltoft, Denmark, October 27-28, 2000 and also at Danske Bank Symposium on Credit Risk Modeling: Theory and Practice, Middelfart, Denmark, January 15-16, 2001. I am grateful to David Lando for useful comments and ideas.

3.1 Introduction

An m-to-default contract is a type of OTC credit derivative whose payoff depends on the occurrence and timing of the first m defaults in an underlying basket of defaultable bonds. It is a generalization of a first-to-default swap or a basket default swap. The valuation of m-to-default contracts is useful for the contracts themselves as well as for pricing each of the tranches in collateralized debt obligations (CDO's) (for more on CDO's see Duffie and Gârleanu [20]). In a CDO the tranches are prioritized according to the number of defaults in the underlying collateral. For instance the lowest prioritization will only receive payments until a specific number of defaults have occurred in the collateral. This is exactly the structure of the payments in an m-to-default contract.

First-to-defaults are treated in Duffie [18], Kijima [36], and Kijima and Muromachi [38]. In the latter paper they also treat second-to-defaults. The general m-to-default is much more complicated than the first-to-default since in a heterogeneous pool the valuation will depend on the order in which firms default. As m increases the number of different default orderings explodes.

For a pool of 50 firms there are more than 10^{16} different orderings in which the first 10 defaults can happen. This problem is often attached by approximating the underlying pool with a homogeneous pool in which firms are independent, so that, the ordering of defaults is no longer important. Furthermore, in this case the number of defaults at any time has a binomial distribution and the valuation problem simplifies. Moody's has developed a way to construct an approximating pool based on so called diversity scores, see Cifuentes and O'Connor [11]. However, even if we can create an approximating pool where the number of defaults have the same expectation and variance as in the original pool the distribution of defaults in the original pool might be very different from the binomial distribution. Hence, we can not be sure how well the m-to-default contract is approximated.

In Duffie [18] default is described by an affine intensity process. This leads to analytical[1] pricing of the first-to-default contracts. We will decompose the m-to-default into a portfolio of first-to-defaults each of which we will price using affine intensities. Unfortunately, for a heterogeneous pool we will need a large number of first-to-default contracts to price an m-to-default analytically. The particular default ordering is no longer important only the number of ways m firms can default from the underlying pool. Again, if we have 50 firms and $m = 10$ then this number is $\binom{50}{10} \simeq 10^{10}$. We will propose a simple way to approximate the value of an m-to-default contract. Instead of making the entire pool homogeneous we will divide the pool into a number of subsets and make firms in each of the subsets homogeneous. Imagine a pool of 32 heterogeneous firms. This pool can be divided into 4 "buckets" each containing 8 firms. The new approximating pool is still a pool with 32 firms but now we only have 4 types of firms. This speeds up the pricing considerably. As we will see this approximation works really well even when we are only using a few buckets. One advantage over an approximating pool where the number of defaults is binomially distributed is that here firms in the approximating pool do not need to be independent.

The paper is organized as follows. First we will define the setup in Section 3.2. Then in Section 3.3 we will define first-to-default contracts and give some results from the literature on the pricing of these credit derivatives. In Section 3.4 we extend this to include more general m-to-default contracts which can be evaluated as a sum of first-to-default contracts. Section 3.5 adapts an affine setup to the setting of Section 3.3 such that the pricing can be done efficiently. Section 3.6 presents a way to approximate a large pool of different firms by using buckets. In Section 3.7 we study this approximation

[1] Analytical in the sense that the price of a first-to-default contract is the solution to a set of ODE's which can be solved very fast.

when valuating two different contracts. Section 3.8 extends the analysis to allow default intensities to depend on other defaults in the pool. Finally, Section 3.9 concludes. All the proofs are in Appendix 3.10.1 and in Appendix 3.10.2 we give the analytical solutions of the ODE's used in this paper.

3.2 The setup

We consider a filtered probability space $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\geq 0}, P)$ and let the filtration be defined as the natural filtration generated by the price processes. Let P be an equivalent martingale measure and all probabilities and expectations will be calculated with respect to this equivalent martingale measure.

We will also assume the existence of an interest rate process

$$dr_t = \mu(t, r_t)dt + \sigma(t, r_t)dB_t$$

which represents a (locally) riskless investment opportunity. Now, the value of any security, S_t with cumulative dividend process D_t, can be written as

$$S_t = E\left[\int_t^T e^{-\int_t^s r_u\,du} dD_s + e^{-\int_t^T r_u\,du} S_T\right] \tag{3.1}$$

see for example Duffie[17].

Furthermore, consider n stopping times $\tau_1, \ldots, \tau_n$ which can be thought of as the arrival of defaults in the underlying pool of firms. Assume the existence of n non-negative, $\mathcal{F}_t$-predictable processes $\lambda_1, \ldots, \lambda_n$ which satisfy

$$\int_0^t \lambda_i(s)\,ds < \infty \quad \text{for every } t \text{ a.s.}$$

and for which

$$1_{\{\tau_i \leq t\}} - \int_0^t \lambda_i(s)\left(1 - 1_{\{\tau_i \leq s\}}\right)\,ds$$

is a martingale for every $i = 1, \ldots, n$. Then we say that $\lambda_i(t)\left(1 - 1_{\{\tau_i \leq t\}}\right)$ is an intensity for τ_i.[2] For more details see Brémaud[8]. In the following we will refer to λ_i as the pre-default intensity which is very important for pricing defaultable claims.

Loosely speaking, the pre-default intensity process for a default time τ_i is a first order approximation of the probability that default occurs within a small time interval Δt given survival till time t

$$P\left(\tau \in (t, t + \Delta t) \middle|\, \tau > t\right) \simeq \lambda(t)\Delta t$$

[2]Notice, that the intensities are predictable with respect to $\mathcal{F}_t$ and *not* $\mathcal{F}_t \vee \sigma\left\{1_{\{\tau_1 \leq s\}}, \ldots, 1_{\{\tau_n \leq s\}}, 0 \leq s \leq t\right\}$.

3.3 Valuation of First-to-Default Contracts

We will decompose the general m-to-default contracts into a collection of first-to-defaults. Therefore, we will first review the theory of pricing first-to-default contracts in an intensity based framework. This approach is described in Duffie[18] and the most important results for this paper are summarized below. Also, the proofs of the assertions in this section can be found in Duffie[18].

The first result is a generalization of the fact that the intensity of the sum of two independent Poisson processes is the sum of the individual intensities. Therefore, the first default from a pool of firms can be described as a new default process with a higher intensity. Hence, pricing a first-to-default contract becomes very similar to pricing a defaultable bond.

Lemma 1 *Let τ_i be a stopping time with pre-default intensity process λ_i for $i \in \{1, \ldots, n\}$ and assume that $P(\tau_i = \tau_j) = 0$ for every $i \neq j$. Then $\sum_{i=1}^{n} \lambda_i$ is an pre-default intensity for the stopping time $\tau = \min(\tau_1, \ldots, \tau_n)$*

If the pre-default intensity processes λ_i all depend on a set of exogenously given state variables, X_t, then the survival probability is

$$P(\tau_i > t) = E\left[e^{-\int_0^t \lambda_i(s, X_s)\, ds}\right] \tag{3.2}$$

For more on valuation using stochastic default intensities see e.g. Lando[43]. Generally, we have to be careful when calculating survival probabilities using this formula. Kusuoka[40] gives a counter example where the survival probability under an equivalent probability measure is *not* given by (3.2).[3]

Assume we have a claim, S, which pays Z conditional upon survival of a firm till time T. In the event of default at time t the claim pays $Y(t)$. Let λ be the pre-default intensity for this firm and τ the time of default. In Duffie & Singleton[26] and Lando[43] it is shown that the promised payments no longer should be discounted with the riskless interest rate, but a risk adjusted rate. This adjustment is exactly the pre-default intensity.

Proposition 1 *Define*

$$V(t) = E_t\left[\int_t^T e^{-\int_t^s r_u + \lambda(u)\, du} \lambda(s) Y(s)\, ds + e^{-\int_t^T r_u + \lambda(u)\, du} Z\right]$$

with $V(T) = 0$. If $\Delta V(\tau) = 0$ almost surely then $S(t) = V(t)$ for every $t < \tau$.

[3] The problem with calculating survival probabilities arises when the pre-default intensities are predictable with respect to $\mathcal{F}_t \vee \sigma\left\{1_{\{\tau_1 \leq s\}}, \ldots, 1_{\{\tau_n \leq s\}}, 0 \leq s \leq t\right\}$ and not predictable with respect to $\mathcal{F}_t$. Hence, in our setup we do not have that sort of problem.

To price a first-to-default contract we combine Lemma 1 and Proposition 1, since the default of such a claim is exactly the first default from a group of firms. Hence, for a first-to-default contract the risk adjustment is the sum of the pre-default intensities from the entire group. Assume we have a group with n firms and now let S be a claim which pays Z conditional on the survival of *all* the n firms. In case of default the contract pays $Y_i(\tau_i)$ if firm i is the first firm to default and the time of default for firm i is τ_i. The pricing formula for this claim is very similar to that of Proposition 1

Proposition 2 *Define* $\lambda(t) = \sum_{i=1}^{n} \lambda_i(t)$ *and*

$$V(t) = E_t\left[\int_t^T e^{-\int_t^s r_u+\lambda(u)\,du} \sum_{i=1}^{n} \lambda_i(s) Y_i(s)\,ds + e^{-\int_t^T r_u+\lambda(u)\,du} Z\right]$$

and $V(T) = 0$. *If* $\Delta V(\tau) = 0$ *almost surely then* $S(t) = V(t)$ *for every* $t < \tau$ *where* $\tau = \min_{i=1,\dots,n} \tau_i$.

For these types of contracts default correlation is an important issue. We allow for correlation through the state variables. In some cases this might not be enough. In Davis and Lo[16] they use an infectious default model where default probabilities depend on other defaults in the pool. Here, we have implicitly assumed that each λ is a pre-default intensity for a specific firm to default. Some of the λ's, say λ_i, could also be pre-default intensities for a simultaneous default with a compensation of $Y_i(\tau_i)$ in that case. The λ's are, basically, pre-default intensities for any credit event that will terminate the contract. This means that the assumption in Lemma 1 that $P(\tau_i = \tau_j) = 0$ does not exclude simultaneous defaults of firms.

3.4 First m to default

We will define an m-to-default contract as a contract defined on an underlying pool with n firms. Let $U^{m,n}$ be a contract that pays out an amount, Y, for each of the first m defaults where Y might depend on the defaulted firm. Furthermore, let $W^{m,n}$ be a contract which only has a payoff, Z, at time T conditional on less than m defaults. Using (3.1) we find

$$\begin{aligned} U^{m,n} &= E_t\left[\int_t^T e^{-\int_t^s r_u\,du} 1_{\{N(s-)<m\}} \sum_{i=1}^{n} Y_i(s)\,dN(s)\right] \\ W^{m,n} &= E_t\left[e^{-\int_t^T r_s\,ds} 1_{\{N(T)<m\}} Z\right] \end{aligned}$$

In Proposition 2 $m = 1$ and we would have

$$V(t) = U^{1,n}(t) + W^{1,n}(t)$$

where $U^{1,n}$ and $W^{1,n}$ can be evaluated as

$$\begin{aligned} U^{1,n}(t) &= E_t\left[\int_t^T e^{-\int_t^s r_u+\lambda(u)\,du}\sum_{i=1}^n \lambda_i(s)Y_i(s)\,ds\right] \\ W^{1,n}(t) &= E_t\left[e^{-\int_t^T r_u+\lambda(u)\,du}Z\right] \end{aligned}$$

First, we will assume that there is no additional payment at maturity, T. Let firm i be one of the first m to default, and let the time of this default be $\tau_i < T$. Then the contract pays $Y_i(\tau_i)$ at time τ_i.

An m-to-default contract can be computed recursively in the following way. After the time of the first default there are only $n - 1$ firms left so the remaining contract is an $(m - 1)$-to-default of $n - 1$ firms. We will create a portfolio of $(m - 1)$-to-default contracts that exactly matches the m-to-default contract.

First, consider the case $m = 2$ and assume for simplicity that there are no simultaneous defaults. Then we can create a second to default contract from a basket of first-to-default contracts. First, buy n first-to-default contracts. In each of the contracts one party should be excluded from the underlying pool i.e. each contract is a first-to-default of $n - 1$ firms. When one party defaults $n - 1$ of the n contracts are paid out. The only one not paid out is the one where the defaulting party has been excluded. Since we only want one payment at the time of the first default we sell $n - 2$ first-to-default contracts including all n firms. This leaves one payment at the time of the first default and one contract which is a first to default out of the $n-1$ firms which are still alive. The payments are shown in Table 3.1.

Define $U_l^{m,j}(t)$ as the value of an m-to-default of j contract. The sub-vector, l, signifies which firms have been excluded of the original pool of firms. If the k'th entry is 1 then the k'th firm has been excluded otherwise the entry is 0 and the firm is still in the pool. That is $U_{e_k}^{m,n-1}(t)$ is the price of an m-to-default of $n - 1$ firms where firm k has been excluded. Now, we find

Proposition 3 *Assume that firms can not default simultaneously. Then an m-to-default contract satisfies the recursion*

$$U^{m,n}(t) = \frac{1}{m-1}\left(\sum_{k=1}^n U_{e_k}^{m-1,n-1}(t) - (n-m)U^{m-1,n}(t)\right) \tag{3.3}$$

Position	Firm excluded	Number of contracts	Payment at first default	Payment at second default
	1	1	$Y_i(t_1)$	0
	$\vdots$	$\vdots$	$\vdots$	$\vdots$
	$j-1$	1	$Y_i(t_1)$	0
	j	1	$Y_i(t_1)$	0
	$j+1$	1	$Y_i(t_1)$	0
Long	$\vdots$	$\vdots$	$\vdots$	$\vdots$
	$i-1$	1	$Y_i(t_1)$	0
	i	1	0	$Y_j(t_2)$
	$i+1$	1	$Y_i(t_1)$	0
	$\vdots$	$\vdots$	$\vdots$	$\vdots$
	n	1	$Y_i(t_1)$	0
Short	None	$n-2$	$-(n-2)Y_i(t_1)$	0
Total			$Y_i(t_1)$	$Y_j(t_2)$

Table 3.1: Portfolio of first-to-default contracts to hedge a second-to-default. Here, we assume that the first firm to default is i at time t_1 and the second is j at time t_2. We also assume that both defaults occur before maturity such that $t_1 < t_2 < T$, otherwise, we would need to multiply the payments with an indicator function $1_{\{t_. < T\}}$.

The recursion gives a way to calculate an m-to-default contract as a portfolio of first-to-default contracts, which can be priced as in Section 3.3.

For example the price of a third-to-default is

$$
\begin{aligned}
& U^{3,n}(t) \\
= & \frac{1}{2}\left(\sum_{k=1}^{n} U_{e_k}^{2,n-1}(t) - (n-3)U^{2,n}(t)\right) \\
= & \frac{1}{2}\left(\sum_{k=1}^{n}\left(\sum_{p\neq k} U_{e_k+e_p}^{1,n-2}(t) - (n-3)U_{e_k}^{1,n-1}(t)\right)\right. \\
& \left. -(n-3)\left(\sum_{k=1}^{n} U_{e_k}^{1,n-1}(t) + (n-2)U^{1,n}(t)\right)\right) \\
= & \frac{1}{2}\left(\sum_{k=1}^{n}\sum_{p\neq k} U_{e_k+e_p}^{1,n-2}(t) - 2(n-3)U_{e_k}^{1,n-1}(t) + (n-2)(n-3)U^{1,n}(t)\right) \\
= & \sum_{k=1}^{n}\sum_{p=k+1}^{n} U_{e_k+e_p}^{1,n-2}(t) - (n-3)\sum_{k=1}^{n} U_{e_k}^{1,n-1}(t) + \frac{(n-2)(n-3)}{2}U^{1,n}(t)
\end{aligned}
$$

In the last equation we use that $U_{e_k+e_p}^{1,n-2}(t) = U_{e_p+e_k}^{1,n-2}(t)$. We can see that (3.3) leads us to calculate the same first-to-default contracts over and over again. A more efficient expression is given below.

Proposition 4 *Let n be the number of defaultable firms in the pool and assume that they cannot default simultaneously. Then an m-to-default contract can be priced as a portfolio of first-to-default contracts and the price is given by*

$$
U^{m,n}(t) = \sum_{j=0}^{m-1}(-1)^{m-j-1}\binom{n-j-2}{m-j-1}\sum_{\{l\in\{0,1\}^n \mid \sum_k l_k = j\}} U_l^{1,n-j}(t) \qquad (3.4)
$$

It is possible to extend (3.4) to allow for simultaneous defaults. This can be done by letting some of the marginal pre-default intensities be a pre-default intensity for simultaneous default. In this paper we will not consider this case.

Notice that in Proposition 4 we do not need the reduced form pricing framework described in Sections 3.2 and 3.3 but as we will see in Section 3.5 this allows for a very efficient pricing technique. This result is purely based on an arbitrage argument so any m-to-default contract can be priced from a set of first-to-default contracts no matter which pricing model we choose.[4]

[4] The result is also true if we let the pre-default intensities be predictable with respect to $\mathcal{F}_t \vee \sigma\left\{1_{\{\tau_1\leq s\}},\ldots,1_{\{\tau_n\leq s\}}, 0\leq s\leq t\right\}$. However, the pricing of each of the first-to-default contracts will be more difficult.

Instead if we are interested in a payment Z at maturity conditioning on less than m defaults have occurred we can change the recursion (3.3) a little.[5] Now, the recursion is

$$W^{m,n}(t) = \frac{1}{m-1}\left(\sum_{k=1}^{n} W_{e_k}^{m-1,n-1}(t) - (n-m+1)W^{m-1,n}(t)\right) \quad (3.5)$$

To see this, if less than $m-1$ defaults occur all the contracts pay out Z and we get $\frac{n-(n-m+1)}{m-1}Z = Z$ and in case of exactly $m-1$ defaults we get $\frac{m-1}{m-1}Z = Z$ from the $m-1$ contracts where we have excluded one of the defaulted firms

Similarly to Proposition 4 we have

Proposition 5 *Define $N(t)$ as the number of defaults in the pool of n firms and let $W^{m,n}$ be the price of a contract that pays $Z1_{\{N(T)<m\}}$ at time T then*

$$W^{m,n}(t) = \sum_{j=0}^{m-1}(-1)^{m-j-1}\binom{n-j-1}{m-j-1}\sum_{\{l\in\{0,1\}^n \mid \sum_k l_k=j\}} W_l^{1,n-j}(t)$$

3.5 Analytical Solutions

The pricing of the m-to-default contracts is based upon a decomposition into a large number of first-to-default contracts. Hence, we will need an efficient method to calculate prices of these contracts. One way is to assume an affine dependence on a set of state variables which have an affine specification. In such a setting we have (almost) analytical solutions.

Let X be a p dimensional vector of state variables and define the dynamics of X by

$$dX_t = \mu(X_t)\,dt + \sigma(X_t)\,dB_t + dJ_t \quad (3.6)$$

where $\mu, \sigma\sigma^t$ are affine in X_t. B_t is a q dimensional Brownian motion and J_t is a pure jump process with jump intensity $\nu(X_t)$ which is also affine in X_t. Furthermore, the jump size distribution of J_t is independent of X_t. X_t is called an affine jump-diffusion.

[5] Another approach is to keep the recursion (3.3) and let $W_{e_k}^{m-1,n-1}$ pay out Z and $W^{m-1,n}$ pay out $\frac{n-m+1}{n-m}Z$ both conditioning on less than $m-1$ defaults. In total, if less than $m-1$ defaults occur we get $\frac{nZ-(n-m+1)Z}{m-1} = Z$ and in case of exactly $m-1$ defaults we get $\frac{m-1}{m-1}Z = Z$ from the $m-1$ contracts where we have excluded one of the defaulted firms.

