

Modeling the Term Structure of Interest Rates with General Diffusion Processes: A Moment Approximation Approach

Hideyuki Takamizawa *
Graduate School of Humanities
and
Social Sciences,
University of Tsukuba

Isao Shoji
Graduate School of Systems
and
Information Engineering,
University of Tsukuba

April 18, 2008

Abstract

We propose an analytical approximation of the term structure of interest rates under general diffusion processes of the short-rate and state variables. A method of approximating conditional moments as the solution to a system of ordinary differential equations is applied to the pricing of bonds. Numerical experiments based on two illustrative models show that the second-order approximation is accurate for maturities of up to five years and the third-order approximation is effective for longer maturities. We also show the possibility of improving the second-order approximation without much increasing the computational burden.

JEL classification: C63, G12

Keywords: Short-rate, Term structure, Approximation, Conditional moment

*the corresponding author: Graduate School of Humanities and Social Sciences, University of Tsukuba, Tsukuba Ibaraki 305-8571, JAPAN. Phone & Fax: +81-29-853-4071. E-mail: takamiza@social.tsukuba.ac.jp

1 Introduction

The dynamics of the instantaneous risk-free rate, the short-rate, have been intensively studied. Yet, their empirical findings are not sufficiently incorporated into the modeling of the term structure of interest rates despite the fact that the time-series behavior of the short-rate substantially affects the cross-section of bond prices. In fact, since the volatility of changes in the short-rate is invariant to changes in probability measures, statistical properties of the volatility should in principle be inherited from the actual to some pricing measures. Nevertheless, volatility specifications preferred in time series analysis have often been abandoned in term structure modeling.

This is simply because analytical models of the term structure cannot generally be obtained for such specifications, which is inconvenient for practical purposes of pricing and estimation. Examples of such specifications include constant elasticity of volatility (CEV) (e.g., Chan et al., 1992, and Nowman, 1997),¹ and CEV combined with some persistent factors, which are modeled by GARCH processes (e.g., Bali, 2000, and Brenner et al., 1996) or by additional stochastic processes (e.g., Andersen and Lund, 1997, Ball and Torous, 1999, Durham, 2003, and Gallant and Tauchen, 1998). A common feature of the GARCH and stochastic volatility models is that the volatility is specified as the product of the short-rate and persistent factors. These multiplicative volatility models, however, do not admit closed-form expressions of the term structure.

In this paper, we propose an analytical approximation of the term structure based on general models of the short-rate and state variables including those preferred in time series analysis. We adopt a method originally proposed by Shoji (2002), which approximates a vector containing conditional moments of a multidimensional diffusion process as the

¹A notable exception is Ahn and Gao (1999), who derived a closed-form term structure model under the CEV process with the elasticity parameter of 1.5.

solution to a system of ordinary differential equations. Since the price of a discount bond is basically given by the conditional expectation of the stochastic discount factor, this method is readily applicable.

Using the Monte Carlo (MC) method as a benchmark pricing tool, the accuracy of the approximation is examined for two illustrative models. One is a stochastic volatility (SV) model in which the short-rate volatility has a multiplicative form, and the other is a stochastic central tendency (SCT) model, as proposed by Balduzzi et al. (1996, 1998), in which all state variables are instantaneously correlated and have level-dependent volatilities. Numerical results based on the two models are that the second-order approximation is accurate for maturities of up to five years and that the third-order approximation is effective for maturities longer than this. Also, we show the possibility of improving the second-order approximation without much increasing the computational burden.

As for related previous studies, Takamizawa and Shoji (2003, 2004) also proposed a method of approximating the term structure. These studies rely on a local *linear* approximation, which is applied to the drift and diffusion terms of the short-rate and state variables to utilize a framework of affine term structure models. Accordingly, in a multi-factor setting, models to which the previous method is applicable are limited, such that they originally have a similar structure as affine models.² On the other hand, the current method can achieve *higher-order* approximations, which may be more beneficial when the drift and diffusion terms exhibit a high degree of nonlinearity. Besides, no particular restriction on models (other than some technical conditions listed in Section 2) is necessary for applying the current method since the price of a discount bond is directly computed as the conditional expectation. In terms of computational efficiency, the previous method

²See Takamizawa and Shoji (2004, p.155). Hence, the previous method cannot be applied to the models considered in Section 3 of this paper for deriving the term structure.

has an advantage: with a single calculation, the previous method obtains a vector of model-implied yields, whereas the current method obtains a single yield.

