A generalized one-factor term structure model and pricing of interest rate derivative securities

Jiang, George J.

Publication date:
1997

Citation for published version (APA):
Jiang, G. J. (1997). *A generalized one-factor term structure model and pricing of interest rate derivative securities*. s.n.

A Generalized One-Factor Term Structure Model and Pricing of Interest Rate Derivative Securities

George J. Jiang [1]

May, 1997

Abstract

The purpose of this paper is to propose a nonparametric interest rate term structure model and investigate its implications on term structure dynamics and prices of interest rate derivative securities. The nonparametric spot interest rate process is estimated from the observed short-term interest rates following a robust estimation procedure and the market price of interest rate risk is estimated as implied from the historical term structure data. That is, instead of imposing *a priori* restrictions on the model, data are allowed to speak for themselves, and at the same time the model retains a parsimonious structure and the computational tractability. The model is implemented using historical Canadian interest rate term structure data. The parametric models with closed form solutions for bond and bond option prices, namely the Vasicek (1977) and CIR (1985) models, are also estimated for comparison purpose. The empirical results not only provide strong evidence that the traditional spot interest rate models and market prices of interest rate risk are severely misspecified but also suggest that different model specifications have significant impact on term structure dynamics and prices of interest rate derivative securities.

Keywords: Interest Rate Term Structure, Nonparametric Estimation, Pricing of Derivative Securities, Numerical Solution of PDE, Monte Carlo Simulations

1   Department of Econometrics, University of Groningen, P.O. Box 800, 9700 AV Groningen, The Netherlands, e-mail: G.Jiang@eco.rug.nl

1. Introduction

The framework of modeling term structure dynamics of interest rates in modern continuous-time finance theory is to postulate the whole yield curve as determined by a small set of state variables (factors) which are assumed to follow diffusion processes.[2] Dynamics of term structure and prices of interest rate derivative securities can then be projected from either solving (analytically or numerically) a partial differential equation (PDE) which is derived through Itô's lemma following a no-arbitrage argument or in a general equilibrium framework, or performing Monte Carlo simulations along a risk-neutral process of the underlying state variables.

Over the past few decades, theoretical development of modeling term structure dynamics has been mainly along the following two directions. One direction is, while keeping a simple, tractable, and parsimonious structure, to extend the model through more flexible specification in order to better describe the dynamics of state variables and project the term structure movements. Development along this direction is evidenced in various one-factor models.[3] Extension of the one-factor model reflects both the desire to incorporate nonlinearity in the spot rate process and to avoid the difficulty involved in solving high-dimensional PDEs. However, one-factor models have been criticized for: first, it implies perfect correlation of the local price move-

---

2   See Vetzal (1994) for a comprehensive survey of parametric continuous-time interest rate term structure models and also Chan, et al (1992) for an empirical comparison of various parametric interest rate term structure models.

3   For example, the parametric models by Merton (1973), Cox (1975), Vasicek (1977), Dothan (1978), Courtadon (1982), Marsh and Rosenfeld (1983), Cox, Ingersoll and Ross (1985) (hereafter CIR), Brown and Schaefer (1991), Chan, et al (1992), Constantinides (1992), and Duffie and Kan (1993), the semi-parametric model by Aït-Sahalia (1996), and the nonparametric models by Jiang and Knight (1995) and Stanton (1996).

1

ments of bonds of all maturities, and second, the implied yield curves are constrained in terms of their shapes due to particular functional specifications of the model. Hence the one-factor models sometimes only provide a poor fit to the observed term structure. In order to better model the term structure dynamics, several authors have extended the model along a different direction by including more state variables in the term structure representation. Development along this direction can be seen from many multiple-factor models.[4] But the gains of more generality are achieved at the cost of greater complexity which is reflected in the general lack of analytic solutions for the valuation PDE. It appears that those assume Ornstein-Uhlenbeck processes for all factors or unrealistically assume stochastic independence among factors are the only models which have closed form solutions. The difficulty of solving higher than two-dimensional PDEs has prevented general multi-factor models from being implemented.

Goes even further along the second direction are the non-Markov and time-inhomogeneous models which are designed to perfectly replicate the current term structure.[5] The non-Markov model due to Heath, Jarrow and Morton (1992) has the advantage that it can fit forward rate volatilities at all times, but this advantage is achieved at considerable cost, making the implementation of such model a formidable task. The trick used in the time-inhomogeneous models is to specify the parameters of

---

4   Examples are the Brennan and Schwartz (1979), Langetieg (1980), Schaefer and Schwartz (1984), CIR (1985), Chen and Scott (1992), Longstaff and Schwartz (1992), Chaplin and Sharp (1993), and Kraus and Smith (1993).

5   Examples are the models by Ho and Lee (1986) (a discrete time model), Heath, Jarrow, and Morton (1992), Black, Derman and Toy (1990), Hull and White (1990, 1993), Jamshidian (1990, 1991), and Black and Karasinski (1991).

2

stochastic processes as time-dependent, which can be adjusted to fit the current term structure as accurately as desired. Although the time-inhomogeneous models are often used by practitioners, they have been proved to fail providing a consistent fundamental model for the future (out-of-sample) behavior of interest rates and term structure. They are also criticized for ignoring the evidence that there are persistent arbitrage opportunities present in the observed term structure of interest rates. By re-estimating the model every day in order to retain the exact fit to the current yield curve, the model is prone to undermining the fundamental arbitrage-free assumptions and misprice interest rate options (see e.g. Backus, Foresi and Zin, 1995, and Canabarro, 1995). On the contrary, models which do not take the *entire* yield curve as given but are based on the no-arbitrage argument have a potential advantage of detecting such arbitrage opportunities. Moreover, since models of any kind have to be estimated from the sampling observations of the stochastic variables, never their populations, a procedure which promises to fit exactly into input data is liable to be severely misleading.

Contribution of this paper is along the combination of aforementioned both directions: First, we extend the spot rate process through nonparametric specifications of both the drift function and diffusion function to better model the dynamics of spot rates; Second, we specify the market price of interest rate risk as an implied nonparametric function so that the model generated term structure has the best fit into the historical term structures. In other words, instead of imposing *a priori* restrictive functional forms for the drift function, the diffusion function, and the market price

3

of risk, nonparametric estimation allows data to speak for themselves. The model precludes arbitrage opportunities, preserves a simple structure and the computational tractability, and at the same time allows for maximal flexibility in fitting into the data.

The paper is organized as follows. Section 2 outlines the spot rate approach of modeling term structure dynamics; Section 3 summarizes two well known one-factor models, i.e., the Vasicek (1977) model and the CIR (1985) model, and examines the behavior of these models and their closed form solutions for bond and bond option prices; In Section 4, consistent estimators of the nonparametric drift function, diffusion function and market price of risk are proposed. Procedures to obtain nonparametric prices of interest rate derivative securities by either solving the PDE numerically or performing Monte Carlo simulations along the risk-neutral process are proposed as well. In Section 5, the nonparametric model is implemented using historical Canadian interest rate term structure data. Empirical results not only provide strong evidence that the traditional spot interest rate models and market price of interest rate risk are misspecified but also suggest that different model specifications have significant impact on the term structure dynamics and prices of interest rate derivative securities. A brief conclusion is contained in Section 6.

2. Spot Rate Term Structure Model and Pricing of Derivative Securities

The spot interest rate term structure modeling approach assumes that spot interest rates are sufficient statistics for the stochastic movement of current term structure, and therefore the prices of interest rate derivative securities can be derived in terms of

4

the spot interest rates. Although this framework can allow in principle for an arbitrary finite number of state variables, in practice the number of factors is usually restricted to one [6], i.e., the short-term interest rate $r(t)$. The basic assumptions on the market in a one-factor model can be summarized as: (a) the spot interest rate follows a diffusion process; (b) the price of a pure discount bond depends only on the spot rate over its term; and (c) the market is efficient.

2.1. Term Structure Model: The Spot Rate Process and Market Price of Risk

Consider a continuous trading market with no taxes, transactions costs, or short sale constraints, uncertainty in this economy is represented by the complete filtered probability space $(\Omega, F, \{F_t\}, Q)$, where Ω is the sample space, F is the σ-algebra of measurable events, $\{F_t\}$ is a right-continuous filtration $\{F_t, t \geq 0\}$ generated by a standard Brownian motion in R, and Q is a probability measure. The dynamics of the spot interest rate process is assumed to be represented by the following time-homogeneous stochastic differential equation (SDE):

$$dr(t) = \mu(r(t))dt + \sigma(r(t))dW(t) \tag{1}$$

with initial condition $r(0) = r_0$, where $\mu(\cdot)$ and $\sigma^2(\cdot)$ are respectively the instantaneous mean and variance of the process, and $W(t)$ is the standard Brownian motion or Wiener process. In traditional spot interest rate models, $\mu(\cdot)$ and $\sigma^2(\cdot)$ are specified

---

6   As Vetzal (1994) pointed out, the general problem with multi-factor models lies not in their construction but in their implementation. It is very difficult to solve the valuation PDE when there are more than two state variables. Strictly speaking, the problem in the case of the term structure is excessive computer time. The prices of interest rate derivative securities may be solved in principle by using Monte Carlo methods. The situation is, however, more difficult when Monte Carlo methods are not applicable to the pricing of some types of derivatives.

as simple parametric functions for pure simplicity and tractability. Most parametric specifications of the spot interest rate models are nested in the model by Chan, et al (1992) which specifies $\mu(r(t)) = \alpha_0 + \alpha_1 r(t), \sigma(r(t)) = \sigma r(t)^\gamma$. Special cases of this model are the Vasicek (1977) model with restriction $\gamma = 0$, the Cox, Ingersoll and Ross (1985) (hereafter, CIR) model, the Brown and Dybvig (1986) model, and the Gibbons and Ramaswany (1993) model with $\gamma = 1/2$, the Courtadon (1982) model with $\gamma = 1$, the Merton (1973) model with $\alpha_1 = 0, \gamma = 0$, the Dothan (1978) model with $\alpha_0 = \alpha_1 = 0, \gamma = 1$, and the Cox (1975) and Cox, Ingersoll and Ross (1980) model with $\alpha_0 = \alpha_1 = 0, \gamma = 3/2$. Aït-Sahalia (1996) extends the model by specifying $\mu(\cdot)$ as a linear mean-reverting function, and $\sigma^2(\cdot)$ as a semi-parametric function determined by $\mu(\cdot)$ and the marginal density function of the process. Stanton (1996) proposed nonparametric estimation of the drift function and diffusion function based on their approximations. In this paper we further extend the spot interest rate model by assuming both $\mu(\cdot)$ and $\sigma^2(\cdot)$ are robust nonparametric functions. That is, no *a priori* restrictions are imposed on the structure of spot interest rate process, data are allowed to speak for themselves. Moreover, the model can be either strictly stationary or stationary only in the asymptotic sense.

The strongest implication of the one-factor term structure model is that the whole yield curve is endogenous. Even though the one-factor model is criticized for various reasons, it is still very attractive to both practitioners and academics mainly because: First, it promises to offer a consistent model, with parsimonious structure, for the fundamental behavior of interest rates and term structure; Second, it provides an

6

unifying tool for the pricing of many interest rate derivative securities; and Third, most importantly, the model is easy to implement from a computational point of view since the underlying model is a one-dimensional Markov process. Given the spot rate $r(t)$ at time t, $r(t) = r$, and its dynamics described by (1), let $P(r(t), t, T)$ represent the price at t of any interest rate derivative security maturing at T. From Itô's Lemma, the instantaneous return on the bond is

$$dP/P = [\mu(r)\frac{\partial P}{\partial r} + \frac{\partial P}{\partial t} + \frac{1}{2}\sigma^2(r)\frac{\partial^2 P}{\partial r^2}]/Pdt + \sigma(r)\frac{\partial P}{\partial r}/PdW(t) \qquad (2)$$

In efficient markets, the instantaneous expected rate of return (including the cash flow rate) for any asset can be written as the risk free return, $r(t)$, plus a risk premium [7]. Let $\xi(r(t), t, T)$ be the instantaneous expected rate of return and $\nu(r(t), t, T)$ be the instantaneous standard deviation of the interest rate derivative security, the absence of arbitrage implies that

$$\xi(r(t), t, T) = r(t) + \lambda(r(t))\nu(r(t), t, T) - d(r(t), t) \qquad (3)$$

where $\lambda(\cdot)$ is the risk premium factor or market price of interest rate risk, the existence of which is a necessary condition for absence of arbitrage, and $d(r(t), t)$ is the cash flow rate of the security. It is noted that since there is only a single factor which affects bond returns, this implies that the instantaneous returns on bonds of all maturities are *locally* perfectly correlated.