In Duffie, Pan, & Singleton[23] it is shown that for $a_1, a_2, a_3 \in \mathbb{R}$ and $b_1, b_2, b_3 \in \mathbb{R}^p$

$$\begin{aligned} & E\left[e^{-\int_t^T a_1 + b_1 \cdot X_s \, ds} e^{a_2 + b_2 \cdot X_T} \left(a_3 + b_3 \cdot X_T\right) \Big| X_t \right] \\ = & \left(\alpha_1(t) + \beta_1(t) \cdot X_t\right) e^{\alpha_2(t) + \beta_2(t) \cdot X_t} \end{aligned} \tag{3.7}$$

where $\alpha_i : [0,T] \to \mathbb{R}$ and $\beta_i : [0,T] \to \mathbb{R}^p$ solve a set of ODE's for $i = 1, 2$. In that paper they also give the set of ODE's which needs to be solved. In Appendix 3.10.2 we solve the ODE's considered in this paper analytically.

We can use this to price first-to-default contracts. If we assume that $r_t, \lambda_j(t), \log Y_j(t)$, and $\log Z$ are all affine in X_t we can see from Proposition 2 that all the first-to-default contracts can be priced using (3.7).

Correlation between the default times is induced by the common state variables. The intuition behind this is that some events e.g. a recession or an earthquake are likely to make *all* the firms more volatile.

3.6 A Pool of Buckets

From (3.4) we see that to calculate the price of an m-to-default contract we need to calculate $\sum_{j=0}^{m-1} \binom{n}{j}$ first-to-default prices. For $n = 50$ and $m = 10$ this is more than 3 billion. Instead if we assume that all firms have the same pre-default intensity then expression (3.4) simplifies to

$$U^{m,n}(t) = \sum_{j=0}^{m-1} (-1)^{m-j-1} \binom{n-j-2}{m-j-1} \binom{n}{j} U^{1,n-j}(t) \tag{3.8}$$

Notice, that the number of calculations needed to evaluate this expression does not depend on the number of firms n in the pool since the number of terms in the sum is independent of n.

In general firms are not symmetric but Moody's has developed a principle in which the original pool of firms is compared to a hypothetical pool of firms in which all firms are independent and identical.[6] Now, the m-to-default can be approximated by a similar m-to-default written on the hypothetical pool and (3.8) can be used. However, in general we do not know how well we have approximated the distribution of the number of defaults.

It is difficult to find an approximating pool and even though we match the expected number of defaults with the original pool we can not be sure

[6] The number of firms in the hypothetical pool is called the diversity score and this pool should have the same loss distribution as the original pool. For more on this see Cifuentes and O'Connor[11].

that the distribution of defaults is close to the original distribution. For the contracts we are valuating in Section 3.7 we will not only be interested in a single point in the default distribution, so we need a good approximation for the entire default distribution. This can not necessarily be achieved by a binomial distribution therefore we propose a different approximation.

Instead of approximating with a homogeneous pool of firms we will collect firms in different buckets and make firms within each bucket identical. Let the number of buckets be K then K can be any number between 1 and n. If $K = 1$ all firms are identical and if $K = n$ no firms are identical. Now, we can approximate the original pool by a pool with less buckets. E.g. let $n = 4$ and let the pre-default intensity of firm i be $\lambda_i = 0.1i$ for $i = 1, \dots, 4$. This pool could be approximated by a pool with 2 buckets. In the first bucket collect firm 1 and 2 and define the pre-default intensity for bucket 1 as 0.15 which is the average. In the second bucket collect firm 3 and 4 and define their pre-default intensity as 0.35. Hence, in the new pool there are 2 firms with pre-default intensity 0.15 and 2 with pre-default intensity 0.35. This approximation will be investigated further in Section 3.7.

If we assume that the pool is divided into K buckets and let n_k be the number of firms in bucket k, i.e. $n = \sum_{k=1}^{K} n_k$. Assume for simplicity that $n_k > m - 1$ for every k,[7] then we find

$$\begin{aligned} & U^{m,n}(t) \\ = & \sum_{j=0}^{m-1} (-1)^{m-j-1} \binom{n-j-2}{m-j-1} \sum_{\substack{i_1,\dots,i_K \\ \sum_{k=1}^{K} i_k = j}} \binom{n_1}{i_1} \cdots \binom{n_K}{i_K} U^{1,n-j}_{i_1,\dots,i_K}(t) \end{aligned} \tag{3.9}$$

where the sub-indices denote the number of firms excluded from each bucket. Notice, that the number of elements in the sum depends on the number of buckets and not on the total number of firms n. For the special case $K = 2$ we have

$$\begin{aligned} & U^{m,n}(t) \\ = & \sum_{j=0}^{m-1} (-1)^{m-j-1} \binom{n-j-2}{m-j-1} \sum_{i=0}^{j} \binom{n_1}{i} \binom{n-n_1}{j-i} U^{1,n-j}_{i,n-i}(t) \end{aligned} \tag{3.10}$$

Notice, that we do not need to assume that firms are independent to use (3.9).

[7] Alternatively, we could define $\binom{n}{k} = 0$ if $k > n$.

3.7 Examples

In this section we will use the techniques described in the previous sections for pricing. The idea is to study how well we can approximate a heterogeneous pool with a pool of buckets.

Let n be the number of firms in the pool. Let $T_0, T_1, \dots, T_I$ be payment dates where $\Delta T_i \equiv T_i - T_{i-1}$ for $i = 1, \dots, I$ is assumed to be constant and $T_0 = 0, T_I = T$. For all our examples we have $I = 10$, $T = 5$, and $\Delta T_i = \frac{1}{2}$ for every i.

In both of our examples we will be interested in $\sum_{j=1}^{m} W^{j,n}(t)$ which can be calculated the following way

Proposition 6

$$\sum_{j=1}^{m} W^{j,n}(t) = \sum_{j=0}^{m-1} (-1)^{m-j-1} \binom{n-j-2}{m-j-1} \sum_{\{l \in \{0,1\}^n \mid \sum_k l_k = j\}} W_l^{1,n-j}(t)$$

We will model the pre-default intensity as

$$\lambda_i(t) = a_i + b_i^X X_i(t) + b_i^r r_t$$

where X_i for $i = 1, \dots, n$ are independent state variables all independent of r_t. In the case of identical firms we define $a = a_i, b^X = b_i^X, b^r = b_i^r$ for every i. This is just a simple way of introducing correlation in the pool. Each of the pre-default intensities depend on a firm specific variable X_i and a common variable which in this case is the interest rate.

Assume that the X_i's are affine jump-diffusion processes as defined in (3.6) with

$$\begin{aligned} \mu_X(x) &= \mu_X^0 + \mu_X^1 x \\ \sigma_X(x)^2 &= \sigma_X^0 + \sigma_X^1 x \\ \nu_X(x) &= \nu_X^0 + \nu_X^1 x \end{aligned}$$

and r_t has a CIR specification

$$\begin{aligned} \mu_r(r) &= \mu_r^0 + \mu_r^1 r \\ \sigma_r(r)^2 &= \sigma_r^1 r \\ \nu_r(r) &= 0 \end{aligned}$$

In all our examples we let $\mu_r^0 = 0.005, \mu_r^1 = -0.1, \sigma_r^1 = 0.0025$.

Now, all the first-to-default contracts can be priced using the technique described in Section 3.5. For example

$$\begin{aligned} V^{1,j}(0) &= j\delta \int_0^T e^{\alpha_r^j(t)+\beta_r^j(t)r_0+j\alpha_x(t)+j\beta_x(t)x_0} \left(A_x(t)+B_x(t)x_0\right) dt \\ &\quad + \sum_{i=1}^{p} e^{\alpha_r^j(T_i)+\beta_r^j(T_i)r_0+j\alpha_x(T_i)+j\beta_x(T_i)x_0} c_i \end{aligned}$$

where $\alpha_r^j, \alpha_x : [0,T] \to \mathbb{R}$ and $\beta_r^j, \beta_x : [0,T] \to \mathbb{R}^p$ are solutions to a set of ODE's for every j. We have chosen $\nu_X^1 = 0$ and an exponential jump distribution (with parameter 0.1) such that the ODE's can be solved analytically. See Appendix 3.10.2. This is very convenient for our purpose since we need very high accuracy of the solutions since the approximation errors are multiplied by large numbers and then subtracted to give the right result. See Proposition 6.

As we saw in Section 3.6 calculating an m-to-default can be computationally intensive if the number of different firms is large. We will study a pool of 16 firms all with different pre-default intensities. We will valuate two types of contracts defined on this pool. We will compare these values with values of contracts defined on a pool of 16 firms with only 8 different types of firms. I.e. we collect the original 16 firms in 8 different buckets and for each bucket we define a pre-default intensity (we will use a pre-default intensity which is defined by the average of the parameters in the bucket) which is then used to calculate the values of the contracts. We will continue to decrease the number of different types of firms till all firms are identical, to see the effect of using an "average" pre-default intensity for pricing instead of using the marginal pre-default intensities.

Let the parameters be given as $a_i = b_i^r = 0$, $b_i^X = 1$. Since, $b_i^r = 0$ the pre-default intensities are all independent. Furthermore, $\mu_{X_i}^1 = -0.6, \sigma_{X_i}^1 = 0.02, \sigma_{X_i}^0 = 0, \nu_{X_i}^0 = 10\mu_{X_i}^0$, and $\lambda_i(0) = -\frac{\mu_{X_i}^0 + 0.1\nu_{X_i}^0}{\mu_{X_i}^1}$ which is the long run mean.[8]

In group I we let

$$\mu_{X_i}^0 = \frac{0.012(2i-1)}{n}$$

which gives us a group of evenly distributed pre-default intensities. In group II we have clustered the pre-default intensities more around the mean keeping the minimum, maximum, and mean constant. This is done by using an

[8] With this specification of pre-default intensities the average of such pre-default intensities will be of the same type, see Duffie an Gârleanu[20]. The parameters are very similar to those used in Duffie an Gârleanu[20].

inverse normal distribution function on μ_X^0 from group I. In Figure 3.1 we have plotted $\theta_i = \frac{\mu_{X_i}^0}{\mu_{X_i}^1}$ which is the mean reversion level for each group. Table

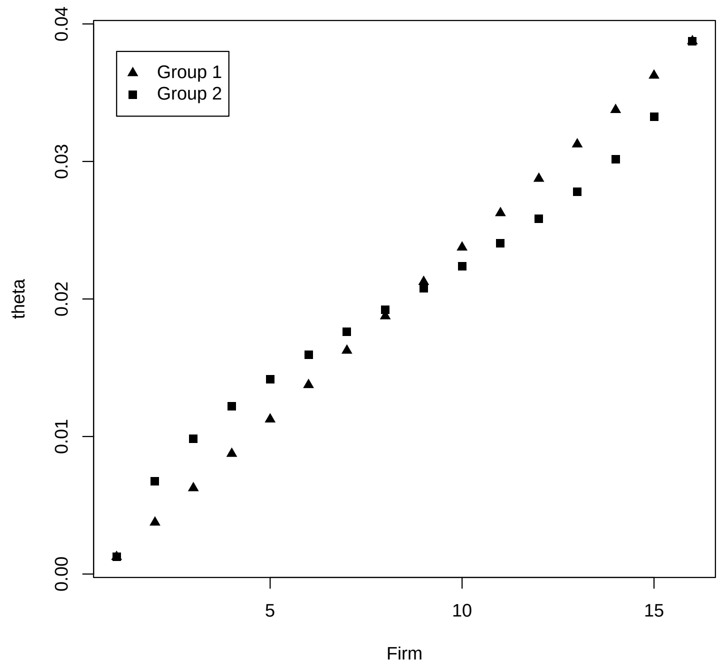


Figure 3.1: A plot of θ for both Group I and II.

3.2 and 3.3 show how the firms are collected in buckets. In each bucket we use the average parameters of the pre-default intensities of all the firms in that bucket. We have calculated the 1-year default probabilities which are also shown in Table 3.2 and 3.3. In both groups the 1-year default probabilities for firms vary from 0.2% to a maximum of 7.4%.

In Figure 3.2 we have plotted the distribution of the number of defaults for Group I after 10 years. The same figure for Group II is very similar, therefore it has not been included. As we can see the distribution does not change much for a different number of buckets. Especially, for $K = 4, 8, 16$ the distributions are very similar. Hence, we will not expect prices to differ that much.

Firm	1-year default probabilities (%)				
	$K = 16$	$K = 8$	$K = 4$	$K = 2$	$K = 1$
1	0.2476	0.4946	0.9868	1.9638	3.8891
2	0.7410				
3	1.2320	1.4765			
4	1.7205				
5	2.2066	2.4487	2.9312		
6	2.6903				
7	3.1716	3.4114			
8	3.6505				
9	4.1271	4.3645	4.8375	5.7766	
10	4.6013				
11	5.0731	5.3082			
12	5.5427				
13	6.0099	6.2426	6.7063		
14	6.4748				
15	6.9373	7.1678			
16	7.3977				

Table 3.2: Collection of firms in buckets for Group I.

3.7.1 Basket m-to-default swap

For this basket m-to-default swap we will receive δ at the time of each of the first m defaults from the underlying pool. At each of the payment dates we will pay a coupon. The coupon at T_i is defined as

$$\begin{aligned} c_i \Delta T_i &= \frac{c}{2}\,[m - N(T_i)]^+ \\ &= \frac{c}{2}\sum_{j=1}^{m} 1_{\{N(T_i)<j\}} \end{aligned}$$

such that each time a default occurs the total coupon payment is reduced.

Let U denote the default leg and W the fixed leg, i.e.

$$\begin{aligned} U^{m,n}(t) &= E\left[\int_t^T e^{-\int_t^s r_u\,du} 1_{\{N(s-)<m\}}\,dN(s)\right] \\ W^{m,n}(t) &= E\left[\sum_{i=1}^{I} e^{-\int_t^{T_i} r_s\,ds} 1_{\{N(T_i)<m\}}\right] \end{aligned}$$

<table>
<tr><th rowspan="2">Firm</th><th colspan="5">1-year default probabilities (%)</th></tr>
<tr><th>$K = 16$</th><th>$K = 8$</th><th>$K = 4$</th><th>$K = 2$</th><th>$K = 1$</th></tr>
<tr><td>1</td><td>0.2476</td><td>0.2476</td><td rowspan="3">1.1710</td><td rowspan="8">2.3748</td><td rowspan="16">3.8891</td></tr>
<tr><td>2</td><td>1.3265</td><td rowspan="2">1.6295</td></tr>
<tr><td>3</td><td>1.9315</td></tr>
<tr><td>4</td><td>2.3877</td><td rowspan="2">2.5797</td><td rowspan="5">3.0901</td></tr>
<tr><td>5</td><td>2.7714</td></tr>
<tr><td>6</td><td>3.1142</td><td rowspan="3">3.4289</td></tr>
<tr><td>7</td><td>3.4329</td></tr>
<tr><td>8</td><td>3.7385</td></tr>
<tr><td>9</td><td>4.0394</td><td rowspan="3">4.3472</td><td rowspan="5">4.6815</td><td rowspan="8">5.3799</td></tr>
<tr><td>10</td><td>4.3432</td></tr>
<tr><td>11</td><td>4.6578</td></tr>
<tr><td>12</td><td>4.9939</td><td rowspan="2">5.1809</td></tr>
<tr><td>13</td><td>5.3674</td></tr>
<tr><td>14</td><td>5.8077</td><td rowspan="2">6.0968</td><td rowspan="3">6.5325</td></tr>
<tr><td>15</td><td>6.3851</td></tr>
<tr><td>16</td><td>7.3977</td><td>7.3977</td></tr>
</table>

Table 3.3: Collection of firms in buckets for Group II.

Now, the value of the basket m-to-default swap is

$$DS^m(t) = \delta U^{m,n}(t) - \frac{c}{2}\sum_{j=1}^{m} W^{j,n}(t)$$

hence, the fair coupon is

$$c = 2\frac{\delta U^{m,n}(t)}{\sum_{j=1}^{m} W^{j,n}(t)}$$

In Table 3.4 we have calculated the fair coupon for $m = 5$ and $m = 8$ with 16 firms in the underlying pool. We can see that using 4 buckets we are only off by 1.5 bps. Similarly, in Table 3.5 we have calculated the fair fixed rate for an underlying pool with 32 firms. Here, the errors using 4 buckets are approximately 2 bps.

3.7.2 CDO

Let the underlying pool of corporate bonds be defined such that at payment dates each firm pays a constant rate $c\Delta T_i$ till default or time T whichever

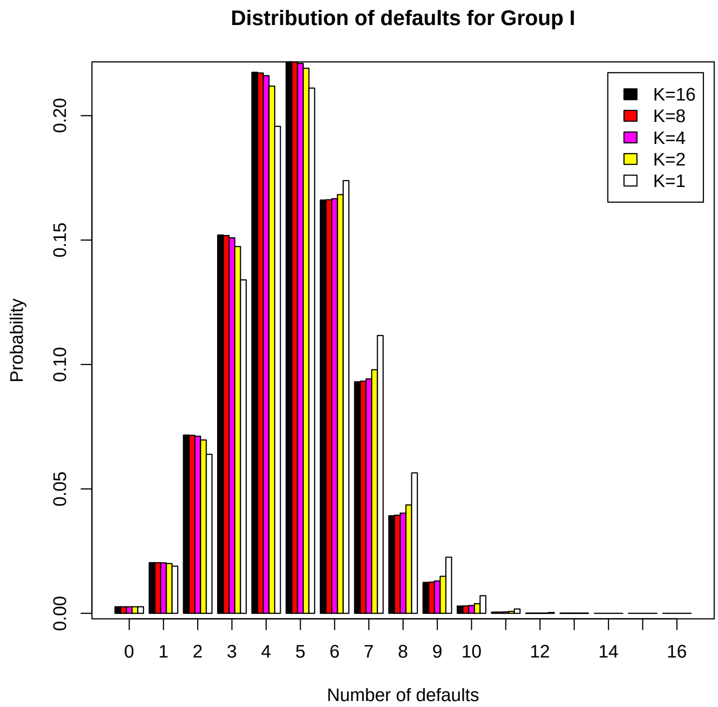


Figure 3.2: A plot of the distribution of the number of defaults for Group I after 10 years with $n = 16$. Notice, that the distributions are very similar especially for $K = 4, 8, 16$.

comes first. In case of a default we recover $Y_i(\tau) \equiv \delta$ for every $i = 1, \ldots, n$. Notice, that we are using a fraction of face value as recovery instead of a fraction of market value.[9]

The value of the entire CDO is

$$V_{CDO}(t) = \sum_{j=1}^{n} E\left[\delta \int_0^T e^{-\int_0^t r_s + \lambda_j(s)\, ds} \lambda_j(t)\, dt + \sum_{i=1}^{I} e^{-\int_0^{T_i} r_t + \lambda_j(t)\, dt} c_i\right]$$

where $c_i = c\Delta T$ for $i = 1, \ldots, I-1$ and $c_I = 1 + c\Delta T$.

Let the CDO be divided into 3 parts. An equity, junior bondholders, and

[9]Recovery of market value can be done by multiplying the pre-default intensity by 1 minus the recovery rate. See Duffie and Singleton[26]. So recovery of market value would actually be easier computationally.