The rest of the paper is organized as follows. Section 2 explains and implements the proposed method for bond pricing. Section 3 performs numerical experiments on the accuracy of the approximation and presents their results. Section 4 provides concluding remarks.

2 Approximation Formula of Conditional Moments and its Application to Bond Pricing

We practically explain the approximation method originally proposed by Shoji (2002). The computation of up to the conditional second moment is illustrated first for a one-dimensional process, which is then extended to multi-dimensional processes. The computation of higher-order moments is also possible and is a straightforward extension.

2.1 A one-dimensional process

Let X_t be a stochastic process, which evolves according to the following stochastic differential equation (SDE):

$$dX_t = f(X_t)dt + \sigma(X_t)dW_t, \quad (1)$$

where W_t is standard Brownian motion, and where the drift and diffusion functions, f and σ , satisfy certain technical conditions for the solution to (1) to exist for an arbitrary X_0 . We also assume that f and $g = \sigma^2$ are appropriately smooth.³

³More specifically, when computing the first through third conditional moments by the proposed method, we have only to assume that f and g belong to C^4 -class. This assumption is needed for the convergence of truncated Taylor series to the true ones as $t \downarrow s$ and at the same time for the convergence of the approximate conditional moments. We give a simple sketch of the proof about convergence of Tay-

Let $\Psi_s(t)$ be a vector containing the first and second moments of an increment of X_t conditioned on time $s < t$. That is,

$$\Psi_s(t)' = E_s \left(X_t - X_s \quad (X_t - X_s)^2 \right) . \quad (2)$$

We then show that $\Psi_s(t)$ is approximated as the solution to a system of ordinary differential equations.

By integrating (1) and taking the conditional expectation, we have

$$E_s[X_t - X_s] = E_s \left[\int_s^t f(X_u) du \right] . \quad (3)$$

By applying the Taylor expansion to $f(X_u)$ around X_s up to the third order and substituting this into (3), we have

$$\begin{aligned} E_s[X_t - X_s] &= f(X_s)(t - s) \\ &+ E_s \left[\int_s^t \left\{ f^{(1)}(X_s)(X_u - X_s) + \frac{1}{2} f^{(2)}(X_s)(X_u - X_s)^2 \right\} du \right] \\ &+ E_s \left[\int_s^t \frac{1}{6} f^{(3)}(\xi)(X_u - X_s)^3 du \right] , \end{aligned} \quad (4)$$

where $f^{(k)} = \frac{\partial^k f}{\partial X^k}$, and $\xi = \alpha X_t + (1 - \alpha)X_s$ for some $\alpha \in [0, 1]$. (4) can be expressed in a vector form as

$$E_s[X_t - X_s] = f(X_s)(t - s) + \left(f^{(1)}(X_s) \quad \frac{1}{2} f^{(2)}(X_s) \right) \int_s^t \Psi_s(u) du + R_1 , \quad (5)$$

where R_1 is a residual term corresponding to the last term in (4).

Next, by applying the Ito formula to $(X_t - X_s)^2$ and taking the conditional expectation, we have

$$E_s[(X_t - X_s)^2] = E_s \left[\int_s^t \{ 2f(X_u)(X_u - X_s) + g(X_u) \} du \right] , \quad (6)$$

lor series in Appendix A. The assumption does not seem too restrictive since most interest-rate models actually used lie in such a family.

where $g = \sigma^2$. By applying the Taylor expansion to $f(X_u)$ and $g(X_u)$ around X_s up to the second and third orders, respectively, and then substituting these into (6), we have

$$\begin{aligned} E_s[(X_t - X_s)^2] &= g(X_s)(t - s) \\ &+ E_s \left[\int_s^t \left\{ (2f(X_s) + g^{(1)}(X_s))(X_u - X_s) + (2f^{(1)}(X_s) + \frac{1}{2}g^{(2)}(X_s))(X_u - X_s)^2 \right\} du \right] \\ &+ E_s \left[\int_s^t \left\{ (f^{(2)}(\xi_1) + \frac{1}{6}g^{(3)}(\xi_2))(X_u - X_s)^3 \right\} du \right], \end{aligned} \quad (7)$$

where $g^{(k)}$ is defined analogously with $f^{(k)}$, and $\xi_i = \alpha_i X_t + (1 - \alpha_i)X_s$ for some $\alpha_i \in [0, 1]$ ($i = 1, 2$). (7) can also be expressed in a vector form as

