CREDIT RISK
FROM THEORY TO APPLICATION

CREDIT RISK
FROM THEORY TO APPLICATION

By

CHUANG YI, B.Sc.

A Thesis

Submitted to the School of Graduate Studies

in Partial Fulfilment of the Requirements

for the Degree

Master of Science

McMaster University

MASTER OF SCIENCE (2005) McMaster University
(Mathematics) Hamilton, Ontario

TITLE: Credit Risk
From Theory to Application

AUTHOR: Chuang Yi, B.Sc.
(Beijing University of Aeronautics and Astronautics)

SUPERVISOR: Dr. Matheus Grasselli

NUMBER OF PAGES: xvi, 96

To my dearest parents and sister

Abstract

In this thesis, we calibrated a one factor CIR model for interest rate and a two factor CIR model for each hazard rate of 21 firms. The time series of the interest rate and each hazard rate for 21 firms are also obtained. Extended Kalman Filter and Quasi-Maximum Likelihood Estimation are used as the numerical scheme. The empirical results suggest that multifactor CIR models are not better than multifactor Hull-White model. Positive correlations between hazard rate and interest rate are discovered, although most hazard rates are found to be negatively correlated with the default-free interest rate. The 21 filtered time series of the hazard rates suggest that there maybe a hidden common factor shared only by the intensities. Monte Carlo Simulation is conducted both for interest rate and hazard rates. The simulation indicate that both the SKF and the EKF work pretty well as a filter tool but may produce bad estimation for the value of the likelihood function. QMLE works fine in linear state space form model, but it does a poor job in the case of non-linear state space form.

Acknowledgements

I am very grateful to Dr. Matheus Grasselli, Dr. Tom Hurd and Dr. Peter Miu for serving on my supervisory committee. Thanks to my supervisor Dr. Matheus Grasselli for his attentive guidance. Thanks to Dr. Tom Hurd for his sparkling discussions as well as exhaustive guidance. Thanks to Dr. Peter Miu, who is a professor at Michael G. Droote Business School at McMaster University, for strengthening my background in finance and in empirical calibration of the models.

Special thanks to Dr. Shui Feng, who taught me Brownian Motion and Stochastic Differential Equations, which inspired my interests in the field of financial mathematics. Thanks to Dr. Lia Bronsard, Dr. Aaron Childs, Dr. Fred M. Hoppe, Dr. Bartosz Protas, Dr. Roman Viveros-Aguilera, Dr. Eric T. Sawyer who have instructed my graduate courses. Thanks also to Dr. Maung Min-Oo and Dr. Hans U. Boden who are close friends of PhiMac.

I would like to thank Dr. Jingping Yang, a visiting professor from Peking University, for the enlightening discussions we had.

I would like to express my sincere gratitude to all of my colleagues in PhiMac: Xuping Zhang, Lingling Wang, Mengmeng Ma, Bogdan Traicu, Maria Betcheva, Elena Alexandru-Gajura, Mirela E. Cara, Kirill Mayorov, Raphael Yan, William Sajko. You are my sisters and brothers in this PhiMac family. I also appreciate the help from my friends in the department, such as Tao Li, Qihao Xie, Dahai He, Guihong Fan and

Zhenbin Yan.

Last but not the least, I want to say thanks to my dearest parents and sister for their perpetual love and support.

Contents

List of Tables

List of Figures

Introduction

According to Giesecke's (2004) definition, credit risk is the distribution of financial losses due to unexpected changes in the credit quality of a counterparty in a financial agreement. The changes in the credit quality is so unexpected that it is hardly to predict the default events, which usually will bring catastrophic financial losses. Voices from the financial markes are asking: how to avoid or at least minimize our expected losses, how to price defaultable financial instruments?

These challenges attract not only the people in the financial industries, but also the intellegent minds in the academia. Tools from other disciplines are widely used to tackle credit risk problems, such as mathematics, statistics, economics and computer programming. People have been answering these problems for years, vaguely however promissingly.

Fischer Black and Myron Scholes are in fact the first to publish their answers to these questions, although they are well known because of their contributions to option pricing. In 1973 they published their paper *The Pricing of Options and Corporate Liabilites* which led them to the laureate of the Nobel prize in economics in 1997. It is in that very paper, that they built the foundation of structure approach to analyzing credit risk in the second part *Corporate Liabilities*, although it has not gained much attention as it deserves. Ironically, following Black-Scholes idea, in 1974, Robert Merton independently published his paper on the pricing of corporate debt and which is now always cited as the first model of credit risk: *Merton's Model.* Black-Scholes and Merton's models are refered as classical models nowadays as a counterparty of the first-passage approach proposed by Black and Cox (1976).

Several drawbacks of structural models turned out to be noticed, which lead people to think about other methods. Reduced form credit models are then introduced as an alternative by Artzner and Delbaen (1995), Jarrow and Turnbull (1995), Lando (1998), Duffie and Singleton (1999).

Later on, incomplete information framwork is introduce by Duffie and Lando (2001), Giesecke (2001) and Cetin, Jarrow, Protter and Yildirim (2002), which is a combination of the first two approaches.

In this thesis, we will do a review of the development of the credit theory. More important, specific examples are given to illustrate how the theory is implemented and applied to real problems. Technical tools such as Extended Kalman Filter (EKF) and Quasi-Maximum Likelihood Estimation (QMLE) are used.

The contents are organized as follows. Chapter 1 is a short review of interest rate theory, focussing on modeling short rates under risk-neutral measure. Chapter 2 explains the terminologies at the beginning, then analyzes the three approaches to credit risk. In Chapter 3, focussing on intensity based models (reduced form approach), multi-factor Guassian models and CIR models are studied in detail. Chapter 3 is the conjunction of the previous two chapters and Chapter 4. In Chapter 4, we introduce the techniques to use the real market data to implement the model built in the previous chapter. Our aim is to calibrate the parameters in the model and back out the time series of the hazard rate processes.

Chapter 1

Interest Rate Theory

Interest rates play a crucial role in the pricing of both default-free and defaultable securities. As a result, one can not avoid considering interest rates when talking about credit risk. Also, the theory of credit risk has many similar characteristics with the theory of interest rates (for example, the adjusted interest rate in credit risk and the default-free interest rate). In order to keep the context complete, talking about the interest rate theory becomes a necessity.

Both practitioners in industry and the theorists in academia are enthusiastic about modeling the dynamics of the interest rate. This enthusiasm has never been cooled down ever since the first study on short rates by Vasiček (1977). Other papers on short rates are Dothan (1978), Ho and Lee (1986), Cox, Ingersoll and Ross (1985), Hull and White (1990). Forward interest rates are well studied in the monumental work of Heath, Jarrow and Morton (1992) which is one of the most important papers on interest rates ever published. A pricing kernel approach on positive interest rates is conducted by Flesaker (1993), Rogers (1997), Flesaker and Hughston (1997), Jin and Glasserman (2001). Other contributors on the modern interest rate theory include Duffie and Kan (1996), Brennan and Schwartz (1982), Longstaff and Schwartz (1992), just to mention a few. For a more complete list of the contributors and their related work, please refer to *The New Interest Rate Models* edited by Lane Hughston (2000).

In this chapter, $\{W(t); t \geq 0\}$ is a standard Brownian motion on some probability space $(\Omega, \mathcal{F}, \mathcal{P})$, and $\{\mathcal{F}(t); t \geq 0\}$ is the filtration generated by $W(t)$. Let $\mathcal{Q}$ denote the equivalent martingale measure, or risk-neutral measure. The time horizon is assumed to be finite through out this thesis, which means the maturity date T is bounded above. We review the short rate models, the forward rate framework and the affine term structures.

1.1 Zero Coupon Bonds

In this chapter, no credit risk is involved. The primary securities we are considering here are zero coupon bonds, also known as pure discount bonds. The following definition is from *Arbitrage Theory in Continuous Time* by Bjork (1998).

Definition 1.1.1. *A* ***zero coupon bond*** *with* ***maturity date*** *T, also called a T-* ***bond****, is a contract which guarantees the holder to be paid one dollar on the maturity date T. The price at time t of a T-bond is denoted by $P(t, T)$.*

Particularly, $P(t, t) = 1$ holds for all t by the definition. The bond price $P(t, T)$ could be analyzed from two different perspectives.

First, let's consider when t is fixed. The price $P(t, T)$ turns out to be a deterministic function of the maturity T. The graph of this function in T is called the term structure of the bond price at t. Typically it is a very smooth graph and thus it is reasonable to assume that $P(t, T)$ is differentiable w.r.t T.

On the other hand, let's consider $P(t, T)$ for fixed maturity date T. It is a stochastic process in t, however not a deterministic function in t. This process gives the price at different time t of a T-*bond.*

1.2 Interest Rates

There are many different kinds of interest rates with different definitions, however two of them are mostly concerned, namely **the forward rate** and **the short rate**. They are defined respectively as follows.

Definition 1.2.1. *The* ***instantaneous forward rate with maturity*** T*, or simply* ***forward rate*** *at time* t*, is defined by*

$$f(t,T) = -\frac{\partial \ln P(t,T)}{\partial T}. \tag{1.1}$$

The ***short rate*** *at time* t*, also known as* ***spot rate*** *is defined by*

$$r(t) = f(t,t). \tag{1.2}$$

It is easy to see from the definition that the forward rate $f(t,T)$ is positive if and only if the term strucuture of the bond prices is a decreasing funtion in T. The Figure 1.1 illustrates a typical zero coupon bond price with positive interest rate.

Indirect modeling the bond prices by modeling the forward rates or short rates becomes a fashion in recent thirty years. The following two sections will discuss modeling these two kinds of interest rates respectively.

1.3 Modeling Short Rates Under Risk-Neutral Measure

Short rates could be modelled under the physical measure $\mathcal{P}$, however the bond prices are not uniquely determined by the $\mathcal{P}$*-dynamics* of the short rate $r(t)$ because of the lack of specification of the market price of the risk which is determined by the market. Therefore, it is favored to model the short rate directly under the risk-neutral measure $\mathcal{Q}$.

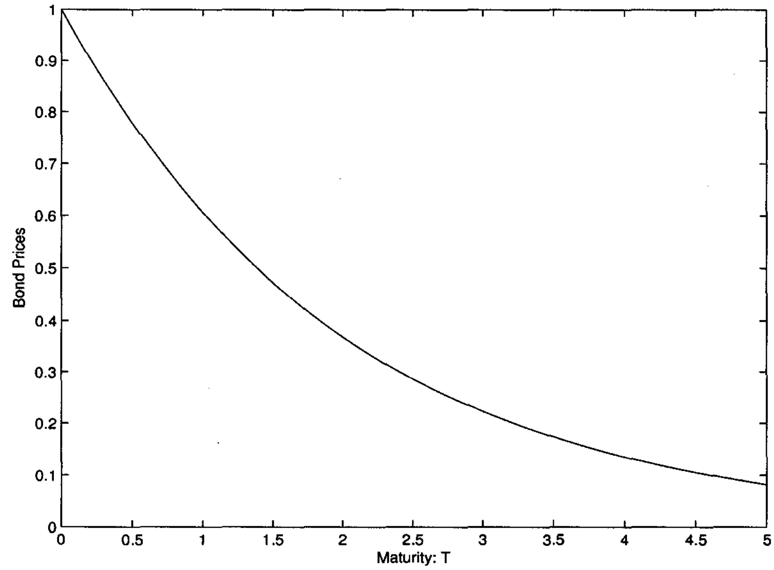

Figure 1.1: Term Structure of The Bond Price at Time $t = 0$ with Positive Interest Rate

A general model of the short rate r_t (from now on, we will use r_t to denote r(t)) under $\mathcal{Q}$-*dynamics* is written as a stochastic differential equation (SDE)

$$dr_t = \mu(t, r_t)dt + \sigma(t, r_t)dW_t, \tag{1.3}$$

where $\mu(t, r_t)$ and $\sigma(t, r_t)$ are deterministic functions in t and r_t. They are called the drift and the volatility respectively of the $\mathcal{Q}$-*dynamics* of the interest rate r_t.

Once the $\mathcal{Q}$-*dynamics* of the short rate r_t is specified, under the no-arbitrage framework, the zero-coupon bond prices could be given by

$$P(t, T) = E^Q[\exp(-\int_t^T r_s ds) \mid \mathcal{F}_t] \tag{1.4}$$

Ever since Vasiček (1977), a great number of proposals have been made on how to specify the $\mathcal{Q}$-*dynamics* for r_t. So far, no standard model for short rates is unanimously accepted. Among the existing models, the following ones are mostly favored.

- Vasiĉek

$$dr_t = (\alpha - \beta r_t)dt + \sigma dW_t,$$

- Cox-Ingersoll-Ross

$$dr_t = \kappa(\theta - r_t)dt + \sigma\sqrt{r_t}dW_t,$$

- Hull-White (Extended Vasiĉek)

$$dr_t = (\alpha(t) - \beta r_t)dt + \sigma dW_t,$$

The parameters β and κ are required to be positive in the above models so as to keep the mean-reverting property of the dynamics. The function $\alpha(t)$ in Hull-White, is not random but deterministic function in t.

The SDEs of Vasiĉek and Hull-White models could be solved analytically and it turns out that they are Gaussian processes. This allows negative interest rates because of the properties of Gaussian processes. However, they are very popular in industry since they can fit the market data comparatively well given positive initial value of their SDEs. On the other hand, the CIR does not have negativity problem, but it does not fit the market data so well.

1.4 Affine Term Structures

Affine term structure modeling is an extention of the above models. It is well developed and used in interest rates modeling as well as in credit risk because of its pleasing analytical tractability.

Definition 1.4.1. *If the term structure of the bond prices could be written into*

$$P(t,T) = \exp(A(t,T) - B(t,T)r_t) \tag{1.5}$$

then the model is said to have an affine term structure.

The functions $A(t,T)$ and $B(t,T)$ above are functions of the two real variables t and T, however, in most cases, they are functions of the difference between the two variables, i.e. they are functions of $(T-t)$.

Current research puts a lot of efforts on the application of affine processes in finance. It deserves to mention an important work made by Duffie, Filipovic and Schachermayer (2001).

The affine term structure models do exist. A special case is that when the drift and the square of the volatility term of the short rate r_t under *Q-dynamics* are linear in r_t. In this case, the problems shrink to solving the Ricatti equations for the function $A(t,T)$. This could be described by the following proposition.

Proposition 1.4.1. ***(Affine Term Structure)*** *If the drift μ and volatility σ in the Q-dynamics of the short rate r_t have the following form*

$$\begin{cases} \mu(t,r_t) & = & \alpha(t)r_t + \beta(t), \\ \sigma(t,r_t) & = & \sqrt{a(t)r_t + b(t)}. \end{cases} \tag{1.6}$$

then the model admits an affine term structure of the form (1.5), where A and B satisfy the system

$$\begin{cases} \frac{\partial A(t,T)}{\partial t} & = & \beta(t)B(t,T) - \frac{1}{2}b(t)B^2(t,T), \\ A(T,T) & = & 0. \end{cases} \tag{1.7}$$

$$\begin{cases} \frac{\partial B(t,T)}{\partial t} & = & \frac{1}{2}a(t)B^2(t,T) - \alpha(t)B(t,T) - 1, \\ B(T,T) & = & 0. \end{cases} \tag{1.8}$$

The Vasiček, CIR and Hull-White all fall in this situation.

Proposition 1.4.2. ***(Vasičk)*** *In the Vasičk model, the bond prices are given by*

$$P(t,T) = \exp(A(t,T) - B(t,T)r_t),$$

where

$$B(t,T) = \frac{1 - \exp(-\beta(T-t))}{\beta},$$

$$A(t,T) = \frac{(B(t,T) - T + t)(\alpha\beta - \frac{1}{2}\sigma^2)}{\beta^2} - \frac{\sigma^2 B^2(t,T)}{4\beta}.$$

Proposition 1.4.3. ***(CIR)*** *In the CIR model, the bond prices are given by*

$$P(t,T) = A(t,T)\exp(-B(t,T)r_t),$$

where

$$A(t,T) = \left\{\frac{2\gamma\exp[(T-t)(\kappa+\gamma)/2]}{(\kappa+\gamma)(\exp[\gamma(T-t)]-1)+2\gamma}\right\}^{2\kappa\theta/\sigma^2},$$

$$B(t,T) = \frac{2(\exp[\gamma(T-t)]-1)}{(\kappa+\gamma)(\exp[\gamma(T-t)]-1)+2\gamma},$$

and

$$\gamma = \sqrt{\kappa^2 + 2\sigma^2}.$$

Proposition 1.4.4. ***(Hull-White)*** *In the Hull-White model, the bond prices are given by*

$$P(t,T) = \exp(A(t,T) - B(t,T)r_t),$$

where

$$B(t,T) = \frac{1 - \exp(-\beta(T-t))}{\beta},$$

$$A(t,T) = \int_t^T \{\frac{1}{2}\sigma^2 B^2(s,T) - \alpha(s)B(s,T)\}ds.$$

In Hull-White model, the function $\alpha(t)$ is chosen so as to fit the observed forward rate curve at $t = 0$, $\{f(0,T); T > 0\}$. The above propositions will be needed to implement the calibration of those models for short rates.

1.5 Heath-Jarrow-Morton

In their 1992 paper, Heath, Jarrow and Morton studied the interior relationships of a general framework for forward rates. Although their work is sometimes refered as **HJM model**, it in fact should really be refered as **HJM framework** because of its great generality built in. Their main contribution in the 1992 paper could be summarized as following.

Theorem 1.5.1. ***(Heath-Jarrow-Morton)*** *For each positive T, let $\alpha(\mu, T)$ and $\sigma(\mu, T)$ be adapted processes, where $0 \leq \mu \leq T$. Assume $\sigma(\mu, T) > 0$ for all μ and T. Let $f(0, T)$ be a deterministic function, and define*

$$f(t, T) = f(0, T) + \int_0^t \alpha(\mu, T) d\mu + \int_0^t \sigma(\mu, T) dW(u). \tag{1.9}$$

Then $f(t, T)$, for $0 \leq t \leq T$ is a family of forward rate processes for a term-structure model without arbitrage if and only if there exists an adapted process $\theta(t)$, such that

$$\alpha(t, T) = \sigma(t, T) \int_t^T \sigma(t, \mu) d\mu + \sigma(t, T)\theta(t). \tag{1.10}$$

where $\theta(t)$ is called the market price of risk.

A simple scenario is when the market price of risk is zero, mathematically $\theta(t) = 0$. This simply implies that the model is built in an equivalent martingale measure $\mathcal{Q}$, in which the no-arbitrage condition is satisfied. Note that the market price of risk does not depend on the maturity date T of the bond.

More detailed discussion could be found in *Stochastic Calculus and Finance* by Steven Shreve, 1997.

This framework is very general, however it allows negative forward rates. In order to enforce positivity on the forward rate $f(t, T)$, more conditions on the drift and the volatility of the dynamics $f(t, T)$ should be satisfied. This discussion can be found in Jin and Glasserman (2001).

1.6 Yield to Maturity

The short rate is the objective what we model, however we are not able to observe the short rate directly from the market. What we do able to observe is the *yield to* *maturity* which is defined as follows.

Definition 1.6.1. *For any fixed time t, the **yield to maturity** of a T-bond is given by*

$$Y(t,T) = -\frac{1}{T-t}\ln P(t,T).$$

The short rate and the yield to maturity are thus related through the bond prices. For affine term structure models of the short rate, it is easy to see from the above definition that the yield to maturity $Y(t,T)$ is linear in the short rate r_t.

Yield curve is a graph of the yield to maturity given a fixed t. Alternatively, we can think about the yield curve as a function of T given any fixed t. The following figure of the yield curve of the U.S. government strips on September 22nd, 2004, is from Bloomberg.

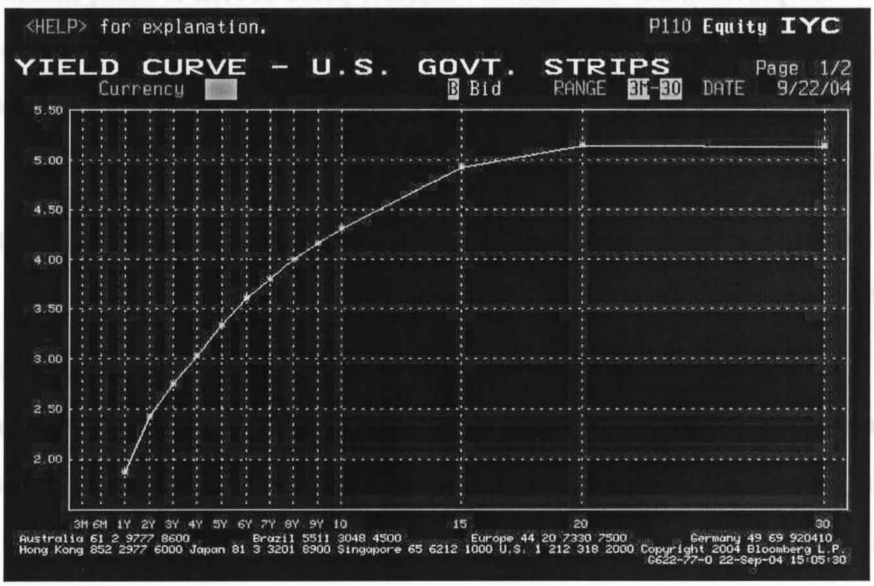

Figure 1.2: Yield Curve - U.S. Government Strips, 9/22/04

Chapter 2

Credit Risk Theory

In the first chapter, we have discussed the interest rate theory where no credit risk is involved. Defaut free zero coupon bonds play an essential role in analyzing the default free interest rate. In this chapter, defaultable bonds will be added into our scenario and credit risk is thus induced.

Although people have been facing credit risk ever since early ages, credit risk has not been widely studied until recent 30 years. Early literature(before 1974) on credit risk uses traditional actuarial methods of credit risk, whose major difficulty lies in their complete dependence on historical data.

Up to now, there are three main quantitative approaches to analyzing credit risk: **structural approach**, **reduced form approach** and **incomplete information approach**.

Merton (1974) firstly builds a model based on the capital structure of the firm, which becomes the basis of the structural approach. In his approach, the company defaults at the bond maturity time T if its assets value falls below some fixed barrier at time T. Thus the default time τ is a discrete random variable which picks T if the company defauls and infinity if the company does not default. As a result, the equity of the firm becomes a contingent claim of the assets of the firm's assets value. Black and Cox (1976) extends the definition of default event and generalize Merton's method into the first-passage approach. In Black and Cox (1976), the firm defaults when the history low of the firm assets value falls below some barrier D. Thus, the default event could take place before the maturity date T.

Intensity-based approach, also known as reduced form approach, as a counterparty of the structure approach, is introduced by Artzner & Delbaen (1995), Jarrow & Turnbull (1995) and Duffie & Singleton (1999). In this approach, the default event is modelled as either a stopped Possion process or a stopped Cox process with intensity h_t. The intensity h_t is then called hazard rate in reduced form approach, since the product of h_t and an infinitesimal time period dt is the default probability of the firm at that infinitesimal time period dt given the firm has not default yet before time t. It was showned in Lando (1998) and Duffie & Singleton (1999) that the defaultable bonds can be calculated as if they were default-free using an interest rate that is the risk-free rate adjusted by the intensity.

Incomplete information approach is developed by Duffie & Lando (2001), Giesecke (2001), Cetin, Jarrow, Protter & Yildirim (2002), which is a combination of structural approach and intensity-based approach. In this approach, the default event is directly modelled as a point process N_t with one jump of size of one at default. This point process N_t is a positive submartingale and could be decomposed into a martingale plus its compensator A_t by Doob-Meyer decomposition theorem. In reduced form approach the compensator A_t of the default process N_t can be represented as a definite integration of the hazard rate h_t. Incomplete information approach generalizes the forms of the compensator A_t which may not be represented as an integration of the hazard rate h_t. People turn to model the compensator A_t directly from the definition of default instead of modeling h_t.

This chapter will simply introduce the concept of credit risk and the common three approaches to analyzing it. Intensity-based approach will be used as our empirical modeling study. Therefore more detailed discussion about the second approach is conducted in a seperate chapter, chapter 3.

2.1 Terminologies in Credit Risk

Risk is almost everywhere at every second on everyone. Although normally the probability of risky events is comparatively small, people are widely aware that it should not be ignored because of its catastrophic aftermath, such as September 11, 2001. Financial risks can also lead to tragic results, such as the bankcruptcy of Long Term Capital Management in 1998.

