No. 918

Modeling the Term Structure of Interest Rates with General Short-Rate Models

by

Hideyuki Takamizawa and Isao shoji

April 2001

Modeling the Tern Structure of Interest Rates with General Short-Rate Models

Hideyuki Takamizawa and Isao Shoji

Institute of Policy and Planning Sciences

University of Tsukuba

Tsukuba Ibaraki 305-8573, Japan

TEL:81-298-53-5094

FAX:81-298-53-5094

e-mail:shoji@sk.tsukuba.ac.jp

April 6, 2001

Modeling the Tern Structure of Interest Rates with General Short-Rate Models

Abstract

We propose models of the term structure of interest rates with general processes of the short-rate. As documented by many studies, it is important to model the short-rate generally to capture its observed behavior. This adequate specification is essential for term structure models to explain yields accurately. With such general conditions, however, term structure models cannot be obtained explicitly. In this paper we derive an analytical model of the term structure by locally approximating a short-rate process which has arbitrarily specified drift and diffusion term. The empirical analysis confirms that our model outperforms the existing affine term structure model.

1 Introduction

In this paper, we derive an analytical model of the term structure of interest rates, allowing for a general process of the short-rate. The reason we model the short-rate generally is to capture its actual behavior which is consistent with empirical findings by many studies. For example, Chan, Karolyi, Longstaff, and Sanders (hereafter CKLS) (1992) introduce a more general short-rate model which has the volatility proportional to r^{γ} while γ is a free parameter. This model nests most of the well-known models and empirical performance of these alternative models is examined. Their major findings, using data on the U.S. interest rate, is that the volatility of short-rate changes is highly elastic to the level of the short-rate; the estimate of γ is about 1.5. This implies that the specification of the volatility is important in order to capture the observed behavior of the short-rate hence that we need to mode the diffusion term generally. Similar results are also presented by other studies.

Moreover, recent works which employ various statistical techniques find nonlinearity in the drift term as well as in the diffusion term. For example, Aït-Sahalia (1996) proposes test statistics to measure "distances" between a marginal distribution implied by a parametric model and that estimated directly by the nonparametric method. Conley et al. (1997) use efficient test functions obtained from infinitesimal changes in the stationary density of the short-rate to examine its behavior. Stanton (1997) derives the approximate formula for the drift and diffusion function and estimates their functional forms nonparametrically. One of the key common findings of these studies is that the short-rate behaves like a random walk in its mean range but that it reverts to the range rapidly when it is at extremely high or low levels. In other words, the degree of mean reversion of the short-rate depends on its level, hence the drift term is highly nonlinear. [1] Therefore, we also need

to model the drift term generally.

It is natural to think that adequately specified short-rate models are essential for term structure models to explain yields accurately since by construction the term structure depends on the drift and diffusion term of a short-rate process. However, existing models of the term structure do not consider actual behavior of the short-rate. For example, in affine term structure models proposed by Vasicek (1977) and Cox, Ingersoll, and Ross (hereafter CIR) (1985), both the drift and the local variance of short-rate changes are assumed to be linear in the short-rate; see Duffie and Kan (1995), and Dai and Singleton (2000) for more profound discussion on affine models. However, as discussed by many studies, the affine specification is restrictive for actual data. This problem remains unsolved in nonlinear term structure models proposed by Longstaff (1989) and Beaglehole and Tenney (1992). Their short-rate models have the volatility which is proportional to square root of the short-rate. Besides, although each drift term is nonlinear, the sharp mean reversion observed at extreme levels cannot be modeled with their specification. A more recent nonlinear model proposed by Ahn and Gao (1999) assume the volatility to be proportional to r^{γ} while $\gamma = 1.5$, which is consistent with the empirical findings in CKLS (1992). However, this elasticity parameter γ may differ in data of various countries; see, for example, Dahlquist (1996), Nowman (1997), and Shoji and Ozaki (1996). Thus fixing this parameter is not good enough to derive term structure models which explain actual yields in many countries.

The main drawback of all these term structure models is restricted assumption on behavior of the short-rate. As noted before, this poor specification leads to the lack of descriptive power of term structure models. Therefore we need a term structure model, allowing for a generally specified process of the short-rate by which we incorporate its actual behavior. However, with such general conditions, an analytical model of the term

structure cannot be obtained since the partial differential equation for a discount bond price has no closed-form solution. That is, we always face a fundamental trade-off between general specification and analytical tractability.

In this paper, we propose a model of the term structure of interest rates which solves this trade-off. The advantage of our model is that we can capture actual behavior of the short-rate by specifying its process generally and that we obtain an analytical solution for the term structure of interest rates. The key technique for deriving a model is a local linearization method (LLM) developed by Shoji and Ozaki (1997) and Shoji (1998). We apply the LLM to the short-rate process by which arbitrary specified mean and variance are locally approximated by linear functions. With the approximated process, we can solve the partial differential equation for a discount bond price. To examine the descriptive power of our model, we compare the empirical performance between the CIR model and ours. Our empirical analysis consists of two steps. The first step is to select an appropriate model of the short-rate among models studied in the existing literature by using in-sample data. The second step is to compare pricing errors of term structure models by using out-of-sample data.

The rest of this paper is organized as follows. Section 2 introduces our model. Section 3 discusses appropriate models for dynamics of the short-rate. Section 4 provides empirical results of performance comparison between the CIR and our term structure model based on mean squared pricing errors. Section 5 concludes.

2 The Model

For deriving our model, we follow the general equilibrium framework in CIR (1985). CIR endogenously derive the short-rate and factor risk premiums, both of which determine

price of any contingent claims with payoff conditions. In order to derive their closed-form model of the term structure of interest rates, several conditions are assumed. Particularly, the following assumptions are important: (i) preference of the representative consumer is logarithmic, (ii) a single factor X governing stochastic nature of the economy follows the square-root process, (iii) the means and variances of rates of return on production technologies are proportional to X. By the assumption (iii), the short-rate is derived as a proportional function of a state as $r = cX$, which implies that the short-rate also follows the square-root process by the assumption (ii). However, as documented by many empirical studies, the square-root model is not adequate to describe actual behavior of the short-rate. Hence, we do not assume (ii) in our model. Instead, we specify the drift and diffusion term of a short-rate process generally while holding the proportional relation, $r = cX$. Since by this relation the stochastic property of the short-rate is the same as that of a state variable, our model begins with the dynamics of the short-rate given by an arbitrary specified stochastic differential equation (SDE) as,

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

where $\{W_t, 0 \leq t \leq T\}$ is one-dimensional standard Brownian motion, and where μ and σ are drift and diffusion functions defined on $\mathbf{R}_+ \times [0,T]$ into $\mathbf{R}$ and $\mathbf{R}_+$, respectively. We assume that both μ and σ are well specified so that the solution of the SDE (1) exists for an arbitrary position of the initial state.