It must be noted that the above representations depend on certain technical conditions

---

7   The main concern here is simply the absence of arbitrage opportunities. It is usually assumed that agents are restricted to self-financing strategies that are adapted to $\{F_t\}$ (which means that at any point in time a strategy can depend only on information known at that time) and for which the discounted value of the portfolio under an equivalent probability measure is a martingale.

on $\mu(\cdot)$, $\sigma(\cdot)$, and $\lambda(\cdot)$. Regularity conditions on the drift and diffusion terms are that $\mu(\cdot)$ and $\sigma(\cdot)$ each must be Borel measurable and satisfy Lipschitz conditions and growth conditions (see e.g. Wong (1971), p.150). These conditions are imposed to ensure the existence of a unique strong (non-exploding) solution to the SDE of $r(t)$. Similarly, restrictions must be placed on the market price of risk to guarantee the existence of the equivalent martingale measure. In the traditional models, the functional forms of the market price of risk are also specified for pure simplicity and tractability, examples are $\lambda(r(t)) = \lambda$ in Vasicek (1977), $\lambda(r(t)) = \lambda r(t)^{1/2}/\sigma$ in CIR (1985), $\lambda(r(t)) = 0$ in Chan, et al (1992), and $\lambda(r(t)) = \lambda$ in Aït-Sahalia (1996a). As CIR (1985) point out, the arbitrage approach does not imply that every choice of the functional form for the market price of risk will lead to bond prices which do not admit arbitrage opportunities. Indeed CIR (1985) showed with an example that a linear functional form for $\lambda(r(t))\sigma(r(t))$ can lead to internal inconsistency. In addition to condition (3), Duffie (1988, pp.227-228) provided regularity conditions that the market price of risk must satisfy in order for the model to be consistent, i.e., $\lambda(\cdot)$ must be a predictable process satisfying $\int_t^u \lambda^2(r(s))ds < \infty$, $a.s. \forall \quad u \geq t$, and $E[exp\{\frac{1}{2}\int_t^T \lambda^2(r(s))ds\}] < \infty$. It is easy to verify that the example in CIR (1985) fails to satisfy the regularity conditions. In this paper, instead of imposing *a priori* restrictions on the functional form of the market price of risk, we will specify the market price of risk as a general function which satisfies the above regularity conditions. Its specific form or shape is determined by forcing the term structures implied by the model to best fit the observed term structures. Therefore instead of totally ignoring the information contained in the current term structure or to another

8

extreme taking the entire yield curve as given, the major information contained in the historical term structures is extracted to determine the functional form of the market price of risk.

2.2. Pricing of Bonds and Other Derivative Securities: The PDE Approach and Monte Carlo Approach

Equalling both the expected rate of return and instantaneous standard deviation of the interest rate derivative security in (3) and those in the SDE of $P(r, t, T)$ in (2) yields

$$\frac{\partial P(r, t, T)}{\partial t} + \frac{1}{2}\sigma^2(r)\frac{\partial^2 P(r, t, T)}{\partial r^2} + \{\mu(r) - \lambda(r)\sigma(r)\}\frac{\partial P(r, t, T)}{\partial r}$$
$$-r(t)P(r, t, T) + d(r, t)P(r, t, T) = 0 \qquad (4)$$

This is the fundamental equation for the price of any contingent claim whose value depends solely on the spot rate, $r(t)$, and the time to maturity, $\tau = T - t$. The PDE can be solved for the price $P(r, t, T)$ with certain given conditions: the continuous payment rate $d(r, t)$, the initial (or final) condition, and boundary condition(s) depending on the particular security considered (e.g., call, put, cap, floor, swap, etc.).

Solutions of PDEs of the parabolic or elliptic type, such as (4), can be represented in an integral form in terms of an underlying stochastic process. Under regularity conditions, the PDE for the derivative security prices has a unique solution or representation given by

$$P(r, t, T) = \int_0^{+\infty} \Gamma(r, t, T : \omega, 0, T)\phi(\omega)d\omega$$
$$+ \int_T^t \int_0^{+\infty} \Gamma(r, t, T; \omega, s, T)\psi(\omega, s, T)d\omega ds \qquad (5)$$

where $\Gamma(r, t, T; \omega, s, T)$ is the fundamental solution of $LP(r, t, T) = \psi(r, t, T)$ in

9

the sense that for $r, \omega > 0$ and $T \geq t \geq s \geq 0$, $\Gamma(r, t, T; \omega, s, T)$, as function of r and t, solves $LP = 0$ for every (ω, s) in $R_+ \times [0, T]$, and $\phi(r) = P(r, T, T)$, where L is a natural differential operator defined as $LP(r, t, T) = \frac{\partial P(r,t,T)}{\partial \tau} - \frac{1}{2}\sigma^2(r)\frac{\partial^2 P(r,t,T)}{\partial r^2} - \{\mu(r) - \lambda(r)\sigma(r)\}\frac{\partial P(r,t,T)}{\partial r} + rP(r, t, T)$.

An alternative way of solving for the prices of derivative securities is based on performing Monte Carlo simulations of the sample paths of the risk-neutral process

$$d\tilde{r}(t) = (\mu(\tilde{r}(t)) - \lambda(\tilde{r}(t))\sigma(\tilde{r}(t)))dt + \sigma(\tilde{r}(t))dW(t) \qquad (6)$$

which is also a time-homogeneous diffusion process, with the drift term modified for the market price of interest rate risk. The sample paths, all starting at $r(t) = r$ at date t and finishing at date T, can be simulated from (6). The conditional expectation under the risk-neutral dynamics gives the prices

$$P(r(t), t, T) = E_t[b(\tilde{r}(T))exp\{-\int_t^T \tilde{r}(u)du\}$$
$$+ \int_t^T exp\{-\int_t^\tau \tilde{r}(u)du\}d(\tilde{r}(\tau), \tau)d\tau | \tilde{r}(t) = r(t)] \qquad (7)$$

where $b(\cdot)$ is the final payoff at maturity. The price $P(r(t), t, T)$ can then be obtained by averaging the argument of the conditional expectation over the simulated sample paths. For example, a zero-coupon bond price with face value $P(r(t), T, T) = 1$ is given by

$$P(r(t), t, T) = E_t[exp\{-\int_t^T \tilde{r}(u)du\} | \tilde{r}(t) = r(t)]$$

and the price of a call option on a zero coupon bond of maturity $T - t$ with strike

price K and exercise date t_1, $t \leq t_1 \leq T$, is given by

$$C^0(P(r(t), t, T), K, t_1) = E_t[exp\{-\int_t^{t_1} \tilde{r}(u)du\}$$

$$\max[P(r(t_1), t_1, T) - K, 0]|\tilde{r}(t) = r(t)]$$

Simulations of the sample path can be performed using the Euler scheme or alternatively the Milshtein scheme with discretization step T/n over the time interval $[0, T]$. It is noted that the Milshtein scheme has better convergence rates than the Euler scheme for the convergence in $L^p(\Omega)$ and the almost sure convergence (see Talay, 1996). In financial applications of the Monte Carlo simulation methods, a number of variance reduction methods have been proposed, e.g. the control variate approach, the antithetic variate method, the moment matching method, the importance sampling method, the conditional Monte Carlo methods, and quasi-random Monte Carlo methods (see, e.g. Boyle, Broadie and Glasserman, 1996). It is noted that Monte Carlo simulation is also one of the often used approach to solving the PDE in (4). The Monte Carlo method seems to be disadvantageous when a finite difference method, a finite element method, a finite volume method, or a suitable deterministic algorithm is numerically stable and does not require too long computational time. However, in the financial applications a Monte Carlo algorithm may be interesting when one wants to compute the option prices at only a few points or when the PDE approach is difficult to implement due to the complexity of the problem.

3. Parametric Models and Closed Form Solution of Bond and Bond Option Prices

In order to have closed form solutions for bond or bond option prices, most authors have chosen simple parametric specifications for the spot interest rate model and the market price of risk. The following are two well known parametric models which lead to closed form solutions of bond and bond option prices. The first set of parametric specifications is as following:

(i) $\mu(r(t)) = \beta(\alpha - r(t)); \sigma(r(t)) = \sigma;$ and $\lambda(r(t)) = \lambda$

where $\beta(> 0), \alpha, \sigma(> 0)$ and λ are constants, i.e. the spot interest rate model is specified as an Ornstein-Uhlenbeck process (as in Vasicek (1977) and Jamshidian (1989)), and the market price of risk as a constant. Under certain initial and boundary conditions, the PDE (4) has known solutions. For instance, the price of pure discount bond, i.e. $P(r, T, T) = 1$, with maturity $T - t$, as derived in Vasicek (1977), is given by

$$P(r, t, T) = exp[A(T - t)(R(\infty) - r) - (T - t)R(\infty) - \frac{\sigma^2}{4\beta}A(T - t)^2] \quad (8)$$

where $t \leq T$, $R(\infty) = \alpha + \sigma\lambda/\beta - \frac{1}{2}\sigma^2/\beta^2$, $A(T - t) = \frac{1}{\beta}(1 - e^{-\beta(T-t)})$. Or equivalently, the term structure of interest rates takes the form

$$R(t, T) = R(\infty) + (r(t) - R(\infty))\frac{A(T - t)}{(T - t)} + \frac{\sigma^2 A(T - t)^2}{4\beta(T - t)}, \quad t \leq T \quad (9)$$

The yield curve is monotonically increasing for values of $r(t)$ smaller or equal to $R(\infty) - \frac{1}{4}\rho^2/\alpha^2$, and monotonically decreasing for values of $r(t)$ larger than or equal to $R(\infty) + \frac{1}{2}\rho^2/\alpha^2$. For values of $r(t)$ larger than $R(\infty) - \frac{1}{4}\rho^2/\alpha^2$ but below $R(\infty) + \frac{1}{2}\rho^2/\alpha^2$, the yield is a humped curve. The value at time t of an European call option, given that $r(t) = r$, on a pure discount bond with maturity $T - t$ with

12

exercise price K and expiration date S, $t \leq S \leq T$, as derived in Jamshidian (1989), is given by:

$$P(r, t, S; T, K) = P(r, t, T)N(h) - K P(r, t, S)N(h - \sigma_p) \qquad (10)$$

where $h = ln[P(r, t, T)/P(r, t, S)K]/\sigma_p + \sigma_p/2$, $\sigma_p = \sigma(1 - e^{-\beta(T-S)})(\frac{1 - e^{-2\beta(T-t)}}{2\beta^3})^{1/2}$, and $N(\cdot)$ is the CDF of standard normal distribution.

The discount yield curve implied by this model is obviously restricted in its shape. Apart from that, the Ornstein-Uhlenbeck process is also often criticized for allowing negative interest rates, as $r(t)$ is defined over $(-\infty, +\infty)$. Moreover, the following analysis will show that the Ornstein-Uhlenbeck process imposes very strong unrealistic restrictions on the structure of spot interest rate. If $r(t)$ follows an O-U process, then we have (see Appendix for a brief derivation), $E[r(t)] = \alpha$, $Var[r(t)] = \frac{\sigma^2}{2\beta}$, $Cov(r(t + \tau), r(t)) = \frac{\sigma^2}{2\beta}e^{-\beta\tau}$, or $Corr(r(t + \tau), r(t)) = e^{-\beta\tau}$ which is positive for all τ, decreases as τ increases, and approaches 1 as τ goes to 0. For δ-period difference of the stochastic process, $\Delta_\delta r(t) = r(t) - r(t - \delta)$, we have $E[\Delta_\delta r(t)] = 0$, $Var[\Delta_\delta r(t)] = \frac{\sigma^2}{\beta}(1 - e^{-\beta\delta})$, and $Cov(\Delta_\delta r(t + \tau), \Delta_\delta r(t)) = -\frac{\sigma^2}{2\beta}e^{-\beta(\tau-\delta)}(1 - e^{-\beta\delta})^2$, where $\tau \geq \delta$, or

$$Corr(\Delta_\delta r(t + \tau), \Delta_\delta r(t)) = -\frac{1}{2}e^{-\beta(\tau-\delta)}(1 - e^{-\beta\delta}), \quad \tau \geq \delta$$

which is always negative. When $\tau = \delta$, the first-order autocorrelation of the δ-period difference of the spot interest rate is $Corr(\Delta_\delta r(t + \delta), \Delta_\delta r(t)) = -\frac{1}{2}(1 - e^{-\beta\delta})$ which is negative and bounded between 0 and $-1/2$. When $\tau = 2\delta$, the second-order autocorrelation is $Corr(\Delta_\delta r(t + 2\delta), \Delta_\delta r(t)) = -\frac{1}{2}e^{-\beta\delta}(1 - e^{-\beta\delta})$ which is also

13

negative and bounded between 0 and $-1/8$.

