THE UNIVERSITY OF WARWICK

University of Warwick institutional repository: http://go.warwick.ac.uk/wrap

A Thesis Submitted for the Degree of PhD at the University of Warwick

http://go.warwick.ac.uk/wrap/36336

This thesis is made available online and is protected by original copyright.

Please scroll down to view the document itself.

Please refer to the repository record for this item for information to help you to cite it. Our policy information is available from the repository home page.

Three Models of the Term Structure of Interest Rates

Joonhee Rhee

Submitted in accordance with the requirements for the degree of Ph.D.

University of Warwick

Warwick Business School

August 1998

Acknowledgement

First of all, I wish to express my appreciation to my parents for giving me the opportunity to continue to study. Without their unending love, I could not have completed this thesis. My only anxiety is that for the last three years I may not have been a very good son.

I also wish to thank my supervisor Dr. Nick Webber. I have always been grateful to him for his excellent advice and helpful comments, without which my work would not have progressed.

I also thank Dr. Jack Broyles, who encouraged me whenever I was in despair and whose suggestions were always greatly appreciated

Professor Sang Yong Park of Yonsei University, Seoul, Korea supported me financially and psychologically from the start of my studies. I shall never forget his kindness.

I also thank Mr. Keith Taylor for English corrections of my thesis.

Last but not least, I cannot express the love I feel for to my son, Seungmin. Since he was born in May 1994, I have seen him only for a few month. But he is always in my thought.

Abstract

In this dissertation, we consider the stochastic volatility of short rates, the jump property of short rates, and market expectation of changes in interest rates as the crucial factors in explaining the term structure of interest rates. In each chapter, we model the term structure of interest rates in accordance with these factors.

CONTENTS

Lists of Tables

Lists of Figures

Chapter 1. Introduction

The term structure of interest rates plays a pivotal role in financial economics, from a simple NPV (Net Present Value) calculation to advanced options pricing. The interest rate, especially the short rate, is determined by complex economic forces. Models of the term structure of interest rates should reflect some, if not all, parts of this complexity. In this dissertation, we consider the stochastic volatility of short rates, the jump property of short rates, and market expectation of changes in interest rates as the crucial factors in explaining the term structure of interest rates. In each chapter, we model the term structure of interest rates in accordance with these factors.

The most natural way to model the term structure of interest rates in finance is to choose a general equilibrium framework. This approach involves a kind of microeconomic problem in a macroeconomic setting. Lucas (1978) and Sargent (1987) follow this approach in pure economics, and Cox, Ingersoll and Ross (1985b) and Longstaff and Schwartz (1992) do so in financial economics. Unfortunately, the approach has some limitations. The spot rate and the price of financial derivatives are expressed as an indirect utility function. When we implement a model from the general equilibrium framework, we have to specify or estimate the parameters of the utility function. Recently, Duffie (1992), Duffie and Epstein (1992) and Duffie and Lions (1992) have directly modelled the utility function in a continuous time framework. Hence they have considerably reduced the inconvenience associated with the general equilibrium modelling. However, they have not resolved all the problems involved in solving the associated Partial Differential Equation and, most importantly, in solving the problem of a utility function in the presence of jumps.

A very significant revolution in finance is represented by the application of martingale theory to check the no-arbitrage condition in pricing financial derivatives or modelling the term structure of interest rates. This step distinguishes finance from economics. If a model of financial derivatives satisfies the no-arbitrage condition, it is not necessary to model in a general equilibrium framework. Harrison and Kreps (1979) obtain this result, and Harrison and Pliska (1981) develop it further. They call this characteristic "viability", which means "always supportable" by a general equilibrium framework.

We apply this latter approach to model the term structure of interest rates in a continuous time framework. Three main aspects of the dissertation will be highlighted :

First, we present a three-factor affine model of the term structure of interest rates, in which the factors are the spread rate, its volatility and the long rate. The long rate that we consider in the first case is the consol rate. In the second case, the long rate used is a bench-mark rate affecting the level toward which the short rates converges. In the case of using the consol yield, we extend a two-factor model proposed by Schaefer and Schwartz (hereafter SS, 1984) to a three-factor case. Our model adds the stochastic volatility of the spread rates process. We also provide an approximate solution to the fundamental valuation equation using orthogonal state variables. In a similar context, we use a bench-mark rate affecting a level toward which the short rates converges as a factor. We shall explain this further in Chapter 4. We successfully provide a closed form solution of a pure discount bond price, using three factors in the latter case.

Secondly, we investigate the affine model of the term structure in the presence of jumps, and extend Duffie and Kan's (1996) model to the presence of jumps. We present approximated solutions for a pure discount bond price in a two-factor model. The first factor used in our model is a pure diffusion process, and the second factor is a pure jump process.

Thirdly, we investigate a model of the term structure of interest rates in the presence of expectations of changes in the regime of interest rates, or of bubbles. We express the model of the term structure of interest rates in terms of forward rate processes in a similar way to the frameworks of Brace, Gatarek, and Musiela (1997, BGM), and Jeffrey (1994). Assuming a Gaussian and time-homogeneous volatility structure, we present an explicit solution for forward rates in terms of fundamentals[1], and consequently for a pure discount bond price under the expectation of changes in regimes. In this procedure, we re-interpret the Sargent (1972)-type interest rate model in the continuous-time framework of modern finance.

The structure of the dissertation is as follows.

In Chapter 2, two approaches to pricing the financial assets are explained: the general equilibrium approach and the no-arbitrage approach (martingale approach). To clarify the concept of the martingale in finance, the martingale measure is expressed in terms of time and state preference. In addition, some important concepts in a stochastic model, such as the Radon-Nikodym derivative and the Girsanov theorem, are explained.

In Chapter 3, we explain the theory of the term structure of interest rates both in the general equilibrium model and the martingale context in Sections 2 and 3. In the general equilibrium context, we explain how to obtain a pure discount bond price and pricing kernel from the the Euler equation. In particular, the pricing kernel approach in finance is quite important because it overcomes the arbitrary specification problems of the market price of risk in the martingale approach. Following Jamshidian (1991), we explain the relationships between spot rate process, forward rate process, and bond price process in the martingale measure context (this relationship is used in Chapter 6). In Section 4, we explain Duffie and Kan's (1996) affine model of the term structure of interest rates. Two chapters of the dissertation (Chapters 4 and 5) are in fact based on the affine framework. We also review affine

1. We shall provide a formal definition of the fundamental in Chapter 6.

type models of the term structure of interest rates, such as that of Fong and Vasicek (1991). In Section 5, we explain how the three main Chapters 4, 5 and 6 can be integrated as one topic. The empirical distribution of changes in interest rates displays the leptokurtosis of interest rates. The distributional kurtosis can be explained by stochastic volatility (Chapter 4), the jumps of interest rates (Chapter 5), and the market expectation of interest rates (Chapter 6).

In Chapter 4, we investigate a three-factor affine model of the term structure, assuming stochastic volatility. We present an approximate solution of a pure discount bond price and a closed form solution of a pure discount bond price. Two models are examined. In the first model, the drift term of the consol rate is derived from the no-arbitrage condition. In the risk-neutral measure, this model is not an affine model. To obtain an approximate solution of a pure discount bond price, we assume that the drift term of the consol rate is linear with respect to the consol rate. In the second model, we use a long-term rate which affects the convergence of short rates as a factor. We successfully provide a closed form solution for a pure discount bond price. In an empirical estimation, our model is discretized in terms of a GARCH-X specification. In addition, we compare the model with Schaeffer and Schwartz's (1984) model, using the LR test.

In Chapter 5, we extend the affine model of the term structure of interest rates to the presence of jumps. We also present an approximate solution of a pure discount bond price in a two-factor case. First, we show how the affine framework can be extended to the jump diffusion case. In the next section, as examples, we demonstrate the two-factor jump-affine model of the term structure in two cases and obtain approximate solutions. In an empirical estimation, we use a maximum-likelihood method for the approximate density.

Recently, many models of the term structure of interest rates have had a macroeconomic foundation (Tice and Webber (1997)). Unfortunately, these models are based on the Keynesian IS-LM framework. We do not know whether they are viable in the sense described by Harrison and Kreps (1979). The Lucas-type

economy model is known to be viable in this sense. It is well-known, however, that the IS-LM model cannot be compatible with a Lucas-type economy. To reflect this, in Chapter 6, we combine the rational expectations model from monetary macroeconomics with a modelling of the term structure of interest rates. In Section 1, we explain the motivation of this chapter. In Section 2, we explain the Cagan-type monetary model, the Krugman and Miller (1992) target zone model, the Sargent-type interest rate model and the BGM model. We demonstrate how the Krugman and Miller framework can be applied to the Sargent-type interest rates model. We express the model of the term structure of interest rates in terms of the BGM framework. In Section 3, we choose a bubble path as a solution.

To obtain tractability or a closed form solution for interest rate derivatives, we sometimes permit the interest model to allow negative interest rates: for instance, the Gaussian model and our two-factor jump-diffusion model. In Appendix III, we construct a theoretical model of the term structure of interest rates in the presence of jumps to see what kinds of term structure model do not allow negative interest rates. Following Flesaker and Hughston (1996), the interest rate positivity property is incorporated into the discount bond price.

Chapter 2. Asset Pricing in Financial Economics: General Equilibrium and Martingales

Uncertainty in financial economics plays a pivotal role in determining asset pricing. If there is no uncertainty, asset pricing is very simple. For instance, in the discrete time framework under certainty, any asset price is just a summation of known discounted future cash flows. In reality, however, the future cash flows and discount factors are unknown. So, how can we reasonably reflect uncertainty in modelling asset prices? The straightforward way to do this is to assume uncertainty in the form of a probability density. Modelling uncertainty as a normal distribution (Gaussian) is the usual approach in financial economics as well as in econometrics. There are several reasons for doing this. The uncertainty can be represented as symmetric about the mean. More fundamental is the advantage of analytical simplicity. Both the unconditional and conditional densities of the normal distribution are normal distributions. In addition, it is rather easy to extend a model to the multivariate case. For instance, when we analyse the relation between the risk premium of the market portfolio and the investors' optimal portfolio decision, we can use Stein's lemma.[2] However, this can be applied only to the normal distribution. Unfortunately, however, the normal distribution assumption does not conform closely to reality. It is well-known that stock returns are more nearly lognormally distributed, and that the underlying stochastic process of this and other economic variables, e.g. interest rates, exhibits skewness and excessive kurtosis

2. Let X and Y be bivariate normally distributed. Then we have

$$\mathrm{Cov}(g(X), Y) = E[g'(X)]\ \mathrm{Cov}(X,Y)$$

where g is differentiable. This is called Stein's lemma.

(leptokutosis). This means that the distribution is not symmetric and that the return distribution of some assets is fat-tailed. Accordingly, modelling the uncertainty of some assets requires careful attention, and must take into consideration the empirical evidence of the distribution of asset returns.

On the other hand, the continuous-time framework has been a fundamental tool since Merton's (1971) and Black-Sholes' (1973) option pricing. The continuous-time approach is one way of reducing the inaccuracy resulting from the discrete-time approach[3]. The continuous-time approach has the strong advantage that it expresses uncertainty in the form of a stochastic differential equation instead of as a distribution. Actually these two approaches are the same. In pricing a stock option, the stock price is assumed to follow a lognomal distribution. Equivalently, we could express it as a geometric Brownian motion. When we try to find the density function from the stochastic differential equation, we have to solve the Kolmogorov forward equation. Unfortunately, closed-form solutions are few. In some cases, for instance in Sun (1992), we can to solve a stochastic difference equation. However, solving a differential equation is easier than solving a difference equation. The attraction of the continuous-time approach is mainly due to the simplicity of modelling asset pricing.

One of the most important papers in asset pricing is Lucas (1978), which offers a theoretical examination of the stochastic behaviour of equilibrium asset prices in a pure exchange economy. Lucas investigates the relationship between the exogenously determined productivity changes and market movements in asset prices. Futhermore, he generalizes the martingale property of stochastic asset prices, which is a characteristic of market efficiency (we will discuss the martingale theory in the next section). Lucas models the uncertainty of productivity (in other words, changing the opportunity set) as a probability distribution rather than as a stochastic differential equation. In contrast to Lucas, Merton (1971) derives a relationship

3. Expressing uncertainty in the form of the stochastic difference equation is also possible. Sun (1992) uses a stochastic difference equation approach in analyzing the general equilibrium theory of the term structure of CIR (1985). In his paper, the price formula of a pure discount bond converges to that of CIR (1985).

between the equilibrium expected rates of return on assets in a series of stochastic differential equations. He extends a classical CAPM to the multi-factor CAPM in a continuous-time general equilibrium model. Synthesizing these two approaches, Cox, Ingersoll and Ross (hereafter, CIR,1985a, 1985b) published a pathbreaking paper. CIR develop a general equilibrium asset pricing model in the continuous-time framework of Merton (1971) and in the Lucas-type economic structure. In CIR, the model endogenously determines the stochastic process followed by the equilibrium price of any financial asset and shows how this process depends on the underlying real variables, i.e. for a term structure of interest rates, the economy determines the type of stochastic process for interest rates. This is different from the no-arbitrage approach of term structure modelling used by, for example, Vasicek (1977). One of the principle results of CIR (1985a, 1985b) is a derivation of a partial differential equation which asset prices must satisfy. The solution of the equation determines the equilibrium price of an asset in terms of the underlying real variables.

Asset pricing models must not allow arbitrage opportunities. The papers mentioned above derive asset pricing in the general equilibrium context. This approach automatically guarantees the no-arbitrage condition. In finance, there is another way to derive asset pricing models which does not admit to arbitrage opportunities.

Asset pricing in a general equilibrium context starts from the assumptions of the utility function and a budget constraint. Eventually, we have to solve the Hamilton-Jacobi-Bellman equation. The solution (the investment proportions of the riskless asset and the risky asset) is obtained in terms of an indirect utility function. It requires quite strong assumptions such as a specific utility function and the market clearing condition. The no-arbitrage approach seeks to develop pricing models in an economy with risk aversion and time preference. The no-arbitrage approach assumes only that "people prefer more to less" and is not concerned with the utility functions of individual investors.

In order to clarify the concept of the martingale, which is frequently mentioned

in this thesis, we briefly sketch asset pricing in a no-arbitrage economy[4]. To explain some concepts in a more intuitive way, we use both a continuous-time and a discrete-time approach throughout this thesis. Assume initially, 1) risk neutral investors, 2) no time preference, 3) state space Ω with a finite number of states, 4) information structure (filtration), $\xi = \{\xi_t : t=0,1,\ldots,T\}$, $\xi_0 = \{\Omega\}$, 5) N+1 securities (the first one is a riskless security), 6) cash flows of security j : $X_j = \{X_j(t): t=1,\ldots,T\}$, where $j=0,1\ldots. \quad N+1$. Then the ex-dividend prices S_j of a riskless security and the risky securities are

$$
\begin{aligned}
(1) \qquad S_j &= \{S_j(t);, t=1,\ldots,T\} \\
S_0(t) &= B(t) \\
S_j(t) &= E(S_j(t+1)+X_j(t+1) \mid \xi_t\}
\end{aligned}
$$

where $B(t)$ is a T-period discount bond with face value equal to one, assuming that there are no coupons. As the price processes are ex-dividend, $B(T)=0$, $S_j(T)=0$. Following Huang and Litzenberger (1988), we define the accumulated dividend process for security j to be

$$D_j(t) = \sum_{s=0}^{t} X_j(s), \text{ for all } t=0,1\ldots T$$

Adding $D_j(t)$ to both sides of the third line of (1) gives

$$
\begin{aligned}
(2) \qquad S_j(t)+D_j(t) &= E(S_j(t+1)+D_j(t+1) \mid \xi_t\} \\
&= E(S_j(s)+D_j(s) \mid \xi_t\}, \text{ for all } s > t
\end{aligned}
$$

where the second line of (2) follows from repeated substitution of the first line of (2)

4. This is based on lecture note of Subrahmanyam (1996) in an EDEN Seminar.

into itself. Thus, in the absence of risk aversion and time preference, prices plus accumulated dividends are martingales. No-arbitrage, therefore, is automatically guaranteed.

Then, what happens when the representative economic agent has a time preference? We can express the asset price in line with the marginal utility of the representative agent[5]. Then,

(3)
$$B(t) = E\left(\frac{U'_{t+1}(C_{t+1})}{U'_t(C_t)}(B(t+1)) \mid \xi_t\right)$$
$$S_j(t) = E\left(\frac{U'_{t+1}(C_{t+1})}{U'_t(C_t)}(S_j(t+1)+X_j(t+1)) \mid \xi_t\right)$$

for all $t < T-1$, and where U denotes the utility function, U' the marginal utility, and C_t the consumption at time t. As in the expression (3), in the presence of time preference, the spot price is not a martingale, even with the assumption of risk neutrality. To make the spot price a martingale, we rewrite (3) as

(4)
$$S_j(t) = E\left(\frac{U'_{t+1}(C_{t+1})}{E(U'_{t+1}(C_{t+1}))} \; \frac{E(U'_{t+1}(C_{t+1}))}{U'_t(C_t)}(S_j(t+1)+X_j(t+1)) \mid \xi_t\right)$$

Defining,

(5)
$$\phi(t+1) = \frac{U'_{t+1}(C_{t+1})}{E(U'_{t+1}(C_{t+1}))}$$

where ϕ is interpreted as a risk adjustment factor. Substituting (5) into (4), we obtain,

(6)
$$S_j(t) = E\left(\frac{E(U'_{t+1}(C_{t+1}) \mid \xi_t)}{U'_t(C_t)} \; \phi(t+1)\;(S_j(t+1)+X_j(t+1)) \mid \xi_t\right)$$

5. See *Foundation of Financial Economics* by Huang and Litzenberger (1988), p. 227.

$$= E((\frac{B(t)}{B(t+1)})\ \phi(t+1)\ (S_j(t+1)+X_j(t+1))\ |\ \xi_t)$$

where the second line of (6) comes from the first line of (3). Let $S_j^*(t)=\frac{S_j(t)}{B(t)}$ and $X_j^*(t)=\frac{X_j(t)}{B(t)}$. For instance, setting $N=1$, $T=1$, then we obtain:

(7) $$S^*(0) = E(\phi\ (S^*(1)+X^*(1))\ |\ \Omega)$$

$$= \sum_{w\in\Omega} \pi_w \phi_w\ (S^*(1)+X^*(1)).$$

We define the conditional probability,

(8) $$\pi_w^*(t) = \pi_w(t)\phi_w(t+1)$$

where π is the conditional probability at time t, given the filtration ξ_t. From (7), we can obtain the following relationship:

(9) $$E(\phi\ |\ \xi_t) = \sum_{w\in\xi_t} \pi_w^*(t) = \sum_{w\in\xi_t} \pi_w(t)\phi_w(t+1)$$

and, using the definition of ϕ in (5), we find that equation (9) becomes equal to 1. Hence π_w^* behaves like a probability. It is called the Equivalent Probability Measure (Huang and Litzenbeger, 1988). We rewrite equation (6) as

(10) $$S_j^*(t) = E(\phi(t+1)\ (S_j^*(t+1)+X_j^*(t+1))\ |\ \xi_t)$$

$$= \sum_{w\in\xi_t} \pi_w(t)\phi_w(t+1)\ (S_j^*(t+1)+X_j^*(t+1))$$

$$= \sum_{w\in\xi_t} \pi_w^*(t)\ (S_j^*(t+1)+X_j^*(t+1))$$

$$= E^*((S_j^*(t+1)+X_j^*(t+1))\ |\ \xi_t).$$

We obtain an expression in the same form as the third line of equation (1). The security price plus the accumulated dividend are martingales under the new probability measure. Since this measure makes the security price into a martingale process, this probability measure is called the Equivalent Martingale Measure (EMM).

We derive the EMM from the assumption of risk aversion and time preference. However, the security price under EMM behaves like equation (1). Equation (1) is based on the assumption of risk neutrality. Hence the investor is risk neutral under EMM. This is why many authors, including Duffie (1992), call EMM the risk neutral measure. However, risk aversion and time preference are captured by the EMM distribution shift from the original distribution. This martingale measure π_w^* is just the Arrow-Debreu (1954) state price vector. Actually, no-arbitrage is guaranteed if and only if there is a state price vector[6]. Since the unique existence of a state price vector means the unique existence of EMM, there is no arbitrage if and only if there exists an EMM. Martingale technology in asset pricing has the same result as the general equilibrium framework but is more straightforward. This is why many asset pricing models are derived from the martingale technology. Harrison and Kreps (1979), Harrison and Pliska (1981), and Huang (1987) are included in this category of authors. However, as Harrison and Kreps point out, financial models from the martingale approach are viable, which means "always supportable" by the general equilibrium model.

How can we be sure of the unique existence of EMM? To understand this, we resort to the Riesz Representation Theorem.

The Riesz Representation Theorem (RRT)[7] : Given a vector space L, with an inner product defined by $(x, y) = E_{Emm}[\ x \cdot y\ ]$, then for each linear functional $F: L \rightarrow \Re$, there is a unique π in L, called the Riesz representation of F, such that,

6. See the proof in Duffie (1992), p. 3.
7. See Duffie (1992), p. 25.

$$F(x) = E_{EMM}[\ \pi \cdot x\ ] \quad \text{where } x \in L.$$

If F is strictly increasing, then π is strictly positive.

We show how the RRT can be applied. Following the notation of the existing literature, we denote the original probability by P and the EMM by Q. As in Harrison and Pliska (1981), we define a price system for an asset, in particular a contingent claim, to be a map $F: L \rightarrow [\ 0, \infty)$. Then by RRT, there is a one-to-one correspondence between the price system F and probability measure Q i.e., $F(x) = E_Q(\sum_{t=0}^{T} \pi_t x_t)$. As in Duffie (1992), the π is called a state-price deflator. If L is assumed to be dense, then the space L is understood to be a Banach space[8].

The RRT provides the unique existence of EMM or probability measure Q. The next problem is how we find the EMM or probability measure Q. To understand this, we need two mathematical concepts: Radon-Nikodym derivatives and the Girsanov Theorem[9]. We first introduce the Radon-Nikodym Theorem.

Theorem (Radon-Nikodym)[10]: If the EMM Q is absolutely continuous with respect to the original probability measure P, then there is a non-negative random variable ρ such that for any $A \in \xi$, $Q(A) = \int_A \rho dP$. The random variable ρ is unique with P-probability one.

The random variable ρ is called the density of the probability measure Q with respect to probability measure P, or the Radon-Nikodym derivative of Q. We

8. Banach space is a complete normed vector space.
9. See Dothan (1990), p. 208.
10. ee Billingsley (1995), p. 423.

denote this by $\rho = \frac{dQ}{dP}$[11]. Actually, since the Radon-Nikodym derivative ρ is a kind of likelihood ratio, it should not deviate away from 1 if the two probability measures are to be equivalent. In the preceding theorem, if, for instance $E(\rho) = 1$, then we can create an EMM Q from the original probability measure P by defining $Q(A) = E(1_A\rho)$ for any event A where 1_A is a characteristic function, i.e. if event A happens, then $1_A = 1$, otherwise, $1_A = 0$. Then, for any random variable X, $E_Q(X) = E_P(\rho X)$ and the condition $Q(A) > 0$ is satisfied whenever $P(A) > 0$, and vice versa. Futhermore, the probability space should be equivalent. This means that P and Q have the same events of probability 0. Intuitively, for instance, when today's stock price is £1, getting, say, £1000 tomorrow is very improbable. If the probability of £1000 is 0 under probability P, it is also 0 under Q. Generally, if ς is a sub-information set of ξ and Q is equivalent to P, then

$$E_Q(X|\varsigma) = \frac{E_P(\rho X|\varsigma)}{E_P(\rho|\varsigma)}.$$

We are now in a position to state the Girsanov Theorem. The Theorem provides a martingale process under the new probability measure Q from an original process under the old probability measure P in *a stochastic continuous time process* such as the *Itô* process. As mentioned above, the new probability measure Q is an EMM as well as a risk neutral measure. Hence, any processes under Q admit no-arbitrage opportunities even without reference to general equilibrium.

A more general version and rigorous proof of the Girsanov Theorem is available in Oksendal (1995). We just apply the result to an *Itô* process.

11. The Radon-Nikodyum derivative is a martingale.

Girsanov Theorem : Let x be an *Itô* process in $\Re^n$

(11) $$dx_t = \mu_t\, dt + \sigma_t\, dz$$

where dz is a standard Brownian motion in $\Re^d$, and where μ: $\Re^n$ x $[\ 0, \infty) \rightarrow \Re^n$ and σ: $\Re^n$ x $[\ 0, \infty) \rightarrow \Re^{nxd}$. Suppose v is an n dimensional vector process in some Banach space[12] L^1 such that $\sigma_t \theta_t = \mu_t - v_t$. Then there exists an equivalent martingale measure Q such that

(12) $$d\hat{z}_t = dz_t + \theta_t\, dt$$

defines a standard Brownian motion $\hat{z}$ in $\Re^d$ on the probability measure space (Ω, ξ, Q) and, then

(13) $$dx_t = v_t\, dt + \sigma_t\, d\hat{z}.$$

Finally, for any random variable z,

$$E_Q(z) = E_P(\rho_T\, z)$$

where $\rho_t = \exp^{(-\int_0^t \theta_s dz_s - \frac{1}{2}\int_0^t \theta_s \theta_s ds)}$.

Intuitively, the Girsanov Theorem states that if we change the drift term of a given *Itô* process, then the probability law of the process does not change dramatically. As mentioned above, the probability law of the new process is absolutely continuous with respect to the law of the old process (by Radon-Nikodym). Futhermore, the new process is a martingale process.

12. Duffie (1992) defines v in L^1. This is the largest space of the Banach spaces.

The Riesz Theorem and the Girsanov Theorem provide an insight and justify the use of martingale technology in asset pricing. Most asset pricing models have been based on the continuous-time framework under the no-arbitrage condition. Since the theory of martingales and semimartingales in finance is quite lengthy, we do not explain the semimartingale and its application further. Among many papers, Harrison and Pliska (1981), Madan (1988) and Back (1991) explain the general theory of semimartingale in an intuitive way. More detailed explanations, including many proofs, are available in Dothan (1990), Chung and Willams (1990), and Protter (1996).

Chapter 3. The Theory of the Term Structure of Interest Rates and a Review of the Existing Literature

3.1. Introduction

Prior to the development of the general equilibrium approach or the risk-neutral valuation technique (martingale approach), the theory of interest rates seems to have been derived primarily from portfolio theory or basic asset pricing. For instance, Stiglitz (1970) obtained the demand for financial assets and bonds in one-and two-period equilibrium models.

To clarify some important relationships between the martingale approach (or pricing kernel) and the general equilibrium approach in models of the term structure of interest rates, we will explain two approaches to the theory of interest rates in this chapter. This is crucial for expressing the market price of risk in terms of a utility function. In addition, the risk-neutral approach in Section 3.2 is essential to understanding Chapter 6. The term structure of interest rates in this chapter deals with a zero coupon bond or discount bond price. As shown in Jamshidian (1989), a coupon bond can be decomposed into a portfolio of zero coupon bonds, each with a face value equal to the coupon, as if it were stripped from the coupon bond and priced separately.

3.2. The General Equilibrium Approach

We will start with the Euler equation, which is similar to the result in Chapter 2, but slightly different in terms of the notation and the asset type. In a standard setting, as presented in Samuelson (1969), Merton (1971), Rubinstein (1976), Lucas (1978) and Breeden (1986), representative economic agents are assumed to base their consumption and portfolio selection on the maximization of the sum of time additive expected utilities of consumption at each time t in known state $\overline{\omega}$, $u(c_t, \overline{\omega})$. A necessary condition for an interior maximization at time t is that the decrease in an agent's satisfaction from selling a marginal pound's worth of assets at t and thereby giving up expected future consumption, must equal the increase in marginal utility from consuming the proceeds at t. Assuming that the sold asset is the T maturity bond at price $P(t, T, \overline{\omega})$, then the marginal increase in satisfaction of consuming the proceeds from the sale is $P(t, T, \overline{\omega})u'(c_t, \overline{\omega})$, where $u'(.)$ is the marginal utility with respect to consumption. The expected loss in utility from selling the bond is the expectation of its liquidation value $P(t+1, T, \omega)$ at t+1, times the marginal utility from consuming the uncertain proceeds $u'(c_{t+1}, \omega)$ at time t+1. Mathematically, this can be expressed as

$$(1) \qquad P(t, T, \overline{\omega})u'(c_t, \overline{\omega}) = \int_{w \in \Omega} P(t+1, T, \omega)u'(c_{t+1}, \omega)\pi(\omega)\,d\omega$$

where $\pi(\omega)$ is the probability density function of state ω at time t+1, and Ω is the set of all states. By definition, assuming that $\{w\}$ is an ordered set, for instance, the real numbers, then the density π is given by

$$\mathrm{Prob}(\ \omega_{t+1} \le \omega | \omega_t = \overline{\omega}\ ) = \int_{-\infty}^{\omega} \pi(u, \overline{\omega}\ )du.$$

We suppress the notation dependency on $\overline{\omega}$. We define $\phi_\omega(t, t+1, \overline{\omega}) = \dfrac{u'(c_{t+1}, \omega)}{u'(c_t, \overline{\omega})}$, which can be interpreted as the price at which, at the margin, time t consumption in state $\overline{\omega}$ is traded off against time t+1 stochastic consumption. The definition of ϕ_w here is different from that in Chapter 2. Using the definition of ϕ_ω, equation (1) can be rewritten as :

$$(2) \qquad P(t, T, \overline{\omega}) = \int_{w \in \Omega} [\ \phi_\omega(t, t+1, \overline{\omega}) P(t+1, T, \omega)\ ]\ \pi(\omega)\, d\omega\ .$$

This is what we call the Euler equation. We define

$$\pi^*(\omega) = \phi_\omega(\ t, t+1, \overline{\omega}\ )\pi(\omega)$$

where π^* is non-negative and $\int_{w \in \Omega} \pi^*(\omega)\, d\omega = 1$. Since $\pi^*(\omega)$ can be regarded as a probability measure, we can write

$$(3) \qquad P(t, T, \overline{\omega}) = E_{\pi^*}[\ P(t+1, T, \overline{\omega})]\ .$$

The transformation from probability measure π to π^*, which makes the discount bond price a martingale as seen in (3) under the new probability measure, is relatively straightforward in the case of the Von Neumann-Morgenstern utility function. The transformation involves just weighting the probability π of each outcome by the marginal utility of wealth $\phi_\omega(t, t+1, \overline{\omega})$ of that outcome.

To find a risk neutral measure, we define $Y(t, \overline{\omega})$ to be the one-period riskless return. Since the one-period riskless return $Y(t, \overline{\omega})$ is known at time t, equation (3) becomes

(4) $$P(t, T, \overline{\omega}) = \frac{1}{Y(t, \overline{\omega})} \int_{w \in \Omega} [\ Y(t, \overline{\omega})\phi_{\omega}(t, t+1, \overline{\omega})P(t+1, T, \overline{\omega})\ ]\ \pi(\omega)d\omega$$

We define a new measure

$$\pi^{**}(w) = \phi_w \pi(w) Y(t, \overline{w})$$

then, from (4)

(5) $$Y(t, \overline{w}) = E_{\pi^{**}}[\ \frac{P(t+1, T, \overline{\omega})}{P(t, T, \overline{\omega})}\ ]\ .$$

That is, the bond's expected return computed using the probability measure π^{**}, just as it would be if investors were risk neutral, equals the riskless return. The measure π^{**} is called the risk neutral measure.

To explain the concept of the pricing kernel, the pricing formula (5) can be expressed under probability measure π as

(6) $$P(t, T, \overline{\omega}) = E_{\pi}[\ \phi_{\omega}P(t+1, T, \overline{\omega})\ ]\ .$$

Dividing both sides (6) by $P(t, T, \overline{\omega})$ gives

(7) $$1 = E_{\pi}[\ \phi_w P^*_{\omega}]$$

where $P^*_{\omega} = \frac{P(t+1, T, \overline{\omega})}{P(t, T, \overline{\omega})}$.

Cox, Ingersoll and Ross (1985b), and Breeden (1986) have derived endogenous bond returns P^*_{ω} with optimal paths of consumption and ϕ_w in the model where a more detailed specification about exchange and production uncertainty is superimposed. In

these models, the real bond yields are positively related to production and consumption growth rates, and are negatively related to uncertainty about future production opportunities. On the other hand, Constantinides (1992) proposed that the ϕ_w, which is called the pricing kernel (we call this approach the semi-equilibrium approach), should be modelled directly as a stochastic process. This definition is slightly different from that of Sargent (1987), who defines the pricing kernel as $\phi_\omega \pi(\omega)$[13]. Das and Foresi (1996) follow the pricing kernel approach in deriving a one-factor jump-diffusion model of the term structure of interest rates. As Constantinides (1992) points out, the direct time-series representation for the pricing kernel makes the procedure of equilibrium specification unnecessary.

This is the main framework of the term structure model of interest rates in the general equilibrium setting.

3.3. The Risk Neutral Approach

Instead of modelling the term structure of interest rates in the general equilibrium context, we can directly model any asset price under the risk-neutral measure or martingale measure. This implies that pricing modelling can be considerably simplified. If we try pricing an asset in a general equilibrium setting, we have to define a utility function and budget constraint, etc. However, as Duffie (1992) pointed out, a martingale is a black box, which makes specifying these unnecessary.

In this section we explain the modern framework of the model of the term structure of interest rates in the continuous-time setting. We also express the term structure model in accordance with the HJM (1992) framework. For simplicity we confine the story to a single-factor Gaussian model in which the uncertainty is

13. See *Dynamic Macroeconomic Theory* by Sargent (1987), p. 111.

driven by a one-dimensional Brownian motion. In particular, to simplify our analysis, we consider the process Markovian in some cases. An intuitive extension and application to the multi-factor case is available in Karoui and Lacoste (1992).

To explain three seemingly different modellings of the term structure model of interest rates (see equations 9-1, 9-2, and 9-3), we start with some notation. Let $P = P(t, T)$ and $f = f(t, T)$ be the price of a T maturity zero coupon bond at time t and the instantaneous forward rate at time t, respectively. Then the following relations are true:

(8)
$$P(t, T) = \exp\left(-\int_t^T f(t, s)ds\right)$$
$$f(t, T) = \frac{-\partial[\ \ln P(t, T)]}{\partial T}$$
$$P(t, t) = 1, \quad r(t) = f(t, t)$$

where $r(t)$ is the spot rate. We assume that the instantaneous spot rate, the drift and volatility of bond price, and the market price of risk which shall be defined later are state-independent.

Assuming that the bond price, forward rate and short rate follow *Itô* processes, then:

(9-1)
$$\frac{dP}{P} = \mu(t, T)dt - \sigma(t, T)dz$$

(9-2)
$$df = \alpha(t, T)\ dt + v(t, T)\ dz$$

(9-3)
$$dr = \alpha(t)dt + v(t)\ dz,$$

following the conventional notation for bond price process, we take minus sign in the diffusion term in equation (9-1). Using *Itô*'s lemma, the following relations are satisfied:

(10-2) $$\alpha(t, T) = \frac{-\partial\mu(t, T)}{\partial T} + v(t, T)\sigma(t, T),$$

(10-3) $$v(t) = v(t, t), \ \sigma(t, t) = 0, \text{ and}$$

(10-4) $$\mu(t, T) = r(t) - \int_t^T \alpha(t, s)ds + \frac{\sigma(t, T)^2}{2}.$$

(10-5) $$\alpha(t) = \frac{\partial f(t, T)}{\partial T}\Big|_{T=t} - v(t)\lambda(t)$$

where $\lambda(t)$ is the market price of risk. Equation (10-4) is obtained by integrating equation (10-2) with respect to T, and using the fact that $\sigma(t, t) = 0$ and $\mu(t, t) = r(t)$. The proof is available in Hull and White (1993), and Jamshidian (1991).

The ratio $\lambda(t) = \frac{\mu(t, T) - r(t)}{\sigma(t, T)}$, which is independent of T for all $0 \le t \le T$, is called the market price of risk. Hereafter, we omit the time t for $\lambda(t) = \lambda$. Then the pure discount bond price at time t is given by

$$P(t, T) = E_t^Q \exp(-\int_t^T r(s)ds)$$

where the probability measure Q is a risk neutral measure, and by the Girsanov Theorem, $d\hat{z} = dz - \lambda dt$. Using the money market account $\beta(t) = \exp(\int_0^t r(s)ds)$ as the numeraire[14], we can obtain the new process,

$$P^*(t) = P(t, T)\exp(-\int_0^t r(s)ds).$$

14. Any asset in the model whose price is always strictly positive can be taken as the numeraire. We then denominate all other assets in units of this numeraire. The money market account $\beta(t)$, for instance, could be the numeraire. At time t, the bond is worth $\frac{P(t, T)}{\beta(t)}$ units of money market. We could also use the T-maturity bond as the numeraire.

Since this discounted process follows $\frac{dP^*}{P^*} = (\mu(t, T) - r)dt - \sigma(t, T)dz$, $\mu(t, T) = r + \sigma(t, T)\lambda$ holds if and only if $\frac{dP^*}{P^*} = -\sigma(t, T)d\hat{z}$. Hence, the discounted process is a martingale with respect to the probability measure Q. As Jamshidian (1991) points out, it is usually difficult to evaluate the expected discounted payoff for complex derivatives.

On the other hand, if we can assume that the spot rate follows diffusion (9-3), then using the no-arbitrage condition, we can obtain the fundamental PDE which is satisfied by bond prices or the prices of interest rate derivatives. Vasicek (1977), Brennan and Schwartz (1979), Fong and Vasicek (1991), and our three-factor model which is introduced in Chapter 4 are included in this category. In this approach, the arbitrage-free interest rate model is determined by the spot rate and the market price of risk.

Equivalently, we can express the term structure of interest rates as forward rates in terms of the initial yield curve, the bond return volatility, and the market price of risk (9-2). This approach was first introduced by Heath, Jarrow and Morton (HJM) (1992). To see the HJM framework, the no-arbitrage condition $\mu(t, T) = r + \sigma(t, T)\lambda$ is applied. Differentiating this condition with respect to T, then from (10-1) and (10-2), we obtain $\alpha(t, T) = v(t, T)(\sigma(t, T) - \lambda)$ (In particular, $\lambda = -\frac{\alpha(t, t)}{v(t)}$, evaluating at $T = t$)[15]. Thus (9-2) becomes

15. By assumption, the market price of risk $\lambda(t)$ is independent of T. Here we show why two definitions of the market price of risk are equivalent. To do this, we find a limit value of $\lim_{T \to t} \frac{\mu(t, T) - r(t)}{\sigma(t, T)}$. To find the value, we use L'hospital rule. Differentiating both the numerator and denominator with respect to T, we obtain

$$\lim_{T \to t} - \frac{(\alpha(t, T) - \sigma(t, T)v(t, T))}{v(t, T)}$$

where the numerator comes from (10-3). using $\sigma(t, t) = 0$, we can obtain the limit value $-\frac{\alpha(t, t)}{v(t)}$.

(11) $$df = v(t, T)[\ (\sigma(t, T) - \lambda)dt + dz\ ]$$

$$= v(t, T)[\ \sigma(t, T)\ dt + d\hat{z}\ ]$$

$$= v(t, T)\ d\bar{z}$$

where $d\bar{z}$ is what we call the forward risk-adjusted probability measure, or T-measure defined as

$$d\bar{z} = (\sigma - \lambda)\ dt + dz = \sigma\ dt + d\hat{z}\ .$$

Integrating the second line of (11), we find

(12) $$f(t, T) = f(t_0, T) + \int_{t_0}^{t} v(s, T)[\ \sigma(s, T)ds + d\hat{z}]\ .$$

This is the HJM term structure model. As seen above, given the market price of risk, the initial yield curve and the volatility structure determine the arbitrage-free term structure model.