2.1.1 Credit Risk

Financial risks are the risks people facing in the financial markets, such as the fluctuations of the bonds' prices, the changes of the default-free interest rate, or unexpected defaults, etc. Credit risk is one of the most important financial risks in the markets.

Credit risk is the risk induced from credit events such as credit rating change, restructuring, failure to pay, repudiation and bankruptcy etc. A more mathematical definition is given by Giesecke (2004): *credit risk is the distribution of financial losses due to unexpected changes in the credit quality of a counterparty in a financial agreement.*

In the case of the last three events mentioned above, they are also refer as **default events**. Default events induce **default risk** which is a risk that your counterparty may fail to honor a financial agreement. Since we are mostly concern with default events, credit risks are always speaked of as default risk as well.

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

Table 2.1: Credit Rating Categories by S&P's and Moody's

Credit risk is believed to have two components for a given company. One is called **systematic risk** which is from the common factors in the market, while the other one is refered as **idiosyncratic risk** which is from specific factors of the firm itself. Systematic risk is market-wide factors affecting all the firms and which capture the contagious characteristics of credit risk in the market. On the other hand, idiosyncratic risk is firm-specific which depicts the health of the firm. Systematic and idiosyncratic risks are usually assumed to be independent when we model the credit dynamics.

2.1.2 Credit Ratings

The health of the firm is reflected from the **credit ratings**. Standard & Poor's and Moody's are the major two credit rating agencies in North America. Table 2.1 is the symbols used by Standard & Poor's and Moody's respectively for rating the obligors (this table is also used in Li (2002)). Each firm is assigned to one of the categories based on their creditworthiness.

AAA is the highest rating assigned by S&P's, while Aaa is the corresponding highest assigned by Moody's. Generally speaking, the healthier a firm is the higher rating it will get. For a AAA company, it is less probable to default than a BBB company in the same time period. Table 2.2 depicts the average cumulative default probabilities for different rating classes over different time periods (this tale is also used in Giesecke (2002)). It can also be seen from the table that the firm is more likely to default in a longer term than a short period. If we let time goes to infinity, we would like to think that all the firms will default eventually.

Years	1	2	3	4	5	10
AAA	0.00	0.00	0.07	0.15	0.24	1.40
AA	0.00	0.02	0.12	0.25	0.43	1.29
A	0.06	0.16	0.27	0.44	0.67	2.17
BBB	0.18	0.44	0.72	1.27	1.78	4.34
BB	1.06	3.48	6.12	8.68	10.97	17.73
B	5.20	11.00	15.95	19.40	21.88	29.02
CCC	19.79	26.92	31.63	35.97	40.15	45.10

Table 2.2: Average Cummulative Default Probabilites of Different Rating Classes (in %). Standard and Poor's, 2001.

The credit rating for each firm is not static but dynamically changing over time depending on the changes of creditworthiness of the firm. Alternatively speaking, companies may transfer from one credit category to another. However, the probability of remaining in the same category is comparatively larger than the transition probability to another. Table 2.3 is the transition matrix of U.S industries from 1981 to 2001.

Different models of the transition matrix are proposed in the literature. For example, Lando (1998) suggests using a generalized Markovian model for the credit rating transition matrix, in which the last state of rating is absobing state in Markovian Chain representing the default event.

2.1.3 Liquidity Risk

Liquidity risk is the lack of supply or demand when you intend to buy or sell a large amount of goods. There are many papers in the literature on liquidity risk. However, in this thesis, we assume that we are living in a liquid market where you can sell or buy as many as you want.

	AAA	AA	A	BBB	BB	B	CCC	D
AAA	89.55	8.80	0.83	0.17	0.08	0.00	0.00	0.00
AA	0.50	86.79	6.43	0.54	0.07	0.06	0.03	0.02
A	0.06	1.91	87.35	4.74	0.42	0.15	0.05	0.05
BBB	0.04	0.24	4.41	85.43	4.06	0.60	0.17	0.26
BB	0.04	0.07	0.43	6.48	79.00	7.33	0.88	1.18
B	0.00	0.10	0.26	0.34	5.23	79.32	4.04	6.28
CCC	0.14	0.00	0.29	0.82	1.85	9.60	54.37	27.84
D	0.00	0.00	0.00	0.00	0.00	0.00	0.00	93.96

Table 2.3: One Year Bivariate Transition Matrix-U.S All Industries (in %), 1982-2001. Source: Giesecke, Kay (2002).

2.1.4 Recovery Rate and Loss Rate

So far, we have not specify how much the bond investors lose in the event of default. They may receive nothing when default happens. In this case, we refer it as **zero recovery rate**. However, in practice, investors usually receive some recovery payment upon default. **Loss Rate** is defined by one minus recovery rate. In the case of zero recovey rate, the loss rate thus is 1.

A variety of ways on modeling the recovery rate of defaultable claims have been proposed in the literature: recovery of face value, recovery of an equivalent default free bond and recovery of market value.

Under reduced form framefork, as in Duffie & Singleton (1999), we model the recovery of market value as a constant $1 - L$, in which L denotes the corresponding loss rate.

2.1.5 Credit Spreads

As in interest rate theory, we use $P(t,T)$ to denote the price of a default-free zero-coupon bond paying \$1 at maturity date T. Defaultable bonds are added into our consideration when credit risk is involved. Empirically, we assume treasuary bonds issued by the government are default-free while corporate bonds issued by firms are defaultable. A defaultable bond is always mentioned together with its issuer. Thus we use $P_j(t,T)$ denotes the price of a defaultable zero-coupon bond issued by firm j paying \$1 at maturity date T given there is no default. We will see in this section that credit risk theory share many similarities with interest rate theory.

As the term structure of default free bond prices in interest rate theory, the term structure of defaultable bond price $P_j(t,T)$ is a function of T, for fixed t. We only consider the bond prices when $T \geq t$. We assume both bond prices vanish to 0 as T goes to infinity.

Recall that the **Default-Free Yield Spread** $Y(t,T)$ is defined by

$$Y(t,T) = -\frac{\ln P(t,T)}{T-t}. \tag{2.1}$$

We could similarly define **Defaultable Yield Spread** $Y_j(t,T)$ of firm j through

$$Y_j(t,T) = -\frac{\ln P_j(t,T)}{T-t}. \tag{2.2}$$

Credit Yield Spread $S_j(t,T)$ for defaultable bond $P_j(t,T)$ is defined by the difference of those two yields mentioned above

$$S_j(t,T) = Y_j(t,T) - Y(t,T) = -\frac{1}{T-t}\ln\frac{P_j(t,T)}{P(t,T)}. \tag{2.3}$$

The term structure of credit yield spread $S_j(t,T)$ is a function of T, for fixed t. Giesecke (2004) points out that it should tend to increase with increasing maturity, reflecting the fact that uncertainty is greater in the distant future than in the near term.

Recall that the **Default-Free Forward Rate** $f(t,T)$ is defined by

$$f(t,T) = -\frac{\partial}{\partial T} \ln P(t,T). \tag{2.4}$$

and the **Default-Free Spot Rate** $r(t)$ is defined by $r(t) = f(t,t)$. Usually, we informally call $r(t)$ default-free interest rate or simply interest rate.

Similarly we could define **Forward Default Rate** $R_j(t,T)$ for bond $P_j(t,T)$

$$R_j(t,T) = -\frac{\partial}{\partial T} \ln P_j(t,T). \tag{2.5}$$

The **Spot Default Rate** $R(t)$ is defined by $R(t) = R(t,t)$.

The **Credit Forward Spread** $H_j(t,T)$ is defined by the difference of the defaultable forward rate and the default-free forward rate

$$H_j(t,T) = R_j(t,T) - f(t,T) = -\frac{\partial}{\partial T} \ln \frac{P_j(t,T)}{P(t,T)}.$$

The **Spread** $H_j(t)$ is defined by $H_j(t) = H_j(t,t)$. People usually use $h(t)$ to denote hazard rate. In reduced form models, it could be shown that the credit spot spread is the product of the hazard rate and the loss rate L. Since empirically the hazard rate and the loss rate can not be observed seperately, the product (credit spot spread) of the two is usually modelled directly. Hazard rates will be discussed in detail in the section of *Reduced Form Approach.*

Credit Spreads (credit yield spread, credit forward spread, credit spot spread) of different firms are correlated through time in the market. Firstly, the correlation means that these firms share common economic factors in the market. Secondly, people believe that the credit spreads of other firms will jump upon one firm defaults. Therefore, we often say credit risk is *contagious.*

Ideally, we would like to assume positivity as a characteristic of both default free forward rate and credit forward spread which is equivalent to impose positivity on interest rate and hazard rates. If we impose this positivity, it can be easily shown that the term structures of default-free and defaultable bond prices should satisfy the following conditions (for fixed t):

- Both $P(t,T)$ and $P_j(t,T)$ are smoothly decreasing in terms of T
- $P_j(t,T)$ is always less than $P(t,T)$ for the same T
- $P_2(t,T)$ is always less than $P_1(t,T)$ given firm 2 has a higher credit forward spread

These restrictions also guarantee positive credit yield spread. Credit spreads are used to measure **credit premium**, which compensates risk-averse investors for assuming credit risk. Therefore, the credit spreads should remain positive. The higher credit risk assumed by the investors, the higher credit premium got be payed by them.

These conditions mentioned above are depicted by Figure 2.1.

2.1.6 Default Time and Default Processes

Default time τ is when the default happens. It is modelled as a random process which is specified in a model assumption. In Merton's method, which we will study later on, the default time is a predictable discrete random variable. But in reduced form approach, the default time τ is an unpredictable process.

Default process N_t is defined through default time τ by

$$N_t = 1_{\{\tau \leq t\}} = \begin{cases} 1 & if \quad \tau \leq t \\ 0 & if \quad else. \end{cases}$$

In general, the default process N_t is just a point process. But this is too general to conduct research on credit risk. Therefore, in different models, it is specified under

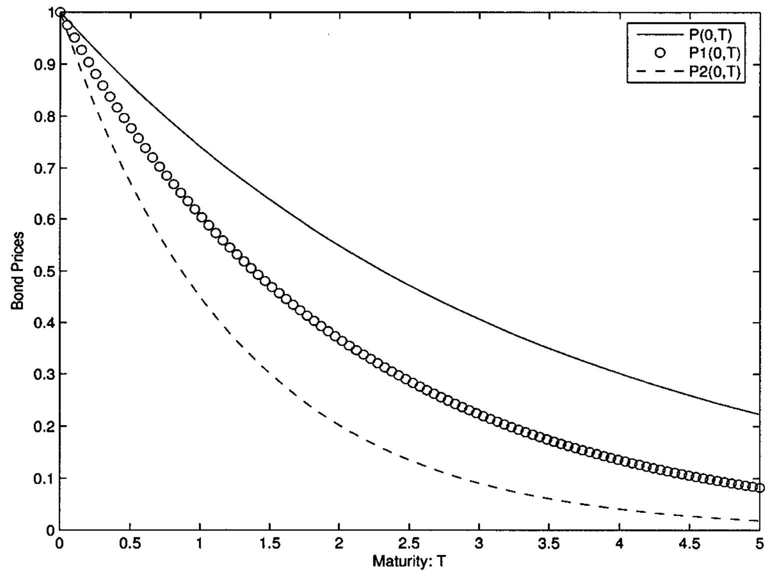

Figure 2.1: Term Structure of Bond Prices at $t = 0$, with Positivity Restrictions on Interest Rate and Hazard Rates. Firm 2 Has a Higher Credit Spread Than Firm 1.

different assumptions. Such as reduced form models, the default process is assumed to be either a stopped Poisson or a stopped Cox.

Default time and default process will be discussed in detail under different scenarios when we talk about different approaches to analyzing credit risk.

2.1.7 Default Probabilities

The portfolio manager would like to know how likely a firm will default before some fixed time. Thus the default probabilities play an essential role in managing credit risk as well as pricing credit claims. Consider firm j, let τ_j denote the default time of that firm. The physical default probability of the firm j before time T_j is mathematically calculated by $P[\tau_j < T_j]$ which is called the marginal default probability for firm j in multi-firm case. After a change of measure, the physical probability could be transfered into risk-neutral default probability $Q[\tau_j < T_j]$. The survival probability for firm j is defined by the difference of one and its default probability.

Consider multi-firm cases, say n firms, the joint probability of each firm j defaults before time T_j, for $j = 1, 2, ..., n$ is calculated by

$$P[\tau_1 < T_1, \tau_2 < T_2, ..., \tau_n < T_n].$$

If all the n firms are mutually independent, then the joint default probability is thus given by the product of their marginal default probabilities. However, as we have mentioned a lot previously that the default events may highly correlated. Then, the **default correlation** is introduced.

2.1.8 Default Correlation

Default correlation is far from exhausted understanding by us, but it is one of the most interesting subject in credit risk theory. In fact, default correlation is an ambiguous expression to describe the default dependency due to the short history of research on this subject. In Merton's method, default correlation is modelled by assuming the dependency of the assets of the firms. Some structure models use the equity correlations to approximate the asset correlation. In reduced form models, default correlation is thought as the correlation of the hazard rates.

Whatever correlation you are talking about, the correlation of the default time τ_j is among the most important. Because when the correlation of the default time τ_j is given, as well as the marginal default probabilty, then the joint default probability can be derived. This is one way of modeling, namely starting from marginal default probabilities plus correlation to obtain joint default probability. A typical example is **factor models** which we will discuss in detail in the next chapter. Some researchers like Hull-White propose to use the other way around, namely beginning from joint default probability to marginal default probabilities and correlation. For the second approach, **copula functions** become a major mathematical tool on dependency modeling. Gaussian, student-t, and Archimedean copulas are among the most popular. Please refer to Nelsen (1999) for details on copulas.

2.1.9 Credit Derivatives

One of the reason why modeling on default dependency becomes popular is the emergence of multi-name credit instruments involving several reference entities, such as basket Credit Default Swaps (CDS) and Collateralized Debt Obligations (CDO).

Credit derivatives can separate the credit risk from an underlying and thereby enable investors to reduce the exposure to credit risk. As a result, the credit derivatives market had grown from $200Bn in 1997 to more than $1600Bn in 4 years.

One of the main tasks of credit risk theory is to price those credit derivatives. Another one is to effectively use these derivatives to hedge credit risk. These practical problems arising from the real world are, however, dealt with by using abstract mathematics.

2.2 Structure Approach

Although the title of this chapter is called *Credit Risk Theory*, the problems of credit risk are originally from the real life of practicing in financial markets rather than theoretic reasoning itself. However mathematics and theoretical reasoning are the major tools to tackle these problems. From now on, more mathematics will be involved in analysing instead of literally description.

Structure approach is firstly pioneered by Merton (1974), and then is extended to **first-passage approach** by Black and Cox (1976).

2.2.1 Merton's Method

Consider a firm j with market value V, which is financed by equity and a single issue of T-*bond* with face value K. Suppose the market we live in is frictionless in which we could trade continuously.

The total market value of the firm V is modelled by a geometric Brownian motion with constant drift μ and constant volatility σ, i.e.

$$dV_t = \mu V_t dt + \sigma V_t dW_t, \quad V_0 > 0.$$

We know that the solution of this SDE is simply

$$V_t = V_0 \exp((\mu - \frac{1}{2}\sigma^2)t + \sigma W_t).$$

Default-free interest rate is assumed to be a positive constant r, therefore the risk-free bond prices are given by

$$P(t,T) = K \exp(-r(T-t)).$$

Default time τ is the next crucial ingredient we need to model. The firm is assumed to default at the bond maturity date T, if the total market value of the firm is not sufficient to pay its obligation to the bond holders. In this situation, the bond holders immediately take over the firm. Thus the default time τ is a discrete random variable given by

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

We have thus set up a model for the credit risk of the firm j. The following analysis is for two objectives: firstly, the default probability, second, the pricing of the equity E and the defaultable bond $P_j(0,T)$.

Situation	Assets	Bonds	Equity
No Default	$V_T \geq K$	K	$V_T - K$
Default	$V_T < K$	V_T	0

Table 2.4: Payoffs at maturity of a firm

Since W_T is normally distributed with mean zero and variance T, default probability $p(T)$ is given by

$$p(T) = P[V_T < K] = P[\sigma W_T < \log l - mT] = \Phi\left(\frac{\log l - mT}{\sigma\sqrt{T}}\right)$$

where we define $l = \frac{K}{V_0}$ as the initial leverage ratio, and $m = \mu - \frac{1}{2}\sigma^2$.

From debt covenants priority assumption and limited liability of the firm, it is easy to see that the payoffs (equity) at maturity of the firm is $(V_T - K)^+$ which is equivalent to the payoff of a European call option on the assets of the firm with strike K and maturity T. Thus, pricing equity and credit risky debt is reduced to pricing European options. Table 2.4 depicts the situation of the equity of the firm.

The equity value thus is given by the Black-Scholes call option formula

$$E_0 = V_0\Phi(d_1) - \exp(-rT)K\Phi(d_2)$$

where

$$d_1 = \frac{(r + \frac{1}{2}\sigma^2)T - \log L}{\sigma\sqrt{T}}, \quad d_2 = d_1 - \sigma\sqrt{T}.$$

The bond payoff is $K - (K - V_T)^+$ which is equivalent to hold a portfolio composed of a long T-*bond* with face value K and a short European put on the asset of the firm with strike price K and maturity T. Therefore the defaultable bond price at time $t = 0$ is

$$P_j(0, T) = K\exp(-rT) - P(\sigma, T, K, r, V_0)$$

where $P(\sigma, T, K, r, V_0)$ is the Black-Scholes formula for put options. Using the *put-call* *parity*, we could also write the above equation into

$$P_j(0,T) = V_0 - E_0.$$

Recall the definition of *credit yield spread*, or simply credit spread, from the previous section

$$S_j(t,T) = -\frac{1}{T-t}\ln(\frac{P_j(t,T)}{P(t,T)}).$$

Thus, the credit spread in Merton's model at time $t = 0$ is given by

$$S_j(0,T) = -\frac{1}{T}\log(\Phi(d_2) + \frac{1}{L}\exp(rT)\Phi(-d_1)),$$

which is a function of the maturity date T, the asset volatility σ, and the leverage ratio L.

The short spread is the credit spread for maturities going to zero. It is the credit premium over an infinitesimal period of time demanded by the bond holders as a compensation for bearing the default risk. It is easy to see from the graphs (also easy to check from mathematical point of view) that the short spreads tend to zero when the leverage ratio is less than one, while go to infinity for the leverage ratio not less than one.

The Figure 2.2 and the Figure 2.3 show the credit spreads in basis points with different parameters.

2.2.2 First-Passage Method

In Merton's model, the firm can only default at the maturity date T. However, it is recognized that, as in Black-Cox (1976), the firm may default before the maturity date T. As in Merton's model, we still use geometric Brownian motion to model the total market value of the firm V_t, but the definition of default time τ is modified to

$$\tau = \min\{t > 0 : V_t \leq D\}$$

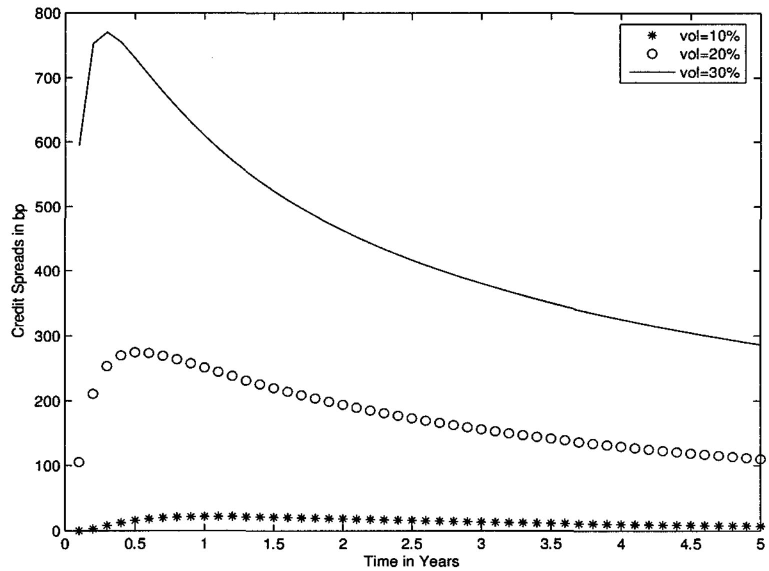

Figure 2.2: Term Structure of Credit Spreads at time 0 with the interest rate $r = 6\%$ per year and leverage ratio $L = 0.9$, for different volatilities.

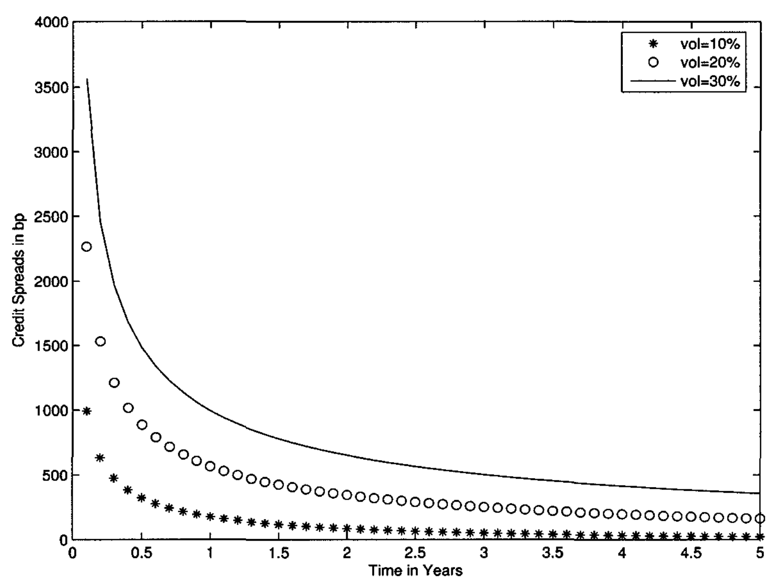

Figure 2.3: Term Structure of Credit Spreads at time 0 with the interest rate $r = 6\%$ per year and leverage ratio $L = 1.0$, for different volatilities.

where D is called the barrier of the firm's assets.

This definition says a default takes place when the assets of the firm fall to some positive level D for the first time. The firm is assumed to take the position of not default at time $t = 0$. Thus the barrier D is bounded above by V_0, i.e. $D < V_0$.

Since $\{\tau \leq t\} = \{\min_{s \leq t} V_s \leq D\}$, the default probability is given by

$$P[\tau \leq t] = P[\min_{s \leq t}(ms + \sigma W_s) \leq \ln(D/V_0)],$$

where $m = \mu - \frac{1}{2}\sigma^2$.

We know that the historical low $M_t = \min_{s \leq t}(ms + \sigma W_s)$ follows the inverse Gaussian distribution, and it follows that the default probability can be calculated explicitly by

$$P[\tau \leq t] = \Phi\left(\frac{\ln(D/V_0) - mt}{\sigma\sqrt{t}}\right) + \exp(\frac{2m\ln(D/V_0)}{\sigma^2})\Phi\left(\frac{\ln(D/V_0) + mt}{\sigma\sqrt{t}}\right).$$

Consider a defaultable zero coupon bond with zero recovery rate which means that the bond holders get nothing when a default happens. Under the no arbitrage argumentation, this bond price is given by the discounted payoff under the risk-neutral measure

$$P_j(0,T) = E^Q[\exp(-rT)1_{\{\tau > T\}}] = \exp(-rT)(1 - Q[\tau \leq T]).$$

The risk-neutral probability $Q[\tau \leq T]$ can be calculated by using the previous result for $P[\tau \leq T]$ if we set $\mu = r$.

The term structure of credit spreads in first-passage method is pretty much similar to the structures in Merton's model. But the short spreads are always zero in first-passage method. This is not plausible, since the bond holders pay a non-zero premium to compensate the risk that the firm may default in the next infinitecimal time.

Also, the credit spreads decrease as the maturities increase both in Merton's method and first-passage method which contradicts to empirical observation that spreads tend to increase with increasing maturity, reflecting the fact that uncertainty is greater in the distant future than in the short term, pointed by Giesecke (2004).

Because of these drawbacks in structural approach, people turn to figure out some other new method such as reduced form approach, which is the main theme of the next section.

2.3 Reduced Form Approach

Reduced form approach or *intensity-based approach* goes back to Artzner and Delbaen (1995), Jarrow and Turnbull (1995), Lando (1998), Duffie and Singleton (1999). The basic idea is based on modeling the default process as a stopped possion process (Giesecke, 2002).