Group	K	5-to-default c (%)	8-to-default c(%)
I	16	6.211	3.254
	8	6.214	3.256
	4	6.227	3.266
	2	6.276	3.304
	1	6.471	3.456
II	16	6.289	3.314
	8	6.292	3.316
	4	6.306	3.327
	2	6.351	3.362
	1	6.471	3.456

Table 3.4: Fair coupon for a basket 5-to-default and 8-to-default swap with $n = 16$ and $\delta = 1$.

senior bondholders. At payment dates the equity part receives

$$[E - N(T_i)]^+ \, c\Delta T = c\Delta T \sum_{m=1}^{E} 1_{\{N(T_i)<m\}}$$

Furthermore, the equity holders will receive the recovery of the first E defaults. In total the value of the equity part, V_E, is

$$V_E(t) = \delta U^{E,n}(t) + \frac{c}{2} \sum_{m=1}^{E} W^{m,n}(t)$$

where $U^{E,n}, W^{m,n}$ are defined as in Section 3.7.1. Similarly, we find the value of the junior bondholders and senior bondholders as

$$\begin{aligned} V_{JB}(t) &= \delta U^{JB+E,n}(t) + \frac{c}{2} \sum_{m=1}^{JB+E} W^{m,n}(t) - V_E(t) \\ V_{SB}(t) &= V_{CDO}(t) - V_E(t) - V_{JB}(t) \end{aligned}$$

In Table 3.6 we have calculated the value of the equity part of a CDO with $n = 16$ and $E = 5$. For this example there are no junior bondholders so $JB = 0$. For every CDO we use the coupon which gives the entire CDO a value of n. As we can see the value of the equity part does not change much. The relative change in $V(0)$ from $K = 16$ to $K = 2$ is approximately 0.16%.

Group	K	5-to-default c (%)	8-to-default c(%)
I	32	15.406	8.932
	16	15.407	8.933
	8	15.410	8.937
	4	15.424	8.953
	2	15.478	9.015
	1	15.694	9.260
II	32	15.526	9.071
	16	15.527	9.072
	8	15.531	9.076
	4	15.545	9.092
	2	15.585	9.137
	1	15.694	9.260

Table 3.5: Fair coupon for a basket 5-to-default and 8-to-default with $n = 32$

In Table 3.7 we have done the same exercise with $n = 32$ firms and an equity part of $E = 5$ and junior bondholders $JB = 5$. Again, we see very small changes as we change the number of buckets. E.g. the relative change of V_{E+JB} from $K = 32$ to $K = 4$ is approximately 0.06%.

In Table 3.8 we can see that the time used to calculate the prices explodes after $K = 8$ for $n = 32$, whereas the gain in accuracy is less than 0.1%.

3.8 Supplementary remarks

A major concern for derivatives on a pool of firms is the default correlation in the pool. As the model is presented in this paper we are able to model correlation through a joint dependence on some of the state variables. In some cases this might not be enough. Davis and Lo[16] develops an infectious default model where correlation is modeled through a dependence on other defaults in the pool. In an intensity setting we have to be careful if intensities depend on other defaults. Consider the example from Kusuoka[40] with two firms. Let τ_1, τ_2 be stopping times (for default) with intensities $\lambda_1(t, \tau_2), \lambda_2(t, \tau_1)$ and define $N_i(t) = 1_{\{t \geq \tau_i\}}$ for $i = 1, 2$. Let the intensities be defined as

$$\lambda_1(t, \tau_2) = \lambda_1^0(t)(1 - N_2(t)) + \lambda_1^1(t)N_2(t) \quad (3.11)$$

$$\lambda_2(t, \tau_1) = \lambda_2^0(t)(1 - N_1(t)) + \lambda_2^1(t)N_1(t) \quad (3.12)$$

Group	K	c(%)	U_E	W_E	V_E
I	16	6.8283	1.7159	1.8971	3.6130
	8	6.8295	1.7163	1.8964	3.6127
	4	6.8345	1.7181	1.8934	3.6115
	2	6.8544	1.7251	1.8819	3.6069
	1	6.9334	1.7520	1.8372	3.5892
II	16	6.8592	1.7269	1.8787	3.6056
	8	6.8604	1.7274	1.8780	3.6053
	4	6.8661	1.7293	1.8748	3.6041
	2	6.8843	1.7355	1.8645	3.6000
	1	6.9334	1.7520	1.8372	3.5892

Table 3.6: Value of equity part with $E = 5$ and $n = 16$

This captures the idea that after a default the market changes, hence the default intensity of the surviving firm also changes.

To calculate the survival probability of firm 1 we would usually use the formula

$$P(\tau_1 > t) = E\left[e^{-\int_0^t \lambda_1(s,\tau_2)\,ds}\right] \tag{3.13}$$

however, implicitly this probability depends on the intensity of a default for firm 2. From (3.12) we see that after a default of firm 1 $\lambda_2(t, \tau_1) = \lambda_1^1(t)$. When calculating firm 1 survival probabilities it is clearly not interesting what happens after a default of firm 1. Hence, (3.13) is wrong. Instead we define the firm 1 pre-default intensities (i.e. $N_1(t) \equiv 0$ for every t) as

$$\begin{aligned} \bar{\lambda}_1(t, \bar{\tau}_2) &= \lambda_1^0(t)(1 - \bar{N}_2(t)) + \lambda_1^1(t)\bar{N}_2(t) \\ \bar{\lambda}_2(t) &= \lambda_2^0(t) \end{aligned}$$

where $\bar{\lambda}_1(t, \bar{\tau}_2)$ is the intensity for the stopping time $\bar{\tau}_1$ and $\bar{\lambda}_2(t)$ is the intensity for the stopping time $\bar{\tau}_2$ with $\bar{N}_2(t) = 1_{\{t \geq \bar{\tau}_2\}}$. Now, we can calculate the survival probability as

$$P(\tau_1 > t) = E\left[e^{-\int_0^t \bar{\lambda}_1(s,\bar{\tau}_2)\,ds}\right]$$

Group	K	c (%)	U_E	W_E	V_E	U_{E+JB}	W_{E+JB}	V_{E+JB}
I	32	6.8280	2.2014	0.7318	2.9332	3.5676	3.5011	7.0685
	16	6.8283	2.2014	0.7318	2.9332	3.5679	3.5008	7.0687
	8	6.8295	2.2014	0.7318	2.9332	3.5690	3.4989	7.0678
	4	6.8345	2.2017	0.7316	2.9333	3.5734	3.4910	7.0644
	2	6.8544	2.2027	0.7311	2.9338	3.5908	3.4599	7.0407
	1	6.9334	2.2067	0.7289	2.9356	3.6561	3.3440	7.0000
II	32	6.8717	2.2036	0.7305	2.9341	3.6060	3.4327	7.0387
	16	6.8720	2.2036	0.7305	2.9341	3.6062	3.4323	7.0385
	8	6.8736	2.2037	0.7305	2.9342	3.6075	3.4300	7.0375
	4	6.8786	2.2040	0.7303	2.9343	3.6117	3.4225	7.0342
	2	6.8934	2.2047	0.7299	7.0346	3.6239	3.4009	7.0248
	1	6.9334	2.2067	0.7289	2.9358	3.6561	3.3440	7.0000

Table 3.7: Value of equity part and junior bondholders part with $E = 5$, $JB = 5$, and $n = 32$

To see this

$$
\begin{aligned}
& P(\tau_1 > t) \\
= \; & P(\tau_1 > t, \tau_2 > t) + P(\tau_1 > t, \tau_2 \le t) \\
= \; & P(\bar{\tau}_1 > t, \bar{\tau}_2 > t) + P(\bar{\tau}_1 > t, \bar{\tau}_2 \le t) \\
= \; & P(\bar{\tau}_1 > t |\, \bar{\tau}_2 > t)\, P(\bar{\tau}_2 > t) + \int_0^t P(\bar{\tau}_1 > t |\, \bar{\tau}_2 = s)\, P(\bar{\tau}_2 \in \, ds) \\
= \; & e^{-\int_0^t \lambda_1^0(u)\, du} P(\bar{\tau}_2 > t) + \int_0^t e^{-\int_0^s \lambda_1^0(u)\, du - \int_s^t \lambda_1^1(u)\, du} P(\bar{\tau}_2 \in \, ds) \\
= \; & \int_0^\infty e^{-\int_0^{s \wedge t} \lambda_1^0(u)\, du - \int_{s\wedge t}^t \lambda_1^1(u)\, du} P(\bar{\tau}_2 \in \, ds) \\
= \; & \int_0^\infty e^{-\int_0^t \bar{\lambda}_1(u,s)\, du} P(\bar{\tau}_2 \in \, ds) \\
= \; & E\left[e^{-\int_0^t \bar{\lambda}_1(u, \bar{\tau}_2)\, du}\right] \qquad (3.14)
\end{aligned}
$$

Here, we have assumed that λ_1^0, λ_1^1 are time dependent functions. They could also have been stochastic processes. Then conditioning on the path of the process it is just a time dependent function. After conditioning take expectations on both sides and the result is still valid.

The transformation to $\bar{\tau}$ simplifies the conditional probability in the fourth equality significantly. For example if λ_1^0, λ_1^1 are constants then for

K	W_E		W_{E+JB}	
	Δ	CPU	Δ	CPU
32	0	2.6	0	2699.9
16	0	0.2	3	33.1
8	0	0.0	23	0.4
4	2	0.0	102	0.0
2	8	0.0	412	0.0
1	29	0.0	1572	0.0

Table 3.8: CPU is the CPU-time (in sec) used to calculate prices for each K, and Δ is the difference (in bps) to the $K = 32$ price (in Group I).

the stopping times $\bar{\tau}_1, \bar{\tau}_2$ we have

$$P\left(\bar{\tau}_1 > t \middle| \bar{\tau}_2 > t\right) = e^{-\lambda_1^0 t}$$

whereas for τ_1, τ_2

$$P\left(\tau_1 > t \middle| \tau_2 > t\right) = \frac{\lambda_1^0 + \lambda_2^0 - \lambda_2^1}{\lambda_1^0 e^{(\lambda_1^0 + \lambda_2^0 - \lambda_2^1)t} + \lambda_2^0 - \lambda_2^1}$$

Hence, we can not expect the conditional survival probability of τ_1 given τ_2 to be of an exponential form. The difference is that $\bar{\tau}_2$ has no information about $\bar{\tau}_1$ whereas τ_2 clearly depends on τ_1 and therefore the conditional probability gets more complicated. (3.14) is also shown more generally in Kusuoka[40].

When pricing derivatives on a pool of firms the problem is the same. Therefore, the intensities used for pricing first-to-default contracts are pre-default intensities. A pre-default intensity is an intensity assuming no default has occurred in the pool. Remember, for this contract it is not interesting what happens after the first default. For general m-to-default contracts the problem is more complicated since after the first default the intensities of the surviving firms will change. Hence, the decomposition of an m-to-default contract into a portfolio of first-to-default's is very convenient in this setup. In most of the first-to-default contracts some of the firms have been excluded from the pool. Default of the excluded firms will then work as state variables. In the example with two firms $\bar{N}_2(t)$ worked as a state variable when calculating the survival probability of firm 1.

For example let the number of firms be three and define the default intensities by

$$\lambda_t^i = X_t^i + \sum_{j=1}^{3} Y_t^{ij} 1_{\{\tau_j < t\}} \text{ for } i = 1, 2, 3$$

To calculate the price of a second-to-default we decompose it into first-to-default's. This approach was not based on any model framework so it is also applicable here, hence

$$W^{2,3}(t) = \sum_{i=1}^{3} W_{e_i}^{1,2}(t) - 2W^{1,3}(t)$$

$W^{1,3}(t)$ can be priced using the regular pre-default intensities which are X_t^i for $i = 1, 2, 3$ in this case. Therefore,

$$W^{1,3}(t) = E_t^Q \left[e^{-\int_t^T r_s + \sum_{i=1}^{3} X_s^i, ds} Z \right]$$

Define

$$\bar{\lambda}_t^i = X_t^i + Y_t^{i1} 1_{\{\bar{\tau}_1 < t\}} \text{ for } i = 2, 3$$

where $\bar{\tau}_1$ is a stopping time with intensity X_t^1 then

$$W_{e_1}^{1,2}(t) = E_t^Q \left[e^{-\int_t^T r_s + \bar{\lambda}_s^2 + \bar{\lambda}_s^3 \, ds} Z \right]$$

We can price $W_{e_2}^{1,2}(t), W_{e_3}^{1,2}(t)$ in a similar way.

3.9 Concluding Remarks

We have given a way to decompose an m-to-default into a portfolio of first-to-defaults. We have been working in a reduced form setting where pre-default intensities are affine combinations of exogenously given state variables. In this type of setting we can achieve analytical solutions for the first-to-default contracts. The decomposition is only based on an arbitrage argument so this result does not depend on the choice of model chosen to price the first-to-default contracts.

For a heterogeneous pool of firms the number of calculations needed might be too large to be handled efficiently. Therefore we approximated the heterogeneous pool with pools with less types of firms but more firms of each type. It turned out that this approximation works well. In our examples the approximation was both fast and accurate.

3.10 Appendix

3.10.1 Proofs

Proof of Proposition 3.

To see the recursion, consider the time of the first default and let i be the defaulted firm. Then $U_{e_k}^{m-1,n-1}$ pays out a premium of $Y_i(\tau_i)$ if $i \neq k$, i.e. we receive $\frac{n-1}{m-1}Y_i(\tau_i)$. This is partly cancelled by the premium of $\frac{n-m}{m-1}Y_i(\tau_i)$ we have to pay for shorting $\frac{n-m}{m-1}$ of $U^{m-1,n}$. In total we receive $\frac{n-1-n+m}{m-1}Y_i(\tau_i) = Y_i(\tau_i)$. Until default number $m-1$ the situation is the same since all the contracts are still paying out premiums. After this time only the $m-1$ contracts $U_{e_k}^{m-1,n-1}$ where k has already defaulted pay out dividends. At this time these contracts are all first-to-defaults out of the remaining $n-m+1$ firms. I.e. at the time of the m'th default we receive the premium $\frac{m-1}{m-1}Y_j(\tau_j) = Y_j(\tau_j)$ where j is the m'th firm to default, and after this time no more premiums are paid out.

■

Proof of Proposition 4.
We will show (3.4) by induction of m. For $m = 1$ (3.4) is obvious. Consider $m = 2$

$$\begin{aligned} U^{2.i}(t) &= \sum_{\substack{j=0 \\ j=\sum l_k}}^{1} (-1)^{1-j} \binom{i-j-2}{1-j} U_l^{1,i-j}(t) \\ &= \sum_{k=1}^{i} \binom{i-3}{0} U_{e_k}^{1,i-1}(t) - \binom{i-2}{1} U^{1,i}(t) \\ &= \sum_{k=1}^{i} U_{e_k}^{1,i-1}(t) - (i-2)U^{1,i}(t) \end{aligned}$$

Now, assume that (3.4) is true for $U^{m-1,n}$ then

$$
\begin{aligned}
U^{m,n}(t) &= \frac{1}{m-1}\left(\sum_{k=1}^{n} U_{e_k}^{m-1,n-1}(t) - (n-m)U^{m-1,n}(t)\right) \\
&= \frac{1}{m-1}\left(\sum_{k=1}^{n} \sum_{\substack{j=0 \\ j=\sum l_i \\ l_k=0}}^{m-2} (-1)^{m-j-2}\binom{n-j-3}{m-j-2} U_{l+e_k}^{1,n-j-1}(t)\right. \\
&\qquad \left. -(n-m)\sum_{\substack{j=0 \\ j=\sum l_i}}^{m-2} (-1)^{m-j-2}\binom{n-j-2}{m-j-2} U_l^{1,n-j}(t)\right) \\
&= \frac{1}{m-1}\left(\sum_{\substack{j=1 \\ j=\sum l_i}}^{m-1} j(-1)^{m-j-1}\binom{n-j-2}{m-j-1} U_l^{1,n-j}(t)\right. \\
&\qquad \left. -(n-m)\sum_{\substack{j=0 \\ j=\sum l_i}}^{m-2} (-1)^{m-j-2}\binom{n-j-2}{m-j-2} U_l^{1,n-j}(t)\right)
\end{aligned}
$$

For the last equality we have used that $U_{e_k+e_i}^{1,n-2}(t) = U_{e_i+e_k}^{1,n-2}(t)$ and changed the limits on the sum.

Now, we evaluate the sums for each j separately. For $j = m-1$

$$
\begin{aligned}
& \frac{1}{m-1}\sum_{\sum l_k = m-1} (m-1)U_l^{1,n-m+1}(t) \\
=& \sum_{\sum l_k = m-1} U_l^{1,n-m+1}(t)
\end{aligned}
$$

For $j = 0$

$$
\begin{aligned}
& (-1)^{m-1}\frac{n-m}{m-1}\binom{n-2}{m-2} U^{1,n}(t) \\
=& (-1)^{m-1}\binom{n-2}{m-1} U^{1,n}(t)
\end{aligned}
$$

and finally for $0 < j < m-1$

$$\begin{aligned}
&\frac{1}{m-1}\left(j(-1)^{m-j-1}\binom{n-j-2}{m-j-1}U_l^{1,n-j}(t)\right.\\
&\quad\left.-(n-m)(-1)^{m-j-2}\binom{n-j-2}{m-j-2}U_l^{1,n-j}(t)\right)\\
=\;& (-1)^{m-j-1}\left(\frac{j}{m-1}\binom{n-j-2}{m-j-1}+\frac{n-m}{m-1}\binom{n-j-2}{m-j-2}\right)U_l^{1,n-j}(t)\\
=\;& (-1)^{m-j-1}\binom{n-j-2}{m-j-1}\left(\frac{j}{m-1}+\frac{m-j-1}{m-1}\right)U_l^{1,n-j}(t)\\
=\;& (-1)^{m-j-1}\binom{n-j-2}{m-j-1}U_l^{1,n-j}(t)
\end{aligned}$$

■

Proof of Proposition 5.
Substitute the old recursion (3.3) with the new recursion (3.5) and use the proof of Proposition 4.

■

Lemma 2

$$\sum_{k=0}^{p}(-1)^k\binom{n}{k} = (-1)^p\binom{n-1}{p}$$

Proof.

$$\begin{aligned}
&\sum_{k=0}^{p}(-1)^k\binom{n}{k}\\
=\;& \sum_{k=0}^{p}(-1)^k\left(\binom{n-1}{k-1}+\binom{n-1}{k}\right)\\
=\;& \sum_{k=0}^{p}(-1)^k\binom{n-1}{k}+\sum_{k=0}^{p-1}(-1)^{k+1}\binom{n-1}{k}\\
=\;& (-1)^p\binom{n-1}{p}
\end{aligned}$$

■

Proof of Proposition 6.
We will only prove Proposition 6 for identical firms. Using Lemma 2 and Proposition 5 we find that

$$\begin{aligned}
&\sum_{j=1}^{m} W^{j,n}(t) \\
&= \sum_{j=1}^{m}\sum_{k=0}^{j-1}(-1)^{j-k-1}\binom{n-k-1}{j-k-1}\binom{n}{k}W^{1,n-k}(t) \\
&= \sum_{k=0}^{m-1}\binom{n}{k}W^{1,n-k}(t)\sum_{j=k+1}^{m}(-1)^{j-k-1}\binom{n-k-1}{j-k-1} \\
&= \sum_{k=0}^{m-1}(-1)^{m-k-1}\binom{n-k-2}{m-k-1}\binom{n}{k}W^{1,n-k}(t)
\end{aligned}$$

■

3.10.2 Analytical solution of ODE's.

Consider the ODE's

$$\begin{aligned}
\beta'(t) &= -\rho_1 + K_1\beta(t) + \frac{1}{2}H_1\beta^2(t) + l_1\left(\phi(\beta(t)) - 1\right) \\
\alpha'(t) &= -\rho_0 + K_0\beta(t) + \frac{1}{2}H_0\beta^2(t) + l_0\left(\phi(\beta(t)) - 1\right) \\
B'(t) &= K_1B(t) + H_1\beta(t)B(t) + l_1\phi'(\beta(t)) \\
A'(t) &= K_0B(t) + H_0\beta(t)B(t) + l_0\phi'(\beta(t))B(t)
\end{aligned}$$

with initial conditions $\beta(0) = \alpha(0) = 0, B(0) = B_0, A(0) = A_0$.