The second set of parametric specifications is as following:

(ii) $\mu(r(t)) = \beta(\alpha - r(t))$; $\sigma(r(t)) = \sigma r^{1/2}(t)$; and $\lambda(r(t)) = \lambda r^{1/2}(t)/\sigma$

where $\beta(> 0)$, α, $\sigma(> 0)$ and λ are constants, which is specified in CIR (1985). The PDE (4) with certain initial and boundary conditions also has known solutions. For instance, the price of a pure discount bond with maturity $T - t$, as derived in CIR (1985), is of the form

$$P(r, t, T) = A(t, T)e^{-B(t,T)r} \tag{11}$$

with $A(t, T) \equiv [\frac{\phi_1 exp(\phi_2(T-t))}{\phi_2[exp(\phi_1(T-t))-1]+\phi_1}]^{\phi_3}$, $B(t, T) \equiv \frac{exp(\phi_1(T-t))-1}{\phi_2[exp(\phi_1(T-t))-1]+\phi_1}$, where $\phi_1 \equiv [(\beta + \lambda)^2 + 2\sigma^2]^{1/2}$, $\phi_2 \equiv (\beta + \lambda + \phi_1)/2$, and $\phi_3 \equiv 2\beta\alpha/\sigma^2$. The bond price is a decreasing convex function of the spot interest rate. Or equivalently, the term structure of interest rate takes the form

$$R(t, T) = \frac{1}{T - t}(B(t, T)r(t) - ln A(t, T)) \tag{12}$$

which is an increasing function of $r(t)$ cross section and is either an increasing, humped or decreasing function of $T - t$ cross maturity depending on the value of $r(t)$. Similarly, the value at time t of an European call option, given that $r(t) = r$, on a pure discount bond of maturity $T - t$ with exercise price K and expiration date S, $t \le S \le T$, as derived in CIR (1985), is given by:

$$P(r, t, S; T, K) = P(r, t, T)\chi^2(2r^*(\phi + \psi + B(S, T)); \tfrac{4\beta\alpha}{\sigma^2},$$

$$\tfrac{2\phi^2 re^{\gamma(S-t)}}{\phi+\psi+B(S,T)}) - K P(r, t, S)\chi^2(2r^*(\phi + \psi); \tfrac{4\beta\alpha}{\sigma^2}, \tfrac{2\phi^2 re^{\gamma(S-t)}}{\phi+\psi}) \tag{13}$$

14

where $\gamma \equiv \phi_1$, $\phi \equiv \frac{2\gamma}{\sigma^2(e^{\gamma(S-t)}-1)}$, $\psi \equiv (\beta + \lambda + \gamma)/\sigma^2$, $\chi^2(\cdot; \cdot, \cdot)$ is the noncentral chi-square distribution with $\frac{4\beta\alpha}{\sigma^2}$ degrees of freedom and parameter of noncentrality $\frac{2\phi^2\gamma e^{r(S-t)}}{\phi+\psi}$, and $r^* \equiv [ln(\frac{A(S,T)}{K})]/B(S,T)$ is the critical interest rate below which exercise will occur, i.e., $K = P(r^*, S, T)$.

Different from the O-U process, the CIR squared-root process only allows for non-negative interest rates as zero is a reflecting barrier of the process. However, similar to the O-U process, the CIR squared-root process also imposes very strong unrealistic restrictions on the structure of the spot interest rates. If $r(t)$ follows a CIR squared-root process, then we have (see Appendix for a brief derivation), $E[r(t)] = \alpha$, $Var[r(t)] = \frac{\alpha\sigma^2}{2\beta}$, $Cov(r(t+\tau), r(t)) = \frac{\alpha\sigma^2}{2\beta}e^{-\beta\tau}$, and $Corr(r(t+\tau), r(t)) = e^{-\beta\tau}$ which is always positive, decreases to 0 as τ increases, approaches 1 as τ goes to 0. For δ-period difference of the process, $\Delta_\delta r(t) = r(t) - r(t-\delta)$, we have $E[\Delta_\delta r(t)] = 0$, $Var[\Delta_\delta r(t) = \frac{\alpha\sigma^2}{\beta}(1-e^{-\beta\delta})$, and $Cov(\Delta_\delta r(t+\tau), \Delta_\delta r(t) = -\frac{\alpha\sigma^2}{2\beta}e^{-\beta(\tau-\delta)}(1-e^{-\beta\delta})^2$, where $\tau \geq \delta$, or

$$Corr(\Delta_\delta r(t + \tau), \Delta_\delta r(t)) = -\frac{1}{2}e^{-\beta(\tau-\delta)}(1 - e^{-\beta\delta}) \qquad \tau \geq \delta$$

It imposes the same restrictions as the O-U process on the pattern of the autocorrelation of the δ-period difference of the spot interest rate. That is, the autocorrelation of the δ-period difference of the process is always negative with the first-order autocorrelation bounded between $-1/2$ and 0, the second-order autocorrelation bounded between $-1/8$ and 0, and so on.

Apart from the above two sets of parametric specifications, the models specified by

Dothan (1978) and Brown and Schaefer (1991) also give closed form solutions for bond and bond option prices. However, Dothan's model implies that both the interest rate itself and its log returns are nonstationary processes. The Brown and Schaefer model restricted the closed form solution of bond prices to the Vasicek and CIR type, i.e. $P(r, t, T) = A(t, T)exp[-B(t, T)r]$, and the underlying model of the state variable is defined as the risk-neutralized process based on its equivalent probability measure. Even so, $A(t, T)$ and $B(t, T)$ do not always have explicit solutions given the stochastic process of the spot interest rate. In addition to models with closed form solutions, bond and bond option prices have been computed by Hull and White (1990) based on time-inhomogeneous models, and by Aït-Sahalia (1996) based on a spot rate model with parametric drift and semi-parametric diffusion, to list only a few. Since assumptions made on different models are different from one to another, the pricing formulas are different as well. In Hull and White (1990), prices of call options on a 5-year bond are computed based on different parametric models [8]. For instance, the prices of out-of-the-money call options with strike price 105.00 (par value of bond is 100.00) and maturities of 0.5, 1.0, 1.5, 2.0, 3.0, and 4.0 years are respectively 0.05, 0.16, 0.22, 0.22, 0.12, and 0.01 for extended Vasicek model and 0.04, 0.13, 0.17, 0.17, 0.08, and 0.00 for CIR model. The differences are generally over 20 %. Aït-Sahalia (1996) also found that the prices based on his model are significantly different from those calculated from the Vasicek model and CIR model, especially for deep out-of-the-money long-maturity options.

---

8   The parameters are set as arbitrary values (not estimated from actual data) but are ensured that the initial short-term interest rate volatilities are equal among different models.

16

4. Nonparametric Model and Numerical Solution of Bond and Bond Option Prices

The fact that different models generate different prices for bond and other interest rate derivative securities naturally raises the question that which of the models should be employed. Or in other words, which of the competing models made the most reliable assumptions about the spot interest rate process and the market price of risk. Before looking at the data, this question cannot be answered with any credibility as the data set is different from one to another and one model fitting one data set well does not necessarily mean fitting another data set also well. That is, either we can first specify the model with *a priori* restrictions, then subject the model to empirical test to determine whether the model should be rejected or not, or alternatively we can impose least restrictions on the model and let the data determine what kind of structure the model should have. The approach adopted in the traditional spot interest modeling is exactly the first one. Unfortunately the models are implemented in most cases without subjecting to empirical tests. The nonparametric approach proposed in this paper to modeling term structure dynamics is the alternative. That is, instead of specifying simple functional forms for the drift function, diffusion function and the market price of risk for pure simplicity and tractability, the model specification is determined by the specific data set.

With the coefficient functions $\mu(r(t))$ and $\sigma(r(t))$ in (1) specified as elements of a family of general functions, the diffusion function $\sigma(r(t))$ can be estimated using

17

the following consistent nonparametric estimator from discretely observed high frequency data (see Jiang and Knight, 1995, for proof of consistency):

$$\hat{\sigma}^2(r) = \frac{\sum_{i=1}^{n-1} n\, K(\frac{r_i-r}{h_n})(r_{i+1} - r_i)^2}{\sum_{i=1}^{n} T\, K(\frac{r_i-r}{h_n})} \tag{14}$$

where, without loss of generality, $\{r_i : i = 1, 2, ..., n\}$ are assumed to be equispaced n observations over the time period $[0, T]$ with $T > 0$ and sampling interval $\Delta_n = T/n$, $K(\cdot)$ is a positive kernel density function satisfying regularity conditions, and h_n is the window-width of the nonparametric estimator (see Appendix for admissible window-width conditions of the estimator).[9] The above nonparametric estimator of diffusion function requires only mild regularity conditions and works for very general (both stationary and nonstationary) diffusion processes and the drift term is a nuisance coefficient function. As $h_n \to 0, n \to \infty, nh_n \to \infty$, and $nh_n^3 \to 0, \hat{\sigma}^2(r)$ is a pointwise consistent estimator of $\sigma^2(r)$ and asymptotically normally distributed. The variance of $\hat{\sigma}^2(r)$ can be consistently estimated by $\hat{V}[\hat{\sigma}^2(r)] = \hat{\sigma}^4(r)/\sum_{i=1}^{n} K(\frac{r_i-r}{h_n})$.

Direct estimation of the drift function $\mu(r(t))$ without any restrictions on the underlying structure of the diffusion process is impossible in general from discretely observed data either over a short time interval (no matter how frequent the data is) or with fixed frequency (no matter how long period it is spanned). It's *approximations* from high frequency data are possible, as suggested by Stanton (1996), but they can be extremely non-robust in that the estimates are very sensitive to the sampling path.

9　Choice of h_n based on cross-validation (CV) rule is proved to be asymptotically optimal w.r.t. the averaged squared error (ASE) rule or integrated squired error (ISE) rule by Kim and Cox (1996). A condition required in their proof is strong-(α-) mixing which is less restrictive than the often required uniform-(ϕ-) mixing condition. While the uniform-mixing condition is quite satisfactory for most Markovian processes, it is too strong to apply to Gaussian processes.

A robust nonparametric estimator of the drift function is developed in Jiang and Knight (1995) as well using information contained in the marginal density function through the following relationship. A consistent estimator of the nonparametrically specified drift function $\mu(r(t))$ can be obtained from

$$\hat{\mu}(r) = \frac{1}{2}[\frac{d\hat{\sigma}^2(r)}{dr} + \hat{\sigma}^2(r)\frac{\hat{p}'(r)}{\hat{p}(r)}] \tag{15}$$

which is derived from the Kolmogorov forward equation under the condition that the diffusion process is either strictly stationary or has a limiting probability density function, where $\hat{p}(r)$ is the kernel estimator of the marginal density function of the stochastic process.

Following above identification and estimation procedure, both the nonparametric diffusion function and drift function can be identified and estimated. The empirical results in this paper not only confirm that the above identification and estimation procedure provides robust estimation results for the spot interest rate process but also suggest there are strong evidence that the traditional spot interest rate models are misspecified.