Let $T_1, ..., T_n$ denote the arrival times of some event, say default. We call the sequence $\{T_i\}$ a homogeneous Poisson process with intensity λ if the inter-arrival times $T_{i+1} - T_i$ are independent and exponentially distributed with parameter λ.

Equivalently, letting $N(t) = \sum_i 1_{\{T_i \leq t\}}$ count the number of event arrivals in the time interval $[0, t]$, we say that $N(t)$, for $t \geq 0$, is a homogeneous Poisson process with intensity λ if the increments $N(t) - N(s)$ are independent and have a Poisson distribution with parameter $\lambda(t - s)$ for $s < t$, i.e.

$$P[N(t) - N(s) = k] = \frac{1}{k!}(\lambda(t-s))^k e^{-\lambda(t-s)}.$$

Suppose λ is a constant, in intensity-based approach, we set the default time to be the first jump of the Poisson process $N(t)$. Therefore, default time τ is exponentially distributed with parameter λ and the default probability can be expressed as

$$F(t) = P[\tau \leq t] = 1 - e^{-\lambda t}.$$

The intensity is the conditional default arrival rate given no default up to time t, i.e.

$$\lim_{h \downarrow 0} \frac{1}{h} P[\tau \in (t, t+h] | \tau > t] = \lambda.$$

In probability theory, people call λ hazard rate and it is the density $f(t)$ over survival probability

$$\lambda = f(t)/(1 - F(t)).$$

Survival probability is defined by $S(t) = 1 - F(t)$. Figures 2.4 and 2.5 depicts default probability and survival probability respectively for different constant λ.

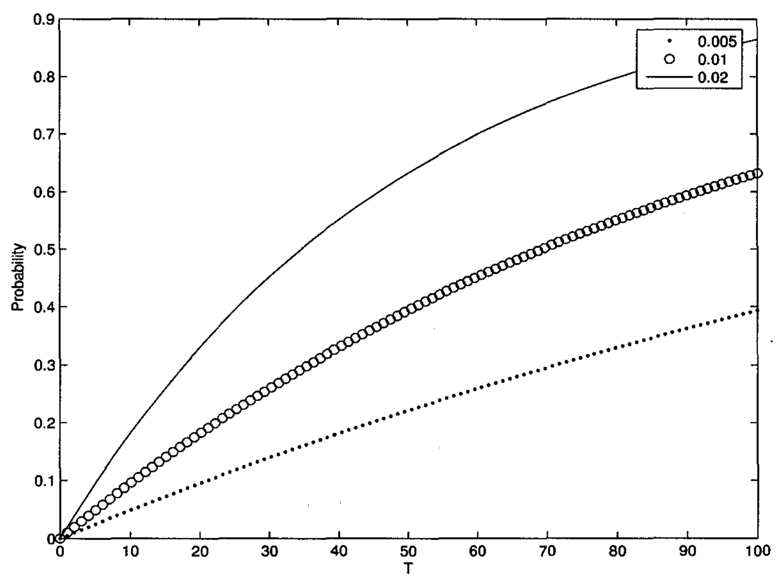

Figure 2.4: Default Probabilities with Different Constant λ

Suppose $\lambda = \lambda(t)$ is a deterministic function of time t. Then $N(t)$ is an inhomogeneous Poisson process with intensity function $\lambda(t)$. Therefore the default probability is

$$F(t) = 1 - e^{-\int_0^T \lambda(u) du}$$

Empirically, people use parametric intensity model as

$$\lambda(t) = h_i \quad t \in [T_{i-1}, T_i), \quad i = 1, 2, ...$$

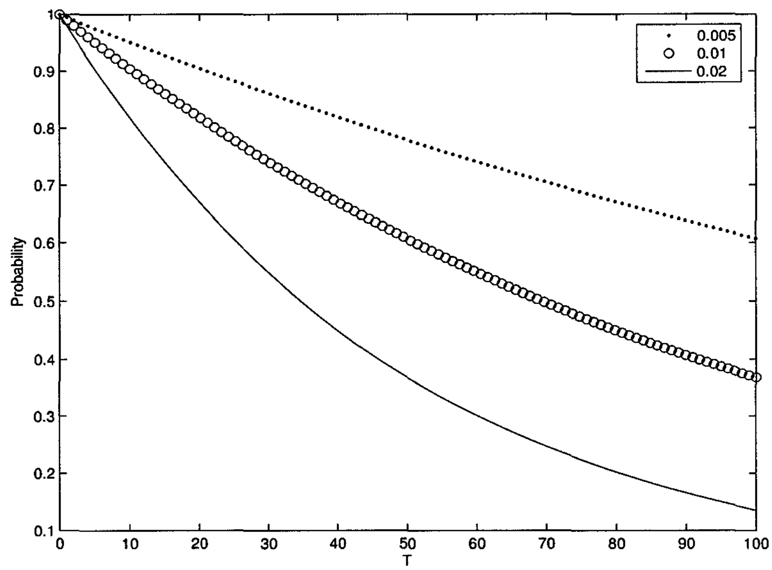

Figure 2.5: Survival Probabilities with Different Constant λ

for constants h_i and T_i, which can be calibrated from market data.

In general, the hazard rate λ is not a constant either deterministic function, but a stochastic variable itself. In this case, this Poisson process is called *doubly stochastic* or *Cox* process with parameter λ_t, which is stochastic. Conditional on the realization of λ_t, it becomes an in-homogeneous Possion.

A Cox process $N(t)$ with intensity λ_t, for $t \geq 0$, is a generalization of the inhomogeneous Poisson process in which the intensity is allowed to be a stochastic process itself such that conditional on the realization of the intensity. Lando (1998) studied credit risky securities on Cox processes and built up the pricing blocks.

The conditional and unconditional default probabilities are given by

$$F_\lambda(t) = P[\tau \leq t \mid \lambda] = 1 - e^{-\int_o^t \lambda(u)du}$$

$$F(t) = P[\tau \leq t] = 1 - E[e^{-\int_o^t \lambda(u)du}].$$

Let L_t denote the loss fraction in market value if the firm were to default at time

t. Duffie & Singleton (1999) shows that the defaultable zero-coupon bond $P_j(t,T)$ can be priced as a default-free bond using the adjusted default-free interest rate with intensity. This is given in the fomula

$$P_j(t,T) = E_t^Q[e^{-\int_0^T R(u)du}], \quad R_j(u) = r(u) + L\hat{\lambda}_j(u)$$

where the expectation is taken under the risk-neutral measure Q given the information up to time t and the $\hat{\lambda}_j(u)$ denotes the hazard rate under risk-enutral measure.

Different models of the hazard rates or default intensities will produce different results from the above pricing formula. Particularly, we are interested in those which could give closed form formulae for the defaultable bond prices, one of which is affine term structure models. Intensity based modeling will not be discussed in detail until the next chapter.

Since generally the defaultalbe bond pricing formulae are very complex, it is impossible to analyse the credit spreads in a generalized fashion. Here, we simply give an example to illustrate how the credit spreads in reduced form approach differ from those in structure approach.

Example 1. Let's consider firm j with constant hazard rate λ. Assume its risk-neutral hazard rate is $\hat{\lambda}$. Default free interest rate is assumed to be a constant r which is positive. A defaultable zero coupon bond with maturity T issued by firm j with zero recovery rate is priced at time 0 through

$$P_j(0,T) = E^Q[e^{-rT} 1_{\{\tau > T\}}] = e^{-rT} Q[\tau > T] = e^{-(r+\hat{\lambda})T}.$$

That is, the defaultable bond could be priced using the default-adjusted discounting rate $r + \hat{\lambda}$ instead of discounting with the risk-free interest rate r.

The term structure of the credit spread at time 0 is thus given by

$$S_j(0,T) = -\frac{1}{T} \ln \frac{e^{-(r+\hat{\lambda})T}}{e^{-rT}} = \hat{\lambda}.$$

This tells us that the credit spread is therefore given by the risk-neutral intensity $\hat{\lambda}$. The term stucture of credit spread is flat with a constant intensity. This is quite different from the term structure depicted in Figure 2.2 and Figure 2.3.

We consider n firms with respective intensities $\lambda_1, \lambda_2, ..., \lambda_n$ forming a multivariate Cox process driven by some state process which includes both systematic and idiosyncratic economic factors driving the credit risk of firms.

The joint survival probability is given by

$$P[\tau_1 > T_1, ..., \tau_n > T_n] = E^Q[e^{-\sum_i A_{T_i}}].$$

where $A_t = \int_0^t \lambda_s ds$ is the upward trend of the increasing default process N_t, and it is also called compensator, as the difference of N_t and A_t is a martingale. In incomplete information approach, Giesecke(2001), Duffie & Lando (2001) gives more general reduced form formula in terms of trend.

Incomplete Information Approach is firstly studied by Duffie and Lando (2001), Giesecke (2001) and Cetin, Jarrow, Protter and Yildirim (2002). The mathematical fundation for this approach is Doob-Meyer decomposition theory. We will not discuss in detail here.

Chapter 3

Intensity Based Models

In previous two chapters, we have talked about the theory of interest rates and the theory of credit risk. As the title of this thesis indicates, from now on, we are going to talk about practical issues in credit risk management. There are two aspects with which we need to concern. The first is modeling. Models are the bridge of the relationships between theoretical and empirical study of this field. Theoretical modeling is based on practical issues arising from the market. Available information from the market needs to be taken care of when we start to set up the model. Although you will find that it is impossible to incorporate all the ingredients required by the real situation, people try to improve their models so as to satisfy as much requirements as possible. The second aspect is the calibration of the model, which is saved for the next chapter.

In this chapter, we first talk about *Two-Factor Gaussian* model under reduced form framwork to clarify the calculations in this kind of modeling. Then, we will talk about *Multi-Factor Gaussian* models and *Multi-Factor CIR*. Examples are given respectively to illustrate how to improve the models.

3.1 Two-Factor Gaussian Model

In this section, we study the two factor Gaussian hazard rate model for two companies. Our aim is to find the joint survival distribution of the two companies. Firstly, we study the case with constant coefficients without considering default-free interest rate. Then, we extend the first case to a model with coefficients which are deterministic functions. Finally, we set up a model with consideration of the default-free interest rate. In the previous chapter, since we start with Poisson distribution, we use λ to denote the hazard rate. But here we use h as the hazard rate notation which is more widely used in the literature.

3.1.1 Model with Constant Coefficents

Assume the hazard rates of firm 1 and firm 2 follow Vasicek model

$$dh_t^i = (\alpha_i - \beta_i h_t^i)dt + \sigma_i(\rho_i dW + \sqrt{1-\rho_i^2} dZ_i) \quad i = 1,\, 2 \tag{3.1}$$

Where α_i, β_i, σ_i, ρ_i are constants, and α_i, β_i, σ_i are positive while ρ_i could be negative; dW is systematic shock (standard Brownian motion) while dZ_i (standard Brownian motion) are idiosyncratic shocks. We assume that dW, dZ_1, dZ_2 are independent.

If we define $dW_i = \rho_i dW + \sqrt{1-\rho_i^2} dZ_i$, then $dW_i dW_i = dt$, so W_i are Brownian motions. Therefore, equation (3.1) could be rewriten as

$$dh_t^i = (\alpha_i - \beta_i h_t^i)dt + \sigma_i dW_i \quad i = 1,\, 2 \tag{3.2}$$

with correlation relationships

$$dW_1 dW_2 = \rho_1 \rho_2 dt$$

. We know that the solution of this SDE is given by

$$h_t^i = e^{-\beta_i t}\left[h_0^i + \frac{\alpha_i}{\beta_i}(e^{\beta_i t} - 1) + \int_0^t e^{\beta_i u}\sigma_i dW_i(u)\right] \quad i = 1, 2 \tag{3.3}$$

The mean of h_t^i is given by

$$E[h_t^i] = \frac{\alpha_i}{\beta_i} + (h_0^i - \frac{\alpha_i}{\beta_i})e^{-\beta_i t} \quad i = 1, 2 \tag{3.4}$$

Notice that when $t = 0$, it is just h_0^i , when $t \to +\infty$, it tends to $\frac{\alpha_i}{\beta_i}$. This property is called **mean-reverting**.

The variance of h_t^i is given by

$$Var[h_t^i] = \frac{\sigma_i^2}{2\beta_i}(1 - e^{-2\beta_i t}) \quad i = 1, 2 \tag{3.5}$$

Notice that when $t = 0$, it is just 0 , when $t \to +\infty$, it tends to $\frac{\sigma_i^2}{2\beta_i}$.

Let $A_i(T) = \int_0^T h_t^i dt$. In order to study $A_i(T)$, we define

$$X_i(t) = \int_0^t e^{\beta_i u}\sigma_i dW_i(u), \quad Y_i(T) = \int_0^T e^{-\beta_i t}X_i(t)dt$$

Integrating equation (3.3) from 0 to T, we obtain

$$A_i(T) = \int_0^T e^{-\beta_i t}\left[h_0^i + \frac{\alpha_i}{\beta_i}(e^{\beta_i t} - 1)\right]dt + Y_i(T) \quad i = 1, 2 \tag{3.6}$$

It is easy to show that $Y_i(T)$ is Gaussian with mean 0. Hence $A_i(T)$ is a Gaussian too with mean and variance given by

$$E[A_i(T)] = \frac{\alpha_i}{\beta_i}T + \frac{1}{\beta_i}(h_0^i - \frac{\alpha_i}{\beta_i})(1 - e^{-\beta_i T}) \quad i = 1, 2 \tag{3.7}$$

$$Var[A_i(T)] = \frac{\sigma_i^2}{\beta_i^2}\left[T + \frac{1}{2\beta_i}(1 - e^{-2\beta_i T}) - \frac{2}{\beta_i}(1 - e^{-\beta_i T})\right] \quad i = 1, 2 \tag{3.8}$$

Notice that when $T = 0$, the mean is 0 and the variance is 0; when $T \to +\infty$, the mean tends to infinity and the variance tends to infinity too.

From equation (3.6) and (3.7), we have the following equation

$$A_i(T) = E[A_i(T)] + Y_i(T) \quad i = 1,\, 2 \tag{3.9}$$

It follows that $Cov[A_1(T_1), A_2(T_2)] = E[Y_1(T_1)Y_2(T_2)]$. Without loss of generality, we assume that $T_1 < T_2$

$$\begin{aligned}
E[Y_1(T_1)Y_2(T_2)] &= \int_0^{T_1}\int_0^{T_2} e^{-\beta_1 t-\beta_2 s} E[X_1(t)X_2(s)]dsdt \\
&= \int_0^{T_1}\int_0^{T_2} e^{-\beta_1 t-\beta_2 s}\left(\int_0^{t\wedge s} \sigma_1\sigma_2\rho_1\rho_2 e^{(\beta_1+\beta_2)u}du\right) dsdt \\
&= \frac{\sigma_1\sigma_2\rho_1\rho_2}{\beta_1+\beta_2}\int_0^{T_1}\int_0^{T_2} e^{-\beta_1 t-\beta_2 s}(e^{(t\wedge s)(\beta_1+\beta_2)}-1)dsdt \\
&= \frac{\sigma_1\sigma_2\rho_1\rho_2}{\beta_1+\beta_2}\left[\int_0^{T_1}\int_0^{t} e^{-\beta_1 t-\beta_2 s}(e^{s(\beta_1+\beta_2)}-1)dsdt\right. \\
&\quad \left. +\int_0^{T_1}\int_t^{T_2} e^{-\beta_1 t-\beta_2 s}(e^{t(\beta_1+\beta_2)}-1)dsdt\right] \\
&= \frac{\sigma_1\sigma_2\rho_1\rho_2}{(\beta_1+\beta_2)\beta_1^2\beta_2^2}\left[\beta_1\beta_2(\beta_1+\beta_2)T_1 - \beta_2(\beta_1+\beta_2) + \beta_2(\beta_1+\beta_2)e^{-\beta_1 T_1}\right. \\
&\quad \left. +\beta_1(\beta_1+\beta_2)e^{-\beta_2 T_2} - \beta_1^2 e^{-\beta_2(T_2-T_1)} - \beta_1\beta_2 e^{-\beta_1 T_1-\beta_2 T_2}\right].
\end{aligned}$$

The next important ingredient we need to model is the default time τ. Following Schonbucher (2003), we assume the default trigger variables v_i are independent exponentially distributed random variables with parameter 1. Furthermore, we assume v_i are independent of $A_i(T_i)$.

Define default time τ_i

$$\tau_i = inf\,\{T_i \,:\, A_i(T_i) \geq v_i\} \quad i = 1,\, 2 \tag{3.10}$$

Then, the conditional and unconditional marginal survival probabilities are given by

$$P[\tau_i > T_i \mid h^i_{T_i}] = \exp[-A_i(T_i)] \quad i = 1,\, 2 \tag{3.11}$$

$$P[\tau_i > T_i] = E[\exp[-A_i(T_i)]] \quad i = 1,\, 2 \tag{3.12}$$

Next, we want to calculate the joint survival distribution function. Suppose the joint distribution density function of $(A_1(T_1), A_2(T_2))$ is $f(x, y)$. Since $A_1(T_1), A_2(T_2)$ are normal random variables, the joint density function $f(x, y)$ is determined by its mean vector and the covariance matrix which we have already calculated. In fact, $f(x, y)$ is given by

$$f(x,y) = \frac{\exp\left\{-\frac{1}{2(1-\rho^2)}\left[\frac{(x-a_1)^2}{\delta_1^2} - 2\rho\frac{(x-a_1)(y-a_2)}{\delta_1\delta_2} + \frac{(y-a_2)^2}{\delta_2^2}\right]\right\}}{2\pi\delta_1\delta_2\sqrt{1-\rho^2}}$$

where a_1, a_2, δ_1, δ_2, ρ are given by (refer to equations (3.7), (3.8), (3.9))

$$a_1 = E[A_1(T_1)] \qquad a_2 = E[A_2(T_2)]$$

$$\delta_1 = \sqrt{Var(A_1(T_1))} \qquad \delta_2 = \sqrt{Var(A_2(T_2))}$$

$$\rho = Cov[A_1(T_1), A_2(T_2)]/\sqrt{Var(A_1(T_1))Var(A_2(T_2))}$$

Let $\xi = (x - a_1)/\delta_1$, $\zeta = (y - a_2)/\delta_2$; then ξ and ζ are correlated (with ρ) unit normal random variables. We simplify $f(x, y)dxdy$ into

$$\begin{aligned} f(x,y)dxdy &= f(\delta_1\xi + a_1, \delta_2\zeta + a_2)\delta_1\delta_2 d\xi d\zeta \\ &= \frac{1}{2\pi\sqrt{1-\rho^2}}\exp\left\{-\frac{1}{2(1-\rho^2)}\left[\xi^2 - 2\rho\xi\zeta + \zeta^2\right]\right\} d\xi d\zeta \\ &\equiv g(\xi,\zeta)d\xi d\zeta \end{aligned}$$

Mathematically, we calculate the joint survival probability

$$\begin{aligned} P[\tau_1 > T_1, \tau_2 > T_2] &= E[1_{(\tau_1 > T_1)}1_{(\tau_2 > T_2)}] \\ &= E[1_{(A_1(T_1) < v_1)}1_{(A_2(T_2) < v_2)}] \\ &= \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}\left(\int_{0\vee x}^{+\infty}\int_{0\vee y}^{+\infty} e^{-(v_1+v_2)}dv_1 dv_2\right) f(x,y)dxdy \\ &= \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} e^{-(0\vee x + 0\vee y)} f(x,y)dxdy \\ &= \int_{-\infty}^{0}\int_{-\infty}^{0} f(x,y)dxdy + \int_{-\infty}^{0}\int_{0}^{+\infty} e^{-x} f(x,y)dxdy \\ &\quad + \int_{0}^{+\infty}\int_{-\infty}^{0} e^{-y} f(x,y)dxdy + \int_{0}^{+\infty}\int_{0}^{+\infty} e^{-(x+y)} f(x,y)dxdy \end{aligned}$$

$$
\begin{aligned}
= & \int_{-\infty}^{-a_2/\delta_2}\int_{-\infty}^{-a_1/\delta_1} g(\xi,\zeta)d\xi d\zeta + \int_{-\infty}^{-a_2/\delta_2}\int_{-a_1/\delta_1}^{+\infty} e^{-(\delta_1\xi+a_1)}g(\xi,\zeta)d\xi d\zeta \\
& + \int_{-a_2/\delta_2}^{+\infty}\int_{-\infty}^{-a_1/\delta_1} e^{-(\delta_2\zeta+a_2)}g(\xi,\zeta)d\xi d\zeta \\
& + \int_{-a_2/\delta_2}^{+\infty}\int_{-a_1/\delta_1}^{+\infty} e^{-(\delta_1\xi+a_1+\delta_2\zeta+a_2)}g(\xi,\zeta)d\xi d\zeta
\end{aligned}
$$

Practically, we can simply use the last term $\int_0^{+\infty}\int_0^{+\infty} e^{-(x+y)}f(x,y)dxdy$ to approximate the joint survival probability. Since this probability highly depends on the initial value h_0^i, given rightly chosen positive h_0^i, this approximation should be quite accurate. Hazard rate could be negative from mathematical point of view in Hull-White model, which is not quite applausable from financial point of view. Considering this, CIR model is better than Hull-White model.

This approximation enables us to compute the joint survival distribution function in a closed form

$$
\begin{aligned}
P[\tau_1 > T_1, \tau_2 > T_2] &= E[\exp(-A_1 - A_2)] \\
&= \exp(-E[A_1] - E[A_2] + Var[A_1 + A_2]/2) \\
&= \exp(-E[A_1] - E[A_2] + Var[A_1]/2 \\
&\quad +Var[A_2]/2 + Cov[A_1, A_2]) \\
&= \exp\Big\{ -\frac{\alpha_1}{\beta_1}T_1 - \frac{1}{\beta_1}(h_0^1 - \frac{\alpha_1}{\beta_1})(1 - e^{-\beta_1 T_1}) \\
&\quad -\frac{\alpha_2}{\beta_2}T_2 - \frac{1}{\beta_2}(h_0^2 - \frac{\alpha_2}{\beta_2})(1 - e^{-\beta_2 T_2}) \\
&\quad +\frac{\sigma_1^2}{2\beta_1^2}\left[T_1 + \frac{1}{2\beta_1}(1 - e^{-2\beta_1 T_1}) - \frac{2}{\beta_1}(1 - e^{-\beta_1 T_1})\right] \\
&\quad +\frac{\sigma_2^2}{2\beta_2^2}\left[T_2 + \frac{1}{2\beta_2}(1 - e^{-2\beta_2 T_2}) - \frac{2}{\beta_2}(1 - e^{-\beta_2 T_2})\right] \\
&\quad +\frac{\sigma_1\sigma_2\rho_1\rho_2}{(\beta_1+\beta_2)\beta_1^2\beta_2^2}\left[\beta_1\beta_2(\beta_1+\beta_2)T_1 - \beta_2(\beta_1+\beta_2) + \beta_2(\beta_1+\beta_2)e^{-\beta_1 T_1}\right. \\
&\quad \left. +\beta_1(\beta_1+\beta_2)e^{-\beta_2 T_2} - \beta_1^2 e^{-\beta_2(T_2-T_1)} - \beta_1\beta_2 e^{-\beta_1 T_1 - \beta_2 T_2}\right]\Big\}
\end{aligned}
$$

Unfortunately, this messy formula can not be simplified. However if β_i are sufficiently small, we can simplify this formula by approximating $e^{-\beta_i T_i}$ as $(1-\beta_i T_i)$. In this situation

$$E[N_i] \approx h_0^i T_i \quad Var[N_i] \approx 0 \quad Cov[N_1, N_2] \approx 0$$

$$P[\tau_1 > T_1, \tau_2 > T_2] \approx \exp[-h_0^1 T_1 - h_0^2 T_2] = P[\tau_1 > T_1] P[\tau_2 > T_2]$$

This tells us when β_i are sufficiently small, we can simplify regards τ_i as independent variables.

3.1.2 Model with Coeffients of Deterministic Functions

We can extend the above case into a model with coefficients which are deterministic functions in time. Hence the equation (3.1) should be modified into

$$dh_t^i = (\alpha_i(t) - \beta_i(t) h_t^i) dt + \sigma_i(t)(\rho_i dW + \sqrt{1-\rho_i^2} dZ_i) \quad i = 1, 2 \tag{3.13}$$

Where $\alpha_i(t)$, $\beta_i(t)$, $\sigma_i(t)$ are positive functions in time, while ρ_i remains constant.