$$\begin{aligned} E_s[(X_t - X_s)^2] &= g(X_s)(t - s) \\ &+ \begin{pmatrix} 2f(X_s) + g^{(1)}(X_s) & 2f^{(1)}(X_s) + \frac{1}{2}g^{(2)}(X_s) \end{pmatrix} \int_s^t \Psi_s(u) du + R_2, \end{aligned} \quad (8)$$

where R_2 is a residual term corresponding to the last term in (7). Expressing (5) and (8) together in a vector form leads to

$$\Psi_s(t) = A(X_s) \int_s^t \Psi_s(u) du + b(X_s)(t - s) + R, \quad (9)$$

where

$$A(X_s) = \begin{pmatrix} f^{(1)}(X_s) & \frac{1}{2}f^{(2)}(X_s) \\ 2f(X_s) + g^{(1)}(X_s) & 2f^{(1)}(X_s) + \frac{1}{2}g^{(2)}(X_s) \end{pmatrix} \quad \text{and} \quad b(X_s) = \begin{pmatrix} f(X_s) \\ g(X_s) \end{pmatrix}.$$

(9) can be developed to

$$\Psi_s(t) = \int_s^t e^{A(X_s)(t-u)} b(X_s) du + \hat{R}, \quad (10)$$

where \hat{R} is a residual vector. If, furthermore, A is invertible, we have

$$\Psi_s(t) = A^{-1}(X_s) \{ e^{A(X_s)(t-s)} - I \} b(X_s) + \hat{R}. \quad (11)$$

(10) or (11) without the residual vector is referred to as the approximation formula, which is used for computing the prices of discount bonds.

According to Shoji (2002), the rate of convergence of both residual vectors, R and \hat{R} , is $O((t-s)^{(n+3)/2})$ for up to n -th conditional moments. Therefore, n can also be considered as the order of approximation. Conversely, the formula is exact if f and g are at most linear and quadratic functions, respectively. In this case, $f^{(2)}$ and $g^{(3)}$ are zero, and so are the elements of the residual vectors.

2.2 A two-dimensional process

Let $X'_t = (x_{1,t} \ x_{2,t})$ be a two-dimensional stochastic process, which follows

$$dx_{i,t} = f_i(X_t)dt + \sigma_i(X_t)'dW_t \quad (i = 1, 2), \quad (12)$$

where W_t is two-dimensional standard Brownian motion. A moment vector can be specified here as

$$\Psi_s(t)' = E_s \left(x_{1,t} - x_{1,s} \quad x_{2,t} - x_{2,s} \quad (x_{1,t} - x_{1,s})^2 \quad (x_{2,t} - x_{2,s})^2 \quad (x_{1,t} - x_{1,s})(x_{2,t} - x_{2,s}) \right). \quad (13)$$

Similar to the one-dimensional case, the conditional first moments are approximated using the Taylor expansion of $f_i(X_t)$ around X_s up to the third order. Likewise, the conditional second moments are approximated using the Taylor expansion of $f_i(X_t)$ and $g_{ij}(X_t) = \sigma_i(X_t)'\sigma_j(X_t)$ around X_s up to the second and third orders, respectively. Then, expressing these moments in a vector form leads to (9), where (X_s is abbreviated for simplicity)

$$A = \begin{pmatrix} f_1^{(1,0)} & f_1^{(0,1)} & \frac{1}{2}f_1^{(2,0)} & \frac{1}{2}f_1^{(0,2)} & f_1^{(1,1)} \\ f_2^{(1,0)} & f_2^{(0,1)} & \frac{1}{2}f_2^{(2,0)} & \frac{1}{2}f_2^{(0,2)} & f_2^{(1,1)} \\ 2f_1 + g_{11}^{(1,0)} & g_{11}^{(0,1)} & 2f_1^{(1,0)} + \frac{1}{2}g_{11}^{(2,0)} & \frac{1}{2}g_{11}^{(0,2)} & 2f_1^{(0,1)} + g_{11}^{(1,1)} \\ g_{22}^{(1,0)} & 2f_2 + g_{22}^{(0,1)} & \frac{1}{2}g_{22}^{(2,0)} & 2f_2^{(0,1)} + \frac{1}{2}g_{22}^{(0,2)} & 2f_2^{(1,0)} + g_{22}^{(1,1)} \\ f_2 + g_{12}^{(1,0)} & f_1 + g_{12}^{(0,1)} & f_2^{(1,0)} + \frac{1}{2}g_{12}^{(2,0)} & f_1^{(0,1)} + \frac{1}{2}g_{12}^{(0,2)} & f_1^{(1,0)} + f_2^{(0,1)} + g_{12}^{(1,1)} \end{pmatrix},$$

and $b' = (f_1 \quad f_2 \quad g_{11} \quad g_{22} \quad g_{12})$.