Next, we denote by $P(r,t,T)$ price of a default-free discount bond at time t maturing in time T. CIR derive the fundamental valuation equation for price of a contingent claim which holds in the general equilibrium (CIR (1985), p.390). In this case, a discount bond pays 1 (dollar) only at maturity date with probability 1. Thus,

$$\tfrac{1}{2}\sigma^2(r,t)\frac{\partial^2 P}{\partial r^2}(r,t,T) + \{\mu(r,t) - \lambda(r,t)\}\frac{\partial P}{\partial r}(r,t,T) + \frac{\partial P}{\partial t}(r,t,T) - rP(r,t,T) = 0 \,, \tag{2}$$

$$P(r,T,T) = 1\,,$$

where $\lambda(r,t)$ is the risk premium function, which is also derived endogenously in equilibrium.

It is important to note that from equation (2) price of a discount bond is fundamentally determined by the three functions, μ, σ and λ. Since λ is determined endogenously by following the CIR framework, it is particularly important to provide adequate specification for μ and σ for valuing a discount bond accurately.

If we want to obtain the solution without restriction on μ and σ, we need to employ numerical methods. But the numerical solution has the following drawback. In pricing a discount bond by a model, we need to estimate its parameters. Since the pricing equation is given only by the from of a partial differential equation, we have to run two algorithms simultaneously, the one for solving the partial differential equation and the other for estimating parameters. These procedures not only require intensive computation but have no guarantee of convergence of the solution. Thus the numerical solution is not practical in this situation.

On the other hand, our model has two advantages. First, we specify μ and σ generally and determine their adequate functional forms by data. Second, we derive a closed-form model of the term structure, which reduces computational time substantially in estimating parameters as well as in calculating yields. The key technique for deriving our model is a local linearization method (LLM) developed by Shoji and Ozaki (1997) and Shoji (1998). By applying the LLM to the process, we locally approximate the mean and variance of changes in the short-rate by linear functions in r and t. With the approximated process the partial differential equation can be solved explicitly, that is, a closed-form model of the term structure is obtained.

Actually, the LLM is applied to the risk adjusted process of the short-rate. Note that from equation (2) we can interpret that a discount bond is priced in the following stochastic system.

$$\begin{aligned} dr &= (\mu(r,t) - \lambda(r,t))\,dt + \sqrt{\sigma^2(r,t)}\,d\tilde{W}_t \\ &\equiv \hat{\mu}(r,t)dt + \sqrt{\sigma^2(r,t)}\,d\tilde{W}_t\,, \end{aligned} \tag{3}$$

where

$$\tilde{W}_t = W_t + \int^t \frac{\lambda(r_u,u)}{\sigma(r_u,u)}du\,.$$

By applying the LLM, $\hat{\mu}$ and σ^2 are replaced by linear functions in r and t. Under the risk adjusted process of the short-rate, the partial differential equation of a discount bond price is obtained as

$$\tfrac{1}{2}\sigma^2(r,t)\frac{\partial^2 P}{\partial r^2}(r,t,T) + \hat{\mu}(r,t)\frac{\partial P}{\partial r}(r,t,T) + \frac{\partial P}{\partial t}(r,t,T) - rP(r,t,T) = 0\,. \tag{2'}$$

Therefore, coefficients of partial derivatives of P become linear by the approximation to the process, thus we have an analytical solution for a discount bond price.

The LLM is applied as follows. First, we approximate the drift in equation (3). By applying Ito formula to $\hat{\mu}(r,t)$, we have

$$\begin{aligned} d\hat{\mu}(r,t) &= \frac{\partial\hat{\mu}}{\partial r}(r,t)dr + \tfrac{1}{2}\frac{\partial^2\hat{\mu}}{\partial r^2}(r,t)(dr)^2 + \frac{\partial\hat{\mu}}{\partial t}(r,t)dt \\ &= \frac{\partial\hat{\mu}}{\partial r}(r,t)dr + \left(\tfrac{1}{2}\frac{\partial^2\hat{\mu}}{\partial r^2}(r,t)\sigma^2(r,t) + \frac{\partial\hat{\mu}}{\partial t}(r,t)\right)dt\,. \end{aligned} \tag{4}$$

Integrating both sides of equation (4) on $[t,u]$ gives

$$\int_t^u d\hat{\mu}(r,s) = \int_t^u \frac{\partial\hat{\mu}}{\partial r}(r,s)dr_s + \int_t^u \left(\tfrac{1}{2}\frac{\partial^2\hat{\mu}}{\partial r^2}(r,s)\sigma^2(r,s) + \frac{\partial\hat{\mu}}{\partial t}(r,s)\right)ds\,.$$

By assuming that the first order derivatives in r and t and the second order derivative in

r multiplied by $\sigma^2(r,t)$ are locally constant during this time interval, we obtain

$$\hat{\mu}(r,u) - \hat{\mu}(r,t) = \frac{\partial \hat{\mu}}{\partial r}(r,t)(r_u - r_t) + \left(\tfrac{1}{2}\frac{\partial^2 \hat{\mu}}{\partial r^2}(r,t)\sigma^2(r,t) + \frac{\partial \hat{\mu}}{\partial t}(r,t) \right)(u-t) \, .$$