Define: $K_i(t) = \int_0^t \beta_i(u) du$, then we can do calculations as above. But it will be more complicated to compute those formulae. In the constant case, $K_i(t)$ is just $\beta_i t$ which simplifies the computations. Also, this two-dimentional case can be extended to n-dimentional case.

3.1.3 Model Combining Default-Free Interest Rate

Following Miu (2003), we use Hull-White model to capture the dynamics of default free interest rate r_t . We know that r_t is Gaussian. In the above cases we model h_t^i without considering default-free interest rate. In Miu (2003), the author models h_{it} (here we will use Miu's notation h_{it} to denote h_t^i) as affine

$$h_{it} = a_{ih} + h_{it}^* + b_{ir} r_t \quad i = 1, 2 \tag{3.14}$$

where a_{ih} and b_{ir} are constants, and h_{it}^* is modelled as a constant coefficient Hull-White hazard rate model. The stochastic process driving h_{it}^* is assumed to be independent of the stochastic process driving the default-free interest rate r_t. Here, h_{it} is still Gaussian. Therefore we can do the same calculations for the joint survival probability. But, generally speaking, a closed form can not be obtained.

3.2 Multi-Factor Gaussian Model

In the previous section, we investigate a Gaussian Model with two companies. This can be extended to multi-firm case. The hazard rates of each firm are assumed to follow a Gaussian dynamics with some correlation. Let's consider k firms. The model could be set up in the following way

$$dr_t = (a - br_t)dt + \sigma dW_t$$

$$dh_{it} = (\alpha_i - \beta_i h_{it})dt + \sigma_i dW_i(t), i = 1, 2, ..., k$$

$$dWdW_i = \rho_i dt.$$

Since mostly empirical studies show that the interest rate and the hazard rates are negatively correlated, most of ρ_i should be expected to be negative. As a result, the correlations between two hazard rates of two firms are likely to be positive. However, empirical studies also implies that negaitve correlations between hazard rates do exist.

Also, the interest rate and the hazard rates could be negative due to the Gaussian dynamics assumed in the model. In short, this model has its shortcomings as well as advantages which are summarized as follows.

The shortcomings are: negative interest rates and hazard rates allowed in the model, and only positive correlations allowed between default intensities.

The advantages are: analytical tractability because of the closed-form solutions for bond prices, negative correlation between interest rate and hazard rates are satisfied, the model can be calibrated to full term structures of bond prices and spreads.

3.2.1 An Example: Miu (2003)

In this section, we give an example of multi-factor Gaussian model studied by Miu (2003). In his paper, the default-free interet rate r_t is assumed to follow Hull-White model. The hazard rates $h_i(t)$ for each firm i are decomposed into three parts: a constant α_{ih}, an r_t independent term h_{it}^* and an r_t dependent term $\beta_{ir}r_t$. Finally, we assume h_{it}^* follows mean-reverting process and we decompose default randomness into two resources: one comes from the market (systematic), the other comes from the firm itself (unsystematic). The model is like this

$$\textit{under physical measure}: \quad dr_t = [\theta_r(t) - \kappa_r r_t]dt + \sigma_r dz_{rt} \tag{3.15}$$

$$\textit{under risk-neutral measure}: dr_t = [\theta_r(t) - \kappa_r r_t - \lambda_r \sigma_r]dt + \sigma_r d\hat{z}_{rt}$$

where λ_r is the market price of default-free interest rate risk, $\theta_r(t)$ is a function so chosen to ensure the fitting of all observed default-free bond prices, and $\hat{z}_{rt}$ is Q-Brownian motion. Hazard rates are modelled as follows

$$h_{it} = \alpha_{ih} + h_{it}^* + \beta_{ir} r_t \tag{3.16}$$

$$\textit{under physical measure}: \quad dh_{it}^* = \kappa_{ih}(\theta_{ih} - h_{it}^*)dt + \sigma_{ih} dz_{iht} \tag{3.17}$$

$$\textit{under Q measure}: \quad dh_{it}^* = \kappa_{ih}(\theta_{ih} - h_{it}^*)dt - \lambda_{ih}\sigma_{ih}dt + \sigma_{ih} d\hat{z}_{iht}$$

where λ_{ih} is the market price of default risk for firm i.

The data chosen for calibration of the model is composed of the US Treasury prices and month-end corporate bond prices. The observation is from January 1993 to December 2001. The US Treasury prices are used to calibrate the interest rate by

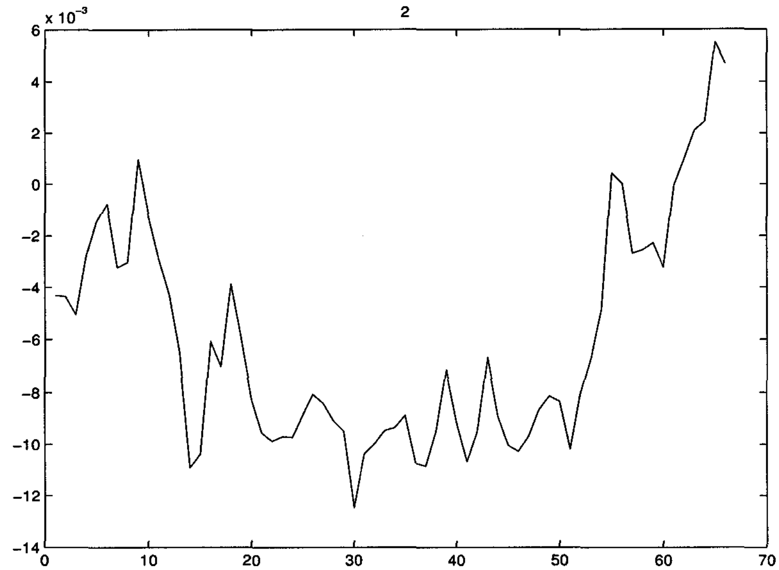

Figure 3.1: The Hazard Rate Filtration of Firm 4 with Aa Rated. Time 0 is January 1993. The unit of X-axis is monthly. The time 108 is December 2001.

polynomial spline specification proposed by Adams and Van Deventer (1994). About 72 bonds issured by 21 firms are observed. These data are used to calibrate the hazard rates for each firm. Figure 3.5 and Figure 3.6 are the filtrations of the hazard rates of those firms obtained by Kalman Filtering. You can see that there are many negative values of the hazard rates in the graphics.

3.3 Multi-Factor CIR Model

For multi-factor Gaussian model, we notice that the interest rate and hazard rates could be negative which is not plausible. However, if we use CIR instead of Vasicek

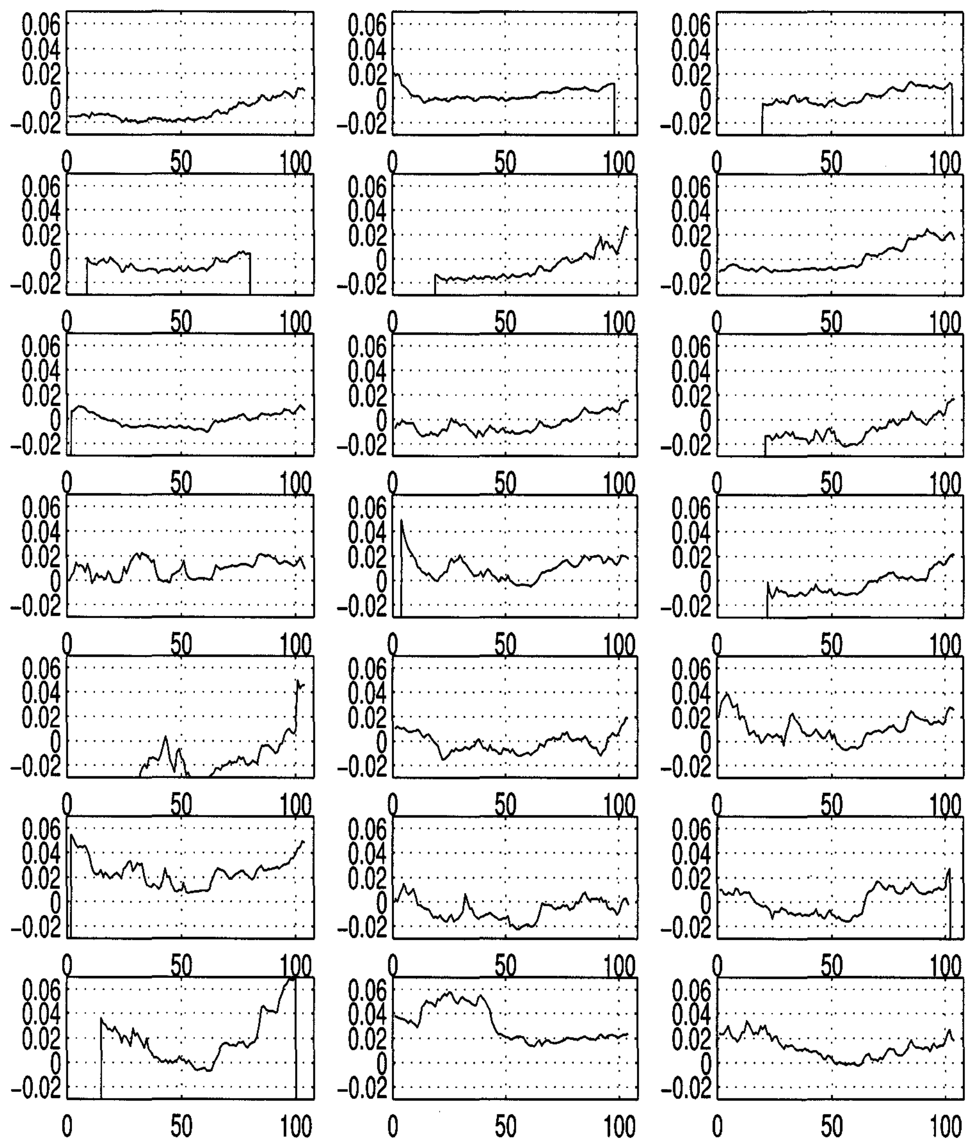

Figure 3.2: The Hazard Rate Filtrations of Firms 1 to 20. Time 0 denotes January 1993. The unit of X-axis is montly. The last observation is December 2001.

as the factors, the negativity problem will be solved easily. The model is like this

$$\begin{aligned} dx_i &= (\alpha_i - \beta_i x_i)dt + \sigma_i \sqrt{x_i} dW_i, \\ &\quad i = 1, 2, ..., n, \ \ \alpha_i > \sigma_i^2/2, \\ r(t) &= \sum_{i=1}^{n} a_i x_i(t), \ \ a_i \geq 0, \\ h_j(t) &= \sum_{i=1}^{n_j} b_{i,j} x_i(t), \ \ b_{i,j} \geq 0. \end{aligned}$$

where W_i are independent.

Although, the positivity of the interest rate and hazard rates is satisfied in the model, some shortcomings also arise as a by-product. The correlations between the interest rate and the hazard rates are all positive. And still there are no negative correlations between the hazard rates of different firms.

Since it has closed-form solutions too for the bond prices, it can be calibrated by fitting the market data. In fact, this multi-factor framework is not only restricted for Vasicek or CIR factors, we can use some other affine factors, which will still produce closed form solutions for the bond prices.

Some specific models, however, showed by the calibration that some of the coefficients a_i and $b_{i,j}$ are negative. The following example is studied by Duffee (1999).

3.3.1 An Example: Duffee (1999)

In Duffee (1999), interest rate is assumed to follow two-factor CIR, while the hazard rates follow three-factor CIR model.

$$
\begin{aligned}
r_t &= \alpha_r + s_{1,t} + s_{2,t}, \\
ds_{i,t} &= \kappa_i(\theta_i - s_{i,t})dt + \sigma_i\sqrt{s_{i,t}}dZ_{i,t}, \;\; i = 1, 2. \\
h_{j,t} &= \alpha_j + h^*_{j,t} + \beta_{1,j}(s_{1,t} - \bar{s}_{1,t}) + \beta_{2,j}(s_{2,t} - \bar{s}_{2,t}) \\
dh^*_{j,t} &= \kappa_j(\theta_j - h^*_{j,t})dt + \sigma_j\sqrt{h^*_{j,t}}dW_{j,t}
\end{aligned}
$$

where $\bar{s}_{i,t}$ are the averages of the time serious of $s_{i,t}$ respectively, and $dZ_{1,t}$, $dZ_{2,t}$ and $dW_{j,t}$ are independent, and α, κ, θ, β, σ are constants.

The calibration of this model is easy to implement. The interest rate model is calibrated first using the Treasury bills. Then the hazard rates model for each firm are calibrated seperately using the corporate bond prices of each firms.

Empirical study by Duffee (1999) shows that $\beta_{1,j}$ and $\beta_{2,j}$ are all negative. Theoretically speaking this may cause producing negative hazard rates $h_{j,t}$. However, the firm specific shocks $h^*_{j,t}$ are comparatively larger than the negative terms in the hazard rates model. The interest rate and hazard rates are thus negatively correlated, because their correlations, if any, are entirely captured by the $\beta_{1,j}$ and $\beta_{2,j}$.

Chapter 4

Calibration: Extended Kalman Filter with Quasi-Maximum Likelihood Estimation

For affine term structure models, there are many numerical techniques available in the literature to implement the calibration. Here we use the Extended Kalman Filter with Quasi-Maximum Likelihood Estimation proposed in Duan & Simonato (1995) to illustrate how a calibration is conducted. In order to make the text easy to read, the model we choose is very simple with one factor CIR as the interest rate, while two factor CIRs as the hazard rates. In fact, the calibration method explained below is also effective when we extend our model by adding more CIR common factors shared both by the interest rate and the hazard rates, such as Duffee (1999).

The structure of this chapter is organized as follows. We first introduce the model and study the term stucture of the model. Then, the calibration method is explained. Time series of the interest rate is obtained by Standard Kalman Filter while the time series of the hazard rates are obtained by Extended Kalman Filter. The coeffiecients of the model are obtained by Quasi-Maximum Likelihood Estimation with Extended Kalman Filter. Some empirical results are given in the end and Monte Carlo Simulation is also analysed.

4.1 The Model

Here, I follow Miu (2003) and Duffee (1999) by using one CIR factor to model the default-free interest rate while using two CIR factors to capture the dynamics of the hazard rates for each firm.

$$r_t = \alpha_r + s_t$$

$$h_{j,t} = \alpha_{j,h} + s_{j,t} + \beta_j s_t$$

where α_r, $\alpha_{j,h}$ and β_j are constants while s_t and $s_{j,t}$ follow CIR, i.e.

$$ds_t = \kappa(\theta - s_t)dt + \sigma\sqrt{s_t}dZ_t$$

$$ds_{j,t} = \kappa_j(\theta_j - s_{j,t})dt + \sigma_j\sqrt{s_{j,t}}dZ_{j,t}$$

dZ_t, $dZ_{j,t}$ and $dZ_{i,t}$ are independent for $i \neq j$.

Under the equivalent martingale measure, these SDEs can be written as

$$ds_t = (\kappa\theta - (\kappa + \lambda)s_t)dt + \sigma\sqrt{s_t}d\hat{Z}_t$$

$$ds_{j,t} = (\kappa_j\theta_j - (\kappa_j + \lambda_j)s_{j,t})dt + \sigma_j\sqrt{s_{j,t}}d\hat{Z}_{j,t}$$

where $d\hat{Z}_t$, $d\hat{Z}_{j,t}$ and $d\hat{Z}_{i,t}$ are independent for $i \neq j$; λ and λ_j are the market price of the risks.

4.2 Term-Structure of The Model

For zero-coupon bonds, this model has closed forms of pricing formula, which has exponential affine term structure. For details of the derivation, please refer to Pearson and Sun (1994) or Duffie (1996). Here we simply present the results.

4.2.1 Default-Free Zero-Coupon Bonds

The pricing formula of default-free zero-coupon bond $P(t,T)$ which pays 1$ at time T can be represented as

$$P(t,T) = A(t,T)e^{-B(t,T)s_t-(T-t)\alpha_r}$$

where

$$A(t,T) = \left\{ \frac{2\gamma \exp[(T-t)(\kappa+\lambda+\gamma)/2]}{(\kappa+\lambda+\gamma)(\exp[\gamma(T-t)]-1)+2\gamma} \right\}^{2\kappa\theta/\sigma^2}$$

$$B(t,T) = \frac{2(\exp[\gamma(T-t)]-1)}{(\kappa+\lambda+\gamma)(\exp[\gamma(T-t)]-1)+2\gamma}$$

$$\gamma = \sqrt{(\kappa+\lambda)^2+2\sigma^2}$$

The yield to maturity of this default-free zero-coupon bond is defined as

$$Y(P,t,T) = -\frac{\ln P(t,T)}{T-t}$$

It turns out that the yield is linear in the underlying factor s_t

$$Y(P,t,T) = \frac{B(t,T)s_t}{T-t} + (\alpha_r - \frac{\ln A(t,T)}{T-t}).$$

4.2.2 Defaultable Zero-Coupon Bonds

We follow Duffie & Singleton (1999) framework and suppose that the loss rate is the same constant L for all firms. Duffee (1999) chose 0.56 as the loss rate in line with Moody's evidence for the data he used, while Miu (2003) selected 0.5116 according to Moody's (2000). Since the corporate bonds data I use here is the same as Miu's (2003) data, I also use 0.5116 as the loss rate. Thus, the pricing formula for defaultable zero-coupon bond of firm j could be represented as

$$P_j(t,T) = E_t^Q[\exp(-\int_t^T r_\zeta + h_{j,\zeta} L d\zeta)]$$

where

$$r_t + h_{j,t}L = a_j + b_j s_t + L s_{j,t}$$

$$a_j = \alpha_r + \alpha_{j,h} L$$

$$b_j = 1 + \beta_j L$$

Both $b_j s_t$ and $L s_{j,t}$ are CIRs and they are independent because of the independence of s_t and $s_{j,t}$. It follows from the multi-factor CIR interest rate model theory that the defaultable bond price $P_j(t,T)$ has closed form

$$P_j(t,T) = A_{1,j} A_{2,j} \exp[-B_{1,j} b_j s_t - B_{2,j} L s_{j,t} - (T-t)a_j]$$

$$A_{1,j} = \left\{ \frac{2\gamma_{1,j} \exp[(T-t)(\kappa b_j + \lambda + \gamma_{1,j})/2]}{(\kappa b_j + \lambda + \gamma_{1,j})(\exp[\gamma_{1,j}(T-t)] - 1) + 2\gamma_{1,j}} \right\}^{2\kappa b_j \theta/(\sigma b_j)^2}$$

$$B_{1,j} = \frac{2(\exp[\gamma_{1,j}(T-t)] - 1)}{(\kappa b_j + \lambda + \gamma_{1,j})(\exp[\gamma_{1,j}(T-t)] - 1) + 2\gamma_{1,j}}$$

$$\gamma_{1,j} = \sqrt{(\kappa b_j + \lambda)^2 + 2(\sigma b_j)^2}$$

$$A_{2,j} = \left\{ \frac{2\gamma_{2,j} \exp[(T-t)(\kappa_j L + \lambda_j + \gamma_{2,j})/2]}{(\kappa_j L + \lambda_j + \gamma_{2,j})(\exp[\gamma_{2,j}(T-t)] - 1) + 2\gamma_{2,j}} \right\}^{2\kappa_j L \theta_j/(\sigma_j L)^2}$$

$$B_{2,j} = \frac{2(\exp[\gamma_{2,j}(T-t)] - 1)}{(\kappa_j L + \lambda_j + \gamma_{2,j})(\exp[\gamma_{2,j}(T-t)] - 1) + 2\gamma_{2,j}}$$

$$\gamma_{2,j} = \sqrt{(\kappa_j L + \lambda_j)^2 + 2(\sigma_j L)^2}$$

The yield to maturity of this defaultable zero-coupon bond $Y_j(P_j, t, T)$ is thus linear in s_t and $s_{j,t}$

$$Y_j(P_j, t, T) = \frac{B_{1,j} b_j s_t + B_{2,j} L s_{j,t}}{T - t} + (a_j - \frac{\ln A_{1,j} + \ln A_{2,j}}{T - t}).$$

4.2.3 Defaultable Coupon Bonds

Let $G_j(t, T)$ denote the price at time t of the defaultable coupon bond issued by firm j bearing coupon c_j which is assumed to be the same at different date of payment. We further suppose that there are n remaining number of coupon payments. Therefore, the defaultable coupon bond price $G_j(t, T)$ could be represented by linear combination of defaultable zero-coupon bond prices $P_j(t, T_k)$ of different maturities T_k that are the coming payment date of the remaining coupons.

$$G_j(t, T_n) = \sum_{k=1}^{n} c_j P_j(t, T_k) + 100 P_j(t, T_n)$$

In this situation, the yield of the defaultable coupon bond is not linear in the CIR factors any more. However, if we employ a first-order Taylor approximation of the right hand side of the above equation, then, $G_j(t, T_n)$ itself is linear in the state variate s_t. We will discuss it in detail in the section 4.7.

4.3 State Space Models and Kalman Filter

Kalman Filter is applied on state space models which are very powerful tools for handling a wide range of time series models.

4.3.1 The State Space Form

The state space form consists of two main equations which are called Transition Equation and Measurement Equation respectively. Transition equation captures the dynamics in time of the state vector x_k (here we use k to denote $t = k$) which is unobservable. Measurement equation builds the relationship between the unobservable state vector x_k and the observable measurement y_k. The general state space form applies to multivariate time series and it could be non-linear. For Standard Kalman Filter (SKF), these equations are linear in state vector x_k.

The state space model of the SKF is

$$\begin{cases} x_{k+1} &= F_{k+1,k}x_k + c_k + \omega_k \\ y_k &= H_k x_k + d_k + \nu_k \end{cases} \tag{4.1}$$

where the state vector x_{k+1} is K dimentional if the model has K factors, $F_{k+1,k}$ is a $K \times K$ transition matrix, c_k is a K dimentional array; where the observation y_k is an N dimentional array, H_k is $N \times K$ matrix, d_k is also an N dimentional array; where ω_k and ν_k are independent Gaussian noise with mean zero and covariance matrix Q_k and R_k respectively.

The first equation is called the **transition equation** which describes how the unobserved state variable x_k change dynamicly. The second equation is called the **measurement equation** which builds up the relationship between the unobserved state variable x_k and the observation y_k.

4.3.2 Standard Kalman Filter

Given the state space form above, our objective is to use the entire observation of $y_1, y_2, ..., y_k$ up to time k to optimally estimate the state vector x_k in the sense of minimum mean square error. Kalman filter is implemented in two steps. The first step is prediciton in which we use the information $y_1, y_2, ..., y_{k-1}$ to get the best estimate of x_k. Let $\hat{x}_k^-$ denote this best estimate, which is called prior estimate. The next step is to use the lately information observed in y_k to update the first estimate $\hat{x}_k^-$. Let $\hat{x}_k$ denote this optimal estimate, which is called posteriori estimate. From the state space form, we have that

$$\hat{x}_k^- = F_{k,k-1}\hat{x}_{k-1} + c_{k-1} \tag{4.2}$$

Correspondingly, the prior covariance matrix P_k^- and posterior covariance matrix P_k are defined by

$$P_k^- = E[(x_k - \hat{x}_k^-)(x_k - \hat{x}_k^-)^T]$$

$$P_k = E[(x_k - \hat{x}_k)(x_k - \hat{x}_k)^T]$$

The error covariance propagation from P_{k-1} to P_k^- is given by

$$P_k^- = F_{k,k-1}P_{k-1}F_{k,k-1}^T + Q_{k-1} \tag{4.3}$$

Equation (4.2) and Equation (4.3) are called prediction equations. Once the new information y_k is observed, the prior estimate of x_t could be updated through the following updating equations

$$G_k = P_k^- H_k^T[H_k P_k^- H_k^T + R_k]^{-1} \tag{4.4}$$

$$\hat{x}_k = \hat{x}_k^- + G_k(y_k - H_k\hat{x}_k^- - d_k) \tag{4.5}$$

$$P_k = (I - G_k H_k)P_k^- \tag{4.6}$$

Equation (4.4) is called Kalman Gain Matrix. Equation (4.5) is called state estimate update equation, while Equation (4.6) is called error covariance update equation.