2.3 Application to bond pricing

Let

$$z_{s,t} = \exp\left(-\int_s^t r(X_u)du\right), \quad (14)$$

and the price of a discount bond at time t maturing at time T is equal to the conditional first moment of $z_{t,T}$ under the risk-neutral measure. This (actually $E_t[z_{t,T} - z_{t,t}]$) is computed as one of the elements of the moment vector, $\Psi_t(T)$.

In applying the approximation formula (11), the SDE for $z_{s,t}$ is required:

$$dz_{s,t} = -r(X_t)z_{s,t}dt \quad \text{with} \quad z_{s,s} = 1. \quad (15)$$

Therefore, if X_t is a two-dimensional process, as is the case for standard SV models in which the short-rate and volatility are driven by two stochastic factors, a moment vector for a three-dimensional process actually needs to be computed. Table 1 provides the elements of $A(X_s)$ and $b(X_s)$ for a moment vector containing up to conditional second moments of a three-dimensional process. Likewise, if X_t is a three-dimensional process, as for the SCT model considered in the next section, a moment vector for a four-dimensional process is required. The derivation of $A(X_s)$ and $b(X_s)$ is a straightforward extension of the lower-dimensional cases. Also, the derivation of $A(X_s)$ and $b(X_s)$ for higher-order approximations is a straightforward extension of the lower-order cases. Therefore, tables listing the elements of them are omitted to save space. It is noted that filling in $A(X_s)$ is not as tedious as it seems. Since $z_{s,t}$ has no diffusion term, there are actually many zero elements in $A(X_s)$.

3 Numerical Experiments

To examine whether the approximation formula can in practice be utilized for pricing discount bonds, we perform numerical experiments on the accuracy using two illustrative models. We first check the accuracy of the second-order approximation. If it performs poorly, we explore possible causes of the inaccuracy, and then consider the third-order approximation. The identification of the causes also helps us improve the second-order approximation with much less computational burden than that of the third-order approximation.

3.1 Models

The first is an SV model, which has a multiplicative volatility specification, similar to the models in Andersen and Lund (1997), and Gallant and Tauchen (1998):

$$dr_t = (a_0 + a_1 r_t + a_2 v_t)dt + v_t r_t dW_{1,t}, \quad (16)$$

$$dv_t = (b_0 + b_2 v_t)dt + \sigma v_t dW_{2,t}, \quad (17)$$

where $\{W_{1,t}, W_{2,t}\}$ are Brownian motions in the risk-neutral measure with the correlation coefficient ρ_{12} . It is also assumed that the drift of r_t depends on v_t to incorporate the volatility feedback effect on expected changes in the short-rate. Note that given $b_0 > 0$ and $v_0 > 0$, v_t does not reach zero in finite time: see Karlin and Taylor (1981, Chapter 15.6). This property is favorable for the volatility factor.

The second is a three-factor SCT model, which is similar to the models originally proposed by Balduzzi et al. (1996, 1998):

$$dr_t = k_1(\theta_{2,t} - r_t)dt + \sigma_1 r_t dW_{1,t}, \quad (18)$$

$$d\theta_{2,t} = k_2(\theta_{3,t} - \theta_{2,t})dt + \sigma_2 \theta_{2,t} dW_{2,t}, \quad (19)$$

$$d\theta_{3,t} = k_3(\bar{\theta} - \theta_{3,t})dt + \sigma_3\theta_{3,t}dW_{3,t}, \quad (20)$$

where $\{W_{1,t}, W_{2,t}, W_{3,t}\}$ are Brownian motions in the risk-neutral measure with the correlation coefficients ρ_{ij} . Since each process can be considered as some interest rate, it is natural to assume both that the diffusion term of each process depends on its level and that the instantaneous correlations are non-zero.

To obtain reasonable parameter values, we estimate the models using weekly data on U.S. interest rates over the period from January 4, 1990 to December 28, 2005 (835 observations). The details of the estimation procedure are presented in Appendix B. Panels A and B of Table 2 present parameter values for the SV and SCT models, respectively. The actual estimates are in the columns labeled SV-P1 and SCT-P1, which serve as base cases. To examine the accuracy in relation to various aspects of the models, we change some of the parameter values from the base cases, while keeping the others unchanged. SV-P2 is characterized by faster mean-reversion of r_t : a_1 is doubled and a_0 is adjusted to keep the long-term mean of r_t , $-a_0/a_1 + (a_2b_0)/(a_1b_2)$, unchanged. SV-P3 is characterized by faster mean-reversion of v_t : both b_2 and b_0 are doubled, by which the long-term mean of v_t is unchanged. SV-P4 is characterized by lower volatility: σ is reduced to half. SCT-P2 exhibits faster mean-reversion: (k_1, k_2, k_3) are all doubled. SCT-P3 exhibits lower volatilities: $(\sigma_1, \sigma_2, \sigma_3)$ are all reduced to half.