Following Jamshidian (1991), in order to examine the volatility structure, the following two volatility functions are defined:

(13-1) $$x(t_0, t) = \sqrt{\int_{t_0}^{t} v(s, t)^2 ds}$$

(13-2) $$y(t_0, t, T) = \sqrt{\int_{t_0}^{t} (\sigma(s, T) - \sigma(s, t))^2 ds}.$$

where $0 \le t_0 \le t \le T$. Assuming coefficients are Markovian, we can obtain the following relation[16] from equation (2-2) in Carverhill (1994):

(14) $$\sigma(t, T) = \frac{v(t)}{v(s, t)}(\sigma(s, T) - \sigma(s, t))$$

where $0 \le s \le t \le T$. After rearranging, squaring and integrating (14) with respect to s, we obtain,

$$\int_{t_0}^{t}(\sigma(s, T) - \sigma(s, t)))^2 ds = (\frac{\sigma(t, T)}{v(t)})^2 \int_{t_0}^{t} v(s, t)^2 ds.$$

Hence,

(15) $$y(t_0, t, T) = (\frac{\sigma(t, T)}{v(t)})x(t_0, t).$$

since $\sigma(t, T)$ and $v(t)$ are always assumed to be positive. We can rearrange (14) as,

(16) $$\frac{v(s, t)^2 \sigma(t, T)}{v(t)} = v(s, t)[\ \sigma(s, T) - \sigma(s, t)]\ .$$

Integrating equation (16) with respect to s, the left side of equation (16) becomes $y(t_0, t, T)x(t_0, t)$. Hence (16) becomes

(17) $$x(t_0, t)y(t_0, t, T) = \int_{t_0}^{t} v(s, t)[\ \sigma(s, T) - \sigma(s, t)]\ ds.$$

Comparing (13) with equation (17), this looks like the Cauchy-Schwartz inequality. The equality holds only if the two parts of the integrand of the right side of (17) are linearly independent. This implies that for fixed $T \le \overline{T}$, we can find $g(T, \overline{T})$ such that $v(t, T) = g(T, \overline{T})v(t, \overline{T})$. Now let $h(T) = g(T, \overline{T})$ and

16. See the proof in Carverhill (1994), p. 308.

$q(t) = v(t, \overline{T})$. Thus we obtain

(18) $$v(t, T) = q(t)h(T).$$

This is the main result of proposition 2-1 of Carverhill. The above result implies that the volatility structure is a deterministic function, and the spot rate process is Markovian. To examine further the volatility structure from Theorems 2 and 3 of Jamshidian (1991), we can express

(19) $$v(t, T) = v(t)\frac{v(t_0, T)}{v(t_0, t)}.$$

This expression is obtained from equation (18). From (c) and (d) in Theorem 2 of Jamshidian (1991), we can find a function $a(T)$ such that for all $0 \le t \le T$,

$$v(t, T) = v(t)\exp(-\int_t^T a(u)du)$$

where $a(T) = -\dfrac{\partial \ln v(t_0, T)}{\partial T}$.

From (12) and (19), we have

(20) $$f(t, T) = f(t_0, T) + \frac{v(t, T)}{v(t)}\int_{t_0}^{t} v(s, t)[\ \sigma(s, T)ds + d\hat{z}\ ].$$

For a spot rate, we have,

(12-1) $$r(t) = f(t_0, t) + \int_{t_0}^{t} v(s, t)[\ \sigma(s, t)ds + d\hat{z}].$$

Hence (12-1) can be rewritten as

(21) $$\int_0^t v(s,t)d\hat{z} = r(t) - f(t_0,t) - \int_0^t v(s,t)\sigma(s,t)ds .$$

Substituting equation (21) into (20), we obtain,

$$f(t,T) = f(t_0,T) + \frac{v(t,T)}{v(t)}(r(t) - f(t_0,t) + \int_{t_0}^t v(s,t)[\ \sigma(s,T) - \sigma(s,t)ds\ ]\ ).$$

Finally using (17), we express the forward rate as

(22) $$f(t,T) = f(t_0,T) + \frac{v(t,T)}{v(t)}(r(t) - f(t_0,t) + x(t_0,t)y(t_0,t,T)).$$

To express the spot rate process from (22), differentiating (22) with respect to T and using the relation that $dr(t) = \frac{\partial f(t,T)}{\partial T}|_{T=t}dt + v(t)d\hat{z}$, we have:

(23) $$dr(t) = (\frac{\partial f(t_0,t)}{\partial t} + x(t_0,t)^2 + a(t)[\ f(t_0,t) - r(t)]\ )\ dt + v(t)\ d\hat{z}$$

$$= (c(t) - a(t)r(t))\ dt + v(t)\ d\hat{z}$$

where $c(t) = \frac{\partial f(t_0,t)}{\partial t} + x(t_0,t)^2 + a(t)f(t_0,t)$. Hence, the function $a(t)$ is a mean reversion parameter for the Gaussian model. As seen in equation (23), the initial forward rate curve controls the trends of the future spot rate. Futhermore, the mean reversion of spot rates is towards the forward rate with a rate of speed equal to $a(t)$. Finally, integrating and exponentiating equation (22) gives a bond pricing formula similar to that in Merton (1973) and Vasicek (1977).

These are the relationships between the bond price, the forward rate, and the spot rate. Actually, these three seemingly different approaches to modelling the term

structure of interest rates are eventually closely related. In the next section, we review the existing term structure models. Since these models are discussed in many papers (e.g. Strickland, 1994), we simply describe the term structure model and the available applications instead of offering a detailed explanation.

3.4. The Affine Model of the Term Structure of Interest Rates

We can classify models of the term structure of interest rates in many different ways, for instance in terms of the number of factors, or the underlying framework (the general equilibrium approach and the martingale approach), or, the stochastic process of bond prices, spot rates, and forward rates. Since the illustration of these models is provided in the major textbooks such as in Hull (1993), we will not reproduce the details here. However, since our models of the term structure of interest rates in Chapters 4 and 5 correspond to the affine type model published by Duffie and Kan (1996), we shall briefly explain the affine framework.

The Affine Model of the Term Structure of Interest Rates

An Affine Model is one in which bond yields are affine functions of the underlying dynamic factors. Duffie and Kan's (1996, hereafter DK) affine model of the term structure of interest rates encompasses many interest rates models. DK show that if the bond price takes the form of the exponential-affine

$$P(X, \tau) = \exp(A_1(\tau) + A_2(\tau)'X)$$

where X is an nx1 vector, $A_1(\tau)$ is a function, $A_2(\tau)$ is a column vector of an n function, and τ is the bond maturity, then the risk neutral process for X satisfies

$$dX_t = (aX_t+b)\ dt+\Sigma\begin{pmatrix} \sqrt{v_1(X\ )},0, & \dots & ,0 \\ 0, \sqrt{v_2(X\ )},0 & \dots & ,0 \\ & & \\ 0, & \dots & 0,\sqrt{v_n(X\ )} \end{pmatrix} dz_t$$

where a and Σ are (n x n) matrices, b is an (n x 1) column vector and

$$v_i(X\ ) = \alpha_{1i}+\alpha_{2i}X$$

where, for each i, α_{1i} is a scalar, α_{2i} is an (n x 1) vector, and dz_t are independent (n x n) Brownian increments. Since the yield $-\frac{\ln P(X,\tau)}{\tau}$ is affine in X, this is called an affine model of the term structure of interest rates. The affine class of term structure of interest rates includes the Vasicek (1977), CIR (1985b), Longstaff and Schwartz (1992), Fong and Vasicek (1991).

The Affine Yield Factor Model

As Duffie and Kan (1996) make clear, yields at fixed maturities are chosen as factors. This is called an affine yield factor model. On the other hand, any state process X_t

$$dX_t = \mu(X_t)dt + \sigma(X_t)dz$$

which may be unobservable is allowed to be an affine factor, and a change of variables from the original state factor X_t to a new yield state Y_t (observable) is attempted, where Y_t is defined by

$$Y_t = -\frac{A(\tau)+B(\tau)X_t}{\tau}.$$

Let $k = -\frac{B(\tau)}{\tau}$ and $h = \frac{A(\tau)}{\tau}$, then $X_t = \frac{1}{k}(Y_t+h)$. In this case, we can write

$$dY_t = \hat{\mu}(Y_t)dt + \hat{\sigma}(Y_t)dz$$

where

$$\hat{\mu}(Y) = k\mu(k^{-1}(Y+h)), \quad \hat{\sigma}(Y) = k\sigma(k^{-1}(Y+h)).$$

Duffie and Kan show that this change of variables preserves the affine framework. This is also called an affine yield factor model.

Before DK, Brown and Schaefer (BS, 1994b) analyzed a one-factor and a two-factor affine model of the term structure of interest rates. They transformed the bond pricing formula into forward rates and investigated the relationship between the yield and the interest rate volatility. DK extend the model of BS into the multi-factor case. The advantage of this approach is that the yield can easily be expressed in terms of the covariance structure of its dynamic factors. This means that we can express the affine model as a multi-factor Markov parameterization of an HJM model. In Chapter 4 we shall show how to fit the affine model of the term structure of interest rates into the HJM framework (a three-factor case). Another advantage is that we can transform the state variable of the affine model (which is possibly unobservable) into yield factors of an affine model (observable). Actually, most non-affine multi-factor models do not allow direct observation of the state variables from the yield curve. Filtering may be useful in this case. However, the affine model of the term structure is easily changed into an affine yield model without the use of filtering. Brown and Schaefer (1994b), for example, estimate the parameters of the two-factor model (assuming the two-factor Vasicek process), using the changes of factors into yields.

We will briefly review some important affine models of the term structure of interest rates.

(1) Vasicek (1977)

The best known model of the term structure of interest rates is Vasicek model. which is is based on the assumption that the short rate r process is as follows:

$$dr = \alpha(\mu - r)dt + \sigma dz.$$

In this specification, a closed form solution of a pure discount bond price and a European type bond option formula (Jamshidian, 1989) are available. Chen (1992) also derives the closed-form solution for futures prices and European options on futures and on pure discount bonds. Unfortunately, this specification of the interest rate process allows negative interest rates.

(2) Schaeffer and Schwartz (1984)

In Chapter 4, we extend this model to the three-factor case. Schaeffer and Schwartz model the term structure model with two factors: the spread rates (short rate - consol rate) and the consol rates. Their model is given by :

$$ds = \alpha(\bar{s} - s)dt + \sqrt{v}dz$$
$$dl = \beta(s, l, t)dt + \sigma\sqrt{l}dz$$

where $\beta(s, l, t) = \sigma^2 - sl$ under the risk-neutral measure.

Strictly speaking, their model is not an affine model of the term structure of interest rates. However, since they use an approximation for the drift of the consol rate as $\sigma^2 - \hat{s}l$, their solution for a pure discount bond price has an exponential affine form. They obtain an approximate solution of a pure discount bond price, dividing the PDE into two parts.

(3) Cox, Ingersoll and Ross (CIR, 1985b)

CIR assume that the short rate process is as follows:

$$dr = \alpha(\mu - r)dt + \sigma\sqrt{r}dz.$$

The short rate can here be guaranteed to be positive. Using this dynamic of the short rate, CIR obtain a closed form for a European call option on a pure discount bond. Longstaff (1993) extends the CIR model, deriving formulas for European call and put options on coupon bonds. Chen and Scott (1992) also show that the CIR model can be used to price options on bond futures.

(4) Longstaff and Schwartz (LS, 1992)

Using the framework of CIR, LS develop the dynamics of two factors which are independent of each other:

$$dx = (\gamma - \delta x)dt + \sqrt{x}dz_1$$
$$dy = (\eta - \theta y)dt + \sqrt{y}dz_2.$$

The spot rate and the the volatility of the spot rate are given by a weighted sum of the factors:

$$r = \alpha x + \beta y$$
$$v = \alpha^2 x + \beta^2 y.$$

LS obtain a closed form for a pure discount bond and a call option on a pure discount bond.

(5) Fong and Vasicek (FV, 1991)

FV obtain a closed form for a pure discount bond price under the dynamics of the short rate and the volatility of the short rate:

$$dr = \alpha(\bar{r}-r)dt + \sqrt{v}dz_1$$
$$dv = \gamma(\bar{v}-v)dt + \delta\sqrt{v}dz_2.$$

Unfortunately, the short rate can be negative. The closed form solution of an option price on bonds is not available.

(6) Chen (1996)

Chen obtains a closed form solution for a pure discount bond price and for other interest rate derivatives under a three-factor dynamic: the short rate, the short-term mean, and the volatility of the short rate, which are given by:

$$dr = k(\theta-r)dt + \sqrt{\sigma}\sqrt{r}dz_1$$
$$d\theta = v(\bar{\theta}-\theta)dt + \delta\sqrt{\theta}dz_2$$
$$d\sigma = \mu(\bar{\sigma}-\sigma)dt + \eta\sqrt{\sigma}dz_3 .$$

where k, v, μ, σ, δ, and η are constant.

(7) Das and Foresi (1996)

Das and Foresi model the term structure of interest rates under the assumption of the possibility of short rate jumps. Their model is given by:

$$dr = \alpha(\mu-r)dt + \sigma dz + ydN.$$

where dN denotes the Poisson process with constant jump intensity and y is the jump distribution. The authors assume that the sizes of jumps follow the exponential distribution. They obtain a closed-form solution of a pure discount bond price. In addition, using the Fourier transformation of the probability distribution of interest rates, they obtain the the price of an option on a bond. Unfortunately, as in the Vasicek model, this model produces negative interest rates depending on the parameter values.

In the following section, we shall explain how the next chapters of the dissertation are related.

3.5. Issues for the Modelling of the Term Structure of Interest Rates

The empirical distribution of changes in interest rates displays leptokurtosis. (Das and Foresi, 1996). This implies that the volatility of interest rates follows a stochastic dynamic. Actually, an ARCH specification of a short-rate process fits the market data well (Steeley, 1990). Many two-factor or three-factor term structure models of interest rates assume that the second factor is the volatility of interest rates, as in Fong and Vasicek (1991) and in the two-factor CIR model. Recently, Chen (1996) obtained a closed form for a pure discount bond price and for a bond option price using a three-factor model with stochastic volatility. The stochastic dynamics of interest rates is regarded as crucial in explaining the empirical distribution of changes in interest rates. We shall study this in Chapter 4.

On the other hand, from the empirical distribution of changes in interest rates, we may question whether an assumption of the continuity of the interest rates process is a good approximation of reality. Empirically observed kurtosis can be explained by jumps in interest rates. A jump-diffusion process implies the opposite effects to those of the stochastic volatility approach: The kurtosis tends to be smaller as the

of the stochastic volatility approach: The kurtosis tends to be smaller as the sampling interval becomes smaller in the stochastic volatility model. In this thesis, we model the term structure of interest rates to reflect these distributional effects. Studies based on the jump diffusion term structure models of interest rates are relatively few compared with those of the diffusion model. We shall extend the DK affine model of the term structure of interest rates to the presence of jumps in Chapter 5.

As Shiller (1989) pointed out, the non-stationary property for the volatility of interest rates could be the result of the market expectation of regime changes in interest rates. We shall investigate the possibility of expectations of the regime changes of interest rates in Chapter 6. As in the jump-diffusion model, we may model discrete shifts of interest rates with a jump process such as a Poisson process. With finite numbers of jumps in any open interval, jump diffusion models of interest rates assume that interest rates follow a diffusion process in between jumps. However, if the market anticipates possible regime changes in interest rates, this can be expected to affect the process. We investigate how the market expectation for regime change in interest rates affects the non-linear property of the volatility of interest rates. As we shall explain, regime changes in exchange rates or interest rates mean changes in the central bank's target rates. Accordingly, depending on the expectation of changing target rates, the exchange rates or interest rates are affected. The regime change models are quite well developed in exchange rate modelling, for example in Krugman and Miller (1992). However, mathematical modelling of regime change is not available in the model of the term structure of interest rates. In Chapter 6, we shall apply the method of Krugman and Miller to the model of the term structure of interest rates. As mentioned in Sections 3.1 and 3.2, two approaches have been taken to the modelling of the term structure of interest rates: the general equilibrium approach and the no-arbitrage approach. We choose the no-arbitrage approach. However, interest rates, in reality, are determined by a complex economic process. To reflect this, we employ the rational expectations model from monetary macroeconomics in modelling the term structure of interest rates.

3.6. Conclusion

In this chapter, we have briefly reviewed two important ways of modelling the term structure of interest rates (the general equilibrium approach and the martingale approach). As explained in the previous sections, in spite of the complete specifications of the economy in the general equilibrium approach, it requires many complicated procedures, for instance, the specification of the individual utility and a solution to the Hamilton-Jacobi-Bellman equation. To overcome the complications, modern finance usually employs the simple modelling of the martingale approach. Using this martingale approach, we model the term structure of interest rates in the next chapters. In particular, in order to reflect the distributional features of interest rates, we model interest rates to include the stochastic volatility of short rates (Chapter 4), the jump factor (Chapter 5), and the market expectation of changes in interest rates (Chapter 6).

Chapter 4. A Three-Factor Model of The Term Structure of Interest Rates

4.1 Introduction

Many model of the term structure of interest rates have either one or two factors, and vary the assumption of the specific type of processes (e.g. square root process or Ornstein-Uhlenbeck process). In a two-factor model, the second factor may be the volatility of an interest rate. The Fong and Vasicek (1991), and the two-factor CIR models are examples of these. Recently, Chen (1996) obtained a closed form solution for a pure discount bond price and for a bond option price, using a three-factor model. However, Chen actually obtained the closed-form solution of a pure discount bond price under the assumption that the short rate followed a Vasicek process, allowing interest rates to be negative.

The explanatory power of the term structure model may increase with the number of factors. However the solution of a pure discount bond price and bond option price is difficult to obtain for non-Gaussian models. In particular, obtaining the probability density in pricing an option, and estimating parameters using the maximum likelihood method are quite difficult. Although the trinomial approximation of Hull and White (1990b) or numerical algorithms such as the finite difference method are available, non-Markovian models (in the case of stochastic volatility) require much computation. In this chapter, we obtain one closed-form and one approximate solution for a pure discount bond price using three factors. Since we choose the spread rate rather than the short rate as one of the factors, the use of the Vasicek process for the spread rate allows spread rates to be negative.

In this chapter, we develop two three-factor affine term structure models of interest rates. In the first model, the factors are: 1) the spread rate (short rate - consol rate), 2) the volatility of the spread rate and 3) the consol rate. In the second model[17], the factors are: 1) the spread rate (short rate - long rate), where the long rate is defined to be a bench-mark rate affecting the level toward which short rate converges, 2) the volatility of spread rate and 3) the long rate. All factors have mean reversion. To preclude negative values of the consol rate, of the long rate and of the volatility of the spread rate, these processes are assumed to follow a CIR process. The spread rate, however, is assumed to follow a Vasicek process because negative values should be permitted. Hence, our model is not a non-negative affine model (Pang and Hodges, 1995).

The chosen factors and the number of factors are based entirely on previous empirical studies.

Steeley (1990) finds that three factors possibly determine 95% of the term structure of interest rates. He also finds that the short-term interest rate and the spread rate are characterized by the ARCH effect. This provides a clear evidence of the stochastic volatility of the short rates and spread rates (short rate - long rate). Steeley suggests that the long rate (for level), the spread rate (for the slope), and the volatility of the spread rate (for curvature) might provide a better description of the term structure of interest rates.

In view of these studies, constructing a three-factor model by combining the long rate with either of the two rates (the short rate or the spread and its volatility) seems a promising path to follow. However, because the use of the spread rate can simplify the solution of partial differential equations (the spread is known to be orthogonal to the long rate (Steeley, 1990), we choose the spread rather than the short rate as one of the three factors. However, it is important to note here that the sentence in Steeley (1990) "The use of the spread rate by Schaefer and Schwartz, which is essentially no more than a redefinition of variables compared to

17. We provide a closed form solution in Appendix I.

the Brennan and Schwartz,...." is quite wrong. First, as seen in the next paragraph, since Schaefer and Schwartz use the consol rate rather than the long rate as a factor, their model does not admit an arbitrage opportunity. However, Brennan and Schwartz model may admit arbitrage opportunities. Secondly, as shall be seen in Appendix I, the use of short rate as a factor implies that the mean reversion level of the short rate is fixed. However, the use of the spread rate as a factor allows the mean reversion level of the short rates to vary as it depends on the long rate[18]. Accordingly, the two models are quite different, and the spread rate is not the redefinition of the short rate.

Actually, the consol rate itself might also be regarded as the bench-mark rate. If the consol rate is used as a factor, however, some care should be taken not to admit arbitrage opportunities. Dybvig, Ingersoll and Ross[19] (1989) and Hogan (1993) have shown that in some cases, including the consol rate as a factor may cause the short rate to fail to have a finite-value solution, and consequently must offer the possibility of arbitrage. However, our model does not include these cases. Furthermore,

---

18. To see the effect of the long rate on the mean reversion of the short rate, we shall compare two models. Ignoring the stochastic term, we set up two-factor interest model as

Model A :

$$dr = a(b-r)dt, \qquad dl = m(c-l)dt$$

Model B:

$$ds = k(n-s)dt, \qquad dl = m(c-l)dt$$

where r is the short rate, l is the long rate, and $s = r - l$ is the spread rate. All parameters a, b, m, c, k and n are assumed to be constant. In Model A, the mean reversion level b is fixed.

· On the other hand, if we substitute the long rate equation into spread rate equation in Model B, we obtain :

$$dr = k\left(\frac{kn+mc}{k} - \frac{(m-k)l}{k} - r\right)dt$$

Hence the mean reversion of the short rate in Model B depends on the level of the long rate.

19. Dybvig, Ingersoll and Ross (1989) analyze the behavior of the long term coupon bond rate, including the zero coupon when the maturity goes infinity. This is the consol rate.

Duffie, Ma and Yong (1995) obtain regularity conditions by which the short rate and the consol rate are consistent with the definition of the consol rate as the yield on a perpetual annuity under an equivalent martingale measure. The derivation of the drift term for the consol rate in our model satisfies their regularity conditions, especially their equation (4.31).

The plan of this chapter is as follows. In Section 2, we derive the price of a pure discount bond using our three underlying factors. In Section 3, we provide the empirical results of the three-factor model of interest rates, and we present a conclusion in Section 4.

4.2 The Model and Its Derivation

4.2.1 The Model

Assume that uncertainty is specified by the probability space (Ω, ξ, P), where ξ is a σ-algebra subset of Ω, and P is the probability measure on ξ. An adapted filtration $\{\xi_t; t \geq 0\}$ defines the information set available to the market agent. The filtration is assumed to be right continuous and P-complete. We assume the following two three-factor models:

(Model 1)

$$ds = \alpha(\bar{s} - s)dt + \sqrt{v}dz_1$$

$$dv = \gamma(\bar{v} - v) + \delta\sqrt{v}dz_2$$

$$dl = \beta(s, l, t)dt + \sigma\sqrt{l}dz_3$$

(Model 2)

$$ds = \alpha(\bar{s} - s)dt + \sqrt{v}dz_1$$

$$dv = \gamma(\bar{v} - v) + \delta\sqrt{v}dz_2$$

$$dl = m(\bar{l} - l)dt + \sigma\sqrt{l}dz_3$$

where $s = r - l$ is the spread, r is short rate, l is the consol rate in model 1 and a bench-mark long rate in model 2, and v is the volatility of the spread rate. We assume that dz_1, dz_2 and dz_3 are standard Gauss-Wiener processes and in particular, that $dz_1 dz_2 = \rho dt$ in model 1, and other instantaneous correlations between processes are equal to zero, and $\beta(s, l) \in C^2([\ 0, \infty))$ i.e. a twice differentiable function. As in Chen (1996), the assumption of $dz_1 dz_2 = \rho dt$ might be not necessary because the spread rate s and its volatility v are already correlated through the stochastic differential equation for the spread rate. We shall assume that all correlations between the Brown motion are zero in model 2. Using this fact, we obtain a closed-form of solution for a pure discount bond price in model 2 in Appendix I. For model 1, however, we assume that $dz_1 dz_2 = \rho dt$, following Fong and Vasicek. We shall specify β under the equivalent martingale measure. In this chapter, we explain and derive a pure discount bond price only from model 1. The derivation of a pure discount bond price from model 2 is in Appendix I.

A consol bond is defined here as a bond paying a continuous coupon at a rate of £1 per period with its value equal to $\frac{1}{l}$. Our model specifies the spread rate instead of the short rate. Hence the spread should not follow a CIR type process. In our model, the spread process can be negative but has the mean reverting

property. Similarly, the volatility equation of the spread rate also has a mean reverting tendency. The consol rate also follows a CIR process but the drift term will be specified under the no-arbitrage condition.

4.2.2 Derivation of the Partial Differential Equation

Under the three-factor description of the term structure of interest rates, it follows from *Itô*'s lemma that the instantaneous return on a bond is given by,

(1) $$\frac{(dP+cdt)}{P} = \mu dt - \phi dz_1 + \psi dz_2 + \zeta dz_3$$

where P is the bond price,

(2) $$\mu = \frac{1}{P}(\alpha(\bar{s}-s)P_s + \gamma(\bar{v}-v)P_v + \beta P_l + \frac{1}{2} vP_{ss} + \frac{1}{2}\delta^2 vP_{vv} + \frac{1}{2}\sigma^2 lP_{ll} + \rho\delta vP_{sv} + P_t + c),$$

(3) $$\phi = -\sqrt{v}\frac{P_s}{P}, \quad \psi = \delta\sqrt{v}\frac{P_v}{p}, \quad \zeta = \sigma\sqrt{l}\frac{P_l}{P},$$

where c is a continuous coupon rate and the subscripts represent partial derivatives. Since we are concerned with a pure discount bond, we will set $c=0$ in deriving a PDE.

The signs of ϕ, ψ, and ζ are chosen arbitrarily to correspond to the direction of the relationship between each factor and the price.

An arbitrage argument applied to the equation for bond prices leads to the equilibrium condition:

(4) $$\mu = r + \lambda_1 \phi + \lambda_2 \psi + \lambda_3 \zeta$$

where λ_1, λ_2 and λ_3 are the market prices of risk due to changes of the spread, the volatility of the spread rate and the consol rate, respectively. To simplify the pricing formula for a pure discount bond, following Fong and Vasicek (1991), we will assume that λ_1 and λ_2 are proportional to the level of risk

$$\lambda_1 = \lambda\sqrt{v}$$

$$\lambda_2 = \eta\sqrt{v}.$$

To derive the partial differential equation satisfied by the value of all default free bonds, we use the fact that the consol rate is inversely proportional to the price of the consol bond as in Brennan and Schwartz (1982). From $P = \frac{1}{l}$

$$P_s = 0,\ P_{ss} = 0,\ P_t = 0,\ P_v = 0,\ P_{vv} = 0,\ P_l = \frac{-1}{l^2},\ P_{ll} = \frac{2}{l^3}\ .$$

Substitution of these derivatives for the consol bond into (2) and (4) yields the following expression for the market price risk of consol rate[20] :

(5) $$\beta - \lambda_3 \sigma\sqrt{l} = \sigma^2 - sl.$$

As shall be seen in the PDE (6) of a pure discount bond price, $\beta - \lambda_3 \sigma\sqrt{l}$ is the coefficient of P_l. We can replace this with $\sigma^2 - sl$, and do not need to specify either β and λ_3.

The drift term (5) and diffusion term of the consol rate in our model is consistent with the equation of (4.31) in Duffie, Ma and Yong (1995). The authors obtain

20. Since $l\,(\beta\frac{-1}{l^2} + \frac{\sigma^2}{l^2} + 1) = r - \lambda_3\frac{\sigma\sqrt{l}}{l}$, we immediately obtain (5).

regularity conditions by which the definition of the short rate and the consol rate is consistent with the definition of the consol rate as the yield on a perpetual annuity under an equivalent martingale measure. Under their regularity condition, they show that the consol rate process under an equivalent martingale measure Q is[21]

$$dl = (-sl + l^3 \|A\|^2)dt + \hat{A}dz$$

where, $A = \frac{\hat{A}}{-l^2}$.

Since the drift term of the consol rate in our case is also derived from the probability measure Q, in our model the consol rate process conforms to this condition[22].

Finally, substitution for μ from (2) and for the market price of the consol rate from (5) into the equilibrium condition (4) yields the partial differential equation satisfied by the value of a default-free bond. Therefore, we obtain the partial differential equation for a discount bond $P = P(s, v, l, \tau)$:

---

21. In the original paper, they mistakenly write $\frac{l^3}{2}\|A\|^2$, in which $\frac{1}{2}$ cancels out by the coefficient of the second derivative of the consol price.

22. To show this, we derive the consol price process from our consol rate process. First, in our model we know that the consol rate process under Q is

$$dl = (\sigma^2 - sl)dt + \sigma\sqrt{l}\, dz.$$

Using Ito's lemma, we can obtain the consol price process Y under Q is,

$$dY = \frac{s}{l}\, dt - \frac{\sigma}{l\sqrt{l}}\, dz$$

$$= (rY-1)dt - \frac{\sigma}{l\sqrt{l}}\, dz.$$

The drift term of the consol price process is exactly the same as (1.3) in Duffie, Ma and Yong. Hence, our model conforms to their condition.

(6) $(\alpha\ \bar{s}-\alpha\ s+\lambda\ v)\ P_s+(\gamma\ \bar{v}-(\gamma+\delta\ \eta)\ v)\ P_v$

$$+\frac{1}{2}\ v\ P_{ss}+\rho\ \delta\ v\ P_{sv}+\frac{1}{2}\delta^2\ v\ P_{vv}+P_l(\sigma^2-l\ s)+\frac{1}{2}\ \sigma^2\ lP_{ll}-\ (l+s)\ P-P_\tau=0$$

subject to the terminal boundary condition $P(s,l,v,0)$ = 1.

Following Schaefer and Schwartz (1984), we are able to solve a related equation by separating it into two parts. One part depends only on s and v, and the other depends only on l. The related equation is identical to (6) except that the term s in the coefficient of P_l is a constant $\hat{s}$,

(7) $(\alpha\ \bar{s}-\alpha\ s+\lambda\ v)\ \overline{P}_s+(\gamma\ \bar{v}-(\gamma+\delta\ \eta)\ v)\ \overline{P}_v$

$$+\frac{1}{2}v\ \overline{P}_{ss}+\rho\delta v\ \overline{P}_{sv}+\frac{1}{2}\delta^2 v\ \overline{P}_{vv}+\ \overline{P}_l(\sigma^2-l\hat{s})+\frac{1}{2}\sigma^2 l\ \overline{P}_{ll}-(l+s)\ \overline{P}-\ \overline{P}_\tau=0$$

with boundary condition

(8) $\overline{P}(s,l,v,0)=1.$

Because of our assumption that the drift of l is $\sigma^2-\hat{s}l$, our model is an affine-type term structure model, and the solution to (7) subject to (8) can be written as

$$\overline{P}(s,l,v,\tau)=X(s,v,\tau)Y(l,\tau)$$

where $X(s,v,t)$ is itself the solution to

(9) $(\alpha\ \bar{s}-\alpha\ s+\lambda\ v)X_s+(\gamma\ \bar{v}-(\gamma+\delta\ \eta)\ v)X_v$

$$+\frac{1}{2}\ v\ X_{ss}+\rho\ \delta\ v\ X_{sv}+\frac{1}{2}\ \delta^2\ v\ X_{vv}-s\ X-X_\tau=0$$

with boundary condition $X(s,v,0)=1$,

and $Y(l,t)$ is solution to

$$(10)\qquad \frac{1}{2}\ \sigma^2\ l\ Y_{ll}+(\sigma^2-l\ \hat{s}\ )\ Y_l-l\ Y-Y_\tau=0$$

with $Y(l,0)=1$.

Equation (9) with boundary condition in the spread rate s is the same as the PDE, in the short rate r, derived by Fong and Vasicek (1991). Similarly, equation (10) in l is given by CIR (1985b) in r. The product of X and Y is the solution to equation (7), i.e.

$$(11)\qquad \overline{P}\ (s,l,v,\tau)=A(\tau)\ \exp(-s\ D(\tau)+v\ F(\tau)+G(\tau)-l\ B(\tau)\ )$$

where

$$A(\tau)=\left(\frac{2\ \theta\ \exp((\hat{s}+\theta)\ \frac{\tau}{2})}{(\hat{s}+\theta)\ (\exp(\theta\ \tau)-1)+2\ \theta}\right)^2$$

$$B(\tau)=\frac{2\ (\exp(\theta\ \tau)-1)}{(\hat{s}+\theta)\ (\exp(\theta\ \tau)-1)+2\ \theta}$$

$$D(\tau)=\frac{1}{\alpha}(1-\exp\ (-\alpha\ \tau))$$

$$G(\tau)=-\alpha\ \bar{s}\int_0^\tau D(t)\ dt+\gamma\ \bar{v}\int_0^\tau F(t)\ dt$$

$$F(\tau)=-\frac{2\ i\ \alpha\ k\ z}{\delta^2}\ \exp\ (-\alpha\ \tau)$$

$$+2\frac{\alpha}{\delta^2}[\ \sum_{j=1}^{2}K_j\ \exp(-\beta_j\alpha\tau)[\ \beta_j\ M(d_j,e_j,i\ k\ \exp(-\alpha\tau))$$

$$+i\ k\ \exp(-\alpha\tau)\ \frac{d_j}{e_j}\ M(d_j+1,e_j+1,i\ k\ \exp(-\alpha\tau))]\ ]$$

$$/\ \sum_{j=1}^{2}K_j\ \exp(-\beta_j\ \alpha\ \tau)\ M(d_j,e_j,i\ k\ \exp(-\alpha\ \tau))$$

and where

$$k=\frac{\delta}{\alpha^2}\sqrt{1-\rho^2}$$

$$z=\frac{1}{2}-\frac{i}{2}\frac{\rho}{\sqrt{1-\rho^2}}$$

$$d_j=\frac{1}{2}e_j-\frac{i}{2}\frac{\rho}{\sqrt{1-\rho^2}}[\ \frac{\delta}{\alpha^2}(1-\alpha\lambda)-\rho(\mathfrak{h}-1)]$$

$$e_j=2\pi_j-\varkappa+1$$

$$\pi_1=\pi$$

$$\pi_2=\varkappa-\pi$$

$$i=\sqrt{-1}$$

$$\pi=\frac{1}{2}\varkappa-\ \frac{1}{2}\sqrt{\varkappa^2-\frac{\delta^2}{\alpha^4}+2\lambda\frac{\delta^2}{\alpha^3}}$$

$$x = \frac{\gamma + \delta\,\eta}{\alpha} + \frac{\rho\,\delta}{\alpha^2}$$

$$M(d, e, z) = 1 + \sum_{n=1}^{\infty} \frac{d(d+1)\ldots(d+n)z^n}{e(e+1)\ldots(e+n)n!}$$

$$\theta = \sqrt{s^2 + 2\,\sigma^2}\ .$$

We can compute the functions $F(\tau)$ and $G(\tau)$ using the approximation (Frobenius solution) suggested by Selby and Strickland (1993).

4.2.3 The Features of the Model

The functions $D(\tau)$, $A(\tau)$, $B(\tau)$ and F(t) are real with a finite limit. The function $D(\tau)$ is from Vasicek (1977), and $A(\tau)$ and $B(\tau)$ are given by CIR (1985b). The limits of the functions $F(\tau)$ and $G(\tau)$ are given by Fong and Vasicek[23] (1991). The bond pricing formula (11) shows the following realistic properties :

$$\lim_{\tau \to \infty} P(s, l, v, \tau) < \infty,$$

$$\lim_{s \to \infty} P(s, l, v, \tau) = 0,$$

23. The limit of function $F(\tau)$ is

$$\lim_{\tau \to \infty} F(\tau) = \frac{2\,\alpha\,\beta}{\delta^2},$$

and the function $G(t)$ is just calculated by integration of $D(t)$ and $F(t)$ with finite limit over τ. Hence the limit of function of $G(\tau)$ is finite.

$$\lim_{l \to \infty} P(s, l, v, \tau) = 0.$$

We express our model in terms of the forward rate. The forward rate at time t for the instantaneous future period at time $t+\tau = T$ is $\frac{P_t}{P}$. Since our model belongs to the affine class and is time homogeneous[24], the instantaneous forward rate can easily be obtained from equation (11) as

$$(12) \qquad f(t, T) = r + \mu_s D(\tau) - \mu_v F(\tau) + \mu_l B(\tau) - \frac{1}{2} v D(\tau)^2$$

$$+ \rho \delta v D(\tau) F(\tau) - \frac{1}{2} \delta^2 v F(\tau)^2 - \frac{1}{2} \sigma^2 l^2 B(\tau)^2$$

where μ_s, μ_v and μ_l are the risk-adjusted drift of the spread rate process, the volatility of the spread rate, and the consol rate process, respectively.

The quantities $D(\tau)$, $B(\tau)$ and $F(\tau)$ can be interpreted as two duration measures and the exposure to volatility, respectively. The forward rate is therefore expressed as functions of the exposure to interest rates (duration) and volatility. Equation (12) shows that the forward rate is a concave function of two duration measures and the volatility exposure measure. We can confirm that the volatility parameters, like some members of the affine model, affect the forward rate curve in two directions. First, the functions, $F(\tau)$ and $B(\tau)$ are themselves functions of volatility. Secondly, changes in volatility affect the convexity of the forward rate curve with respect to two duration measures and the volatility exposure measure[25].

---

24. Time homogeneous means that the drift and diffusion coefficients of the fundamental valuation equation can be functions of the level of the state variables and do not explicitly depend on time.

25. The convexities of the forward rate curve with respect to $B(\tau)$, $D(\tau)$ and $F(\tau)$ are $-\frac{\sigma^2 l^2}{2}$, $-\frac{v}{2}$, and $-\frac{\delta^2 v}{2}$ respectively.

The term structure of interest rates is determined from the pricing equation (11). Defining $R(s, l, v, \tau)$ as the spot rate, then,

$$(13) \qquad R(s, l, v, t) = -\frac{1}{\tau}\ln(\overline{P}(s, l, v, \tau))$$

$$= -\frac{1}{\tau}(-sD(\tau)+vF(\tau)+G(\tau)-B(\tau)+\ln A(\tau)) \ .$$

Equation (13) demonstrates the behavior of interest rates as a function of the three factors.

To use equation (11), it is necessary to find a value for $\hat{s}$. For the purpose of computing $\hat{s}$, Schaefer and Schwartz (1984) ignore the diffusion terms in their two equations, and then solve the two pairs of the deterministic differential equations. Ignoring uncertainty, the price of a pure discount bond is:

$$\text{(i)} \qquad \overline{P} = \exp(-\int_0^t s(\tau)d\tau)\ \exp(-\int_0^t l(\tau)d\tau).$$

Then, the risk adjusted processes are

$$\text{(ii)} \qquad \frac{ds}{dt} = \alpha(\tilde{s}-s), \quad \frac{dl}{dt} = \sigma^2 - s\ l.$$

where $\tilde{s} = \bar{s} - \frac{\lambda\ v}{\alpha}$.

On the other hand, under equation (7), the pair of deterministic differential equations are

$$\text{(iii)} \qquad \frac{ds}{dt} = \alpha(\tilde{s}-s), \quad \frac{dl}{dt} = \sigma^2 - \hat{s}\ l.$$

SS choose $\hat{s}$, so that the value of $\overline{P}$, under (ii) and under (iii), are the same. Since in our case, the stochastic dynamic of the volatility of the spread is omitted in calculating $\hat{s}$, computing $\hat{s}$ is expected to have a relatively large error compared with that of SS.

Since the numerical solution itself is an approximation, we test our model by another method. Following SS, we investigate the internal consistency of the model, and compare the computed yield from the model with the input consol yield. For simplicity and to avoid a numerical calculation, we assume that the initial value of s is equal to $\hat{s}$ [26]. In this case, it can be determined that $\hat{s} = \hat{s}$. To calculate the consol yield, we compute the price of a 200-year annuity. This is an approximation as well. The formula used is as follows:

$$l = \frac{1}{\sum_{n=1}^{200} \overline{P}(s, l, v, n)} \tag{14}$$

where $\overline{P}(s, l, v, n)$ is from equation (11). We compare the calculated consol yield from equation (14) with the input consol yield. Table 1 shows the base case of the parameters, which is similar to the case of SS.

26. See Schaefer and Schwartz (1984), pp. 416-418.

Table 1. Base Case of Parameters

$\alpha = 0.72, \quad \gamma = 0.1$

$\delta = 0.001, \quad \rho = 0.3$

$\eta = \lambda = 0, \quad v = \bar{v} = 0.0005$

$s = \hat{s} = \tilde{s} = \bar{s} = -0.007, \; l = 0.05$

Model

(risk-adjusted process)

$$ds = \alpha(\tilde{s} - s)dt + \sqrt{v}dz_1$$

$$dv = \gamma(\tilde{v} - v) + \delta\sqrt{v}dz_2$$

$$dl = (\sigma^2 - \hat{s}\, l)dt + \sigma\sqrt{l}dz_3$$

We compare two yields both in our model and in the SS model. We are interested only in the reasonable consol rate ranging from 5% to 11%. Table 2 shows the result. Surprisingly, the two models produce large errors. The SS model fits the lower consol rate better than does ours. In the high consol rate, our model fits better. In the original paper, SS reported that the error of their model was negligible in the low variance case. However, even in this case, the $\hat{s}$, which is computed from the SS suggestion, does not seem to provide a good guide. Actually, discovering an heuristic to enable us to compute an approximate value of $\hat{s}$ is difficult, especially in our case. We iterate values for $\hat{s}$ between -0.01 and 0.01 with interval 0.001, which are the values around the value (-0.007) suggested by SS. We

can not find the value $\hat{s}$ which produces a reasonable consol price in both cases.