We need to initialize this algorithm to get it started. Without any observation of y, the best estimate of x_0 at time $k = 0$ is chosen to be its unbiased mean

$$\hat{X}_0 = E[x_0]$$

$$P_0 = E[(x_0 - E[x_0])(x_0 - E[x_0])^T]$$

For the details of derivation of the SKF, please refer to A. Harvey, 1989, *Forecasting, Structural time series models and the Kalman filter;* S. Haykin, *2001, Kalman Filtering and Neural Networks.*

4.3.3 Extended Kalman Filter

For SKF, the measurement equation and transition equation are both linear in the state vector x_k. However, in most cases the measurement equation is nonlinear which we may extend the classical SKF through a linearization procedure. This linearized filter is referred as the Extended Kalman Filter (EKF). For general EKF, both the measurement equation and the transition equation could be nonlinear. Since the situation we have here (and most situations we have) is that nonlinearality only happens in the measurement equation, we will derive the EKF in this particular case.

The state space model is

$$\begin{cases} x_{k+1} &= F_{k+1,k}x_k + c_k + \omega_k \\ y_k &= \varphi(k, x_k) + \nu_k. \end{cases} \tag{4.7}$$

where φ is a nonlinear function in x_k and we also require that φ is differentiable with respect to x. The other assumptions needed are the same as in SKF.

The first order derivative of φ in x is constructed as

$$H_k = \frac{\partial \varphi(k, x_k)}{\partial x} |_{x=\hat{x}_k^-}$$

The nonlinear function φ is linearized through first order Taylor expansion around $\hat{x}_k^-$

$$\varphi(k, x_k) \cong \varphi(k, \hat{x}_k^-) + H_k(x_k - \hat{x}_k^-)$$

Substituting this approximation into the measurement equation, we obtain

$$y_k \cong H_k x_k + d_k + v_k$$

where

$$d_k = \varphi(k, \hat{x}_k^-) - H_k \hat{x}_k^-$$

Then, the SKF is implemented.

4.4 Maximum Likelihood Estimation

4.4.1 Likelihood Function

In the classical theory of Maximum Likelihood Estimation (MLE), the T sets of observations, $\{y_1, y_2, ..., y_T\}$, are required to be independently indentically distributed (i.i.d.). However, for a time series model, the observations are typically not independent. Thus, a conditional probability density function is used to write the joint density function. The log likelihood function is therefore defined as

$$\mathcal{L}(y; \psi) = \sum_{t=1}^{T} \ln p(y_t | Y_{t-1})$$

where ψ denotes the parameters vector which is to be determined; Y_{t-1} denotes the observations up to time $t-1$; $p(y_t|Y_{t-1})$ is the (joint) probability density function of the t-th set of observations.

If the initial state vector and the noise in the state space model has multivariate Gaussian distributions, the distribution of y_t conditional on Y_{t-1} is also Gaussian. Recall the section of SKF, the measurement equation could be rewritten as

$$y_t = H_t \hat{x}_t + H_t(x_t - \hat{x}_t) + d_t + \nu_t$$

It follows that the distribution of y_t conditional on Y_{t-1} is also Gaussian with mean $\hat{y}_t$

$$E_{t-1}[y_t] \equiv \hat{y}_t = H_t\hat{x}_t + d_t$$

and the covariance matrix D_t

$$Cov_{t-1}[y_t] \equiv D_t = H_t P_t^- H_t^T + R_t$$

For a Gaussian model, therefore the log likelihood function can be written explicitly as

$$\mathcal{L}(y;\psi) = -\frac{NT}{2}\ln 2\pi - \frac{1}{2}\sum_{t=1}^{T}\ln|D_t| - \frac{1}{2}\sum_{t=1}^{T}\varepsilon_t^T D_t^{-1}\varepsilon_t$$

where $\varepsilon_t = y_t - \hat{y}_t$.

4.4.2 Maximum Likelihood Estimator

The maximum likelihood estimator of the parameters vector ψ based on a sample (observation) Y is the parameter value $\hat{\psi}$ at which the likelihood function $\mathcal{L}(y;\psi)$ attains its maximum as a function of ψ, with y fixed. If the likelihood function is differentiable in ψ, the likely candidates for the maximum likelihood estimator are the values at which the gradient vector of the likelihood function is 0. This possible estimator is given by solving

$$\frac{\partial}{\partial\psi}\mathcal{L}(y;\psi) \equiv g(\psi) = 0.$$

The solution of the above equation is the local maximum, provided that the second derivatives in ψ_i of the likelihood function $\mathcal{L}(y;\psi)$ exist and the function $\mathcal{L}(y;\psi)$ is concave, which means that -1 times the matrix of second derivatives is everywhere definite positive, i.e.

$$\frac{-\partial^2}{\partial\psi\partial\psi^T}\mathcal{L}(y;\psi) \equiv Z(\psi) > 0$$

However, the boundary must be checked separately for extrema in order to get the global maximum.

In stead of finding the estimator, an alternative way of obtaining the estimates of the parameters is to use Newton-Raphson method which has a better convergence property.

4.4.3 Newton-Raphson Method

Newton-Raphson Method requires that $Z(\psi)$ to be definite positive. Let's consider the second-order Taylor series expansion of $\mathcal{L}(y;\psi)$ (from now on, we will simply use $\mathcal{L}(\psi)$ to denote $\mathcal{L}(y;\psi)$ given no confusion) around $\psi^{(0)}$

$$\mathcal{L}(\psi) \cong \mathcal{L}(\psi^{(0)}) + [g(\psi^{(0)})]^T[\psi - \psi^{(0)}] - \frac{1}{2}[\psi - \psi^{(0)}]^T Z(\psi^{(0)})[\psi - \psi^{(0)}]$$

In order to maximize $\mathcal{L}(\psi)$, we set the first order derivative of the above approximation with respect to ψ equal to zero. It turns out

$$g(\psi^{(0)}) - Z(\psi^{(0)})[\psi - \psi^{(0)}] = 0$$

This recursive equation allows us to get an improved estimate of ψ (denoted $\psi^{(1)}$) using $\psi^{(0)}$ as the initial guess.

$$\psi^{(1)} = \psi^{(0)} + [Z(\psi^{(0)})]^{-1} g(\psi^{(0)})$$

Once $\psi^{(1)}$ is obtained, we could calculate the gradient and Hessian at $\psi^{(1)}$ in order to find another improved estimate of ψ (denoted $\psi^{(2)}$). We could continue this iterating fashion until the desired estimate is obtained. The general formula for this iteration is

$$\psi^{(m+1)} = \psi^{(m)} + [Z(\psi^{(m)})]^{-1} g(\psi^{(m)})$$

Since the second-order Taylor expansion of $\mathcal{L}(\psi)$ is only an approximation of it, this iteration is often modified as following

$$\psi^{(m+1)} = \psi^{(m)} + \delta[Z(\psi^{(m)})]^{-1}g(\psi^{(m)})$$

where δ is a step control scalar chosen so as to find the best choice of $\psi^{(m+1)}$ which maximizes the log likelihood $\mathcal{L}(\psi^{(m+1)})$.

4.4.4 Quasi-Maximum Likelihood Estimation

The SKF is based on the normality assumption which is violated in some non-Gaussian models. However MLE may still be a consistent way to estimate parameters even if the normality assumptions are relaxed. The case we have here is CIR whose innovations are not Gaussian. The MLE we impliment in this situation is called Quasi-Maximum Likelihood Estimation. More detailed discussions could be find in Duan and Simonato (1997).

4.5 Data Description

Month-end American strips (Treasury Bills) from January 1993 to December 2004 are obtained from Bloomberg, totally 132 months. Maturities are chosen as 1-year, 2-year, 3-year, 5-year, 7-year, 10year and 30-year. Figure 4.1 is the term structure of the zero coupon bond prices observed at time January 29th, 1993. Figure 4.2 is the time series of the yield to maturity 1 year, observed monthly from January 29th, 1993 to December 31, 2003. We have totally 7 time series for the yield to maturity, since we have chosen 7 different maturities, pictured by Figure 4.2 to Figure 4.8.

Treasury Bills are assumed to be non-defaultable, and thus are used to calibrate the parameters in the model for interest rates.

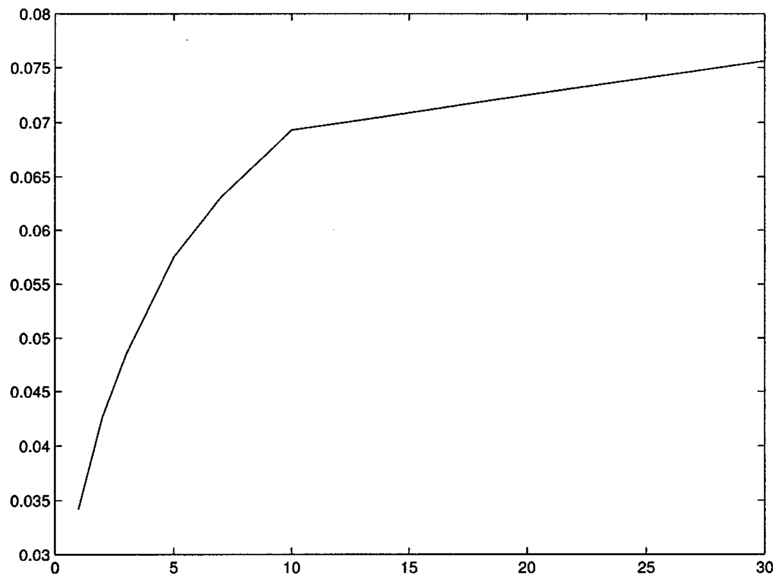

Figure 4.1: Term Structure of The Bond Price at Time $t = 0$(Jan. 29, 1993).

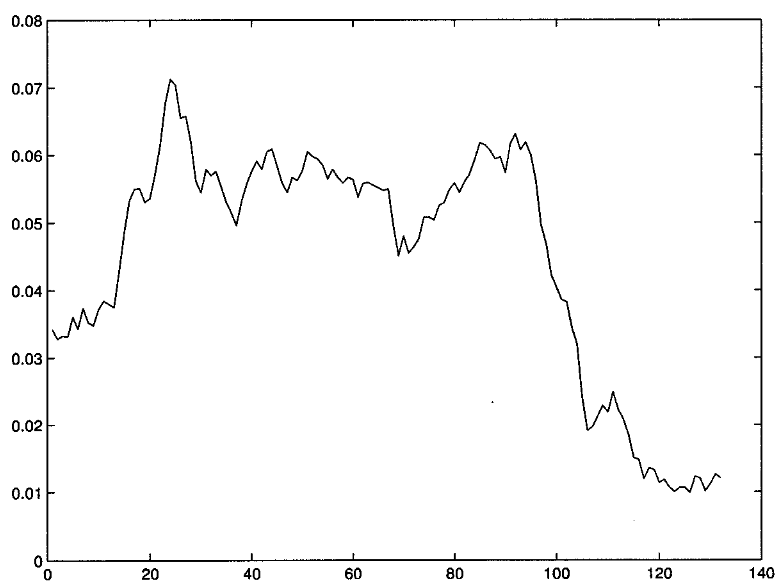

Figure 4.2: Time Series of Yield to Maturity 1 Year (time 0 is Jan. 29, 1993, time 132 is Dec. 31, 2003, monthly observation).

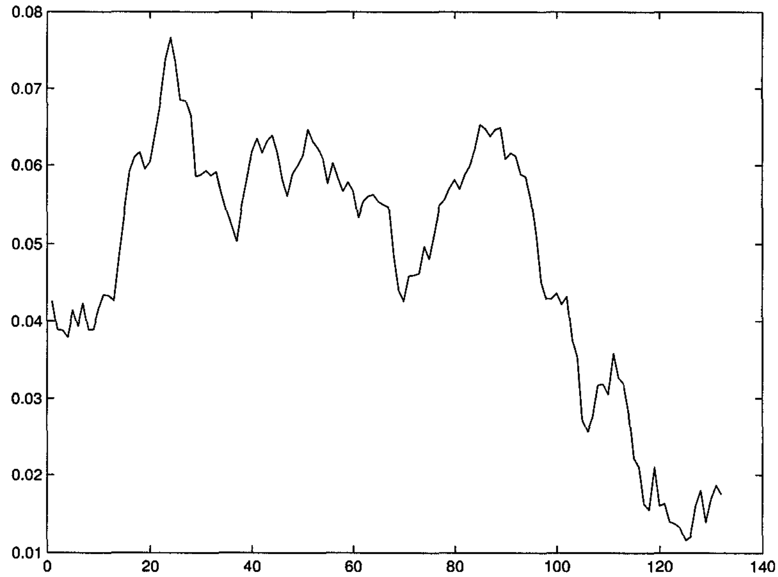

Figure 4.3: Time Series of Yield to Maturity 2 Year (time 0 is Jan. 29, 1993, time 132 is Dec. 31, 2003, monthly observation).

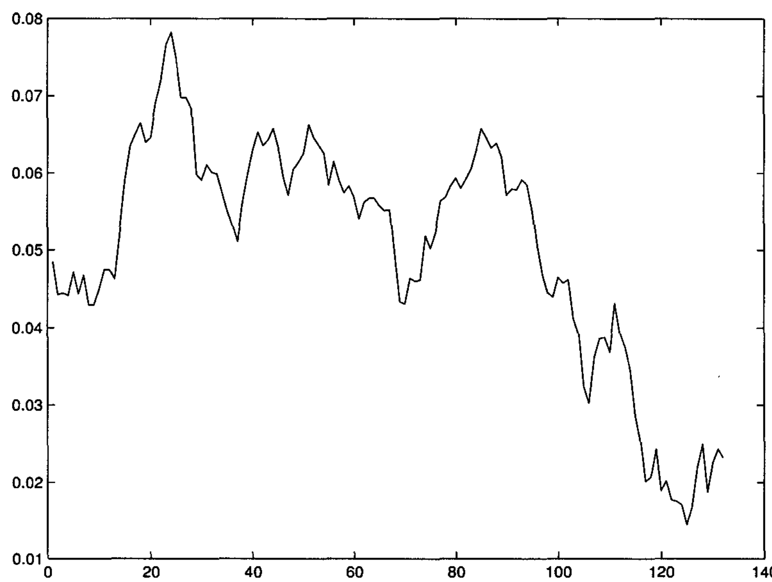

Figure 4.4: Time Series of Yield to Maturity 3 Year (time 0 is Jan. 29, 1993, time 132 is Dec. 31, 2003, monthly observation).

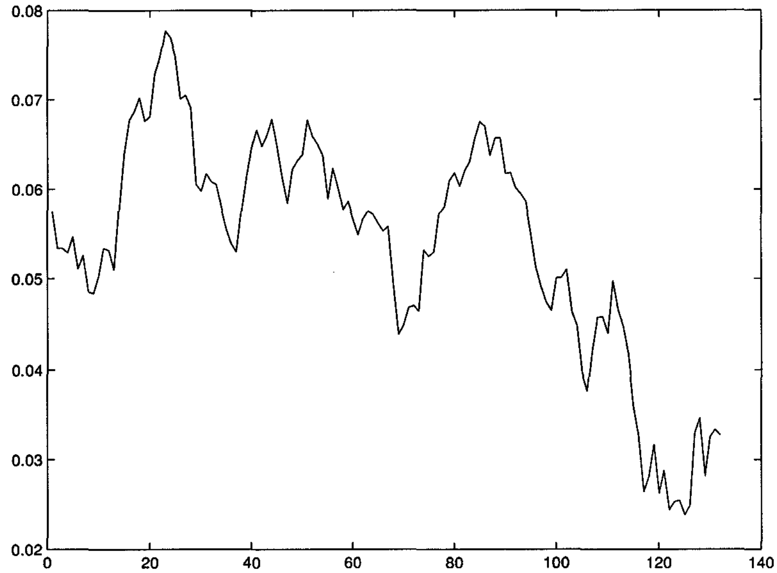

Figure 4.5: Time Series of Yield to Maturity 5 Year (time 0 is Jan. 29, 1993, time 132 is Dec. 31, 2003, monthly observation).

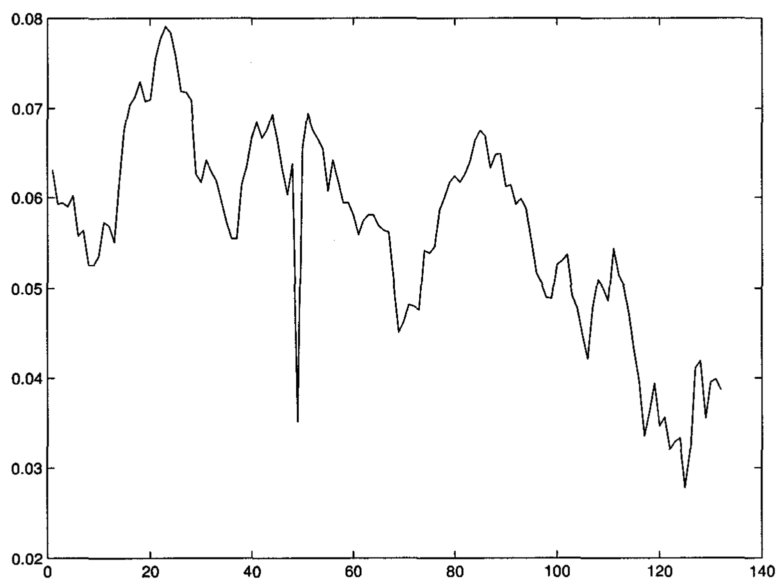

Figure 4.6: Time Series of Yield to Maturity 7 Year (time 0 is Jan. 29, 1993, time 132 is Dec. 31, 2003, monthly observation).

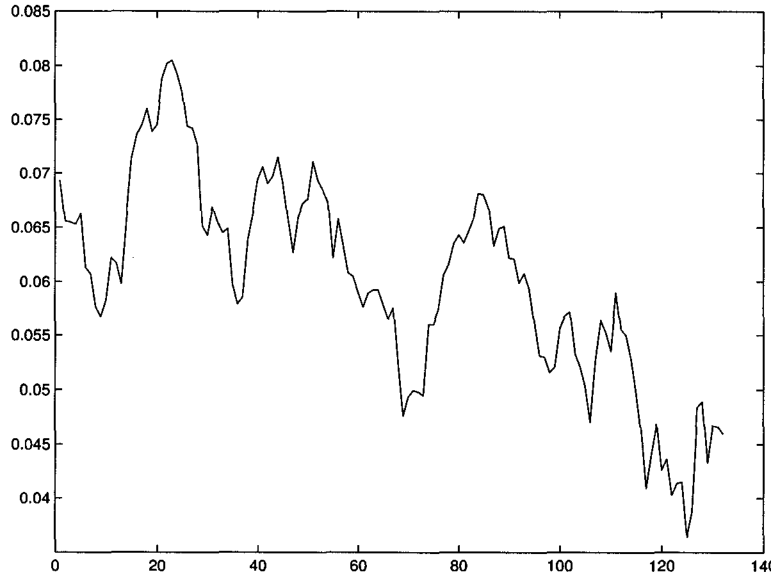

Figure 4.7: Time Series of Yield to Maturity 10 Year (time 0 is Jan. 29, 1993, time 132 is Dec. 31, 2003, monthly observation).

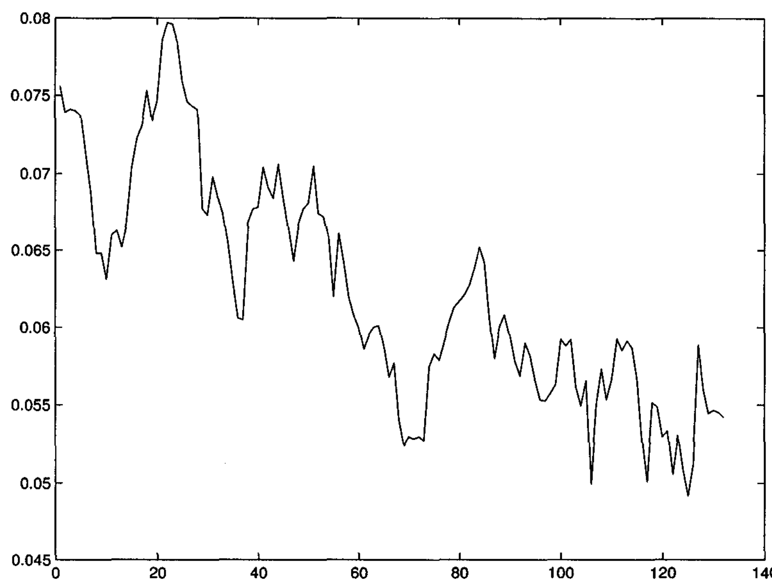

Figure 4.8: Time Series of Yield to Maturity 30 Year (time 0 is Jan. 29, 1993, time 132 is Dec. 31, 2003, monthly observation).

Corporate bonds data were obtained from Dr. Peter Miu from the Michael G. DeGroote School of Business at McMaster University. There are totally 82 bonds issued by 21 firms (around 4 bonds per firm) observed monthly from January 1995 to June 2001 (totally 78 observations for each bond). These corporate bonds are assumed to be defaultable, and thus are used to conduct the calibration for hazard rates. Coupons are paid half-yearly. Table 4.1 lists all the names of the companies observed and their rating classes. Figure 4.9 depicts the time series of three bond prices issued by Atlantic Richfield Company.

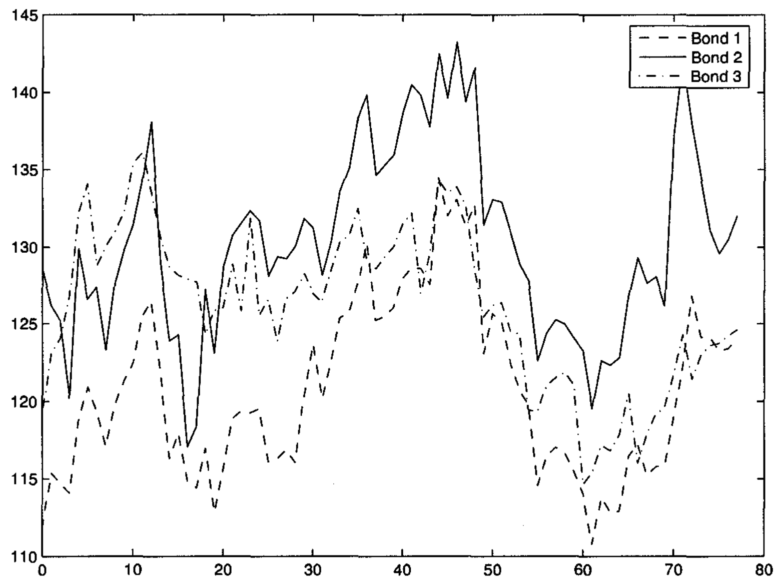

Figure 4.9: Time Series of Three Bond Prices Issued by Atlantic Richfield Company. The Coupons and Maturities are \$4.565, \$4.94, \$5.44 and 16.09 year, 21.09 year, 10.609 year for Bond 1, Bond 2 and Bond 3 respectively(time 0 is Jan. 29, 1995, time 77 is June, 2001, monthly observation).

Notice that we only have the data of corporate bonds data from January 1995 to June 2001. Therefore we will only need the time series of interest rate in that time horizon.

Name of the company	Rating Class	Number of Bonds
Atlantic Richfield Company	AA+	3
AT&T	#NA	5
BellSouth Telecommunication Inc	AA-	4
Burlington	BBB+	6
Carterpillar	A+	5
Coastal Corp	BBB	5
Consolidated Natural Gas Company	BBB+	3
Eli Lilly	AA	2
Illinois Power Company	BBB+	3
Mcdonnel Douglas	AA-	3
New York Telephone	A+	7
Pacific Bell	AA-	4
Philadelphia Electric	A	3
Phillips Petroleum	BBB	3
Public Service Electricity & Gas Company	A-	9
Ralston Purina Company	BBB+	3
Pockwell International Corporation	AA-	2
Safeway Plc	BBB-	3
Southern California Gas	AA-	4
Southwestern Bell Telephone Company	AA-	3
Time Warner Incorporated	BBB+	2
Total: 21 firms		82 bonds

Table 4.1: Issuers of Corporate Bonds Observed, Data is from Peter Miu.