Rearranging terms gives

$$\hat{\mu}(r,u) = a_2(t) r_u + a_1(t) u + a_0(t) \, , \tag{5}$$

where

$$\begin{aligned} a_2(t) &\equiv \frac{\partial \hat{\mu}}{\partial r}(r,t) \, , \\ a_1(t) &\equiv \tfrac{1}{2}\frac{\partial^2 \hat{\mu}}{\partial r^2}(r,t)\sigma^2(r,t) + \frac{\partial \hat{\mu}}{\partial t}(r,t) \, , \\ a_0(t) &\equiv \hat{\mu}(r,t) - a_2(t) r_t - a_1(t) t \, . \end{aligned}$$

Similarly, given information on time t we locally approximate the conditional variance of changes in the short-rate, $\sigma^2(r,u)$, by a linear function as

$$\sigma^2(r,u) = b_2(t) r_u + b_1(t) u + b_0(t) \, , \tag{6}$$

where

$$\begin{aligned} b_2(t) &\equiv \frac{\partial \sigma^2}{\partial r}(r,t) \, , \\ b_1(t) &\equiv \tfrac{1}{2}\frac{\partial^2 \sigma^2}{\partial r^2}(r,t)\sigma^2(r,t) + \frac{\partial \sigma^2}{\partial t}(r,t) \, , \\ b_0(t) &\equiv \sigma^2(r,t) - b_2(t) r_t - b_1(t) t \, . \end{aligned}$$

Note that approximation to the conditional variance by the LLM is valid until the linearized function hits zero. Thus, we obtain the approximated process of the short-rate in the risk neutral probability as

$$dr_u = \{a_2(t) r_u + a_1(t) u + a_0(t)\} \, du + \sqrt{b_2(t) r_u + b_1(t) u + b_0(t)} \, d\tilde{W}_u \, , \tag{7}$$

and the partial differential equation valuing a discount bond as

$$\frac{1}{2}\left\{b_2(t)r_u + b_1(t)u + b_0(t)\right\}\frac{\partial^2 P}{\partial r^2}(r,u,T)$$
$$+\left\{a_2(t)r_u + a_1(t)u + a_0(t)\right\}\frac{\partial P}{\partial r}(r,u,T) + \frac{\partial P}{\partial u}(r,u,T) - rP(r,u,T) = 0\,, \qquad (8)$$

with the boundary condition $P(r,T,T) = 1$.

Solving equation (8) on time interval $[t,T]$, we have (see Appendix for detail)

$$P(r,\tau) = A(\tau)\exp(-B(\tau)r)\,, \qquad (9)$$

where

$$\begin{aligned}
B(\tau) &= \frac{2(e^{\gamma\tau}-1)}{g(\tau)}\,,\\
A(\tau) &= \left[\frac{2\gamma e^{k_1\tau/2}}{g(\tau)}\right]^{\frac{2}{b_2^2}(b_2\hat{\mu}(r,t)-a_2\sigma^2(r,t)+b_1)} \times (2\gamma)^{2k_3\tau/b_2^2}\\
&\quad \times \exp\left\{\frac{1}{b_2}(\sigma^2(r,t)-b_2 r)(\tau - B(\tau)) + \frac{\tau^2}{2b_2^2}(b_1b_2+k_1k_3)\right\}\\
&\quad \times \exp\left\{-\frac{2k_3}{b_2^2}\int_t^T \log g(T-u)du\right\}\,,\\
\tau &= T-t\,,\\
\gamma &= (2b_2+a_2^2)^{0.5}\,,\\
g(\tau) &= k_1e^{\gamma\tau}+k_2\,,\\
k_1 &= \gamma - a_2\,,\\
k_2 &= \gamma + a_2\,,\\
k_3 &= b_2a_1 - b_1a_2\,.
\end{aligned}$$

Yield to maturity of a discount bond is given by $Y(r,\tau) = \frac{-1}{\tau}\log P(r,\tau)$. Therefore, by equation (9),

$$Y(r,\tau) = -\frac{1}{\tau}\left(\ln A(\tau) - B(\tau)r\right) . \tag{10}$$

Note that $A(\tau)$ and $B(\tau)$ are nonlinear functions in r at time t. Hence yield to maturity in our model is highly nonlinear in r. Later in the empirical analysis presented in section 4, we refer to our model as a nonlinear model for simplicity.

3 Selecting an Appropriate Short Rate Model

In this section, we determine appropriate specification for dynamics of the short rate. Many short-rate models are proposed in the existing literature, and among them we choose the following models for comparing the empirical performance.

$$\text{(CIR)} \qquad dr = (\alpha_1 r + \alpha_0)dt + \sigma\sqrt{r}\,dW_t \tag{11}$$

$$\text{(CKLS)} \qquad dr = (\alpha_1 r + \alpha_0)dt + \sigma r^{\beta}\,dW_t \tag{12}$$

$$\text{(Aït-Sahalia)} \qquad dr = (\alpha_2 r^2 + \alpha_1 r + \alpha_0 + \alpha_{-1} r^{-1})dt + \sigma r^{\beta}\,dW_t \tag{13}$$

Note that the CKLS model alone nests most of the well-known models. For example, Merton (1973), Vasicek (1977), Dothan (1978), Brenner and Schwartz (1979), as well as CIR; see CKLS (1992), p.1211. Aït-Sahalia (1996) considers more general specification for the diffusion term than that given by (13) to match the marginal density implied by the parametric model to that estimated by nonparametric procedures. But in this study, we restrict the short-rate model to a class of constant elasticity of the local variance since parameters in the more complex model do not successfully converge by conventional estimation methods. The same model as equation (13) is also examined by Conley et al. (1997).

We use the U.S. Treasury bill three-month rate as a proxy for the short-rate because of idiosyncratic behavior of short-term yields less than two months to maturity documented by Duffee (1996). [2] Our data is monthly form CRSP 12-month file and the sample period is form January 1965 to December 1989, which is almost the same period as that studied in CKLS (1992). Figure 1 displays plots of the historical data on the three-month rate.