3.2 Experimental setup

We employ the Monte Carlo (MC) method as a benchmark pricing tool. For the MC method, the continuous-time models are discretized by the Euler method with a step size of 1/250. The number of replications is set to 5,000. Antithetic variates are used for variance reduction.

The approximation error is defined as the difference between the yield computed by

the proposed method, y_τ^{ap} , and that computed by the MC method, y_τ^{mc} : $e_\tau = y_\tau^{ap} - y_\tau^{mc}$ for $\tau = 0.5, 1, 2, 3, 5, 7$, and 10.

State variable values at which the yields are evaluated are selected from the actual data. In the estimation procedure presented in Appendix B, the time series of the state variables are generated: r_t is proxied by the three-month yield for both models; v_t for the SV model is backed out from the five-year yield using the approximation formula; and $(\theta_{2,t}, \theta_{3,t})$ for the SCT model are proxied by the two-year and ten-year yields.⁴ Using these data, a kind of arbitrariness on the choice of state variable values seems to be avoidable, such as providing them with model-implied long-term means. More importantly, the time-series of the approximation errors can be generated over the same period as the data, which enables us to highlight when and in what condition the approximation is (in)accurate. For reference, the time-series of the state variables are plotted in Figure 1.⁵ Of particular note is the similarity of the plots between the SV factor, v_t , and the spread of the five-year yield over the three-month yield. In fact, the correlation between the two is 0.975. Hence, the SV factor implied by our model represents the slope factor, i.e., one of the important factors driving the term structure: see, e.g., Knez et al. (1994), and Litterman and Scheinkman (1991).

⁴We also consider various combinations of yields used for inversion or as proxies. Although the estimation results change somewhat, the accuracy results, which are of primary interest, change little from those reported below.

⁵To reduce the computational burden on the MC simulations, the sample size of the state variables is reduced to one fifth by picking up every five observations. This does not change the accuracy results from those using the full-sample data, which is checked for the base cases, SV-P1 and SCT-P1.

3.3 Numerical results for the second-order approximation

Table 3 presents mean absolute errors (MAE) expressed in basis point (bp, 1bp = 0.01%): $10^4 \times \frac{1}{T} \sum_t |e_{\tau,t}|$ for $\tau = 0.5, 1, 2, 3, 5, 7, 10$, and $T = 167$ ($= 835/5$). First, we report the results for the SV model, shown in Panel A of Table 3. The second-order approximation is very accurate for yields with maturities of up to three years: all the MAEs are within 0.4 bp. For the five-year yield, the MAEs are still within 2 bp. For maturities beyond five years, however, the accuracy becomes worse: for SV-P1, the MAE at $\tau = 10$ reaches nearly 30 bp. This inaccuracy seems to be somewhat alleviated when the speed of mean reversion of either r_t or v_t is faster (SV-P2, -P3), and when the variations in both r_t and v_t are smaller (SV-P4).

In Panels (a) and (b) of Figure 2, the time-series of the approximation errors, $e_{\tau,t}$ with $\tau = 5, 10$, are plotted for SV-P1. Obviously, the second-order approximation undervalues yields throughout the period, indicating that it fails to generate high spreads at long maturities (a possible cause of this undervaluation is explored in the next subsection). While the maximum deviation is at most -5.7 bp for the five-year yield, it is nearly -80 bp for the ten-year yield, which is much larger than the one-standard deviation of the MC error. The undervaluation is particularly evident during 1991–94 and around 2002. Looking at Figure 1(a), these periods correspond to high spreads and high volatility. In fact, when we calculate the correlations between the SV factor and approximation errors, $corr(v_t, e_{\tau,t})$ with $\tau = 5, 10$, they are -0.85 and -0.91 , respectively, indicating that the higher the volatility, the undervaluation becomes more substantial. The difficulty of the second-order approximation thus arises in a high volatility regime, which results in the failure of generating high spreads at long maturities.