4.6 Calibrate the Default-free Interest Rate

The state space form of the model for the default-free interest rate could be represented as

$$s_{k+1/12} = \theta(1 - e^{-\kappa/12}) + e^{-\kappa/12}s_k + \omega_k, \quad k = 0, \frac{1}{12}, \frac{2}{12}, \ldots, \frac{M}{12}$$

$$Y(P, k, T) = \frac{B(k,T)s_t}{T-k} + (\alpha_r - \frac{\ln A(k,T)}{T-k}) + \nu_k$$

The first equation is transition equation which is obtained from discretizing the CIR of s_t which is unobserved state variable. Since the data we choose is month-end, time steps we choose here is monthly too, or $\frac{1}{12}$ yearly. However, s_t itself is non-central chi-square distributed but non-Gaussian. The innovation ω has conditional expected mean of zero and a conditional variance of

$$Q_k = \sigma^2 \left(\frac{1 - e^{-\kappa/12}}{\kappa}\right) \left(\frac{1}{2}\theta(1 - e^{-\kappa/12}) + e^{-\kappa/12}s_{k-1/12}\right)$$

Noticing that Q_k is linear in $s_{k-1/12}$, we construct $\hat{Q}_k$ by replacing $s_{k-1/12}$ in Q_k with $\hat{s}_{k-1/12}$

$$\hat{Q}_k = \sigma^2 \left(\frac{1 - e^{-\kappa/12}}{\kappa}\right) \left(\frac{1}{2}\theta(1 - e^{-\kappa/12}) + e^{-\kappa/12}\hat{s}_{k-1/12}\right)$$

The second equation is the measurement equation in which the yield $Y(P, t, T)$ is observed monthly. The noise iterm v_k is added to allow some imperfect measurements from observing the yeilds' data.

Given the state space form, we could perform SKF with QMLE to get the filtration of the state variable s_k as well as the parameters in the model. The calibrated result and analysis could be found in section 4.8.

4.7 Calibrate the Hazard Rates

After we have calibrated the parameters in interest rate model, we assume the estimated value are exact value when we conduct the estimation for hazard rates. We do not need to run the KF for the interest rate again. For each firm j, we need to run the EKF with QMLE once. There are totally 21 firms, thus we need to run 21 times EKF with QMLE.

As in section 4.2.3, we know that observed corporate coupon bond prices are nonlinear in the unobserved variable $s_{j,t}$. As a result, we could not implement Standard Kalman Filter directly to conduct this calibration. Extended Kalman Filter is therefore needed to solve this nonlinear problem. Since defaultable corporate bond prices could be represented into a linear combination of zero-coupon bond price which followes term structure, this makes linearization much easier to do.

The transition equation is of the same form as in the calibration of interest rates. Following the discussion in section 4.3.3, $\phi(t, x_t)$ is the linear combinations of the zero-coupon bond prices for $G_j(t, T_n)$. At every observation time t, the H_t and d_t in Kalman Filter have to be updated each time which are given as follows

$$H_t = \sum_{k=1}^{n} c_j \frac{dP_j(t, T_k)}{dx} + 100 \frac{dP_j(t, T_k)}{dx}$$

$$d_t = \sum_{k=1}^{n} c_j P_j(t, T_k) + 100 P_j(t, T_n) - H_k \hat{x}_t^-$$

where $\frac{dP_j(t,T_k)}{dx}$ is given by

$$\frac{dP_j(t, T_k)}{dx} = -A_{1,j} A_{2,j} B_{2,j} L \exp[-B_{1,j} b_j s_t - B_{2,j} L \hat{x}_t^-]$$

and where $\hat{x}_t^-$ denotes the priori estimate for $s_{j,t}$ at time t.

4.8 Empirical Results

The empirical results are based on the previous two sections. The constant parameters α_r and $\alpha_{j,h}$ as in the model (section 4.1) are unable to be pinned down with any reliability by the data. Pearson and Sun (1994), Duffee (1999) also faced this problem. Here, I follow their approach and set both of them equal to zero. Therefore, the parameters we care about are: for interest rate, $(\kappa, \theta, \sigma, \lambda)$; for hazard rates, $(\kappa_j, \theta_j, \sigma_j, \lambda_j, \beta_j)$ for each company j. Since time here is measured in year, all the parameters calibrated are expressed on an annual basis.

4.8.1 Results for Interest Rates

Estimation results for the default free interest rates are displayed in Table 4.2. The parameter κ, which represents the rate of mean reversion is around 0.3790. The parameter θ is around 0.0365 which is the limit mean of this CIR model. The volatility σ is around 0.0666 which is very low. The product of κ and θ is 0.01383 which should remain the same even the two parameters are under risk-neutral measure. The sum of κ and λ is around 0.1931 which represents the kappa in the CIR under risk-neutral measure which remains positive. Table 4.3 shows the Root of Mean Square Error of the yields for seven different maturities. They are computed by using the calibrated parameters and the filtered time series of the default-free interest rate. The average of the RMSE is around 0.0041 which is comparatively high. Duffee (1999), Chen & Scott (2003) have calibrated multi-factor CIR model for the default-free interest rate using the same technique over different time periods. Duffee (1999) claims that the RMSE he found is less than 10 basis points using two factor CIR model for the default-free interest rate. This is the reason why two factor models are prefered than one factor models. Figure 4.10 is the filtered time series of the interest rate.

Since we only have corporate bond data available from January 1995 to June 2001, only the time series of the interest rate in that time horizon is used when we calibrate hazard rate models. Figure 4.11 shows the time series of the interest rate from January 1995 to June 2001, which is very similar to the filtration Miu got in his paper (2003).

κ	θ	σ	λ	$\kappa\theta$	$\kappa + \lambda$
0.3790	0.0365	0.0666	-0.1859	0.0138	0.1931
(0.0044)	(0.0050)	(0.0045)	(0.0543)		

Table 4.2: Estimation Results for Interest Rate. The numbers in the parentheses are the standard errors.

1-year	2-year	3-year	5-year	7-year	10-year	30-year
0.0044	0.0027	0.0026	0.0032	0.0045	0.0049	0.0061

Table 4.3: Root of Mean Square Error of the yields for different maturities.

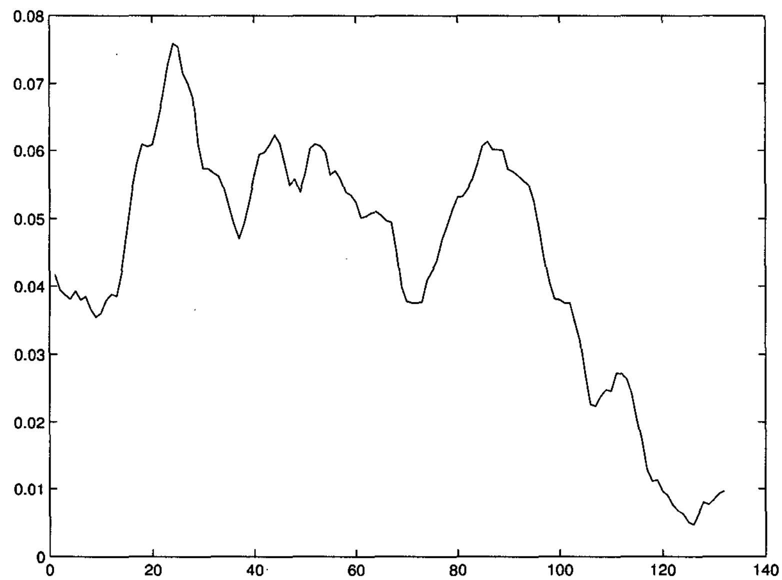

Figure 4.10: Time Series of interest rate (time 0 is Jan. 29, 1993, time 132 is Dec. 31, 2003, monthly observation).

But my result is a little bit higher than his, around 50 basis point higher. Miu used Hull-White model for the interest rate and calibrated the parameters by spline fitting. This time series of interest rate is somewhat robust, given the consideration that we are using different models for the default free interest rate.

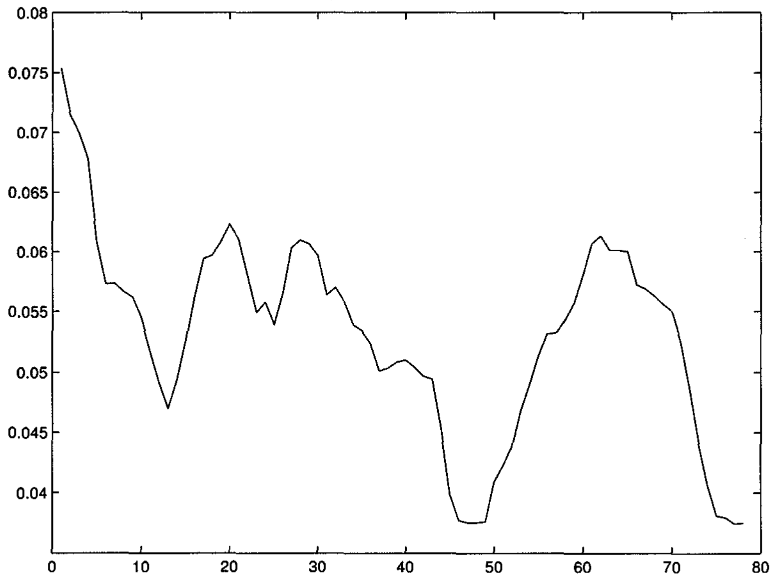

Figure 4.11: Time Series of interest rate (time 1 is Jan 1995, time 78 is June 2001, monthly observation).

4.8.2 Results for Hazard Rates

Let's firstly concentrate on the results of one company say Atlantic Richfield Company (the results of all other 20 firms will be given in shortly in the section *Results by Ratings*). There are total three bonds issued by Atlantic Richfield Company observed. The coupons of the three bonds are \$4.565, \$4.94, \$5.44 respectively paid half yearly. The time to maturity at the first observation time Januray 1995 are 16.09 years, 21.09 years, 10.609 years respectively. The observation time horizon is from January 1995 to June 2001 with monthly observation (refer to Figure 4.9).

κ_j	θ_j	σ_j	λ_j	β_j	$\kappa_j\theta_j$	$\kappa_j + \lambda_j$
0.2884	0.0007	0.0713	-0.2256	-0.0274	2.02e-4	0.0628
(0.1217)	(0.0001)	(0.0141)	(0.0663)	(0.0424)		

Table 4.4: Estimation Results for Hazard Rate of Atlantic Richfield Company. The numbers in the parentheses are the standard errors. AA+ Rated.

Table 4.4 shows the estimated parameters in the hazard rate model for the Atlantic Richfield Company. The calibrated results show that the Atlantic Richfield Company has a median mean-revearting coefficient κ_j and a low volatility. The long run mean θ_j is around 0.0007 which is very small. Small long run mean indicates small hazard rates. This is consistant with the rating class of the Atlantic Richfield Company which is AA+ rated. The correlation coefficient β_j is found to be negative. This is also found in some other papers, such as Duan & Simonato (1995), Chen & Scott (2003), Geyer and Pichler (1998). Duffee (1999), Miu (2003). Since the correlation between interest rate and hazard rates are fully determined by β_j in our set up, negative β_j means that the default-free interest rate is negatively correlated with the hazard rates. However, there are also some papers in the literature claiming that positive correlation between those two are found, such as Tauren (1999). This correlation relationship is still a controversial problem nowadays. For some other firms, I also find positive β_j.

Because we are using relatively a few bond data to fit each firm's term structure, the individual parameters estimated are subjected to a certain uncertainty. This is also noticed by Duffee (1999), in which he focuses his discussion on the median parameter across all 161 firms.

Figure 4.12 is the filtered time series of the firm specific factor for the Atlantic Richfield Company. Figure 4.13 depicts the corresponding filtered time series of the hazard rate process. Notice that most the hazard rates are positve during the observed time horizon, but still some of them are below zero. Negative hazard rates are due to the small firm specific factor $s_{j,t}$ and the negative coefficient β_j of the common factor. Since the hazard rates in the model are risk-neutral hazard rates which is usually much higher than the real hazard rates, the time series of the real hazard rates should be much lower than the one have got here in Figure 4.13.

The RMSE of the three different coupon bonds are \$2.0132, \$2.5000 and \$1.3730 respectively. The average RMSE is \$ 1.9620 which is less than two dollars.

4.8.3 Results by Ratings

The results for the all 21 firms are given by Figure 4.14 to Figure 4.34 in the order of decreasing ratings. In these figures, the numbers in the vector "Para" are $[\kappa_j, \theta_j, \sigma_j, \lambda_j, \beta_j]$. Notice that, there are 7 firms(out of 21 firms), whose "Para" is not available. This is due to the drawbacks of the numerical scheme we have used. QMLE does not work very well when it is combined with EKF. Although the filtered time series does not depend on the small change of the initialized parameters very much, QMLE is very sensitive on the initial parameters. We do not have a prototype to select the initialization. If we choose a very bad parameter vector as a start, it may cause the MLE diverges or it may produce a singular matrix or matrices which are badly scaled. Details will be discussed in next section.

Figure 4.14 is the time series of the hazard rates for a AA+ rated company. Comparing Figure 4.14 to the Figure 4.34 which is the time series of the hazard rates for a BB- rated company, we clearly see that the hazard rates of a lower rated company is much lower than a higher rated firm. Since θ_j indicates the level of the hazard rates, it is expected to be smaller for a higher rated firm. This is not that clear in the results. The changes of the hazard rates over the observation periods for all the 21 firms have something in common. For most of them, the time series decrease in the first half period and then increase again in the second half period. But the time series of filtered interest rates have a reverse phase, meaning that it increase first and then decrease. This realization shows the negative correlation between the interest rates and the hazard rate. However, for some of the firms, negative β_j are found, indicating the negative correlations between the hazard rates and the interest rate. The similarity of the changes of the time series across different firms captures the systematic influence of the common factors on all the firms in the market.

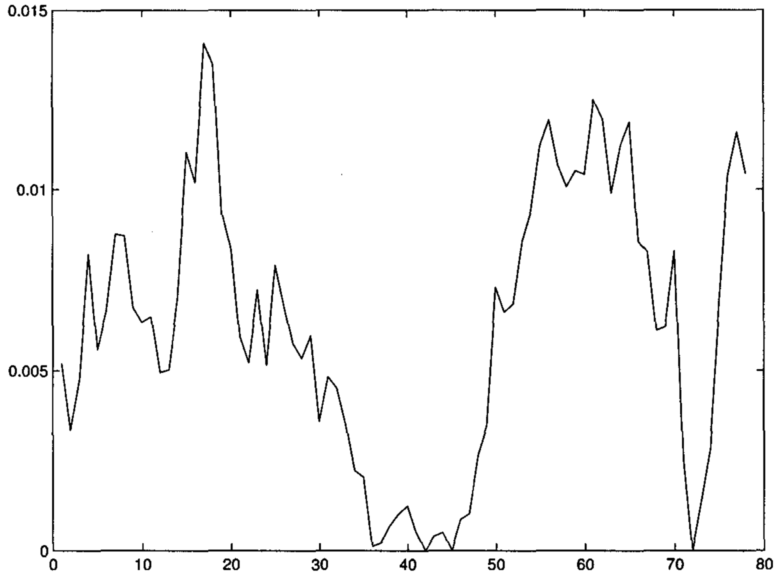

Figure 4.12: Time series of the firm specific factor of the Atlantic Richfield Company (AA+ rated). Time 1 is Jan. 1995, time 78 is June 2001, monthly observation.

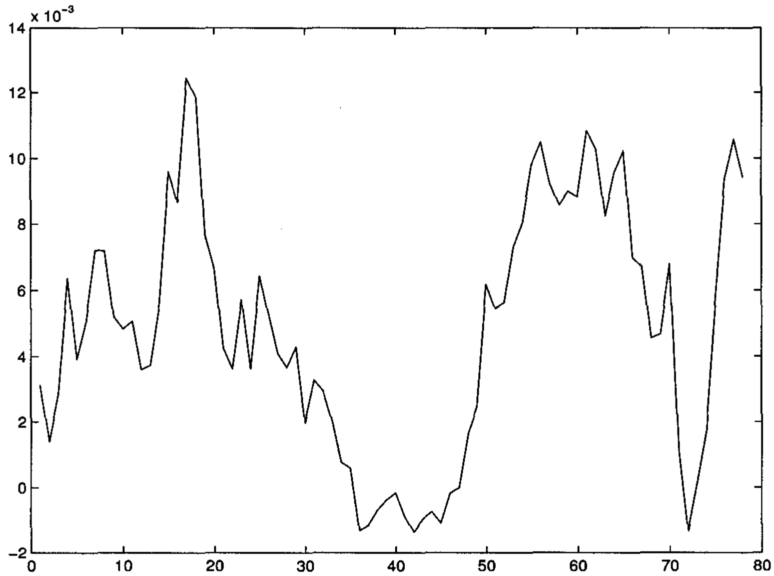

Figure 4.13: Time series of the hazard rates of the Atlantic Richfield Companytime (AA+ rated). Time 1 is Jan. 1995, time 78 is June 2001, monthly observation.

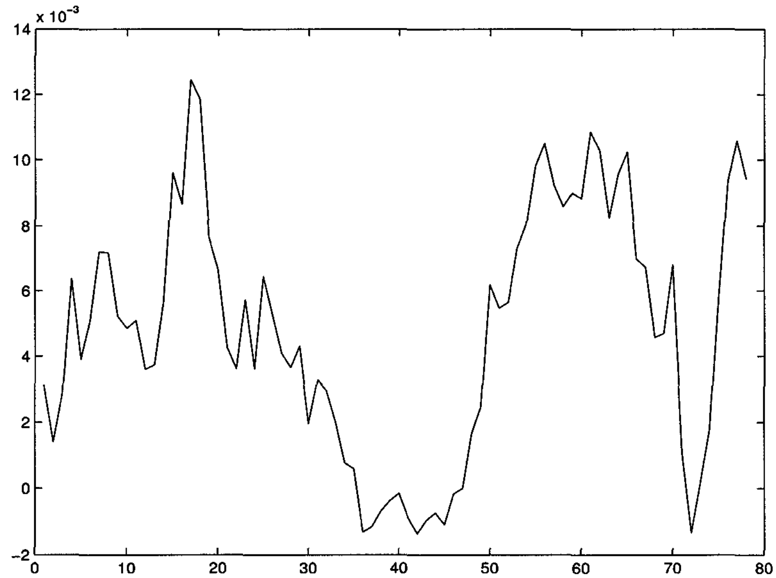

Figure 4.14: AA+ rated; Para = [0.2884, 0.0007, 0.0713, -0.2256, -0.0274]; RMSE = 1.9. Atlantic Richfield Company.

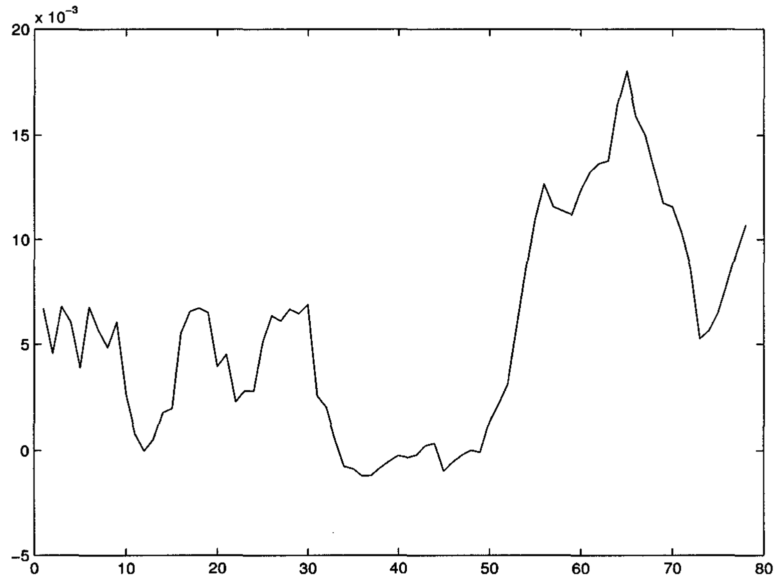

Figure 4.15: AA rated; Para = [0.3124, 0.0020, 0.0855, -0.1800, -0.0240]; RMSE = 4.6992. Eli.

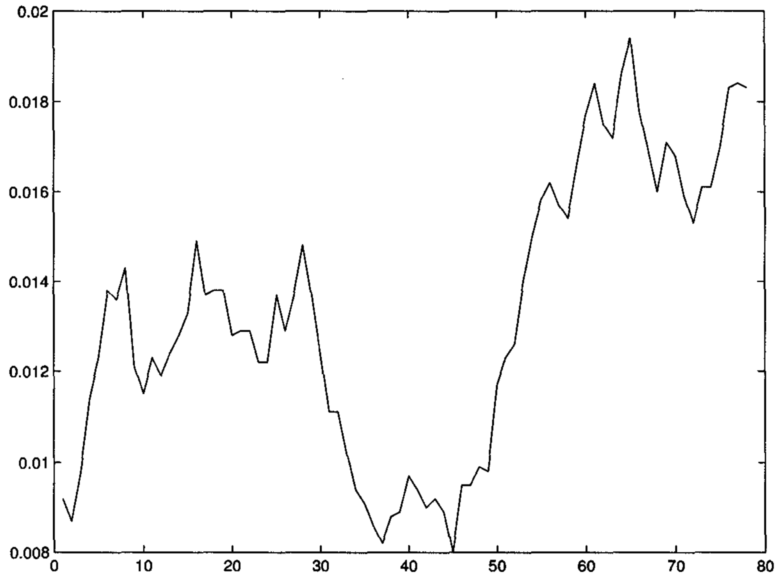

Figure 4.16: AA- rated; Para = []; RMSE = 7.0115. Southcal.

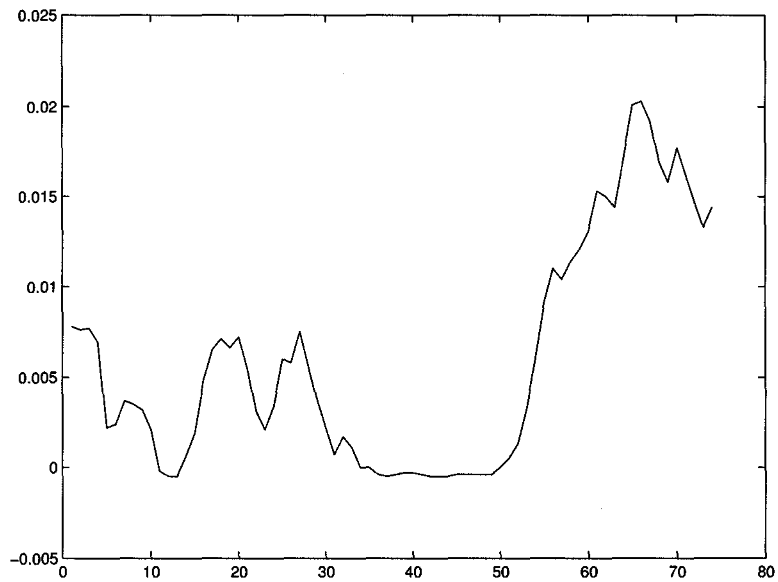

Figure 4.17: AA- rated; Para = []; RMSE = 0.8229. Bellsouth.

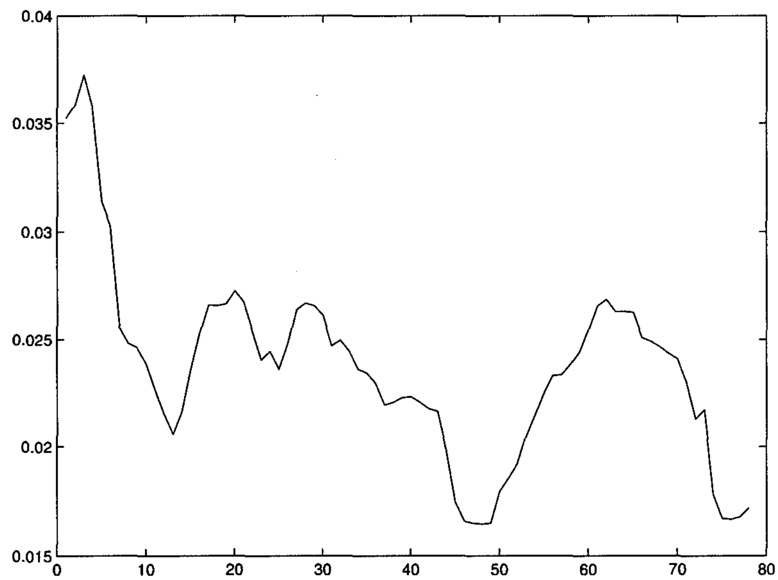

Figure 4.18: AA- rated; Para = [0.0846, 0.0176, 0.3835, -0.5606, 0.4379]; RMSE = 16.2830. Mcdonnell.