To learn about the functional form of the market price of interest rate risk, we have to resort to the information contained in the term structure data of interest rates or the information contained in the cross-sections data of any other interest rate derivative security. A direct way of observing $\lambda(\cdot)$ empirically is proposed in Vasicek (1977) using the following equality:

$$\frac{\partial Y(r(t), \tau)}{\partial \tau}|_{\tau=0} = \frac{1}{2}(\mu(r(t)) - \sigma(r(t))\lambda(r(t))) \tag{16}$$

19

where $Y(r(t); \tau)$ is the yield at t with instantaneous risk free rate $r(t)$ and maturity τ. The advantage of using above equality is that the coefficient of $\frac{\partial P(r,t,T)}{\partial r}$ in the valuation PDE (4) can be replaced by $2\frac{\partial Y(r(t),\tau)}{\partial \tau}|_{\tau=0}$. By doing so, we can actually avoid directly estimating the market price of risk $\lambda(r(t))$ and even the drift function $\mu(r(t))$ of the spot rate process. However, estimating $\frac{\partial Y(r(t),\tau)}{\partial \tau}|_{\tau=0}$ requires observations of the yields with very short maturities for different levels of instantaneous risk free rates. Apart from the fact that this kind of observations is very difficult to collect, such data set is not desirable for estimation and statistical inference due to the spurious microstructure effects associated with yields of short maturities and unavoidable measurement errors. Moreover, this equality only uses the information of the yield curve close to the origin, some important information contained in the entire yield curve might be ignored. In this paper, we first observe the market price of risk $\lambda(r(t), t)$ by fitting the implied yield curve of the model to the historical yield curve [10], then estimate the time-stationary market price of risk curve $\lambda(r(t))$ using the smoothing technique:

$$\hat{\lambda}(r) = \frac{\sum_{t=t_0}^{t=t_n} K(\frac{r(t)-r}{h_n})\lambda(r(t), t)}{\sum_{t=t_0}^{t=t_n} K(\frac{r(t)-r}{h_n})} \tag{17}$$

where $\{t_0, t_1, ..., t_n\}$ are the points of time at which the historical yield curves are observed, $K(\cdot)$ is a kernel function satisfying regularity conditions, and h_n is the smoothing parameter which can be chosen to minimize the IMSE of $\hat{\lambda}(r)$ (see Appendix for the algorithm of calculating the smoothing parameter). As the non-parametric estimator is smooth and bounded, it is easy to verify that the estimated

---

10 In practice, $\lambda(r(t), t)$ can be obtained by minimizing, e.g., the sum of squared deviations across maturities between the given historical yield curve and the yields produced by the model.

20

market price of risk satisfies the regularity conditions in section 2.1, and hence the model is internally consistent and precludes any arbitrage opportunities.

With nonparametrically estimated $\mu(\cdot)$, $\sigma(\cdot)$ and $\lambda(\cdot)$, the prices of discount bonds and other derivative securities $P(\cdot, \cdot)$ can be obtained by either solving the PDE numerically or performing Monte Carlo simulations along the sample path of risk neutral process, as discussed in section 2.2. By assuming $\lambda(\cdot)$ is given or conditional on $\lambda(\cdot)$, the asymptotic distribution and asymptotic variance of the derivative security prices $P(\cdot; \cdot)$ are derived (see Jiang (1996)). In practice, the block-wise bootstrap technique proposed in Künsch (1989) can be used to compute the standard derivations. Similarly, standard derivations in the case of Monte Carlo simulations can also be straightforwardly computed, which can be used to monitor the errors of the Monte Carlo simulations.

The advantages of the above nonparametric term structure model include: (i) the model is nonparametrically specified, allowing for non-linearity and maximal flexibility in fitting into the data; (ii) the model precludes any arbitrage opportunities; (iii) the model is time-homogeneous and provides a consistent framework to study the fundamental behavior of interest rates and term structure; and (iv) the model retains a parsimonious structure and the computational tractability.

5. Implications of Nonparametric Model on Term Structure Dynamics and Option Prices

In this section, we will estimate the nonparametric term structure model using the

21

historical Canadian interest rate term structure data and investigate its implications on term structure dynamics and prices of interest rate derivative securities. The Vasicek (1977) and CIR (1985) models are also estimated using GMM for comparison purpose. The market prices of risk for both the Vasicek (1977) and CIR (1985) models are estimated by fitting into the average yield curve over the sample, while the market price of nonparametric risk is estimated using the smoothing technique proposed in Section 4.[11] The estimates of $\mu(\cdot)$, $\sigma(\cdot)$ and $\lambda(\cdot)$ then are plugged into the PDE (4) which is solved analytically or numerically for bond prices and the risk-neutral process (6) based on which the Monte Carlo simulations are performed to compute the bond option prices.

5.1. The Data

The observed interest rate term structure data is provided by Statistics Canada and plotted in Figure 1, which presents the historical movements of Canadian Treasury bill rates with maturities of 1-month, 3-month, 6-month, and 1-year, as well as Canadian bond yields with maturities of 2-year, 3-year, 5-year, 10-year, and 30-year. The data are weekly and cover the period from June 2, 1982 to March 1, 1995, providing 666 observations in total. All the data are quoted as the average rates of the business days in a week and are expressed in annualized form as continuously compounded yield

---

11   $\lambda(\cdot) = 0$ under the assumption that the local expectations hypothesis holds, i.e., the expected return on all interest rate-sensitive contingent claims is the riskless rate. Many theories have been presented to explain the relation among interest rates on bonds of various maturities in an uncertain economy. One of the earliest theories is known as the expectations hypothesis. In continuous-time models, the expectations hypothesis plays the same pivotal role that risk neutrality does for option pricing. Of the various versions of expectations hypotheses, the local expectations hypothesis is the only one acceptable as a statement of equilibrium in continuous-time models. Other versions of expectations hypothesis all lead to arbitrage opportunities.

22

Table 1.
Summary Statistics of the Term Structure Data and Stationarity Test

(a) Summary Statistics [b]

Variables[a]	N	Mean	Std. Dev.	ρ_1	ρ_3	ρ_5	ρ_7	ρ_9	ρ_{11}
r_t^{1m}	666	0.0891	0.0267	0.990	0.967	0.942	0.916	0.887	0.856
$r_{t+1}^{1m} - r_t^{1m}$	665	-1.24E-4	2.61E-3	0.014	0.067	0.058	0.081	0.036	-0.046
r_t^{3m}	666	0.0910	0.0260	0.991	0.967	0.940	0.913	0.883	0.853
$r_{t+1}^{3m} - r_t^{3m}$	665	-1.17E-4	2.33E-3	0.189	0.045	0.075	0.099	0.033	0.003
r_t^{6m}	666	0.0927	0.0252	0.989	0.961	0.931	0.900	0.870	0.838
$r_{t+1}^{6m} - r_t^{6m}$	665	-1.20E-4	2.79E-3	0.097	0.028	0.074	0.052	0.063	0.013
r_t^{1y}	666	0.0946	0.0240	0.988	0.957	0.922	0.887	0.854	0.819
$r_{t+1}^{1y} - r_t^{1y}$	665	-1.19E-4	2.80E-3	0.090	0.028	0.020	0.019	0.063	0.009
r_t^{2y}	666	0.0937	0.0210	0.987	0.954	0.915	0.876	0.836	0.798
$r_{t+1}^{2y} - r_t^{2y}$	665	-1.10E-4	2.48E-3	0.068	0.109	0.037	0.053	0.030	0.016
r_t^{3y}	666	0.0943	0.0196	0.987	0.952	0.909	0.866	0.823	0.782
$r_{t+1}^{3y} - r_t^{3y}$	665	-1.02E-4	2.31E-3	0.103	0.131	0.020	0.034	0.008	0.027
r_t^{5y}	666	0.0962	0.0187	0.987	0.954	0.916	0.875	0.834	0.796
$r_{t+1}^{5y} - r_t^{5y}$	665	-1.04E-4	2.09E-3	0.051	0.113	0.031	0.041	-0.011	0.015
r_t^{10y}	666	0.0994	0.0180	0.987	0.957	0.921	0.883	0.847	0.812
$r_{t+1}^{10y} - r_t^{10y}$	665	-1.02E-4	1.83E-3	0.066	0.096	0.059	0.036	-0.007	0.060
r_t^{30y}	666	0.1024	0.0175	0.988	0.960	0.927	0.891	0.858	0.826
$r_{t+1}^{30y} - r_t^{30y}$	665	-1.01E-4	1.72E-3	0.038	0.085	0.070	0.023	-0.025	0.031

(b) Augmented Dickey-Fuller Stationarity Test of the 1-Month T-Bill Rates

H_0:	Test Statistic [c]	Critical Value (10%)
Nonstationarity	-2.83	-2.57

Note: a. $r_t^{1m}, r_t^{3m}, r_t^{6m}, r_t^{1y}, r_t^{2y}, r_t^{3y}, r_t^{5y}, r_t^{10y}, r_t^{30y}$ denote respectively the weekly observations of 1-, 3-, 6-month, and 1-year Canadian Treasury bill rates, as well as 2-, 3-, 5-, 10-, and 30-year Canadian bond yields. $r_{t+1}^{1m} - r_t^{1m}, r_{t+1}^{3m} - r_t^{3m}, r_{t+1}^{6m} - r_t^{6m}, r_{t+1}^{1y} - r_t^{1y}, r_{t+1}^{2y} - r_t^{2y}, r_{t+1}^{3y} - r_t^{3y}, r_{t+1}^{5y} - r_t^{5y}, r_{t+1}^{10y} - r_t^{10y}, r_{t+1}^{30y} - r_t^{30y}$ denote the corresponding week-to-week change of the weekly Treasury bill rates and bond yields;
b. ρ_j denotes the autocorrelation coefficient of order j;
c. The lag order for the augmented Dickey-Fuller nonstationarity test of the 1-month t-bill rates is 9, see Harvey (1993) for description of the test statistic, and Phillips (1987) for the justification of using the Dickey-Fuller table when the residuals are heteroskedastic and possibly serially dependent.

to maturity [12]. The short-term interest rate used for estimating the spot rate process is

---

12 When calculating Treasury bill rates and bond yields, the rate of coupon bonds is first converted to the zero coupon bond rate, then the bond equivalent yields are transformed to the continuously compounded yields to maturity. It is also noted that the measurement error of Treasury bill rates or bond yields for different days within a week due to different maturity lengths which is caused by fixed auction day is averaged out and therefore minimal.

the weekly Canadian 1-month Treasury bill rates which capture the volatility of the short-term interest rates from week to week. The rates as a time series are the first intersection of the term structure data as plotted in Figure 1.

Table 1 reports the means, standard deviations, and part of the first eleven autocorrelations of the weekly rates and the weekly changes in the rates for all different maturities. The unconditional average level of the weekly observations for 1-month, 3-month, 6-month, 1-year Treasury bill rates and 2-year, 3-year, 5-year, 10-year, 30-year bond yields are respectively 0.0891, 0.0910, 0.0927, 0.0946, 0.0937, 0.0943, 0.0962, 0.0994, and 0.1024, which is generally increasing with maturity, with standard deviations of 0.0267, 0.0260, 0.0252, 0.0240, 0.0210, 0.0196, 0.0187, 0.0180, and 0.0175, which is generally decreasing with maturity. Although the autocorrelations in the interest rate level decays very slowly, those of the week-to-week changes are generally small and are not consistently positive or negative. It is also noted that the autocorrelations of different orders of the first difference of the Treasury bill rates and bond yields seem more likely to be positive than negative, suggesting that neither the O-U process nor the CIR process can be used as a reasonable representation of any of the series, as both models only allow negative autocorrelations for the δ-period difference of the process. The result of a formal augmented Dickey-Fuller nonstationarity test for the one month Treasury bill rates is also reported in Table 1. The null hypothesis of nonstationarity is rejected at the 10% significance level. Since the test is known to have low power which is the probability of rejecting the null hypothesis when it is not true, even a slight rejection means that stationarity of the series is very

24

Table 2.
Correlation Matrix and Principal Components of Weekly Changes in OTR Yields

(a) Correlation Matrix

	r_t^{1m}	r_t^{3m}	r_t^{6m}	r_t^{1y}	r_t^{2y}	r_t^{3y}	r_t^{5y}	r_t^{10y}	r_t^{30y}
r_t^{1m}	1.000	0.994	0.977	0.955	0.909	0.879	0.840	0.770	0.705
r_t^{3m}		1.000	0.993	0.978	0.934	0.905	0.863	0.789	0.718
r_t^{6m}			1.000	0.994	0.959	0.933	0.893	0.820	0.747
r_t^{1y}				1.000	0.978	0.957	0.921	0.853	0.781
r_t^{2y}					1.000	0.994	0.976	0.930	0.873
r_t^{3y}						1.000	0.990	0.955	0.907
r_t^{5y}							1.000	0.984	0.951
r_t^{10y}								1.000	0.989
r_t^{30y}									1.000

(b) Eigenvalues' Absolute (10E04) and Relative (%) Magnitudes

Absolute	1.70	2.23	2.84	4.16	5.28	7.65	20.10	31.70	424.36
Relative	0.34	0.46	0.57	0.83	1.06	1.53	4.02	6.34	84.87

(c) Orthonormal Basis of Eigen-vectors

Maturity	v9	v8	v7	v6	v5	v4	v3	v2	v1
0.083	0.0	-0.02	-0.14	0.29	0.21	0.05	0.52	-0.73	0.22
0.25	0.02	0.03	0.37	-0.71	-0.29	-0.00	-0.08	-0.40	0.32
0.50	0.04	-0.05	-0.19	0.44	-0.51	-0.29	-0.47	-0.16	0.43
1.0	-0.05	0.12	-0.15	-0.13	0.75	-0.08	-0.43	0.00	0.44
2.0	-0.01	-0.05	0.62	0.36	0.03	0.54	0.01	0.20	0.38
3.0	0.07	-0.42	-0.57	-0.24	-0.14	0.45	0.17	0.24	0.35
5.0	-0.37	0.74	-0.16	-0.04	-0.18	-0.03	0.32	0.26	0.30
10	0.79	0.08	0.09	-0.02	0.05	-0.36	0.31	0.26	0.25
30	-0.48	-0.50	0.20	-0.01	0.07	-0.53	0.30	0.23	0.22

Notes:
The variance-covariance matrix of changes in yields over one-week intervals is decomposed into an orthonormal basis of eigen-vectors. The directions of the eigen-vectors are chosen to maximize the associated variance.

likely.