If $\rho_1 \neq 0$ we have

$$B(t) = -\frac{B_0}{\rho_1}\beta'(t)$$

since

$$\begin{aligned}
B'(t) &= -\frac{B_0}{\rho_1}\beta''(t) \\
&= -\frac{B_0}{\rho_1}\left(K_1\beta'(t) + H_1\beta(t)\beta'(t) + l_1\beta'(t)\phi'(\beta(t))\right) \\
&= K_1B(t) + H_\beta(t)B(t) + l_1\phi'(\beta(t))B(t)
\end{aligned}$$

and $B(0) = -\frac{B_0}{\rho_1}\beta'(0) = B_0$.

Also,

$$A(t) = A_0 - \frac{B_0}{\rho_1}\left(\alpha'(t) + \rho_0\right)$$

since

$$\begin{aligned}
A'(t) &= K_0 B(t) + H_0 \beta(t) B(t) + l_0 \phi'(t) B(t) \\
&= -\frac{B_0}{\rho_1}\left(K_0 \beta'(t) + H_0 \beta(t)\beta'(t) + l_0 \phi'(\beta(t))\beta'(t)\right) \\
&= -\frac{B_0}{\rho_1}\alpha''(t)
\end{aligned}$$

hence,

$$\begin{aligned}
A(t) - A_0 &= \int_0^t A'(s)\, ds \\
&= -\frac{B_0}{\rho_1}\left(\alpha'(t) - \alpha'(0)\right) \\
&= -\frac{B_0}{\rho_1}\left(\alpha'(t) + \rho_0\right)
\end{aligned}$$

If $l_1 = 0$ then β is the solution to the Riccati equation which is

$$\beta(t) = -\frac{2\rho_1(e^{\gamma t} - 1)}{2\gamma + (\gamma - K_1)(e^{\gamma t} - 1)}$$

where

$$\gamma = \sqrt{K_1^2 + 2H_1\rho_1}$$

In general we have that

$$\begin{aligned}
& \int_0^t \frac{c_1 + c_2 e^{\gamma s}}{c_3 + c_4 e^{\gamma s}}\, ds \\
&= \int_0^t \frac{c_1}{c_3 + c_4 e^{\gamma s}}\, ds + \int_0^t \frac{c_2}{c_4 + c_3 e^{-\gamma s}}\, ds \\
&= \frac{c_1 t}{c_3} + \frac{c_1}{c_3 \gamma}\log\frac{c_3 + c_4}{c_3 + c_4 e^{\gamma t}} + \frac{c_2 t}{c_4} - \frac{c_2}{c_4 \gamma}\log\frac{(c_3 + c_4)\, e^{\gamma t}}{c_3 + c_4 e^{\gamma t}} \\
&= \frac{c_1 t}{c_3} + \frac{c_1}{c_3 \gamma}\log\frac{c_3 + c_4}{c_3 + c_4 e^{\gamma t}} - \frac{c_2}{c_4 \gamma}\log\frac{c_3 + c_4}{c_3 + c_4 e^{\gamma t}} \\
&= \frac{c_1 t}{c_3} + \frac{c_1 c_4 - c_2 c_3}{c_3 c_4 \gamma}\log\frac{c_3 + c_4}{c_3 + c_4 e^{\gamma t}}
\end{aligned}$$

which gives us an analytical form of $\int_0^t \beta(s)\,ds$ as

$$\begin{aligned}
&\int_0^t \beta(s)\,ds \\
= \quad & -2\rho_1 \int_0^t \frac{e^{\gamma t}-1}{\gamma+K_1+(\gamma-K_1)\,e^{\gamma t}} \\
= \quad & 2\rho_1 \left(\frac{t}{\gamma+K_1} + \frac{\gamma-K_1+\gamma+K_1}{\gamma\,(\gamma+K_1)\,(\gamma-K_1)} \log \frac{\gamma+K_1+\gamma-K_1}{\gamma+K_1+(\gamma-K_1)\,e^{\gamma t}} \right) \\
= \quad & \frac{2}{H_1} \log \frac{2\gamma e^{\frac{1}{2}(\gamma-K_1)t}}{\gamma+K_1+(\gamma-K_1)\,e^{\gamma t}}
\end{aligned}$$

To solve for asset prices as described in Section 3.5 $\phi(x)$ is defined as $\phi(x) = \int e^{xz}\,d\psi(z)$ where $\psi(z)$ is the jump distribution. See Duffie, Pan, & Singleton[23]. If the jump distribution is exponential with parameter η then $\phi(x) = \frac{1}{1-\eta x}$. In this case we also have an analytical form of $\int_0^t \phi(\beta(s))\,ds$ as

$$\begin{aligned}
&\int_0^t \frac{1}{1-\eta\beta(s)}\,ds \\
= \quad & \int_0^t \frac{\gamma+K_1+(\gamma-K_1)\,e^{\gamma s}}{(2\eta\rho_1+\gamma-K_1)\,e^{\gamma s}+\gamma+K_1-2\eta\rho_1}\,ds \\
= \quad & \frac{\gamma+K_1}{\gamma+K_1-2\eta\rho_1} t + \frac{4\eta}{2H_1+4K_1\eta-4\eta^2\rho_1} \log \frac{2\gamma}{\delta(t)}
\end{aligned}$$

where

$$\delta(t) = (2\eta\rho_1+\gamma-K_1)\,e^{\gamma t}+\gamma+K_1-2\eta\rho_1$$

In this paper we only consider the case $H_0 = 0$ but for completeness we also give $\int_0^t \beta^2(s)\,ds$ which is necessary if $H_0 \neq 0$.

$$\int_0^t \beta^2(s)\,ds = \frac{(\gamma-K_1)^2\,t + 4K_1 \log \frac{\gamma+K_1+(\gamma-K_1)e^{\gamma t}}{2\gamma} + 2H_1\beta(t)}{H_1^2}$$

Chapter 4

Swap pricing with two-sided default risk in a rating-based model

Acknowledgments
This chapter is based on a joint work with David Lando with the same title published in European Finance Review 3, pp 239-268, 1999. We are very grateful for a number of constructive comments from our discussant at the 1998 EFA meetings, Georges Hübner. We are also grateful for discussions with Klaus Toft, and for comments from an anonymous referee, Mette Juel Eriksen, and seminar participants at Den Danske Bank Symposium on Embedded Options, the 1998 Nordic Derivatives Forum in Copenhagen, The First International Conference on Risk Management in Munich, and at the Norwegian School of Economics and Business Administration. We also thank Claus Carøe, Jens Hugger, and Niels Jørgen Kokholm for helpful discussions on implementation and the Faculty of Natural Sciences, University of Copenhagen for funding a high speed computing environment which greatly reduced our computation time.

4.1 Introduction

This paper analyzes the pricing of defaultable securities in rating based models where the default of more than one agent is involved. We extend the model of Duffie and Huang[21] to a framework which explicitly takes into account the rating of each party. Although our method is by no means restricted to swap contracts we will use as our illustrative example a plain vanilla interest rate swap[1]. Our extension allows us to investigate the effects on swap spreads

[1]For an introduction to swap valuation, see for example Sundaresan[60]. For discussions on reasons for the growth of swap contracts, see Smith, Smithson, and Wakeman[55],

of early termination provisions, i.e. credit triggers, which are linked to the ratings of the contracting parties. Clearly, a credit trigger will make each counterparty look less risky, as illustrated for example in Wakeman[62], simply because the trigger eliminates those defaults that occur after a sequence of downgrades. How much this affects swap spreads can be studied using the technique presented in this paper. We also consider the following questions:

- How does the degree of rating asymmetry affect swap spreads?
- How does the swap spread vary with rating when the two parties have the same rating?
- How important is a stochastic specification of the transition intensities (as in Lando [41],[43])?

A second application of the numerical technique is to the study of default swaps. Here we study how the ratings of the reference security and the protection seller affect default swap premiums.

The inclusion of ratings is conveniently handled in an 'intensity-based' or 'reduced form' approach in which one focuses on modeling the default intensity of the parties directly.[2] In our setup, the rating category and the state variable (which is the short rate on the money market account) determine the default intensities. Introducing more state variables would complicate the numerical solution, but the reduced form approach can easily handle more firm-specific variables related to, for example, the asset value of the firm. Indeed, as shown by Duffie and Lando[22], reduced-form modeling is consistent with a full modeling of a firm's assets and liabilities, if bondholders have incomplete information on issuers' assets.

Rating-based models of default risk are popular for modeling defaultable bonds and credit derivatives since they use readily observable data which enable a financial institution to control credit risk without having to build detailed models for each counterparty. Using ratings alone may, however, result in a too crude approximation and approaches which allow stochastic variations in default intensities within each rating category are called for. To handle such extensions we will be working with the framework presented in Lando[41],[43] which extends the model presented in Jarrow, Lando, and

---

Litzenberger[46] and Brown, Harlow, and Smith[10]

[2]Papers using this technique include Artzner and Delbaen[3], Das and Tufano[15], Duffie and Singleton[26], Duffie, Schroder, and Skiadas[24], Jarrow, Lando, and Turnbull[33], Jarrow and Turnbull[34], Lando[41], Lando[43], Madan and Unal[47], Schönbucher[54]. For a survey on the various approaches to modeling default risk, see for example Lando[42].

Turnbull[33][3]. This framework allows for rating transition intensities to depend on random state variables, but requires a somewhat restrictive assumption on the stochastic transition intensities to obtain closed-form solutions for prices on defaultable bonds. And even if there may be analytical expressions for prices of corporate bonds many types of derivative prices, including swaps with settlement based on pre-default market value, may not have analytical solutions. Therefore, to allow for more complicated derivatives and to handle more realistic specifications of the stochastic transition intensities, a model based on ratings must have a numerical implementation to be of practical use. This paper provides such an implementation.

An elegant way of studying two-sided default risk in a reduced-form setting is presented in Duffie and Huang[21]. Here the authors show how to value swaps with a settlement payment depending on the pre-default market value of the contract - a problem which cannot be solved simply by analyzing the values of the floating-rate side and the fixed-rate side separately. Using a finite state space Markov chain as an additional factor in the default intensities, the solution equations become a system of quasi-linear PDEs. We pay special attention to an ADI-method which is well-suited for this problem which initially seems large due to the fact that both a spot rate and a two-dimensional rating process are involved. We also show in this paper a derivation of the valuation PDEs which does not build on recursive methods.

Earlier papers dealing with the pricing of one-sided default risk in swaps include Abken[1], Artzner and Delbaen[2], Cooper and Mello[12], Rendleman[53], Sundaresan[60], Solnik[57], and Turnbull[61]. Rendleman[53] also considers two-sided risk in an analysis based on the asset values of two firms entering into the contract and Sorensen and Bollier[58] consider one- and two-sided default risk in general terms without stating an explicit term structure model or model for default risk. Hübner[32] starts from different assumptions to obtain swap spreads analytically, and Jarrow and Turnbull[35] present a discrete-time implementation which like the above mentioned approaches does not model ratings changes before default. For empirical evidence on swap spread behavior, see the papers by Brown, Harlow, and Smith[10], Dufresne and Solnik[27], Duffie and Singleton[25], Minton[50], and Sun, Sundaresan, and Wang[59]. In this paper, we do not estimate swap spreads from actual data, so it is difficult to compare the size of our spreads with the empirical findings of these papers. Our model agrees with Duffie and Huang[21] in showing small swap spread sensitivities to changes in credit quality. This is consis-

[3]More recent contributions using rating-based modeling include Arvanitis, Gregory, and Laurent[4], Das and Tufano[15], Kijima[36], Kijima and Komoribayashi[37], Li[45], and Nakazato[51].

tent with the study of Sun, Sundaresan, and Wang[59] in which there is no significant difference in the midpoint prices (i.e. the average of bid and offer prices) for identical swap contracts bought and sold by a AAA-rated and an A-rated firm.

Since our paper is primarily concerned with implementation and comparing different specifications of the rating process, we have chosen very simple contracts to illustrate our methods. For example, the floating leg of the swap contract is tied to the default-free term structure, and not to a LIBOR rate.[4] If one viewed the AA category as being representative of LIBOR-rates, the swap spreads would increase by an amount roughly equal to the default adjusted spot rate for LIBOR. Working with a spread to treasuries allows us to focus on variations in the spread induced by variations in counterparty risk as opposed to variations induced by fluctuations in an underlying rate containing a default-risk adjustment.

We analyze one swap contract but our framework can easily handle a whole portfolio of contracts between two parties and effects of netting provisions may then be taken into account. Also, effects of using collateral could be handled in our framework but we have chosen not to include that into our analysis. Indeed, the spread results that we are getting can be used to compare the expenses of setting up a detailed system of collateral to the cost of offering a lower rated entity the same terms as, say, an AA-rated counterparty, or alternatively to say how much the spread should be increased in a contract with a lower rated counterparty posting no collateral.

The outline of the paper is as follows: In Section 4.2, we fix the notation and reiterate the model for recovery proposed by Duffie and Huang[21] and set up the relevant system of quasi-linear PDEs which arises from our model formulation. We also propose an alternative derivation which does not build upon recursive methods. Section 4.3 sets up the ADI-method for computing prices in our framework. In Section 4.4 we address various problems in swap pricing with two-sided default risk: How sensitive are spreads to credit ratings if both parties have the same rating? How sensitive are swap spreads to differences in rating? The results we obtain in this section are similar to those obtained in Duffie and Huang[21]. In Section 4.5 we consider how varying the specification of the stochastic transition intensities changes the conclusions. In Section 4.6 we look at how the inclusion of a credit trigger affects the fair swap rate. Predictably, credit triggers reduce swap spreads so the contribution is to get a feel for the magnitude of the reduction. Section 4.7 considers an application of our numerical procedure to the valuation

[4]For more on the impact on swap spreads of using LIBOR rates to determine the floating leg, see Dufresne and Solnik[27].

of default swaps. We investigate in this section the premium of a default swap as a function of the joint credit quality of the default protection seller and the underlying reference security. In Section 4.8 convergence of the numerical scheme is established and the non-default fixed rate, which is used to calculate spreads, is found. Also, different settlement rules are analyzed. Finally, Section 4.9 concludes.

4.2 The model

We consider a filtered probability space $(\Omega, \mathbb{F}, (\mathbb{F}_t)_{t\geq 0}, P)$ and assume the existence of an equivalent martingale measure Q which may or may not be uniquely defined. Jarrow, Lando, and Turnbull[33] propose one way of determining a martingale measure from an empirically observed transition matrix and an observed term structure of credit spreads.

The filtration will be defined as the natural filtration of the price and rating processes defined below. Given is a spot rate process

$$dr_t = \mu(r_t, t)dt + \sigma(r_t, t)dW_t$$

and from the associated money market account we define the discount factor

$$B_{t,s} = \exp\left(-\int_t^s r_u\, du\right)$$

such that under the martingale measure prices of zero coupon bonds can be computed as

$$P(t, T) = E^Q\left(B_{t,T}|\mathbb{F}_t\right)$$

We consider in this paper a two-dimensional stochastic process of ratings

$$\eta_t = (\eta_t^A, \eta_t^B)$$

where η is a continuous-time process with state space

$$\mathbb{L}^{K-1} = \{1, \ldots, K-1\} \times \{1, \ldots, K-1\}$$

describing the joint evolution of the rating of firms A and B. Here, η represents non-default states and the default intensities for firm i, when the joint rating of the companies is η, is given under Q by $\lambda^i(\eta)$. Later, we will also work with the formulation adopted in for example Jarrow, Lando, and Turnbull[33] and Lando[43] where default is given by adding the absorbing state K to the state space. In both cases, the modeling of the simultaneous rating process η is convenient for notational purposes and, more importantly, for allowing correlation between the two rating processes. The joint rating process allows us to include cases in which, say, default of one party triggers the default of the other.

To specify the stochastic evolution of η we use a $(K-1)^2 \times (K-1)^2$ matrix of transition intensities which we refer to (with a slight abuse of language) as a generator matrix. In the case where Λ is constant it is simply the generator of a time-homogeneous continuous-time Markov chain. When Λ is time-dependent, but deterministic, it models a time-inhomogeneous Markov chain, as used for example in Jarrow, Lando, and Turnbull[33]. When Λ depends on the stochastic short rate, we use the setup of Lando[43] where, after conditioning on a sample path of the short rate process, Λ is the generator of a non-homogeneous Markov chain.

In the case where the two rating processes are independent, we may construct Λ from the generator matrices of the chains η_t^A and η_t^B by remembering that the chain may only jump in such a way that one of the ratings changes. We illustrate this in Appendix 4.10.2. Two continuous-time chains which are independent do not jump simultaneously.

We are now interested in pricing an interest rate swap struck between the two parties, and to compute swap spreads in the usual fashion by setting the swap's price at initiation equal to zero.

A critical ingredient of the pricing of the swap is the rules for settlement in default. We work, as do Duffie and Huang[21], with the following setup: In the event of default, if the contract has positive value to the non-defaulting party, the defaulting party pays a fraction of the pre-default market value of the swap to the non-defaulting party. If the contract has positive value to the defaulting party, the non-defaulting party will pay the full pre-default market value of the swap to the defaulting party.

Mathematically, this can be stated as follows: Assume that B defaults first. Assume also for simplicity that there are no lumpy payments planned on the swap at the default date. The payment received by A at the default date is given by

$$S(\tau-, r_\tau, \eta_{\tau-}) \left(1_{\{S(\tau-,r_\tau,\eta_{\tau-})<0\}} + \phi_{\tau-}^B 1_{\{S(\tau-,r_\tau,\eta_{\tau-})>0\}}\right) \tag{4.1}$$

where S is the price of the swap contract seen from counterparty A. Then $S(\tau-, r_\tau, \eta_{\tau-})$ is the price right before default[5] and $\phi_{\tau-}^B$ is the recovery paid by the defaulting party B. A similar expression can be written stipulating what happens if A defaults first.

In terms of arbitrage pricing technology, if we think of a security as a claim to a cumulative (actual) dividend process D, (which for defaultable

[5]Note that $r_{\tau-} = r_\tau$ due to the continuity of r.

securities may be different from the promised dividend process) then in a no arbitrage setting (see for example Duffie[17], p.118), the price of the security satisfies

$$S_t = E_t^Q \left(B_{t,T} S_T + \int_t^T B_{t,s} dD_s \right) \tag{4.2}$$

If the contract we are pricing has a maturity of T, then we are left with the expression

$$S_t = E_t^Q \int_t^T B_{t,s} dD_s \tag{4.3}$$

but note that by the definition of the swap contract the cumulative dividends depend through the settlement provisions on the future values of S and also on the random default time. A key result stated for two-sided default problems in various forms in Duffie and Huang[21] and Duffie, Schroder, and Skiadas[24] tells us how to find S in terms of a pre-default process V and how to compute V as a function of *promised* cash flows, i.e. the cash flow paid if there is no default before expiration. Define

$$V(t, r_t, \eta_t) = E_t^Q \left[\int_t^T -R(s, V_s, \eta_s) V(s, r_s, \eta_s) ds + dD_s^p \right] \tag{4.4}$$

where the D_t^p is the *promised* cumulative dividend received by A up to time t and

$$\begin{aligned} R(t, v, k) &= r_t + s^A(t,k) 1_{\{v<0\}} + s^B(t,k) 1_{\{v>0\}} \\ s^i(t,k) &= (1 - \phi^i) \lambda^i(k) \end{aligned} \tag{4.5}$$

Assume that

$$\Delta V(\tau, r_\tau, k) \equiv V(\tau, r_\tau, k) - V(\tau-, r_{\tau-}, k) = 0$$

almost surely for each category k.[6] Then the result of Duffie and Huang[21] implies that the swap price is given as

$$S(t, r_t, k) = V(t, r_t, k) 1_{\{t<\tau\}}$$

[6] Note that this is a requirement that the default event is not occurring at the same time (with positive probability) as an event which would have caused a jump in the price of the security even if no change in rating had occurred. If, for example, we had a jump component in interest rates, then it would be easy to construct an example in which interest rates jump at the same time as a default. One could also imagine that transition intensities (including default intensities) are affected by the same jump process as the one triggering default, in which case the pre-default value would jump at the time of a default.

for each joint rating category k. Let $V_t = V(t, r_t, \eta_t)$ and $R_t = R(t, V_t, \eta_t)$ then the solution to (4.4) is

$$V_t = E_t^Q \left[\int_t^T e^{-\int_t^s R_u du} dD_s^p \right]. \tag{4.6}$$

Hence the equation which gives us the swap price as an expected discounted value of actual cash flows has been translated into an expression involving an expected, discounted value of promised cash flows, but with a more complicated discount factor.