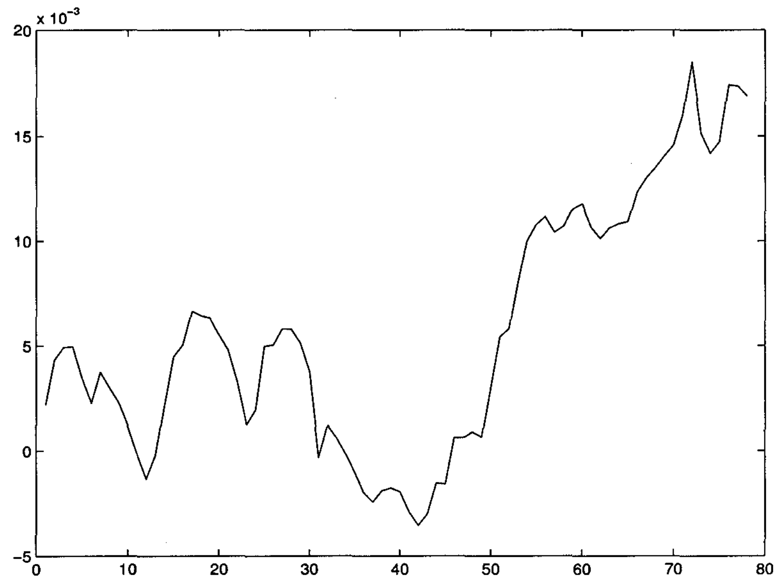

Figure 4.19: AA- rated; Para = [0.2198, 0.0005, 0.0385, -0.1368, -0.1012]; RMSE = 2.1666. Pacific.

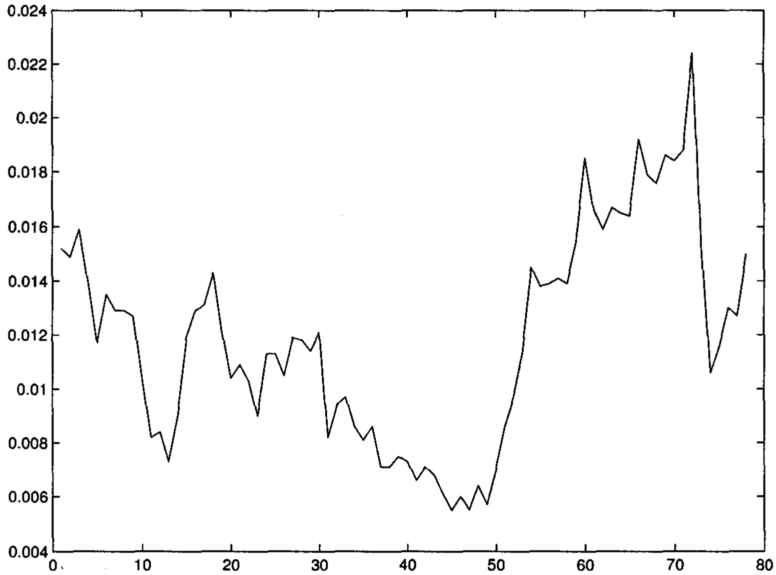

Figure 4.20: AA- rated; Para = [0.2635, 0.0012, 0.0455, -0.2407, 0.1214]; RMSE = 5.4235. Southwest.

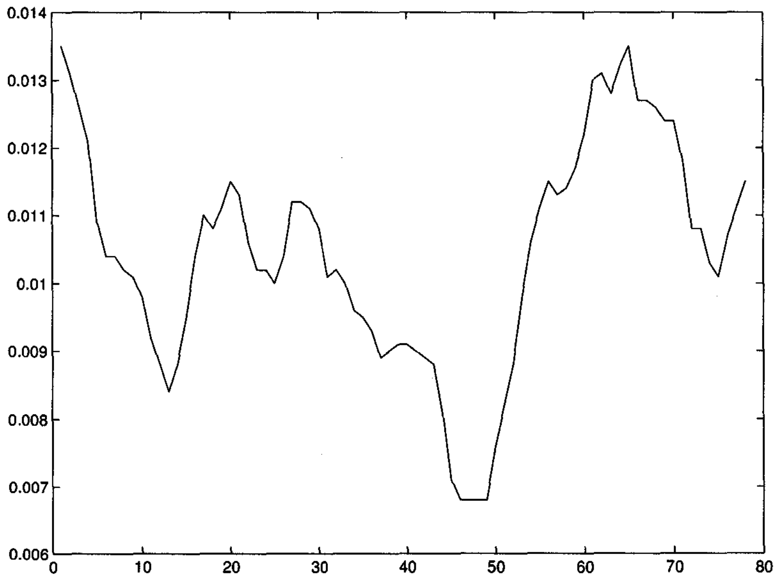

Figure 4.21: AA- rated; Para = [0.0327, 0.0032, 0.0343, -0.6227, 0.1784]; RMSE = 39.7250. Rockwell.

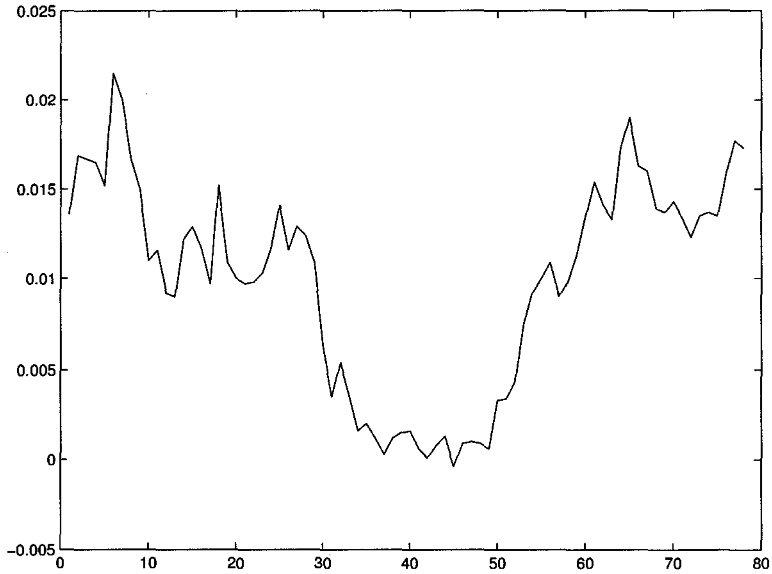

Figure 4.22: A+ rated; Para = []; RMSE = 4.3658. Newyork.

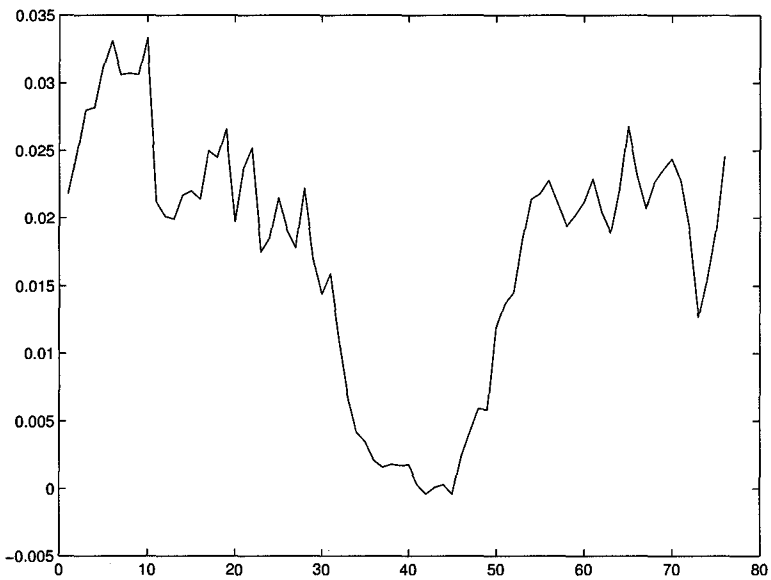

Figure 4.23: A+ rated; Para = []; RMSE = 4.3271. Caterpillar.

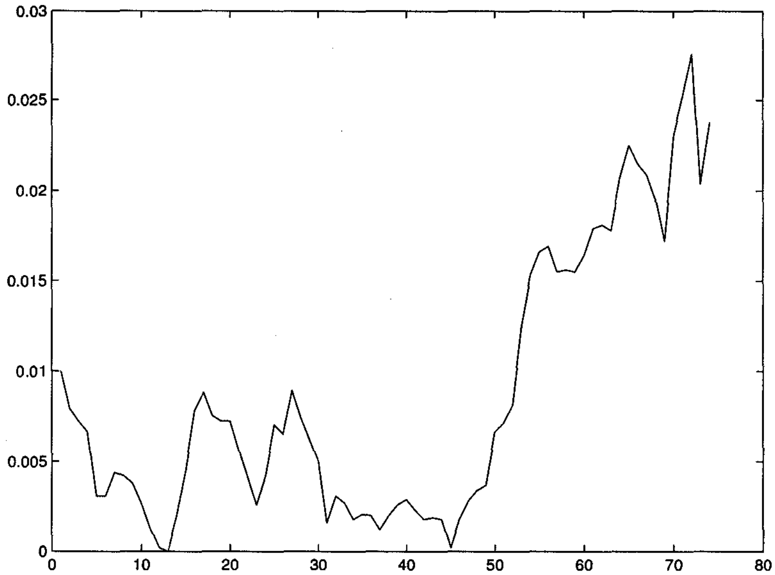

Figure 4.24: A rated; Para = []; RMSE = 0.9935. AT&T.

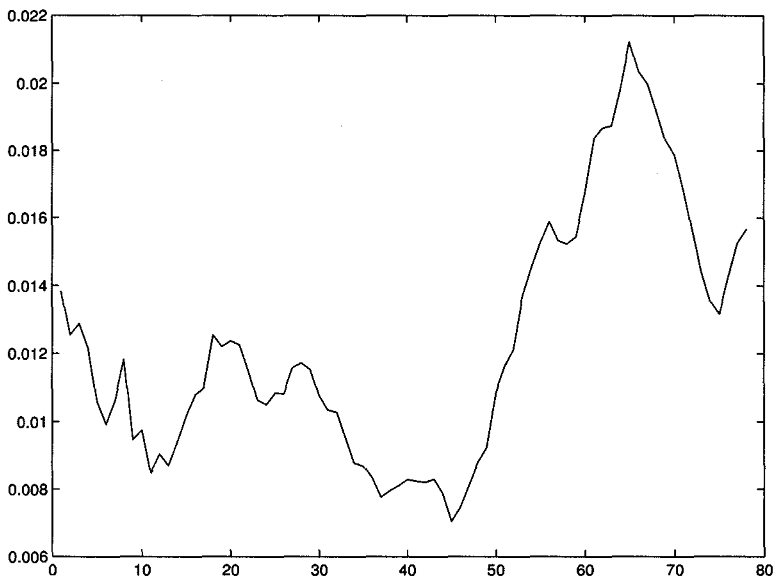

Figure 4.25: A rated; Para = [0.2837, 0.0026, 0.0423, -0.3780, 0.1473]; RMSE = 21.8183. Philadelphia.

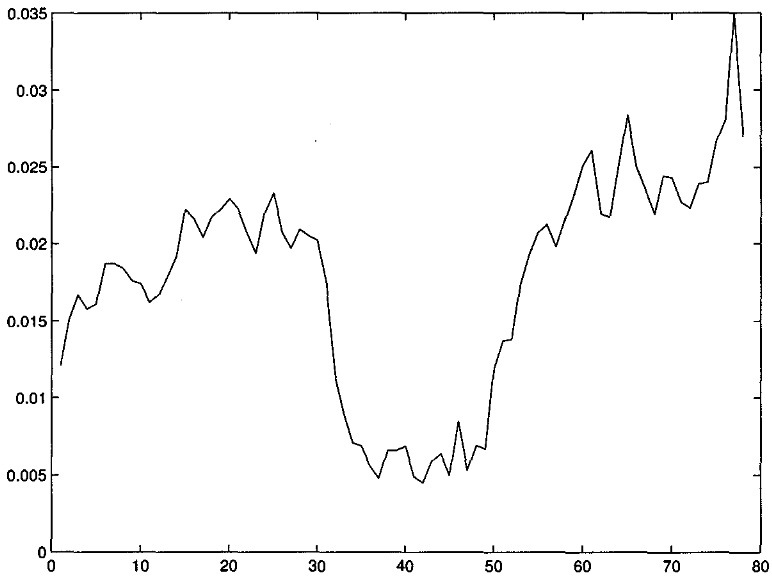

Figure 4.26: A- rated; Para = []; RMSE = 3.4751. Public.

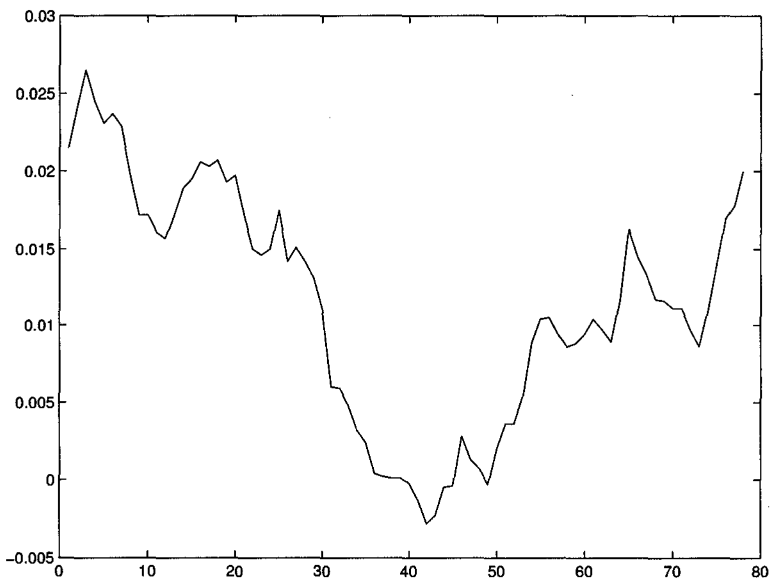

Figure 4.27: BBB+ rated; Para = [0.0176, 0.0006, 0.0309, -0.0279, -0.1885]; RMSE = 7.3986. Time.

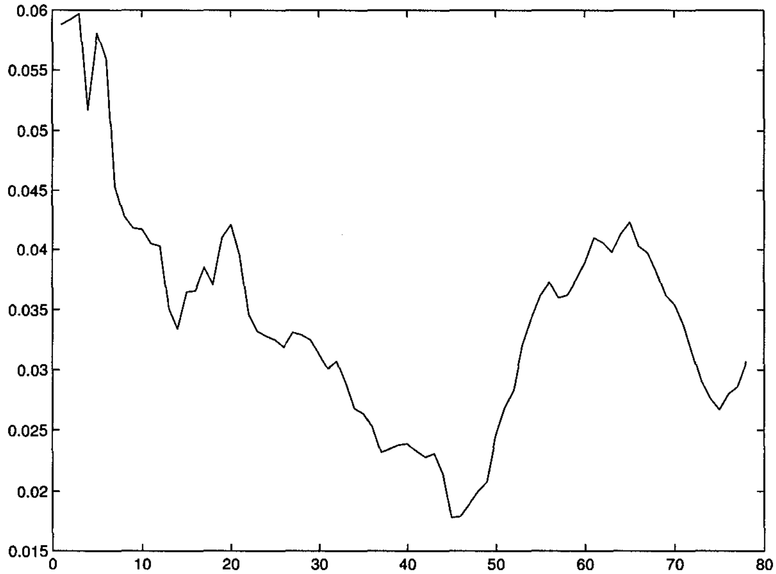

Figure 4.28: BBB+ rated; Para = [0.0366, 0.0011, 0.0478, -0.0116, 0.3973]; RMSE = 2.6671. Illinois.

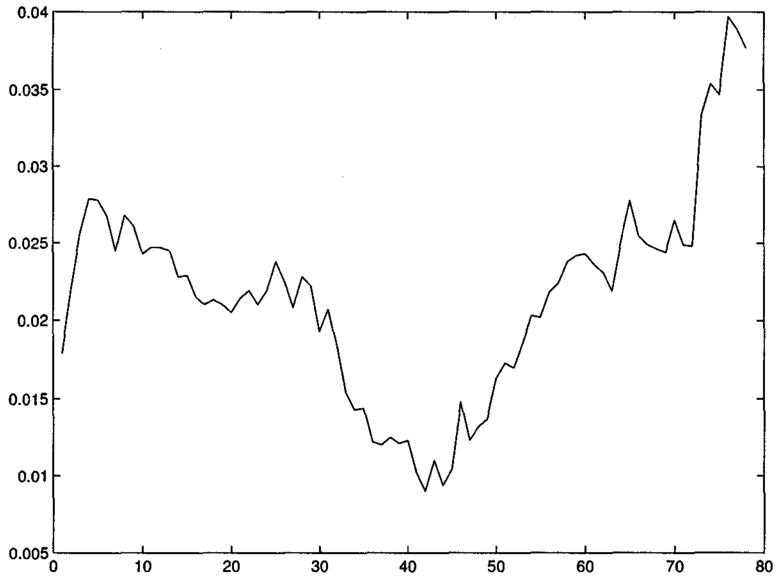

Figure 4.29: BBB+ rated; Para = []; RMSE = 2.2893. Burlington.

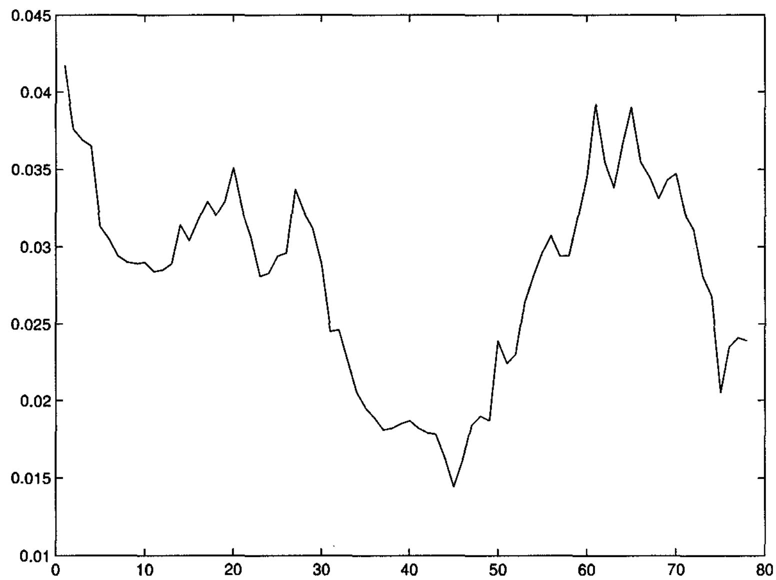

Figure 4.30: BBB+ rated; Para = [0.2719, 0.0051, 0.0595, -0.1851, 0.3607]; RMSE = 10.0519. Ralston.

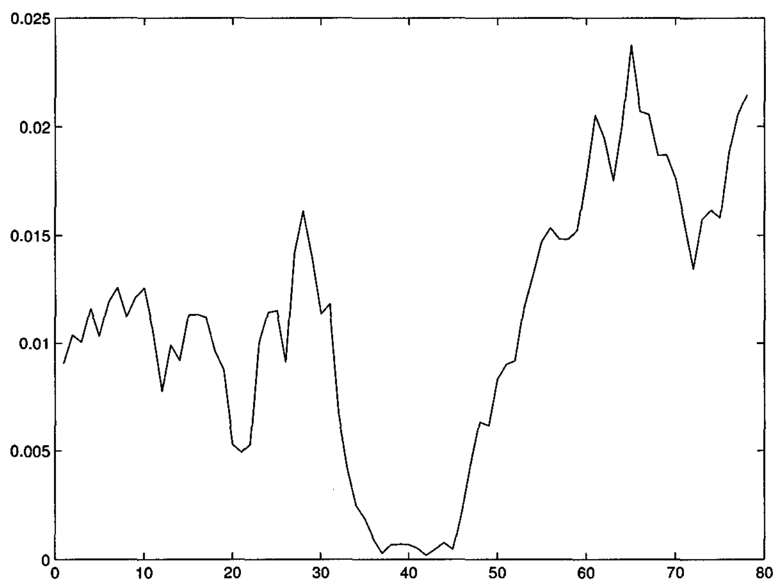

Figure 4.31: BBB+ rated; Para = [0.3000, 0.0007, 0.0679, -0.2047, -0.0082]; RMSE = 4.8428. Consol.

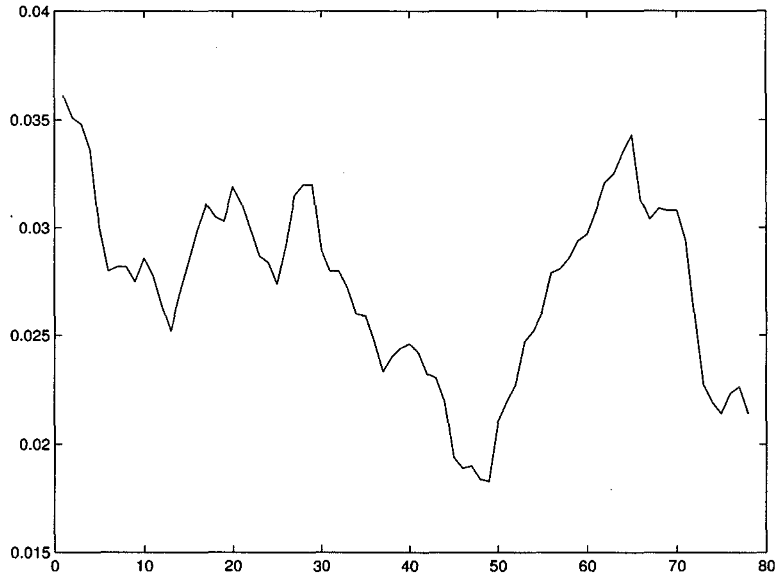

Figure 4.32: BBB rated; Para = [0.2769, -0.0004, 0.0240, -0.4296, 0.4568]; RMSE = 20.3572. Phillips.

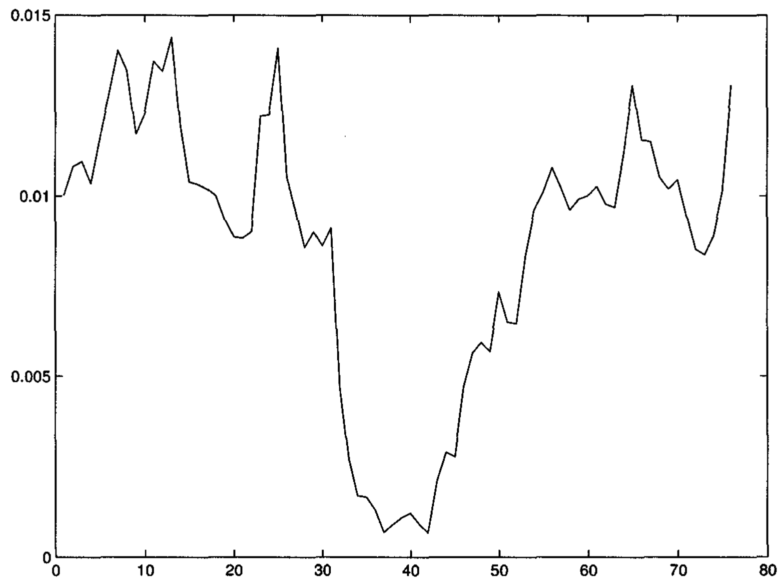

Figure 4.33: BBB rated; Para = [0.2695, 0.0015, 0.0281, -0.2651, -0.02]; RMSE = 3.0630. Coastal.

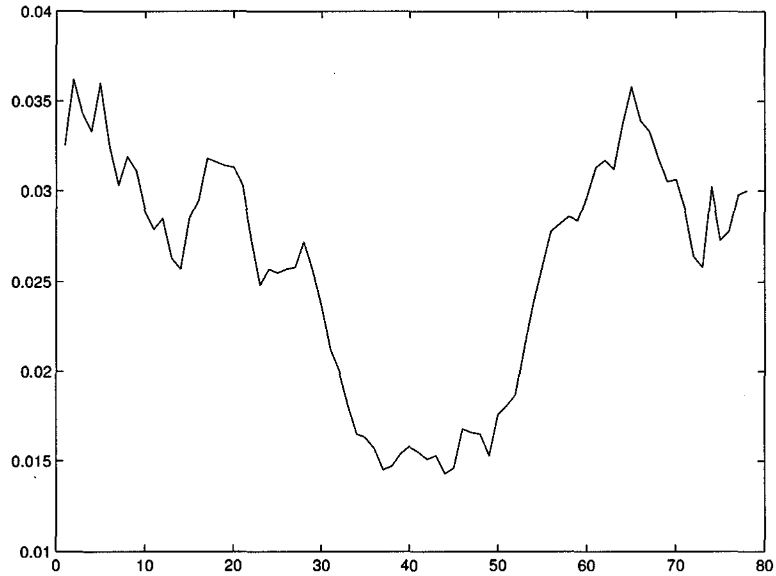

Figure 4.34: BBB- rated; Para = [0.1970, 0.0014, 0.0357, -0.2073, 0.2556]; RMSE = 2.6624. Safeway.