Table 2. Comparison of Models

($\hat{s}$ from SS formula) (unit, %)

Low Variance Case

(σ = 0.02)

Consol Yield (Input)	Our Model	SS Model
5	5.50	4.93
8	8.23	7.69
11	10.92	10.41

High Variance Case

(σ = 0.03)

Consol Yield (Input)	Our Model	SS Model
5	5.76	5.19
8	8.33	7.77
11	10.95	10.43

Figure 1 shows the yields setting $\hat{s}=-0.007$ in Table 1. The long term-yield

seems to converge to a higher value than the input consol yield. This might be the result of our approximation. Since our model is basically a combination of the CIR model and Fong and Vasicek Model, we do not reproduce the detailed graph of the sensitivity with respect to the changes of parameters values. Figure 2 shows the curves of three factor loadings $\frac{D(\tau)}{\tau}$, $-\frac{(D(\tau)-B(\tau))}{\tau}$ and $-\frac{F(\tau)}{\tau}$, which shows sensitivity of the short rate, the consol rate and the volatility of the spread rate with respect to the bond yields. As seen in Figure 2, the effect of the short rate is more significant for short maturity yields, while the effect of the volatility of the spread rate is relatively high for intermediate maturity, but slowly dies out. The effects of the consol rates are rapidly increase as the maturity increase. For the long maturity, the effect of the consol rate dominates others.

4.2.4 Fitting to the Term Structure under the HJM framework

To fit the term structure model to a given yield curve, we merge our model with the HJM framework, following Chen (1996). Suppressing the notational dependence of the forward rate on the three-factors, let $f(0, T)$ be the initial forward rate, maturing at time T. The forward rate at time t is defined as,

$$f(t, T) = -\frac{\partial \ln P(t, T)}{\partial T}. \tag{15}$$

Following HJM, we assume that the forward rate is as follows:

$$f(t, T) = f(0, T) + \int_0^t a(s, T, f(s, T))ds + \sum_{i=1}^{3} \int_0^t b_i(s, T, f(s, T))dz_i(s). \tag{16}$$

As in HJM (1992), we assume that the functions $a(.)$ and $b_i(.)$ satisfy the Lipschitz condition and the growth condition.

From equation (15), we know that the bond price is given by

$$(17) \qquad P(t, T-t) = e^{-\int_t^T f(t,s)ds} \quad \text{for all } T \in [\ 0, \tau]\ , t \in [\ 0, T]\ .$$

To show the no-arbitrage condition, we introduce a money market account, or a numeraire security $B(t)$ as in HJM (1992) and Babbs (1991), with

$$(18) \qquad B(t) = e^{\int_0^t r(s)ds}.$$

Let $Z(t, T-t)$ be the relative bond price for a T-maturity bond, which is given by

$$(19) \qquad Z(t, T-t) = \frac{P(t, T-t)}{B(t)}\ .$$

Applying Ito's lemma to $Z(t, T-t)$ yields

$$(20) \qquad \frac{dZ(t, T-t)}{Z(t, T-t)} = -\int_t^T a(t, s, f(s, T))dsdt + \frac{1}{2}\sum_{i=1}^{3}[\ \int_t^T b_i(t, s, f(s, T))ds]\ ^2 dt\ .$$
$$- \sum_{i=1}^{3}\int_t^T b_i(t, s, f(s, T))dsdz_i(t).$$

As in HJM and in Babbs, we know that the process $Z(t, T-t)$ is a martingale under the equivalent probability measure Q. This implies that the drift term in (20) should be zero:

$$(21) \qquad \int_t^T a(t, s, f(s, T))ds = \frac{1}{2}\sum_{i=1}^{3}[\ \int_t^T b_i(t, s, f(s, T))ds]\ ^2.$$

From HJM (1992), the bond price follows the process:

(22) $$\frac{dP(t, T-t)}{P(t, T-t)} = \mu(.)dt + \sum_{i=1}^{3} \int_t^T b_i(t, s, f(s, T))dsdz_i(t).$$

On the other hand, from equation (3) in our model, the bond price follows the process:

(23) $$\frac{dP(t, T-t)}{P(t, T-t)} = \mu(.)dt - D(\tau)\sqrt{v}\, dz_1 + F(\tau)\delta\sqrt{v}\, dz_2 - B(\tau)\sigma\sqrt{l}\, dz_3.$$

To fit our model into the HJM framework, the volatility of bond returns from (22) is computed and compared with (23):

$$\int_t^T b_1(t, s, f(s, T))dsdz_1(t) = -D(\tau)\sqrt{v}\, dz_1.$$

$$\int_t^T b_2(t, s, f(s, T))dsdz_2(t) = -F(\tau)\,\delta\sqrt{v}\, dz_2.$$

$$\int_t^T b_3(t, s, f(s, T))dsdz_3(t) = -B(\tau)\,\sigma\sqrt{l}\, dz_3.$$

Assuming that the forward rates are known at every time t, then the bond price is given by

(24) $$P(t, T) = e^{-\int_t^{t+T} f(t,s)ds}.$$

To express (24) in terms of initial forward rate and volatility structure, we take the derivative of (21) with respect T. Then we obtain:

$$a(t, s, f(s, T)) = \sum_{i=1}^{3} b_i \int_t^T b_i(t, s, f(s, T))ds.$$

Substituting this into equation (16), then we can express (24) as

$$P(t, T) = \exp\Big(\int_t^{t+T} f(0, s)ds + \sum_{i=1}^{3}\int_t^{t+T}\int_0^t b_i(s, T)\int_t^T b_i(s, T, f(s, T))dsdT$$
$$+ \sum_{i=1}^{3}\int_t^{t+T}\int_0^t b_i(s, T, f(s, T))dz_i ds\Big).$$

Substituting each b_i into this gives the solution for the bond price under the HJM framework. The formula is consistent with the given initial yield curve.

4.3. An Estimation of the Parameters

The empirical investigation of a continuous time diffusion model is not straightforward since a closed form expression for the density in the diffusion model is difficult to obtain. Among the many different types of diffusion processes, the Gaussian and square root (CIR type) processes have known densities. For instance, Brown and Dybvig (1986), Pearson and Sun (1991), Sun (1992), and Brown and Schaefer (1994a) use the maximum likelihood method in estimating the CIR model. Since the closed form solution of a pure discount bond price in the CIR model is known, it is straightforward to implement a Maximum Likelihood Estimation (MLE) method even if the state variable is unobserved, since it can be expressed in terms of the observable yields using the CIR formula.

One way to avoid specifying the density is to use the Generalized Method of Moment (GMM). Chan, Karolyi, Longstaff and Sanders (1992, hereafter, CKLS) analyse a variety of models for the short rate and its volatility, using a GMM. Gibbons and Ramaswamy (1993) also use GMM in estimating the CIR model in a similar way to Pearson and Sun (1991). Duffie and Singleton (1993) use the method of Simulated Moments. However, to use the full information set available, the conditional moment is more desirable than the unconditional moment. He (1990) applies this approach to estimate the CIR model. However, as with the difficulty in

using MLE, this approach requires the determination of the whole density. As He (1990) states, when we use MLE or CGMM (Conditional GMM), we have to specify the density of the diffusion model. Upon increasing the number of factors, it is more difficult to get an analytic solution of the density. Binomial or trinomial approximation of the density, as in Hull and White (1990b), is one possible way to use MLE or CGMM in multi-factor cases. Of course, in the case of changing volatility some care, as in Nelson and Ramaswamy (1989), should be taken to make the tree recombining or Markovian. However, the required calculation in using this approach is enormous, for instance, in estimating the multi-factor model in which the factors follow different processes.

On the other hand, Nelson (1990) and Duan (1997) show that some families of discrete ARCH or GARCH models converge in distribution to *Itô* processes as the length of the discrete time interval goes to zero. This result can broaden the possible application of some types of stochastic diffusion models with a stochastic volatility, as with our model, to the maximum likelihood method in estimating parameters. However, as Longstaff and Schwartz (1992) point out, this discretised ARCH specification is only an approximation to the continuous time model.

More recently, Ait-Sahalia (1996a, 1996b) applies a non-parametric technique to estimate a diffusion model. He estimates the drift term from Ordinary Least Squares and the diffusion parameters by a Kernel density method. The non-parametric approach also has a problem if the number of factors is increased.

On the other hand, Oakes (1997) estimates parameters for multivariate diffusion models which is constructed by applying standard GMM methods to a moment restriction vector evaluated using numerical approximations to the conditional moments of the underlying diffusion. We explain his method in Appendix IV. He argues that the moment approximation method offers a potential improvement over the application of GMM or ML to discretized versions of the diffusion model. A general algorithm is presented for approximating the conditional moments of the diffusion by constructing a trinomial tree for each underlying variable, after an

orthogonal transformation to remove the dependency on each underlying variable. The algorithm he employed is a theoretically good approximation of the diffusion model. However, it requires a lot of computation on the increasing the number of factor.

Until now, it is thought that there is no one best method of estimating a diffusion model. Futhermore, accuracy requires a highly complicated calculation. We estimate our model from the time series data, using a GARCH-X model similar to Brener, Harjes, and Kroner (hereafter, BHK, 1996). Even if the parameters of the GARCH model do not map directly into the parameters of the continuous time process, this model has an advantage in easily implementing the stochastic volatility, and in using MLE. Specifically, the specification of our model is GARCH-X. Before we specify our model, we acknowledge that as in Longstaff and Schwartz (1992), our model does not exactly map into the following discretization.

From our model, the risk-adjusted processes are:

(25-1) $$ds = \alpha(\tilde{s} - s)dt + \sqrt{v}dz_1$$

(25-2) $$dv = \gamma(\tilde{v} - v) + \delta\sqrt{v}dz_2$$

(25-3) $$dl = (\sigma^2 - \hat{s}\ l)dt + \sigma\sqrt{l}dz_3\ .$$

where a tilde denotes the risk-adjusted parameters, and where $\tilde{s} = \bar{s} - \lambda_1$, $\tilde{v} = \bar{v} - \lambda_2$, and λ_1 and λ_2 are the market price of risk of the spread rate and the volatility of the spread rate, respectively.

Following Longstaff and Schwartz (1992), we discretize our model as follows: Using $s = r - l$, we can rewrite (25-1) as

(26) $$dr - dl = \alpha(\tilde{s} - s)dt + \sqrt{v}dz_1.$$

Substituting (25-3) into (26), then

$$dr = \alpha(\tilde{s} - s)dt + (\sigma^2 - \hat{s}\, l)dt + \sigma\sqrt{l}dz_3 + \sqrt{v}dz_1 .$$

Expressing the above equation in discrete terms, we can express it in the following form:

(27) $$\Delta r_t = \beta_0 + \beta_1 s_t + \beta_2 l_t + \varepsilon_t, \quad \varepsilon_t \sim N(0, V_t \Delta t)$$

where

$$V_t = v_t + \sigma^2 l_t$$

due to the orthogonality of dz_1 and dz_3, and

$$\beta_0 = (\alpha\tilde{s} + \sigma^2)\Delta t,$$
$$\beta_1 = -\alpha\Delta t,$$
$$\beta_2 = -\hat{s}\,\Delta t.$$

To express our model in the GARCH specification, we discretize our stochastic volatility as,

$$\Delta v = \gamma(\tilde{v} - v)\Delta t + \delta\sqrt{v}\, Z\sqrt{\Delta t}$$

where Z is a standard normal variable, and $V_{t-1} = v_{t-1} + \sigma^2 l_{t-1}$. Then, following LS and Chen (1994),

$$V_t = v_t + \sigma^2 l_t + c_4\varepsilon_{t-1}^2$$

$$= \gamma(\tilde{v}-v_{t-1})\Delta t+v_{t-1}+\sigma^2 l+c_4\varepsilon_{t-1}^2$$

$$= \gamma\tilde{v}\Delta t+(1-\gamma\Delta t)\ v_{t-1}+\sigma^2 l+c_4\varepsilon_{t-1}^2{}'$$

$$= \gamma\tilde{v}\Delta t\ +\ (1-\gamma\Delta t)(V_{t-1}-\sigma^2 l_{t-1})\ +\ \sigma^2 l_t\ +\ c_4\varepsilon_{t-1}^2.$$

$$= \gamma\tilde{v}\Delta t\ +\ (1-\gamma\Delta t)V_{t-1}-(1-\gamma\Delta t)\sigma^2 l_{t-1}\ +\ \sigma^2 l_t\ +\ c_4\varepsilon_{t-1}^2.$$

where the parameter c_4 comes due to GARCH specification. Since the parameter δ does not appear in the above specification, δ can not be specified, which is the similar econometric specification problem of the models in Longstaff and Schwartz (1992) and Chen (1994). As we acknowledged in the beginning of this section, the GARCH specification is not a one-to-one mapping into our continuous model.

We can express the above equation in a compact way, in the GARCH form as

(28) $$V_t\ =\ c_0\ +\ c_1 l_t\ +\ c_2 V_{t-1}\ +\ c_3 l_{t-1}\ +\ c_4\varepsilon_{t-1}^2$$

where

$$c_0\ =\ \gamma\tilde{v}\ \Delta t,$$

$$c_1\ =\ \sigma^2,$$

$$c_2\ =\ (1-\gamma\Delta t),$$

$$c_3=-c_1 c_2$$

This specification of (26) and (27) is similar to the GARCH-X model of Brenner, Harjes, and Kroner (1996). We estimate our model using the GARCH-X framework. We use the Berndt-Hall-Hall-Hausman (BHHH) algorithm to find the maximum likelihood parameter estimates[27]. For parsimony, when we maximize the likelihood

27. We use the RATS package.

function, we explicitly restrict the coefficient $c_3 = -c_1 c_2$ for parsimony.

Our empirical study covers monthly interest rate data for US Treasury securities from January 1988 to August 1996. As Dahlquist (1996) states, the interest rates with shorter maturities sometimes contain idiosyncrasies (i.e. seasonalities and high bid-ask spreads). The maturities of these are one-month and ten-years. The one-month yields are used as a proxy for the short rate, and the ten-year yield for the consol rate. The data are obtained from Datastream. Tables 3 and 4 show the empirical result. We calculate the covariance of the spread rate and the long rate as 0.096. As in Steeley (1990), the correlation between the two processes are very low. Our assumption of zero correlation between two processes is reasonable.

Figure 4 plots the OLS residual of our model (27). We first test to see if there are non-linear effects in the model and to compare our specification of the model of the term structure with that of SS. We perform the LM test for the presence of ARCH effects, as was proposed by Engle (1982). We obtain the series of residuals from the SS specification of spread rates. Then, we regress the squared residuals on the first four lags. The results are as follows :

$$\begin{array}{lllllll} \varepsilon_t^2 = & 0.11 & + \; 0.37\varepsilon_{t-1}^2 & -0.05\varepsilon_{t-2}^2 & +0.02\varepsilon_{t-3}^2 & -0.05\varepsilon_{t-4}^2 \\ & (3.15) & (3.68) & (-0.51) & (0.16) & (-0.54) \end{array}$$

$$TR^2 = 35.278, \qquad DW = 2.00$$

where the values in parenthesis denote t-values.

Except for the first lag, the coefficients are not significant. In addition, TR^2 is 35.279, whereas the critical value of $\chi^2(4)$ at the 5% significant level is 14.86. This implies that the stochastic volatility of the spread rate has explanatory power and, in particular, the GARCH(1,1) specification is reasonable.

Next, we perform the LR test for two models (the SS model and our model).

This is an indirect test. We cannot specify the exact likelihood functions of the two models. However, as Hamilton (1994) indicates, the maximization of the Gaussian MLE of ARCH specification still offers a consistent estimator. We impose zero-restriction on the coefficients c_2 and c_4. This restriction implies the specification of SS. We compare the two values of the likelihood function. Let us call the parameters of the unrestricted model (our model) A, and of the restricted model (SS model) B. The value of 2 [log (A)-log(B)] is 13.968. From $\chi^2(2)$ = 10.60 at the 5% significant level, we might conclude that our three-factor model can reject the SS model at the 5% significance level, although the likelihood value does not increase dramatically.

Table 3. Estimates Using the GARCH-X Process[28]

	β_0	β_1*	β_2	c_0*	c_1*	c_2	c_4
Estimate	-0.191	-0.072	0.007	-0.212	0.049	0.172	0.314
t-value	-1.142	-2.756	0.277	-2.637	4.717	0.752	1.706

Note : * indicates significance at the 1 % level.

Table 4. Estimates of Annualized Parameters

	γ*	α*	$\hat{s}$	$\tilde{s}$	σ^2*	$\tilde{v}$*
Estimate	9,936	0.864	-0.084	-2.709	0.049	-0.021
t-value	3.6316	2.769	-0.280	n.a	4.717	-2.830**

Note : * indicates significance at the 1 % level.
** t value is obtained by Delta method

28. The estimation result of SS model is as follows

$$\Delta r_t = -0.137 - 0.069\ s_t - 0.001\ l_t + \varepsilon_t$$
(-0.654) (-2.798) (-0.049)

$$var(\varepsilon_t) = -0.373 + 0.073\ c_1.$$
(-4.721) (5.468)

Tables 3 and 4 show the estimation results for our model. Since we cannot calculate the t-value of $\tilde{s}$, we report only the estimate. The t-value for $\tilde{v}$ is obtained by Delta method[29]. Other parameters can be recovered from the original parameters. The spread rate and the volatility of the spread rate is a stable process since the two coefficients (γ and α) are significant and positive. In particular, γ, the mean reversion speed of the volatility is very high in our data set. $\tilde{s}$ ($= \bar{s}-\lambda_1$), the mean reversion level of risk-adjusted process s has a high negative value. As seen in footnote (28), the mean reversion speed of our model is slightly higher than that of SS. $\tilde{v}$ ($= v-\lambda_2$), the mean reversion level of the volatility has a negative value as well. This might be due to a high market price of risk (λ_1 and λ_2) of the spread rate and the volatility. β_2 or $\hat{s}$ is not significant. Our model assumes implicitly that the volatility of the short rate depends on the long rate, which was interpreted as a factor affecting the level toward which the short rate converges. As explained in Section 4.1, the use of the spread rate as a factor implies that the volatility of the short rate is affected by the long rate, while the use of the short rate as a factor, as in Fong and Vasicek (1991), does not. As seen in Table 3, the volatility of the short rate heavily depends on the level of the long rates (c_1). It seems that due to the effects of the long rates on the conditional variance of the short rate, the moving average component (c_4) and the autoregressive term (c_2) does not much affect the volatility of the short rates.

The LR test for comparison between our model and SS model reflect this as well. The value of the likelihood function does not dramatically increase even if our model performs better than that of SS. This implies that the non-linear effect of the volatility of the short rate in our data set depends more on the level of the long rate rather than the ARCH effect of the short rate. Our three-factor specification does not dramatically improve the explanatory power of the term structure of interest rates in this data set. Further studies should be carried out interest rates of UK

29. See *A Course in Econometrics* by Goldberger (1991), p110.

and other countries.

The result of our mathematical model and consequently the empirical model is quite similar to that of a GARCH-X estimation proposed by BHK (1996). BHK (1996) find that the specification of the volatility of the short rates as a GARCH-X model rather than as a pure GARCH or ARCH model increases the explanatory power of the model. We do not directly compare our model with the pure GARCH model, however, the level of the interest rates has a great effect on the behavior of the volatility of short rates.

In the empirical test of our model, the ten-year yield, which is used as a proxy of the consol rate, is assumed to affect directly the volatility of the short rate. This is also similar to the specification of Litterman, Scheinkman and Weiss (1988) (LSW). LSW argue that the volatility of the short rate is quite well explained by the combination of mid-and long-term yields. They estimate the volatility of the short rate, using simple OLS. As in equation (27), the volatility of the short rate in our three-factor model is also expressed as the long-term yield (consol rate). The use of the spread rate rather than the short rate, as a factor, mathematically supports the GARCH-X modelling of the short rate and the LSW empirical work.

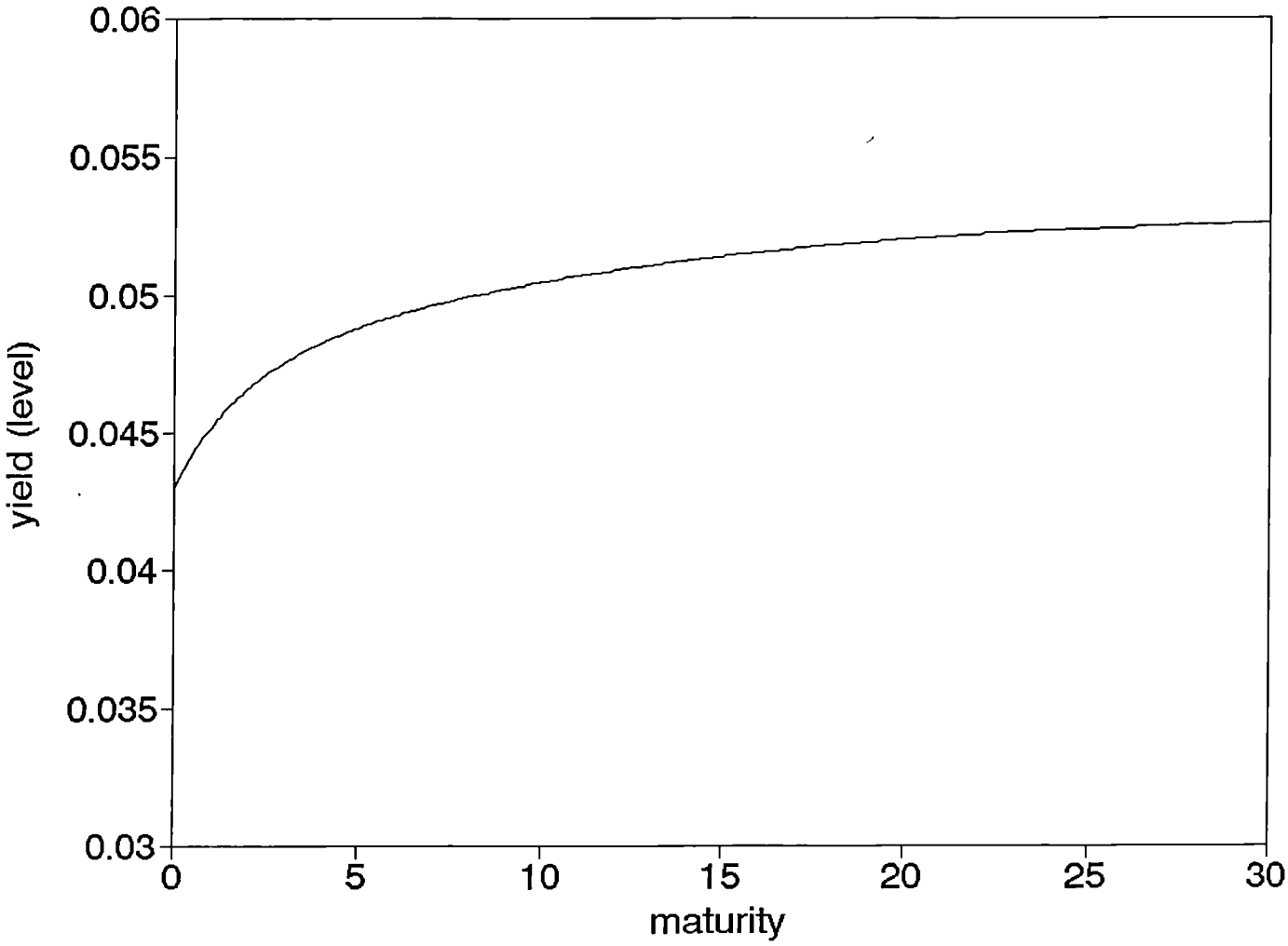
4.4 Conclusion

We present an approximate solution of a pure discount bond price using the three-factor model of the term structure of the interest rates. Model 1 is an extension of the SS two-factor model. The chosen factors (Model 1 and Model 2) are based on previous empirical studies such as that of Steeley (1990). The accuracy of the approximation technique suggested by SS is quite limited and for certain parameter values the solution produces quite large errors. As stated in the previous section, discussing an heuristic to enable us to compute an approximate value of $\hat{s}$ is difficult. However, in Appendix I, we present a closed-form solution

of a pure discount bond price based on three factors (Model 2). Considering the fact that there are only a few closed-form solutions of a pure discount bond price using three factors, this new model might show increased explanatory power for the behaviour of interest rates.

In an empirical result for our model for US data, we find that the conditional variance of the short rate can be explained by the level of the long rate. From the LR test, we demonstrate that our model performs better than the SS model. As mentioned in the previous section, further studies should be carried out interest rates for the UK and for other countries. In particular, our model supports theoretically the use of GARCH-X specification of short rate by BHK and the specification of the volatility of the short rate by LSW. As LS (1992) indicate, however, the GARCH-X discretization of a continuous time model does not map into our model. This can be expected to affect the empirical result. The approach employed by Oakes (1997) could improve the problems caused by the approximation of the continuous-time model into the discrete time model. However, the approach would be computationally very difficult to implement. We shall provide a example in Appendix IV.

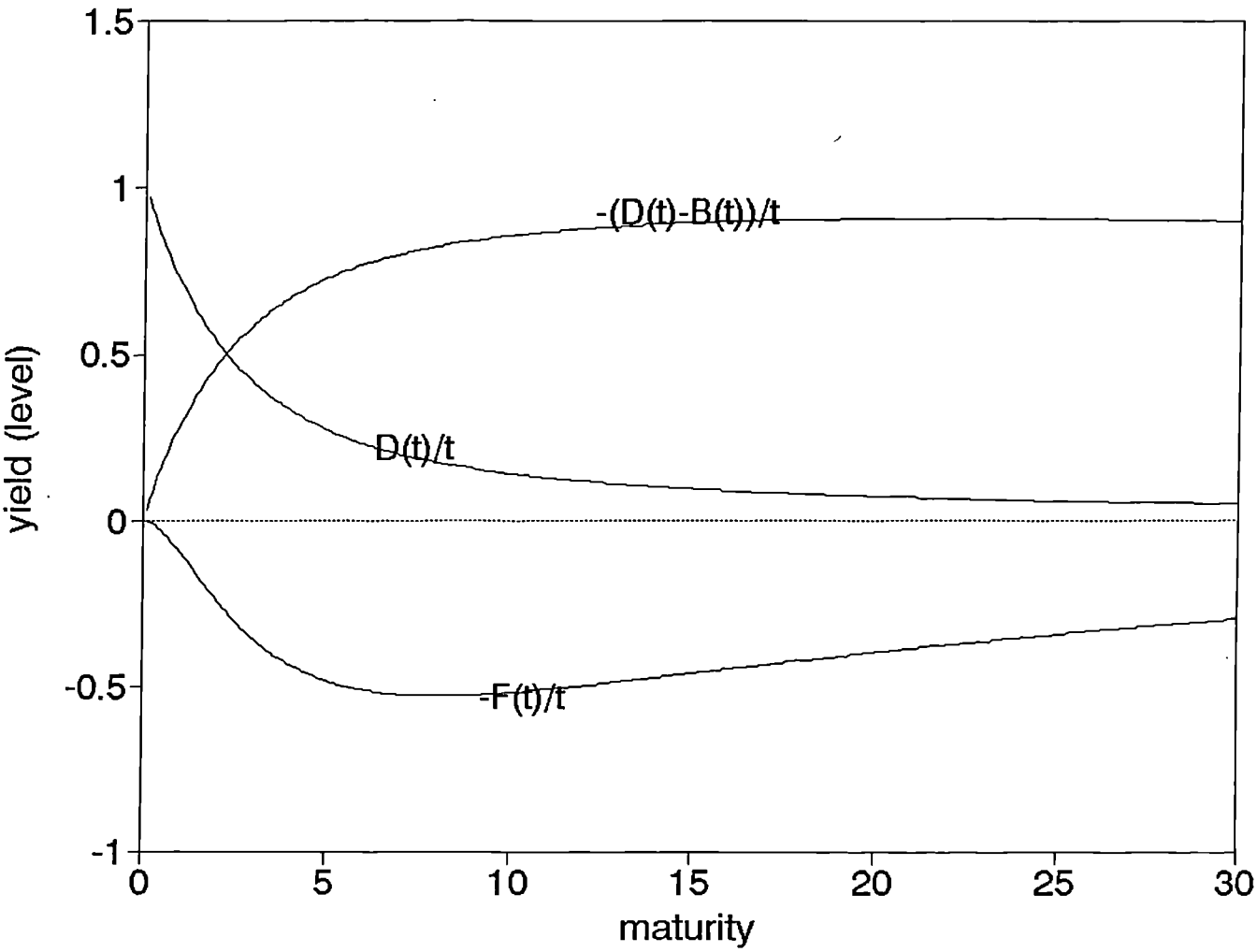
Figure 1. The Term Structure of Interest Rates



$\alpha = 0.72,\ \ \gamma = 0.1,\ \ \delta = 0.001, \rho = 0.3, \eta = \lambda = 0,$

$v = \bar{v} = 0.0005, s = \hat{s} = \bar{s} = -0.007,\ \ l = 0.05$

Figure 2. Factor Loadings



$\alpha=0.72,\ \ \gamma=0.1,\ \ \delta=0.001,\ \rho=0.3,\ \eta=\lambda=0,$

$v=\bar{v}=0.0005,\ s=\hat{s}=\bar{s}=-0.007,\ \ l=0.05$

(i) $\frac{D(\tau)}{\tau}$: factor loading of the short rate

(ii) $-\frac{(D(\tau)-B(\tau)}{\tau}$: factor loading of the consol rate

(iii) $-\frac{F(\tau)}{\tau}$: factor loading of the volatility of the spread rate

Figure 3. Interest Rates

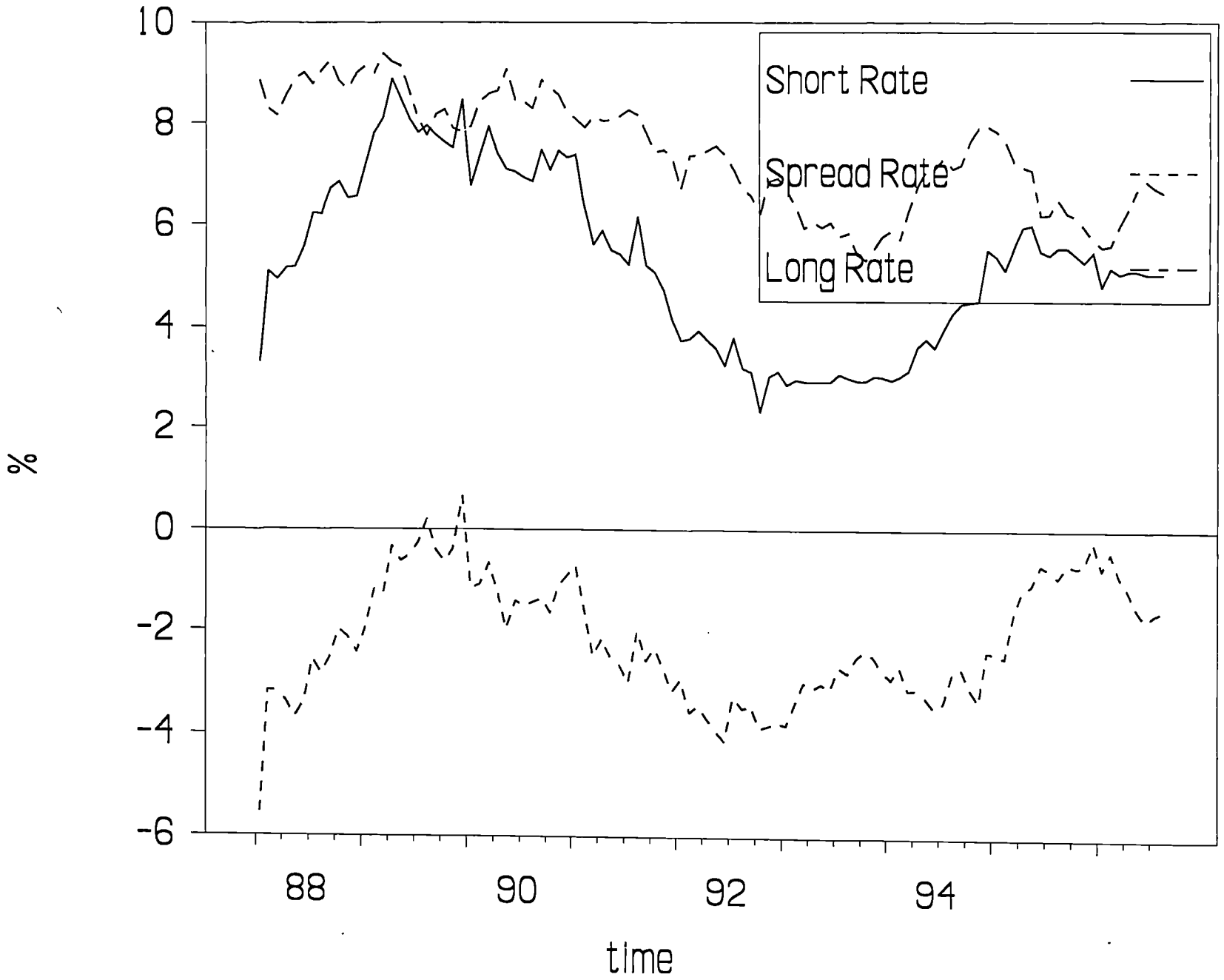
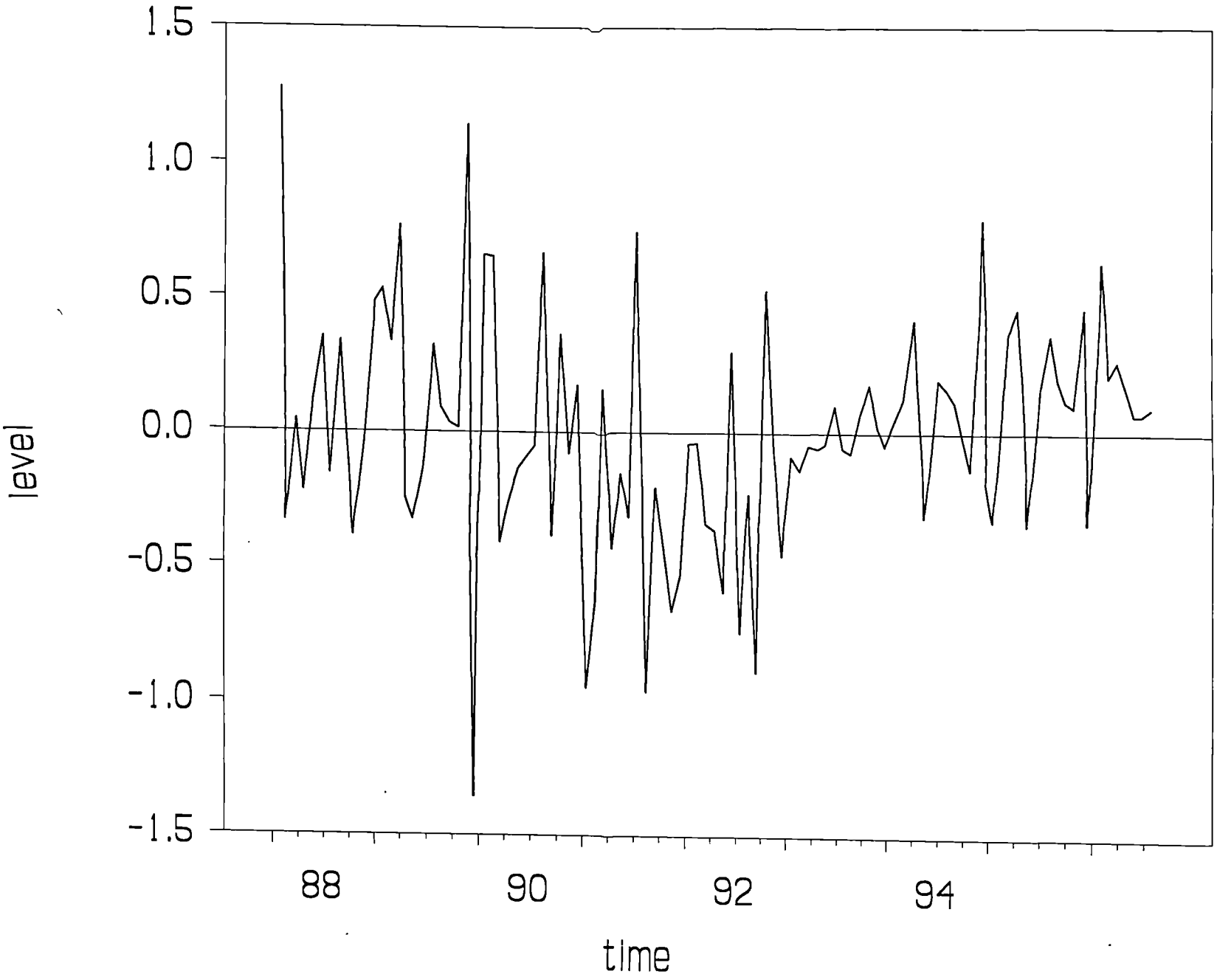


Figure 4. Residual



Chapter 5. An Affine Term Structure Model with Jumps

5.1. Introduction

Most continuous-time finance has been based on the assumption of *Itô* type processes for underlying state variables. However, one may question whether an assumption of the continuity of the price process is a good approximation of reality. Many studies such as Back (1991) have shown that the announcement of certain information is strongly related to jumps in the asset price. Back studied the term structure of interest rates, when state variables follow jump diffusion processes.

Studies based on the jump-diffusion models of the term structure of interest rates are relatively few compared with those of the diffusion model. One reason is that exact closed-form solutions for bond prices and derivative prices when state variables follow a mixed jump diffusion are hard to derive. A general equilibrium model allowing for jumps was first published by Ahn and Thompson (1989). Their model is an extension of CIR (1985b). They present an approximate solution of a pure discount bond using a square root process with jumps Recently, Das and Foresi (1996) assume that the short rate follows the Ornstein-Uhlenbeck process of Vasicek (1977), adding a jump shock. They derive an exact closed-form solution for the price of a pure discount bond. They follow the pricing kernel approach, which is also used in Constantinides (1992). Most studies of jump diffusion models assume that the local martingale part of a state variable has discontinuities (Back, 1991; Dothan, 1990). Basak (1995) examines the consumption-portfolio choice problem in a continuous time framework with jumps. He assumes that the state space is a continuous semimartingale, but allows for a price jump in the finite variation part.

In this chapter we study the general framework for a jump diffusion model of the term structure of interest rates with the affine term structure. We extend the Duffie and Kan (1996) affine model of the term structure of interest rates to the presence of jumps. The affine model of the term structure of interest rates is numerically tractable because this class of model is characterized by an exponential affine bond pricing function, rather than numerically solving a partial differential equation for the term structure. Our study investigates the conditions which must be satisfied for an affine type of term structure model of interest rates obtained in the presence of the jumps (Sections 2 and 3). In Section 4 we present approximate solutions of a pure discount bond price using our framework. In Section 5 we estimate a two-factor model with a jump term using a simplified approximate maximum likelihood method.

5.2. An Affine Model of the Term Structure With Jumps

We consider the possibility of sudden changes in a state vector X. Let (Ω, ζ, P) be the probability space, where Ω denotes a set of states of nature, ζ the σ-algebra subset of Ω, and P the probability measure on ζ. The filtration, $\{\zeta_t : t \geq 0\}$ to which the set of random variables is adapted defines the information set available to the agents. The filtration is right continuous and P-complete. We suppose that the state vector X is a "special" semimartingale, which means that the finite variation process in the decomposition can be taken to be predictable as in Back (1991)[30]. As in Duffie and Kan (1996), we assume further that the state vector X has values in some open set D of the n dimensional real space $\Re^n$. The market value $p_{t,\tau}$ at time t of a zero coupon bond maturing at time $t+\tau$ is given by $f(X_t, \tau)$, where $f \in C^2(D \times [0, \infty))$, i.e a twice differentiable function. The short

30. We consider only the case where the discontinuity of the state variables is allowed in the local martingale part, not in the finite variation part. This concept is best illustrated by reference to Poisson processes. Hence, we exclude the type of semimartingale whose finite variation part contains jumps as in Basak (1995).

rate process is a mapping $R : D \rightarrow \Re$ defined as the limit of yields as maturity goes to zero, or,

$$(1) \qquad R(x) = \lim_{\tau \to 0} \frac{-\ln f(x, \tau)}{\tau}, \quad x \in D .$$

As in Harrison and Kreps (1979) and Harrison and Pliska (1981), as well as Babbs and Webber (1994), some technical regularity[31] is required for the equivalence between the absence of arbitrage and the existence of an equivalent martingale measure Q.