We now derive two representations of the price process - one using a theorem in Duffie and Huang and one using the enlarged Markov chain which includes default categories (i.e. categories in which at least one of the parties has defaulted).

First, let $f(t, r_t, i)$ denote the (pre-default) price at time t, when the joint rating of the parties is given by i, of a contingent claim whose only promised payoff from B to A is $d(T, r_T, \eta_T)$ at time T. Let $\bar{f}$ be a $(K-1)^2$ dimensional vector of time t prices

$$\bar{f}(t, r_t) \equiv \begin{bmatrix} f(t, r_t, (1,1)) \\ \vdots \\ f(t, r_t, (K-1, K-1)) \end{bmatrix} \tag{4.7}$$

The following proposition gives a system of PDEs that this vector function must satisfy.

Proposition 7 *Define $\tilde{\Lambda}$ as a $(K-1)^2 \times (K-1)^2-$matrix with elements (whose dependence on t, r_t is suppressed)*

$$\begin{aligned} \tilde{\lambda}_{ij} &= \lambda_{ij}, \qquad i, j \in \mathbb{L}^{K-1}, \quad i \neq j \\ \tilde{\lambda}_{ii} &= \lambda_{ii} - (1-\phi^A)\lambda^A(i) 1_{\{f(t,r_t,i)<0\}} - (1-\phi^B)\lambda^B(i) 1_{\{f(t,r_t,i)\geq 0\}} \end{aligned}$$

where the transition intensities λ_{ij} and the diagonal elements λ_{ii} may depend on time and the short rate process r. Assume that for t fixed the process

$$f(s, r_s, \eta_s) \exp\left(-\int_t^s R_u du \right) - f(t, r_t, \eta_t)$$

is a semimartingale whose local martingale part is a martingale.

Then $\bar{f}$ is the solution to

$$\frac{\partial}{\partial t}\bar{f} + \mu \frac{\partial}{\partial r}\bar{f} + \frac{1}{2}\sigma^2 \frac{\partial^2}{\partial r^2}\bar{f} + \tilde{\Lambda}\bar{f} - r_t \bar{f} = 0 \tag{4.8}$$

with the boundary condition

$$f(T, r, i) = d(T, r, i) \quad \textit{for every } i \in \mathbb{L}^{K-1}.$$

The proof of the proposition is given in Appendix 4.10.1.

Note that the system of PDEs is linked through the matrix $\tilde{\Lambda}$. Intuitively, the term $\tilde{\Lambda}\bar{f}$ takes into account the price changes that can occur due to rating changes and default whereas the remaining terms keep track of changing spot rates and time. To compute $f(t, r_t, i)$ for a claim with promised lumpy dividends at time point $T_1, \ldots, T_N = T$, we use the proposition above for $T_{N-1} \leq t < T_N$ and work our way backward in the usual way: For $T_{N-2} \leq t < T_{N-1}$ use the procedure above with $d(T, r, i)$ in the right hand side of the boundary condition replaced by

$$f\left(T_{N-1}, r, i\right) + d\left(T_{N-1}, r, i\right)$$

and so forth.

It is instructive and convenient for handling general settlement rules to see the derivation of the pricing formula starting with (4.2) but using a Markov chain *including* the default categories. Again, consider a two-dimensional Markov process of ratings, $\tilde{\eta}$, with a state space $\mathbb{L}^K$ which includes the default state represented by K. Define the default space

$$\mathbb{D} = \mathbb{L}^K \backslash \mathbb{L}^{K-1}$$

consisting of states where at least one party has defaulted. Let $t \in [T_{N-1}, T)$ and consider the discounted gain process for $t < u < T$

$$G_{t,u} = B_{t,u} S_u + \int_t^u B_{t,s} dD_s \tag{4.9}$$

where $\Delta D_s = \delta(s, r_s, \tilde{\eta}_{s-}, \tilde{\eta}_s)$ is a lump sum dividend payment in the event of a transition at time s from $\tilde{\eta}_{s-}$ to $\tilde{\eta}_s$. Note that this allows us to include payments made upon transitions between non-default states, which will be useful for contracts with credit triggers. The promised dividend at time T is $d(T, r_T, \tilde{\eta}_T)$. Holding t fixed we know from Duffie[17] that the discounted gain process, $G_{t,u}$ is a martingale under Q and we have the pricing formula $S_t = E_t[G_{t,u}]$ for $t < u$. This time, let the function g be defined by

$$\begin{aligned} g(t, r_t, \tilde{\eta}_t) &= S_t, \quad t < T \\ g(T, r_T, \tilde{\eta}_T) &= d(T, r_T, \tilde{\eta}_T). \end{aligned}$$

Notice that for $\tilde{\eta}_t \in \mathbb{D}$, $g(t, r_t, \tilde{\eta}_t) = 0$ since we assume that there is a settlement payment when entering into the absorbing set of states $\mathbb{D}$, and after that no dividends are paid out. We will now find an expression for $dG_{t,u}$ and due to the martingale property of $G_{t,u}$, the drift must equal 0.

Again (see (4.30)) for $\tilde{\eta}_t \notin \mathbb{D}$ we find

$$dS_u = \left(\mathcal{D}g(u, r_u, \tilde{\eta}_u) + (\Lambda \bar{g}(u, r_u))_{\tilde{\eta}_u}\right) du + M_u \qquad (4.10)$$

where M is a martingale, $\bar{g}$ is a K^2- dimensional vector obtained by stacking the functions $g(\cdot, \cdot, i), \quad i \in \mathbb{L}^K$ and $\mathcal{D}g$ is given as in (4.29). This will give us $B_{t,u}dS_u + S_u dB_{t,u}$ easily. Now, all we need is the second part of (4.9). For $t < u < T$ there are no dividends paid out unless a transition occurs so this contribution is

$$\begin{aligned}
& \int_t^u B_{t,s} dD_s \\
= & \sum_{\{s: t < s \leq u, \Delta\tilde{\eta}_s \neq 0\}} B_{t,s}\delta(s, r_s, \tilde{\eta}_{s-}, \tilde{\eta}_s) \\
= & \hat{M}_u + \int_t^u \lambda_{\tilde{\eta}_s} \sum_{i \neq \tilde{\eta}_s} \frac{\lambda_{\tilde{\eta}_s, i}}{\lambda_{\tilde{\eta}_s}} \delta(s, r_s, \tilde{\eta}_s, i) B_{t,s} ds \\
= & \hat{M}_u + \int_t^u \sum_{i \neq \tilde{\eta}_s} \lambda_{\tilde{\eta}_s, i} \delta(s, r_s, \tilde{\eta}_s, i) B_{t,s} ds
\end{aligned}$$

where $\hat{M}$ is a martingale[7]. This can be written with matrix notation

$$\hat{M}_u + \int_t^u \left[diag\left(\Lambda \Xi_s'\right)\right]_{\tilde{\eta}_s} B_{t,s} ds \qquad (4.11)$$

where $diag$ is a vector of the diagonal elements and Ξ is a $K^2 \times K^2$ matrix whose element i, j represents the dividend payment in case of a transition from i to j.[8]

Now, insert (4.10) and (4.11) in (4.9) to see that a necessary condition for $G_{t,u}$ to be a martingale is

$$\mathcal{D}\bar{g}(s, r_s) + \Lambda \bar{g}(s, r_s) + diag(\Lambda \Xi_s') - r_s \bar{g}(s, r_s) = 0 \qquad (4.12)$$

with boundary condition $g(T, r, i) = d(T, r, i)$, $i \in \mathbb{L}^K$. If dividends associated with transition to default are given as in (4.1) and Ξ_{ij} is zero for $j \notin \mathbb{D}$, then the equations for g in the non-default categories are the same as in (4.8).

[7] After the second equality sign we have replaced $\tilde{\eta}_{s-}$ by $\tilde{\eta}_s$ which does not change the integral with respect to Lebesgue measure.

[8] The diagonal of Ξ is zero since no transitions are associated with the diagonal. As an example, we could have $\Xi_{ij} = g(t, r_t, i) - g(t, r_t, j)$, which would capture a payment at each transition date compensating for the change in contract value due to a rating change.

4.3 Numerical Solution

To justify our solution method, consider first using an ordinary implicit finite difference method to approximate the solution to (4.8). For each state of the Markov chain we would need a number of grid points to approximate the interest rate, say M. We would now have to solve MK^2 equations so for (say) $M = 50$ we would have 2450 equations for a Markov chain with 49 states. Using Gaussian elimination (the method could probably be improved by taking the special structure of the system into consideration) we would need in the neighborhood of $2450^3 \simeq 15$ billion multiplications and additions cf. Golub and Ortega[30]. On a moderate size system it takes about 13 minutes to solve such a system. In order to price a defaultable zero-coupon bond with maturity one year from today using a time step of $\Delta t = 0.1$ would take about 2 hours. Instead, we have chosen to approximate the solution of (4.8) using an *Alternating Direction Implicit Finite-Difference Method* (ADI-method) which seems to be more efficient when working with two state variables. The basic idea is to alternate which variable is implicit and which is explicit: In the first step the interest rate is implicit and the Markov chain is explicit and in the next step it is the other way around. This way we only have to solve two equation systems of dimension M and K, respectively. Think of this as first solving for the interest rate and then solving for the change in the Markov chain.

Define the approximating grid points as

$$S_n^{m,k} \simeq S(\Delta tn, \Delta rm + r_0, \Delta\eta k) = S(t, r, \eta)$$

where $n = 0, \ldots, N \equiv \frac{T}{\Delta t}$, $m = -M, \ldots, -1, 0, 1, \ldots, M$, and $k = 1, \ldots, L$, where L is the number of states in $\mathbb{L}^{K-1}$, i.e. 49 in this illustration. So for each time point we have an $(2M+1) \times L$ matrix with values of S. Strictly speaking, the values of S we compute in the following are pre-default values of the swap but this is equal to the value of the swap as long as there is no default.

Define the $(2M+1) \times L$ matrix S_n as the asset price at time point n in every grid point i.e.

$$S_n \equiv \begin{bmatrix} S_n^{M,1} & \ldots & S_n^{M,L} \\ \vdots & & \vdots \\ S_n^{-M,1} & \ldots & S_n^{-M,L} \end{bmatrix}$$

Notice, that the row numbers of S_n are opposite to the usual matrix notation.

Furthermore, define a discrete time version of $\bar{S}_t$ as

$$\bar{S}_n^m \equiv \begin{bmatrix} S_n^{m,1} \\ \vdots \\ S_n^{m,L} \end{bmatrix}$$

Instead of keeping track of which variable is implicit and which is explicit, we say that each step consist of two iterations. To do this we introduce half a step by, $t + \frac{1}{2}\Delta t$. Assume we know all the values at time point $n+1$ of S (that is for all values of m and k) then for r implicit and η explicit we approximate S at $n + \frac{1}{2}$ by using

$$\frac{\partial^2 \bar{S}_{\Delta t(n+\frac{1}{2})}}{\partial r^2} \simeq \frac{\bar{S}^{m+1}_{n+\frac{1}{2}} - 2\bar{S}^{m}_{n+\frac{1}{2}} + \bar{S}^{m-1}_{n+\frac{1}{2}}}{\Delta r^2} \tag{4.13}$$

$$\frac{\partial \bar{S}_{\Delta t(n+\frac{1}{2})}}{\partial r} \simeq \frac{\bar{S}^{m+1}_{n+\frac{1}{2}} - \bar{S}^{m-1}_{n+\frac{1}{2}}}{2\Delta r} \tag{4.14}$$

$$\tilde{\Lambda}\bar{S}_{\Delta t(n+\frac{1}{2})} \simeq \tilde{\Lambda}\bar{S}^m_{n+1} \tag{4.15}$$

$$\frac{\partial \bar{S}_{\Delta t(n+\frac{1}{2})}}{\partial t} \simeq \frac{\bar{S}^{m}_{n+1} - \bar{S}^{m}_{n+\frac{1}{2}}}{\frac{1}{2}\Delta t} \tag{4.16}$$

Using these approximations (4.8) leads to a system of equations

$$\left(I_{2M+1} + \tfrac{\Delta t}{2} D\right) S_{n+\frac{1}{2}} = S_{n+1}\left(I_{(K-1)^2} + \tfrac{\Delta t}{2}\tilde{\Lambda}'\right) \tag{4.17}$$

where D is a tridiagonal matrix and I_n denotes the $n-$dimensional identity matrix. The advantage of the ADI-method is now apparent: S is a $(2M+1) \times (K-1)^2$ matrix instead of a $(2M+1)(K-1)^2$ dimensional vector. Therefore, the equation system to be solved is much smaller.

For the next iteration we use the Markov chain as implicit and the interest rate as explicit. I.e. we change the approximations (4.15) and (4.16) to

$$\tilde{\Lambda}\bar{S}_{\Delta t\left(n+\frac{1}{2}\right)} \simeq \tilde{\Lambda}\bar{S}^m_{n} \tag{4.18}$$

$$\frac{\partial \bar{S}_{\Delta t\left(n+\frac{1}{2}\right)}}{\partial t} \simeq \frac{\bar{S}^{m}_{n+\frac{1}{2}} - \bar{S}^{m}_{n}}{\frac{1}{2}\Delta t} \tag{4.19}$$

Insert the approximations (4.13), (4.14), (4.18), and (4.19) into (4.8) which leads to a similar system of equations

$$\begin{aligned} \left(I_{(K-1)^2} - \tfrac{\Delta t}{2}\tilde{\Lambda}\right)(S_n)' &= \left(S_{n+\frac{1}{2}}\right)'\left(I_{2M+1} - \tfrac{\Delta t}{2}D\right)' \\ \Leftrightarrow \quad S_n\left(I_{(K-1)^2} - \tfrac{\Delta t}{2}\tilde{\Lambda}\right)' &= \left(I_{2M+1} - \tfrac{\Delta t}{2}D\right)S_{n+\frac{1}{2}} \end{aligned} \tag{4.20}$$

Solving (4.20) involves finding the inverse of $I_{(K-1)^2} - \frac{\Delta t}{2}\tilde{\Lambda}$ which causes a little problem since $\tilde{\Lambda}$ contains an indicator function. This can be done with the approximation

$$\begin{aligned} & S_n \left(I_{(K-1)^2} - \frac{\Delta t}{2}\tilde{\Lambda} \right)' \\ = \; & S_n \left(I_{(K-1)^2} - \frac{\Delta t}{2}\Lambda \right)' + S_n D_n^{\phi} \\ \simeq \; & S_n \left(I_{(K-1)^2} - \frac{\Delta t}{2}\Lambda \right)' + S_{n+\frac{1}{2}} D_{n+\frac{1}{2}}^{\phi} \end{aligned} \tag{4.21}$$

where D_n^{ϕ} is a diagonal $(K-1)^2 \times (K-1)^2$ matrix containing the indicator functions. Evaluated at time $n+\frac{1}{2}$ these indicator functions are known. This part can now be moved to the right hand side of (4.20). D^{ϕ} contains jumps so we have to make sure that this approximation is appropriate. $D_n^{\phi} \neq D_{n+\frac{1}{2}}^{\phi}$ only if the swap price changes sign in that time period. If the swap price changes sign over a very small period of time it must be close to zero, which justifies the approximation (4.21)[9].

To simplify notation we define

$$\begin{aligned} \mathtt{A}_1 &= \left(I + \frac{\Delta t}{2} D \right) \\ \mathtt{A}_2 &= \left(I - \frac{\Delta t}{2} \Lambda' \right) \\ \mathtt{B}_1 &= \left(I + \frac{\Delta t}{2} \tilde{\Lambda}' \right) \\ \mathtt{B}_2 &= \left(I - \frac{\Delta t}{2} D \right) \end{aligned}$$

Now it follows that the two systems needing to be solved are

$$\begin{aligned} \mathtt{A}_1 S^{n+\frac{1}{2}} &= S^{n+1} \mathtt{B}_1 \\ S^n \mathtt{A}_2 &= \mathtt{B}_2 S^{n+\frac{1}{2}} - S^{n+\frac{1}{2}} D^{\phi} \end{aligned}$$

In Appendix 4.10.3 we have included a description of the implementation and a discussion of the complications that arise when considering time-dependent generators of even stochastic (e.g. spot rate dependent) transition intensities.

[9] Another way of eliminating the indicator functions is to use the method described at the end of Section 4.2. However, this is a slower solution due to the higher dimension of the intensity matrix. Our simulations indicate that the convergence properties of the two approaches are approximately the same. Therefore, for swap pricing we recommend using the $(K-1)^2 \times (K-1)^2$ intensity matrix.

4.4 Pricing Swaps with Default Risk

We have selected a 5 year "plain vanilla" swap with semiannual exchanges of fixed rate payments for floating rate payments. More precisely, consider 11 dates, $T_0, T_1, T_2, \ldots, T_{10}$, where $T_i - T_{i-1} = \frac{1}{2}$ for every $i = 1, \ldots, 10$. The payment at time T_i is defined as

$$X_i = F\left(\frac{1}{P(T_i, T_i + \frac{1}{2})} - 1 - \frac{c}{2}\right)$$

for $i = 1, \ldots, 10$ where c is the fixed rate and F is the constant notional amount. Without loss of generality we will assume that $F = 1$. Note, however, that we only study a version without payment in arrear. Also, note that the floating payment is based on the yield of a six-month treasury bond.

For our interest rate model we have chosen the Vasicek model, but of course other models of the term structure could be used. Hence, under the risk neutral measure, the spot rate is described as

$$dr = \kappa(\mu - r)dt + \sigma dW_t$$

where W_t is a standard Brownian motion.

For our computations we have chosen typical values of parameters $\mu = 0.05$, $\kappa = 0.15$, and $\sigma = 0.015$. Furthermore, we have set $r_0 = 0.05$ and chosen a recovery rate for both counterparties equal to 0.4. The ratings of the two counterparties are assumed independent and follow Markov chains described by the constant generator matrix used in Jarrow, Lando, and Turnbull[33](Table 4 page 507)[10].

It is straightforward to calculate c such that the initial value of a default free swap is 0. Using closed form solutions for bond prices in the Vasicek model, we find that in a riskless swap the fixed side would have to pay a fixed rate of

$$c = 5,0125\%$$

to make the initial value of the swap zero. Swap spreads are computed with respect to this quantity.