For estimating parameters, we approximate these continuous-time models by the Euler scheme as,

$$r_{t+\Delta t} - r_t = f(r_t)\Delta t + \epsilon_{t+\Delta t}\,, \qquad \epsilon_{t+\Delta t} \sim N(0,\, \sigma^2 r_t^{2\beta})\,,$$

where Δt is a time interval between successive observations, and where $f(r) = (\alpha_1 r + \alpha_0)\Delta t$ and $\beta = 0.5$ for the CIR model, $f(r) = (\alpha_1 r + \alpha_0)\Delta t$ for the CKLS model, and $f(r) = (\alpha_2 r^2 + \alpha_1 r + \alpha_0 + \alpha_{-1} r^{-1})\Delta t$ for the Aït-Sahalia's model. By this discretization scheme, the random component is approximated to be a variable generated from a conditionally normal distribution. Thus we estimate parameters of approximated models by the maximum likelihood method. Since the Aït-Sahalia's model nests the other two models, we compare the short-rate models based on likelihood ratio tests (LRT).

Table 1 shows the results of parameter estimation. All the estimates are consistent with behavior of the short-rate, that is, the short-rate does not explode or reach zero in finite time. For example, the estimates in the CIR model satisfy inequalities as $2\alpha_0 - \sigma^2 \geq 0$ and $\alpha_1 < 0$. Similarly, the estimates in the CKLS model satisfy $\beta > 1$ and $\alpha_0 > 0$, and in the Aït-Sahalia's model $0 < \beta < 1.5$, $\alpha_{-1} > 0$ and $\alpha_2 < 0$. All these inequalities assure appropriate behavior of the short-rate at both boundaries, *i.e.*, zero and infinity; see Karlin and Taylor (1981), chapter 15, section 6 for behavior at boundaries and related restrictions on parameters.

The drift indicates mean reversion of the short-rate except in the CKLS model, in which

the estimate of α_1 is not significant. In both the CKLS and Aït-Sahalia's model, estimates of β are around 1.35, which are significantly different from 0.5 in the CIR model and less than the estimate in CKLS (1992). The last column denoted by LRT displays statistics of likelihood ratio tests with associated p-values in square brackets. The figure in the first (second) row is the statistic of the test in which the null and the alternative hypothesis are the CIR (CKLS) and the CKLS (Aït-Sahalia's) model, respectively. According to the test statistics, the CIR model is strongly rejected against the CKLS model while the CKLS model cannot be rejected against the Aït-Sahalia's model at conventional significance levels. Therefore we choose the CKLS model as an appropriate process of the short-rate.

4 Comparison of Term Structure Models

4.1 Estimating the market risk parameter

In the previous section, we specify functional forms of the drift and diffusion term and estimate their parameters. The remaining information that determines price of a discount bond is the risk premium. In the CIR framework, the risk premium is derived as covariance between changes in a state X and rates of return on production technologies. Remember that the short-rate is derived as $r = cX$ by assuming that the means and variances of rates of return on production technologies are proportional to X. Therefore, the risk premium is proportional to covariance between changes in the short-rate and rates of return on production technologies. Since variance of the former is proportional to $r^{2\beta}$ and variance of the latter is proportional to r, we have the risk premium function $\lambda(r)$ as

$$\lambda(r) = \lambda r^{\beta+0.5} . \tag{14}$$

Note that the same approach for deriving the risk premium function is taken in Longstaff (1989) and Beaglehole and Tenney (1992).

For estimating the parameter of the risk premium, we have the following regression equation by using equation (10).

$$Y(r,\tau) = -\frac{1}{\tau}(\ln A(\tau;\lambda) - B(\tau;\lambda)r) + \varepsilon_\tau \,, \tag{15}$$

where we assume that the mean of ε_τ is zero and that ε_τ is independent from r_t.

We employ the GMM by Hansen (1982), which is convenient for incorporating heteroskedastic structure of the covariance matrix of error terms. The orthogonality conditions we put are $E[\varepsilon_\tau] = 0$ and $E[\varepsilon_\tau r_t] = 0$. In other words, we use a constant and r as instrumental variables and the number of conditions are equal to two times the number of yields in our data set.

Similarly, we estimate λ in the CIR model. CIR assume the short-rate model given by (11) and derive the risk premium function as $\lambda(r) = \lambda r$. Therefore, their model for a discount bond is

$$P(r,\tau) = A(\tau)\exp(-B(\tau)r)\,, \tag{16}$$

where

$$\begin{aligned}
A(\tau) &= \left[\frac{2\gamma e^{(\lambda-\alpha_1+\gamma)\tau/2}}{(\lambda-\alpha_1+\gamma)(e^{\gamma\tau}-1)+2\gamma}\right]^{2\alpha_0/\sigma^2}, \\
B(t,T) &= \frac{2(e^{\gamma\tau}-1)}{(\lambda-\alpha_1+\gamma)(e^{\gamma\tau}-1)+2\gamma}, \\
\gamma &= ((\lambda-\alpha_1)^2+2\sigma^2)^{1/2}\,, \\
\tau &= T-t\,.
\end{aligned}$$

The data is monthly form CRSP 12-month file. The sample period is from January 1965 to December 1989, the same period in estimating parameters of the short-rate models. The

three-month rate is a proxy for the short-rate, as noted before, and yields of four through eleven months to maturity are explained by models. We exclude data on the twelve-month yield since almost a quarter of the data are missing in our sample period. Therefore, we use sixteen orthogonality conditions for estimating λ by the GMM.

Table 2 shows estimates of the market risk parameter. The estimates in both models are significant and negative, which implies positive term premium. The estimate in our nonlinear term structure model is much greater in absolute value than that in the CIR model. This is simply because the risk premium function is proportional to $r^{\beta+0.5}$ in our model and the estimate of β is greater than 1.

4.2 Comparing empirical performance of term structure models

We compare empirical performance between the CIR model and ours by using out-of-sample data. The dataset is the same from CRSP 12-month file and the out-of-sample data cover the period from January 1990 to March 2000 (123 observations). Table 3 provides results of comparison based on root mean squared errors (RMSE), which is decomposed into bias and standard deviation (Std Dev). We find that our nonlinear term structure model outperforms the CIR linear model for all yields. The differences of RMSE between the two models are raging from 4.6 basis point (bp) (four-month yield) to 7.7 bp (eleven-month yield). The components of RMSE indicate that better performance of our model comes from substantial reduction of the bias of pricing errors. For example, for eleven-month yield, the bias of our model is nearly half as large as that of the linear model while the variances are almost the same magnitude. The last column in Table 3 shows the proportion that the nonlinear model outperforms the linear model in prediction. For example, for the four-month yield, absolute pricing errors of the nonlinear model are

smaller than those of the linear model at 119 observations out of total 123 observations. As a whole, our model describes actual yields better at more than 80% of total observations in the out-of-sample period.