Next, we report the results for the SCT model, shown in Panel B of Table 3. Similarly,

the second-order approximation is accurate for maturities of up to five years, and becomes worse for the ten-year yield. Consistent with the previous results, the accuracy improves when the volatilities of the state variables are lower (SCT-P3). However, faster mean reversion results in deterioration of the accuracy (SCT-P2), which is in contrast to the SV model. One explanation is as follows. The state variables are proxied by the 3-month, 2-year, and 10-year yields for the SCT model, which normally form a positive slope as seen in Figure 1. The increase in k_i then increases the drift and as a result the level of each process. This also increases the volatility of each process by the CEV specification (with the elasticity of one). Therefore, the performance of the second-order approximation becomes worse. Also, the finding that high interest rate levels lead to deterioration of the accuracy for the SCT model can be confirmed in Panels (c) and (d) of Figure 2, where the time-series of the approximation errors are plotted for SCT-P1. Large negative errors appear during 1990–92, when all of $(r_t, \theta_{2,t}, \theta_{3,t})$ are high as seen in Figure 1.⁶

3.4 A possible cause of the undervaluation of yields

The undervaluation of yields is equivalent to the overvaluation of bond prices, $E_t[z_{t,T}]$. Recall that $z_{t,u}$ evolves according to $dz_{t,u} = -r_u z_{t,u} du$. By integrating both sides on $u \in [t, T]$ and taking the conditional expectation, we have

$$\begin{aligned} E_t[z_{t,T}] - z_{t,t} &= - \int_t^T E_t[r_u z_{t,u}] du \\ &= - \int_t^T E_t[r_t z_{t,t} + z_{t,t}(r_u - r_t) + r_t(z_{t,u} - z_{t,t}) + (r_u - r_t)(z_{t,u} - z_{t,t})] du, \end{aligned} \quad (21)$$

Looking at the right hand side (RHS) of (21), the first term is deterministic, whereas the second term is exactly calculated for our models. Therefore, the overvaluation of $E_t[z_{t,T}]$

⁶For the SV model, on the other hand, faster mean-reversion results in improvement of the accuracy (SV-P2, -P3). This is because doubling a_1 or b_2 does not much increase interest rate levels, whereas it makes the short-rate more stable, as is the case of reducing σ to half (SV-P4).

implies the undervaluation of the last term, $E_t[(r_u - r_t)(z_{t,u} - z_{t,t})]$, at least for large u , which exceeds the overvaluation of the third term, $E_t[r_t(z_{t,u} - z_{t,t})]$.

The undervaluation of $E_t[(r_u - r_t)(z_{t,u} - z_{t,t})]$ is directly attributable to the approximation to the drift of $z_{t,u}$: $f_z(r_u, z_{t,u}) = -r_u z_{t,u}$. Specifically, in getting to the approximation formula, we apply the Ito formula to $(r_u - r_t)(z_{t,u} - z_{t,t})$ and take the conditional expectation to obtain

$$E_t[(r_u - r_t)(z_{t,u} - z_{t,t})] = \int_t^u E_t[(z_{t,a} - z_{t,t})f_r(X_a)]da + \int_t^u E_t[(r_a - r_t)f_z(r_a, z_{t,a})]da, \quad (22)$$

where $f_r(X_a)$ is the drift function of r_a , which is linear in the state vector, X_a , in our models. Hence, the first expectation on the RHS of (22) is taken to a quadratic function of $(X_a, z_{t,a})$. In other words, this can be expressed by a combination of conditional first and second moments of increments of $(X_a, z_{t,a})$. Therefore, the Taylor expansion of $f_r(X_a)$ is unnecessary for the second- or higher-order approximations. On the other hand, the second expectation on the RHS of (22) is taken to a higher than quadratic function of $(r_a, z_{t,a})$, as $f_z(r_a, z_{t,a}) = -r_a z_{t,a}$. In fact, when this is to be expressed by conditional moments of increments of $(r_a, z_{t,a})$, the following third moment is necessary: $E_t[(r_a - r_t)^2(z_{t,a} - z_{t,t})]$. For the second-order approximation, however, this is not included. Then, we need to approximate $f_z(r_a, z_{t,a})$ by the first-order Taylor expansion, so that the second expectation can be expressed by a combination of conditional first and second moments. By replacing f_z with its linear approximation, f_z^{ap} , in (22), the approximation formula is now effective.

f_z^{ap} is given by

$$\begin{aligned} f_z^{ap}(r_a, z_{t,a}) &= f_z(r_t, z_{t,t}) + f_z^{(1,0)}(r_t, z_{t,t})(r_a - r_t) + f_z^{(0,1)}(r_t, z_{t,t})(z_{t,a} - z_{t,t}) \\ &= -(r_a + r_t z_{t,a} - r_t). \end{aligned} \quad (23)$$