We first consider the fair fixed rate to be paid if the initial ratings of the parties are the same and we consider what happens if this common rating varies. To be precise, the rating at time 0 is assumed to be the same for the two parties, but of course rating transitions can bring the two parties into

[10]It is straightforward to construct the generator matrix for the Markov chain of joint ratings, using the fact that the only non-zero entries off the diagonal are those corresponding to transition of one of the firms, see Appendix 4.10.2

different categories during the life of the contract. As can be seen from Figure 4.1 the spreads are very small (all smaller than 0.8 basis points) and do not vary a lot across ratings at an absolute level (but do show significant variation at a relative level). One may wonder why there is *any* spread in the case with

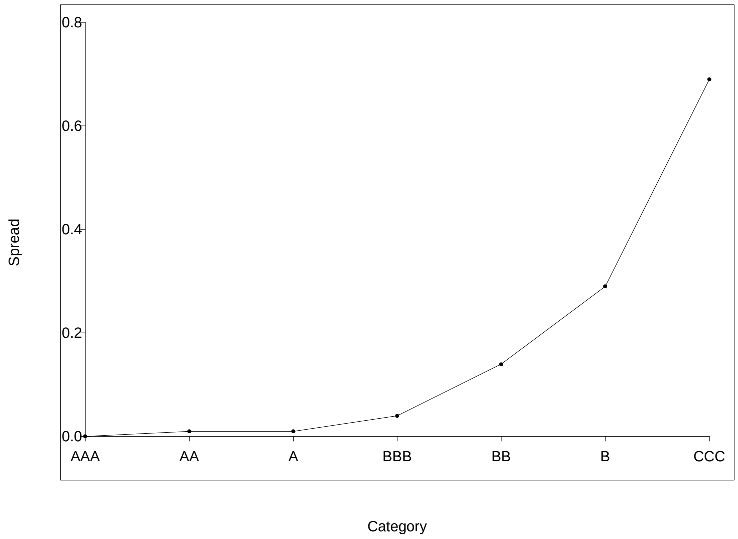


Figure 4.1: Swap spreads in basis points when the initial category is the same for both counterparties.

symmetric ratings. It turns out that the asymmetry in payments (one pays floating, the other fixed) does indeed produce a small spread. Hence the small spreads we see are not due to numerical error. To see this, consider the case of identical, constant default intensities for two counterparties who can be in only one non-default category, and with a constant fractional recovery of $\phi = 0.4$. Using the same Vasicek model for riskless interest rates as above, we may compute analytically the value of the swap spread.[11] The size of the spread in the symmetric case as a function of the default intensity λ for different values of the initial short rate is shown in Figure 4.2.

Spreads change significantly, however, when we allow for rating asymmetries. Hence we next fix the (initial) rating of party A at AA and let the (initial) rating of B vary. Both the case where A pays fixed and A pays floating are shown in Figure 4.3. It is worth noting that there is a high degree

[11]The analytic expression is easily obtained using equation (4.6) and the fact that with symmetric default intensities we have $R(t) = r_t + (1 - \phi)\lambda$, hence the indicator functions vanish. The resulting expression for the spread is not zero.

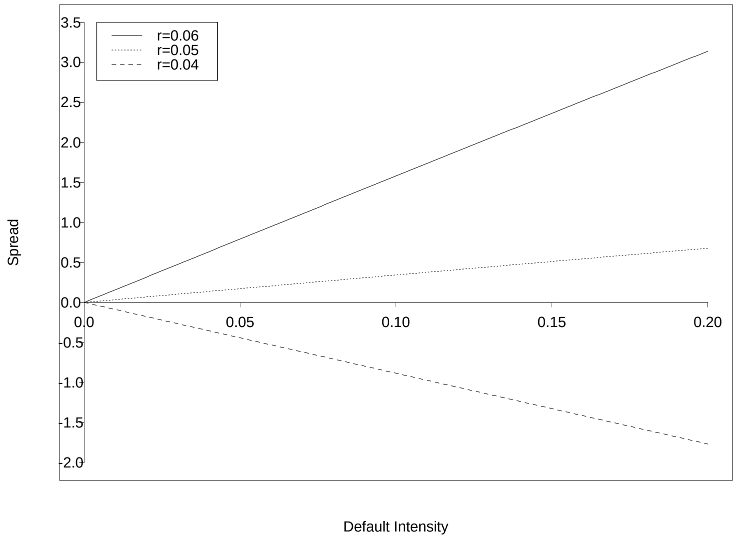


Figure 4.2: Swap spreads in the case of symmetric intensities as a function of the intensity. Recovery is fixed at 0.4 and the riskless rate follows a Vasicek model. Three different values of the initial short rate are used.

of symmetry in the spreads for the two cases. This symmetry is however sensitive to choice of initial spot rate level and spot rate model.

In Figure 4.4, we plot the spreads against spreads on defaultable bonds with same recovery and maturity as the swap and the graph seems to suggest an almost linear relationship between the two spreads. The same linear relationship is found in Duffie and Huang[21]. We find a bond yield spread of 100 basis points translates into a swap spread of approximately 1.7 basis points, which is very similar to Duffie and Huang[21], who find a translation to 1 basis point.

4.5 Time-dependent and stochastic generators

The numerical method we have outlined can easily be adapted to the case where the generator of the Markov transition matrix is non-homogeneous and to the case where default intensities depend on the driving state variables also. The interest in non-homogeneous matrices arises primarily because of calibration issues: As shown in Jarrow, Lando, and Turnbull[33] one can use

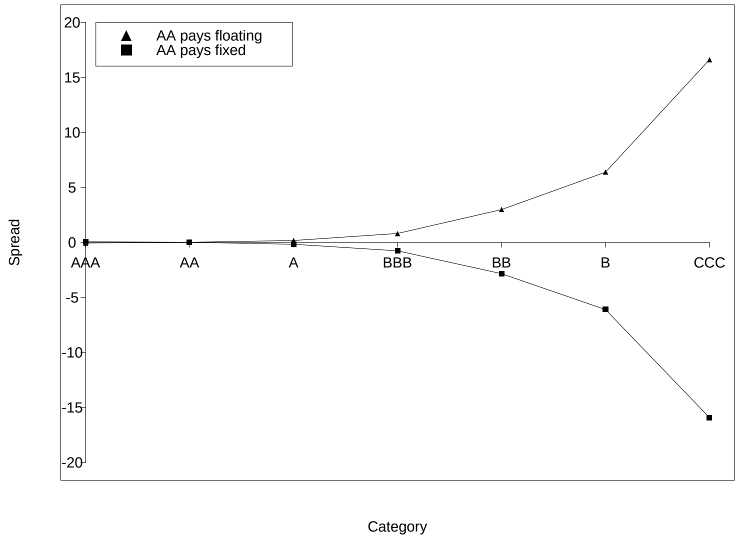


Figure 4.3: Swap spreads in basis points when the initial rating of one counterparty is fixed to AA and the rating of the other counterparty varies. Both the case where AA pays fixed and AA pays floating is considered. Notice the symmetry.

an empirical generator matrix and a time-dependent risk adjustment matrix to fit an initial structure of zero coupon bonds. The product of these two matrices gives a non-homogeneous generator. Another possibility designed to take into account changes in business cycles and correlation between rating migrations and interest rates is to let the generator matrix have elements which are functions of the interest rate (or, more generally, the state variables driving the interest rates). This approach is described in Lando[41],[43]. In this section we will present some illustrations of the latter approach applied to swap pricing. We consider two cases of stochastic intensities. The first case is the 'affine' case studied in Lando[43]: The generator of the Markov chain which includes the default category is now given for each party as

$$A(r_s) = B\mu(r_s)B^{-1}$$

where B is a constant matrix of eigenvectors for the generator used in Jarrow, Lando, and Turnbull[33], and where $\mu(r_s)$ is a diagonal matrix whose elements are

$$\begin{aligned} \mu_i(r_s) &= \gamma_i + \kappa_i(r_s - r_0), \qquad i = 1, \ldots, K-1 \\ \mu_K(r_s) &= 0. \end{aligned}$$

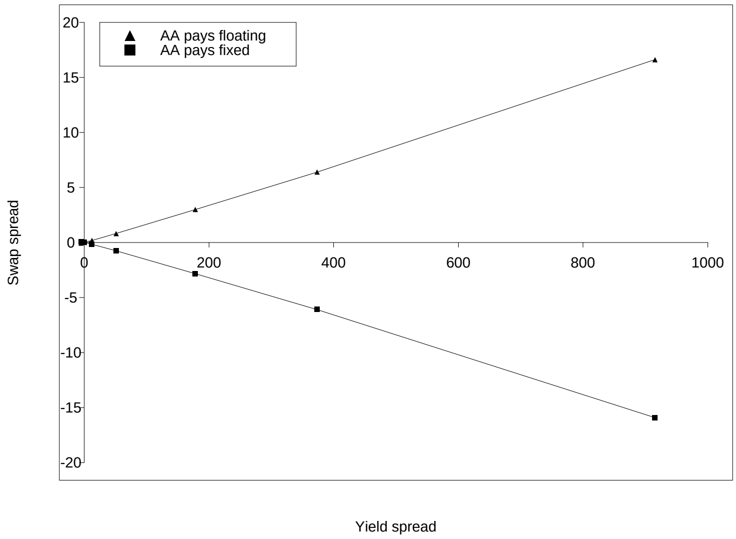


Figure 4.4: Swap spreads as a function of corporate bond yield spread for each category. Notice the almost linear relationship.

and where the coefficient vectors γ, κ are chosen to calibrate the level and the spot rate sensitivities of the initial corporate bond spreads to observed values[12]. Below, we will refer to this generator as the affine generator. In our example we have used the initial spreads (for zero recovery bonds)

$$\hat{s} = (16, 20, 27, 44, 89, 150, 255)'$$

and initial spread sensitivities

$$\hat{ds} = (-0.1, -0.15, -0.2, -0.25, -0.3, -0.5, -1.0)'$$

Now, the initial spreads and spot rate sensitivities are calibrated (see Lando[43]) by defining

$$\begin{aligned} \gamma &= \beta^{-1}\hat{s} \\ \kappa &= \beta^{-1}\hat{ds} \end{aligned}$$

where β is a $K-1 \times K-1$ matrix with entries

$$\beta_{ij} = B_{ij}B_{jK}^{-1} \quad \text{for } i,j = 1,\ldots,K-1.$$

[12]Since we work with intensities which are affine functions of the spot rate, there is a positive probability of having negative intensities in this framework. We solve this in our numerical implementation by setting the intensity equal to zero in such cases.

This gives us

$$\begin{aligned}\gamma &= (-0.0057, -0.0161, -0.0200, -0.0227, -0.0272, -0.0328, -0.0345)' \\ \kappa &= (0.2674, 0.5600, 0.6320, 0.7069, 0.9013, 1.3590, 1.4002)'\end{aligned}$$

We study the effect of interest rate sensitivity in rating transitions and default intensities in this setting by comparing the affine generator with the 'base case' obtained by setting $\kappa = 0$. This base case corresponds to having zero interest rate sensitivity in the affine generator. Figure 4.5 shows that swap spreads change considerably when the interest rate sensitivity is taken into account. For example, the spread on a swap with an AA-rated floating payer

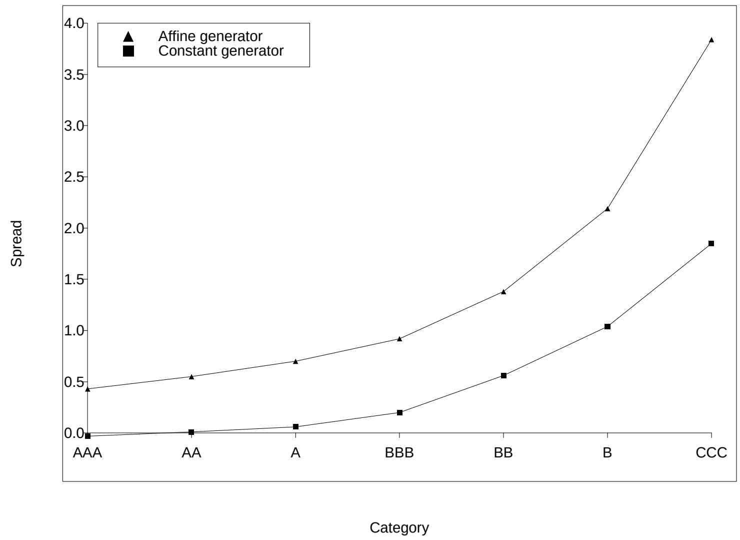


Figure 4.5: A comparison between swap spreads calculated with an affine generator and a constant generator. For both generators AA pays floating and fixed pay side varies. The constant generator is equal to the affine generator for $r_s = r_0$.

and a B-rated counterparty paying fixed, increases from approximately 1 to 2.2 basis points. The intuition is the following: When interest rates go up, cash flows become less uncertain because the default intensities are calibrated to fall in this case and the opposite is true when rates go down. This means that the floating payer sees uncertain positive cash flows and less uncertain negative cash flows and consequently wants compensation for this through a higher spread. Also, the symmetry that we had in the non-stochastic case between spreads when AA paid fixed and AA paid floating vanishes. This is shown in Figure 4.6.

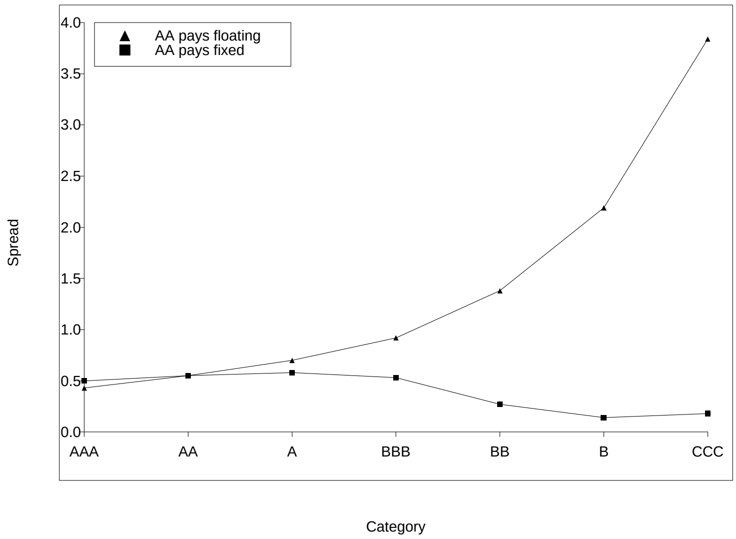


Figure 4.6: Swap spreads calculated with an affine generator. Notice that the spreads are no longer symmetric.

A problem with the affine generator specification is that it is hard to control the sign of all entries in the generator and it may take relatively small fluctuations in the riskless interest rate to produce negative entries in the off-diagonal elements of the generator[13].

Therefore, we consider a second case in which rating migrations are held constant but default intensities are spot rate dependent. The default intensity from category i is assumed to be of a 'logit' form

$$\lambda_i(r_t) = \frac{a_i}{1 + \exp(b_i + c_i(r_t - r_0))} \qquad i = 1, \dots, K-1$$

and negative sensitivity of intensities to changes in spot rates is obtained by having $c > 0$. This form is similar to the form proposed in Wilson[63] and also used in Das and Sundaram[14]. We choose the parameter vectors a, c to match initial spreads and spread sensitivities[14] and we impose the condition

[13]In our calibration, already at a spot rate level of 8% will some entries become negative.

[14]Note that with zero recovery the initial spread for category i is $\lambda_i(r_0)$ and the sensitivity is the derivative $\lambda_i'(r_0)$.

$b = 0$ which gives symmetric dependence of intensities on spot rates around the initial level with our choice of a, c :

$$\begin{aligned} a_i &= 2\lambda_i(r_0) \\ c_i &= -\frac{2\lambda_i'(r_0)}{\lambda_i(r_0)} \end{aligned}$$

In Figure 4.7 we compare the results to those obtained using an affine generator. In both cases the short spreads and the sensitivity of changes in the

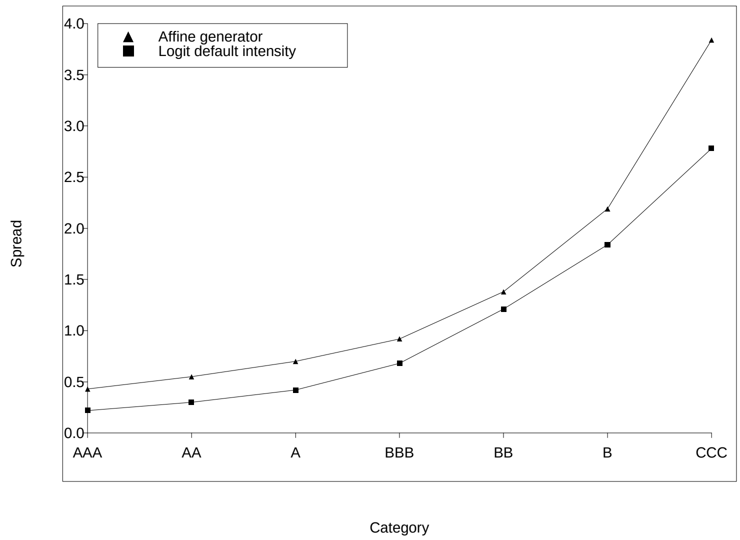


Figure 4.7: Swap spreads calculated with 'logit' default intensities compared with the results using an affine generator. Both models fit the initial spreads and spread sensitivities.

short spreads to changes in the spot rate have been matched but the logit generator avoids negative intensities. The difference in spreads is not huge, suggesting that for some cases the computations using the affine generator are not invalidated by the positive probability of negative intensities.

4.6 Credit triggers

As an example of how our methods can take into account provisions in contracts based on ratings, we will consider a credit trigger in the swap contract. The credit trigger is formulated as follows: If the triggering category is k then in the event that either of the counterparties drops to k or below, the contract terminates. The settlement payment in that case is defined to be equal to the value of a swap without a credit trigger if the downgraded party has a rating one level above the triggering category, k. In the event that a counterparty defaults we use the partial settlement described in Section 4.2 where it is partial to the swap *with* a credit trigger. These conventions are for illustrative purposes only and it is easy to fit in other market conventions.

In Table 4.1 we have chosen the triggering category to be BB, (which often is said to mark the transition to 'speculative grade'). We compare the

		AAA	AA	A	BBB
JLT Generator	no trigger	-0.06	0.01	0.18	0.81
	BB-trigger	-0.02	0.00	0.10	0.38
Affine Generator	no trigger	0.43	0.55	0.70	0.92
	BB-trigger	0.42	0.54	0.69	0.90
Logit Generator	no trigger	0.22	0.30	0.42	0.68
	BB-trigger	0.22	0.29	0.39	0.57

Table 4.1: A comparison of spreads on a swap with a BB credit trigger and regular swap spreads as studied above. The credit trigger has a significant effect in the case of a constant generator, little effect when using the affine and logit generator due to the calibration shifting probability mass from downgrade transitions to direct default transitions.

effect of a trigger in three different settings: The first uses the (empirical) constant generator used in Jarrow, Lando, and Turnbull[33]. The second uses a stochastic generator with the affine specification which has been calibrated to match short spreads and spread sensitivities, and the third specification uses a stochastic generator where only the default intensities are changed using a logit functional form for calibration. In the constant case we see that introducing a credit trigger reduces the spreads to somewhere between one third and two thirds. depending on ratings. For a B-trigger the spreads are reduced to between 20 and 40% of the original spreads, again depending on ratings. Similarly, the reduction for a BBB-trigger is between 60 and 80%.

In the second case (with the affine generator) a danger in using this calibration shows up: To compensate for the fact that the observed spreads are far from those generated by the empirical transition probabilities, there is a large risk adjustment, and this affects also the non-default transition probabilities. In this example, these take low values after calibration. The low dimension of the risk adjustment parameters only allows us to control for initial spreads and spread sensitivities but there is little control over non-default transition intensities. And in our example, the intensities of direct default become relatively much more important after calibration and this in turn renders the credit trigger much less efficient. This is true also in the final example with the logit generator.