Figure 2 displays the time-series property of relative performance between the two models. Plots are absolute pricing errors of the linear model minus those of the nonlinear model, hence the nonlinear model shows better (worse) performance in a positive (negative) region than the linear model. Clearly, our nonlinear model has better descriptive power for most of the sample period.

5 Concluding Remarks

We derive a term structure model with a generally specified short-rate model. In existing models, behavior of the short-rate is restricted. However, our model shows that we do not have to restrict the drift and diffusion term of a short-rate process for the sake of the closed-form solution for the term structure. In our model, the functional forms of the drift and diffusion term and their parameters are simply determined by actual data. With such general conditions, we still derive an explicit model by applying the local linearization method (LLM) to the diffusion process of the short-rate. And the better descriptive power of our model is confirmed by the empirical analysis, showing that our model outperforms the affine model which has sever restriction on behavior of the short-rate. Since our model is a closed-form, computational time is substantially reduced comparing to numerical methods.

In this paper we derived our term structure model in the CIR general equilibrium framework. By following this framework, the market risk function is automatically determined and we can focus only on appropriate specification for the drift and diffusion term

of a short-rate process. But it is important to note that our model can be constructed by employing arbitrage-free arguments. In this case, the market risk function should be provided exogenously. By finding an appropriate functional form of the market risk which does not violate arbitrage-free conditions and by incorporating this result into term structure models, the descriptive power may be improved. This is one of the future courses of this study.

Another future study is related to models in the multi-factor framework. Throughout this paper we stressed the importance of the term structure model which is consistent with observed behavior of the short-rate. And more recent works reveal that two or three factors are needed to capture the dynamics of the short-rate more adequately; see, for example, Andersen and Lund (1997), Balduzzi et al. (1998), Bali (2000), Brenner et al. (1996), Gallant and Tauchen (1998). As for term structure models in the multi-factor framework, Longstaff and Schwartz (1992) assume another factor which derives the volatility of the short-rate. Balduzzi et al. (1996) and Chen (1996) introduce three factors which are projected onto the short-rate, the volatility of short-rate changes, and the long-term mean of the short-rate; see also Dai and Singleton (2000) for more general three-factor models. However, their term structure models are based on affine specification of factor processes, which may be inadequate for actual behavior. Therefore, it is important to find appropriate factors which affect bond prices and to examine the whole system of these stochastic factors by data. And it is our future task that we derive a term structure model with such more general conditions.

Appendix

We are going to solve equation (8)

$$\tfrac{1}{2}\left\{b_2(t)r_u + b_1(t)u + b_0(t)\right\}\frac{\partial^2 P}{\partial r^2}(r,u,T)$$
$$+\left\{a_2(t)r_u + a_1(t)u + a_0(t)\right\}\frac{\partial P}{\partial r}(r,u,T) + \frac{\partial P}{\partial u}(r,u,T) - rP(r,u,T) = 0\,,$$

with the boundary condition $P(r,T,T) = 1$ as follows.

Since each coefficient of partial derivatives of P becomes linear in r, though dependent explicitly on time u, we assume a form of the solution to be $P(r,u,T) = A(u,T)\exp(-B(u,T)r)$. By differentiating this function properly and substituting partial derivatives into equation (8), we have ordinary differential equations (ODE) for A and B,

$$\frac{dB}{du} = \tfrac{1}{2}b_2B^2(u,T) - a_2B(u,T) - 1\,, \qquad (A1)$$

and

$$\frac{dA/du}{A} = -\tfrac{1}{2}(b_0 + b_1u)B^2(u,T) + (a_0 + a_1u)B(u,T)\,, \qquad (A2)$$

with boundary conditions $B(T,T) = 0$ and $A(T,T) = 1$. The ODE for B is almost the same as that in the original CIR model, thus integrating both sides of equation (A1) on $[t,T]$ provides

$$B(t,T) = \frac{2(e^{\gamma(T-t)} - 1)}{(\gamma - a_2)e^{\gamma(T-t)} + (\gamma + a_2)}\,, \qquad (A3)$$

where

$$\gamma = (2b_2 + a_2^2)^{0.5}\,.$$

We rewrite equation (A3) just for notational convenience as

$$B(\tau) = \frac{2(e^{\gamma\tau} - 1)}{g(\tau)} \qquad (A3')$$

where

$$\tau = T - t\,, \quad g(\tau) = k_1e^{\gamma\tau} + k_2\,,$$

$$k_1 = \gamma - a_2 \,, \quad k_2 = \gamma + a_2 \,.$$

Next, solving the ODE for A is not straightforward because it depends on the current time u explicitly. Before integrating both sides of equation (A2), we decompose B and B^2 as follows.

$$B(u,T) \equiv L_1 + L_2 \frac{g'(T-u)}{g(T-u)} \,, \tag{A4}$$

and

$$B^2(u,T) \equiv M_1 + M_2 \frac{g'(T-u)}{g(T-u)} + M_3 \frac{g'(T-u)}{g(T-u)^2} \,, \tag{A5}$$

where $g' = dg/du$. Simple algebra gives

$$L_1 = \frac{-k_1}{b_2} \,, \qquad L_2 = \frac{-2}{b_2} \,, \tag{A6}$$

and

$$M_1 = \frac{k_1^2}{b_2^2} \,, \qquad M_2 = \frac{-4a_2}{b_2^2} \,, \qquad M_3 = \frac{4\gamma k_2}{b_2^2}. \tag{A7}$$