Taking the difference between f_z^{ap} and f_z leads to

$$f_z^{ap} - f_z = -(r_a + r_t z_{t,a} - r_t) + r_a z_{t,a} = (r_a - r_t)(z_{t,a} - 1). \quad (24)$$

The above difference is more likely to be negative. That is, f_z^{ap} is more negative on average than f_z . This is because $(r_u - r_t) > 0$ is more likely given that the risk-neutral drift of the short-rate is positive over a sufficiently wide range of the short-rate, while $(z_{t,u} - 1) < 0$ holds certainly. Therefore, when f_z is replaced with f_z^{ap} in (22), $E_t[(r_u - r_t)(z_{t,u} - z_{t,t})]$ is undervalued by the approximation formula.

As seen, the undervaluation of $E_t[(r_u - r_t)(z_{t,u} - z_{t,t})]$ is a cause of the overvaluation of the bond price by (21), which in turn is attributed to the linear approximation of f_z . Conversely, this approximation is unnecessary for evaluating $E_t[(r_u - r_t)(z_{t,u} - z_{t,t})]$ if we consider up to conditional third moments. Then, the third-order approximation can be expected to improve the accuracy. In addition, among the conditional third moments required, the above argument suggests that $E_t[(r_u - r_t)^2(z_{t,u} - z_{t,t})]$ may have a fundamental role for the improvement, if it exists. The next subsection verifies these predictions.

3.5 Numerical results for the third-order approximation

We perform the same numerical experiments to examine the accuracy of the third-order approximation. To highlight its effect, we continuously use the same parameter values as presented in Table 2 and, more importantly, the same time-series of the state variables as used for the second-order approximation.⁷

Panels A and B of Table 4 present the MAEs of the third-order approximation for the SV and SCT models, respectively, showing that the improvement of the accuracy is substantial: for the seven-year yield, all MAEs are within 0.5 bp, and for the ten-year yield the MAEs range from 0.4 bp to 4.4 bp. Figure 3 also graphs the time-series of errors of the

⁷For the SV model, the third-order approximation generates slightly different values of the implied SV factor from those already obtained by the second-order approximation. We also run MC simulations using the newly obtained SV values and calculate errors of the third-order approximation. The results are unchanged from those using the already obtained SV values presented in Table 4.

third-order approximation for SV-P1 and SCT-P1. The errors for the five-year yield do not exhibit a systematic pattern but fluctuate in a narrow range of ± 0.5 bp. The errors for the ten-year yield now turn to be positive, indicating that the third-order approximation overvalues longer-term yields. However, the magnitude of errors is much smaller than that of the second-order approximation.

Next, we examine whether $E_t[(r_T - r_t)^2(z_{t,T} - z_{t,t})]$ has a significant impact on the improvement of the accuracy. Aiming at reducing the computational burden, we add this moment alone to the moment vector for the second-order approximation, rather than extract this moment alone from the moment vector for the third-order approximation. Panels C and D of Table 4 present the MAEs of the extended second-order approximation for both models, showing that our prediction is correct. Though the magnitude of MAEs increases from that of the third-order approximation, it is much smaller for the SV model and marginally smaller for the SCT model than that of the second-order approximation. Figure 4 also graphs the time-series of errors of the extended second-order approximation. For SV-P1, by adding $E_t[(r_T - r_t)^2(z_{t,T} - z_{t,t})]$ alone, the undervaluation of yields is substantially corrected: occasionally even positive errors appear. For SCT-P1, although the pattern of errors is similar to that shown in Panels (c) and (d) of Figure 2, the magnitude of errors is somewhat decreased.

The last finding is particularly useful for high-dimensional models, where the length of a moment vector rapidly increases by higher-order approximations: in general, the length of $\Psi_{t,T}$ consisting of up to n -th conditional moments of a d -dimensional process is $\binom{n+d}{n} - 1 = (n+d)!/(n!d!) - 1$. However, by carefully selecting effective moments, it may be possible that this increase is marginal while higher accuracy is achieved.

4 Concluding Remarks

We proposed an analytical approximation of the term structure of interest rates under general diffusion processes. This is obtained as one of the elements of a moment vector, which is approximated as the solution to a system of ordinary differential equations. Based on two illustrative models, we showed that for maturities of up to five years, the second-order approximation is accurate, which seems to justify matching the approximation formula exactly with the five-year yield to obtain a model-implied latent process. Beyond five years to maturity, the third-order approximation is effective. We also showed that inclusion of $E_t[(r_T - r_t)^2(z_{t,T} - z_{t,t})]$ is the key to the improvement of the accuracy without much increasing the computational burden.