Suppose that the state vector X satisfies a stochastic differential equation of the form,

$$(2) \qquad dX_t = v(X_t)\, dt + \sigma(X_t) dZ_t^* + y\, dN_t$$

where $v : D \rightarrow \Re^n$, $\sigma : D \rightarrow \Re^{n \times n}$ and $y : D \rightarrow \Re^n$ is a state independent distribution or fixed number. If the jumps depend on the state variable, we cannot construct an affine model of the term structure of interest rates[32]. Z_t^* is an n dimensional Wiener process and N_t is a Poisson distribution with intensity $\lambda^*(X_t)$ (i.e. , Prob($dN = 1$) = $\lambda^* dt$), where $\lambda^* : D \rightarrow IR^+$. One could add additional independent Poisson processes, but this would not change fundamentally our analysis.

The technical regularity condition implies that there exist standard Wiener processes and intensities under Q such that

$$(3) \qquad dX_t = \mu(X_t)\, dt + \sigma(X_t) dZ_t + y\, dN_t$$

31. See Babbs and Webber (1994).
32. See Section 5.3.

where

(3-1) $$Z^* = Z - \int \theta_i ds, \quad 1 \le i \le n, \quad \lambda = (1-\theta_{n+1})\lambda^*$$

and $\mu(X_t) : D \to \Re^n$, $\lambda(X_t) : D \to \Re^+$. θ_i, θ_{n+1} are assumed to be essentially bounded, predictable and adapted processes and specifically θ_{n+1} is strictly less than unity. Explicit formulas for θ_i, θ_{n+1}, which are functions of the indirect utility function, can be obtained in a general equilibrium context. A special case is available in the general equilibrium model of Bates (1989, 1996).

Following Duffie and Kan (1996), we investigate choices for (f, μ, σ, y) that are compatible, in that we have

(4) $$f(X_t, T-t) = E[\ \exp(-\int_t^T R(X_s)\ ds)\ ] \quad \textit{for almost all}, \quad 0 \le t \le T \le \infty$$

where E denotes expectation under the probability measure Q.

For example, Ahn and Thompson (1989) derived a term structure model in the presence of jumps in the CIR type economy setting. Following a different method from Ahn and Thompson, however, we build an affine model of the term structure of interest rates in the no-arbitrage context.

5.3. An Affine Factor Model with Jumps

We consider a class of compatible (f, μ, σ, y) with

(5) $$f(x, \tau) = \exp[\ A(\tau) + B(\tau)x\ ]$$

with boundary conditions

$$A(0)=0, \quad B(0)=0$$

where A and B are C^1 functions on $[\ 0,\infty)$. This is called an exponential-affine model of the term structure of interest rates. From (1), we know that R is an affine function on D.

Following Duffie and Kan (1996), we consider the zero coupon bond price process $p_t = F(X_t, t) \equiv f(X_t, \tau)$ for a fixed maturity T, where $\tau = T-t$ and $t \le T$. By $It\hat{o}$'s lemma,

(6) $$dp_t = D^*F(X_t, t)\ dt + F_x(X_t, t)\sigma(X_t)\ dZ + (F(x+y, t)-F(x, t))dN$$

where

$$D^*F(x, t) = F_t(x, t) + F_x(x, t)\ \mu(x) + \frac{1}{2}tr[\ F_{xx}(x, t)\sigma(x)\sigma(x)']$$

(6-1) $$E(dp_t) = DF(X_t, t)\ dt$$

where

$$DF(x, t) = F_t(x, t) + F_x(x, t)\ \mu(x) + \frac{1}{2}tr[\ F_{xx}(x, t)\sigma(x)\sigma(x)'] +$$
$$\lambda(x)\int_D[\ F(x+y, t)-F(x, t)]\ \ dv(\ y\ )$$

where $\lambda(X_t)$ is the risk adjusted arrival intensity of jumps in X at time t as in (3-1), and v is a fixed probability measure in $\Re^n$ defining the distribution of jumps, and $E([\ f(x+y, t)-f(x, t)\ ]\ dN) = \lambda(x)\int_D[\ F(x+y, t)-F(x, t)]\ \ dv(\ y\ )dt.$

From (5), we can calculate that

(7) $$DF(x,t)=F(x,t)\ (-A'(\tau)-B'(\tau)\ x)\ +\ F(x,t)B(\tau)\ \mu(x)x\ +$$
$$F(x,t)\frac{1}{2}\sum_{i=1}^{n}\sum_{j=1}^{n}B_i(\tau)B_j(\tau)\sigma_i(x)\sigma_j(x)'\ +$$
$$\lambda(x)\int_D[\ F(x+y,t)-F(x,t)]\ \ dv(\ y\ ).$$

By (4), F solves the following PDE:

(8) $$DF(X_t,t)-R(X_t)\ F(X_t,t)=0,$$

with a boundary condition $F(X_T,T)=1$.

From (7) we have

(9) $$-R(x)-A'(\tau)-B'(\tau)x+B(\tau)\mu(x)+$$
$$\frac{1}{2}\sum_{i=1}^{n}\sum_{j=1}^{n}B_i(\tau)B_j(\tau)\sigma_i(x)\sigma_j(x)'\ +$$
$$\lambda(x)\int_D[\ \frac{F(x+y,t)}{F(x,t)}\ -1]\ \ dv(\ y\ )=0.$$

From (5) we can calculate;

$$\frac{F(x+y,t)}{F(x,t)}\ =\exp\sum_j B_j(\tau)\ y_j.$$

Hence, (9) becomes

(10) $$-R(x)-A'(\tau)-B'(\tau)x+B(\tau)\mu(x)+$$
$$\frac{1}{2}\sum_{i=1}^{n}\sum_{j=1}^{n}B_i(\tau)B_j(\tau)\sigma_i(x)\sigma_j(x)' +$$
$$\lambda(x)\int_D[\ \exp(\sum_j B_j(\tau)y_j)\ -1\ ]\ dv(\ y\ )=0.$$

As in Duffie and Kan (1996), (10) implies that μ, $\sigma\sigma'$, and λ are affine functions on D under a mild nondegeneracy condition[33]. Following Duffie and Kan, we rewrite (10) as,

(11) $$\alpha(x,\tau)=\sum_i B_i(\tau)\mu_i(x)+\frac{1}{2}\sum_i\sum_j B_i(\tau)B_j(\tau)\beta_{i\,j}(x) +$$
$$\lambda(x)\int_D[\ \exp(\sum_j B_j(\tau)y_j)\ -1\ ]\ dv(\ y\ )$$

where $\alpha(x,\tau)=R(x)+A'(\tau)+B'(\tau)x$. Since R is affine, α is also affine. We let G be the function on D into $\mathcal{R}^N$, for $N = 2n+\frac{n^2-n}{2}+1$, denoted by

$$G(x)=(\mu_1(x),\mu_2(x),\ldots\ldots,\beta_{11}(x),\ldots\ldots\ldots\beta_{nn}(x),\lambda(x)).$$

We can regard (11) as a system of equations in τ and x of the form,

(12) $$\alpha(x,\tau)=\gamma(\tau)'G(x),\quad (x,\tau)\subseteq D\ x\ [\ 0,\infty)$$

where $\gamma:[\ 0,\infty)\rightarrow\mathcal{R}^N$. For example $\gamma_N=\int_D[\ \exp(\sum_j B_j(\tau)y_j)\ -1\ ]\ dv(\ y\ )$.

We can stack each of any N maturities $s_1,\ldots.s_N$ to obtain

(13) $$\gamma(s_1,\ldots s_N)'\ G(x)=(\alpha(x,s_1),\ldots\alpha(x,s_N))'.$$

33. The covariance matrix of the diffusion terms of the factors is non-singular.

If $\gamma(s_1, \ldots s_N)$ can be chosen to be non-singular, then G must be affine[34]. As in (11), if the jump distribution depends on the state variables, the state variables are contained in the exponential function. Accordingly, it is clear that we cannot construct an affine model of the term structure of interest rates in that case. Therefore, we require that the distribution of jump size does not depend on the state variables.

If $\mu(x)$, $\sigma(x)\sigma(x)'$, $\lambda(x)$ are affine in x, we can collect all terms in x_i into a form $[\ -B_i'(\tau)+\delta_i(B(\tau))]\ x_i$, where $\delta_i(B(\tau))$ have the form,

(14)

$$\delta_i(B(\tau)) = a + \sum_j b_j B_j(\tau) + \sum_{ij} c_{ij} B_i(\tau) B_j(\tau) + d_i \int_D [\ \exp(\sum_j B_j(\tau) y_j) - 1\ ]\ dv(\ y\ ).$$

where a, b_j are fixed coefficients for the drift term, c_j is for diffusion term, and d_j is a coefficient for a jump intensity. By (10) and the matching principle[35], we must have $-B_i'(\tau)+\delta_i(B(\tau))=0$ for i and τ. Accordingly, we get the following ODE (Ordinary Differential Equation):

(15) $$B_i'(\tau)=\delta_i(B(\tau)), \quad B(0)=0.$$

The term in (10) not involving x has the form $-A'(\tau)+\pi\ (B(\tau))$, where $\pi_i(B(\tau))$ has the same form as equation (14), except for different coefficients:

(16)

$$\pi_i(B(\tau))=a' + \sum_j b'_j B_j(\tau) + \sum_{ij} c'_{ij} B_i(\tau) B_j(\tau) + d'_i \int_D [\ \exp(\sum_j B_j(\tau) y_j) - 1\ ]\ dv(\ y\ ).$$

34. See the proof of the diffusion case in Duffie and Kan (1996) p. 386.
35. The matching principle means that for $\alpha \in \mathcal{R}^n$, $\beta \in \mathcal{R}^{n \times n}$, some set U which is a subset of $\mathcal{R}^n$, then if $\alpha+\beta x = 0$ for all $x \in U$, then $\alpha=0$ and $\beta=0$.

where a', b', c' and d' are fixed coefficients. Again, by the matching principle,

(17) $$A'(\tau)=\pi\ (B(\tau)), \qquad A(0)=0$$

The solution for $A(\tau)$ is

(18) $$A(\tau)=\int_0^{\tau} \pi\ (B(s))ds$$

where B solves (15).

In the next section, following Babbs and Webber (1994, 1995), we model a jump diffusion term structure model of official interest rates such as discount rates and Lombard rates. As Babbs and Webber (1994, 1995) argue, the influence of the monetary authority's interest rate policy on the term structure of interest rates has long been of interest to economists. In the next section, we try to model the affine term structure of interest rates, including official set rates[36] such as the discount rates. The official set rates sometimes play a role in the monetary authority's signalling to the markets its intentions on the future level of short rates. It is well known that the official rates follow a pure jumps process rather than a diffusion process.

5.4. An Example of an Affine Term Structure model with Jumps

In this section we apply the affine term structure model in the presence of jumps to the pricing of pure discount bonds. We present approximate solutions of a two-factor model in two cases. Recent papers by Babbs and Webber (1994, 1995) emphasize the rate-setting role of the monetary authorities in modelling the term structure of interest rates. Babbs and Webber (BW) find that changes in official

36. See the definition of official rates in Babbs and Webber (1994,1995)

interest rates are associated with jumps in the market rate. BW model the market's perceptions of the chances of changes in the officially-set short rate, r, in terms of stochastic jump intensities depending on the current level of r, and of a single additional state variable, x. They assume that x follows a diffusion, representing some indicator of the condition of the real or financial economy. Veruete and Webber (VW, 1994) also present a jump-diffusion model of the UK sterling short rate (LIBOR). They decompose LIBOR into the sum of a mean-reverting continuous part, and a jump process.

Following the results in BW and VW, we model a two-factor affine term structure model of interest rates in the presence of jumps as follows:

(19) $$dx_1 = (ax_1 + \delta x_2 + b)dt + \sigma \sqrt{\alpha x_1 + \beta}\, dz.$$

$$dx_2 = c\, dt + y\, dN.$$

where y is the distribution of jump size. We assume that the jump intensity λ is constant. The short rate is determined by :

$$r = \eta_0 + \eta_1 x_1 + \eta_2 x_2$$

In BW the first factor takes the form of the market's perception of a prospective or shadow level of interest rates, given the condition of the economy and the policy stance of the Government. We do not identify the factor as in BW. We can assume that the state variable follows a CIR or a Vasicek type diffusion process, depending on the choice of parameters α and β. The second factor is an officially set rate following a jump process. Here, we omit the diffusion term and the state dependent drift term. If $c=0$, then the state space is disconnected. We model the jump part as a Poisson process, assuming the size of jumps is either exponentially distributed or fixed in absolute value[37], and is independent of our two

37. We shall give these examples later.

factors. In addition, we assume that the jump intensity is constant. We can change the drift term of the first factor in an intuitive way as in BW. Let us suppose that the first factor takes the form

$$dx_1 = v(\mu - x_1)dt + \sigma\sqrt{\alpha x_1 + \beta}\, dz \tag{20}$$

where $\mu = (1-w)\bar{\mu} + w x_2$, $0 \le w \le 1$. This assumes that x_1 reverts to the level of a weighted average of a long run rate $\bar{\mu}$ and the current level of x_2. Then, comparing (19) with (20), we obtain $a = -v$, $\delta = v w$, $b = v(1-w)\bar{\mu}$.

From (19), we model the Ornstein-Uhlenbeck process of Vasicek ($\alpha = 0$, $\beta = 1$)[38]. We price a pure discount bond using the affine framework. We assume that the processes in (19) are the risk-adjusted processes. This means that we assume the θ_1 and θ_2 are given by (3-1)). Since we model an affine term structure of interest rates to derive bond prices, we can easily guess the functional form of the price of a pure discount bond as

$$f(x_1, x_2, \tau) = \exp[\, A(\tau) + B_1(\tau)\, x_1 + B_2(\tau)\, x_2] \,. \tag{21}$$

In this case, the short rate $R(x) = \eta_0 + \eta_1 x_1 + \eta_2 x_2$, where η_0, η_1, η_2 are determined in the general equilibrium context[39]. Then, from equations (14) and (16),

38. In the case of CIR process, we failed to obtain a closed form solution for a pure discounted bond prices as well as an approximate solution.

39. When we model the term structure of interest rates in the no-arbitrage setting, we usually assume that the market prices of risk of factors are given. When we fit the affine factor model of Duffie and Kan (1996) into the general equilibrium setting such as CIR and Bakshi and Chen (1996), the market price of factors is a function of the parameters of the utility function and of the economic variables (see p. 27, Nunes (1996)). Furthermore, the short rate can be expressed as a linear combination of factors. The coefficients of factors are not determined arbitrarily. The coefficients are determined by the other economic variables under the assumption (see p. 26 Nunes). However, we cannot know the coefficients of the factors from the knowledge of the market price of risk. Hence, as in Hodges and Pang (1995), once we model the term structure of interest rates under the no-arbitrage setting, we have to regard the coefficients as parameters. On the other hand, we can transform the affine factor model with the non-unity coefficients model into a model with

the functions $B_1(\tau)$, $B_2(\tau)$ and $A(\tau)$ are the solutions of the following differential equations,

$$(22) \quad B_1(\tau)' = -\eta_1 + a\, B_1(\tau), \quad B_1(0) = 0$$

$$(23) \quad B_2(\tau)' = -\eta_2 + \delta\, B_1(\tau)\ , \quad B_2(0) = 0$$

$$(24) \quad A(\tau)' = -\eta_0 + \frac{\sigma^2\beta}{2} B_1(\tau)^2 + b\, B_1(\tau) + c\, B_2(\tau) + \lambda Q(\tau),$$

$$A(0) = 0$$

where $Q(\tau)$ is involved with the jump term. We assume that jump intensity λ is constant. Hence, the jump term appears only in equation (24). Since the first factor is a Vasicek type process (in particular, the diffusion does not depend on the state variable), the square term of $B_1(\tau)$ does not appear in equation (22). But, since the first factor depends on the second factor, the $B_1(\tau)$ appears in equation (23). To obtain $Q(\tau)$, we take expectation operator in (6), assuming our two factors. Then

$$E(df(x_1, x_2, \tau)) = (\quad)dt + E((f(x_1, x_2+y, \tau) - f(x_1, x_2, \tau))dN).$$

The last term can be rewritten as follows:

$$E((f(x_1, x_2+y, \tau) - f(x_1, x_2, \tau))dN) = E(f(x_1, x_2+y, \tau) - f(x_1, x_2, \tau)|dN=1)E(dN)$$

$$= E(f(x_1, x_2+y, \tau) - f(x_1, x_2, \tau)|dN=1)\lambda dt$$

$$= Q(\tau)\lambda f(x_1, x_2, \tau)dt$$

---

unity coefficients, using the market price of risk (Babbs and Nowman, 1997). However, the numbers of parameters to be estimated econometrically do not decrease.

where $Q(\tau)$ is the instantaneous expected rate of change of the bond price if a jump occurs. Comparing the second line with the third line of the above equation, we get:

$$(25) \qquad Q(\tau) = E\left(\frac{(f(x_1, x_2+y, \tau)-f(x_1, x_2, \tau))}{f(x_1, x_2, \tau)} \,|\, dN=1\right)$$

$$= \int_D \left[\frac{(f(x_1, x_2+y, \tau)-f(x_1, x_2, \tau))}{f(x_1, x_2, \tau)} \right] dv(y).$$

We give two examples.

Example 1:

Following Das and Foresi (1996), the distribution of the jump size y, in "absolute value", is assumed to be exponential (mean jump size is $\frac{1}{p}$). The sign of the jump is assumed to be positive with the probability q and negative with the probability $1-q$. Then, we can obtain $Q(\tau)$ as :

$$Q(\tau) = \frac{\int_0^\infty [\ q(f(x_1, x_2+y, \tau)-f(x_1, x_2, \tau))]\ p\exp(-py)dy}{f(x_1, x_2, \tau)}$$

$$+ \frac{\int_0^\infty [\ (1-q)(f(x_1, x_2-y, \tau)-f(x_1, x_2, \tau)]\ p\exp(-py)dy}{f(x_1, x_2, \tau)}.$$

Substituting equation (21) into the above equation, we obtain the following expression:

$$Q(\tau) = \int_0^\infty q(\exp(B_2(\tau)y)-1)p\exp(-py)dy$$
$$+ \int_0^\infty (1-q)(\exp(-B_2(\tau)y)-1)p\exp(-py)dy.$$

If we calculate the integral in the above expression explicitly, assuming[40] that $|B_2(\tau)| < p$,

(26) $$Q(\tau) = \frac{(1-q)\ B_2(\tau)}{p-B_2(\tau)} - \frac{q\ B_2(\tau)}{p+B_2(\tau)},$$

Note 1: When we choose a distribution of the jump size, there are only a few distributions which are integrable in a closed form in $Q(\tau)$. For instance, when we assume the normal distribution of the jump size, we cannot obtain a closed-form expression for $Q(\tau)$.

Example 2:

Similarly, if the jump size y is fixed in absolute value, and the sign of the jump is assumed to be positive with the probability q and negative with the probability $1-q$, then

(27) $$Q(\tau) = q\ e^{B_2(\tau)\ y} + (1-q)\ e^{-B_2(\tau)\ y} - 1.$$

We price a pure discounted bond under our two different set of assumptions of jump distribution (example 1 and example 2). Proposition 1 assumes that the size in absolute value of jumps is distributed exponentially. Proposition 2 assumes that the size of jumps is fixed in absolute value. In Proposition 1 and 2, we present approximate solutions for the price of a pure discounted bond. We shall see the

40. This assumption is required in the derivation of the formula of Das and Foresi (1996) for a pure discount bond price, although they did not state the assumption explicitly.

accuracy of the two approximation solutions after the proof of Proposition 2.

Proposition 1:

Suppose that our two semimartingale factors have processes

$$dx_1 = (ax_1+\delta x_2+b)dt + \sigma dz.$$

$$dx_2 = c\, dt + y\, dN,$$

i.e. $\alpha=0$ and $\beta=1$ in (19). As in example 1, the distribution of the jump size y, in absolute value, is assumed to be exponential with mean jump size $\frac{1}{p}$. The sign of the jump is assumed to be positive with probability q and negative with probability $1-q$. If the term $a\tau$ is sufficiently small, then the price of a pure discount bond at time t that promises to pay one pound at maturity $t+\tau$ is,

$$f(x_1, x_2, \tau) = \exp[\ A(\tau) + B_1(\tau)\ x_1 + B_2(\tau)\ x_2]$$

where

$$B_1(\tau) = -\frac{\eta_1}{a}(e^{a\tau} - 1)$$

$$B_2(\tau) = (-\eta_2 + \frac{\eta_1 \delta}{a})\ \tau - \frac{\eta_1 \delta}{a^2}\ (e^{a\tau}-1)$$

$$A(\tau) = -\eta_0 \tau + \frac{\sigma^2}{2} \frac{\eta_1^2}{a^2} \left(-\frac{2 e^{a\tau}}{a} + \frac{e^{2a\tau}}{2a} + \tau\right)$$

$$-\frac{b \eta_1}{a}\left(\frac{e^{a\tau}}{a} - \tau\right) + \frac{c}{2}\left(-\eta_2 + \frac{\eta_1 \delta}{a}\right)\tau^2$$

$$-\frac{c \eta_1 \delta}{a^2}\left(\frac{e^{a\tau}}{a} - \tau\right)$$

$$+ \lambda \left[-\tau + \frac{2p\, q\, ArcTanh\left[\frac{\eta_2 + \eta_1 \delta \tau}{\sqrt{\eta_2^2 + 2p\eta_1\delta}}\right]}{\sqrt{\eta_2^2 + 2p\eta_1\delta}} - \frac{2p\,(1-q)\, ArcTanh\left[\frac{\eta_2 + \eta_1 \delta \tau}{\sqrt{\eta_2^2 - 2p\eta_1\delta}}\right]}{\sqrt{\eta_2^2 - 2p\eta_1\delta}} \right]$$

$$+ \frac{3 \sigma^2 \eta_1^2}{4 a^3} + \frac{b \eta_1}{a^2} + \frac{c \delta \eta_1}{a^3}.$$

$$- \lambda \left[\frac{2p\, q\, ArcTanh\left[\frac{\eta_2}{\sqrt{\eta_2^2 + 2p\eta_1\delta}}\right]}{\sqrt{\eta_2^2 + 2p\eta_1\delta}} - \frac{2p\,(1-q)\, ArcTanh\left[\frac{\eta_2}{\sqrt{\eta_2^2 - 2p\eta_1\delta}}\right]}{\sqrt{\eta_2^2 - 2p\eta_1\delta}} \right]$$

Proof:

The first factor is basically a Vasicek type process. Equations (22) to (24) become:

(22-2) $\quad B_1(\tau)' = -\eta_1 + a\, B_1(\tau), \quad B_1(0) = 0$

(23-2) $B_2(\tau)' = -\eta_2 + \delta B_1(\tau)$, $B_2(0) = 0$

(24-2) $A(\tau)' = -\eta_0 + \frac{\sigma^2}{2} B_1(\tau)^2 + b B_1(\tau) + c B_2(\tau) + \lambda Q(\tau)$,

$A(0) = 0$.

First, we solve equation (22-2). This is a first-order linear differential equation. Then, the solution with the boundary condition $B_1(0) = 0$ is

(28) $$B_1(\tau) = -\frac{\eta_1}{a}(e^{a\tau} - 1).$$

Secondly, we solve equation (23-2). The solution with the boundary condition $B_2(0) = 0$ is,

(29) $$B_2(\tau) = (-\eta_2 + \frac{\eta_1 \delta}{a})\tau - \frac{\eta_1 \delta}{a^2}(e^{a\tau} - 1).$$

Finally, we solve the $A(\tau)$. From (26), the jump part is

(30) $$\lambda Q(\tau) = \lambda \left[\frac{(1-q) B_2(t)}{p - B_2(t)} - \frac{q B_2(t)}{p + B_2(t)} \right].$$

Unfortunately, the integration of (30) has no closed-form solution. To integrate equation (30), we use a Taylor approximation. If we use the approximation:

(31) $$B_2(\tau) = (-\eta_2 + \frac{\eta_1 \delta}{a})\tau - \frac{\eta_1 \delta}{a^2}(e^{a\tau} - 1)$$

$$= \frac{\eta_1 \delta}{a^2} + (-\eta_2 + \frac{\eta_1 \delta}{a})\tau - \frac{\eta_1 \delta}{a^2} \sum_{n=0}^{\infty} \frac{(a\tau)^n}{n!}$$

$$\cong \frac{\eta_1\delta}{a^2} + \left(-\eta_2 + \frac{\eta_1\delta}{a} \right) \tau - \frac{\eta_1\delta}{a^2}\left(1 + a\tau + \frac{(a\tau)^2}{2}\right)$$

$$= -\eta_2\tau - \frac{\eta_1\delta}{2}\tau^2,$$

by integrating equation (24-2) using (31), we obtain the solution.

■

Proposition 2:

Suppose that our two semimartingale factors have processes

$$dx_1 = (ax_1 + \delta x_2 + b)dt + \sigma dz.$$

$$dx_2 = c\, dt + y\, dN,$$

and, the jump has a fixed size $+y$ with probability q and $-y$ with probability $(1-q)$. If the $B_2(\tau)y$ term is sufficiently small, then the price of a pure discount bond at time t that promises to pay one pound at maturity $t+\tau$ is

$$f(x_1, x_2, \tau) = \exp[\ A(\tau) + B_1(\tau)\ x_1 + B_2(\tau)\ x_2]\ .$$

where

$$B_1(\tau) = -\frac{\eta_1}{a}(e^{a\tau} - 1)$$

$$B_2(\tau) = \left(-\eta_2 + \frac{\eta_1\delta}{a}\right)\tau - \frac{\eta_1\delta}{a^2}(e^{a\tau} - 1)$$

$$A(\tau) = -\eta_0 \tau + \frac{\sigma^2}{2} \frac{\eta_1^2}{a^2} \left(-\frac{2 e^{a\tau}}{a} + \frac{e^{2a\tau}}{2a} + \tau\right)$$

$$-\frac{b \eta_1}{a}\left(\frac{e^{a\tau}}{a} - \tau \right) + \frac{c+\lambda y(2q-1)}{2} \left(-\eta_2 + \frac{\eta_1 \delta}{a}\right)\tau^2$$

$$-\frac{[\ c+\lambda y(2q-1)]\ \eta_1\ \delta}{a^2} \left(\frac{e^{a\tau}}{a} - \tau \right)$$

$$+ \frac{3\ \sigma^2\ \eta_1^2}{4\ a^3} + \frac{b\ \eta_1}{a^2} + \frac{[\ c+\lambda y(2q-1)]\ \ \delta\ \eta_1}{a^3}.$$

Proof:

Since the solution of $B_1(\tau)$ and $B_2(\tau)$ is the same as in Proposition 2, we solve only for $A(\tau)$. If the jump has a fixed size in absolute value, then

(32) $$\lambda\ Q(\tau) = \lambda\ [\ q\ e^{B_2(\tau)\ y} + (1-q)\ e^{-B_2(\tau)\ y} - 1]\ .$$

If we use a Taylor expansion of $e^{B_2(\tau)y}$, the integration of (32) has a closed form. Since it is an integration of an infinite sum, this is not practical. Using the Taylor expansion of $e^{B_2(\tau)y}$ up to the second term, assuming $B_2(\tau)y$ is sufficiently small, then

$$\begin{aligned} \lambda\ Q(\tau) &= \lambda[\ q(1+B_2(\tau)y)+(1-q)(1-B_2(\tau)y)-1] \\ &= \lambda[\ (2q-1)\ B_2(\tau)y]. \end{aligned}$$

Integrating the remaining term, we obtain the solution.

■

This approximation has the same effect as replacing the jump in a bond price,

$$F(X+y, t) - F(X, t)$$

with $F(X, t) \ln \frac{F(X+y, t)}{F(X, t)}$ as in Basak (1995). The accuracy of this approximation in Proposition 1 and Proposition 2 would be high only if a pure discount bond is priced at a relatively short maturity and the parameters are small.

We investigate the size of errors in the case of Propositions 1 and 2. We compare the values of two spot yields as an exact integration and as the approximate integration for $A(\tau)$. Since $B_1(\tau)$ and $B_2(\tau)$ can be integrated in a closed form, the only difference between the two integrations is the jump part of $A(\tau)$. Let $f(x_1, x_2, \tau)$ be the bond price in integrating $A(\tau)$ numerically, and $\bar{f}(x_1, x_2, \tau)$ be our formula for approximate $A(\tau)$. Hence the percentage difference of the two spot yields is

$$\frac{\ln[\ f(x_1, x_2, \tau)] - \ln[\ \bar{f}(x_1, x_2, \tau)]}{\ln[\ f(x_1, x_2, \tau)]}$$

$$= \frac{[\ A(\tau)+B_1(\tau)x_1+B_2(\tau)x_2] - [\ \bar{A}(\tau)+B_1(\tau)x_1+B_2(\tau)x_2]}{[\ A(\tau)+B_1(\tau)x_1+B_2(\tau)x_2]}.$$

where $\bar{A}(\tau)$ which is a part of the formula $\bar{f}(x_1, x_2, \tau)$ is the approximation solution for equation (24). We integrate $f(x_1, x_2, \tau)$ numerically using Mathematica. To remove the effect of other parameter values, the values of η_0, σ, b, and c are assumed to be zero. Table 1 shows the result of the difference between the two integrations in the first model. Table 2 shows the results in the second model.

Table 1. Errors Percentage of the Interest Rate Difference in Model 1

Stochastic Process

$$r = \eta_0 + \eta_1 x_1 + \eta_2 x_2$$
$$dx_1 = (\ a x_1 + \delta\ x_2 + b\ )dt\ + \sigma\ dz.$$
$$dx_2 = c\ dt\ +\ y\ dN$$

Base Case .

$\delta = 0.05, \eta_0 = 0,\ \eta_1 = 0.7,\ \eta_2 = 0.1,\ \sigma = b = c = 0,$

$q = 1, \frac{1}{p}(\ mean\ jump\ size) = 5\%,\ \lambda = 1,\ x_1 = 0.08,\ \ x_2 = 0.08$

	$\tau = 5$	$\tau = 8$	$\tau = 10$
$a = -0.1$	-0.013 %	-0.069 %	-0.069 %
$a = -0.5$	-0.024 %	-0.175 %	-0.317 %
$a = -1$	-0.032 %	-0.165 %	-0.332 %

Table 2. Errors Percentage of the Interest Rate Difference in Model 2

Stochastic Process

$$r = \eta_0 + \eta_1 x_1 + \eta_2 x_2$$
$$dx_1 = (ax_1 + \delta x_2 + b)dt + \sigma dz.$$
$$dx_2 = c\,dt + y\,dN$$

Base Case .

$a = -0.5, \delta = 0.05, \eta_0 = 0, \eta_1 = 0.7, \eta_2 = 0.1,$
$\sigma = b = c = 0, \quad q = 1, \lambda = 1, \quad x_1 = 0.08 \quad x_2 = 0.08$

	$\tau=5$	$\tau=8$	$\tau=10$
$y = 0.5\%$	4.614 %	6.361 %	10.096 %
$y = 1\%$	8.781 %	15.042 %	10.903 %
$y = 4\%$	34.310 %	40.217 %	55.002 %

As seen in Table 1, the error for the approximate solution of Proposition 1 is negligible for a small value of a. We might use this formula for a pure discounted bond price with small value of a and relatively short maturity However, unlike our expectation, model 2 (Table 2) shows quite large errors, in particular, for the large jumps. The approximation is used by Basak (1995), and Ahn and Tompson (1989), who did not provide the accuracy of their approximation.

5.5. An Estimation of the Two-Factor Jump-Diffusion Model

5.5.1. An Approximate Maximum Likelihood Estimation

The empirical literature on the jump diffusion process is sparse. Beckers (1981) estimates the parameters of the diffusion jump model of stock returns. Since the log likelihood function and its derivative are highly non-linear and involve infinite sums, he employs a version of the method of moments known as the cumulant matching method as an alternative to maximum likelihood. This method sometimes produces negative values for variance parameters. Ball and Torous (1983b) use a similar method. However, they introduce a Bernoulli jump process, and show that the density of a Bernoulli mixture of Gaussian and of a Poisson mixture of Gaussian process are practically indistinguishable for small values of the jumps intensity. Accordingly, they can simplify the density function. They implement both the maximum likelihood method and the method of cumulants for the simplified density. They find that the maximum likelihood estimates of the variance parameters are consistently positive, whereas the estimates of the method of cumulants variance are not. Ball and Torous (1985) use the maximum likelihood method for an original density. Since the log likelihood function involves an infinite sum, they truncate the infinite sum so that sufficient accuracy is achieved. More recently, Hamilton (1988) models a discrete regime shift in the spot rate process. He models a spot interest rate process that can shift randomly between two or more regimes. The model allows there to be different parameters in the different regimes. This approach, however, has difficulties in determining the number of states.

In this section we estimate two-factor jump diffusion model with the approximate maximum likelihood method for the simplified density[41] (Bernoulli mixture Gaussian) as in Ball and Torous (1983b).

41. There are only a few cases with more than two jumps in one month in our data set. To preserve the parsimony of parameters, we use the method of Ball and Torous.

First, in our model, the spot rate has the process

(33) $$r = \eta_0 + \eta_1 x_1 + \eta_2 x_2$$

where

(34) $$dx_1 = (ax_1 + \delta x_2 + b)dt + \sigma\, dz$$

is the continuous part of the spot rate. Equation (34) is a risk-adjusted process. We can change the drift term of the first factor in a intuitive way as in BW. Let us suppose that the first factor takes the form

(34-1) $$dx_1 = v (\mu - x_1)dt + \sigma\, dz$$

where $\mu = (1-w)\bar{\mu} + w\, x_2$, $0 \le w \le 1$. This assumes that x_1 reverts to the level of a weighted average of a long run rate $\bar{\mu}$ and the current level of x_2. Then, comparing (34) with (34-1), we obtain $a = -v$, $\delta = v\, w$, $b = v\,(1-w)\,\bar{\mu}$. We assume that the second factor x_2 has :

$$dx_2 = y\, dN\ ,\ x_1\ , x_2 \in \Re\ ,$$

where the risk-adjusted jump intensity λ is constant. In addition, we assume that the jump size y is normally distributed with mean ν and variance s^2. This assumption is purely for the simplification for the estimation of the model as shall be mentioned in Note 2. In our model, the jump term appears in the second factor.

Note 2 : Depending on the assumption of the distribution of the jump size, we might change the estimation method. For instance, if we assume that the size of jumps is exponentially distributed, then we cannot use MLE, since we do not know the joint density of dx_1 and dx_2. As seen in Figure 2, the jump size of discount

rates in our data is not uniformly distributed. Here, we assume that the jump size y, is normally distributed Most of all, this assumption simplifies the likelihood function. This means that we can use the maximum likelihood method rather than, for example, GMM.

First, equation (34) is substituted into (33) after first differencing of (33). Then we obtain,

(35) $$dr = \eta_1 dx_1 + \eta_2 dx_2$$

$$= \eta_1 (ax_1 + \delta x_2 + b)\, dt + \eta_1 \sigma\, dz + \eta_2 y\, dN .$$

The density of dr is a Poisson mixture of Gaussian densities. However, since the Poisson mixture densities involves an infinite sum, the infinite sum has to be truncated. Following Ball and Torous (1983b, 1985), we approximate dN as a Bernoulli process. Then dr is a Bernoulli mixture of Gaussian densities. It is known that for small values of the jump intensity λ, the two densities are nearly the same. Since a jump is a rare event, we assume that there is at most one jump in any period up to say a month, and we use a Bernoulli mixture of Gaussian densities in applying MLE. We denote the Bernoulli mixture of Gaussian densities by $m(x : \Phi)$, where Φ is a parameter set. Unfortunately, in the specification of (35), we can not estimate η_1 and other parameters separately. To obtain η_0, η_1 and η_2, we assume the following model (33) first.

(36) $$r = \eta_0 + \eta_1 x_1 + \eta_2 x_2 + w$$

where w is assumed to be a white noise[42].

Hamilton (1994) shows that most of interest rates process usually contain unit

42. This is for estimating the model using OLS.

roots. As seen in (36), if short rate, x_1 and x_2 are all I(1) variables and w is I(0) variable by assumption, then the specification (36) becomes a spurious regression problem as mentioned in Granger and Newbold (1974). However, even if r, x_1 and x_2 are all I(1), the specification of (36) can be used if the variables are cointegrated. Before estimating the specification of (36) by OLS, therefore, we have to check the cointegration relationships between the variables. Otherwise, the parameters estimated might have some bias due to the unit root. We will perform these tests.

To explain the estimation technique of (35), we briefly review the approach of Ball and Torous (1983b). Ball and Torous (1983b) show that if a security has the process

$$\frac{dM}{M} = \alpha\, dt + \sigma\, dz + (e^{y}-1)dN$$

and the jump size is distributed as $\ln y \sim N(\nu, s)$, then for a small value for λ, the density function $m(x : \Phi)$ of the security return $\ln \frac{M_t}{M_{t-1}}$ in the unit time interval becomes

$$m(x : \Phi) = (1-\lambda)\phi(\alpha, \sigma^2) + \lambda\phi(\alpha+\nu, \sigma^2 + s^2), \tag{37}$$

where $\phi(\nu, s^2) = \frac{1}{\sqrt{2\pi s^2}} \exp[-\frac{(x-\nu)^2}{2s^2}]$.

To simplify the estimation, Ball and Torous (1983b) assume that the mean value of jump ν is zero. Here, we do not assume ν is zero. To use the method of Ball and Torous, we have to modify (37) since in our model the drift of the short rate depends on the two state variables. To modify (37) for our model, we assume that there are two different regimes or situations for dr, similar to the regime shifting model of Hamilton (1994). The first one is that dr follows the diffusion process,

(38) $$dr = \eta_1(ax_1 + \delta x_2 + b)\ dt + \eta_1 \sigma\ dz.$$

The second one is that dr follows jump-diffusion process

(39) $$dr = \eta_1(ax_1 + \delta x_2 + b)\ dt + \eta_1 \sigma\ dz + \eta_2 y\ dN.$$

The problem is how we can model the jump term econometrically. To see the effect of jumps (the mean value of jumps), we introduce a dummy variable. If there is a jump, we denote it by 1, otherwise 0. Suppose we observe the series at discrete unit time intervals. We discretize (38) and (39) as in the following equations,

(38-1) $$\Delta r = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \varepsilon_{t1}.$$

(39-1) $$\Delta r = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_3 D + \varepsilon_{t2}.$$

where ε_{t1} is normally distributed with mean zero and variance s_1^2, and ε_{t2} is normally distributed with mean zero and variance s_2^2, and where

$$\beta_0 = \eta_1 b$$
$$\beta_1 = \eta_1 a$$
$$\beta_2 = \eta_1 \delta$$
$$\beta_3 = \eta_2 E(y).$$

In (39-1), if there is no jump, then $D=0$ and (39-1) is exactly same as (38-1). On the other hand, if there is a jump, then $D=1$ and (39-1) becomes

(40) $$\Delta r = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_3 + \varepsilon_{t2}.$$

Formally, if there is no jump, then

(41) $$E(\Delta r \mid no\ jump) = \beta_0 + \beta_1 x_1 + \beta_2 x_2$$

and if there is jump,

(42) $$E(\Delta r \mid jump) = \beta_0 + \beta_3 + \beta_1 x_1 + \beta_2 x_2.$$

In this model, β_3 is the difference between the expected change of interest rates in the two models (41) and (42). This is interpreted as the effect of the average value of the jumps of x_2 on Δr. It is important to note that the dummy variable in the specification of (39-1) is actually the etimation of the jump size of Δr when there are jumps in x_2 rather than a estimation of the parameters of the jump size of x_2.

If (39-1) is a true specification of the short rate, the variance of ε_2 is greater than ε_1 since variance of ε_2 contains the variance of the jumps. Assuming that there is no correlation between dz and dN, we can write

(43) $$var(\varepsilon_2) = var(\varepsilon_1) + \eta_2^2 var(y)$$

where var means variance, and y is the random variable of the jump size. Hence $var(\varepsilon_1) = \eta_1^2 \sigma^2$, and $var(y) = \dfrac{var(\varepsilon_1) - var(\varepsilon_2)}{\eta_2^2}$.

We set up a likelihood function as follows

(44) $$m(\Delta r; \Phi) = \frac{(1-\lambda)}{\sqrt{2\pi}s_1} \exp\left(\frac{-(\Delta r - \beta_0 - \beta_1 x_1 - \beta_2 x_2)^2}{2s_1^2}\right)$$

$$+ \frac{\lambda}{\sqrt{2\pi}s_2} \exp\left(\frac{-(\Delta r - \beta_0 - \beta_1 x_1 - \beta_2 x_2 - \beta_3 D)^2}{2s_2^2}\right).$$

The log-likelihood function is

(45) $$\ln L(\Delta r : \Phi) = \sum_{i=1}^{N} \ln[\ m(\Delta r_i : \Phi)]$$

where $\Phi = \{\beta_0, \beta_1, \beta_2, \beta_3, \lambda, s_1, s_2\}$.