4.7 Default Swaps

As a second illustration of our numerical technique we consider default swaps - i.e. contracts in which one party (the default protection buyer) pays a fixed, periodic amount to the other party (the default protection seller) until whichever comes first: Default of an underlying reference security or maturity of the contract. If the reference security defaults before the maturity of the default swap, the protection seller pays an amount to the protection buyer which compensates (in a sense which varies among contracts) for the loss in value of the reference security. A contract specification with physical delivery would for example allow the protection buyer to exchange the defaulted reference security for its principal. For more on default swaps, see Duffie[19].

We study the sensitivity of the price of a default swap to changes in the joint credit quality of the protection seller and the reference security. In particular, the effect of correlation can be studied more carefully. A natural intuition to have on default swaps is that the party buying default protection should worry about correlation in the credit quality between the underlying reference security and the default protection seller. We show that if the correlation is 'weak' in the sense that defaults do not occur simultaneously, then the effect of correlation is negligible.

For the illustration, we have selected a 5 year default swap with semiannual fixed rate payments where the reference security is a zero-coupon bond. Since our focus is on the correlation between the protection seller and the issuer of the reference security, we assume for simplicity that the protection buyer cannot default. It is possible to include default of the protection buyer, in which case we would need a three-dimensional stochastic process to represent the joint ratings movements. Again, consider for our illustration 11 dates, $T_0, T_1, \ldots, T_{10}$ where $T_i - T_{i-1} = \frac{1}{2}$ for every $i = 1, \ldots, 10$ and $T_0 = 0$. The payment at time T_i is defined as

$$(T_i - T_{i-1})c = \frac{c}{2}$$

In case the protection seller defaults the settlement value is defined in terms of the contract value just before default as in the case of a swap. The settlement in case the issuer of the zero-coupon bond defaults is as follows: Assume the default occurs at time $t \in (T_{i-1}, T_i)$ then the payment made to the protection buyer is

$$(1-\phi)P(t,T) - (t - T_{i-1})c$$

where ϕ is the recovery rate of the zero-coupon bond and where we have subtracted from the payment the premium for the period between default and next premium payment date. Other payments could have been considered, e.g. a default payment of $1-\phi$, or the fixed payments could be paid in advance instead. This would change the level of the fixed payments but not alter our conclusions.

Now, c is found as a fair fixed rate such that the contract value is 0 at time 0. In Table 4.2 we have used the JLT generator for two independent firms. As can be seen from the table the initial category of the protection

Ref \ Ps	AAA	AA	A	BBB	BB	B	CCC
AAA	2.66	2.65	2.65	2.62	2.54	2.43	2.12
AA	7.17	7.16	7.15	7.09	6.90	6.63	5.88
A	19.48	19.47	19.43	19.32	18.94	18.38	16.87
BBB	61.05	61.02	60.94	60.67	59.76	58.41	54.73
BB	202.35	202.30	202.17	201.69	200.09	197.68	190.93
B	425.73	425.69	425.58	425.21	423.93	422.00	416.49
CCC	1131.75	1131.72	1131.63	1131.33	1130.31	1128.77	1124.27

Table 4.2: Default swap premiums using the JLT generator for different combinations of credit quality of protection seller and reference security. Protection buyer is assumed default-free. Swap premiums are paid in arrear.

seller is not that important for the fair fixed rate. This is due to the fact that in a setup with no simultaneous defaults, there is only a loss to the protection buyer in the event that the protection seller defaults before the reference security, but when that happens the loss is a fractional loss on a default swap where the reference security has not yet defaulted.

Next, we have modeled correlation through the transition intensities' joint dependence on the interest rate. In Table 4.3 we have used an affine generator to model correlation and Table 4.4 shows the results when using logit default intensities. Again, the credit quality of the protection seller is a 'second order' effect compared to the credit quality of the reference security. It is clear from this, that to get significant effects of correlation we need to have 'strong correlation' in the sense of simultaneous defaults of protection seller and reference security. Finally, the effect of having interest rate sensitive

Ref \ Ps	AAA	AA	A	BBB	BB	B	CCC
AAA	8.90	8.90	8.90	8.90	8.89	8.89	8.89
AA	11.13	11.13	11.13	11.13	11.13	11.13	11.12
A	15.00	15.00	15.00	15.00	15.00	14.99	14.99
BBB	24.18	24.18	24.18	24.18	24.18	24.17	24.16
BB	47.96	47.95	47.95	47.95	47.95	47.94	47.93
B	79.75	79.75	79.75	79.75	79.75	79.74	79.73
CCC	134.59	134.59	134.59	134.58	134.58	134.57	134.56

Ref \ Ps	AAA	AA	A	BBB	BB	B	CCC
AAA	9.66	9.65	9.65	9.63	9.61	9.56	9.47
AA	12.07	12.07	12.05	12.04	12.00	11.94	11.81
A	16.11	16.09	16.08	16.06	16.01	15.93	15.76
BBB	25.46	25.44	25.43	25.40	25.34	25.24	25.04
BB	49.71	49.69	49.67	49.63	49.57	49.45	49.20
B	81.85	81.82	81.78	81.73	81.63	81.45	81.08
CCC	136.69	136.64	136.57	136.48	136.28	135.94	135.24

Table 4.3: Default swap premiums using the affine generator for different combinations of credit quality of protection seller and reference security. The upper table considers a 'base case' in which the generator is calibrated in order to match short spreads but has zero sensitivity to changes in the short rate. The lower table is also calibrated to match short spreads but has spot rate sensitive intensities as well. Protection buyer is assumed default-free. Swap premiums are paid in arrear.

default intensities is small: the relative change in default premiums never exceeds 10% in our example. The relative effects in the plain vanilla swaps could be much larger.

4.8 Supplementary remarks

4.8.1 Convergence of ADI-method

In this section we will prove that the ADI-method used in this paper converges. To see this substitute $S^{n+\frac{1}{2}}$ in (4.20) by the solution found in (4.17) leading to

$$S^n = \left(I - \frac{\Delta t}{2} D\right)\left(I + \frac{\Delta t}{2} D\right)^{-1} S^{n+1}\left(I + \frac{\Delta t}{2}\Lambda'\right)\left(I - \frac{\Delta t}{2}\Lambda'\right)^{-1} \quad (4.22)$$

Ref \ Ps	AAA	AA	A	BBB	BB	B	CCC
AAA	9.80	9.80	9.80	9.80	9.79	9.79	9.78
AA	12.71	12.71	12.71	12.71	12.70	12.69	12.68
A	17.55	17.54	17.54	17.53	17.52	17.51	17.49
BBB	28.44	28.44	28.43	28.42	28.40	28.37	28.34
BB	52.52	52.52	52.51	52.50	52.46	52.43	52.38
B	79.49	79.48	79.48	79.47	79.45	79.43	79.41
CCC	115.78	115.78	115.77	115.77	115.75	115.74	115.71

Ref \ Ps	AAA	AA	A	BBB	BB	B	CCC
AAA	10.14	10.14	10.14	10.13	10.11	10.09	10.06
AA	13.16	13.15	13.15	13.14	13.11	13.08	13.04
A	18.13	18.13	18.12	18.10	18.07	18.03	17.96
BBB	29.28	29.27	29.26	29.23	29.18	29.11	29.02
BB	53.76	53.75	53.73	53.69	53.61	53.51	53.36
B	81.18	81.16	81.14	81.10	81.00	80.88	80.71
CCC	117.96	117.94	117.91	117.85	117.72	117.56	117.32

Table 4.4: Default swap premiums using the 'logit' generator for different combinations of credit quality of protection seller and reference security. The upper table considers a 'base case' in which the generator is calibrated in order to match short spreads but has zero sensitivity to changes in the short rate. The lower table is also calibrated to match short spreads but has spot rate sensitive intensities as well. Protection buyer is assumed default-free. Swap premiums are paid in arrear.

Denote the eigenvalues of D by ν_i and the eigenvalues of Λ by λ_i then since D is positive definite and Λ is negative semidefinite the eigenvalues of $\left(I - \frac{\Delta t}{2} D\right)\left(I + \frac{\Delta t}{2} D\right)^{-1}$ and $\left(I + \frac{\Delta t}{2}\Lambda'\right)\left(I - \frac{\Delta t}{2}\Lambda'\right)^{-1}$ are

$$\begin{aligned} \left| \frac{1 - \frac{\Delta t}{2}\nu_i}{1 + \frac{\Delta t}{2}\nu_i} \right| &< 1 \\ \left| \frac{1 + \frac{\Delta t}{2}\lambda_i}{1 - \frac{\Delta t}{2}\lambda_i} \right| &\leq 1 \end{aligned}$$

Assume that we know the right price at time N then we can find the price at time $N-1$ as

$$S((N-1)\Delta t, r, \eta) = \Theta S(T, r, \eta)\Gamma + o(\Delta r^2) + o(\Delta t)$$

where $\Theta \equiv \left(I - \frac{\Delta t}{2} D\right)\left(I + \frac{\Delta t}{2} D\right)^{-1}$ and $\Gamma \equiv \left(I + \frac{\Delta t}{2}\Lambda'\right)\left(I - \frac{\Delta t}{2}\Lambda'\right)^{-1}$. I.e.

the error at each time point is given as a $M \times K$ matrix and we can successively calculate them as

$$\begin{aligned}
\varepsilon_{N-1} \equiv S((N-1)\Delta t, r, \eta) - S^{N-1} &= o(\Delta r^2) + o(\Delta t) \\
\varepsilon_{N-2} \equiv S((N-2)\Delta t, r, \eta) - S^{N-2} &= \Theta \varepsilon_{N-1} \Gamma + o(\Delta r^2) + o(\Delta t)
\end{aligned}$$

In general the error is given as

$$\varepsilon_{N-i} = \sum_{j=0}^{i-1} \Theta^j (o(\Delta r^2) + o(\Delta t)) \Gamma^j$$

and we would like $\varepsilon \to 0$. Remember, that the eigenvalues of Θ and Γ are ≤ 1, so that $\Theta^n \to 0$ and Γ^n does not explode. Define the spectral radius of a matrix A by $\rho(A)$ then

$$\begin{aligned}
\|\varepsilon_0\|_\infty &= \|\sum_{j=0}^{N-1} \Theta^j (o(\Delta r^2) + o(\Delta t)) \Gamma^j \|_\infty \\
&\leq \sum_{j=0}^{N-1} \|\Theta^j\|_\infty \|o(\Delta t, \Delta r)\|_\infty \|\Gamma^j\|_\infty \\
&\leq \sum_{j=0}^{N-1} \rho(\Theta)^j \rho(\Gamma)^j o(\Delta t, \Delta r) \\
&= \sum_{j=0}^{N-1} \rho(\Theta)^j o(\Delta t, \Delta r) \\
&= \frac{1 - \rho(\Theta)^N}{1 - \rho(\Theta)} o(\Delta t, \Delta r)
\end{aligned}$$

which is what we wanted.

4.8.2 Calculation of fixed rate

We have selected a 5 year "plain vanilla" swap with semiannual exchanges of fixed rate payments for floating rate payments. More precisely, consider 10 dates, $T_1, T_2, \ldots, T_{10}$, where $T_{i+1} - T_i = \frac{1}{2}$ for every $i = 1, \ldots, 9$. The payment at time T_i is defined as

$$X_i = F \left(\frac{1}{P(T_i, T_i + \frac{1}{2})} - 1 - \frac{c}{2} \right)$$

for $i = 1, \ldots, 10$ where F is the constant notional amount. Without any loss of generality we will assume that $F = 1$. Notice, that we in this case study a version without payment in arrear.

For our interest rate model we have chosen the Vasicek model

$$dr = \kappa(\mu - r)dt + \sigma dW_t$$

where W_t is a standard Brownian motion. Then we find $P(t,T)$ as

$$P(t,T) = \alpha(T-t)e^{-\beta(T-t)r_t}$$

where

$$\begin{aligned}\beta(x) &= \frac{1-e^{-\kappa x}}{\kappa} \\ \alpha(x) &= \exp\left(\frac{(\beta(x)-x)(\kappa^2\mu - \frac{1}{2}\sigma^2)}{\kappa^2} - \frac{\sigma^2\beta(x)^2}{4\kappa}\right)\end{aligned}$$

The value of a riskless swap of this type would be calculated as

$$\begin{aligned} & S(t) \qquad (4.23) \\ = & \sum_{i=1}^{10} E_t^Q\left[e^{-\int_t^{T_i} r(s)ds}\left(\frac{1}{P(T_i, T_i+\frac{1}{2})} - 1 - \frac{c}{2}\right)\right] \\ = & \sum_{i=1}^{10} E_t^Q\left[e^{-\int_t^{T_i} r(s)ds}\left(\frac{1}{\alpha(\frac{1}{2})e^{-\beta(\frac{1}{2})r(T_i)}} - 1 - \frac{c}{2}\right)\right] \\ = & \sum_{i=1}^{10} \frac{1}{\alpha\left(\frac{1}{2}\right)} E_t^Q\left[e^{-\int_t^{T_i} r(s)ds} e^{\beta(\frac{1}{2})r(T_i)}\right] - (1+\frac{c}{2})P(t,T_i) \\ = & \sum_{i=1}^{10} \frac{1}{\alpha\left(\frac{1}{2}\right)} P(t,T_i) E_t^{T_i}\left[e^{\beta(\frac{1}{2})r(T_i)}\right] - (1+\frac{c}{2})P(t,T_i) \\ = & \sum_{i=1}^{10}\left(\frac{1}{\alpha\left(\frac{1}{2}\right)} e^{\beta(\frac{1}{2})f(t,T_i)} e^{\frac{\sigma^2}{4}\beta(\frac{1}{2})^2\beta(2(T_i-t))} - (1+\frac{c}{2})\right) P(t,T_i) \qquad (4.24)\end{aligned}$$

where E^{T_i} denotes the T_i forward measure and $f(t,T_i)$ is the T_i forward rate. The last equality is due to the fact that under the conditional T_i forward measure $r(T_i) \sim \mathcal{N}\Big(f(t,T_i), \frac{\sigma^2}{2}\beta\big(2(T_i-t)\big)\Big)$ Björk[5]. In a Vasicek model we can find the forward rate as

$$\begin{aligned} f(t,T) &= -\frac{\partial \log P(t,T)}{\partial T} \\ &= \mu - \frac{\sigma^2}{2\kappa^2} - \frac{\sigma^2}{2\kappa^2}e^{-2\kappa(T-t)} + \left(r(t) + \frac{\sigma^2}{\kappa^2} - \mu\right)e^{-\kappa(T-t)} \qquad (4.25)\end{aligned}$$

For our simulations we have chosen to use $\mu = 0.05$, $\kappa = 0.15$, and $\sigma = 0.015$. With these values we can find c such that the value of a risk free swap is 0. Using (4.24) and (4.25) we find

$$c = 5.0125\%$$

4.8.3 Default settlements

The pricing of a swap is basically as easy as it is to price the settlement payment. This is because the payments in a defaultable swap can be decomposed into the promised payments which are paid until a default occur and the settlement payment at the time of default. The first part which we will call a swap-stop contract, since all the remaining payments are stopped after a default, is easily priced as

$$S^{STOP}(t, r_t) = \sum_{i=j}^{n} E_t \left[e^{-\int_t^{t_i} r_s + \lambda_s^A + \lambda_s^B \, ds} X_i \right] \quad \text{for } t \in [t_{j-1}, t_j)$$

where λ^i is the default intensity for counterparty $i = A, B$ and X_i is the promised payment at time t_i. One example of a payment is $X_i = \frac{1}{P(t_{i-1}, t_i)} - 1 - c$ where $P(t_{i-1}, t_i)$ is the price of a default free zero-coupon bond between t_{i-1} and t_i.

The payment at the time of the settlement can be priced as

$$S^{SET}(t, r_t) = E_t \left[\int_t^{t_n} e^{-\int_t^s r_u + \lambda_u^A + \lambda_u^B \, du} \left(\lambda_s^A S^A(s) + \lambda_s^B S^B(s) \right) ds \right]$$

where $S^i(t)$ is the settlement in case counterparty $i = A, B$ defaults first at time t. Both S^{STOP} and S^{SET} are first-to-default contracts of two firms which can be priced using the method described in Section 3.3.

Now, any type of defaultable swap can be priced as

$$S^D(t, r_t) = S^{STOP}(t, r_t) + S^{SET}(t, r_t)$$

Notice that S^{STOP} can be priced analytically in an affine model. However, depending on S^A and S^B we might have more difficulties with S^{SET}. E.g. we can have a partial settlement as in Duffie and Huang[21] where

$$S^A(t) = S^D(t-, r_t) \left(1_{\{S^D(t-, r_t) > 0\}} + \phi_t^A 1_{\{S^D(t-, r_t) < 0\}} \right) \tag{4.26}$$

where ϕ^A is the recovery rate for counterparty A.

A different type of settlement proposed by Hübner[32] is

$$S^A(t) = S_+(t, r_t) - \phi_t^A S_-(t, r_t)$$

where

$$S_+(t, r_t) = \sum_{i=j}^{n} E\left[e^{\int_t^{t_i} r_s\, ds} X_i^+\right] \quad \text{for } t \in [t_{j-1}, t_j)$$
$$S_-(t, r_t) = \sum_{i=j}^{n} E\left[e^{\int_t^{t_i} r_s\, ds} X_i^-\right] \quad \text{for } t \in [t_{j-1}, t_j)$$

With this type of settlement he finds a closed form solution. One problem with this settlement is that we do not net the payment before settling. Consider a situation where counterparty A defaults with a recovery rate of $\phi^A = 0$. Furthermore, assume that before default occurs the contract has negative value to counterparty A. Now, after the default B ends up paying A in settlement for a contract which "promised" B a payment from A. This is because the payments are not netted before settling. This could be the case for two separate contracts but not within a single contract.

Other settlements could be used. E.g. let

$$S^A(t) = X_i^+ - \phi_t^A X_i^- \quad \text{for every } t \in (t_{i-1}, t_i]$$

and similarly

$$S^B(t) = \phi_t^B X_i^+ - X_i^- \quad \text{for every } t \in (t_{i-1}, t_i]$$

which is basically exponential functions we will be able to price S^{SET} (almost) analytically in an affine setup using the Fourier inversion technique described in Duffie, Pan, and Singleton[23]. This settlement is very similar to the one used in Hübner[32]. Here, only the next payment is settled.

In Jarrow, Lando, and Turnbull[33] they propose a settlement of defaultable zero-coupon bonds which is a proportion of a non defaultable bond. We could adopt this type of settlement and replace S^D in (4.26) with a non defaultable version of the contract, S

$$S^A(t) = S(t, r_t)\left(1_{\{S(t,r_t)>0\}} + \phi_t^A 1_{\{S(t,r_t)<0\}}\right)$$

Now, we no longer have a recursion. All the terms in S^{SET} are known and the pricing becomes much more simple. Since the spreads for defaultable swaps are small the results using this settlement will probably be close to the results using (4.26).

4.9 Conclusion

We have presented a method for computing swap spreads in models of default based on ratings. The results confirmed findings in Duffie and Huang[21] that swap spreads are relatively insensitive to credit quality for interest rate swaps. Of course, for a book with thousands of swap contracts the small spreads will add up and there is still a need to price and control the credit risk of the entire portfolio. The framework presented may also be applied to foreign currency swaps. This would, due to the exchange of principal, produce larger spreads but the computational issues involved would be similar.

Our computations have also shown that using stochastic generators has a large impact on the results. We used two specifications: The affine generator permits analytical expressions for corporate bonds but also allows negative intensities to occur, whereas the logit specification only allows positive intensities but also to our knowledge requires numerical solution of bond prices. Both specifications changed our results for swap prices but mostly in the case with no credit triggers.