4.9 Monte Carlo Simulation

In this section, Monte Carlo Simulation is conducted for both interest rate and hazard rates. For the case of interest rate, KF with QMLE works well at least in the simulated world. Duan and Simonato (1999) also have showed their evidence by Monte Carlo Simulation. However, no evidence of Monte Carlo Simulation has been done yet for the case of hazard rates where the state space form is non-linear. In our study, the results show that the Quasi Likelihood function does not produce a good approximation for the real likelihood function in this non-linear problem. But the EKF still works very well in this situation.

4.9.1 Simulation for Interest Rate

Assume the default free interest rate follows CIR and the parameters under physical measure are given as well as the market price of risk λ. In short, the origial parameters given are $(\kappa, \theta, \sigma, \lambda)$. Our objective is to use the Monte Carlo Simulation to recover the original parameters. This is done through the following logic.

First, we use the discretized CIR with the given parameters to simulate a path of default free interest rate under physical measure. Then, use the given parameters and simulated time series of default free interest rate to calculate the theoretical yields for different maturities. After that, the theoretical yields are entered into the Kalman Filter and an estimated time series of the default free interest rate is produced by KF. Finally, we use QMLE in conjunction with the KF to recover the original parameters.

The step size for simulating CIR is chosen to be monthly, i.e. 1/12 yearly. Totally around 132 observations is simulated (11 years). Strips are simulated for seven different maturities: 1-year, 2-year, 3-year, 5-year, 7-year, 10-year and 30-year.

Figure 4.14 shows the simulated interest rate path and the filtered path. Graphically the Kalman Filter works pretty well. The accuracy of the Kalman Filter depends on the observation error (measurement error). Graphically we observe that the less observation error we have, the more accurate the KF works. In fact, the Root of Mean Square Error (RMSE) calculated for the filtered interest rate with smaller observation error R is smaller than the one that has bigger observation error. In the Figure produced, the RMSEs are 0.0043 for R = 0.001 and 0.0023 for R = 0.0001 respectively.

The calibration of the parameters is conducted under the assumption that the observation error R is 0.0001. The accuracy of the calibrated results depend on three factors: the step size dt, the number of simulated points (observations) for each path M and the number of simulated paths No.

Let's consider the latter two cases by fixing dt to be monthly. Table 4.5 shows

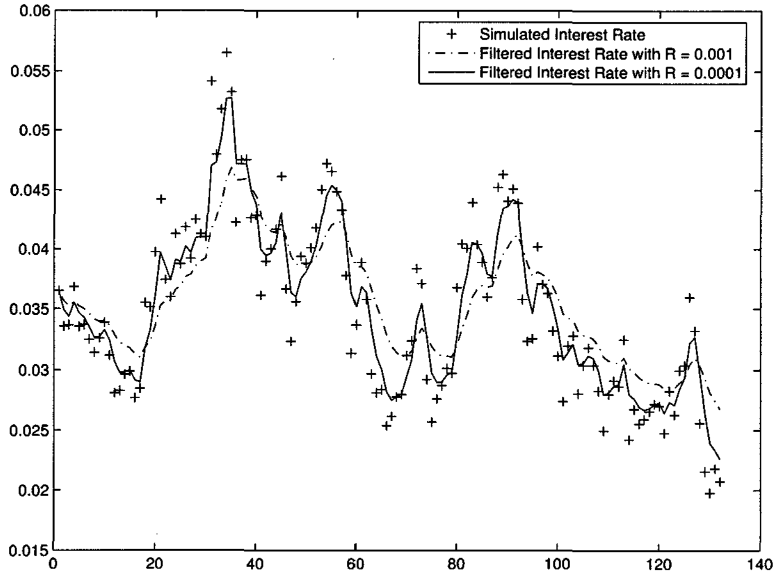

Figure 4.35: Filtered Interest Rates with different observation error R. Parameters used for simulation are $\kappa = 0.3790$, $\theta = 0.0365$, $\sigma = 0.0666$. Step size is chosen to be monthly (1/12 yearly). Around M = 132 points are simulated.

the calibrated results for different M and No we choose. The second row of the table is the origial parameters given. The third row is obtained by simulating one path of interest rate with 132 points. The fourth row is obtained by simulating one path of interest rate with 1000 points. The last row is obtained by doing the second row for 20 times and average the results. It is clear that the more points we simulate the more accurate results we will get; and that the more paths we simulate the more accurate results we will get.

Notice that the second row with M = 132 and No = 1 is exactly the situation we have in calibrating the interest rate parameters using the real observed yields. We have totally 132 observations with one path of interest rate process. The simulated results are not very good comparing to the origial parameters. This is also noticed by Duffee (1999). He mentioned that the individual parameter estimates are subject to substantial uncertainty due to lack of data. For the hazard processes, since we have

	κ	θ	σ	λ
Original	0.3790	0.0365	0.0666	-0.1859
M=132, No=1	0.3310	0.0415	0.0434	-0.1375
M=1000,No=1	0.3779	0.0366	0.0518	-0.1814
M=132, No=20	0.3898	0.0381	0.0599	-0.1877

Table 4.5: Estimated Parameters with different number of simulated points M and different number of simulated paths No.

even less observations, the estimated parameters are even more uncertain. This is why Duffee (1999) studied the median of the calibrated parameters of 161 firms instead of individual results. But if we choose daily observations instead of monthly and hence increase the number of observations, the results will become better numerically, see Duan & Simonato (1999).

Alternatively, we can calibrate those parameters by minimizing the RMSE of the yields. Unfortunately, this does not work not very well due to the fact that the measurements (the yields) are too small. But it works pretty well in the simulation of hazard rates. More detailed discussion is in the next section.

4.9.2 Simulation for Hazard Rate

This part of simulation is based on the assumption that the calibrated interest rate parameters and its filtered time series are exact. Assume the firm specific factor $s_{j,t}$ follows CIR and its parameters are given as well as the market price of risk and the coefficient of the common factor. Thus the parameters we care about for each firm j are $(\kappa_j, \theta_j, \sigma_j, \lambda_j, \beta_j)$. Since we used EKF to calibrate the hazard rate processes, we want to know how well EKF works for non-linear problems. Another objective is also to recover the origial given parameters by QMLE in conjunction with EKF. A simulation to check this has been done in following ways.

Firstly, we use the given parameters to simulate one path of the firm specific factor $s_{j,t}$. Since we already assume the calibrated interest rate is exact, the hazard rate of the firm is totally determined by the time series of the firm specific factor. Then we simulate the prices of several coupon bonds using CIR pricing formula with the relationship between the zero-coupon bond prices and coupon bond prices. Next, we use the simulated bond prices to back out the time series of the firm specific factor. A comparison between the original and the filtered time series is given by the graphs. Finally we use QMLE with EKF to recover the origial parameters. The time step of the simulation is set to be monthly (1/12 yearly). Three coupon bonds are simulated using the coupons and maturities as the same as the Atlantic Richfield Company's.

Figures 4.15 and 4.16 show the original simulated time series of the firm specific factor and the filtered time series with different observation error R. From both of the graphs, we can tell EKF works fine for this non-linear problem. The accuracy does depend on the observation error R but not the same way as in the simulation of interest rate. Here the left hand side of our measurement equation is the bond price which is around 100. This is different from the case in the interest rate simulation in which the left hand side of the measurement equation is the yield which is around 0.06. Thus the error term is at least different in order. Figure 4.15 shows the difference with R = 0.01 and R = 1. The RMSE of the time series of the firm specific factor for R = 0.01 is 0.0030 while for R = 1 is 0.0021 which is smaller. This shows that the EKF does not necessarily become more accurate for less observation error. This realization is at odds with our common sence that the filtered results should be more accurate for more accurate meassurement. This is true in the linear case when the standard KF is used. However, if the model is non-linear, the linearization is not exact when the EKF is executed. Numerical errors may add up and have bad influence as in here. For Figure 4.16, both the RMSE for R = 1 and R = 10 are 0.0023. The accuracy does not change much if you use any number from 1 to 10. But RMSE decreases a lot if we put R = 100, which is around the bond price. Therefore, we use R = 1 to conduct our calibration for the parameters.

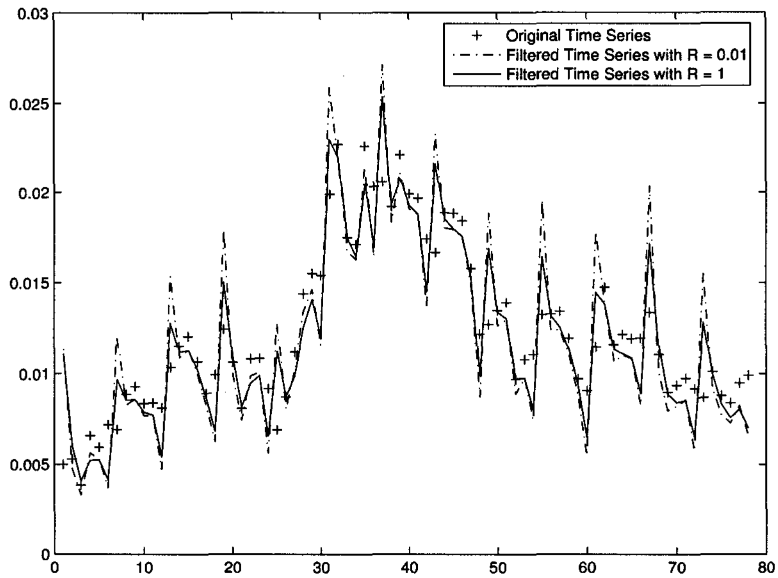

Figure 4.36: (1)Time series for the firm specific factor $s_{j,t}$ with different obervation error R. Parameters for this simulation are $\kappa_j = 0.3244$, $\theta_j = 0.005$, $\sigma_j = 0.0633$, $\lambda_j = -0.1587$, $\beta_j = -0.01$. Coupons are [4.565, 4.94, 5.44] dollars half-yearly. Initial time to maturities are [16.083, 21.083, 10.583] years respectively.

Another way to calibrate those parameters is to minimize the RMSE of the bond prices by simulation. Theoretically, minimizing the root of mean square error is the same as maximizing the likelihood function. However, numerically they are different here. We use the simulated bond prices to achieve our MRMSE, but we calculated the derivatives of the likelihood function in order to get the estimator in MLE. The RMSE works perfectly in the simulation world even for a small number of simulated points. Table 4.6 shows the calibrated results by different methods, where No and M have the same meaning as in the simulation of interest rate. The second row of the table is the original given parameters. The third row is the calibrated results by

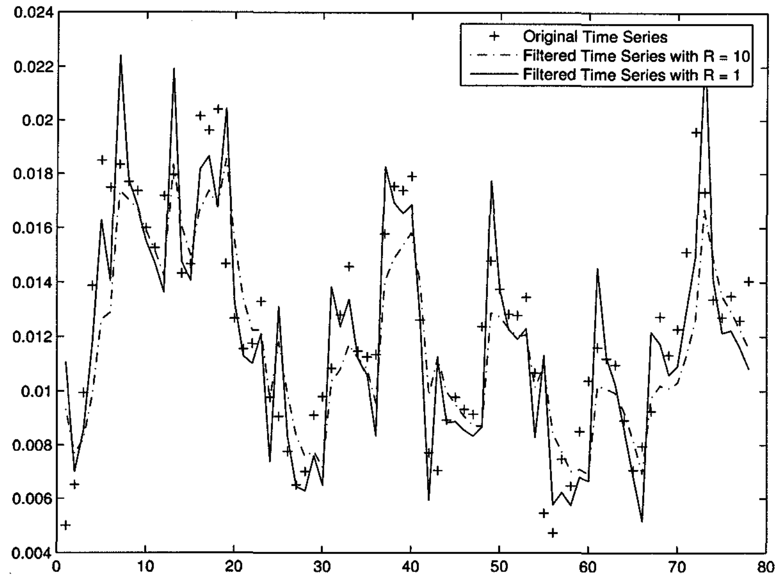

Figure 4.37: (2)Time series for the firm specific factor $s_{j,t}$ with different obervation error R. Parameters for this simulation are $\kappa_j = 0.3244$, $\theta_j = 0.005$, $\sigma_j = 0.0633$, $\lambda_j = -0.1587$, $\beta_j = -0.01$. Coupons are [4.565, 4.94, 5.44] dollars half-yearly. Initial time to maturities are [16.083, 21.083, 10.583] years respectively.

minimizing RMSE. The next row is the estimated results by QMLE with EKF. We can tell that the calibrated results by QMLE with EKF are not very good. This is due to several factors. It is maybe because we have not enough simulated points M, which we used here is only 78. The inaccurate linearization of the EKF may also have bad contribution for this. There is another one factor may have bad influence on the results. That is the degree of the freedom of the CIR. It is defined by $d = 4\kappa_j\theta_j/\sigma_j^2$ which is around 0.4 here. We know that if the degree of freedom d is less than or equal to 1, the process will hit zero with infinite many times with probability one. Many zero hazard rates will add the possibility of producing badly scaled matrix or matrices close to singular. Considering this, we can modify the discretized CIR a little bit by setting the point to θ_j whenever it hits zero. We call this method Modified CIR (MCIR). The last row of the table shows the results by MCIR with QMLE in conjunction with EKF.

	κ_j	θ_j	σ_j	λ_j	β_j
Original	0.3244	0.0050	0.0633	-0.1587	-0.0100
M=78, No=1 by RMSE	0.3172 (0.0346)	0.0051 (0.0100)	0.0637 (0.0656)	-0.1551 (0.0100)	-0.0100 (0.0346)
M=78, No=1 by QMLE	0.2708 (1.3630)	0.0035 (0.0100)	0.0490 (0.0900)	-0.1420 (0.7426)	-0.1086 (0.4287)
M=78, No=1 by MCIR	0.3423 (0.5283)	0.0036 (0.0001)	0.0561 (0.0173)	-0.1752 (0.2867)	-0.0755 (0.1005)

Table 4.6: Estimated Parameters by different approaches.

Unfortunately, we do not get a better off result by modifying the discretized CIR.

Although the minimizing RMSE approach (MRMSE) works nicely in the simulated world, this approach may fail when we use the real bond prices. We do not know exactly how the bond prices are produced in the real market. Let's think that there is a real mechanism which determines the bond prices. What we do is to model that mechanism based on our knowledge. It is impossible that our model will perfectly capture the properties of the real mechanism. Put it in another way that the real mechanism is different from our model. This let the minimizing bond prices approach fail. The reason why it works in the simulated world is pretty simple. Since the simulated bond prices are produced by the same mechanism of our model. Thus we choose QMLE in conjunction with EKF instead of MRMSE.

We can also use the filtered time series to back out the bond prices. If the calibrated parameters are exact for the firm specific factor, then the backed out bond prices should be almost exactly to the original simulated bond prices. This is because both bond prices are calculated using the theoretical bond pricing formula. Numerical result show that the average of the RMSE of the bond prices is less than 30 basis points which is very tiny considering the bond price is around \$100.

Chapter 5

Conclusion

This thesis studies how the theory of credit risk and its applications are combined together. For the theoretical part, a short review of credit risk theory is covered as well as interest rate theory. For the application part, modeling analysis is conducted and the calibrations of the model are implemented by using EKF in conjunction with QMLE.

In the calibration, we use one factor CIR model for default-free interest rate and two factors CIR model for the default intensities. Our data includes Treasury strips yields (monthly) of U.S for seven different maturities from Jan. 1993 to Dec. 2003 (Data source: Bloomberg) and month-end corporate coupon bond prices issued by 21 firms (total 82 bonds) from Jan. 1995 to Jun. 2001 (Data source: MDG Business School). The model (one factor CIR) for default-free interest rate was calibrated by using the yileds data as input and Standard Kalman Filter (SKF) in conjunction with Quasi-Maximum Likelihood Estimation (QMLE) as numerical scheme. The time series of the interest rate from Jan. 1993 to Dec. 2003 was obtained by the filter tool. Root of Mean Square Errors (RMSE) of the theoretical yields were also calculated. The model (two factor CIR) for the hazard rates for each firm was calibrated by using the observed corporate bond prices as input and Extended Kalman Filter (EKF) in conjunction with QMLE as numerical scheme. The time seires of the hazard rates for each of the 21 firms were filtered out. RMSEs of the theoretical corporate bond prices were calculated.

The results of default-free interest rate is not so good compared with two factor CIR model. The RMSE of the yields is around 0.0040, but Duffee (1999) claims it will be decreased to 0.0010 if two factor CIR model for the default-free interest rate is assumed. This could be one of the future studies. The filtered time series of the interest rate is very similar to Miu's (2003) result.

The calibrated parameters for most of the firms are reasonable, however there are seven firms whose parameters are not available. This is due to the drawbacks of the numerical scheme we have used. QMLE does not work very well when it is combined with EKF. Although the filtered time series does not depend on the small change of the initialized parameters very much, QMLE is very sensitive on the initial parameters. We do not have a prototype to select the initialization. If we choose a very bad parameter vector as a start, it may cause the MLE diverges or it may produce a singular matrix or matrices which are badly scaled. Details will be discussed in next section. How to improve the numerical scheme of the calibration may be another direction of future research. Figure 4.6 shows the time series of the yield to maturity with seven years, in which there is a jump around 50 on x-axis. I did not modify this bad data, which may also have bad contributions to the results.

Because we are using relatively a few bond data to fit each firm's term structure, the individual parameters estimated are subjected to a certain uncertainty. This is also noticed by Duffee (1999), in which he focuses his discussion on the median parameter across all 161 firms.

Notice that there are many negative β which causes the hazard rates to go down to zero sometimes. This implies that the CIR factors are not necessarily better than Hull-White factors. The 21 filtered time series for each firm display some properties in common. They look similar in general. Almost all of them firstly go up a little bit and then go down, and then go up and down. And they are dented in the middle. This shows that they are mostly positively correlated to each other. It also suggests that when the interest rate goes up, most hazard rates go down, capturing the negative correlation between the interest rate and the hazard rates. Surprisingly, the ones with positive β show the same dynamics. This is at odds to what we have expected, since positive β in hazard rate process will produce positve correlations with the interest rate. One explaination is that there are maybe some other hidden common factor shared only by the hazard rates. For further study, we could add another common factor for all the hazard rates only.

Monte Carlo Simulation is conducted both for interest rate and hazard rates. The simulation indicates that: both the SKF and the EKF work pretty well as a filter tool but may produce bad estimation for the value of the likelihood function; QMLE works fine in linear state space form model (such as interest rate model here), but it does a poor job in the case of non-linear state space form(such as hazard rates model here).

Bibliography

[1] Bjork, Tomas. *Arbitrage Theory in Continuous Time.* Oxford University Press, 1998.

[2] Black, Fisher and Scholes, Myron. *The Pricing of Options and Corporate Liabilities.* Journal of Political Economy 1973.

[3] Brandimarte, Paolo. *Numerical Methods in Finance.* Wiley Series, 2001.

[4] Casella, George and Berger, Roger L., *Statistical Inference.* Thomson Learning, 2002.

[5] Cox, John C., Ingersoll, Jonathan E. and Ross, Stephen A., *A theory of the term structure of interest rates.* Econometrica 53, 1985.

[6] Crosbie, Peter J. and Bohn, Jeffrey R., *Modeling Default Risk.* Moody's KMV, 2003.

[7] Chen, R.-R. and Scott, L., *Multi-Factor Cox-Ingersoll-Ross Models of the Term Structure: Estimates and Tests from a Kalman Filter Model.* Journal of Real Estate Finance and Economics, 2003.

[8] Dellacherie, Claude and Meyer, Paul A., *Probabilities and Potentials.* Elsevier Sience Pub. Co., 1982.

[9] Duffee, Gregory. *Estimating the Price of Default Risk.* Review of Financial Studies, 1999.

[10] Duffie, Darrell. *Dynamic Asset Pricing Theory.* Princeton University Press, 1996.

[11] Duffie, Darrell and Singleton, Kenneth J., *Modeling Term Structures of Defaultable bonds.* Princeton University Press, 1999.

[12] Duffie, Darrell and Filipovic, D. and Schachermayer, W., *Affine Processes and Applications in Finance.* 2001

[13] Duffie, Darrell and Lando, David. *Term Structure of Credit Spreads with Incomplete Accounting Information.* Econometrica, 2001.

[14] Flesaker, Bjorn, and Hughston, Lane. *Positive Interest.* Financial Derivatives, 1996.

[15] Geyer, Alois L.J., and Pichler, Stefan. *A state-space approach to estimate and test multifactor Cox-Ingersoll-Ross models of the term structure.* Journal of Financial Research, 1998.

[16] Giesecke, Kay. *Credit Risk Modeling and Valuation: An Introduction.* Credit Risk: Models and Management, Vol.2, D. Shimko (Ed.), Wiley, 2004.

[17] Giesecke, Kay. *Credit Derivatives and Structured Credit Products* (lecture notes). Cornell University, 2002.

[18] Giesecke, Kay and Goldberg, L., *Forecasting Default in the Face of Uncertainty.* Journal of Derivatives, 2004.

[19] Giesecke, Kay and Goldberg, L., *Default and Information* (working paper). Cornell University, 2003.

[20] Giesecke, Kay. *Asimple exponential model for dependent defaults.* Journal of Fixed Income, 2003.

[21] Glasserman, Paul. *Monte Carlo Methods in Financial Engineering: Introduction and Foundations.* Springer, 2001.

[22] Harvey, Andrew. *Forecasting, Structural Time Series Models and The Kalman Filter.* Cambridge University Press, 1990.

[23] Hamilton, James D., *Time Series Analysis.* Princeton University Press, 1994.

[24] Hughston, Lane and Turnbull, S.. *A simple derivation of some results for pricing credit derivatives.* King's College London, 2000.

[25] Hanke, M., *State Space Formulation of the Multi-Factor Cox-Ingersoll-Ross-Model.* 1998.

[26] Haykin, Simon. *Kalman Filtering and Neural Networks.* Wiley, 2001.

[27] Jarrow, Robert A. and Deventer, Donald R. van. *Financial Risk Analysis.* Kamakura Corporation, 1994.

[28] Jin, Y. and Glasserman, P., *Equilibrium Positive Interest Rates: A United View.* Review of Financial Studies, 2001.

[29] Lando, David. *On Cox processes and credit risky securities.* Review of Derivatives Research, 1998.

[30] Lund, Jesper. *Non-Linear Kalman Filtering Techniques for Term-Structure Models.* The Aarhus School of Business, 1997.

[31] Li, Tao. *Some Analytical and Numerical Results for Pricing Defaultable Bonds* (Masters Thesis). McMaster University, 2002.

[32] Miu, Peter. *Risk Management of a Credit Portfolio and the Pricing of Systematic Credit Risk.* 2003.

[33] Miu, Peter. *Performances of Alternative Reduced Form Models of Credit Risk: The Role of Fitting the Default-Free Term Structure.* 2001.

[34] Nelsen, Roger B., *An Introduction to Copulas.* Springer-Verlag Telos, 1998.

[35] Oksendal, Bernt K. *Stochastic Differential Equations.* 5th ed. Springer, 2000.

[36] Rogers, L.C.G., *The Potential Approach to the Term Structure of Interest Rates and Foreign Exchange Rates.* Mathematical Finance, 1997.

[37] Steel, J. Michael. *Stochastic Calculus and Financial Applications.* Springer, 2003.

[38] Schonbucher, Philipp. *Term Structure Modelling of Defaultable Bonds.* Review of Derivatives Research, 1998.

[39] Schonbucher, Philipp. *Credit Derivatives Pricing Models.* John Wiley and Sons, 2003.

[40] Shreve, Steven. *Stochastic Calculus and Finace* (electric notes). Carnege Mellon University, 1997.

[41] Tauren, Miikka. *A Comparison of Bond Pricing Models in the Pricing of Credit Risk.* Indiana University, 1999.

[42] Welch, Greg and Bishop, Gary. *An Introduction to the Kalman Filter.* University of North Carolina at Chapel Hill, 2004.

[43] Zhou, Chunsheng. *A Jump-Diffusion Approach to Modeling Credit Risk and Valuing Defaultable Securities.* Federal Reserve Board, 1997.