If the jump intensity is interpreted as an ergodic probability for a Markov chain, the likelihood function (44) is similar to equation (22.3.5) of Hamilton's (1994) change in regime framework. Our model is, as a special case, a two-state regime of Hamilton (1990). Hamilton's changing regime framework is reduced to a simplified specification of the density in the presence of a multi-factor Poisson process model.

The mixture density is known to have the property that a global maximum of the log-likelihood does not exist (Hamilton, 1994). In such a case, numerical maximisation produces a reasonable local maximum (Keifer, 1978). Table 5 reports the results of estimates. We shall explain the empirical result in the next section.

As Dahlquist (1996) states, the interest rates with shorter maturities sometimes contain idiosyncrasies (i.e. seasonalities and high bid-asked spread), the empirical tests are performed on the monthly data of the Federal Fund Rate (FFR) and the Discount Rate (DR) from February 1982 to May 1997. For this variable, there is no change unless there is a jump. Thus, if there is a change, we know there was at least one jump within the period. We cannot distinguish multiple small jumps from a single large jump within such a period, but fortunately multiple jumps within one

month are very infrequent. If there should be jumps of equal size in opposite directions within the month, they would not be detected, but such events are exceedingly rare. In our data, there are only two cases of more than one jump for a month. We acknowledge that we lose this information with the use of monthly data. However, considering the fact that there are so few cases in our data set, this is not expected to affect our result siginificantly. Figure 1. shows the FRR and DR from 1981 to 1997. Following Babbs and Webber (1995), we choose the FFR as a proxy for the short rate r and the DR as a proxy for the second jump factor x_2. The difference between FFR and DR is chosen as a continuous first factor x_1, which is similar to Veruete and Webber (1994), and Balduzzi, Bertola and Foresi (1993).

5.5.2 Empirical Results

In this section, we present our empirical results. Table 3 shows the summary statistics of FFR, DR and FFR-DR. The distribution of the monthly changes in FFR displays excessive kurtosis. This is a natural implication of the jump-diffusion process of FFR. To estimate (36), we perform the unit root test since the OLS on the variables integrated of different orders can produce spurious regression. First, we test for the presence of unit roots[43] for $r = FFR$, $x_1 = FFR - DR$ and $x_2 = DR$, following Augmented Dickey Fuller (ADF) test. There are several methods to test for unit roots. For instance, Phillips and Perron (1988) developed a generalization of the Dickey and Fuller procedure that allows for mild assumptions concerning the distribution of errors such as a first-order moving average process. The discussion of the power of the tests is available in Enders (1995). However, Said and Dickey (1984) showed that the ADF test can be used even when the error follows a moving average process. Hence, we use the ADF test. Following the suggestion of Enders (1995), we start with a relatively large lag length. The estimation process is then

43. x_1 is the difference between FFR and DR. To test unit roots for the variable, we take the first difference of x_1.

repeated, reducing the lag length. It was found that the results for the unit root test were not changed significantly by reducing the lag length. Thus, we report the results of the first test, using lags of six month.

As expected, Table 4 shows that γ_1s for FFR and DR are insignificant based on the Dickey and Fuller statistics. This implies that FFR and DR have unit roots. On the other hand, the difference x_1 between FFR and DR does not have a unit root in our data set. There is a strong cointegration relationship between FFR and DR in our data set although we do not report it here. Unfortunately, because x_1 is an I(0) variable, and r and x_2 are I(1) variables, we cannot use OLS for the specification of (36). If we take the first difference for all variables, x_1 becomes an overdifferenced variable I(-1). This can also produce a spurious regression since the variables are integrated of different orders. Furthermore, we cannot estimate η_0 in the case of OLS for the differenced variables. To check the power of the ADF test, we regress FFR on x_1. As shown in Table 4, FFR is an I(1) variable and x_1 is an I(0) variable. Then the residuals from this regression should contain a unit root. However, the ADF test result, not reported here, reveals stationarity for the residuals. As Enders (1995) mentioned, the power of the ADF and the Phillips and Perron tests for the unit root is very low.

To illustrate an example for identifying the individual parameters of our model, we try OLS without taking the first difference. We acknowledge that this is not a good estimation method[44]. Because of the cointegration relationship between r and x_2, the super-consistent estimate η_2 is a stable parameter. But the estimate η_1 might be changed depending on the period of data set. The dramatic change of the estimated values of the coefficients and the very low value for the Durbin-Watson statistic (around zero) are known to be the usual symptoms of spurious regression. However, the Durbin-Watson statistic (2.01) for our OLS shows no serial correlation

44. Unfortunately, we cannot find any other estimation method for our affine interest rates model.

for the residuals. We also report the test result of stationarity of the residuals for the OLS. As in Table 5, the coefficient a_1 lies between -2 and 0. Accordingly, we can conclude that the residual is stationary. Based on this test and DW, we do not detect clearly the typical symptoms of spurious regression in our model.

Before we report the estimation results, we shall compare our model with a pure diffusion model (without the jump term). To compare our jump-diffusion model (39) with a pure diffusion model (38), we perform the LR test. We compare the two values of the likelihood function (our jump-diffusion model (39) and a pure diffusion model (38)). Table 9 shows the estimation result of (38). Twice the difference between the value of the likelihood functions, LR statistics is 597. Moving from the jump-diffusion model to the pure diffusion model increases the value of the likelihood function. At the 1% significance level, the jump-diffusion model rejects the pure diffusion model. This result implies that our jump-diffusion model performs better than the pure diffusion model.

Table 6. reports the parameter estimates, and the standard errors of the maximum likelihood estimation from our model. We use the BHHH method to maximize the likelihood function. Table 8 reports the recovering of the original parameters, using Table 7. As seen in Table 8, the mean reversion speed v of x_1 is 0.508. We assumed that the first factor x_1 reverts to the level of a weighted average of a long run rate $\bar{\mu}$ and the current level of the second factor x_2. In our data set $\bar{\mu}$ is around 1.3%, and w is 0.781. This implies that discount rate x_2 is more important than the long run rate to explain the mean reversion level. On the other hand, the jump intensity λ is 0.387. The mean and variance of jump size is -0.483 % and 0.002, respectively. The variance of jump size is very small in our data. This implies that the distribution of jump size has a high spike around the mean value. This might conform to the assumption of Babbs and Webber (1995), where the authorities are assumed to change interest rates in equal increments or decrement. Our results are slightly different from the arithmetic mean and variance of the jump size of discount rates. We plot the jump size of discount rates in Figure 2. First,

to simplify our estimation, we assumed that the jump size is normally distributed, and this approximation could affect the estimation results. Second, the arithmetic mean and variance of the jump size of the discount rates are literally the statistics of the jump size of discount rates. However, as mentioned in Section 5.5.1, the dummy variable in the specification of (39-1) is actually analysis of the short rates when there are jumps in discount rates.

Table 3. Statistics of Data (monthly data) (unit,%)

	Mean	Variance	Skewness	Kurtosis
Fed Rate (r)	7.76	12.40	1.18	1.57
Discount Rate (x_2)	6.62	7.23	0.92	0.67
FRR-DR (x_1)	1.14	1.52	2.54	9.82

Table 4. Unit Root Test for FFR, DR and x_1

(i) $$\Delta DR_t = \gamma_1 DR_{t-1} + \sum_{i=2}^{6} \gamma_i \Delta DR_{t-i+1}$$

Parameter	γ_1	γ_2	γ_3	γ_4	γ_5	γ_6
Estimates	-0.008	0.118	0.141	0.062	-0.062	-0.010
Standard Error	0.002	0.074	0.071	0.070	0.070	0.069

(ii) $$\Delta FFR_t = \gamma_1 FFR_{t-1} + \sum_{i=2}^{6} \gamma_i \Delta FFR_{t-i+1}$$

Parameter	γ_1	γ_2	γ_3	γ_4	γ_5	γ_6
Estimates	-0.023	-0.035	0.028	-0.098	-0.014	0.012
Standard Error	0.010	0.079	0.086	0.084	0.081	0.074

(iii) $$\Delta x_{1t} = \gamma_1 x_{t-1} + \sum_{i=2}^{6} \gamma_i \Delta x_{1t-i+1}$$

Parameter	γ_1*	γ_2	γ_3	γ_4	γ_5	γ_6
Estimates	-0.298	-0.049	0.029	-0.058	0.009	0.029
Standard Error	0.050	0.078	0.090	0.093	0.086	0.073

(1) * indicates significance at the 1% level.

(2) ADF statistics is used for the significant test

Table 5. Stationarity Test for the Residuals

$$\hat{e}_t = FFR_t - \widehat{\beta_0} - \widehat{\beta_1} x_1 - \widehat{\beta_2} x_2$$

(iii) $$\Delta \hat{e}_t = a_1 \hat{e}_{t-1} + \sum_{i=2}^{6} a_{i+1} \Delta \hat{e}_{t-i} + \varepsilon_t$$

Parameter	a_1*	a_2	a_3	a_4	a_5	a_6
Estimates	-1.006	0.033	0.138	0.118	0.115	0.104
Standard Errors	0.074	0.069	0.091	0.098	0.087	0.065

(1) * indicates significance at the 1% level.

(2) Assuing the cointegration relation between FFR, x_1 and x_2 Engle-Yoo statistics is used for the significant test

Table 6. Estimation of Parameters (Bernoulli Mixture Gaussian)

Parameter	β_0*	β_1*	β_2*	β_3*	λ*	s_1*	s_2*
Estimate	0.042	-0.149	0.117	-0.561	0.387	0.038	0.041
Standard Error	0.008	0.001	0.001	0.038	0.152	0.001	0.001

(1) Degree of Freedom 188.

(2) * indicates significance at the 1% level.

Table 7. Estimate of η_0, η_1 and η_2 (OLS)

Parameters	η_0	η_1*	η_2*
Estimates	-0.287	0.295	1.162

(1) DW (Durbin-Watson Statistics) = 2.01

Table 8. Estimates of the Original Parameters

$$dr = \eta_1 v\ (\mu - x_1)\ dt + \eta_1 \sigma\ dz + \eta_2 y\ dN\ .$$

where $\mu = (1-w)\bar{\mu} + w\,x_2,\ \ 0 \le w \le 1.$ The first factor x_1 reverts to the level of a weighted average of a long run rate $\bar{\mu}$ and the current level of the second factor x_2, where $a = -v,\ \delta = v\,w,\ b = v\,(1-w)\ \bar{\mu}.$

v	b	a	δ	$\bar{\mu}$	w	σ^2	ν	s
0.508	0.142	-0.508	0.397	1.277(%)	0.781	0.436	-0.483(%)	0.002

Table 9. Estimation of Parameters for Equation (38)

Parameter	β_0*	β_1*	β_2*	s_1*
Estimate	4.480	-0.826	-0.084	0.245
Standard Error	0.023	0.005	0.002	0.002

(1) Degree of Freedom 188.

(2) * indicates significance at the 1% level.

Figure 1. Interest Rates

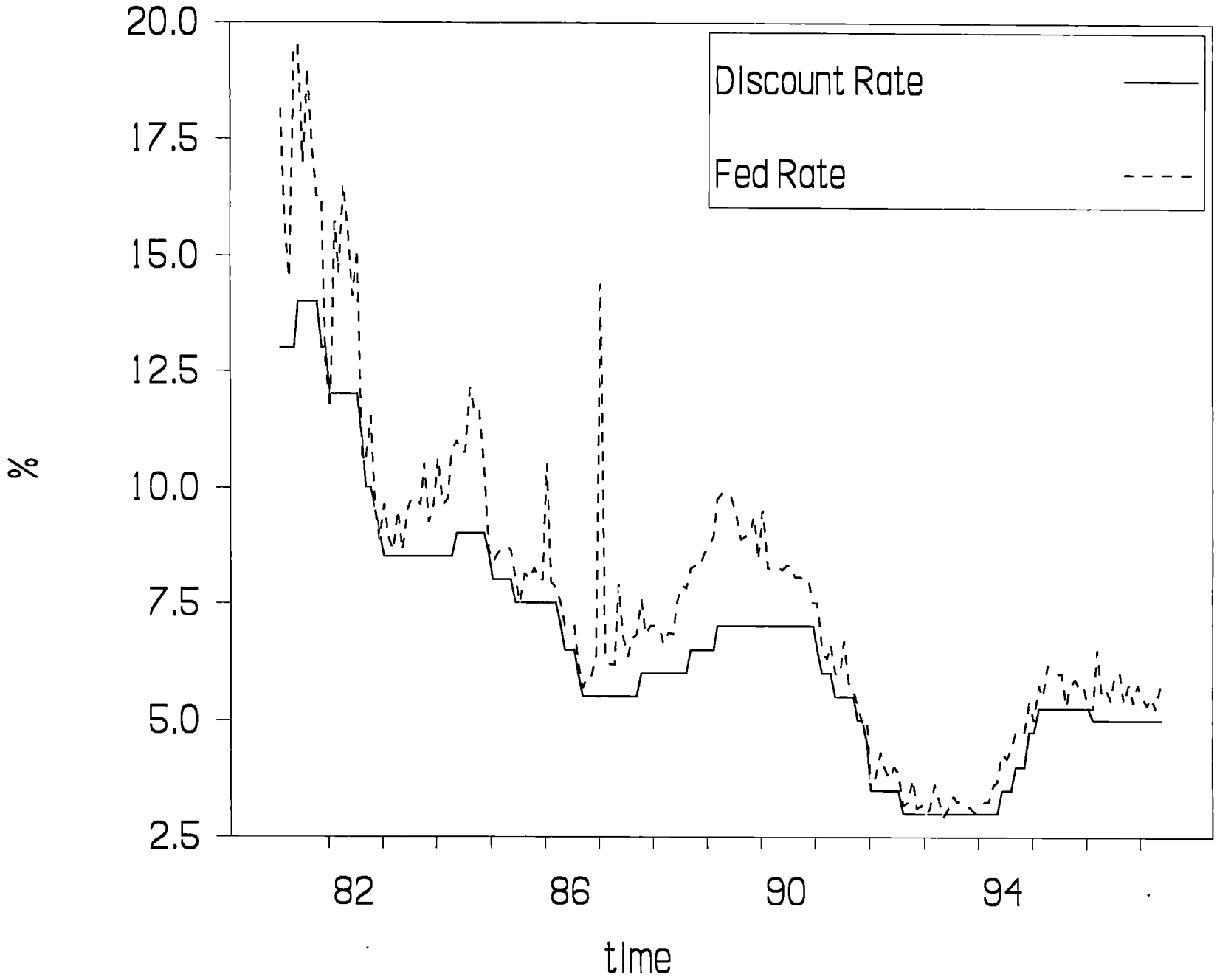
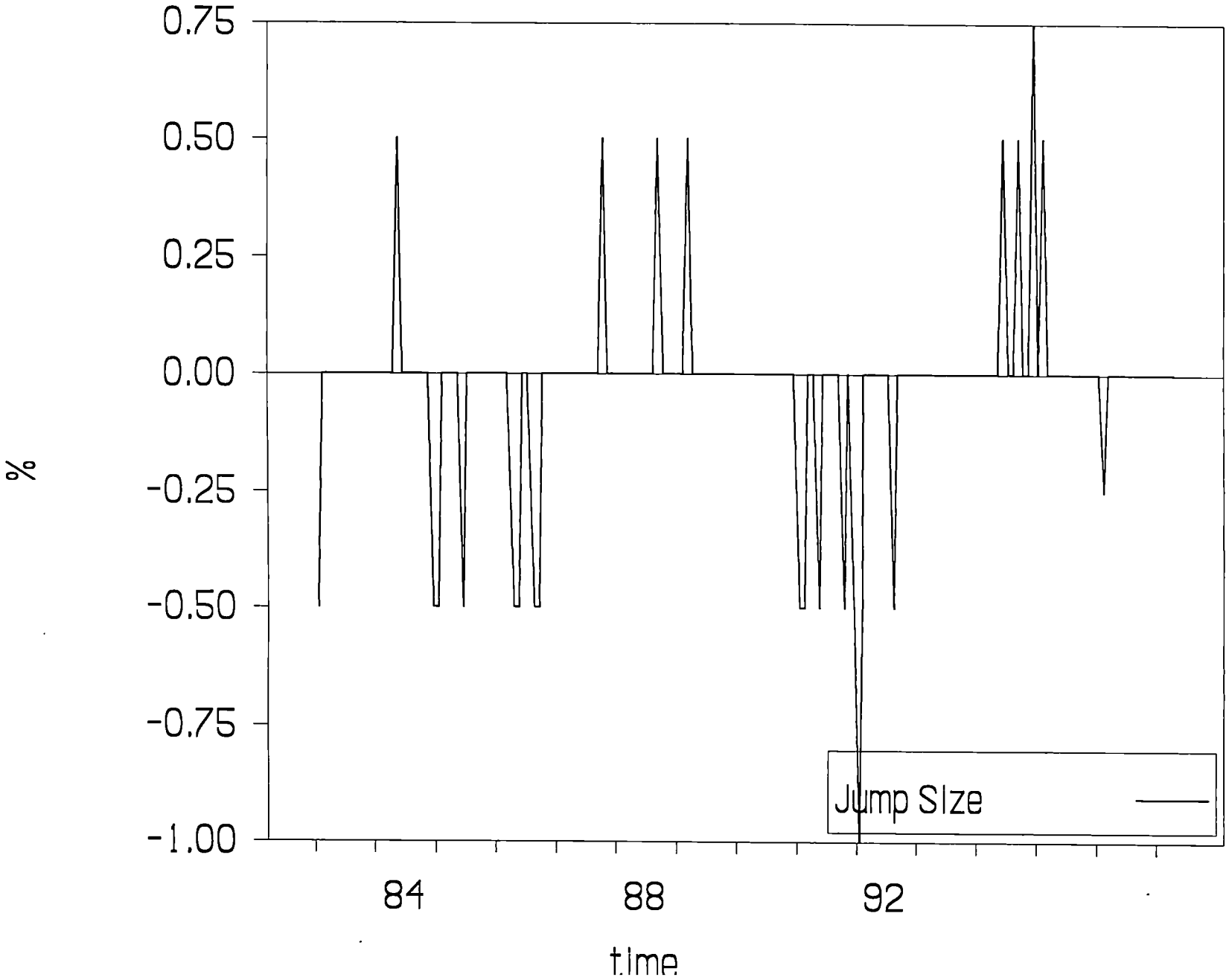


Figure 2. Size of Jump of Discount Rate



5.6. Conclusion

This chapter extends Duffie and Kan's (1996) affine model of the term structure to the presence of jumps. It is shown that the affine model of the term structure with jumps can be obtained if the jumps intensity is affine. We apply this framework to the model of the term structure with the official set interest rates. We present an approximate solution of a two-factor model of the term structure of interest rates. Of the various models of the term structure of interest rates, only Das and Foresi (1996) have so far obtained a closed-form solution for the value of a pure discount bond price. For the model assuming exponentially distributed jump size, the approximate formula is quite accurate for short maturities. The empirical results from the LR test demonstrate that our jump diffusion model fits the U.S. market data better than a pure diffusion model.

Chapter 6. A Model of Term Structure of Interest Rates under the Expectation of Regime Changes

6.1. Introduction

Many models of the term structure of interest rates start with an *Itô* type stochastic process for either the short rates, bond prices, or for forward rates. All these models assume that the term structure of interest rates moves according to its own dynamics, depending on the underlying factors. However, spot rates and forward rates are determined by complex economic processes. Cox, Ingersoll and Ross (CIR (1985b)), Longstaff and Schwartz (1992), and, with a slightly different approach, Tice and Webber (1997) are examples of the incorporation of such factors in models. CIR (1985b) provide a general equilibrium framework for the pricing of discount bonds (the term structure) and of other contingent claims in a continuous time framework. In their framework, the spot rate and its dynamics are endogenously determined. Longstaff and Schwartz (1992) derive a two-factor model of the term structure from the CIR framework. Tice and Webber (1997) model the term structure of interest rates from a standard IS-LM macroeconomics framework.

Additionally, some forward-looking economic variables reflect the expectations of market agents. For instance, in the presence of government intervention in interest rates or exchange rates, the underlying variables are usually determined not only by the fundamental dynamics, but also by the expectations of market participants concerning government interventions or regime changes in the forward-looking variables, such as exchange rates and interest rates. These phenomena (expectations) sometimes generate a speculative bubble[45]. For instance, when we

45. See the definition in *Market Volatility* by Shiller (1989), or in *The Econometrics of*

investigate the stochastic version of a Cagan-type monetary model, which will be explained in the next section, the solution of the stochastic differential equation can be expressed as a bubble function[46]. In the case of exchange rates, many papers, such as that by Krugman and Miller (1992), have tried to explain such a feature of exchange rate movements. Changes in regimes may be also very important for the evolution of interest rates. If the regime can shift, market participants must assign probabilities to such changes. For instance, shifts in the Federal Reserve's target band for the Federal funds rate are associated with dramatic moves in the term structure of interest rates, (Hamilton, 1988, 1990).

We can model discrete shifts of interest rates with a jump process such as a Poisson process (Das and Foresi, 1996). With finite numbers of jumps in any open interval, jump diffusion models of interest rates assume that interest rates follow a diffusion process between jumps. If the market anticipates possible regime changes in interest rates, this can be expected to affect the drift and diffusion of the process. One of the motivations of this chapter is to examine such expectations.

However, it is important to note here that there are many theoretical arguments concerning the existence of bubbles in financial assets and forward-looking variables[47]. Campbell, Lo and MacKinlay (CLM,1997) argue that a bubble cannot exist if there is an upper limit on the forward-looking variables. This clearly contradicts the theoretical framework of Krugman and Miller. CLM also argue that bubbles cannot exist on assets such as bonds which have a fixed value on a terminal date. This theoretical argument requires empirical evidence. We do not intend to test the bubble model here. In this chapter, we shall show how bubbles may be generated in an interest rate model. Importantly, we show that bubbles can create a non-linear volatility structure of interest rates even if the fundamental variable follows a simple Gaussian process.

Financial Market by Campbell, Lo and MacKinlay (1997), p. 258.

46. See *Foundations of International Macroeconomics* by Obstfeld and Rogoff (1996), p. 570.

47. See *The Econometrics of Financial Markets* by Campbell, Lo and MacKinlay (1997), p. 259.

This chapter investigates a situation where interest rates are determined by market expectations of changes in interest rates. We re-express a classical model of interest rates (Sargent, 1972) in the form of a BGM model (Brace Gatarek, and Musiela (1997)). CIR (1981) argue that the model of the term structure of interest rates under uncertainty is incompatible with rational expectations or any other expectation theory of interest rates. However, CIR consider the question only under the objective and the risk-neural probability measures. We do not deal with this conflict between the traditional hypotheses and the modern theory of the term structure of interest rates (CIR, 1981). In this chapter we avoid such problems by using changes of probability measure. However, we follow Campbell's (1986) unified view of the expectation theory of the term structure in the context of modern finance. The crucial implication of rational expectations on the term structure of interest rates is that forward rate processes follow martingales (Sargent, 1979; Samuelson, 1965).

The structure of this chapter is as follows. In section 2, we review several models which we use here, such as Sargent's model of the term structure of interest rates under rational expectations, and we present a model of the term structure of interest rates under the expectations of changes of regimes. Section 3 presents an explicit solution for a forward rate. In Section 4 we discuss boundary conditions. Section 5 offers some conclusions.

6.2. Review of Model

Before formulating a model, we shall explain several concepts used in this chapter. First, to understand the concepts of fundamentals and bubbles, we shall explain a Cagan-type monetary model in the continuous time setting. Second, we shall explain Sargent's model of interest rates. Third, the BGM (Brace, Gatarek, and Musiela) model of the term structure of interest rates shall be explained.

6.2.1 Cagan-Type Model

Here, we briefly study some implications of Cagan's money demand equation. Assume that the supply of money m_t is set exogenously, and the supply equals the demand for money in equilibrium. In continuous time, the Cagan money demand function becomes

(C1) $$m_t - p_t = -\eta \dot{p}_t$$

where p_t is the price level and $\frac{d(\log p)}{dt} = \frac{\dot{p}}{p}$ is the anticipated inflation rate in continuous time and η is the semielasticity of demand for the real balance with respect to expected inflation. Using the method of Obstfeld and Rogoff (1996), one finds that the general solution of (C1) is[48]

(C2) $$p_t = \frac{1}{\eta}\int_t^{\infty} \exp\left(-\frac{(s-t)}{\eta}\right) m_s ds + b_0 \exp\left(\frac{t}{\eta}\right).$$

The solution (C2) is sometimes called a bubble solution. Depending on the limit of the second term of the right hand side of (C2), the price level may or may not have

48. See *Foundations of International Macroeconomics* by Obstfeld and Rogoff (1996), p. 523.

a stable point. If the value of the second term grows exponentially, we have what many authors such as Krugman and Miller (1992), Froot and Obstfeld (1991) and Obstfeld and Rogaff (1996) call a speculative bubble. Next, if we assume that

$$\lim_{t\to\infty} b_0 \exp(\frac{t}{\eta})=0$$

by, for instance, taking $b_0 = 0$ and applying integration by parts, then we obtain

(C3) $$p_t=\frac{1}{\eta}\int_t^\infty \exp(-\frac{(s-t)}{\eta})m_s ds=m_t+\int_t^\infty \exp(-\frac{(s-t)}{\eta})\dot{m}_s ds.$$

If we assume $\dot{m}_t=0$, then

(C4) $$p_t=m_t.$$

The price level is just the supply of money. m_t is so called the fundamental of p_t, and the solution of (C4) is called a saddlepath solution. These are the definitions of the fundamental and of the speculative bubble in this chapter. In the case of stochastic version of (C1), the equation (C1) becomes

(C5) $$m_t-p_t=-\eta E(\dot{p}_t|\xi_t).$$

ξ_t is a sequence of σ-algebra on a probability space or the time t information set. In addition, we assume that the process m_t follows

(C5-1) $$dm_t = \sigma\, dz$$

where σ is constant, and dz is a standard Wiener process.

The saddlepath solution which rules out speculative bubbles is known to be

$$\text{(C6)} \qquad p_t = \frac{1}{\eta}\int_t^{\infty} \exp\left(-\frac{(s-t)}{\eta}\right) E(m_s)ds = m_t + \int_t^{\infty} \exp\left(-\frac{(s-t)}{\eta}\right) E(\bar{m}_s)ds.$$

$$= m_t$$

Equation (C6) is just taking the expectation of equation (C3). The general solution is slightly different from the procedure in (C6). Before we explain other concepts, we formally define the saddlepath solution of this chapter.

Definition 1: A saddlepath solution is a solution that excludes a speculative bubble. The solution (C6) is an example of a saddlepath solution.

To facilitate obtaining a general solution of (C5) with (C5-1), following Krugman and Miller (1992), we shall use *Itô's* lemma. Krugman and Miller (1992) apply the method to exchange rate targets and the currency band model. We briefly explain their model. The model is based on the flexible price exchange model and uncovered interest rate parity. With this model, they express the exchange rate as the sum of two parts : the fundamental and the expectation of a change in the exchange rate. The key problem in the model is how to solve a stochastic differential equation similar to (C5) with (C5-1) where the uncovered interest rate parity is expressed in the form of an expectation of a change in the exchange rate. Krugman and Miller explicitly assume that the general solution depends only on the fundamental m_t. Let the general solution be $G(m_t)$. When we use this technique to equation (C5) with (C5-1), the general solution is

$$\text{(C7)} \qquad G(m_t) = m_t + A\exp(\lambda m_t) + B\exp(-\lambda m_t)$$

where A and B are arbitrary constants and

$$\lambda = \sqrt{\frac{2}{\eta\sigma^2}}$$

We call the solution (C7) the general solution. Here we call $A\exp(\lambda m_t)+B\exp(-\lambda m_t)$ a bubble term[49].

Krugman and Miller explain the expectations of future regime changes in the exchange rate, using the solution (C7). Most papers on this topic, such as Krugman and Miller (1992), investigate two types of regime shifts in exchange rates. The first one is based on a system change, for instance, from a floating rate system to a fixed rate system. The second one is a change of the central bank's target rate. We shall deal with our interest model, using this framework.

6.2.2 Rational Expectation Model of Interest Rates

The forward rate is known to follow a martingale process under the rational expectation model. From the view point of empirical work, a martingale process is a covariance stationary process. In other words, setting up an econometric model, the error or residual should follow white noise. However, as Shiller (1989) pointed out, the rational expectations model cannot explain the high volatility of long-term interest rates. As in Sargent (1972, 1979), the forward rate process can be represented as a martingale process plus a liquidity premium. The liquidity premium is usually described as a reflection of public perceptions and attitudes and is assumed constant (Shiller (1989), Sargent (1972, 1979)). Shiller (1989) argues that the liquidity premium moves very slowly, and cannot explain the volatility of, in particular, long-term interest rates. Shiller conjectures that the volatility of interest

49. Strictly speaking, for $\lambda > 0$, when $B = 0, A \neq 0$, we call it bubble solution. If $A \neq 0, b \neq 0$, we call the solution (C7) a target zone solution. However, here we call both solutions a bubble solution, following Krugman and Miller (1992).

rates can be explained in terms of information about possible regime changes in interest rates. We set up the regime change model in a continuous time framework from the Sargent-type interest rates model.

In a rational-expectations equilibrium in the discrete-time setting, the forward rate processes are known to follow a martingale (Sargent, 1973, 1979)[50]. This may be true even in modern finance when we adopt the approximate linearized framework of Campbell (1986), who shows that an arbitrage argument of CIR (1981) can be compatible with a rational expectations equilibrium in an approximate linearized framework. Although the original papers of Sargent (1979) and Shiller (1981) deal with martingales under an ordinary probability measure P, modern financial theory demands that they be martingales under another measure which we call the Q -rational expectations measure. Later we shall define Q-rational expectations measure properly.

In his empirical papers, Sargent (1972, 1979) found that the forward rate does not follow a martingale process under rational expectations. To explain this, Sargent assumed a liquidity premium. Sargent defines the liquidity premium in the discrete time setting at each time t simply as

$$E(f(t, T) - f(t-1, T))$$

where $f(t, T)$ is the forward rate at time t for fixed date T. This is exactly the same definition of the liquidity premium in Ingersoll (1987).

6.2.3 Continuous Time Version of the Sargent Interest Rate Model

In the discrete time framework, Sargent (1979) modelled the forward rate process as having two components, a martingale part and a liquidity premium part. If we

50. The proof is also available in Shiller (1989) and Begg (1982).

mathematically formulate a model of Sargent-type interest rates in the continuous time framework, then

(1) $$f = k + \Lambda$$

where f is the forward rate process, k is a martingale process, and Λ is liquidity premium. If we express Λ of Sargent (1979) in a continuous time setting, it becomes the expected rate change of the forward rate, $\frac{E(df)}{dt}$. We assume that Sargent modelled (1) under the ordinary measure P. However, the rational expectation can not be the expectation of the measure P, since under P, the forward rate process is known not to follow a martingale process. In the continuous time version of the Sargent model (1), the intertemporal evolution of the forward rate at time t for date T, denoted $f(t, T)$, is assumed to be described by fixing a maturity date, as in the HJM framework.

An alternative perspective of the forward rate dynamics which we shall adopt is to consider the evolution of $f(t, t+\tau)$, denoted by $\tilde{f}(t, \tau)$ for a fixed τ. That is we keep the time to maturity τ constant rather than fixing the maturity date. This alternative approach is used by Brace and Musiela (1994), Jeffrey (1994), and Brace, Gatarek, and Musiela (1997). We call this the BGM framework. The BGM perspective can be always transformed into an equivalent HJM specification. We shall show that f cannot follow a martingale process under the ordinary measure P in the framework of BGM. From Sargent (1979) and Jamshidian (1991), we posit the following stylized fact,

Fact 1: *Under the measure P, f is not a martingale process.*

Note 1:

To clarify the difference between HJM and BGM, we discretize the continuous time setting model. HJM analyze the forward rate process in the following form:

(i) $$\Delta f_{HJM} = f(t, T) - f(t-1, T)$$

where T is fixed maturity date. On the other hand, BGM study the forward rate process in a form:

(ii) $$\Delta f_{BGM} = f(t, T) - f(t-1, T-1)$$

where the time to maturity τ ($T-t = (T-1)-(t-1)$) is fixed. In the Sargent model, the expectation of the change of the forward rates, equation (1) is included in the category (i). However, we shall model interest rates under (ii). Hence, we do not apply the liquidity premium of Sargent to our model.

Note 2:

(i) In Sargent model in a discrete time setting, Λ is $E(\Delta f_{HJM})$, which is the original definition of a liquidity premium

(ii) In our model, Λ will be $E(\Delta f_{BGM})$.

6.2.4. A Sargent-Type Model of Interest Rates under the Probability Measure Q

Following BGM, we write f for the forward rate process as

$$\tilde{f}(t, \tau) = k(t) + \frac{E^{Q}(df \mid \xi_t)}{dt} \tag{2}$$

where k is a martingale process under Q and $\xi = \{\xi_t : t \in [\ 0, \infty)\}$ is a filtration, and ξ_t is a sequence of σ-algebra on a probability space. To develop a Sargent type under the BGM setting, we define a new measure Q.

Definition 2: The probability measure Q is defined as a probability measure under which k in (2) is a martingale.

Definition 3: We say that the measure Q is a Q-rational expectation measure if

(i) k is a martingale under Q

(ii) In the saddlepath solution to (2), f is a martingale process as well.

The condition (ii) implies that $\frac{E^{Q}(df \mid \xi_t)}{dt} = 0$ under the Q-rational expectations measure and $f = k$. We posit the following stylized fact.

Fact 2: *Under the measure Q-rational expectation, f is a martingale process. By a change of measure, we can always make f a non-martingale process under P.*

In Note 2, the expectation term (2) in our model is different from that of Sargent, since we set up the model under the framework of BGM. Following the

conjecture of Shiller (1979), to explain the non-martingale property of the forward rate process in Sargent, we apply Krugman and Miller (1992) framework of regime change to our model. The reason that we use the Krugman and Miller framework is as follows : First, this framework is well established for evaluating stochastic differential equations like equation (2). Second, we can explain the non-martingale property of the forward rate process in a wide range of circumstance such as a speculative bubble phenomenon and regime changes in interest rates. As in Section 6.1, changes in regimes may be very important for the evolution of interest rates. If the regime can shift, market participants must assign probabilities to such changes. We assume that the expectation term in (2) represents such expectations.

Since k is a martingale process under Q, we can express the process k in a continuous time framework using the martingale representation theorem. From the martingale representation theorem, the process $k(t)$ is

(3) $$k(t) = k(0) + \int_0^t \sigma \, d\hat{z},$$

where $\hat{z}$ is a Brownian motion with respect to the probability measure Q. For simplicity, we shall assume that the volatility structure of k is time-homogeneous Gaussian. Setting up the Sargent (1979) type interest rate model in the continuous time framework, we get the following equation system,

(4) $$\hat{f}(t, \tau) = k(t) + \frac{E^Q(df \mid \xi_t)}{dt}$$

$$dk = \sigma d\hat{z}.$$

We solve equation system (4). In other words, we express (4) in terms of (C6) and (C7). First, we try a saddlepath solution. The saddlepath solution is by definition 1 a solution without a bubble. Second, we shall try a general solution or bubble solution in next section.

We shall show in the following proposition that in our setting, for the Sargent type model, the forward rate process is not a martingale under the ordinary probability measure P.

Proposition 1: For the following continuous time version of Sargent's equation system for fixed $\tau = T - t$

$$\hat{f}(t, \tau) = k(t) + \frac{E^P(df \mid \xi_t)}{dt}$$

$$dk = \sigma d\hat{z},$$

f is not a martingale process under the ordinary probability measure P.

Proof:

Assuming that $T-t$ is fixed, we try a saddlepath solution under P. Here, k is a P-martingale. We can obtain a saddlepath solution analogous to (C3)

(5) $$\hat{f}(t, \tau) = \int_t^{\infty} e^{(t-s)} E^P(k(s) \mid \xi_t) ds$$

$$= k(t) \int_t^{\infty} e^{(t-s)} ds$$

$$= k(t)$$

The solution (5) is explained in Froot and Obstfeld (1991) and Obstfeld and Rogaff

(1996). The equation (5) implies that under the ordinary measure P, the forward process f follows a martingale. However, by Jamshidian (1991) and Fact 1, the forward rate does not follow a martingale process under P. This contradicts that f is a martingale under P. The expectation under the measure P is not the rational expectation. Even if we try a bubble solution, which shall be demonstrated in the next section, the equation system (4) does not hold under P. This is simply because the saddlepath solution is a special case of the bubble solution.

■

As seen in proposition 1, under P, f cannot be a martingale process. We try a saddlepath solution under Q. Assuming that $T-t$ is fixed again, we can obtain a saddlepath solution under the measure Q.

(6)
$$\tilde{f}(t,\tau) = \int_t^{\infty} e^{(t-s)} E^Q(k(s) \mid \xi_t) ds$$

$$= k(t) \int_t^{\infty} e^{(t-s)} ds$$

$$= k(t).$$

The measure Q in the saddlepath solution is a Q-rational expectation measure, since the forward rate process f is a martingale process. By a change of measure, we can make f a non-martingale process under the objective measure.

Analogous to the Cagan-type model, the martingale part $k(t)$ in our model is the fundamental part of $\tilde{f}(t,\tau)$. Here, we interpret k as the forward rate process excluding the speculative bubble. Since the fundamental k, which is interpreted as the forward rate excluding the speculative bubble, is always a martingale under Q by definition 2, and the forward rate is a martingale only under the forward risk-adjusted measure, then Q is the forward risk-adjusted measure for the forward

process excluding the speculative bubble. We reinterpreted (3) as

$$(7) \qquad \hat{k}(t,\tau) \equiv k(t,t+\tau) = k_0(t,\tau) + \int_0^t \sigma \, d\hat{z},$$

where $k_0(t,\tau)$ is assumed to be $k(0,t+\tau) - \int_0^t \frac{\partial}{\partial \tau} \hat{k}(s,\tau)ds$. The reason we assume this is to make the drift of $d\hat{k}(t,\tau)$ zero under Q[51].

Note 3:

We briefly explain the differentiation related to equation (7). This Note 3 is from the Appendix 1 from Jeffrey (1994).

In HJM, they express the forward rate process as :

$$\text{(i)} \qquad k(t,T) = k(0,T) + \int_0^t \alpha(s,T)ds + \int_0^t \sigma(s,T)dz.$$

On the other hand, in BGM (1997) and Jeffrey (1994), the forward rate process is

$$\text{(ii)} \qquad k(t,t+\tau) = k(0,t+\tau) + \int_0^t \alpha(s,t+\tau)ds + \int_0^t \sigma(s,t+\tau)dz$$

Differentiating (ii) with respect to t yields the following stochastic differential equation representation for the dynamic behavior :

51. Any stochastic process X_t is a martingale if and only if the drift of dX_t is zero. See p. 79 in Baxter and Rennie (1997). Differentiating $\hat{k}(t,\tau)$ with respect to t yields $d\hat{k}(t,\tau) = \sigma d\hat{z}$. See Appendix I in Jeffrey (1994) for this kind of differentiating or Note 3.

$$dk(t, t+\tau) = \left(\frac{\partial k(0, t+\tau)}{\partial t} + \int_0^t \frac{\partial \alpha(s, t+\tau)}{\partial t} ds + \int_0^t \frac{\partial \sigma(s, t+\tau)}{\partial t} dz\right) dt$$

$$+\alpha(t, t+\tau)dt + \sigma(t, t+\tau)dz$$

which simplifies to :

(iii) $$dk(t, t+\tau) = \left(\alpha(t, t+\tau) + \frac{\partial k(t, T)}{\partial T}\Big|_{T=t+\tau}\right)dt + \sigma(t, t+\tau)dz$$

Conversely, integrating (iii) gives (ii). We can apply this to equation (7).

BGM use slight different notations. They express (iii) as

(iv) $$d\tilde{k}(t, \tau) = \tilde{\alpha}(t, \tau)dt + \tilde{\sigma}(t, \tau)dz$$

$$= \left(\alpha(t, t+\tau) + \frac{\partial \tilde{k}(t, \tau)}{\partial \tau}\right)dt + \tilde{\sigma}(t, \tau)dz$$

Since $k(t, t+\tau) = \tilde{k}(t, \tau)$, from (iv), it is clear that

$$\tilde{\alpha}(t, \tau) = \alpha(t, t+\tau) + \frac{\partial \tilde{k}(t, \tau)}{\partial \tau} \quad \text{and} \quad \sigma(t, t+\tau) = \tilde{\sigma}(t, \tau)$$

We shall follow the notations of BGM.