We also showed that in our setting the effect of the quality of the protection seller in a default swap was relatively small. This has implications for correlation as well: If changing the rating of the protection seller does not drastically alter the contract value, then, if we hold the marginal distribution of default fixed, but change the correlation between protection seller and reference security this will not have a huge effect. This result depends critically on the fact that we did not allow simultaneous default as a possibility in our example. We applied a 'weak' type correlation in which the correlation between defaults is obtained through the default intensities' joint dependence on the state variables. A 'strong' type of correlation in which the defaults are correlated by allowing for simultaneous transitions in the rating process will be considered in future work. Allowing for simultaneous transitions may capture not only default 'contagion' effects but also the widening of credit spreads on non-defaulted bonds following the default of a particular bond - an issue of great concern to risk managers.

4.10 Appendix

4.10.1 Proofs.

Proof of Proposition 7
Consider a time point $t < T$, the date of the contract's final promised cash flow $d_T \equiv d(T, r_T, \eta_T)$. Then

$$V_t = E_t^Q \left[\exp\left(-\int_t^T R_u du \right) d_T \right]$$

In the case where the price processes are driven by Markovian state variables we can compute this expectation by solving an appropriate system of PDEs. Let

$$\begin{aligned} f(t, r_t, \eta_t) &= V_t, \quad t < T & (4.27) \\ f(T, r_T, \eta_T) &= d_T & (4.28) \end{aligned}$$

Define

$$\mathcal{D}f = \frac{\partial}{\partial t} f + \mu \frac{\partial}{\partial r} f + \frac{1}{2}\sigma^2 \frac{\partial^2}{\partial r^2} f \tag{4.29}$$

By Ito's lemma for processes with jumps (see for example Protter[52]) we have

$$\begin{aligned} & f(t, r_t, \eta_t) - f(0, r_0, \eta_0) \\ = & \int_0^t \mathcal{D}f(s, r_s, \eta_s) ds + \sum_{0 < s \leq t} \Big(f(s, r_s, \eta_s) - f(s, r_s, \eta_{s-}) \Big) + M_t \end{aligned}$$

where M_t is a local martingale. The sum can be evaluated as[15]

$$\begin{aligned} & \sum_{0 < s \leq t} f(s, r_s, \eta_s) - f(s, r_s, \eta_{s-}) \\ = & \int_0^t \lambda_{\eta_s} \sum_{l \in \mathbb{L} \setminus \{\eta_s\}} \left(f(s, r_s, l) - f(s, r_s, \eta_s) \right) P(\eta_s, l) ds + M_t^* \end{aligned}$$

[15] We here use the representation of a pure jump process as the sum of a (local) martingale and the compensator of the jump process: For illustration, a counting process N with intensity λ, can be written as

$$N_t = N_t - \int_0^t \lambda_s \, ds + \int_0^t \lambda_s \, ds = M_t + \int_0^t \lambda_s \, ds$$

where M is a local martingale. Also, we use the common notation η_{t-} for the left limit of η at time t.

where $P(\eta_s, l)$ is the conditional probability that $\eta_s = l$ given that the value of η_{s-} and that the Markov chain jumps at time s, λ_{η_s} is the intensity by which the chain leaves the state η_s, and M_t^* is a local martingale. Continuing on, we find

$$\begin{aligned}
&= \int_0^t \left(\lambda_{\eta_s} \sum_{l \in \mathbb{L} \setminus \{\eta_s\}} \frac{\lambda_{\eta_s,l}}{\lambda_{\eta_s}} \Big(f(s, r_s, l) - f(s, r_s, \eta_s) \Big) \right) ds + M_t^* \\
&= \int_0^t \left(\sum_{l \in \mathbb{L} \setminus \{\eta_s\}} \lambda_{\eta_s,l} \Big(f(s, r_s, l) - f(s, r_s, \eta_s) \Big) \right) ds + M_t^* \\
&= \int_0^t \Big(\Lambda \bar{f}(s, r_s) \Big)_{\eta_s} ds + M_t^*.
\end{aligned}$$

Here, $\Big(\Lambda \bar{f}(s, r_s)\Big)_{\eta_s}$ denotes the $\eta_s'th$ element of the vector $\Lambda \bar{f}(s, r_s)$. Also, in the last equality we use the fact that a diagonal element of Λ is minus the sum of the off-diagonal elements in the same row. We can now write the differential of f as

$$df(t, r_t, \eta_t) = \Big(\mathcal{D} f(t, r_t, \eta_t) + \big(\Lambda \bar{f}(t, r_t) \big)_{\eta_t} \Big) dt + d\tilde{M}_t \tag{4.30}$$

where $\tilde{M}_t$ is a local martingale.

Next, consider the discounted pre-default process

$$Y_s = f(s, r_s, \eta_s) \beta_{t,s} \qquad s > t \tag{4.31}$$

where β is defined as

$$\beta_{t,s} = e^{-\int_t^s R_u du}$$

Now, using integration by parts

$$Y_T = f(t, r_t, \eta_t) + \int_t^T f(s, r_s, \eta_s) d\beta_{t,s} + \int_t^T \beta_{t,s} df(s, r_s, \eta_s) \tag{4.32}$$

Inserting (4.30) in (4.32) gives

$$Y_T - f(t, r_t, \eta_t) \tag{4.33}$$

$$= \int_t^T f(s, r_s, \eta_s) d\beta_{t,s} + \int_t^T \beta_{t,s} \Big(\mathcal{D} f(s, r_s, \eta_s) + \big(\Lambda \bar{f}(s, r_s) \big)_{\eta_s} \Big) ds + \hat{M}_t$$

$$= \int_t^T \beta_{t,s} \Big(\mathcal{D} f(s, r_s, \eta_s) + \big(\Lambda \bar{f}(s, r_s) \big)_{\eta_s} - R_s f(s, r_s, \eta_s) \Big) ds + \hat{M}_t \tag{4.34}$$

where $\hat{M}_t$ by assumption is a martingale. We also have from the definition of f and the expression for V that

$$f(t, r_t, \eta_t) = E_t^Q \left(\beta_{t,T} d_T\right) \quad \text{for } t < T \tag{4.35}$$

By taking conditional expectation on both sides of (4.34) we have a second expression for $f(t, r_t, \eta_t)$, and for the two expressions to be equal we must have for all $t < T$ that

$$E_t^Q \left[- \int_t^T \beta_{t,s} \Big(\mathcal{D} f(s, r_s, \eta_s) + \left(\Lambda \bar{f}(s, r_s)\right)_{\eta_s} - R_s f(s, r_s, \eta_s) \Big) \, ds \right] = 0 \tag{4.36}$$

Hence $f(s, r_s, \eta_s) = V_s$ is a solution to the partial differential equation

$$\mathcal{D} f(s, r_s, \eta_s) + \left(\Lambda \bar{f}(s, r_s)\right)_{\eta_s} - R_s f(s, r_s, \eta_s) = 0. \tag{4.37}$$

and from the lumpy dividend paid at time T we have the boundary condition

$$f(T, r, \eta) = d_T.$$

Using a little algebra and writing (4.37) with the same vector notation as in (4.7) gives the result.

■

4.10.2 A generator example

In this appendix, we illustrate the construction of joint generators where independence between the two chains is maintained and we illustrate the distinction between generators with and without the default category. First, let the marginal generator matrices without the default category be given as

$$\Lambda^A = \Lambda^B = \begin{bmatrix} -\alpha & \alpha \\ \beta & -\beta \end{bmatrix}$$

and the default intensities as a 2-dimensional vector

$$\begin{bmatrix} \lambda_1 \\ \lambda_2 \end{bmatrix}$$

Then the generator matrix including the default category, Λ^D is defined as:

$$\Lambda^D = \begin{bmatrix} -(\alpha + \lambda_1) & \alpha & \lambda_1 \\ \beta & -(\beta + \lambda_2) & \lambda_2 \\ 0 & 0 & 0 \end{bmatrix}$$

where the third category (default) is absorbing.

The combined generator matrix without the default category will be a 4×4 matrix with state space $\{(1,1),(1,2),(2,1),(2,2)\}$. Using this sequence of the states and assuming that counterparties A and B are independent (in that case there are no simultaneous jumps) we define the combined generator by

$$\Lambda = \begin{bmatrix} -2\alpha & \alpha & \alpha & 0 \\ \beta & -(\alpha+\beta) & 0 & \alpha \\ \beta & 0 & -(\alpha+\beta) & \alpha \\ 0 & \beta & \beta & -2\beta \end{bmatrix}$$

I.e. Λ_{12} is the intensity that counterparty B jumps from category 1 to 2 and A remains in category 1. Similarly, Λ_{13} is the intensity that counterparty A jumps from category 1 to 2 and B stays in category 1. The 4-dimensional vectors of default intensities are defined by

$$\lambda^A = \begin{bmatrix} \lambda_1 \\ \lambda_1 \\ \lambda_2 \\ \lambda_2 \end{bmatrix} \qquad \lambda^B = \begin{bmatrix} \lambda_1 \\ \lambda_2 \\ \lambda_1 \\ \lambda_2 \end{bmatrix}$$

4.10.3 Implementation

A program solving for the swap price would look something like

```
Initialization
find A_1^{-1}
find A_2^{-1}
for l = 1 to # of payments
     swap = swap + Payment(l)
     for i = 1 to n
          temp = swap*B_1
          swap = A_1^{-1}*temp
          temp = B_2*swap-swap*D^φ
          swap = temp*A_2^{-1}
```

where `n` is the number of time steps between payments and `temp` is the right hand side of the equations, updated for each iteration. The initialization involves reading the matrix Λ from a file and initializing the matrix D.

In the case of a time dependent intensity matrix we need to do the inversion of $(I - \frac{\Delta t}{2}\Lambda')$ in each time step. This means that the program will be slower. Our simulations suggest that the time needed for calculating swap prices with a time dependent intensity matrix is approximately 10% higher than for the constant intensity matrix. A sketch of the program is outlined below.

```
Initialization
find A_1^{-1}
for l = 1 to # of payments
     swap = swap + Payment(l)
     for i = 1 to n
          temp = swap*B_1
          swap = A_1^{-1}*temp
          find A_2^{-1}
          temp = B_2*swap-swap*D^φ
          swap = temp*A_2^{-1}
```

Another complication is an intensity matrix with dependence on the interest rate. This complicates the program, since each row of `swap` is the swap price for different values of the interest rate. This means that we need to split up the matrix `swap` and we will call the `r`'th row for `swap(r)` where `r` is between `rmin` and `rmax`. This idea is outlined below.

```
Initialization
find A_1^{-1}
for r = rmin to rmax by Δr
     find A_2(r)^{-1}
for l = 1 to # of payments
     swap = swap + Payment(l)
     for i = 1 to n
          for r = rmin to rmax by Δr
               temp(r) = swap(r)*B_1(r)
```

```
        swap(r) = A_1^{-1}*temp(r)
        temp(r) = B_2*swap(r)-swap(r)*D^φ(r)
        swap(r) = temp(r)*A_2(r)^{-1}
    collect swap(r) in the matrix swap
```

This program is approximately 50% slower than the original program.

Bibliography

[1] P. Abken. Valuation of default-risky interest-rate swaps. *Advances in Futures and Options Research*, 6:93–116, 1993.

[2] P. Artzner and F. Delbaen. Finem lauda or the credit risk in swaps. *Insurance, Mathematics and Economics*, 9:295–303, 1990.

[3] P. Artzner and F. Delbaen. Default risk insurance and incomplete markets. *Mathematical Finance*, 5:187–195, 1995.

[4] A. Arvanitis, J. Gregory, and J.-P. Laurent. Building models for credit spreads. *Journal of Derivatives*, pages 27–43, Spring 1999.

[5] T. Björk. *Arbitrage Theory in Continuous Time.* Oxford University Press, 1998.

[6] F. Black and J. Cox. Valuing corporate securities: Some effects of bond indenture provisions. *Journal of Finance*, pages 351–367, May 1976.

[7] F. Black and M. Scholes. The pricing of options and corporate liabilities. *Journal of Political Economy*, 3:637–654, 1973.

[8] P. Brémaud. *Point Processes and Queues - Martingale Dynamics.* Springer, New York, 1981.

[9] M. Brennan and E. Schwartz. Corporate income taxes, valuation, and the problem of optimal capital structure. *Journal of Business*, 51:103–114, 1978.

[10] K. Brown, W. Harlow, and D. Smith. An empirical analysis of interest rate swap spreads. *The Journal of Fixed Income*, pages 61–78, March 1994.

[11] A. Cifuentes and G. O'Connor. The binomial expansion method applied to cbo/clo analysis. Special report, Moody's Investors Service, 1996.

[12] I. Cooper and A. Mello. The default risk of swaps. *Journal of Finance*, 46:597–620, 1991.

[13] J. Cox, S. Ross, and M. Rubinstein. Option pricing : A simplified approach. *Journal of Financial Economics*, 7:229–263, 1979.

[14] S. Das and R. Sundaram. A direct approach to arbitrage-free pricing of credit derivatives. Forthcoming in: Management Science, 1998.

[15] S. Das and P. Tufano. Pricing credit-sensitive debt when interest rates, credit ratings and credit spreads are stochastic. *Journal of Financial Engineering*, pages 161–198, 1996.

[16] M. Davis and V. Lo. Infectious defaults. Working Paper. Tokyo-Mitsubishi International plc, 1999.

[17] D. Duffie. *Dynamic Asset Pricing Theory*. Princeton University Press, Princeton, New Jersey, 2nd edition, 1996.

[18] D. Duffie. First-to-default valuation. Working Paper. Stanford University, 1998.

[19] D. Duffie. Credit swap valuation. *Financial Analysts Journal*, pages 73–87, January/February 1999.

[20] D. Duffie and N. Gârleanu. Risk and valuation of collateralized debt obligations. Working Paper. Stanford University, 1999.

[21] D. Duffie and M. Huang. Swap rates and credit quality. *Journal of Finance*, 51(3):921–949, July 1996.

[22] D. Duffie and D. Lando. Term structures of credit spreads with incomplete accounting information. Working Paper, University of Copenhagen and Stanford University, 1999.

[23] D. Duffie, J. Pan, and K. Singleton. Transform analysis and asset pricing for affine jump-diffusions. *Econometrica*, 68(6):1343–1376, 2000.

[24] D. Duffie, M. Schroder, and C. Skiadas. Recursive valuation of defaultable securities and the timing of resolution of uncertainty. *Annals of Applied Probability*, 6:1075–1090, 1996.

[25] D. Duffie and K. Singleton. An econometric model of the term structure of interest-rate swap yields. *The Journal of Finance*, 52(4):1287–1321, September 1997.

[26] D. Duffie and K. Singleton. Modeling term structures of defaultable bonds. *Review of Financial Studies*, 12(4):687–720, 1999.

[27] P. Dufresne and B. Solnik. On the terms structure of default premia in the swap and libor markets. Working Paper, HEC, School of Management, 1998.

[28] E. Elton, M. Gruber, D. Agrawal, and C. Mann. Explaining the rate spread on corporate bonds. Working Paper. Stern School of Business, New York University, 1999.

[29] R. Gill and S. Johansen. A survey of product-integration with a view towards applications in survival analysis. *The Annals of Statistics*, 18(4):1501–1555, 1990.

[30] G. Golub and J. Ortega. *Scientific Computing and Differential Equations*. Academic Press, 1991.

[31] J. Harrison. *Brownian Motion and Stochastic Flow Systems*. Robert E, Krieger Publishing Company, Malabar, Florida, 1990.

[32] G. Hübner. The analytical pricing of asymmetric defaultable swaps. Forthcoming in: Journal of Banking and Finance, 1998.

[33] R. Jarrow, D. Lando, and S. Turnbull. A markov model for the term structure of credit risk spreads. *Review of Financial Studies*, 10(2):481–523, 1997.

[34] R. Jarrow and S. Turnbull. Pricing options on financial securities subject to credit risk. *Journal of Finance*, pages 53–85, March 1995.

[35] R. Jarrow and S. Turnbull. When swaps are dropped. *Risk*, 10(5):70–75, May 1997.

[36] M. Kijima. Default risk and derivative products. *Mathematical Finance*, 8(3):229–247, 1998.

[37] M. Kijima and K. Komoribayashi. A markov chain model for valuing credit risk derivatives. *Journal of Derivatives*, Fall:97–108, 1998.

[38] M. Kijima and Y. Muromachi. Credit events and the valuation of credit derivatives of basket type. *Review of Derivatives Research*, 4(1):55–79, 2000.

[39] C. Koo. The upshot: Investors can't squeeze numbers out of orange. TheStreet.com, 2001.

[40] S. Kusuoka. A remark on default risk models. In *Advances in Mathematical Economics*, volume 1, pages 69–82. Springer-Verlag, 1999.

[41] D. Lando. Three essays on contingent claims pricing. PhD Dissertation, Cornell University, 1994.

[42] D. Lando. Modelling bonds and derivatives with credit risk. In M. Dempster and S. Pliska, editors, *Mathematics of Financial Derivatives*, pages 369–393. Cambridge University Press, 1997.

[43] D. Lando. On cox processes and credit risky securities. *Review of Derivatives Research*, 2:99–120, 1998.

[44] H. E. Leland. Corporate debt value, bond covenants, and optimal capital structure. *Journal of Finance*, 49(4):1213–1252, September 1994.

[45] T. Li. A model of pricing defaultable bonds and credit ratings. Working Paper, John M. Olin School of Business, Washington Univ. St Louis, 1998.

[46] R. Litzenberger. Swaps: Plain and fanciful. *Journal of Finance*, pages 831–850, July 1992.

[47] D. Madan and H. Unal. Pricing the risks of default. *Review of Derivatives Research*, 2:121–160, 1998.

[48] R. C. Merton. A rational theory of option pricing. *Bell Journal of Economics and Management Service*, 4(1):141–183, 1973.

[49] R. C. Merton. On the pricing of corporate debt: The risk structure of interest rates. *Journal of Finance*, 2:449–470, 1974.

[50] B. Minton. An empirical examination of basic valuation models for plain vanilla u.s. interest rate swaps. *Financial Economics*, 44:251–277, 1997.

[51] D. Nakazato. Gaussian term structure model with credit rating classes. Working Paper, The Industrial Bank of Japan, 1997.

[52] P. Protter. *Stochastic Integration and Differential Equations-A New Approach*. Springer, New York, 1990.

[53] R. Rendleman. How risks are shared in interest rate swaps. *The Journal of Financial Services Research*, pages 5–34, 1992.

[54] P. Schönbucher. The term structure of defaultable bond prices. *Review of Derivatives Research*, 2:161–192, 1998.

[55] C. Smith, C. Smithson, and L. Wakeman. The evolving market for swaps. In R. Schwartz and C. Smith Jr., editor, *The Handbook of Currency and Interest Rate Risk Management*, chapter 6. New York Institute of Finance, 1990.

[56] J. Sobehart, R. Stein, V. Mikityanskaya, and L. Li. Moody's public firm risk model: A hybrid approach to modeling short term default risk. 2000.

[57] B. Solnik. Swap pricing and default risk: A note. *Journal of International Financial Management and Accounting*, 2(1):79–91, 1990.

[58] E. Sorensen and T. Bollier. Pricing swap default risk. *Financial Analysts Journal*, May-June:23–33, 1984.

[59] T. Sun, S. Sundaresan, and C. Wang. Interest rate swaps: An empirical investigation. *Journal of Financial Economics*, 34:77–99, 1993.

[60] S. Sundaresan. Valuation of swaps. In S. Khoury, editor, *Recent Developments in International Banking*, chapter XII, pages 407–440. Elsevier Science Publishers (North Holland), 1991.

[61] S. Turnbull. Swaps: A zero sum game. *Financial Management*, pages 15–21, Spring 1987.

[62] L. Wakeman. Credit enhancement. In C. Alexander, editor, *The Handbook of Risk Management and Analysis*, pages 307–327. John Wiley and Sons Ltd, 1996.

[63] T. Wilson. Portfolio credit risk (i). *Risk*, 10(9):111–117, September 1997.