5.2. Principal Components Analysis

In Table 2, the correlation matrix shows a strong correlation between each two of

the Treasury bill rates or the bond yields with different maturities, suggesting the

one-factor spot rate model could very well represent the co-movements of the whole yield curve. Correlation between the front- and back-ends of the yield curve is 0.705. Table 2 also shows the results of a principal components analysis of the variance-covariance matrix of weekly par yield changes over the observation period. There is a dominant first principal component that explains 84.87% of the total variation. There are also two less important second and third principal components accounting respectively for 6.34% and 4.02% of the total variation. The direction of the first principal component is essentially of a parallel shift across maturities, which can be interpreted as the *level* effect. The directions of the second and third principal components are associated respectively with changes in the discrepancy of the long- and short-end rates of the yield curve and the change of the slope of the yield curve, which can be interpreted as the *steepness* effect and *curvature* effect. It is noted that the principal components analysis is essentially based on a static framework, while the one-factor term structure models are dynamic structural models. The one-factor model only implies that the changes of yields with various maturities are *locally* perfectly correlated. Changes of yields with various maturities may not be perfectly correlated over a given non-small time interval due to nonlinearity of both the drift term and diffusion term. Our empirical results show that the nonparametric term structure model can effectively mimic the level effect and steepness effect.

5.3. The Estimation Results

The estimation results of the parametric models as well as nonparametric models are reported in Table 3 and Figures 2.1, 3.1, and 4.1. Both parametric spot interest rate

26

models are estimated using GMM based on the exact moment conditions derived from the continuous-time model rather than the moment conditions derived from the discretized version of the model as in most finance literature (see Appendix for a list of both conditions). The estimators based on the discretized version of the continuous-time model are known to be biased due to misspecification. The nonparametric spot interest rate model is estimated using Gaussian kernels and the window-width which minimizes the IMSE of each functional estimator. The data of the short-term spot interest rate is the weekly observations of the one-month Treasury bill yields. We checked the robustness of the estimation results by using different kernels, different window-width, and different time series of short-term interest rates. The choice of the kernel function proves to be of little importance, and the estimation results are not very sensitive to small changes of window-width around the optimal values. Moreover, the estimation results for both diffusion function and drift function based on the daily 3-month Treasury bill rates are very close to our reported results in terms of their functional shapes.

Figure 2.1 plots the nonparametric estimator, with 95% pointwise confidence band, of the diffusion function and Figure 2.2 plots the diffusion functions estimated from different models. Noticeable features of the nonparametric diffusion function and its important difference from the parametric diffusion functions include: First, the 95 % pointwise confidence band is narrower in the middle but tends to get wider dramatically towards two ends (below 2% and above 18%) for the lack of enough observations around the high and low level of interest rates. Second, the diffusion

Table 3.
Estimation Results of Alternative Interest Rate Term Structure Models

(a) Estimation of the Spot Rate Processes

Model	Estimation Method	Drift Function		Diffusion Function
Vasicek O-U Process	GMM	α =0.0695 (5.047)	β=0.4205 (1.987)	σ^2=3.0682 10^{-4} (6.620)
CIR SR Process	GMM	α =0.0593 (2.758)	β=0.3294 (1.601)	σ^2=2.8586 10^{-3} (6.588)
Nonparametric Process	Nonparametric	Nonp Drift Figure 3.1		Nonp Diffusion Figure 2.1

(b) Estimation of the Market Prices of Interest Rate Risk

Model	Estimation Method	Market Price of	Risk Parameter
Vasicek O-U Process	Curve-fitting	$\lambda_{VAS} =$	-0.9412 (-5.795)
CIR SR Process	Curve-fitting	$\lambda_{CIR} =$	-0.1598 (-10.760)
Nonparametric Process	Nonparametric Smoothing	Nonp Market Price Figure 4.1	of Risk Function

Note: The numbers in brackets are t-ratios of the above estimates.

function exhibits noticeable variations from low to high values, which provides the evidence to reject the constant or flat diffusion function specification in the Vasicek (1977) model. Third, not only does the diffusion function not look flat, it does not look linear either. Rather it shows the shape of a "smile", with both low and high interest rates showing higher volatility and the medium interest rates showing relatively lower volatility. This provides the evidence to reject the CIR (1985) squared-root model specification which expects that high interest rates should vary more than low interest rates.

The nonparametric estimator of the drift function, with its 95% pointwise confidence band which is computed using blockwise bootstrap technique proposed in Künsch (1989), is plotted in Figure 2.1 and compared with parametric drift function estimators

in Figure 2.2. The nonparametric drift function predicts the highest long-term mean for spot interest rate, i.e. around 8.6% which is very close to the unconditional mean of the observation 8.91%, while both the Vasicek and CIR models predicts much lower long-term means at the level of 6.95% and 5.93% respectively. All three drift functions exhibit consistent mean reverting property, i.e. the drift function is consistently positive for interest rate level below the long-term mean and consistently negative otherwise. While the magnitude of the pulling force is proportional to the deviation of the interest rate level from the long-term mean (i.e. β is constant) for both parametric models, the nonparametric drift function exhibits a pulling force whose magnitude has a varying proportion to the deviation of the interest rate level from its long-term mean. The proportion is consistently increasing as the deviation gets larger in its absolute value, first slowly when the interest rate level is close to the long-term mean and then rapidly when the interest rate level is far away from the long-term mean. And in general the pulling force is smaller than that of both parametric models, with Vasicek having the most significant mean reverting property.

The market price of risk λ of the Vasicek model and $\lambda r^{1/2}/\sigma$ of the CIR model are estimated by minimizing the deviation of the model implied yield curve from the average zero-coupon yield curve over the sample. The smoothing functional estimator $\hat{\lambda}(r)$ is obtained using the Gaussian kernel and the smoothing parameter which minimizes its IMSE. The nonparametric estimator of the market price of risk is plotted in Figure 4.1, with its 95% pointwise confidence band, and compared with other market prices of risk in Figure 4.2. It is noted that the nonparametric market

price of risk is essentially non-zero, pointwise significantly negative, and appears to be neither a constant nor a squared-root function of short-term interest rate.

5.4. Implications on Term Structure Dynamics

Having estimated the drift function, diffusion function, and market price of risk, our aim is now to utilize these estimates to investigate implications of model specification on term structure dynamics and prices of interest rate derivative securities. Figures 5.1, 5.2 and 5.3 plot the yield curves computed from the Vasicek model, the CIR model and the nonparametric model with short-term interest rate levels equal to 4%, 10% and 16%. The yields of various maturities are converted from the discount bond prices which are computed using the analytical solutions for the CIR and Vasicek models and by numerically solving PDE (4) for the nonparametric model with the boundary and final conditions specified in the Appendix.[13] Pointwise 95% confidence band of the nonparametric yield curves, computed using the blockwise bootstrap technique, are also plotted [14]. It is easy to see that the nonparametric model and the Vasicek and CIR models imply significantly different yields curves, especially for low- and high-level of short-term interest rates. As indicated previously, the bond and bond option prices are determined through PDE (4) by the drift term, the diffusion term and the market price of risk. For example, the short-term bond prices mostly reflect the differences in the risk-neutral first-moment of the underlying process, as a result,

---

13   Accuracy of the numerical algorithm of solving the parabolic PDE is examined by comparing the numerical solutions with analytical solutions for the Vasicek and CIR models, and the errors are found to be small.

14   The data block size is set as 52, equivalent to one year's observations, based on the autocorrelation coefficients ($\rho_{52} = 0.5461$). The number of replication is 1000.

30

even for the short-end yield curves, the nonparametric model can imply very different results than the Vasicek and CIR models. Since the models are fitted into the observed yield curves differently, consequently the CIR and Vasicek yield curves are relatively close to each other, while the nonparametric prices are mostly significantly different from the parametric yield curves. The difference is more visible for both high and low short-term interest rate level and longer-end yield curves.

Inspection of historical Canadian term structures reveals that the yield curve tends to be upward sloping and steeper at the short end (0-5 years), relatively flat for maturities in excess of 5 years. This is in general consistent with the yield curves computed from the nonparametric term structure model. Moreover, the overall impact of the short rate $r(t)$ on the yield curve is also as expected: an increase in $r(t)$ tends to shift the whole yield curve up but flattens the curve, while a drop in $r(t)$ tends to shift the whole yield curve down but steepens the curve, mimicking both the *level* effect and *steepness* effect.

The term structures replicated by the nonparametric model based on the historical short-term interest rates (i.e. calculating $Y(r(t), \tau)$ by plugging in observed $r(t)$ for various τ) prove to be a very good fit of the historical observations of term structures. As Figure 6.3 indicates, the absolute biases or residuals between the historical term structures and the nonparametric model generated term structures are generally very small, and much smaller than their counterparts in the CIR and Vasicek models, as plotted in Figures 6.1 and 6.2. Further calculation also indicates that the nonparametric model can replicate not only the term structure of yields but also the term structure

31

of yield volatilities, i.e. the yields generated by the nonparametric model also have a downward-slope volatility curve against the term.

5.5. Implications on Option Prices

Table 4 reports the 2- and 4-year call option prices with various strike prices on a 5-year discount bond [15] for the nonparametric, CIR, and Vasicek models based on Monte Carlo simulations. The strike prices are expressed as percentages of current bond prices with corresponding maturities. In performing the simulations, 1,000 risk-neutral interest rate paths are simulated using 100 time periods per day, and the variability of the results is reduced using the antithetic variate approach. The Monte Carlo integration involves two approximations: one is the Monte Carlo error due to replacing the expectation operator with the "sample" average over certain number of replications, and the other is the discretization error due to replacing the continuous-time sample path with a discrete-time sample path using certain discretization interval. In principle, the Monte Carlo error can be reduced to any desired level with significantly large number of replications. However, the rate of convergence is only squared root of the number of replications, so reducing the error by half would require four times of computation. Similarly, the discretization error can be reduced with a smaller discretization step, but due to computational time, this approximation error should be balanced with the Monte Carlo error (see, e.g. Duffie, 1992, pp201-202). Our exercises show that the antithetic variate technique can

---

15   Even if early exercise of the option is allowed (American option), it will never be optimal to exercise it because the underlying pure discount bond pays no coupon. If the underlying bond is a coupon bond, then American options would be valued differently from European options.