Substituting equation (A4) and (A5) into (A2) gives

$$\begin{aligned} \frac{A'(u,T)}{A(u,T)} &= -\tfrac{1}{2}(b_0 + b_1 u)\left\{M_1 + M_2 \frac{g'(T-u)}{g(T-u)} + M_3 \frac{g'(T-u)}{g(T-u)^2}\right\} \\ &\quad +(a_0 + a_1 u)\left\{L_1 + L_2 \frac{g'(T-u)}{g(T-u)}\right\} \,, \end{aligned}$$

where $A' = dA/du$. Integrating both sides on $[t,T]$ and calculating further with the boundary condition $A(T,T) = 1$ gives,

$$\begin{aligned} A(t,T) &= \left[\frac{(2\gamma)^{\frac{1}{2}M_2(b_1T+b_0) - L_2(a_1T+a_0) + \frac{1}{2}M_3(b_1/\gamma k_2)}}{g(\tau)^{\frac{1}{2}M_2(b_1t+b_0) - L_2(a_1t+a_0) + \frac{1}{2}M_3(b_1/\gamma k_2)}}\right] \\ &\quad \times \exp\left[\tau\left\{\tfrac{1}{2}M_1(b_0 + \tfrac{T+t}{2}b_1) - L_1(a_0 + \tfrac{T+t}{2}a_1)\right\}\right] \end{aligned}$$

$$\times \exp\left[\tfrac{1}{2}M_3\left\{\frac{b_1k_1e^{\gamma\tau}\tau}{k_2g(\tau)} - \frac{k_1}{4\gamma}(b_1T+b_0)B(t,T)\right\}\right]$$

$$\times \exp\left[-(\tfrac{1}{2}b_1M_2 - a_1L_2)\int_t^T \log g(T-u)du\right]. \tag{A8}$$

At first glance, $A(t,T)$ depends directly on time t and T, which is in contrast with the original CIR model where $A(t,T)$ depends only on the difference between two points in time, *i.e.*, $\tau = T - t$. However, remember that coefficients a_i, b_i $(i = 0,1,2)$ correspond to suitable derivatives of the drift and diffusion function evaluated at time t. Hence we have the following equations, for example, by equation (5)

$$\begin{aligned} a_0 + a_1(T+t)/2 &= (\hat{\mu}(r,t) - a_2r - a_1t) + a_1(T+t)/2 \\ &= \hat{\mu}(r,t) - a_2r + a_1\tau/2\,. \end{aligned}$$

Similarly,

$$\begin{aligned} a_0 + a_1T &= \hat{\mu}(r,t) - a_2r + a_1\tau\,, \\ a_0 + a_1t &= \hat{\mu}(r,t) - a_2r\,, \end{aligned}$$

and by equation (6)

$$\begin{aligned} b_0 + b_1(T+t)/2 &= \sigma^2(r,t) - b_2r + b_1\tau/2\,. \\ b_0 + b_1T &= \sigma^2(r,t) - b_2r + b_1\tau\,, \\ b_0 + b_1t &= \sigma^2(r,t) - b_2r\,. \end{aligned}$$

Thus bond price depends on t and T only through $\tau = T - t$. Putting these equations into (A8) provides

$$A(t,T) = \left[\frac{(2\gamma)^{\frac{1}{2}M_2(\sigma^2(r,t)-b_2r+b_1\tau)-L_2(\hat{\mu}(r,t)-a_2r+a_1\tau)+\frac{1}{2}M_3(b_1/\gamma k_2)}}{g(\tau)^{\frac{1}{2}M_2(\sigma^2(r,t)-b_2r)-L_2(\hat{\mu}(r,t)-a_2r)+\frac{1}{2}M_3(b_1/\gamma k_2)}}\right]$$

$$\times \exp\left[\tau\left\{\tfrac{1}{2}M_1(\sigma^2(r,t) - b_2 r + b_1\tau/2) - L_1(\hat{\mu}(r,t) - a_2 r + a_1\tau/2)\right\}\right]$$

$$\times \exp\left[\tfrac{1}{2}M_3\left\{\frac{b_1 k_1 e^{\gamma\tau}\tau}{k_2 g(\tau)} - \frac{k_1}{4\gamma}(\sigma^2(r,t) - b_2 r + b_1\tau)B(t,T)\right\}\right]$$

$$\times \exp\left[-(\tfrac{1}{2}b_1 M_2 - a_1 L_2) - \int_t^T \log g(T-u)du\right]$$

Substituting equation (A6) and (A7) and calculating further provides the final result for $A(t,T)$.

Endnote

1 See also Pritsker (1998), and Chapman and Pearson (2000), pointing out estimation problems associated with the nonparametric techniques employed by Aït-Sahalia (1996) and Stanton (1997).

2 Stanton (1996) and Andersen and Lund (1997) also use the three-month yield as a proxy for the short-rate. However, Chapman et al. (1999) discuss the problem of using yields of finite maturity as a proxy for the short-rate. We examined this problem by following their method. Our result indicates that there is no serious problem associated with our model which incorporates behavior of the short-rate by the CKLS model.

Reference

Ahn, D., and B. Gao, 1999, "A parametric nonlinear model of term structure dynamics", *Review of Financial Studies*, 12, 721-762.

Aït-Sahalia, Y., 1996, "Testing continuous-time models of the spot interest rate", *Review of Financial Studies*, 9, 385-426.

Andersen, T. G. and J. Lund, 1997, "Estimating continuous-time stochastic volatility models of the short-term interest rate", *Journal of Econometrics*, 77, 343-377.

Balduzzi, P., S. R. Das, and S. Foresi, 1998, "The central tendency: A second factor in bond yields", *Review of Economics and Statistics*, 80, 62-72.

Balduzzi, P., S. R. Das, S. Foresi, and R. Sundaram, 1996, "A simple approach to three-factor affine term structure models", *Journal of Fixed Income*, 6, 43-53

Bali, T. G., 2000, "Testing the empirical performance of stochastic volatility models of the short-term interest rate", *Journal of Financial and Quantitative Analysis*, 35, 191-215.

Beaglehole, D., and M. Tenney, 1992, "Corrections and additions to 'A nonlinear equilibrium model of the term structure of interest rates'," *Journal of Financial Economics*, 32, 345-353.

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.

Brenner, R., R. Harjes and K. Kroner, 1996, "Another look at models of the short-term interest rate", *Journal of Financial and Quantitative Analysis*, 31, 85-107.

Chan, K. C., G. A. Karolyi, F. A. Longstaff, and A. S. Sanders, 1992, "An empirical comparison of alternative models of the short-term interest rate," *Journal of Finance*, 47, 1209-1227.