Even if we substitute $\tilde{k}(s, \tau)$ rather than $k(s)$ into (6), we obtain the same solution for fixed $T-t$. By expressing k in (7) as a forward rate process, we can simplify the formula for the forward rate process and a pure discount bond price under the speculative bubble. In the next section, we shall find a bubble solution of our model analogous to (C7). By excluding the time dependency on the bubble

solution, the forward rate f in the presence of a bubble can be expressed by the fundamental k, which is the forward rate excluding a bubble, plus a bubble. In our model, for simplifying the analysis, we specify the evolution of k and the bubble path separatedly instead of modelling directly the evolution of f.

Note 4:

The fundamental k in our interest rate model is the forward rate excluding a speculative bubble.

Before we turn to Section 3, we explain the assumptions of the fundamental in our model. the process k is a forward rate process excluding a speculative bubble, or the fundamental part of f. To explain the volatility structure and changes of probability measure, we review a relationship between the bond price, the spot rate and the forward rate which are all assumed to be fundamentals, i.e. processes without a speculative bubble. First, we explain the general framework of BGM.

6.2.5 Structural Framework of BGM

We develop a model of the term structure of interest rates in the framework of BGM. In this section we show several concepts of their framework, which is necessary for obtaining an explicit formula for the forward rate the presence of speculative bubble in section 3. First, we compare the BGM setting with that of HJM. Before comparing the models, we define notations used here.

Definition 4

$\tilde{k}(t, \tau)$: the forward rate at time t with time to fixed maturity τ excluding a

speculative bubble (BGM notation).

$k(t, T)$: the forward rate at time t with fixed maturity T excluding a speculative bubble (HJM notation).

r_f : the spot rate excluding a speculative bubble.

$B(t, \tau)$: the discount bond price at time t with time to maturity τ excluding a speculative bubble.

$\hat{f}(t, \tau)$: the forward rate in the presence of a speculative bubble (BGM notation).

$f(t, T)$: the forward rate in the presence of a speculative bubble (HJM notation).

r : the spot rate in the presence of a speculative bubble.

$P(t, \tau)$: the discount bond price in the speculative bubble.

In the framework of HJM, for fixed $T = t + \tau$, the evolution of $k(t, T)$ is modelled with the following stochastic equation :

(8) $$dk(t, T) = \mu_T(t)dt + \sigma_T(t)dz,$$

where the relation of $u_T(t) = \sigma_T(t)\left(\int_t^T \sigma_s(t)ds - \lambda(t)\right)$ results from the no-arbitrage condition in the bond market where λ represents the market price of risk. As in HJM (1992) and Babbs (1991), this condition can be generated by determining the stochastic evolution of bond prices.

On the other hand, the evolution of the term structure can be represented by the

BGM framework. From BGM (1997), the dynamics $\tilde{k}(t, \tau)$ can be described via the following stochastic differential equation;

(9) $$d\tilde{k}(t, \tau) = \mu_\tau(t)dt + \sigma_\tau(t)dz,$$

where

$$u_\tau(t) = u_T(t) + \frac{\partial \tilde{k}(t, \tau)}{\partial \tau}$$

$$\sigma_\tau(t) = \sigma_T(t).$$

For all $T > 0$, from BGM (1997), the evolution of a zero coupon bond process with maturity $B(t, \tau)$ is described via the following stochastic equation :

(10) $$\frac{dB(t, \tau)}{B(t, \tau)} = \mu^\tau(t)dt - \sigma^\tau(t)dz$$

where

(10-1) $$\mu^\tau(t) = \tilde{k}(t, 0) - \tilde{k}(t, \tau) + \lambda(t)\sigma^\tau(t).$$

From Jamshidian (1991) and BGM (1997), the following relations hold :

(11) $$\sigma_\tau(t) = \frac{\partial \sigma^\tau(t)}{\partial \tau}, \quad \sigma^\tau(t) = \int_0^\tau \sigma_u(t)\, du,$$

and by definition,

$$\sigma(t) = \sigma_0(t), \quad \sigma^0(t) = 0.$$

We defined r_f to be the spot rate process without a speculative bubble. We assume that when there is a speculative bubble, the spot rate contains a bubble as

well[52].

Following BGM (1997), r_f is described via the following equation:

(12) $$dr_f = \frac{\partial}{\partial \tau} \tilde{k}(t, \tau)|_{\tau=0}\, dt + \frac{\partial}{\partial \tau} \sigma^{\tau}(t)|_{\tau=0}\, dz.$$

Assuming that λ is the market price of risk, then similarly to Jamshidian (1991), two risk-adjusted probability measures are defined:

(13) (i) $$d\bar{z} = dz - \lambda(t)dt$$

(ii) $$d\hat{z} = dz + [\ \sigma^{\tau}(t) - \lambda(t) + \frac{1}{\sigma_{\tau}(t)} \frac{\partial}{\partial \tau} \tilde{k}(t, \tau)]\ dt, \qquad T > 0$$

and so

(iii) $$d\hat{z} - [\ \sigma^{\tau}(t) + \frac{1}{\sigma_{\tau}(t)} \frac{\partial}{\partial \tau} \tilde{k}(t, \tau)]\ dt = d\bar{z}$$

where z is a Brownian motion with respect to the original probability measure P, $\bar{z}$ is a Brownian motion with respect to the measure $\bar{Q}$ which we shall call the risk neutral measure, and $\hat{z}$ is a Brownian motion with respect the measure Q which is the forward risk-adjusted measure. We shall prove (ii) in Lemma 3. To develop our model, we introduce several concepts from BGM. In the presence of bubbles, the market price of risk should be changed. We shall show this in the next section.

Lemma 2 : For all $T > 0$, the following relation from Lemma 1 in Jamshidian (1991) still holds in the framework of BGM

52. Here, r_f is different from r, which is the spot rate in the presence of a speculative bubble. Short rates shall be derived from $\tilde{f}(t, \tau)$, setting $\tau = 0$, in the next section,

(i) $$\mu_\tau = -\frac{\partial \mu^\tau}{\partial \tau} + \sigma_\tau(t)\sigma^\tau(t).$$

(ii) $$\mu^\tau(t) = r_f(t) - \int_0^\tau \mu_u(t)du + \frac{1}{2}\sigma^\tau(t)^2.$$

Proof:

proof of (i): By definition,

$$\mu_\tau(t)\,dt = -\frac{\partial}{\partial \tau}(E(dlogB(t,\tau))).$$

Then, from lemma 1 in Jamshidian, we obtain

$$\mu_\tau(t)\,dt = -\frac{\partial}{\partial \tau}(\mu^\tau(t) - \frac{1}{2}\sigma^\tau(t)^2).$$

proof of (ii) : Integrating (i) with respect to τ from 0 to τ, we get

$$\int_0^\tau \mu_u(t)\,du = -\mu^\tau(t) + \mu^0(t) + \frac{1}{2}\sigma^\tau(t)^2.$$

From Jamshidian (1991), $\mu^0(t) = r_f(t)$. If we substitute this into the above equation, we obtain (ii)

■

We can easily check (i) and (ii) from (9), (10) and (11) of the framework of BGM.

Lemma 3 : If z is a Brownian motion under the measure P, then $\hat{z}$ is a Brownian motion under Q, where

$$d\hat{z} = dz + [\ \sigma^{\tau}(t) - \lambda(t) + \frac{1}{\sigma_{\tau}(t)} \frac{\partial}{\partial \tau} \tilde{k}(t, \tau)]\ dt, \qquad T > 0$$

Proof:

We show that the drift of $dk(t, \tau)$ is zero under Q. From (ii) in Lemma 2, we know that $\mu_{\tau} = -\frac{\partial \mu^{\tau}}{\partial \tau} + \sigma_{\tau}(t)\sigma^{\tau}(t)$. Applying (10-1) which is

$$\mu^{\tau}(t) = \tilde{k}(t, 0) - \tilde{k}(t, \tau) + \lambda(t)\sigma^{\tau}(t),$$

to here, it follows that

$$\mu_{\tau} = \frac{\partial}{\partial \tau} \tilde{k}(t, \tau) - \sigma_{\tau}(t)\lambda(t) + \sigma_{\tau}(t)\sigma^{\tau}(t)$$

$$= \frac{\partial}{\partial \tau} \tilde{k}(t, \tau) + \sigma_{\tau}(t)(\sigma^{\tau}(t) - \lambda(t)).$$

Then, we obtain

$$d\tilde{k}(t, \tau) = (\frac{\partial}{\partial \tau} \tilde{k}(t, \tau) + \sigma_{\tau}(t)(\sigma^{\tau}(t) - \lambda(t)))dt + \sigma_{\tau}(t)\, dz.$$

Substituting condition (ii) into above equation, we get

$$d\tilde{k}(t, \tau) = (\frac{\partial}{\partial \tau} \tilde{k}(t, \tau) + \sigma_{\tau}(t)(\sigma^{\tau}(t) - \lambda(t)))dt + \sigma_{\tau}(t)\, dz.$$

$$= (\frac{\partial}{\partial \tau} k(t, \tau) + \sigma_\tau(t)(\sigma^\tau(t) - \lambda(t)))dt.$$

$$+ \sigma_\tau(t)(d\hat{z} - [\ \sigma^\tau(t) - \lambda(t) + \frac{1}{\sigma_\tau(t)} \frac{\partial}{\partial \tau} \hat{k}(t, \tau)]\ dt)$$

$$= \sigma_\tau(t)\, d\hat{z}.$$

■

Proposition 4: For all $T > 0$, the condition on the drift of $d\hat{k}(t, \tau)$ which is,

$u_\tau(t) = u_T(t) + \frac{\partial \hat{k}(t, \tau)}{\partial \tau}$, is equivalent to

$$\hat{k}(t, \tau) = k(0, t+\tau) + \int_0^t \sigma_{\tau+t-s}(s)\, \sigma^{\tau+t-s}(s)\, ds + \int_0^t \sigma_{\tau+t-s}(s) d\bar{z}.$$

Proof:

First, we can write $\mu_T(t)$ as

$$u_T(t) = \sigma_T(t)(\int_t^T \sigma_T(s)ds - \lambda(t)) = \sigma_\tau(t)(\int_0^\tau \sigma_\tau(s)ds - \lambda(t))$$

$$= \sigma_\tau(t)(\sigma^\tau(s) - \lambda(t)).$$

Then, the equation (9) becomes :

$$(14) \qquad d\hat{k}(t,\tau) = \left[\ \sigma_\tau(t)(\sigma^\tau(s) - \lambda(t)) + \frac{\partial \hat{k}(t,\tau)}{\partial \tau}\ \right] dt + \sigma_\tau(t)dz$$

$$= \left[\ \sigma_\tau(t)\sigma^\tau(s) + \frac{\partial \hat{k}(t,\tau)}{\partial \tau}\ \right] dt + \sigma_\tau(t)(-\lambda(t)dt + dz)$$

$$= \left[\ \sigma_\tau(t)\sigma^\tau(s) + \frac{\partial \hat{k}(t,\tau)}{\partial \tau}\ \right] dt + \sigma_\tau(t)d\bar{z}.$$

The last equality is from (13).

From Brace and Musiela (1994) and Jeffrey (1994), the solution of the third line of (14) is known to be

$$(15) \qquad \hat{k}(t,\tau) = k(0, t+\tau) + \int_0^t \sigma_{\tau+t-s}(s)\sigma^{\tau+t-s}(s)\,ds + \int_0^t \sigma_{\tau+t-s}(s)d\bar{z}.$$

■

We can obtain the spot rate process $r_f = \hat{k}(t,0)$ by setting $\tau = 0$ in (15) ;

$$(16) \qquad \hat{k}(t,0) = k(0,t) + \int_0^t \sigma_{t-s}(s)\sigma^{t-s}(s)\,ds + \int_0^t \sigma_{t-s}(s)d\bar{z}.$$

6.2.6 Our Model

We assume that forward rates excluding a speculative bubble under Q is determined by $d\hat{k}(t,\tau) = \sigma d\hat{z}$, where σ is constant. In other words, we assume that the volatility structure in equations (9) and (10) is $\sigma_\tau = \sigma$, $\sigma^\tau = \sigma\tau$. If we substitute

these assumptions into (16), we obtain :

$$(17) \qquad r_f(t) = \hat{k}(t,0) = k(0,t) + \int_0^t \sigma^2(t-s)\,ds + \int_0^t \sigma d\bar{z}.$$

Equation (17) is crucial to simplifying our pricing formula for a pure discount bond price in the presence of a bubble. In the next section, we try a bubble solution to explain the non-martingale property of the forward rate process. The bubble solution is known to be expressed as a saddlepath solution plus a bubble path. In a bubble solution, following Krugman and Miller (1992), we simplify the solution by assuming that the solution depends only on the fundamental k but not on time. The forward rate process f is the fundamental k plus bubble, where the fundamental k is a forward rate process without bubbles.

6.3. An Explicit Form for Forward Rates

In the case where market agents expect the monetary authorities to alter the fundamentals in the future, the forward rate may not satisfy (6), since a bubble is excluded in the solution (6). In such cases, to evaluate the conditional expectation (4) in an easier way, we use *Itô*'s lemma[53], rather than a direct computation of the conditional expectation, following the method of Krugman and Miller. First, we find the family of solutions which satisfies some boundary conditions to the regime switch under consideration. Following the method of Krugman and Miller (1992), we develop explicitly the functional form of the forward rate f from equation (4). Write $g(k)$ for the family of functions that satisfy condition (4) when k evolves as a martingale process under Q. Let $f = g(k)$, then $g(k)$ is a solution of (4), and $E^Q\frac{(df)}{dt}$ can be obtained by *Itô*'s lemma. We shall restrict our attention to solutions that depend on current fundamentals alone. In fact, (4) could have

53. The reason we use *Itô*'s lemma is purely due to make the calculation easier.

solutions that are also functions of variables extraneous to the model. Such solutions are not considered here. Assuming time independence for a stationary solution, we express $\frac{E^Q(df|\xi_t)}{dt}$ in terms of the function $g(k)$, We use $It\hat{o}$'s lemma:

$$df = g(k)'dk + \frac{1}{2}\sigma^2 g(k)''dt. \tag{18}$$

If an expectation is taken under Q, (18) becomes

$$\begin{aligned} \frac{E^Q(df|\xi_t)}{dt} &= g'(k)\frac{E^Q(dk)}{dt} + \frac{\sigma^2 g''(k)}{2} \\ &= \frac{\sigma^2 g''(k)}{2}. \end{aligned} \tag{19}$$

Substitution of (19) into equation (4) yields :

$$g(k) = k + \frac{\sigma^2 g''(k)}{2}. \tag{20}$$

The general solution of (20) is of form :

$$g(k) = k + Ae^{\alpha k} + Be^{-\alpha k}.$$

where $\alpha = \sqrt{\frac{2}{\sigma^2}}$. We find that:

$$\frac{E^Q(df)}{dt} = \frac{\alpha^2\sigma^2}{2}[\ Ae^{\alpha k} + Be^{-\alpha k}] \tag{21}$$

where A and B are determined by boundary conditions. We shall discuss the determination of A and B in the next section.

Using equation (7), equation (4) becomes (suppressing the notational dependence on k)[54] :

(22) $$\hat{f}(t,\tau) = \hat{k}(t,\tau) + [\, Ae^{\alpha k} + Be^{-\alpha k}]$$

$$= k(0,t+\tau) - \int_0^t \frac{\partial}{\partial \tau}\hat{k}(s,\tau)ds + \int_0^t \sigma d\hat{z}$$

$$+ Ae^{-\alpha(\, k(0,t+\tau) - \int_0^t \frac{\partial}{\partial \tau}\hat{k}(s,\tau)ds + \int_0^t \sigma d\hat{z})}$$

$$+ Be^{-\alpha(\, k(0,t+\tau) - \int_0^t \frac{\partial}{\partial \tau}\hat{k}(s,\tau)ds + \int_0^t \sigma d\hat{z})}.$$

Solution (22) is a bubble solution of the expectations for future forward interest rates. Compared with the saddlepath solution (6), the solution (22) has additionally a bubble term. This solution is analogous to (C7).

Combining equations (22) and (13), we can express f under the risk neutral measure $\overline{Q}$ for k as

54. In our general solution, the function $\hat{f}$ depend on only $\hat{k}$. However, $\hat{k}$ is the function of t and τ. We just denote $\hat{f}$ as $\hat{f}(t,\tau)$ instead of $\hat{f}(\hat{k}(t,\tau))$

$$(23) \qquad \tilde{f}(t,\tau) = k(0,t+\tau) - \int_0^t \frac{\partial}{\partial \tau} \tilde{k}(s,\tau)ds$$

$$+ Ae^{\alpha(k(0,t+\tau) - \int_0^t \frac{\partial}{\partial \tau} \tilde{k}(s,\tau)ds + \int_0^t \sigma[\sigma(\tau+t-s)ds + \frac{1}{\sigma}\frac{\partial}{\partial \tau}\tilde{k}(s,\tau)ds + d\bar{z}])}$$

$$+ Be^{-\alpha(k(0,t+\tau) - \int_0^t \frac{\partial}{\partial \tau} \tilde{k}(s,\tau)ds + \int_0^t \sigma[\sigma(\tau+t-s)ds + \frac{1}{\sigma}\frac{\partial}{\partial \tau}\tilde{k}(s,\tau)ds + d\bar{z}])}$$

$$+ \int_0^t \sigma[\sigma(\tau+t-s)\ ds + \frac{1}{\sigma}\frac{\partial}{\partial \tau}\tilde{k}(s,\tau)ds + d\bar{z}\ ]$$

$$= k(0,t+\tau) + [\ Ae^{\alpha(k(0,t+\tau) + \int_0^t \sigma[\sigma(\tau+t-s)ds + d\bar{z}])}$$

$$+ Be^{-\alpha(k(0,t+\tau) + \int_0^t \sigma[\sigma(\tau+t-s)ds + d\bar{z}])}]$$

$$+ \int_0^t \sigma[\sigma(\tau+t-s)\ ds + d\bar{z}\ ]\ .$$

$$= k(0,t+\tau) + [\ Ae^{\alpha(k(0,t+\tau) + \sigma^2(t\tau+\frac{1}{2}t^2) + \int_0^t \sigma d\bar{z})}]$$

$$+ Be^{-\alpha(k(0,t+\tau) + \sigma^2(t\tau+\frac{1}{2}t^2) + \int_0^t \sigma d\bar{z})}]$$

$$+\sigma^2(t\tau+\frac{1}{2}t^2) + \int_0^t \sigma\ d\bar{z}.$$

Similarly, assuming that the spot rate contains a speculative bubble, then for the spot

rate in the presence of a speculative bubble, where setting $\tau = 0$,

(24) $r(t) = f(t,t) = \tilde{f}(t,0)$

$$= k(0,t) + [\ Ae^{\alpha(\ k(0,t) + \frac{1}{2}\sigma^2 t^2 + \int_0^t \sigma d\bar{z})} + Be^{-\alpha(\ k(0,t) + \frac{1}{2}\sigma^2 t^2 + \int_0^t \sigma d\bar{z})}]$$

$$+ \frac{1}{2}\sigma^2 t^2 + \int_0^t \sigma\ d\bar{z}.$$

It seems that the spot rates explode as t goes infinity. This is not related to the bubble solution. This is because we choose Ho-Lee type volatility structure in (7). Since we assumed that the general solution g depends only on the fundamental k, our formula (24) says only that the forward rate process f is expressed in terms of a non-linear function with respect to the fundamental.

From (22) and (13), and *Itô*'s lemma (18), we obtain a stochastic differential equation of the forward rate process in the presence of a bubble under the expectations of changes in regimes,

(25) $$df = \theta(t,\tau)\ dt + \delta(t,\tau)\ d\hat{z}$$

where

$$\theta(t,\tau) = \frac{\alpha^2\sigma^2}{2}\ (Ae^{\alpha\tilde{k}(t,\tau)} + Be^{-\alpha\tilde{k}(t,\tau)}),$$

$$\delta(t,\tau) = [\ \sigma + \alpha\sigma\ (Ae^{\alpha\tilde{k}(t,\tau)} - Be^{-\alpha\tilde{k}(t,\tau)})]\ .$$

Then, under P,

(26) $$df = \psi(t,\tau)\ dt + \delta(t,\tau)\ dz$$

where

$$\psi(t,\tau) = [\ \theta(t,\tau)+\delta(t,\tau)\ (\sigma\tau-\lambda(t))\ ]\ .$$

$$= \frac{\partial \hat{f}}{\partial \tau}+\bar{\psi}(t,\tau)$$

where $\bar{\psi}(t,\tau) = [\ \theta(t,\tau)+\delta(t,\tau)\ (\sigma\tau-\lambda(t))\ ] - \frac{\partial \hat{f}}{\partial \tau}$. Even for simple Gaussian processes for the fundamental, the drift and diffusion function of the forward rate process are highly non-linear stochastic functions with respect to time and fundamentals.

Using the framework of HJM (1992) and Babbs (1991), the necessary condition for the absence of arbitrage opportunities requires the following relationship:

(27) $$-\int_0^\tau \bar{\psi}(t,s)ds + \frac{1}{2}[\int_0^\tau \delta(t,s)ds]^2 = x(t)\int_0^\tau \delta(t,s)ds$$

or $$-\int_0^\tau \psi(t,s)ds + \frac{1}{2}[\int_0^\tau \delta(t,s)ds]^2 = x(t)\int_0^\tau \delta(t,s)ds - \hat{f}(t,\tau)$$

where $x(t)$ is the market price risk in the presence of a speculative bubble. Equation (27) also implicitly guesses a relationship between the two market prices λ and x.

Now, based on the framework of Jamshidian (1991), Babbs (1991) and HJM (1992), the price of a pure discount bond in a speculative bubble follows from the stochastic differential equation

(28) $$\frac{dP(t,\tau)}{P(t,\tau)} = \beta^\tau(t)dt - \eta^\tau(t)dz$$

where

$$\eta^{\tau}(t) = \int_0^{\tau} \delta(t,s)ds$$

$$\beta^{\tau}(t) = r(t) - \int_0^{\tau} \psi(t,s)ds + \frac{1}{2}(\eta^{\tau}(t))^2.$$

Using equation (27), the drift term of a pure discount bond price process in the presence of a speculative bubble can be re-expressed as

$$r(t) + \eta^{\tau}(t)x(t) - \tilde{f}(t,\tau).$$

The drift term of the evolution of a zero coupon bond process in the presence of a speculative bubble has the similar structure as the process excluding the bubble in (10-1).

In the next step, using the Gaussian volatility structure, we explicitly find the value of the forward rates and the pure discount bond price $P(t,\tau)$ from (23). To do this, we have to remove the term $\int_0^t \sigma\, d\bar{z}$ of equation (23). We substitute (17) into (23) for $\int_0^t \sigma\, d\bar{z}$, obtaining

$$(29) \qquad \tilde{f}(t,\tau) = k(0,t+\tau) + [\, Ae^{\alpha(\, k(0,t+\tau)+\sigma^2(t\tau+\frac{1}{2}t^2)+\int_0^t \alpha d\bar{z})}]$$

$$+ Be^{-\alpha(\, k(0,t+\tau)+\sigma^2(t\tau+\frac{1}{2}t^2)+\int_0^t \alpha d\bar{z})}]$$

$$+\sigma^2(t\tau+\frac{1}{2}t^2)+\int_0^t \sigma\, d\bar{z}.$$

$$= k(0, t+\tau)$$

$$+ [\ Ae^{\alpha(\ k(0,t+\tau)+\sigma^2 t\tau + r_f(t) - k(0,t))} + Be^{-\alpha(\ k(0,t+\tau)+\sigma^2 t\tau + r_f(t) - k(0,t))}]$$

$$+\sigma^2 t\tau + r_f(t) - k(0, t).$$

where r_f is given by equation (17). Since $r(t) = f(t, t)$, the spot rate in the presence of a bubble can be expressed in terms of the fundamental r_f which is the spot rate without bubbles:

(30) $$r(t) = r_f + Ae^{\alpha r_f} + Be^{-\alpha r_f}.$$

The same result can be obtained by setting $\tau = 0$ in the first line of equation (22). This confirms that r in our bubble model is really the spot rate process in the presence of bubble, and this justifies equation (24). Unfortunately, we cannot express $\tilde{f}(t, \tau)$ in terms of $r(t)$, since we are not able to give an expression for r_f in terms of r in (30). For each fixed maturity T, the price of a pure discount bond in the presence of a speculative bubble involves the integration of the formidable equation (29):

(31) $$P(t, \tau) = e^{-\int_0^\tau \tilde{f}(t,s)ds}.$$

The corresponding zero-coupon yield for each fixed T is given by

(32) $$R(t, \tau) = -\frac{1}{\tau} \ln P(t, \tau).$$

Forward rates explode as τ moves toward infinity in our model. However, since for fixed T, the term $\sigma^2 t\tau$ which equals to $\sigma^2 t(T-t)$ in (29) has a global maximum

value with respect to t, forward rates, and consequently yields do not explode with finite fundamentals as t moves to maturity T, although the yield can be extremely high value. We call this the maturity effect. Unlike other financial assets such as a stock, the price of a bond having a fixed maturity date does not explode even in a bubble, due to this maturity effect.

6.4. A Discussion of the Determination of A and B

In this section, we shall discuss how to determine A and B in equation (22). In the area of exchange rates, there has been much research such as Bertola and Caballero (1992). In Bertola and Caballero, the authorities are assumed to intervene in the fundamental of the exchange rate with a target rate. With regard to interest-rate intervention, the government does not regulate all types of interest rates. The central bank usually has a target rate. For instance, the Federal Reserve Bank is known to have a target band for the Federal Funds rate in the USA. The study of Bertola and Caballero on target zones for exchange rates might be applied, for instance, to the Federal Reserve's target band for the Federal funds rate. However, the application of the method to our model is difficult, since we did not model the Federal Fund rate. In our model, we assumed that forward rates were affected by the expectation of a regime change in interest rates. To use the method of Bertola and Caballero directly, the forward rate should be used as an instrument of regime change in interest rates by the monetary authorities. This is not the usual practice. The Federal Fund rate may be closely related to the dynamic of forward rates. However, to assume that the government has a target rate for forward rates, would be a very strong assumption. Here, we leave A and B as parameters. In Appendix II, however, we demonstrate the determinant of A and B in the case where the government would use the forward rate as a target rate.

We plot the zero-coupon yields in Figures 1, 2 and 3, using the base case

parameter values given in Table 1, varying the volatility (σ), the fundamental spot rate (r_f), and the bubble amplitude (A), where we assume $A=-B$, respectively. We calculate forward rates from formula (29) and a pure discount bond price from equation (31), using summation rather than integration in the power part of the formula[55]. In Figures 1, 2 and 3, since we use $r_f(t)$ rather than $r(t)$ as an input value, the spot yield with a speculative bubble at $\tau=0$ jumps. The base parameter values are given in Table 1. Since the fundamental process in our model is Gaussian, yields are negative in some parameters. Figure 1 shows that uncertainty (volatility) generally increases the bubble size. However, for short maturity, since $\alpha=\sqrt{\frac{2}{\sigma^2}}$ dominates $\sigma^2 t(T-t)$, the yields in the case of the low σ is slightly higher. In our model, the bubble does not explode, although it can have extremely high values for longer maturities. Figure 2 shows that high fundamental spot rates increase the bubble size as well. The bubble amplitude (A) also affects the bubble positively.

Table 1. Base Case of Parameters

$k(0, t+\tau)=0.08$
$\sigma=0.08$
$T=10$
$r_f=0.08$
$k(0, t)=0.08$
$A=0.002$

55. See HJM (1990), p 422, for the justification of this.

6.5. Conclusion

This chapter analyzes a model of the term structure of interest rates in the presence of expectations of changes in regimes or of a speculative bubble. We apply Krugman and Miller's target zone model of exchange rates to the interest rate model. In this procedure, we re-interpret a traditional model of the model of the term structure, using changing probability measures. From the mathematical derivation of our model, first, we find that a non-linear effect might result from market expectations. This may explain the distributional features of interest rates such as the non-linear effect of volatility. Secondly, the classical Sargent-type model of the term structure of interest rates is reexpressed in the continuous time setting under the forward risk-adjusted measure. Thirdly, CLM conjectures about the non-existence of bubbles on assets such as bonds which have a fixed value on a terminal date is only partially correct. As seen in Section 6.3, for fixed maturity, the bubble cannot explode due to the maturity effects in our model.

Our model has some limitations, however. It allows negative interest rates in some parameter values as in Vasicek (1977). Secondly, in the case where the bubble solution explicitly depends on time, equation (20) has a PDE form. In such case, the solution for the PDE is likely to be very difficult. Hence, the extension of the model to the more general case is quite difficult.

Figure 1. Term Structure vs Volatility

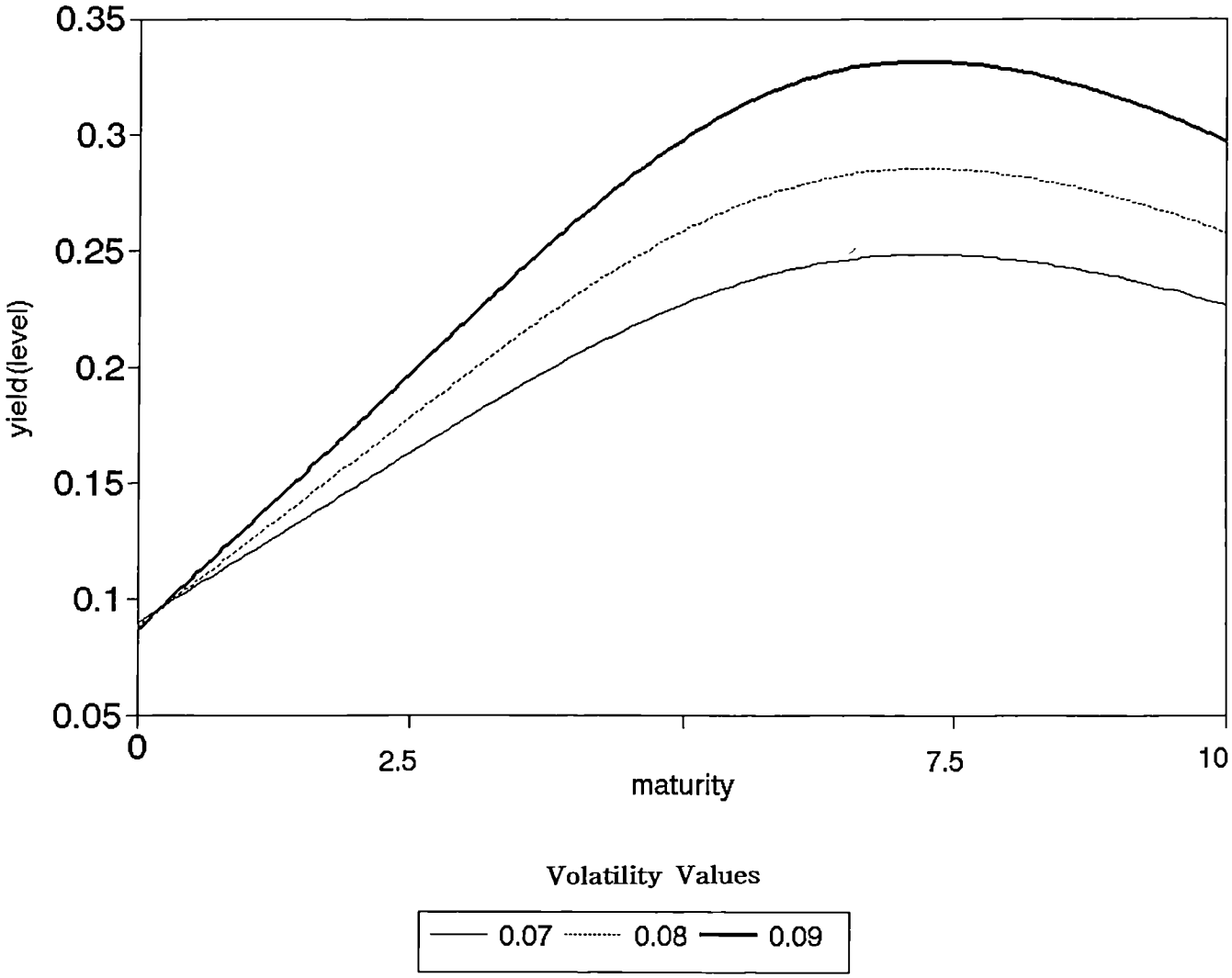


Figure 2. Term Structure vs Spot Rate
(fundmental spot rate)

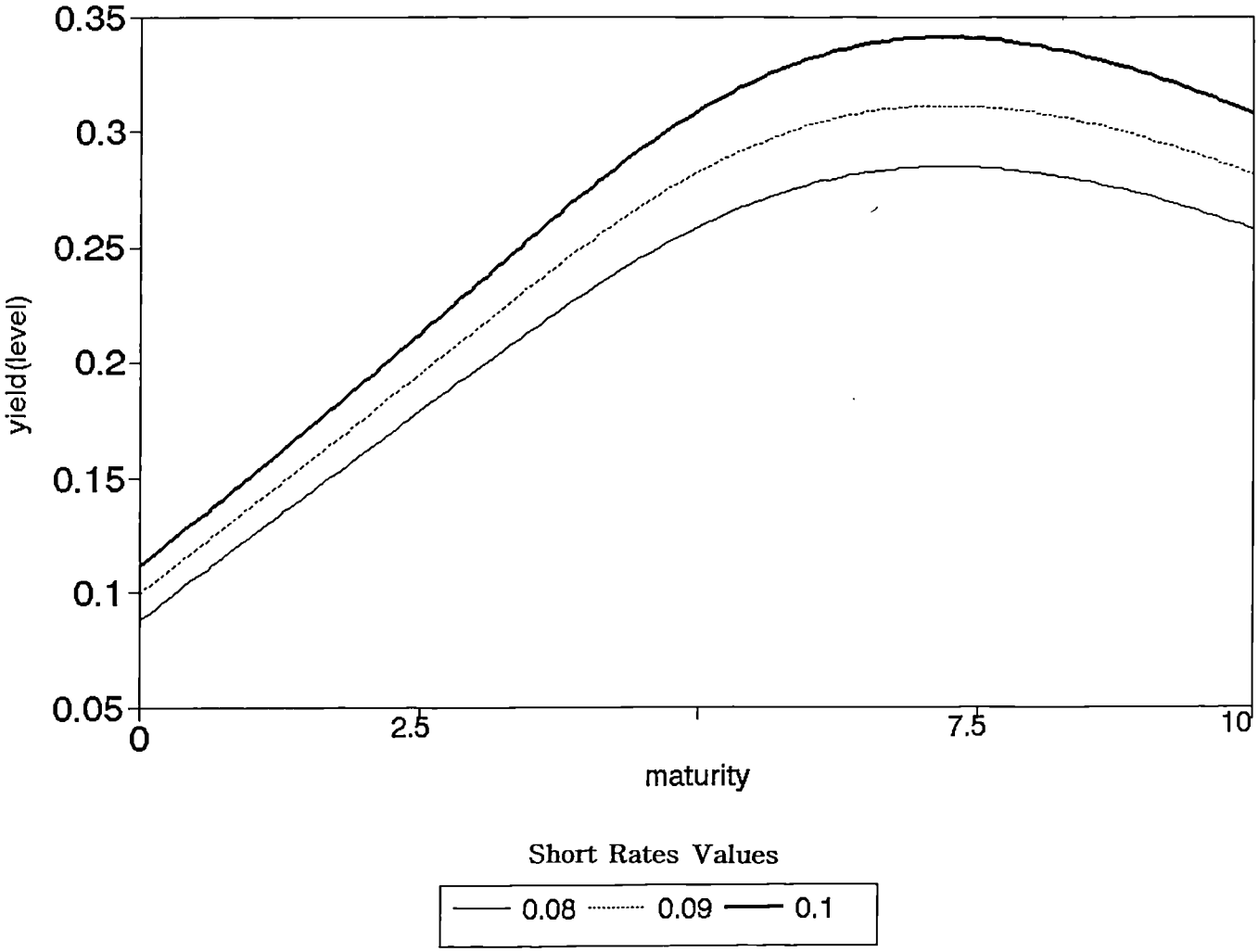
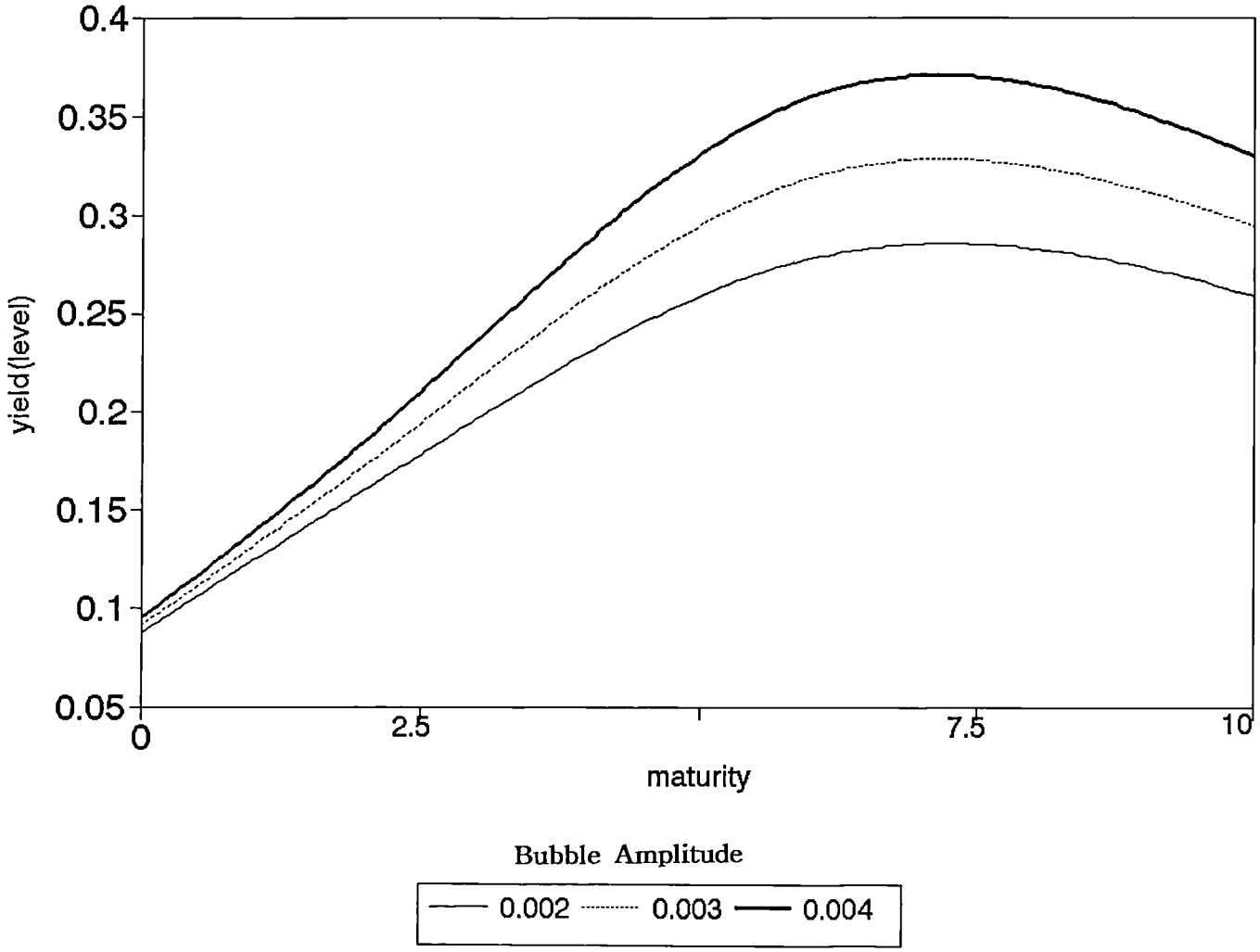


Figure 3. Term Structure vs A
(bubble amplitude)



Chapter 7. Summary and Conclusion

This dissertation examines three models of the term structure of interest rates to explain the distributional features of short rates. All three models incorporate novel features intended to improve their explanatory power. In Chapter 4, a three-factor affine model of the term structure of interest rates incorporating a stochastic volatility was presented. Compared with the two-factor model of Shaefer and Schwartz (1984), our model exhibits quite good explanatory power. In particular, one contribution of Chapter 4 is to present a closed-form solution of a pure discount bond price using three factors[56]. The empirical result of Chapter 4, however, is not satisfactory, since the parameters of the GARCH-X model for the discretization of the continuous time model do not map directly into the parameters of the our three-factor diffusion process.

The frequent jumps in interest rates observable in market data and the statistical properties of interest rates cannot be explained solely by the stochastic volatility model of Chapter 4. In Chapter 5, we extended the affine model of the term structure of DK (Duffie and Kan) to include jumps. The jump-diffusion model can possibly explain more of the empirical properties of interest rates. We presented a two-factor jump-diffusion model. To model the term structure of interest rates with official rates we assumed that the first factor follows a pure diffusion process and the second factor follows a pure jump process. Our two-factor model is especially appropriate for modelling official interest rates. To preserve the affine structure, the jump intensity should be a linear function with respect to the state variables.