32

Table 4

Call Option Prices on a Five-Year Zero-Coupon Bond Under Alternative Models

Option Expiration Time (years)	Annualized Spot Rate	Strike Prices				
		96%	98%	100%	102%	104%
2	0.04	0.2270	0.0541	0.0096	0.0017	0.0006
		(0.0074)	(0.0032)	(0.0012)	(0.0003)	(0.0002)
		0.2104	0.0484	0.0081	0.0008	0.0000
		0.2182	0.0520	0.0088	0.0010	0.0001
	0.16	4.7596	3.6329	2.7557	1.8721	1.1137
		(0.1155)	(0.0814)	(0.0651)	(0.0411)	(0.0307)
		4.9560	3.8662	2.8268	1.9874	1.1961
		4.5607	3.5704	2.5986	1.6941	0.9451
4	0.04	0.3147	0.0287	0.0011	0.0005	0.0002
		(0.0091)	(0.0017)	(0.0003)	(0.0002)	(0.0001)
		0.2875	0.0241	0.0005	0.0000	0.0000
		0.2992	0.0250	0.0006	0.0000	0.0000
	0.16	3.9579	2.9268	1.9978	0.9833	0.2690
		(0.1003)	(0.0874)	(0.0605)	(0.0346)	(0.0079)
		4.2054	3.0663	2.1268	1.1274	0.3071
		3.8289	2.8198	1.8135	0.8664	0.2266

Note: All call option prices correspond to a 5-year discount bond with face value of $100. The exercise price is expressed as percentage of the current bond price with corresponding maturities for each model. The four elements of each cell from top to bottom are: the nonparametric price, its standard error (in parentheses) calculated from simulations, and the Vasicek and CIR prices.

largely reduce the variance of the estimates. Overall, the accuracy of the Monte Carlo simulation can be monitored by computing the standard derivations of the estimates. Table 4 also report the standard derivations computed from the simulations for the nonparametric option prices. [16]

The option prices computed from the nonparametric model and the CIR and Vasicek

---

16 Again, accuracy of the Monte Carlo simulation results are examined by comparing the Monte Carlo results with the analytical solutions of the Vasicek model and the numerical solutions of CIR and nonparametric models, and the errors are also found to be small. Numerically solving the European call option prices from PDE (4) consists in a two-stage procedure. First, the equilibrium value of the underlying bond at the maturity date of the option $P(r, T, S)$ is computed by solving the valuation PDE (4) subject to the terminal and boundary conditions of bond prices. Then this value is substituted into the boundary conditions as specified in the Appendix for the option and the valuation PDE solved a second time subject to this latter condition.

models are in general different, with out-of-the-money options showing the largest percentage discrepancies. By eliminating differences in the prices of the underlying bonds, differences in bond option prices can be attributed to differences in the volatility of the underlying bonds. The volatility of the risk-neutral bond price can be derived as $\sigma^2(r(t))(\frac{\partial P}{\partial r})^2$ according to Itô's lemma or equation (2), which is proportional to the diffusion function of the spot interest rate and the sensitivity of bond price to the spot interest rate. The results show that among all the factors which affect option prices of short-term bonds, the diffusion function, or the second moment, of the spot interest rate appears to play the most important role. For low level of interest rate at 4%, the volatility of the nonparametric model is higher than that of the Vasicek model which is in turn higher than that of the CIR model, as a result, option prices show the same order that the nonparametric option prices are higher than the Vasicek prices which in turn are higher than the CIR prices. For high level of interest rate at 16%, the volatility of the CIR model is higher than that of the nonparametric model which is in turn higher than that of the Vasicek model, as a result, the option prices are ranked accordingly. This second-moment effects are more visible for deep-out-of-the-money options. It is noted that for the deep out-of-the-money options, the CIR and Vasicek prices generally fall outside of two or more standard deviations of the nonparametric prices.

6. Conclusion

In this paper, we extended the interest rate term structure model through nonpara-

34

metric specification of the drift function, diffusion function, and the market price of risk. The model precludes any arbitrage opportunities, retains a simple structure and the computational tractability, and allows for maximal flexibility in fitting into the data. Data are allowed to speak for themselves. The model is implemented using the historical Canadian term structure data. Implications of the nonparametric model on term structure dynamics and prices of interest rate derivative securities are investigated through comparison with the parametric models, i.e. the Vasicek (1977) and CIR (1985) models. The empirical results in this paper not only provide strong evidence that the traditional spot interest rate models and market prices of interest rate risk are misspecified, but also suggest that model specification has significant economic impact on the dynamics of term structure and prices of interest rate derivative securities. Since the model can capture the true volatility of the spot interest rates, and the historical term structures can be best fitted by the model with a flexible functional form of market price of risk, such a model can be used for various tests based on market quotes of option prices. For instance, it can be used to test whether the general assumptions made about the market in the pricing framework are valid or not, or equivalently to test whether the market itself is efficient or not. Moreover, the methodology developed in this paper can easily be employed for the pricing of other options or the options based on other financial instruments. Finally, further research on the dynamics of interest rate term structure can be undertaken via a nonparametric or semi-parametric two-factor models.

Figure 1. Historical Canadian T—Bill Rates and Bond Yields:

maturities 1—, 3—, 6—month, 1—, 2—, 3—, 5—, 10—, and 30—year

36
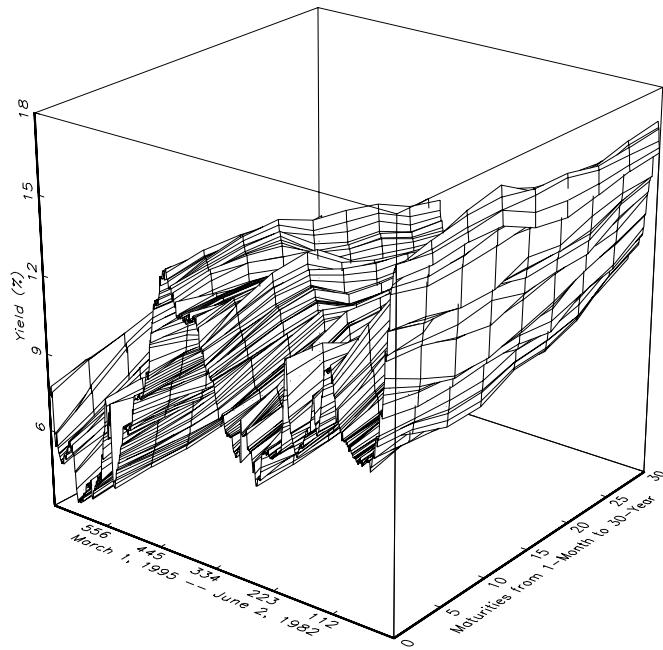
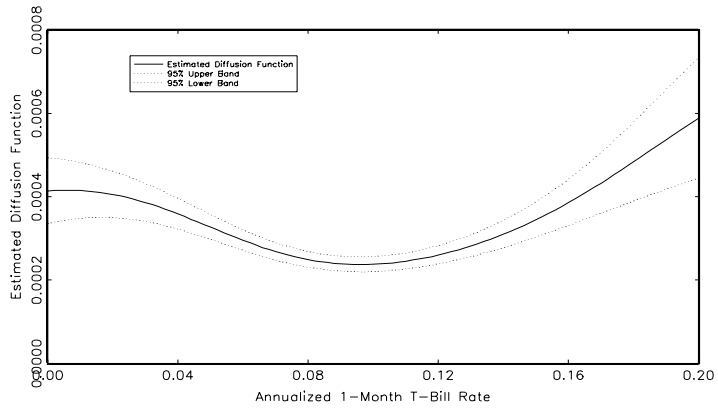Figure 2.1 Nonparametric Diffusion Function

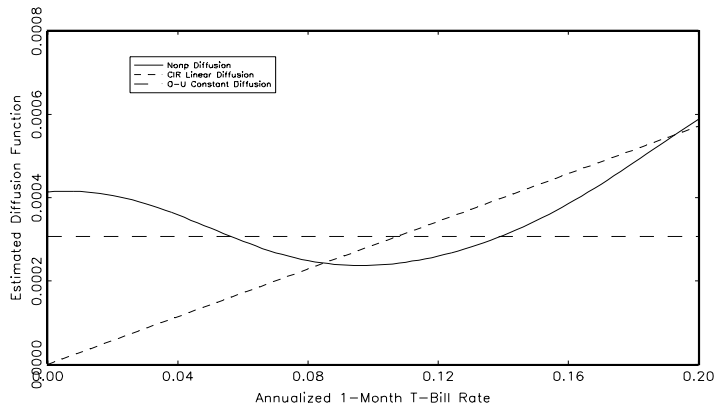Figure 2.2 Alternative Diffusion Function Estimators
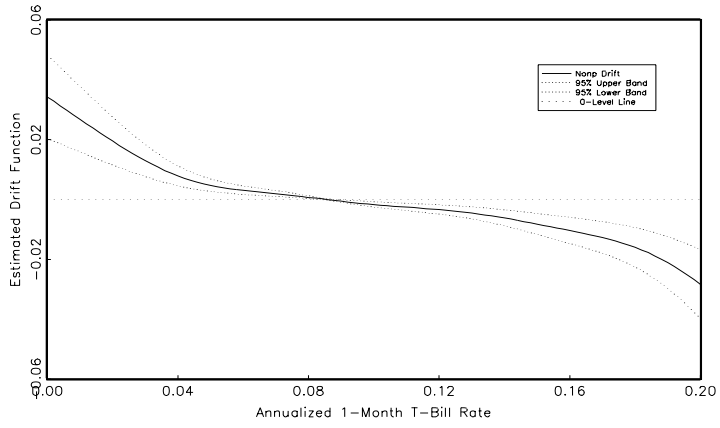
Figure 3.1 Nonparametric Drift Function

Figure 3.2 Alternative Drift Function Estimators
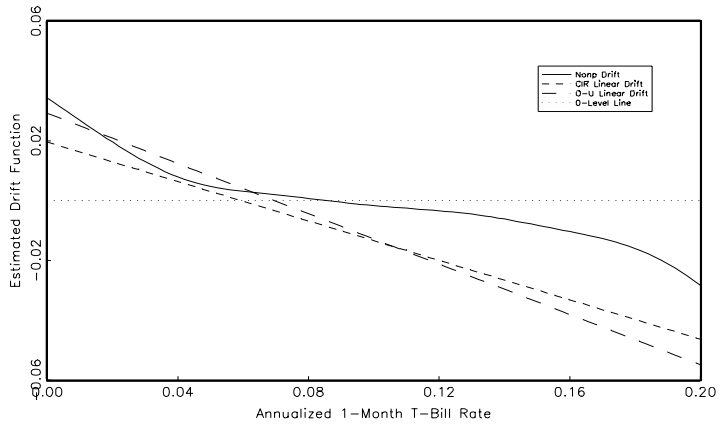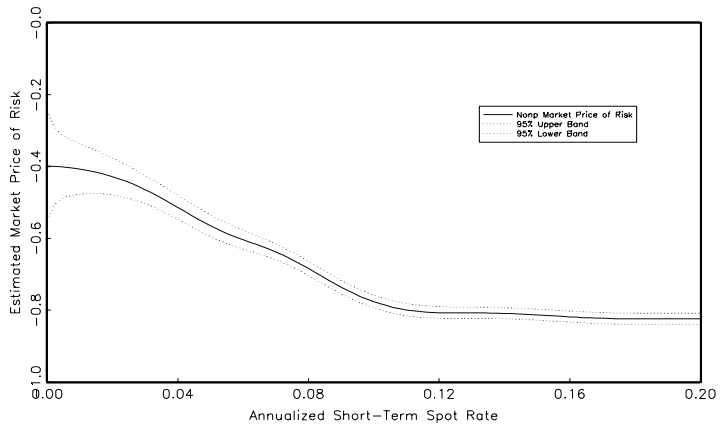

38

Figure 4.1 Nonparametric Market Price of Risk

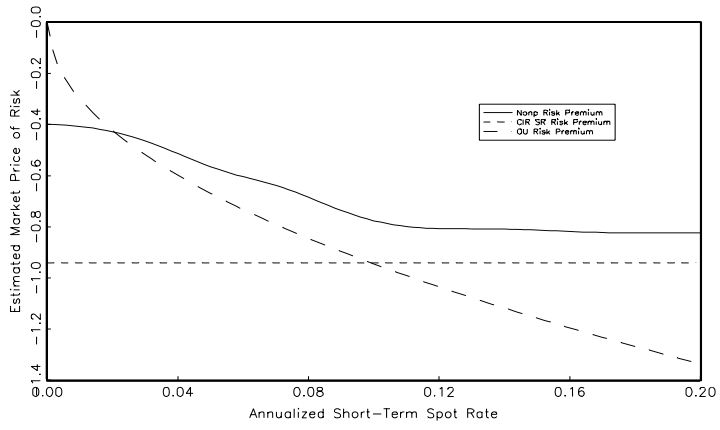
Figure 4.2 Alternative Market Price of Risk Estimators