Chapman, D. A., J. B. Long, Jr., and N. Pearson, 1999, "Using proxies for the short rate: When are three months like an instant?", *Review of Financial Studies*, 12, 763-806.

Chapman, D. A., and N. D. Pearson, 2000, "Is the short rate drift actually nonlinear?", *Journal of Finance*, 55, 355-388.

Chen, Lin, 1996, "A bond pricing formula under a non-trivial, three factor model of interest rates", *Economic Letters*, 51, 95-99.

Conley, T. G., L. P. Hansen, E. G. J. Luttmer, and J. A. Scheinkman, 1997, "Short-term interest rate as subordinated diffusions," *Review of Financial Studies*, 10, 525-577.

Cox, J. C., J. E. Ingersoll, Jr., and S. A. Ross, 1985, "A theory of the term structure of interest rates," *Econometrica*, 53, 385-407.

Dahlquist, M., 1996, "On alternative interest rate processes," *Journal of Banking and Finance*, 20, 1093-1119.

Dai, Q., and K. J. Singleton, 2000, "Specification analysis of affine term structure models," *Journal of Finance*, 55, 1943 - 1978.

Dothan, U. L., 1978, "On the term structure of interest rate", *Journal of Financial Economics*, 6, 59-69.

Duffee, G. R., 1996, "Idiosyncratic variation of Treasury bill yield", *Journal of Finance*, 51, 527-551.

Duffie, D., and R. Kan, 1996, "A yield-factor model of interest rates," *Mathematical Finance*, 6, 379-406.

Karlin, S., and H. M. Taylor, 1981, *A second course in stochastic processes*, Academic Press.

Gallant, A. R. and G. Tauchen, 1998, "Reprojecting partially observed systems with application to interest rate diffusions", *Journal of the American Statistical Association*, 93, 10-24.

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

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," *Journal of Finance*, 47, 1259-1282.

Merton, R. C., 1973, "Theory of rational option pricing", *Bell Journal of Economics and Management Science*, 4, 141-183.

Nowman, K. B., 1997, "Gaussian estimation of single-factor continuous time models of the term structure of interest rates", *Journal of Finance*, 52, 1695-1706.

Pritsker, M., 1998, "Nonparametric density estimation and test of continuous time interest rate models", *Review of Financial Studies*, 11, 449-487.

Shoji, I., 1998, "Approximation of continuous time stochastic process by a local linearization method," *Mathematics of Computation*, 67, 287-298.

Shoji, I., and T. Ozaki, 1996, "A statistical comparison of the short-term interest rate models for Japan, U.S., and Germany", *Financial Engineering and the Japanese Markets*, 3, 263-275.

Shoji, I., and T. Ozaki, 1997, "Comparative study of estimation methods for continuous time stochastic processes," *Journal of Time Series Analysis*, 18, 485-506.

Stanton, R., 1997, "A nonparametric model of term structure dynamics and the market price of interest rate risk," *Journal of Finance*, 52, 1973-2002.

Vasicek, O., 1977, "An equilibrium characterization of the term structure," *Journal of Financial Economics*, 5, 177-188.

	α_2	α_1	α_0	α_{-1}	σ	β	LogL	LR Test
CIR	-	-0.3153	0.0241	-	0.0776	-	1115.5	74.83
		(-1.851)	(2.104)		(24.45)			[0.000]
CKLS	-	-0.2598	0.0206	-	0.6867	1.3532	1152.9	4.49
		(-1.485)	(2.109)		(3.676)	(13.60)		[0.106]
Aït-Sahalia	-25.779	5.5175	-0.3680	0.0080	0.6930	1.3593	1155.2	
	(-1.992)	(1.987)	(-2.009)	(2.123)	(3.610)	(13.41)		

t-statistics are in parenthesis. Figures in square brackets at the last column are p-values

Table 1. Estimates of Parameters in Short-Rate Models

The short-rate models in continuous-time are discretized by the Euler scheme as,

$$r_{t+\Delta t} - r_t = f(r_t)\Delta t + \epsilon_{t+\Delta t}\,, \qquad \epsilon_{t+\Delta t} \sim N(0,\, \sigma^2 r_t^{2\beta})\,,$$

where Δt is a time interval between successive observations, and where $f(r) = (\alpha_1 r + \alpha_0)\Delta t$ and $\beta = 0.5$ for the CIR model, $f(r) = (\alpha_1 r + \alpha_0)\Delta t$ for the CKLS model, and $f(r) = (\alpha_2 r^2 + \alpha_1 r + \alpha_0 + \alpha_{-1} r^{-1})\Delta t$ for the Aït-Sahalia's model.

Parameters in these approximated models are estimated by the maximum likelihood method. LogL denotes values at maximum log-likelihood. The last column LRT displays statistics of likelihood ratio tests. The figure in the first (second) row is the statistic of the test in which the null and the alternative hypothesis are the CIR (CKLS) and the CKLS (Aït-Sahalia's) model, respectively.

	λ
CIR (Linear) Model	-0.0972
	(-18.04)
Nonliear Model	-0.6046
	(-15.84)

t-statistics are in parenthesis.

Table 2. Estimates of the Market Risk Parameter

We use the following regression equation for estimating the parameter of the risk premium function.

$$Y(r,\tau) = -\frac{1}{\tau}(\ln A(\tau;\lambda) - B(\tau;\lambda)r) + \varepsilon_\tau \ ,$$

where $A(\tau;\lambda)$ and $B(\tau;\lambda)$ for our model are given by (9) and for the CIR model by (16), and where the mean of ε_τ is zero and ε_τ is independent from r_t. We employ the GMM by Hansen (1982). The orthogonality conditions we put are $E[\varepsilon_\tau] = 0$ and $E[\varepsilon_\tau r] = 0$.