In Chapter 6, we presented a model of the term structure of interest rates under expectations of changing regimes. Expectations were solved in a form of a bubble

56. The proof is available in Appendix I.

path. Although there have been a number of studies of bubble phenomena in stock prices and exchange rates, there has been no published bubble model of interest rates or bond prices. In Chapter 6, we showed that due to maturity effects, bubbles in the bond market did not explode for fixed maturity dates. In the derivation of the model, we also showed how the Sargent (1972)-type interest rates model could be reconstructed in the continuous-time framework of modern finance. Unfortunately, the approach we used cannot easily be extended to non-Gaussian volatility structures. A further empirical investigation of the model will be required.

To obtain tractability or a closed form solution for interest rate derivatives, we sometimes permit the interest model to allow negative interest rates: for instance, the Gaussian model and our two-factor jump-diffusion model. In Appendix III, we shall construct a theoretical model of the term structure of interest rates in the presence of jumps to see what kinds of term structure model do not allow negative interest rates. Following Flesaker and Hughston (1996), the interest rate positivity property is incorporated into the discount bond price.

Thus, in this thesis, we have presented three different models of the term structure of interest rates and a closed-form and approximate valuation formula for pure discount bond prices. We have tried to explain the current behaviour of interest rate processes in three different ways: stochastic volatility, possible jumps of interest rates, and market expectations of agents. All these models could be used, depending on the situation. For instance, as we demonstrated in Chapter 5, if we want to model interest rates using official set rates such as the discount rates in the US, we could use the jump-diffusion model.

We acknowledge several limitations of this thesis. First, in Chapter 4, we did not suggest or provide a satisfactory guide for a heuristic way to find $\hat{s}$. Futhermore, in our empirical work, we follow the conventional method of using GARCH-X to estimate the parameters of the models. This may affect our empirical results. The approach of Oakes (1997) might solve this problem, although it requires much computation. Secondly, in relation to argument in Chapter 5, before

the thesis was finished, Duffie and Kan (1996) published a paper including remarks on the possible extension of the affine model to the presence of jumps. Hence, the claim to have extended the model might not be correct. Instead, we include examples and empirical work in that chapter. In the empirical part of Chapter 5, we approximate the density, following Ball and Torous. Thirdly, in Chapter 6, when we use our model, we should use the fundamentals of spot rate and forward rate rather than observed rates as input values. This is definitely difficult to implement. In addition, an empirical investigation of our model must be needed.

Further research could be carried out, applying our models to bond option pricing and interest rate derivatives such as caps and swaps. The closed form solution of the transition density is not available. To create a pseudo-density, the use of control variate simulation or the binomial or trinomial method may offer a useful approach.

Appendix I. A Closed-Form Solution of a Pure Discount Bond Price

In this Appendix, we derive a closed-form solution of a pure discount bond price using our three factors. Our model is related to the empirical work of Litterman, Scheinkman, and Weiss (1988). Using OLS, they estimated a model[57] in which future short rates depend on 1) today's short rate ; 2) the level toward which the short rate is expected, as of today, to converge, which they call the long rate ; 3) the volatility of the long rate. They found that the yield curves were well explained according to the three variables. As explained in Chapter 4, however, our model assumes that future short rates depend on 1) the spread rate, which is today's short rate minus a factor affecting the level toward which the short rate is expected, as of today, to converge, which we call the long rate, similar to Litterman, Scheinkman, and Weiss (1988) ; 2) the volatility of the spread rate : 3) the long rate. When we express this mathematically, our three factors are

(1) $$ds = \alpha(\bar{s} - s)dt + \sqrt{v}dz_1$$

(2) $$dv = \gamma(\bar{v} - v) + \delta\sqrt{v}dz_2$$

(3) $$dl = m(\bar{l} - l)dt + \sigma\sqrt{l}dz_3$$

where $s = r - l$ is the spread rate, r is the short rate, l is the long rate, and v is the volatility of the spread rate. We assume that dz_1, dz_2 and dz_3 are standard

57. They did not set up a mathematical model.

Gauss-Wiener processes. As in Chen (1996), the assumption of $dz_1 dz_2 = \rho dt$ with $\rho \neq 0$ might not be necessary, because the spread rate s, its volatility v and the long rate l are already correlated through the stochastic differential equation for the spread rate. For instance, the spread rate is affected by the volatility of the spread rate through the diffusion term of the spread rate. We assume that the correlation between dz_1 and dz_2 is zero in this model. In addition, the correlation between the spread rate and the long rate is very low as shown in Chapter 4. We assume that all correlations between dz_1, dz_2 and dz_3 are zero.

First, to see the effect of the long rate on the mean reversion of the short rate, we substitute (3) into (1), ignoring the diffusion terms. Then setting $\bar{r} = \bar{s} - \bar{l}$, we obtain

$$dr = \alpha(\bar{r} - \bar{l} - r + l)dt + m(\bar{l} - l)dt.$$

$$= \alpha[\ \{\bar{r} - \bar{l} + \frac{m\bar{l}}{\alpha} + (1 - \frac{m}{\alpha})l\} - r]\ dt.$$

$$= \alpha[\ \{a + bl\} - r]\ dt.$$

where $a = \bar{r} - \bar{l} + \frac{m\bar{l}}{\alpha}$ and $b = 1 - \frac{m}{\alpha}$. Hence, our model implies that the stochastic mean reversion level is linear in l. As in the definition of the long rate, the factor l is a bench-mark rate affecting the level toward which the short rate is expected to converge.

Next, we derive a closed-form solution of a pure discount bond price. Let $P(s, v, l, \tau)$ be a pure discount bond price with time to maturity $\tau = (T - t)$, and a terminal payoff 1. Then, following a similar procedure to that in Section 4.2.2. $P(s, v, l, \tau)$ satisfies the following PDE :

(4) $$\frac{1}{2}vP_{vv}+\frac{1}{2}\delta^2 vP_{vv}+\frac{1}{2}\sigma^2 lP_{ll}+(\alpha(\bar{s}-s)+\lambda_s v)P_s$$
$$+(\gamma\bar{v}-\gamma' v)P_v+(m\bar{l}-m' l)P_l+P_t-(s+l)P = 0$$

with the initial condition

$$P(s, v, l, 0)=1$$

where $\gamma'=\gamma-\lambda_v\delta$, $m'=m-\lambda_l\sigma$, and λ_s, λ_v, and λ_l are the market prices of risk of the spread rate, the volatility of spread rate, and the long-term yield respectively assumed to be constant. The market price of risk defined here is different from that of Chapter 4, where market price of risk is assumed to depend on the diffusion parameters.

Proposition AI : Assuming the spot rate dynamics specified by (1), (2) and (3), the value of a pure discount bond promising to one unit at time T, $P(s, v, l, \tau)$ is given by

$$P(s, v, l, \tau) = A(\tau)e^{-B(\tau)s-C(\tau)v-D(\tau)l}$$

where

$$A(\tau)=\exp\left(-\left(\tau+\frac{\exp(-\alpha\tau)}{k}-\frac{1}{k}\right)\bar{s}\right)(c_1\exp(g_1\tau)+c_2\exp(g_2\tau))^{-\frac{2m\bar{l}}{\sigma^2}}$$
$$\times\left(\frac{y^{\rho}\exp(-\psi(y-1))(\Lambda U(Q,S,2\psi y)+M(Q,S,2\psi y))}{\Lambda U(Q,S,2\psi)+M(Q,S,2\psi)}\right)^{-\frac{2\gamma\bar{v}}{\delta^2}}$$

$$B(\tau)=\frac{1-e^{-\alpha\tau}}{\alpha}$$

$$C(\tau)=\frac{2\alpha}{\delta^2}\{-\rho+\psi y+\frac{2y\Lambda Q\psi U(Q+1,S+1,2\psi y)}{\Lambda U(Q,S,2\psi y)+M(Q,S,2\psi y)}$$

$$-\frac{2y\psi\frac{Q}{S}M(Q+1,S+1,2\psi y)}{\Lambda U(Q,S,2\psi y)+M(Q,S,2\psi y)}\}$$

$$D(\tau)=\frac{2}{\sigma^2}\frac{g_1c_1\exp(g_1\tau)+g_2c_2\exp(g_2\tau)}{c_1\exp(g_1\tau)+c_2\exp(g_2\tau)}.$$

where

$$y=\exp(-\alpha\tau)$$

$$\Lambda=-\frac{(\rho-\psi)M(Q,S,2\psi)-2\psi\frac{Q}{S}M(Q+1,S+1,2\psi)}{(\rho-\psi)U(Q,S,2\psi)+2\psi QM(Q+1,S+1,2\psi)}$$

$$Q=-\frac{(\alpha\lambda_s-1)\frac{\delta^2}{2\alpha^4}}{2\psi}+\frac{S}{2}$$

$$S=\frac{\alpha+\sqrt{\gamma'^2-4\theta\alpha^2}}{\alpha}$$

$$\rho=\frac{\gamma'+\sqrt{\gamma'^2-4\theta\alpha^2}}{2\alpha}$$

$$\theta=(1-2\alpha\lambda_s)\frac{\delta^2}{4\alpha^4}$$

$$\psi=i\frac{\delta}{2\alpha^2}$$

$$g_1=\frac{-m+\sqrt{m^2-2\sigma^2}}{2}$$

$$g_2=\frac{-m-\sqrt{m^2-2\sigma^2}}{2}$$

$$c_1=-\frac{g_2}{\sqrt{m^2-2\sigma^2}}$$

$$c_2=\frac{g_1}{\sqrt{m^2-2\sigma^2}}.$$

and where

$$M(a, b, c) = 1 + \frac{ac}{b} + \ldots\ldots + \frac{(a)_n c^n}{(b)_n n!} + \ldots\ldots,$$

where

$M(a, b, c)$ is the Kummer function, and

$$(a)_n = a(a+1)(a+2)\ldots.(a+n-1), \quad a_0 = 1$$

and

$$U(a, b, c) = \frac{1}{\Gamma(a)} \int_0^\infty \exp(-ct) t^{a-1} (1+t)^{b-a-1} dt,$$

where $U(a, b, c)$ is a confluent hypergeometric function, and Γ denotes gamma function.

Proof:

Since our model is an affine type model of the term structure of interest rates, the trial form of solution is:

$$P(s, v, l, \tau) = A(\tau) e^{-B(\tau)s - C(\tau)v - D(\tau)l}.$$

If we substitute the function and its derivatives into (4), we obtain the following PDE:

$$\text{(5)} \qquad \frac{1}{2} vPB^2 + \frac{1}{2} \delta^2 vPC^2 + \frac{1}{2} \sigma^2 lPD^2 + (\alpha(\bar{s} - s) + \lambda_s v)P(-B)$$
$$+ (\gamma \bar{v} - \gamma' v)P(-C) + (m\bar{l} - m' l)P(-D)$$
$$- \frac{A'}{A} P + (B's + C'v + D'l)P - (l+s)P = 0$$

since

$$P_s = P(-B)$$

$$P_{ss} = PB^2$$

$$P_v = P(-C)$$

$$P_{vv} = PC^2$$

$$P_l = P(-D)$$

$$P_{ll} = PD^2$$

$$-P_t = P_\tau = \frac{A'}{A} P + (-B's - C'v - D'l)P.$$

By collecting terms in s, v, l, and remaining terms, (5) yields the system of the following ODEs:

(6) $$aB + B' - 1 = 0, \quad \text{for } s$$

(7) $$\frac{1}{2} B^2 + \frac{1}{2} \delta^2 C^2 - \lambda_s B + \gamma' C + C' = 0, \quad \text{for } v$$

(8) $$\frac{1}{2} \sigma^2 D^2 + m' D + D' - 1 = 0, \text{ for } l$$

(9) $$a\bar{s}B + \gamma\bar{v}C + m\bar{l}D + \frac{A'}{A} = 0, \text{ for the rest}$$

with initial conditions

$$A(0) = 1, \quad B(0) = 0, \quad C(0) = 0, \text{ and } D(0) = 0.$$

Equation (6) has the solution

(10) $$B(\tau)=\frac{1-e^{-\alpha\tau}}{\alpha}$$

Equation (7) is of a Ricatti form. Following Fong and Vasicek (1991) and Chen (1996), the solution to (7) is given by:

(11) $$C(\tau)=\frac{2}{\delta^2}\frac{x'}{x}$$

where $x(\tau)$ satisfies

(12) $$x''+\gamma' x'+\frac{\delta^2}{2}\left(\frac{\lambda_s}{\alpha}(e^{-\alpha\tau}-1)+\frac{1}{2\alpha^2}(e^{-\alpha\tau}-1)^2\right)x=0$$

This is the same form of factor loading function of the volatility of the short rate in Chen (1996)[58]. Hence, the solution is :

$$C(\tau)=\frac{2\alpha}{\delta^2}\{-\rho+\psi y+\frac{2y\Lambda Q\psi U(Q+1,S+1,2\psi y)}{\Lambda U(Q,S,2\psi y)+M(Q,S,2\psi y)}$$
$$-\frac{2y\psi\frac{Q}{S}M(Q+1,S+1,2\psi y)}{\Lambda U(Q,S,2\psi y)+M(Q,S,2\psi y)}\}$$

where

$$y=\exp(-\alpha\tau)$$

$$\Lambda=-\frac{(\rho-\psi)M(Q,S,2\psi)-2\psi\frac{Q}{S}M(Q+1,S+1,2\psi)}{(\rho-\psi)U(Q,S,2\psi)+2\psi QM(Q+1,S+1,2\psi)}$$

58. See *Interest Rate Dynamics, Derivatives Pricing, and Risk Management* (1997) by Chen, p. 133.

$$Q=-\frac{(\alpha\lambda_s-1)\frac{\delta^2}{2\alpha^4}}{2\psi}+\frac{S}{2}$$

$$S=\frac{\alpha+\sqrt{\gamma'^2-4\theta\alpha^2}}{\alpha}$$

$$\rho=\frac{\gamma'+\sqrt{\gamma'^2-4\theta\alpha^2}}{2\alpha}$$

$$\theta=(1-2\alpha\lambda_s)\frac{\delta^2}{4\alpha^4}.$$

$$\psi=i\frac{\delta}{2\alpha^2}$$

Equation (8) also has a Ricatti form. As in Mclachlan (1950), the solution is

(13) $$D(\tau)=\frac{2}{\sigma^2}\frac{u'}{u}$$

Then, $u(\tau)$ satisfies

(14) $$u''+mu'+\frac{\sigma^2}{2}u=0,\quad u(0)=1, u'(0)=0.$$

Since this is an homogeneous second order ODE, we obtain a general solution:

(15) $$u(\tau)=c_1\exp(g_1\tau)+c_2\exp(g_2\tau)$$

where

$$g_1=\frac{-m+\sqrt{m^2-2\sigma^2}}{2}$$

$$g_2=\frac{-m-\sqrt{m^2-2\sigma^2}}{2}$$

$$c_1=\frac{g_2}{g_2-g_1}=-\frac{g_2}{\sqrt{m^2-2\sigma^2}}$$

$$c_2 = \frac{g_1}{g_1 - g_2} = \frac{g_1}{\sqrt{m^2 - 2\sigma^2}}.$$

Then

(16) $$u'(\tau) = g_1 c_1 \exp(g_1 \tau) + g_2 c_2 \exp(g_2 \tau),$$

and

$$D(\tau) = \frac{2}{\sigma^2} \frac{g_1 c_1 \exp(g_1 \tau) + g_2 c_2 \exp(g_2 \tau)}{c_1 \exp(g_1 \tau) + c_2 \exp(g_2 \tau)}.$$

Since each of the functions $C(\tau)$ and $D(\tau)$ has the form $\frac{w(\tau)'}{w(\tau)}$, integrating from 0 to τ, (9) yields

(17) $$\ln A(\tau) = -\left(\tau + \frac{\exp(-\alpha\tau)}{k} - \frac{1}{k}\right)\bar{s} - \frac{2m\bar{l}}{\sigma^2} \ln \frac{u(\tau)}{u(0)} - \frac{2\gamma\bar{v}}{\delta^2} \ln \frac{x(\tau)}{x(0)}.$$

Then

$$A(\tau) = \exp\left(-\left(\tau + \frac{\exp(-\alpha\tau)}{k} - \frac{1}{k}\right)\bar{s}\right)\left(\frac{u(\tau)}{u(0)}\right)^{-\frac{2m\bar{l}}{\sigma^2}}\left(\frac{x(\tau)}{x(0)}\right)^{-\frac{2\gamma\bar{v}}{\delta^2}}$$

■

Appendix II. A Determination of A and B

Here, we assume that the government uses the forward rate as a target rate. We define a 'government central parity process' N_t which can be understood as the central bank's target rate for k. Alternatively, the target band processes (upper and lower bound) can be used.

Suppose monetary authorities may alter the level of k by multiples of a jump size J. The jump size J is assumed to be fixed. First, to simplify the determination of A and B, we give an example of $A=-B$. We assume that if the fundamental reaches some boundaries $N_t + J$, or $N_t - J$, then the authorities may either bring the fundamental back to N_t, or jump it up or down $2J$ from N_t with new target rates $N_{t(+)} = N_t \pm 2J$. The latter case means a change of regimes. For instance, if the fundamental reaches N_t+J, the authorities may bring the fundamental back to N_t or jump it up N_t+2J. Then N_t+2J becomes a new target rate. We assume that the authorities stick to current regime with probability $1-p$, and changes the regime with p, when the interest rate is at the margin of the fluctuation band. The bivariate process $\{k, N\}$ is assumed to be jointly Markov, and the forward rate f is a function of the current level of the fundamental k and the current central parity, $f = f(k, N)$. In the interior of every band, the function $f(k, N)$ has the symmetric form[59] :

(1) $$f(k, N) = k + Be^{-\alpha(k-N)} - Be^{\alpha(k-N)}$$

59. The solution is slightly different from the first line of equation (22). This is for removing the bubble when the fundamental k_t is equal to the central parity N_t.

where $\alpha=\sqrt{\frac{2}{\sigma^2}}$. To determine B, we impose the condition that the forward rate not to be expected to change at times when intervention is known to be imminent. Thus, when $k=N_t+J$, it must be true that

$$(2) \qquad p\, f(N_t+2J\,,\ N_t+2J) + (1-p)f(N_t\,,\ N_t) = f(N_t+J\,,\ N_t).$$

The left-hand side of (2) weights with the respective probabilities the two possible forward rates just after the decision to defend or change the band is revealed. To prevent arbitrage opportunities, this expected forward rate must equal that prevailing at the instant the fundamentals reach the barrier and the authorities's decision is called for.

Using (1) in (2)

$$(3) \qquad B = \frac{(1-2p)J}{e^{\alpha J}-e^{-\alpha J}}.$$

Here, $B\geq 0$ only if $0\leq p\leq \frac{1}{2}$. Then the forward rate f is S-shaped with respect to the fundamental k. When $p=0$, the band is always defended.

Appendix III. Positive Interest Rates in the Presence of Jumps

AIII.1 Introduction

One disadvantage of the HJM framework is the positive probability of negative interest rates. Miltersen (1994) obtains a sufficient condition which assures positive interest rates in the HJM framework. More recently, Flesaker and Hughston (1996) introduced a new framework for pricing interest rate derivatives with an absence of negative interest rates. The purpose of this Appendix is to extend their approach to the presence of jumps. Unlike the approach of Miltersen, Flesaker and Hughston seek to incorporate the condition of positive interest rates into the bond price process. Shirakawa (1991) extends the HJM framework to the multidimensional Poisson-Gaussian process. Unfortunately, his model also allows interest rates to be negative with a positive probability.

AIII.2 Review of the Flesaker and Hughston Approach

In this section, we review the main result of Flesaker and Hughston (1996, hereafter, FH). The bond price process P_{tT} of the bond maturing at T is assumed to be adapted to the filtration ξ_T. The bond market is assumed to include a numeraire security B_t. This security is also adapted to the filtration ξ_T. The probability space (Ω, P, ξ) represents the economy. Following FH, we introduce some notation and assumptions. (1) Let P_{ab} be the value of a discount bond at time

a that matures at time b. (2) P_{ab} is assumed to be differential in b. (3) We consider a family of bond price processes P_{ab} for which $0 \le a \le b \le T$, where T is the fixed terminal date. From Harrison and Pliska (1981), there exists a unique pricing measure with respect to the ratio $\frac{P_{ab}}{P_{aT}}$, which is a martingale in a for any bond in the given family. Following FH, we denote this martingale by N_{ab}:

(1) $$N_{ab} = \frac{P_{ab}}{P_{aT}}, \quad P_{aa} = 1.$$

Then, using some algebra, we obtain

(2) $$P_{ab} = \frac{N_{ab}}{N_{aa}}, \quad N_{aT} = 1.$$

For any c such that $0 \le a \le b \le c \le T$,

(3) $$\frac{P_{ac}}{P_{ab}} = \frac{N_{ac}}{N_{ab}}.$$

Positive interest rates require the condition $\frac{P_{ac}}{P_{ab}} < 1$. From (3), this condition implies that $\frac{N_{ac}}{N_{ab}} < 1$ or $\frac{\partial N_{ab}}{\partial b} < 0$. Hence, there exists a positive martingale M_{ab}, in a, subject to $M_{0b} = 1$ such that

(4) $$\frac{\partial N_{ab}}{\partial b} = \frac{\partial N_{0b}}{\partial b} M_{ab}$$

satisfying the boundary condition $N_{aT} = 1$ and the initial condition $N_{0b} = \frac{P_{0b}}{P_{0T}}$.

The solution of the differential equation (4) is given by:

(5) $$N_{ab} = 1 - \frac{\int_b^T \partial_s P_{0s} M_{as} ds}{P_{0T}}.$$

where $\partial_s P_{0s} = \frac{\partial P_{0s}}{\partial s}$.

Substituting (5) into (2), FH obtain the following formula:

(6) $$P_{ab} = \frac{P_{0T} - \int_b^T \partial_s P_{0s} M_{as} ds}{P_{0T} - \int_a^T \partial_s P_{0s} M_{as} ds}.$$

Since the choice of T is arbitrary, following FH, we take the limit as T goes to infinity to simplify our analysis. Then equation (6) has a natural expression for large T. For the bond price process we obtain:

(7) $$P_{ab} = \frac{\int_b^\infty \partial_s P_{0s} M_{as} ds}{\int_a^\infty \partial_s P_{0s} M_{as} ds}.$$

Following FH, to specify the bond price process (7), we require two conditions. The first is the initial discount function, or

(i) $P_{00} = 1$, $0 < P_{0b} \leq 1$ for all $b \geq 0$

(ii) There exists $\partial_b P_{0b}$ for all $b \geq 0$, and $\partial_b P_{0b}$ is negative.

.The second condition is the family of positive martingales, or

(iii) For $0 \leq a \leq b \leq s$, $M_{as} = E_a(M_{bs})$, $M_{as} > 0$, $M_{0s} = 1$, $\lim_{s \to \infty} M_{as} = 1$.

Conditions (i) and (ii) imply that all rates exist and are positive initially. Condition (iii) requires that the process M_{as} should be a positive martingale.

Instantaneous forward rates $f_{ab} = -\frac{\partial \ln P_{ab}}{\partial b}$ are given by

(8) $$f_{ab} = \frac{\partial_b P_{0b} M_{ab}}{\int_b^T \partial_s P_{0s} M_{as} ds}.$$

Since the numerator and denominator of (8) are both negative according to the conditions (i), (ii) and (iii), the forward rates are positive. Similarly, the short rate $r_a = f_{aa}$ is also positive and is given by

(9) $$r_a = \frac{\partial_a P_{0a} M_{aa}}{\int_a^T \partial_s P_{0s} M_{as} ds}.$$

This is the main result of Flesaker and Hughton (1996).

AIII.3 An Extension to Jumps

In the next step, we incorporate the spot rates positivity property (9) into the drift, volatility and the distribution of jump size of the discount bond process. First, as in the previous section, we define a positive martingale M_{as}. By the martingale representation theorem, we can express M_{as} as

(10) $$\frac{dM_{as}}{M_{as}} = \sigma_{as} dz_a + \gamma_{as}(dQ - \lambda da)$$

where σ_{as} and γ_{as} are adapted processes, and z_a and Q are one-dimensional Wiener

and Poisson processes, respectively. We posit the following proposition for the discount bond price process which guarantees positive interest rate :

Proposition AIII :

With the expression for the short rate (9) and for the positive martingale process (10), for fixed b, the discount bond process P_{ab}, which guarantees positive interest rates has the following stochastic process

$$\text{(11)} \qquad \frac{dP_{ab}}{P_{ab}} = (r_a - V_a \Pi_{ab} - \lambda \Phi_{ab})da + \Pi_{ab} dz_a + \Phi_{ab}(1-\Gamma_a)dQ$$

where

$$\text{(12)} \qquad V_{ab} = \frac{\int_b^\infty \partial_s P_{0s} M_{as} \sigma_{as} ds}{\int_b^\infty \partial_s P_{0s} M_{as} ds}, \qquad \Gamma_{ab} = \frac{\int_b^\infty \partial_s P_{0s} M_{as} \gamma_{as} ds}{\int_b^\infty \partial_s P_{0s} M_{as} ds}$$

and where $\Gamma_{aa} = \Gamma_a$, $V_{aa} = V_a$, $\Phi_{ab} = \Gamma_{ab} - \Gamma_a$, and $\Pi_{ab} = V_{ab} - V_a$.

Proof :

First, we substitute the positive martingale process M_{as} (10) into equation (7). To obtain a stochastic process for a discount bond process, we use the following *Itô* identity:

$$d(\frac{X}{Y}) = \frac{dX}{Y} - \frac{XdY}{Y^2} + \frac{XdY^2}{Y^3} - \frac{dXdY}{Y^2}$$

or

$$\text{(13)} \qquad \frac{d(\frac{X}{Y})}{\frac{X}{Y}} = \frac{dX}{X} - \frac{dY}{Y} + (\frac{dY}{Y})^2 - \frac{dXdY}{XY}$$

If we set :

$$X = \int_b^\infty \partial_s P_{0s} M_{as} ds, \qquad Y = \int_a^\infty \partial_s P_{0s} M_{as} ds, \quad \text{and from (7)}$$

$$P_{ab} = \frac{X}{Y}$$

then, for fixed b,

$$dX = \int_b^\infty \partial_s P_{0s} dM_{as} ds, \qquad dY = \int_a^\infty \partial_s P_{0s} dM_{as} ds - \partial_a P_{0a} dM_{as} da.$$

Then, using[60] $dzdQ=0$, $dz^2=dt$, $dzdt=0$, $dQ^2=dQ$, $dt^2=0$, and $dQdt=0$, and defining

$$V_{ab} = \frac{\int_b^\infty \partial_s P_{0s} M_{as} \sigma_{as} ds}{\int_b^\infty \partial_s P_{0s} M_{as} ds}, \qquad \Gamma_{ab} = \frac{\int_b^\infty \partial_s P_{0s} M_{as} \gamma_{as} ds}{\int_b^\infty \partial_s P_{0s} M_{as} ds}$$

we obtain:

$$\frac{dX}{X} = V_{ab} dz_a + \Gamma_{ab}(dQ - \lambda da),$$

$$\frac{dY}{Y} = V_a dz_a - r_a da + \Gamma_a(dQ - \lambda da),$$

$$\left(\frac{dY}{Y}\right)^2 = V_a^2 da + \Gamma_a^2 dQ,$$

and

$$\frac{dX}{X}\frac{dY}{Y} = V_{ab} V_a da + \Gamma_{ab} \Gamma_a dQ.$$

60. See *Prices in Financial Market* by Dothan (1990), p. 262 for the calculation of the quadratic covariation for the Poisson process. See *Stochastic Integration and Differential Equation* by Protter (1996), p. 90 for the semimartingale process.

If we substitute these four equations into (13), we obtain (11).

■

Next, as in FH, we briefly discuss an appropriate change of measure to move from the terminal measure to the risk neutral measure.

As in Babbs and Webber (1994), in the risk neutral measure, we want to find z_a^0 and λ^0 such that $dz_a^0 = dz_a - \theta_a da$ and $\lambda^0 = (1-\eta_a)\lambda$, where θ and η are the market prices of risk. The θ_a and η_a are determined to ensure that the numeraire security B_a satisfies the following condition:

(14) $$dB_a = r_a B_a da$$

where

(15) $$B_a = \frac{\rho_a}{-\int_a^\infty \partial_a P_{0s} M_{as} ds}.$$

and ρ_a is the Radon-Nikodym derivative[61]. Then we can write

(16) $$B_a = \rho_a P_a$$

where

$$P_a = \frac{1}{-\int_a^\infty \partial_a P_{0s} M_{as} ds}$$

An advantage of the expressions (15) and (16) is that the money market account B_a is identified as ratio of a martingale process ρ_a and the discount factor P_a.

61. See Shirakawa (1991), p. 84.

As explained in FH, the process P_a is viewed as an absolute numeraire. Following the same procedure in Proposition AIII taking $X=1$ and $Y = -\int_a^{\infty} \partial_a P_{0s} M_{as} ds$, the relevant stochastic process for P_a is given by

$$\frac{dP_a}{P_a} = (r_a + V_a^2 + \lambda \Gamma_a) da - V_a dz_a + (\Gamma_a^2 - \Gamma_a) dQ. \tag{17}$$

This formula, without the jump term, is obtained by FH (1996, 1997). According to Flesaker and Hughston (1997), the price process P_a as the natural numeraire has the property that the ratio of any bond price to this numeraire is a martingale under the risk-neutral measure. As in the discount bond price process, which guarantees positive interest rates (11), V_a may be identified as the volatility of the bond price. Similarly, Γ_a can be obtained as the jump distribution of the bond price.

AIII.4 The Pricing of Contingent Claims

To price the contingent claims in the terminal or T-probability measure, we define C_a to be the random payout of an interest rate derivative at time a. Then, for $t \le a \le T$, the conditional expectation, $E[\frac{C_a}{P_{aT}} \mid \xi_t]$ is a martingale under T-measure. The present value of the derivative C_0 can be expressed as

$$C_0 = P_{0T} E[\frac{C_a}{P_{aT}}] . \tag{18}$$

Setting $b = T$ in (6), equation (18) becomes :

$$C_0 = E[\ (P_{0T} - \int_a^T \partial_s P_{0s} M_{as} ds) * C_a\ ] . \tag{19}$$

In a similar way as in (7), (19) can be represented as

$$C_0 = E[\ -(\int_a^{\infty} \partial_s P_{0s} M_{as} ds) * C_a\ ] \tag{20}$$

where the positive martingale process M_{as} is expressed as :

$$\frac{dM_{as}}{M_{as}} = \sigma_{as} dz_a + \gamma_{as}(dQ - \lambda da). \tag{21}$$

This is the main framework in this dissertation for positive interest rates in the presence of jumps.

AIII.5 Conclusion

In this Appendix, we have extended the approach of FH to the presence of jumps in interest rates. Following FH, we have incorporated the positive interest rates property into the discount bond price, but in a different way from that of Miltersen (1994). Miltersen incorporated the condition into the forward rate process. As seen in equation (11), the condition of positive interest rates depends on the types of the volatility structures and the distribution of jump sizes of the bond price process.

Appendix IV. Review on GMM Estimation of the Multi-Factor Diffusion Model

AIV.1 Introduction

The empirical investigation of a continuous model is not straightforward. In the case where underlying state variables are not observable such as stochastic volatility and the time-varying mean of the short rates, the empirical work is more difficult. This means that we do not usually determine the probabilistic behaviors of underlying state variables. If we know the transition density of certain underlying state variables, it is straightforward to apply a Maximum Likelihood method to a series of discretely sampled data.

Except for a limited class of diffusion processes, it is difficult to obtain closed form solutions for the density of the state variables. Among many different types of diffusion processes, only a few have a known density. In the case of the CIR process, the density of the process is known to be a gamma distribution. In the two-factor case (two factors are orthogonal to each other) the density is known to be a non-central χ^2 distribution.

In contrast to the stochastic differential model approach, Nelson (1990) shows that some families of discrete ARCH models converge in distribution to an *Itô* process as the length of the discrete time interval goes to zero. According to his study, in continuous time the stationary distribution for the GARCH(1,1) conditional variance process is an inverted gamma distribution; and in a conditionally normal GARCH(1,1)

process observed at short time intervals, the innovations process is approximately distributed as a Student t. In the Exponential ARCH model of Nelson (1990), the conditional variance in continuous time is lognormal; and in the discrete time model when time intervals are short, the stationary distribution of the innovations is approximately a normal-lognormal mixture. This result can broaden the application of some types of stochastic diffusion models with stochastic volatility (possibly two-factor models) to the Maximum Likelihood method in estimating parameters.

When increasing the number of factors, however, it is almost impossible for one to get an analytic form of the density. Although the transition density is determined by the Kolmogorov forward equation (or the Fokker-Planck PDE), solving the PDE numerically is clearly difficult especially in the multiple-factor case.

An alternative approach is to use a Generalized Method of Moment (GMM) procedure to estimate the parameters of a diffusion process using a discretised model. This approach was used recently by CKLS (1992). An important feature of GMM estimation is the performance of a specification test through a test of over-identifying restrictions. An advantage of the approach over ML is that the transition density is not require to be determined. In this Appendix, we review the moment approximation procedure of the diffusion process in multivariate factors due to Oakes (1997). Oakes (1997) proposes that the GMM estimation be carried out by replacing the true conditional moments by their numerical approximation in the multiple-factor case using an orthogonalizing condition.

AIV.2 GMM Estimation of the Diffusion Model

In this section we summarise an application of GMM to the diffusion model. For simplicity, only a one-dimensional diffusion process is considered. The concepts can easily be extended to the multi-dimensional case.

Let $\{x(t) \in R : t \in (0, \infty)\}$ be a stochastic process on the probability measure space

(Ω, P, ξ). Suppose that $x(t)$ satisfies the stochastic differential equation.

(1) $$dx(t) = \mu(x, \beta)dt + \sigma(x, \beta)dz$$

with $x(0) = \hat{x}$. z is an one-dimensional standard Brownian motion on measure space (Ω, P, ξ). β is the vector of unknown parameters to be estimated. $\mu(x, \beta)$ and $\sigma(x, \beta)$ are assumed to be continuous functions in both x and β, and to satisfy the usual growth and Lipschitz conditions. This means that there exist positive constants c, k such that for all $x, y \in \Re$, and $\beta \in \Theta$ for some compact set in $\Re$,,

(2) $$|\mu(x, \beta)| + |\sigma(x, \beta)| < c(1+|x|)$$

(3) $$|\mu(x, \beta)-\mu(y, \beta)| + |\sigma(x, \beta)-\sigma(y, \beta)| < k|x-y|.$$

The above regularity conditions ensure that a solution to (1) exists and is unique. Note that c and k are assumed to be independent of β. The process x is time-homogeneous. That is, the drift μ and the diffusion σ are independent of time t. This assumption is required for the stationarity of x. Finally, the starting point $\hat{x}$ is a random variable.

We sample x on $T+1$ date $t=0, 1, .., T$. For each t from 1 to T, form a vector of functions of current and past values of x and the parameter vector, denoted $f_t(x, \beta)$, satisfying the following moment condition under the null hypothesis that $\beta = \beta_0$,

(4) $$E[\ f_t(x, \beta_0)] = 0.$$

For any parameter vector $\beta \in \Theta$, let

$$G_T(\beta) = \frac{1}{T}\sum_{t=1}^{T} f_t(x, \beta)$$

If $f_t(x, \beta)$ satisfies a strong or weak law of large numbers, then $plim_T\, G_T(\beta_0) = 0$. We have $plim_T\, G_T(\beta_0) = 0$ if and only if $\beta = \beta_0$. The GMM estimator is obtained as the solution to the non-linear minimization program

$$\hat{\beta}_T = \min_{\beta \in \Theta} G_T(\beta)' W_T G_T(\beta) \tag{5}$$

where W_T is a symmetric, positive semi-definite weighting matrix. Under the standard regularity conditions, $\hat{\beta}_T$ is consistent and asymptotically normal.

In general, the moment function, $f_t(x, \beta)$ can be formed in two different ways. One way is to use an unconditional expectations relation, which sets

$$f_t(x, \beta) = g_t(x, \beta) - E^{\beta}(g_t(x, \beta)) \tag{6}$$

where $g_t = g(x_t, \ldots, x_{t-k}, \beta)$ is a known measurable function. Alternatively, we can use a conditional expectations relation, which sets

$$f_t(x, \beta) = g_t(x, \beta) - E^{\beta}(g_t(x, \beta) \mid x_{t-1})) \tag{7}$$

In both cases, the moment conditions in (4) for f are satisfied. One can see immediately that a complete knowledge of the functional form of $E^{\beta}(g_t(x, \beta))$ or $E^{\beta}(g_t(x, \beta) \mid x_{t-1})$ is essential for the minimization of (5) to be carried out. In general, it is very difficult to calculate the conditional moments of an arbitrary diffusion process.

AIV.3 Example : Two-Factor Model

We write the discrete-time analogue of the model :

(8) $$dx = k_1(\theta_1 - x)dt + \sigma_1 x^{\gamma_1} dz_1$$

$$dy = k_1(\theta_1 - y)dt + \sigma_1 y^{\gamma_1} dz_1$$

where x and y are correlated, $dz_1 dz_2 = \rho dt$

as

$$x_{t+1} = x_t + \tau(k_1(\theta_1 - x_{t-1})) + \varepsilon_{t+1}^{(x)}$$

$$y_{t+1} = y_t + \tau(k_1(\theta_1 - y_{t-1})) + \varepsilon_{t+1}^{(y)}$$

where $\tau = \frac{N}{T}$ is the interval between observation, and where

$$E(\varepsilon_{t+1}^{(x)}) = 0$$

$$E(\varepsilon_{t+1}^{(y)}) = 0$$

$$E((\varepsilon_{t+1}^{(x)})^2) = \sigma_1^2 x_t^{2\gamma_1} \tau$$

$$E((\varepsilon_{t+1}^{(y)})^2) = \sigma_1^2 y_t^{2\gamma_1} \tau$$

$$E(\varepsilon_{t+1}^{(x)} \varepsilon_{t+1}^{(y)}) = \rho \sigma_1 x_t^{\gamma_1} \sigma_2 y_t^{\gamma_2} \tau$$

To estimate this system by *GMM*, we construct for each observation $t = 1, 2, .. T$ the moment vector:

$$
(9) \qquad f_t(x, y, \beta) = \begin{bmatrix} \varepsilon_{t+1}^{(x)} \\ \varepsilon_{t+1}^{(x)} \\ \varepsilon_{t+1}^{(y)} x_t \\ \varepsilon_{t+1}^{(y)} y_t \\ \varepsilon_{t+1}^{(x)\,2} - \tau\sigma_1^2 x_t^{\gamma_1} \\ \varepsilon_{t+1}^{(y)\,2} - \tau\sigma_2^2 y_t^{\gamma_2} \\ (\ \varepsilon_{t+1}^{(x)\,2} - \tau\sigma_1^2 x_t^{\gamma_1}) x_t \\ (\ \varepsilon_{t+1}^{(y)\,2} - \tau\sigma_2^2 y_t^{\gamma_2}) y_t \\ \varepsilon_{t+1}^{(x)} \varepsilon_{t+1}^{(y)} - \rho\sigma_1\sigma_2 x_t^{\gamma_1} y_t^{\gamma_2} \end{bmatrix}
$$

Instead, we seek parameter estimates based on the conditional moment restriction vector :

$$
(10) \qquad f_t(x, y, \beta) = \begin{bmatrix} x_{t+1} - E_\beta(x_{t+1} | x_{t-m}, y_{t-m}) \\ y_{t+1} - E_\beta(y_{t+1} | x_{t-m}, y_{t-m}) \\ (x_{t+1} - E_\beta(x_{t+1} | x_{t-m}, y_{t-m})) x_t \\ (y_{t+1} - E_\beta(y_{t+1} | x_{t-m}, y_{t-m})) y_t \\ (x_{t+1} - E_\beta(x_{t+1} | x_{t-m}, y_{t-m}))^2 - Var_\beta(x_{t+1} | x_{t-m}, y_{t-m}) \\ (y_{t+1} - E_\beta(y_{t+1} | x_{t-m}, y_{t-m}))^2 - Var_\beta(y_{t+1} | x_{t-m}, y_{t-m}) \\ ((x_{t+1} - E_\beta(x_{t+1} | x_{t-m}, y_{t-m}))^2 - Var_\beta(x_{t+1} | x_{t-m}, y_{t-m})) x_t \\ ((y_{t+1} - E_\beta(y_{t+1} | x_{t-m}, y_{t-m}))^2 - Var_\beta(y_{t+1} | x_{t-m}, y_{t-m})) y_t \\ (x_{t+1} - E_\beta(x_{t+1} | x_{t-m}, y_{t-m}))(y_{t+1} - E_\beta(y_{t+1} | x_{t-m}, y_{t-m})) \\ - Cov_\beta(x_{t+1}, y_{t+1} | x_{t-m}, y_{t-m}) \end{bmatrix}
$$

where $E_\beta()$ denotes conditional expectation subject to the parameter vector β. $Var_\beta()$ and $Cov_\beta()$ are similarly defined.