39

Figure 5.1. Nonp, Vasicek and CIR Term Structures with r(t)=0.04
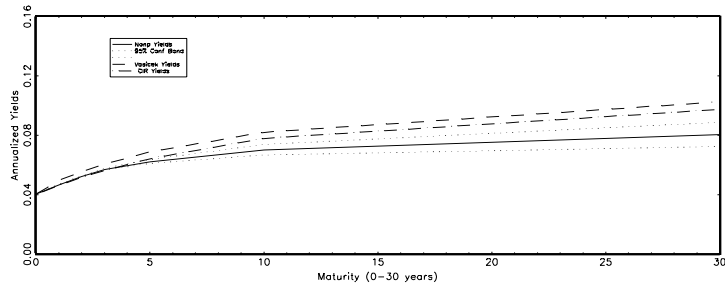
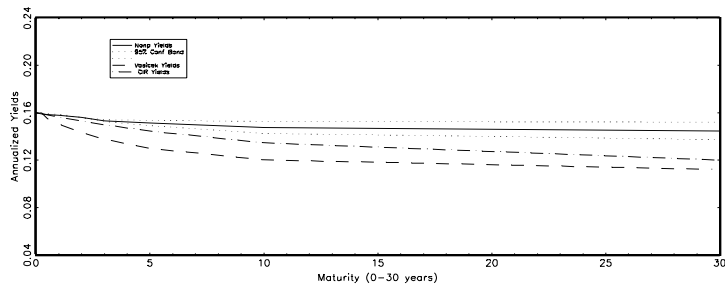Figure 5.2. Nonp, Vasicek and CIR Term Structures with r(t)=0.10
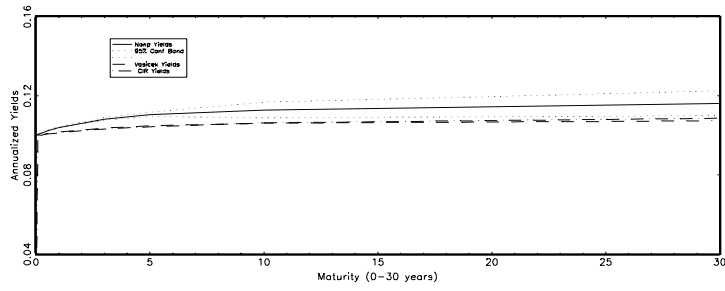
Figure 5.3. Nonp, Vasicek and CIR Term Structures with r(t)=0.16

40

Figure 6.1 Vasicek Model:

Absolute Residuals of Implied Yield Curves
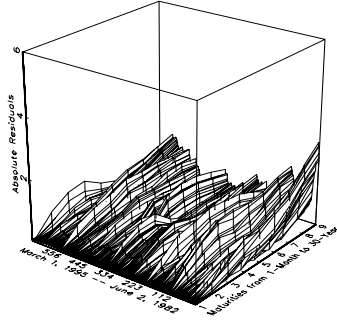
Figure 6.2 CIR Model:
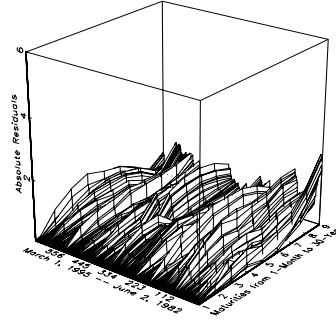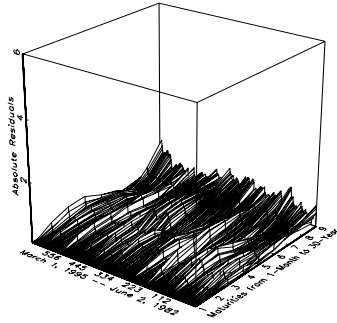
Absolute Residuals of Implied Yield Curves

Figure 6.3 NONP Model:

Absolute Residuals of Implied Yield Curves

41

Appendix:

1. The Ornstein-Uhlenbeck Process: The transitional density function of the O-U process is a Gaussian kernel with conditional mean and variance $E[r(t+\tau)|r(t)] = r(t) + (1 - e^{-\beta\tau})(\alpha - r(t))$, $Var[r(t+\tau)|r(t)] = \frac{\sigma^2}{2\beta}(1 - e^{-2\beta\tau})$. In order for the process to be stationary, the initial marginal density function must be set equal to the final limiting density function of the process, i.e., the marginal density function of the O-U process is also a Gaussian kernel with mean and variance $E[r(t)] = \alpha$, $Var[r(t)] = \frac{\sigma^2}{2\beta}$, the covariance of $r(t+\tau)$ and $r(t)$ can be calculated from

$$
\begin{aligned}
Cov(r(t+\tau), r(t)) &= E[r(t+\tau)r(t)] - \alpha^2 \\
&= E_{r(t)}[r(t+\tau)r(t)|r(t)] - \alpha^2 \\
&= E_{r(t)}[r(t)[r(t) + (1 - e^{-\beta\tau})(\alpha - r(t))]] - \alpha^2 \\
&= \frac{\sigma^2}{2\beta}e^{-\beta\tau}
\end{aligned}
$$

And hence $Corr(r(t+\tau), r(t)) = e^{-\beta\tau}$. For δ-period difference of the O-U process, $\Delta_\delta r(t) = r(t) - r(t-\delta)$, we have $E[\Delta_\delta r(t)] = 0$,

$$
\begin{aligned}
Var[\Delta_\delta r(t)] &= Var[r(t)] + Var[r(t-\delta)] - 2Cov(r(t), r(t-\delta)) \\
&= \frac{\sigma^2}{\beta}(1 - e^{\beta\delta})
\end{aligned}
$$

and for $\tau \geq \delta$,

$$
\begin{aligned}
Cov(\Delta_\delta r(t+\tau), \Delta_\delta r(t)) &= E[(r(t+\tau) - r(t+\tau-\delta))(r(t) - r(t-\delta))] \\
&= -\frac{\sigma^2}{2\beta}e^{-\beta(\tau-\delta)}(1 - e^{-\beta\delta})^2
\end{aligned}
$$

or $Corr(\Delta_\delta r(t+\tau), \Delta_\delta r(t)) = -\frac{1}{2}e^{-\beta(\tau-\delta)}(1 - e^{-\beta\delta})$.

2. The CIR squared-root Process: The transitional density function of the CIR process is a non-central χ^2 distribution with conditional mean and variance $E[r(t+\tau)|r(t)] = r(t) + (1 - e^{-\beta\tau})(\alpha - r(t))$, $Var[r(t+\tau)|r(t)] = \frac{\sigma^2}{\beta}(e^{-\beta\tau} - e^{-2\beta\tau})r(t) + \alpha(\frac{\sigma^2}{2\beta})(1 - e^{-\beta\tau})^2$. Due to the same reason as for the O-U process, the marginal density function of the CIR process is a Gamma distribution with mean and variance $E[r(t)] = \alpha$, $Var[r(t)] = \frac{\alpha\sigma^2}{2\beta}$. Similarly, the covariance of $r(t)$ and $r(t)$ can be calculated from

$$
\begin{aligned}
Cov(r(t+\tau), r(t)) &= E[r(t+\tau)r(t)] - \alpha^2 \\
&= E_{r(t)}[E[r(t+\tau)r(t)|r(t)]] - \alpha^2 \\
&= E_{r(t)}[r(t)[r(t) + (1 - e^{-\beta\tau})(\alpha - r(t))]] - \alpha^2 \\
&= \frac{\alpha\sigma^2}{2\beta}e^{-\beta\tau}
\end{aligned}
$$

And $Corr(r(t+\tau), r(t)) = e^{-\beta\tau}$. For δ-period difference of the CIR process, $\Delta_\delta r(t) = r(t) - r(t-\delta)$, we have $E[\Delta_\delta r(t)] = 0$,

42

$$Var[\Delta_\delta r(t)] = Var[r(t)] + Var[r(t - \delta)] - 2Cov(r(t), r(t - \delta))$$
$$= \frac{\alpha\sigma^2}{\beta}(1 - e^{\beta\delta})$$

and for $\tau \geq \delta$,

$$Cov(\Delta_\delta r(t + \tau), \Delta_\delta r(t)) = E[(r(t + \tau) - r(t + \tau - \delta))(r(t) - r(t - \delta))]$$
$$= -\frac{\alpha\sigma^2}{2\beta}e^{-\beta(\tau-\delta)}(1 - e^{-\beta\delta})^2$$

or $Corr(\Delta_\delta r(t + \tau), \Delta_\delta r(t)) = -\frac{1}{2}e^{-\beta(\tau-\delta)}(1 - e^{-\beta\delta})$.

3. Choices of Kernel Function and Window-Width for Diffusion Function and Drift Function Estimation: The regularity condition of the kernel function of order r for both diffusion function and drift function estimation are as follows:

(i) The kernel $K(\cdot)$ is symmetric about zero, continuously differentiable to order r on R, belongs to $L^2(R)$, and $\int_{-\infty}^{+\infty} K(x)dx = 1$;

(ii) $K(\cdot)$ is of order r: $\int_{-\infty}^{+\infty} x^i K(x)dx = 0, i = 1, ..., r - 1$, and $\int_{-\infty}^{+\infty} x^r K(x)dx \neq 0$, $\int_{-\infty}^{+\infty} |x|^r |K(x)|\, dx < \infty$.

The regularity conditions for the admissible window-width are as follows: as the sample size $n \to \infty$, and the sampling interval $\Delta_n(= 1/ln(n)) \to 0$,

(i) $h_n \to 0, nh_n \to \infty$, and $nh_n^{r+1} \to 0$ for the diffusion function estimation;

(ii) $h_n \to 0, nln(n)h_n \to \infty$, and $nh_n^{2r+1} \to 0$ for the marginal density function estimation; and

(iii) $h_n \to 0, nln(n)h_n^3 \to \infty$, and $nh_n^{2r+1} \to 0$ for the first derivative of the marginal density function estimation, both the marginal density function and its derivative are required in the drift function estimation.

The above conditions ensure that for all cases, the bias in the estimator is asymptotically negligible and at the same time the variance of the estimator goes to zero as sample size increases to infinity. The actual choice of the kernel for all cases is the standard Gaussian kernel with order $r = 2$, i.e $K(x) = \frac{1}{\sqrt{2\pi}}e^{-x^2/2}$. The window-width sequence chosen is $h_n = c_n n^{-1/3}$ for the diffusion function estimation and $h_n = c_n n^{-1/5}$ for the estimation of both the marginal density function and its first derivative, where $c_n = c/ln(n)$, and c is chosen to minimize the IMSE of each functional estimator. Exponential and Epanechnikov kernels were also used for the estimation and produced very similar results. The estimation results are also robust to small changes of the window-width around the optimal values and different time series of short-term interest rates.