The data is monthly form January 1965 to December 1989. The three-month rate is a proxy for the short-rate and yields of four through eleven months to maturity are explained by the models. (sixteen orthogonality conditions)

	CIR (Linear) Model			Nonlinear Model			Difference	Proportion
	RMSE(A)			RMSE(B)			(A)-(B)	(%)
		Bias	Std Dev		Bias	Std Dev		
4-month	18.23	16.64	7.44	13.65	11.66	7.11	4.58	96.7
5-month	20.08	16.45	11.52	15.25	10.28	11.26	4.83	91.9
6-month	22.46	17.21	14.43	17.29	9.90	14.18	5.16	87.8
7-month	26.48	20.00	17.35	20.69	11.58	17.14	5.79	87.0
8-month	29.11	21.14	20.02	22.97	11.62	19.81	6.15	87.0
9-month	31.77	22.31	22.61	25.23	11.73	22.34	6.54	82.9
10-month	34.53	23.69	25.12	27.61	12.07	24.84	6.92	81.3
11-month	36.84	25.72	26.38	29.19	13.09	26.10	7.65	82.9

in basis point except for the last column

Table 3. Comparison of Term Structure Models

Term structure models are compared by using out-of-sample data form January 1990 to March 2000. RMSE denote root mean squared error which is decomposed into bias and standard deviation (Std Dev). Prediction errors are calculated by taking difference between theoretical (model-implying) and actual yields. The last column shows the proportion that our nonlinear model outperforms the linear model in the total out-of-sample period.

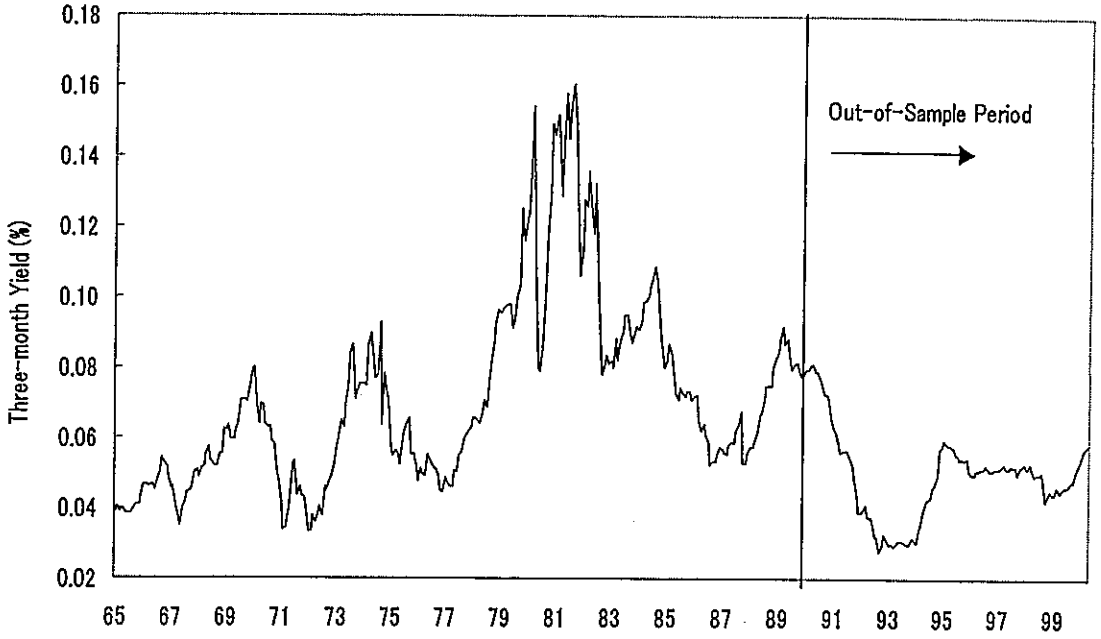

Figure 1: Historical Data on the Treasury Bill Three-month Rate

The data is monthly from January 1965 to March 2000. The sample period is divided into two: The first period is for selecting an appropriate short-rate model and for estimating parameters, and the second period is for comparing performance of our nonlinear term structure model with that of the CIR linear model.

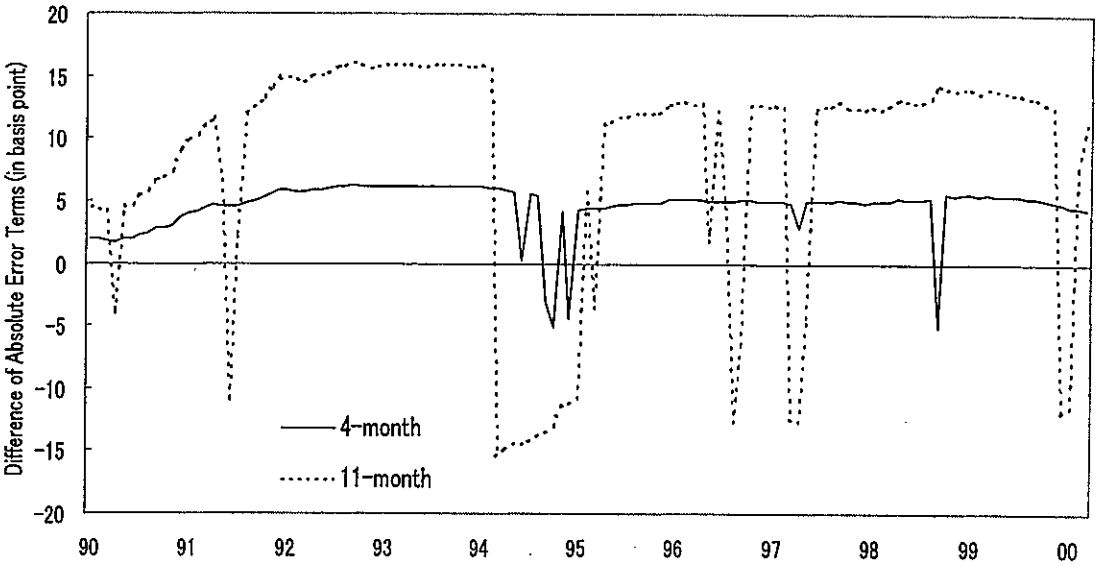

Figure 2: Performance Comparison of Term Structure Models

Plots are absolute pricing errors of the CIR linear term structure model minus those of our nonlinear model, hence the positive (negative) region indicate better performance of the nonlinear (linear) model. The out-of-sample period is form January 1990 to March 2000 (123 observations).