The problem here is how to measure these conditional moment. In some cases, for univariate diffusion models, the exact form of the conditional moments is known[62]. However, in those involving multivariate models, the form of the moments is not known. He (1990) suggests that the conditional moments be estimated at each sample point by construction a binomial lattice with its origin at the current

62. For instance, the conditional moment for the square-root process is available in Oakes (1997) p. 33.

sample value for a univariate case. The probabilities along each path in the lattice are used to generate the conditional expectation which generate the components of the moment restriction vector at each sample point. However, He's method is not practical for multivariate case, since it is based on a non-recombining tree. To overcome the problem, Oakes (1997) suggests a method of a transformation of the original variable that has the property that its diffusion term is constant. Furthermore, to apply the method to a multivariate case, the original variables are orthogonalized using another transformation. We explain this in the next section.

AIV.4 Example : An Algorithm for the Two-Factor Model

Consider the following two-factor model

$$dx = k_1(\theta_1 - x)dt + \sigma_1 x^{\gamma_1} dz_1$$

$$dy = k_1(\theta_1 - y)dt + \sigma_1 y^{\gamma_1} dz_1$$

where z_1 and z_2 are correlated.

First, applying the transformation,

(11)
$$\phi_1 = x^{1-\gamma_1}$$
$$\phi_2 = y^{1-\gamma_2}$$

we obtain new variables which, by *Itô*'s lemma, satisfy :

(12)
$$d\phi_1 = \left[\frac{a}{\phi_1^{\frac{\gamma_1}{1-\gamma_1}}} - b\phi_1 - \frac{c}{\phi_1} \right] dt + (1-\gamma_1)dz_1$$
$$= q_1 dt + v_1 dz_1$$

$$d\psi_2 = \left[\frac{e}{\psi_2^{\frac{\gamma_2}{1-\gamma_2}}} - f\psi_2 - \frac{g}{\psi_2} \right] dt + (1-\gamma_2)dz_2$$

$$= q_2 dt + v_2 dz_1$$

where

$$a = (1-\gamma_1)k_1\theta_1, \quad e = (1-\gamma_2)k_2\theta_2$$

$$b = (1-\gamma_1)k_1, \quad f = (1-\gamma_2)k_2$$

$$c = \frac{\gamma_1(1-\gamma_1)\sigma_1^2}{2}, \quad g = \frac{\gamma_2(1-\gamma_2)\sigma_2^2}{2}$$

The reverse transformation is :

(13) $$x = \psi_1^{\frac{1}{1-\gamma_1}}, \quad y = \psi_2^{\frac{1}{1-\gamma_2}}$$

.

The diffusion terms v_1 and v_2 in (12) are constant. Second, these new variables are orthogonalized to remove the dependency between the variables.

(14) $$\delta_1 = v_2\psi_1 + v_1\psi_2$$

$$\delta_2 = v_2\psi_1 - v_1\psi_2$$

Itô's lemma gives

$$d\delta_1 = (v_2 q_1 + v_1 q_2)dt + v_1 v_2 (dz_1 + dz_2)$$

$$d\delta_2 = (v_2 q_1 - v_1 q_2)dt + v_1 v_2 (dz_1 - dz_2)$$

Since z_1 and z_2 are standard Brownian motions with correlation ρ, the implied covariance matrix for this system is :

$$V=\begin{pmatrix} 2(1+\rho)v_1^2v_2^2dt, & 0 \\ 0, & 2(1-\rho)v_1^2v_2^2dt \end{pmatrix}$$

which is decomposed into $V=B'B$, where:

$$B=\begin{pmatrix} \sqrt{2(1+\rho)dt}\ v_1v_2, & 0 \\ 0, & \sqrt{2(1-\rho)dt}\ v_1v_2 \end{pmatrix}$$

Combining two uncorrelated Brownian motions dz_1^* and dz_2^* with weights as in B gives:

(15)
$$d\delta_1=(v_2q_1+v_1q_2)dt+v_1v_2\sqrt{2(1+\rho)}dz_1^*$$
$$=m_1dt+s_1dz_1^*$$

$$d\delta_2=(v_2q_1-v_1q_2)dt+v_1v_2\sqrt{2(1-\rho)}dz_1^*$$
$$=m_2dt+s_2dz_2^*$$

The reverse transformation is:

(16)
$$\psi_1=\frac{\delta_1+\delta_2}{2v_2},\quad \psi_2=\frac{\delta_1-\delta_2}{2v_1}$$

We explain how to approximate the transformed variable δ_1 and δ_2 by the trinomial scheme.

AIV.5 A Grid Approximation

We summarize the method of Chen (1996) which is slightly more flexible than that of Oakes (1997). The transformed process (15) is approximated by a path independent grid consisting of n steps. We divide the time interval $[t_0, T]$ into n subintervals $[t_i, t_{i+1}]$ of equal length, where $t_{i+1} - t_i = \Delta t = \frac{T}{n}$, $i = \{0, \ldots, n\}$. A grid for δ_1 is constructed by taking δ_2 as fixed $\widehat{\delta_2}$. The value of δ_1 at each node at time t_0 is the current value of δ_1, denoted by δ_1^0. The values of δ_1 at other nodes have the form $\delta_1^j = \delta_1^0 + j\Delta\delta_1$, $j = -i, \ldots, i$, $i = 0, \ldots, n$, for some $\Delta\delta_1$. The relationship between $\Delta\delta_1$ and Δt is

$$\Delta\delta_1 = c\sqrt{\Delta t}$$

and c is a positive constant. The partition space (δ_1, t) forms a grid. Let (i, j) denotes each node of the grid which represents $\delta_1^j = \delta_1^0 + j\Delta\delta_1$ and $t_i = t_0 + i\Delta t$. Given any nodes of the lattice (i, j), the lattice can reach nodes $(i+1, j+1)$, $(i+1, j)$ and $(i+1, j-1)$ at time t_{i+1}. Then, matching the first and second moments, as in Hull and White (1990b), of the process for δ_1 at node (i, j), we find a set of probabilities $q_{j,j+1}(t_i)$, $q_{j,j}(t_i)$, and $q_{j,j-1}(t_i)$. These probabilities are chosen to match the instantaneous drift term and volatility function :

(17)
$$\sum_{l=-1}^{1} q_{j,j+l}\, \delta_1^{i+l} = \delta_1^j + m_\delta \Delta t$$

$$\sum_{l=-1}^{1} q_{j,j+l}\, (\delta_1^{i+l})^2 = (\delta_1^j + m_\delta \Delta t)^2 + s_1^2 \Delta t$$

where $m_\delta = m_1(\delta_i^j, t_i)$. The construction of the lattice is to find a solution to (17).

Hull and White (1990b) show that the following is a set of solutions[63] :

$$q_{j,j+1}(t_i) = \frac{m_\delta \Delta t}{2\Delta\delta_1} + \frac{s_1^2 \Delta t}{2(\Delta\delta_1)^2} + \frac{(m_\delta \Delta t)^2}{2(\Delta\delta_1)^2}$$

(18) $$q_{j,j}(t_i) = 1 - \frac{s_1^2 \Delta t}{(\Delta\delta_1)^2} - \frac{(m_\delta \Delta t)^2}{(\Delta\delta_1)^2},$$

$$q_{j,j-1}(t_i) = -\frac{m_\delta \Delta t}{2\Delta\delta_1} + \frac{s_1^2 \Delta t}{2(\Delta\delta_1)^2} + \frac{(m_\delta \Delta t)^2}{2(\Delta\delta_1)^2}$$

The same procedure can be applied to construct a trinomial lattice for δ_2, given a fixed value of δ_1. The set of probabilities is then used to calculate conditional moments for the original variables.

AIV.6 A Generalization of Transformation in a Multi-Factor Diffusion Model

In this section, we illustrate the application of the transformation of variables in AIV.4 to the multi-factor affine-type diffusion model. Assuming risk-adjusted processes, we express the multi-variate diffusion process as follows;

$$dx = B(x)\ dt + C(x)\ dz$$

where

$$dx = \begin{bmatrix} dx_1 \\ dx_2 \\ \cdot\cdot \\ \cdot\cdot \\ \cdot\cdot \\ dx_n \end{bmatrix},$$

63. More detail procedures and examples are available in Chen (1996).

$$B(x)=\begin{bmatrix} a_1(b_1-x_1) \\ a_2(b_2-x_2) \\ \cdot\cdot \\ \cdot\cdot \\ \cdot\cdot \\ a_n(b_n-x_n) \end{bmatrix},$$

$$C(x)=\begin{bmatrix} c_1\sqrt{x_1} & 0 & \cdots\cdots 0 \\ 0 & c_2\sqrt{x_2} & \cdots\cdots 0 \\ & \cdots\cdots\cdots\cdots\cdots\cdots & \\ & \cdots\cdots\cdots\cdots\cdots\cdots & \\ & \cdots\cdots\cdots\cdots\cdots\cdots & \\ 0 & 0 & \cdots c_n\sqrt{x_n} \end{bmatrix},$$

$$dz=\begin{bmatrix} dz_1 \\ dz_2 \\ \cdot\cdot \\ \cdot\cdot \\ \cdot\cdot \\ dz_n \end{bmatrix},$$

and

$$Var(dz)=\begin{bmatrix} 1 & \rho_{12} & \cdots & \rho_{1n} \\ \rho_{21} & 1 & \cdots\cdots\rho_{2k}\cdots & \rho_{2n} \\ & \cdots\cdots\cdots\cdots\cdots\cdots & & \\ & \cdots\cdots\cdots\cdots\cdots\cdots & & \\ & \cdots\cdots\cdots\cdots\cdots\cdots & & \\ \rho_{n1} & \rho_{n2.} & \cdots\cdots\cdots\cdots & 1 \end{bmatrix}$$

This specification is an example of just the multi-case representation of the equation (1). As in AIV 4, in approximating the diffusion model to the binomial or trinomial tree, we transform the original variables to the variables with constant volatility to make the tree recombining.

To this end, we follow the Nelson and Ramaswamy approach and transform the

original state variable into another variable, which makes the diffusion term constant. We briefly explain the principle of their transformation technique. Let x be the original state variable and y the transformed variable. We choose y to satisfy $y = \int^{x} \frac{dx}{c\sqrt{x}} = \frac{2\sqrt{x}}{c}$. Then, using *Itô*'s lemma, we find that the diffusion term of the transformed variable becomes constant unity.

In our multi-factor case, the same principle is applied. We choose a transform matrix such that the changed variable has the variance one. The appropriate matrix is,

$$T = 2 \begin{bmatrix} \frac{\sqrt{x_1}}{c_1} & 0 & \cdots & 0 \\ 0 & \frac{\sqrt{x_2}}{c_2} & & 0 \cdots 0 \\ & \cdots & \cdots & \\ & \cdots & \cdots & \\ & \cdots & \cdots & \\ 0 & 0 \cdots & \cdots & \frac{\sqrt{x_n}}{c_n} \end{bmatrix}$$

Then we transform $y = T\,x$, by *Itô*'s lemma,

$$dy = Dy(x, t)\ dt + I\ dz$$

where

$$Dy(x, t) = y_x(x, t)\ B(x) + \frac{1}{2} tr[\ y_{xx}(x, t) C(x) C(x)'] \ ,$$

and $tr(.)$ is the trace of the matrix, and I is n-dimensional identity matrix.

In recovering the original diffusion we use the reverse transformation. As stated above, the transformed diffusion term for dy is constant.

However, there is likely to be an instantaneous correlation ρ_{ij} between the Wiener process dz_i and dz_j. Hence, to simplify the approximation procedure, we transform y again to eliminate the correlation. We assume that the correlation is constant. The first transformed diffusion processes are given by

$$dy = Dy(x, t)\ dt + I\ dz$$

$$= \mu(y)\ dt + I\ dz$$

where

$$Var(dy) = \begin{bmatrix} 1 & \rho_{12} & \cdots & \rho_{1n} \\ \rho_{21} & 1 & \cdots\cdots \rho_{2k} \cdots & \rho_{2n} \\ & \cdots\cdots & \cdots\cdots & \\ & \cdots\cdots & \cdots\cdots & \\ & \cdots\cdots & \cdots\cdots & \\ \rho_{n1} & \rho_{n2} & \cdots\cdots\cdots\cdots & 1 \end{bmatrix}.$$

The covariance matrix of y is an nx n symmetric matrix. Therefore, we diagonalize y using the *Jordan canonical form.* Suppose the corresponding characteristic roots are $\lambda_1, \lambda_2, \ldots, \lambda_n$. Assume also that they are distinct. Then, the characteristic vectors of a symmetric matrix are orthogonal. Let $k_1, k_2, \ldots, k_n$ be characteristic vectors of the covariance matrix. Let K be the matrix whose i'th column is the k_i corresponding to λ_i ,

$$K = (k_1, k_2, \ldots, k_n)$$

and let Λ be the matrix whose diagonal is the n characteristic roots in the same order,

$$\Lambda = \begin{bmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & & 0 \cdots 0 \\ \cdots & \cdots & \cdots & \cdots \\ \cdots & \cdots & \cdots & \cdots \\ \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & \cdots & \lambda_n \end{bmatrix}$$

Accordingly, we can obtain the diagonalization of the covariance matrix by premultiplying and postmultipling $Var(dz)$ by K' and K:

$$K' \; Var(dz) \; K = \Lambda.$$

Hence, the second orthogonal transformation is given by

$$p = K' \; y.$$

Itô's lemma gives

$$dp = Dp(y, t) \; dt + p_y \; I \; dz$$

$$= \mu(p) \; dt + K' \; dz$$

where

$$Dp(y, t) = p_y(y, t)\mu(y) + \frac{1}{2} tr(p_{yy}I).$$

$$= K'\mu(y)$$

The implied variance-covariance matrix of dp is Λ. This means

$$dp = \mu(p) \; dt + \sqrt{\Lambda} \; dz.$$

In recovering the transformation, we use $x = (KT)^{-1}p$. Since the vector of transformed variables are orthogonal to each other, we can approximate the multi-factor diffusion by the trinomial method explained in Section AIV.5. The probability for any given node is the just product of the probabilities associated with the corresponding movements in each factor.

REFERENCES

Ahn, C.M. (1992), Option Pricing When Jump Risk Is Systematic, *Mathematical Finance*, 2, no 4, 299-308.

Ahn, C.M., and H.E. Thompson (1989), Jump-Diffusion Processes and the Term Structure of Interest Rates, *Journal of Finance*, XLIII, 155-174.

Ait-Sahalia, Y. (1996a), Testing Continuous Time Models of the Spot Interest Rate, *Review of Financial Studies*, 9, 385-426.

Ait-Sahalia, Y. (1996b), Nonparametric Pricing of Interest Rate Derivative Securities, *Econometrica*, 64, no 3, 527-560.

Amin, K.I. (1991), On the Computation of Continuous time Option Prices Using Discrete Approximations, *Journal of Financial and Quantitative Analysis,* 26, 477-495.

Armstrong, M.A. (1983), *Basic Topology*, Springer-Verlag.

Babbs, S. (1991), A Family of Ito Process Models for the Term Structure of Interest Rates, *FORC*, 91/24.

Babbs, S., and K.B. Nowman (1997), Kalman Filtering of Generalized Vasicek Term Structure Models, *FORC*, 97/80.

Babbs, S., and M.J.P. Selby (1996), Pricing by Arbitrage in Incomplete Markets, *FORC,* 96/69.

Babbs, S., and N. Webber (1994), A Theory of the Term Structure with an Official Short Rate, *FORC,* 94/49.

Babbs, S., and N. Webber (1995), Term Structure Modelling Under Alternative Official Regimes, *FORC,* 95/61.

Back, K. (1991), Asset Pricing for General Processes, *Journal of Mathematical Economics,* 20. 371-395.

Balduzzi, P., G. Bertola, and S. Foresi (1993), A Model of Target Changes and the Term Structure of Interest Rates, *NBER Working paper,* 4347.

Ball, C.A., and W.N. Torous (1983a), Bond Price Dynamics and Options, *Journal of Financial and Quantitative Analysis* 19, 517-531.

Ball, C.A., and W.N. Torous (1983b), A Simplified Jump Process for Common Stock Returns, *Journal of Financial and Quantitative Analysis,* 18, 53-65.

Ball, C.A., and W.N. Torous (1985), On Jumps in Common Stock Prices and Their Impact on Call Option Pricing, *Journal of Finance*, XL, 155-173.

Bakshi, G. S., and Z. Chen (1996), Inflation, Asset Prices, and the Term Structure of Interest Rates in Monetary Economics, *Review of Financial Studies*, 9, 241-275.

Basak, S. (1995), A General Equilibrium Model of Portfolio Insurance, *Review of Financial Studies*, 8, 1059-1090.

Bates, D.S. (1988), Pricing Options under Jump-Diffusion Processes, Wharton School, University of Pennsylvania.

Bates, D.S. (1996), Jumps and Stochastic Volatility : Exchange Rate Processes

Implicit in Deutsche Mark Options, *Review of Financial Studies,* 9, 69-107.

Baxter M., and A. Rennie (1997), *Financial Calculus*, Cambridge University Press.

Becker, S. (1981), A Note on Estimating the Parameters of the Diffusion-Jump Model of Stock Returns, *Journal of Financial and Quantitative Analysis*, XVI, 127-139.

Begg, D. (1982), *The Rational Expectations and the Revolution in Macroeconomics*, Oxford Press.

Bemaud, P. (1981), *Point Process and Queues: Martingale Dynamics*, Springer-Verlag.

Bertola, G., and R. Caballero (1992), Target Zones and Realignments, *The American Economic Review*, 82, 520-536.

Billingsley, P. (1995), *Probability and Measure*, Wiley

Black, F., E. Derman and W. Toy (1990), A One-Factor Model of Interest Rates and Its Application to Treasury Bond Options, *Financial Analysts Journal*, 33-39.

Black, F., and P. Karasinski (1991), Bond and Option Pricing When Short Rates Are Lognormal, *Financial Analysts Journal*, 52-59.

Brace, A., and M. Musiela (1994), A Multifactor Gauss Markov Implementation of Heath, Jarrow, and Morton, *Mathematical Finance*, 4, no 3, 259-283.

Brace, A., D. Gatarek, and M. Musiela (1997), The Market Model of Interest Rate Dynamics, Mathematical Finance, 7, no 2, 127-155.

Breeden, D.T. (1979), An Intertemporal Asset Pricing Model with Stochastic

Consumption and Investment Opportunities, *Journal of Financial Economics,* 7, 265-296.

Breeden, D.T. (1986), Consumption, Production, Inflation and Interest Rates, *Journal of Financial Economics,* 16, 3-39.

Brennan, M.J., and E. S. Schwartz (1979). A Continuous Time Approach to the Pricing of Bonds, *Journal of Banking and Finance*, 3, 133-155.

Brennan, M.J., and E. S. Schwartz (1982). An Equilibrium Model of Bond Pricing and a Test of Market Efficiency, *Journal of Financial and Quantitative Analysis*, 17, 75-100.

Brenner R.J., R.H. Harjes, and K.F. Kroner (1996), Another Look at Models of the Short-Term Interest Rate, *Journal of Financial and Quantitative Analysis*, 31, 85-107.

Brown, B.M., and J.I. Hewitt (1975), Asymptotic Likelihood Theory for Diffusion Processes, *Journal of Applied Probability,* 12, 228-238.

Brown, R.H., and M.S. Schaeffer (1991), Interest Rate Volatility and the Term Structure, Paper present at the Second International Conference of the Centre for Research in Finance, IMI Group, Rome.

Brown, R.H., and M.S. Schaeffer (1994a), The Term Structure of Real Interest Rates and the CIR Model, *Journal of Financial Economics*, 3-42.

Brown, R.H., and M.S. Schaeffer (1994b), Interest Rate Volatility and the Shape of the Term Structure, *Phil. Trans. R.Soc. Lond.*, 347, 563-576.

Brown, S.J., and P.H. Dybvig (1986), The Empirical Implications of the CIR Theory of the Term Structure of Interest Rates, *Journal of Finance*, XLI, 617-632

Campbell, J.Y. (1986), A Defense of Traditional Hypotheses about the Term Structure of Interest Rates, *Journal of Finance*, 183-193.

Campbell, J.Y., A.W. Lo, and A.C. MacKinlay (1997), *The Econometrics of Financial Markets*, Princeton University Press.

Carr, P.P., and R.A. Jarrow (1995), A Discrete Time Synthesis of Derivative Security Valuation Using a Term Structure of Futures Prices, *Handbooks in OR and MS*, Vol 9, Elsevier Science.

Carverhill, A., and C. Stricland (1992), Money Market Term Structure Dynamics and Volatility Expectations, *FORC*.

Carverhill, A. (1994), When Is the Short Rate Markovian ?, *Mathematical Finance*, 4, 305-312.

Chan, K.C., A. Karolyi, A.F. Longstaff, and A.B. Sanders (1992), An Empirical Comparison of Alternative models of Short-Term Interest Rate, *Journal of Finance*, no 3, 1209-27.

Chen, L. (1994), Stochastic Mean and Stochastic Volatility: A Three-Factor Model of the Term Structure and its Application in Pricing of Interest Rate Derivatives, Federal Reserve Board

Chen, L. (1996), *Interest Rate Dynamics, Derivatives Pricing, and Risk Management*, Springer.

Chen, R.R. (1992), Exact Solutions for Futures and European Futures Options on Pure Discount Bonds, Journal of Financial and Quantitative Analysis, 27, 97-107.

Chen, R.R., and L. Scott (1992), Pricing Interest Rate Options in a Two-Factor Cox-Ingersoll-Ross Model of the Term Structure, *The Review of Financial Studies*, 5, 613-636.

Chung, K.L., and R. Williams (1993), *An Introduction to Stochastic Integration*, Birkhauser.

Colwell D.B., and R.J. Elliot (1993), Discontinuous Asset Prices and Non-Attainable Contingent Claims, *Mathematical Finance*, no 3, 295-308.

Constantinides, G.M. (1992), A Theory of Nominal Term Structure of Interest Rates, *Review of Financial Studies,* 5, 531-552.

Cox, J.C., and C. Huang (1989), Optimal Consumption and Portfolio Process When Asset Prices Follow a Diffusion Process, *Journal of Economic Theory,* 49, 33-83.

Cox, J.C., J.E. Ingersoll, and S.A. Ross (1981), A Re-Examination of Traditional Hypotheses about the Term Structure of Interest Rates, *Journal of Finance*, 769-799

Cox, J.C., J.E. Ingersoll, and S.A. Ross (1985a), An Intertemporal General Equilibrium Model of Asset Prices, *Econometrica*, 53, 363-384

Cox, J.C., J.E. Ingersoll, and S.A. Ross (1985b), A Theory of the Term Structure of Interest Rates, *Econometrica*, 53, 385-407.

Cox, D.R., and H.D. Miller (1965), *The Theory of Stochastic Processes*, Chapman and Hall.

Cox, J.C., and S.A. Ross (1976), The Valuation of Options for Alternative Stochastic Processes, *Journal of Financial Economics,* 3, 145-166.

Dahlquist, M. (1996), On alternative Interest Rate Processes, *Journal of Banking and Finance,* 20, 1091-1119.

Das S.R., and S. Foresi (1996), Exact Solutions for Bond and Option Prices with Systematic Jump Risk, *Review of Derivatives Research,* 1, 7-24.

Dothan, M.U. (1990), *Prices in Financial Market*, Oxford Press.

Duan, J.C. (1997), Augmented GARCH(p,q) Process and Its Diffusion Limit, *Journal of Econometrics*, 79, 97-127.

Diffie, D. (1988), *Security Markets : Stochastic Models,* Academic Press.

Diffie, D. (1992), *Dynamics Asset Pricing Theory*, Princeton Press.

Duffie, D., and L.G. Epstein (1992), Stochastic Differential Utility, *Econometrica*, 60, no 2, 353-394.

Duffie, D., and L.G. Epstein (1992), Asset Pricing with Stochastic Differential Utility, *Review of Financial Studies*, 5, no 3, 411-436.

Duffie D., and R. Kan (1996), A Yield Factor Model of Interest Rates, *Mathematical Finance*, 379-406.

Duffie, D., J. Ma, and J. Yong (1995), Black's Consol Rate Conjecture, *Annals of Applied Probability*, 356-382.

Duffie, D., and P.L. Lions (1992), PDE Solution of Stochastic Differential Utility, *Journal of Mathematical Economics,* 21, 577-606.

Duffie, D., and K. Singleton (1993), Simulated Moments Estimation of Markov Models of Asset Prices, *Econometrica*, 61, 929-952.

Dybvig, P.H., J.E. Ingersoll, and S.A. Ross (1989), Long Forward Rates Can Never Fall, Yale School of Management.

Enders, W. (1995), *Applied Econometric Time Series*, John Wiley & Sons.

Engle, R.F. (1982), Autoregressive Conditional Heteroscedasticity with Estimates of the Variance of United Kingdom Inflation, *Econometrica*, 50, 987-1007.

Engle, R.F., M.L. David, and P.R. Russel (1987), Estimating Time Varying Risk Premia in the Term Structure: The ARCH-M Model, *Econometrica*, 55, 391-407.

Fama, E.F. (1976), Forward Rates as Predications of Future Spot Rates, *Journal of Financial Economics*, 3, 361-377.

Figlewski, S., W.L. Silber, and M.G. Subrahmanyam (1990), *Financial Options : from Theory to Practice*, New York University Salomon Center.

Flesaker, B., and L.P. Hughston (1995), Dynamics Models for Yield Curve Evolution : Theory and Practice, Working Paper.

Flesaker, B., and L.P. Hughston (1996), Positive Interest Rate, *RISK*, 9, no 1, 46-49.

Flesaker, B., and L.P. Hughston (1997), International Models for Interest Rates and Foreign Exchange, Working Paper.

Fong, H.G., and A.V. Vasicek (1991), Interest Rate Volatility as a Stochastic Factor, *Gifford Fong Associates*.

Frachot, A., D. Janci, and V. Lacoste (1992), Factor Analysis of the Term Structure :

A Probabilistic Approach, Working Paper, Banque de France.

Froot, K.A., and M. Obstfeld (1991), Exchange-Rate Dynamics Under Stochastic Regimes Shifts, *Journal of International Economics,* 31, 203-229.

Geman, H., N. Karoui, and J. Rochet (1995), Changes of Numeraire, Changes of Probability Measure and Option Pricing, *Journal of Applied Probability* 32, 443-458.

Gibbons, M. R., and K. Ramaswamy (1993), A Test of the Cox, Ingersoll and Ross Model of the Term Structure, *Review of Financial Studies*, 6, 619-658.

Girsanov, I.V. (1960), On Transforming a Certain Class of Stochastic Processes by Absolutely Continuous Substitution of Measures, *Theory of Probability and Its Applications*, 5, no 3, 285-310.

Goldstein, R., and F. Zapatero (1996), General Equilibrium with Constant Risk Aversion and Vasicek Interest Rates, *Mathematical Finance*, 6, no 3,331-340.

Granger, C. W. J., and P. Newbold (1974), Spurious Regressions in Econometrics, *Journal of Econometrics*, 2, 111-120.

Hamilton, J.D. (1988), Rational Expectations Econometric Analysis of Changes in Regime: An Investigation of the Term Structure of Interest Rates, *Journal of Economic Dynamics and Control*, 12, 385-423.

Hamilton, J.D. (1989), A New Approach to the Economic Analysis of Nonstationary Time Series and the Business Cycle, *Econometrica*, 57, 357-384.

Hamilton, J.D. (1990), Analysis of Time Series Subject to Changes in Regime, *Journal of Econometrics*, 45, 39-70.

Hamilton, J.D. (1994), *Time Series Analysis*, Princeton University Press.

Hansen, L.P., and J.A. Schenkman (1995), Back to the Future : Generating Moment Implications for Continuous-time Markov Processes, *Econométrica*, 63, no 4, 767-804.

Harrison, J.M., and D.M. Kreps (1979), Martingale and Arbitrage in Multi-Period Securities Markets, *Journal of Economic Theory*, 20, 381-408

Harrison, J.M., and S.R. Pliska (1981), Martingale and Stochastic integrals in the Theory of Continuous trading, *Stochastic Process and Their Application,* 11, 215-260.

Harvey, A. (1990), *The Econometric Analysis of Time Series*, LSE Handbooks in Economics.

He, H. (1989), Convergence from Discrete to Continuous Time Financial Models, Haas School of Business, University of California at Berkeley.

He, H. (1990), Moment Approximation and Estimation of Diffusion Models of Asset Prices, Haas School of Business, University of California, at Berkeley.

He, H., and H. Leland (1993), On Equilibrium Asset Price Process, *Review of Financial Studies*, 6, 593-617.

Heath, D., R. Jarrow, and A. Morton (1990), Bond Pricing and the Term Structure of The Interest Rates: A Discrete Time Approximation, *Journal of Financial and Quantitative Analysis*, 4, 419-440.

Heath D., R. Jarrow, and A. Morton (1992), Bond Pricing and the Term Structure of Interest Rates : A New Methodology for Contingent Claim Valuation, *Econometrica,* 60, 77-105.

Heston, S. (1988), "Testing Continuous Time Models of GMM Estimators", Working paper, Carnegie Mellon University.

Heston, S. (1993), A Closed-Form Solution of Options with Stochastic Volatility with Applications to Bond and Currency Options, *Review of Financial Studies*, 6, no 2, 327-343.

Ho, T.S., and S.B. Lee (1986), Term Structure Movements and Pricing Interest Rate Contingent Claims, *Journal of Finance*, 41, 1011-1029.

Hogan, M. (1993), Problems in Certain Two-Factor Term Structure Models, *The Annals of Applied Probability*, 3, 576-581.

Huang, C. (1985), Information Structures and Viable Price Systems, *Journal of Financial Economics*, 215-239.

Huang, C. (1987), An Intertemporal General Equilibrium Asset Pricing Model : The Case of Diffusion Information, *Econometrica*, 55, no 1, 117-142.

Huang, C., and R. Litzenberger (1988), *Foundations of Financial Economics,* North-Holland.

Hughston, L.P. (1996), Vasicek and Beyond, Approaches to Building and Applying Interest Rate Models, *RISK*.

Hull, J. (1993), *Options, Futures, and Other derivatives Securities*, 2nd, Prentice-Hall.

Hull, J., and A. White (1990a), Pricing Interest Rate Derivative Securities, *Review of Financial Studies*, 3, 573-92.

Hull, J., and A. White (1990b), Valuing Derivative Securities Using the Explicit Finite

Difference Method, *Journal of Financial and Quantitative Analysis*, 25, no 1, 87-100

Hull, J., and A. White (1993), One Factor Interest Rate Models and the Valuation of Interest Rate Derivative Securities, *Journal of Financial and Quantitative Analysis*, 28, 235-254.

Hull, J., and A. White (1994a), Numerical Procedures for Implementing Term Structure Model I: Single-Factor Models, *Journal of Derivatives*, 7-16.

Hull, J., and A. White (1994b), Numerical Procedures for Implementing Term Structure Model II: Two-Factor Models, *Journal of Derivatives*, 37-48.

Jamshidian, F. (1989), An Exact Bond Option Formula, *Journal of Finance*, XLIV, 205-209.

Jamshidian, F. (1991), Bond and Option Evaluation in the Gaussian Interest Rate Model, *Research in Finance*, 9, 131-170.

Jamshidian, F. (1992), A Simple Class of Square-Root Interest Rate Models, Working Paper.

Jamshidian, F. (1993), Option and Future Evaluation with Deterministic Volatilities, *Mathematical Finance*, 3, no 2, 149-159.

Jarrow, R. (1995), Pricing Interest Rate Options, *Handbooks in OR and MS*, Vol 9, Elsevier Science.

Jeffrey, A. (1994), Construction of a Single Factor Heath-Jarrow-Morton Term Structure Model, School of Banking and Finance, The University of New South Wales.

Karouri, N.E., and V. Lacoste (1992), Multifactor Models of the Term Structure of Interest Rate, AFFI.

Kennedy, D.P. (1995), The Term Structure of Interest Rates as a Gaussian Random Field, *Mathematical Finance*, 4, no 3, 247-258.

Keifer, N.M. (1978), Discrete Parameter Variation :Efficient Estimation of a Switching Regression Model, *Econometrica*, 46, 427-434.

Krugman, P., and M. Miller (1992), *Exchange Rate Targets and Currency Bands*, Centre for Economic Policy Research, NBER.

Langetieg, T.C. (1980), A Mutivariate Model of the Term Structure, *Journal of Finance*, XXXV, 71-97.

Litterman, R., J. Scheinkman, and L. Weiss (1988), Volatility and the Yield Curve, Financial Strategies Group, Goldman Sacks.

Longstaff, F.A., and E.S. Schwartz (1992), Interest Rate Volatility and the Term Structure: A Two-Factor General Equilibrium Model, *The Journal of Finance*, XLVII, 1259-1282.

Lucas, R.E. (1978), Asset Prices in an Exchange Economy, *Econometrica*, 46, no 6, 1429-1445.

Madan, D.B. (1988), Risk Measurement in Semimartingale Models with Multiple Consumption Goods, *Journal of Economic Theory,* 44, 398-412.

Malliaris, A.G., and W.A. Brock (1984), *Stochastic Methods in Economics and Finance,* North-Holland.

Marsh, T.A. (1995), Term Structure of Interest Rates and the pricing of Fixed

Income Claims and Bonds, *Handbooks in OR and MS*, vol 9, Elsevier Science.

Marsh, T.A., and E.R. Rosenfeld (1983), Stochastic Processes for Interest Rates and Equilibrium Bond Prices, *Journal of Finance,* 38, 635-646.

Mclachlan, N.W. (1950), *Ordinary Non-Linear Differential Equations in Engineering and Physical Science*, Oxford.

Mercurio, F., and W.J. Runggaldier (1993), Option Pricing for Jump Diffusions : Approximations and Their Interpretation, *Mathematical Finance*, 3, no 2, 191-200.

Merton, R.C. (1971), Optimum Consumption and Portfolio Rules in a Continuous Time Model, *Journal of Economic Theory*, 3, 373-413.

Merton, R.C. (1976), Option Pricing When Underlying Stock Returns Are Discontinuous, *Journal of Financial Economics,* 3, 125-144.

Merton, R.C. (1993), *Continuous-Time Finance*, Blackwell.

Miltersen, K.R. (1994), An Arbitrage Theory of the Term Structure of Interest Rate, *Annals of Applied Probability*, 3, 953-967.

Modigliani, F., and R.J. Shiller (1973), Inflation, Rational Expectations and the Term Structure of Interest Rates, *Economica,* 12-43.

Ncube, M. (1994), Heterogeneous Discrete Information, Diffusion-Jump Processes and Pricing Derivative Securities, London School of Economics.

Nelson, D.B., and K. Ramaswamy (1989), Simple Binomial Processes as Diffusion Approximations in Financial Models, *Review of Financial Studies*, 3, no, 3, 393-430.

Nelson, D.B. (1990), ARCH Models as Diffusion Approximations, *Journal of Econometrics*, 45 7-38.

Nunes, J.P. (1997), A General Equilibrium Duffie and Kan Term Structure Model, Unpublished paper.

Oakes, D. (1997), A Moment Approximation Method for Parameter Estimation in Multivariate Diffusion Models, Warwick Business School..

Obstfeld, M., and K. Rogoff (1996), *Foundations of International Macroeconomics,* MIT Press.

Oldfield, G.S., and R. Rogalski (1987), The Stochastic Properties of Term Structure Movements, *Journal of Monetary Economics,* 19, 229-254.

Oksendal, B. (1995), *Stochastic Differential Equations*, Springer-Verlag.

Pang, K., and S. Hodges (1995), Non-negative Affine Yield Models of the Term Structure, *FORC,* 95/62.

Pappalarado, L. (1996), Option Pricing and Smile Effect When Underlying Stock Prices are Driven by a Jump Process, *FORC,* 96/70.

Pearson, N.D., and T.S. Sun (1991), An Empirical Examination of the Cox, Ingersoll and Ross Model of the Term Structure of Interest Rate, Working paper, Graduate School of Management, University of Rochester.

Protter, P. (1996), Stochastic Integration and Differential Equations : A New Approach, Springer-Verlag

Revuz, D., and M. Yor (1991), *Continuous Martingales and Brownian Motion,*

Springer-Verlag, New York.

Richard, S.F. (1978), An Arbitrage Model of the Term Structure of Interest Rates, *Journal of Financial Economic,* 6, 33-57.

Rubinstein, M. (1976), The Valuation of Uncertain Income Streams and the Pricing of Options, *Bell Journal of Economics*, 7, 407-425.

Samuelson, P.A. (1965), Proof that Property Anticipated Prices Fluctuate Randomly, *Industrial Management Review*, 6, Spring.

Samuelson, P.A. (1969), Lifetime Portfolio Selection by Dynamic Stochastic Programming, *Review of Economics and Statistics*, 51, 239-246.

Sargent, T.J. (1972), Rational Expectations and the Term Structure of Interest Rates, *Journal of Money, Credit and Banking*, 74-97.

Sargent, T.J. (1979), A Note on Maximum Likelihood Estimation of the Rational Expectation Model of the Term Structure, *Journal of Monetary Economics,* 5, 133-143.

Sargent, T.J. (1987), *Dynamic Macroeconomic Theory*, Harvard University Press.

Sargent, T.J., and N. Wallace (1973), Rational Expectations and the Dynamics of Hyperinflation, *International Economic Review*, 14, no. 2, 328-350.

Schaefer, S.M., and E.S. Schwartz (1984), A Two-Factor Model of the Term Structure: An Approximate Analytical Solution, *Journal of Financial and Quantitative Analysis,* 19, no. 4, 413-424.

Schaefer, S.M., and E.S. Schwartz (1987), Time Dependent Variance and The Pricing of Bond Options, *Journal of Finance*, 42, 1113-28.

Schweizer, M. (1992), Martingale Densities for General Asset Prices, *Journal of Mathematical Economics,* 21, 363-378

Scott, L.O. (1993), Pricing Stock Options in a Jump-Diffusion Model with Stochastic Volatility and Interest Rates : Application of Fourier Inversion Methods, University of Georgia.

Selby, M.J.P., and C. Strickland (1993), Computing the Fong and Vasicek Pure Discount Bond Price Formula, *FORC.*

Shiller, R.J. (1979), The Volatility Structure of the Long-Term Interest Rates and Expectations Models of the Term Structure, *Journal of Political Economy,* 87, 1190-1219.

Shiller, R.J. (1981), Alternative Tests of Rational Expectations Models : The Case of the Term Structure, *Journal of Econometrics,* 16, 71-87.

Shiller, R.J. (1989), *Market Volatility*, MIT Press.

Shirakawa, H. (1991), Interest Rate Option Pricing with Poisson-Gaussian Forward Rate Curve Processes, *Mathematical Finance*, 4, 77-94.

Singleton, K.J. (1980), Expectations Models of the Term Structure and Implied Variance Bounds, *Journal of Political Economy,* 88, 1159-1176.

Smith, G.W. (1991), Solution to a Problem of Stochastic Switching, *Econometrica,* 59, no 1, 237-239.

Stambaugh, R.F. (1988), The Information in Forward Rates : Implications for Model of the Term Structure, *Journal of Financial Economics,* 21, 41-70.

Steeley, J.M. (1990), Modelling the Dynamics of the Term Structure of Interest Rates, *FORC.*

Stiglitz, J.E. (1970), A Consumption-Oriented Theory of the Demand for Financial Assets and the Term Structure of Interest Rates, *Review of Economic Studies* 37, 321-351.

Strickland, C. (1993), Interest Rate Volatility and the Term Structure of Interest Rates, *FORC,* 93/37..

Strickland, C. (1994), A Comparison of Models for Pricing Interest Rate derivative Securities, *FORC,* 94/47.

Sun, T.S. (1992), Real and Nominal Interest Rates: A Discrete-Time Model and Its Continuous-Time Limit, *Review of Financial Studies,* 5, 581-611.

Sutherland, W.A. (1995), *Introduction to Metric Space and Topological Spaces,* Oxford Science Publications.

Tice, J., and N. Webber (1997), A Non-Linear Model of the Term Structure of Interest Rates, *Mathematical Finance* (forthcoming).

Vasicek, O.A. (1977), An Equilibrium Characterization of Term Structure, *Journal of Financial Economics*, 5, 177-88.

Veruete, L., and N. Webber (1994), A Model of UK LIBOR as a Jump-Diffusion Process, *FORC,* 94/48.