4. Choice of Kernel Function and Smoothing Parameter for Market Price of Risk: Choice of the smoothing kernel for the market price of risk is also the standard Gaussian kernel. The algorithm of obtaining the optimal smoothing parameter which minimizes the IMSE of $\hat{\lambda}(r)$ using cross-validation method is as follows (see, e.g. Härdle and Vieu (1987):

Step 1: Compute the leave-one estimate

$$\hat{\lambda}_{h_n,j}(r_j)$$

at the observation points r_j, $j = 1, 2, ..., n$;

Step 2: Construct the cross validation function

$$CV(h_n) = \frac{1}{n}\sum_{j=1}^{n}(\lambda_j - \hat{\lambda}_{h_n,j}(r_j))^2 w(r_j)$$

where w denotes a weight function which is set as a constant in our estimation;

Step 3: Find the optimal window-width as

$$\hat{h} = argmin_{\{h_n\}}|CV(h_n)|$$

5. GMM Estimation of the Vasicek (1977) and CIR (1985) Models: The GMM estimates of $(\alpha, \beta, \sigma^2)$ for the Vasicek (1977) and CIR (1985) models are obtained from the following four exact (first and second order) moment conditions:

$$G_N(\alpha, \beta, \sigma^2) = \frac{1}{N-1}\sum_{i=1}^{N-1} F_i(\alpha, \beta, \sigma^2)$$

with

$$F_i(\alpha, \beta, \sigma^2) = \begin{bmatrix} \epsilon_{i+1} \\ \epsilon_{i+1}r_i \\ \epsilon_{i+1}^2 - E[\epsilon_{i+1}^2|r_i] \\ (\epsilon_{i+1}^2 - E[\epsilon_{i+1}^2|r_i])r_i \end{bmatrix}$$

where $\epsilon_{i+1} = (r_{i+1} - r_i) - E[(r_{i+1} - r_i)|r_i]$ with $E[(r_{i+1} - r_i)|r_i] = (1 - e^{-\beta\Delta t_i})(\alpha - r_i)$ for both models, Δt_i is the ith sampling interval. The exact conditional variance of interest rate changes over time interval of length Δt_i is given by $E[\epsilon_{i+1}^2|r_i] = V[r_{i+1}|r_i]$, with

$$E[\epsilon_{i+1}^2|r_i] = (\sigma^2/2\beta)(1 - e^{-2\beta\Delta t_i})$$

for the Vasicek model; and

$$E[\epsilon_{i+1}^2|r_i] = (\sigma^2/\beta)(e^{-\beta\Delta t_i} - e^{-2\beta\Delta t_i})r_i + (\sigma^2/2\beta)(1 - e^{-\beta\Delta t_i})^2\alpha$$

for the CIR model. These moment conditions correspond to transitions of length Δt_i and are not subject to discretization bias. In our estimation, $\Delta t = 1/52$ for weekly data and $\Delta t = 1/252$ for daily data. Since these GMM systems are overidentified, we weighted the

criterion optimally (see Hansen (1982)).

It is noted that in most financial economics literature, parameter estimation of the diffusion processes using GMM technique consists in first discretizing the continuous-time diffusion process, then based on the discrete model deriving moment conditions. The GMM approach no longer requires that the distribution of interest rate changes be normal, but only requires that the conditional instantaneous variance of the residual is proportional to the length of sampling interval, i.e., $E[\epsilon_{i+1}^2 | r_t] = \sigma^2(\cdot)\Delta t_i$; the asymptotic justification for the GMM procedure requires only that the distribution of interest rate changes be stationary and ergodic and that the relevant expectations exist. The moment conditions are as follows:

$$F_i(\alpha, \beta, \sigma^2) = \begin{bmatrix} \epsilon_{i+1} \\ \epsilon_{i+1} r_i \\ \epsilon_{i+1}^2 - \sigma^2 r_i^\gamma \Delta t_i \\ (\epsilon_{i+1}^2 - \sigma^2 r_i^\gamma \Delta t_i) r_i \end{bmatrix}$$

with $\gamma = 0$ for the Vasicek model, and $\gamma = 1$ for the CIR model, with $\epsilon_{i+1} = r_{i+1} - r_i - \beta(\alpha - r_i)\Delta t_i$ where $\Delta t_i = t_{i+1} - t_i$, and $\Delta t_i = T_0/N$ in the case of equispaced sampling interval. Misspecification of the model and biasedness or inconsistency of the parameter estimators due to "discretization" are known facts in the literature.

6. Final and Boundary Conditions for Bond and Bond Option Prices: The initial (or final) condition and boundary conditions for bond price and bond option price are as follows:

Let $P(r, t, T)$ be the price of a pure discount bond with face value 100 and maturity date T when the spot interest rate is r at date t . It corresponds to the following coupon payment or dividend payment and initial (or final) and boundary conditions: $d(r, t) = 0$(no payment), $P(r, T, T) = 100$ for all $r \geq 0$ (final condition), and $\lim_{r \to +\infty} P(r, t, T) = 0$ for all $0 \leq t \leq T$ (boundary condition).

For a call option which expires at date T, has an exercise (or strike) price K, and the underlying discount bond matures at date S where $T \leq S$. Let $P(r, t, T; S, K)$ be its price at date t when the spot interest rate is r. It corresponds to the following coupon payment or dividend payment and initial (or final) and boundary conditions: $d(r, t) = 0$ (no payment), $P(r, T, T; S, K) = \max(0, P(r, T, S) - K)$ for all $r \geq 0$ (final condition), $\lim_{r \to +\infty} P(r, t, T; S, K) = 0$ for all $0 \leq t \leq T$ (boundary condition).

References:

Aït-Sahalia, Y., 1996a, "Nonparametric Pricing of Interest Rate Derivative Derivative Securities," *Econometrica* Vol.64, No.3, 527-60.

Ames, W.F., 1977, *Numerical Methods for Partial Differential Equations,* New York: Academic Press.

Backus, D.K., S. Foresi, and S.E. Zin, 1995, "Arbitrage Opportunities in Arbitrage-free Models of Bond Pricing," Working paper, New York University.

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*, January-February, 33-39.

Black, F. and P. Karasinski, 1991, "Bond and Option Pricing When Short Rates are Lognormal," *Financial Analysts Journal* (July/August), 52-59.

Boyle, P., M. Broadie, & P. Glasserman, 1996, "Monte Carlo methods for security pricing," forthcoming in *Journal of Economics Dynamics and Control.*

Brennan, M.J. & E.S. Schwartz, 1979, "A continuous time approach to the pricing of bonds," *Journal of Banking and Finance* 3, 133-155.

Brown, S.J. and P.H. Dybvig, 1986, "The Empirical Implications of the Cox, Ingersoll and Ross Theory of the Term Structure of Interest Rates," *Journal of Finance*, 41(3), 617-632.

Brown, R. and S. Schaefer, 1991, "Interest Rate Volatility and the Term Structure," Working paper, London Business School, London.

Canabarro, E., 1995, "Where Do On-Factor Interest Rate Models Fail?" *Journal of Fixed Income*, 5(2), 31-52.

Chan, K.C., G.A. Karolyi, F.A. Longstaff and A.B. Sanders, 1992, "An Empirical Comparison of Alternative Models of the Short-Term Interest Rate," *The Journal of Finance* 47, 1209-1227.

Chaplin, G.B. and K.P. Sharp, 1993, "Analytic Solutions for Bonds and Bond Options under n Correlated Stochastic Processes", Working paper 93-16, University of Waterloo Institute of Insurance and Pension Research, Waterloo, Canada.

46

Chen, R.R. and L.O. 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.

Constantinides, G.M., 1992, "A Theory of the Nominal Term Structure of Interest Rates," *Review of Financial Studies*, 5(4), 531-552.

Courtadon, G., 1982, "The Pricing of Option on Default Free Bond," *Journal of Financial and Quantitative Analysis* 17, 75-100.

Cox, J.C., 1975, "Notes on option pricing I: constant elasticity of variance diffusion," Working Paper, Stanford University.

Cox, J.C., J.E. Ingersoll & S.A. Ross, 1980, "An analysis of variable rate loan contracts," *Journal of Finance* Vol. XXXV, No.2, 389-403.

......, 1985, "A Theory of the Term Structure of Interest Rates," *Econometrica* Vol.53, No.2, 385-407.

Dothan, L.U., 1978, "On the Term Structure of Interest Rates", *Journal of Financial Economics* 6, 59-69.

Duffie, D., 1988, *Security Markets: Stochastic Models.* Academic Press, Boston, MA.

......, 1992, *Dynamic Asset Pricing Theory*, Princeton University Press, Princeton, NJ.

Duffie, D. and R. Kan, 1993, "A Yield Curve Model of Interest Rates," Working paper, Stanford University.

Friedman, A., 1964, *Partial Differential Equations of Parabolic Type,* Englewood Cliff, N.J.: Prentice-Hall.

Gibbons, M.R. and K. Ramaswany, 1993, "A Test of the Cox, Ingersoll and Ross Model of the Term Structure," *Review of Financial Studies,* 6(3), 619-658.

Hansen, L.P., 1982, "Large sample properties of generalized method of moments estimators," *Econometrica* 50, 1029-1054.

Härdle, W., 1990, *Applied Nonparametric Regression*, Cambridge, UK: Cambridge University Press.

Härdle, W. and Vieu, P. 1987, "Nonparametric Prediction by the Kernel Method," Mimeo, Rheinishe-Friedrich-Wilhelms Universitat.

Harvey, A.C., 1993, *Time Series Models* 2nd edition. Cambridge: The MIT Press.

Heath, D.C. R.A. Jarrow, and A. Morton, 1992, "Bond Pricing and the Term Structure of Interest Rates: A New Methodology for Contingent Claims Valuation, " *Econometrica* 60, 77-105.

Ho, T.S.Y. and S.B. Lee, 1986, "Term Structure Movements and Pricing Interest Rate Contingent Claims," *The Journal of Finance* 41, 1011-1029.

Hull, J. and A. White, 1990, "Pricing Interest-Rate-Derivate Securities," *The Review of Financial Studies* Vol. 3, No.4, 573-92.

......, 1993, "One Factor Interest Rate Models and the Valuation of Interest Rate Derivative Securities," *Journal of Financial and Quantitative Analysis* 28, 235-254.

Jamshidian, F., 1989, "An Exact Bond Option Formula," *The Journal of Finance* Vol.XLIV, No.1, 205-209.

......, 1990, "The Preference-Free Determination of Bond and Option Prices from the Spot Interest Rate", *Advances in Futures and Options Research* 4, 51-67.

......, 1991, "Bond and Option Valuation in the Gaussian Interest Model," *Research in Finance* 9, 131-170.

Jiang, G.J. and J.L. Knight, 1995, "A Nonparametric Approach to the Estimation of Diffusion Processes," forthcoming in *Econometric Theory*.

Jiang, G.J., 1996, "Nonparametric Estimation of Diffusion Processes and Applications in Pricing Derivative Securities", Ph.D Dissertation, University of Western Ontario, Canada.

Kim, T.Y. and D.D. Cox, 1996, "Bandwidth Selection in Kernel Smoothing of Time Series," *Journal of Time Series Analysis,* 17(1), 49-63.

Kraus, A. and M. Smith, 1993, "A Simple Multi-Factor Term Structure Model," *The Journal of Fixed Income*, June, 19-23.

Künsch, H.R., 1989, "The Jackknife and the Bootstrap for General Stationary Observations," *Annals of Statistics* Vol.17, No.3, 1217-41.

Langetieg, T.C., 1980, "A Multivariate Model of the Term Structure," *Journal of Finance* 35, 71-97.

48

Litterman, R. and J. Scheinkman, 1991, "Common Factors affecting Bond Returns," *Journal of Fixed Income,* 1, 54-61.

Liu, R.Y.K. and Singh, 1992, "Moving Blocks Jackknife and Bootstrap Capture Weak Dependence," in LePage R. and Billard, Y. eds, *Exploring the Limits of Bootstrap,* Wiley, New York.

Longstaff, F.A. 1989, "A Nonlinear General Equilibrium Model of the Term Structure of Interest Rates," *Journal of Financial Economics* 23, 195-224.

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* 47, September 1992.

Marsh, T.A. and E.R. Rosenfeld, 1983, "Stochastic Processes for Interest Rates and Equilibrium Bond Prices", *The Journal of Finance* Vol.XXXVIII, No.2, 635-46.

Merton, R.C., 1973, "Theory of Rational Option Pricing," *Bell Journal of Economics and Management Science* 4, 141-83.

Phillips, P.C.B., 1987, "Time series regression with a unit root," *Econometrica* 55(2), 277-310.

Prakasa-Rao, B.L.S., 1983, *Nonparametric Functional Estimation,* Academic Press.

Ritchken, P. and K. Boenawan, 1990, "On Arbitrage-Free Pricing of Interest Rate Contingent Claims," *The Journal of Finance* 45(1), 259-64.

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, 413-424.

Stanton, R., 1996, "A nonparametric model of term structure dynamics and the market price of interest rate risk," Working paper, Haas School of Business, U.C. Berkeley.

Talay, D., 1996, "Probabilistic numerical methods for PDEs: Elements of analysis," in D. Talay and L. Tubaro eds., *Probabilistic Methods for Nonlinear Partial Differential Equations,* Berlin: Springer.

Vasicek, O., 1977, "An Equilibrium Characterization of the Term Structure", *Journal of Financial Economics* 5, 177-88.

Vetzal, K.R., 1994, "A Survey of Stochastic Continuous Time Models of the Term Structure of Interest Rates", *Insurance: Mathematics and Economics* 14, 139-161.

Wong, E., 1971, *Stochastic Processes in Information and Dynamical Systems,* McGraw-Hill, New York.