The proposed method can also be utilized for computing the characteristic function, which in turn allows for computing option prices through the Fourier inversion technique: e.g., Bakshi and Madan (2000), and Duffie et al. (2000). The accuracy of the approximation in terms of option prices is worth examining, which is left for future research. Another direction of future research is to estimate the behavior of latent processes implicit in term structure data, using models that have little tension in explaining both time-series and cross-sectional dimensions of the data.

Acknowledgement

We are grateful to the three anonymous referees for comments and suggestions, which enabled us to significantly improve the paper.

Appendix A: Convergence of the Taylor series

Approximation by Taylor series is used to prove the approximate conditional moments proposed by Shoji (2002) converge to the true ones. So, it's also important we should pay our attention to the convergence of the Taylor series when considering how close its approximation is to the true function. Below, we give a simple sketch of how the convergence is achieved. By construction of the proposed method that approximates conditional moments of diffusion processes, we apply the Taylor expansion inside of the conditional expectation. Then, we can prove a truncated Taylor series, taking the first order approximation for example, converges in mean square to the true value as the sampling interval goes to zero. More specifically, let $f(x)$ and $\hat{f}(x)$ be the true function and its first order truncated Taylor expansion around x_0 , respectively. Substituting processes X_t and X_s ($s \leq t$) for x and x_0 , we can show $\lim_{\Delta t \rightarrow 0} E_s[|f(X_t) - \hat{f}(X_t)|^2] = 0$ where $\Delta t = t - s$.

Actually, according to the Taylor theorem, $f(x) = \hat{f}(x) + (f''(x_\theta)/2)(x - x_0)^2$, where $x_\theta = (1 - \theta)x_0 + \theta x$ for some $\theta \in [0, 1]$. So, $E_s[|f(X_t) - \hat{f}(X_t)|^2] = E_s[(f''(X_\theta)/2)^2(X_t - X_s)^4]$. We can show the right hand side converges to zero by (i) the localization argument by a stopping time that is often used in stochastic calculus (e.g., Karatzas and Shreve, 1991, and Rogers and Williams, 1994), and (ii) the fact that $E_s[(X_t - X_s)^{2m}] = O((\Delta t)^m)$ (Shoji, 2002).

The outline of the proof is as follows. Take any positive number K and then consider a stopped process $X_{t \wedge T}$ for a stopping time T defined by $T = \inf\{t \geq s | X_t \geq K\}$. Assume X_t as $X_{t \wedge T}$. Firstly, we show the convergence for the stopped process which is bounded by definition. By the localization of X_t , $f''(X_\theta)$ is bounded as long as f belongs to at least C^2 -class. And then, by the rate of convergence in $E_s[(X_t - X_s)^{2m}]$, we can show $\lim_{\Delta t \rightarrow 0} E_s[|f(X_t) - \hat{f}(X_t)|^2] = 0$. Lastly, letting $K \rightarrow \infty$, the proof is completed. The

statement as above is valid as well when we use the n -th order truncated Taylor expansion of f which belongs to at least C^{n+1} -class.

Appendix B: Estimation Method

Since the estimation is not a main focus of the paper, a simple procedure is employed, which nonetheless seems to provide reasonable values. Following Chen and Scott (1993), Duffee (2000), and Pearson and Sun (1994), we employ the quasi-maximum likelihood method, where both time-series and cross-sectional dimensions of term structure data are utilized. The weekly (Wednesday) data consist of U.S. LIBOR rates with maturities of 3, 6, and 12 months, and U.S. Swap rates with maturities of 2, 3, 4, 5, 7, and 10 years.⁸ Through linear interpolation to these rates, implicit discount bond prices and their yields to maturity are calculated. This yield data, with maturities of 0.25, 1, 2, 3, 5, 7, and 10 years, are actually used for the estimation.

For both models, the short-rate r_t is proxied by the three-month yield. For the SV model, v_t is backed out from the five-year yield, $y_{5,t}$, by assuming that it is exactly explained by the approximation formula at any point in time. Specifically, letting \tilde{Y} denote the second-order approximation of the true yield function, we invert the equation, $y_{5,t} = \tilde{Y}(r_t, v_t, 5)$, for v_t . For the SCT model, $(\theta_{2,t}, \theta_{3,t})$ are proxied by the two-year and ten-year yields, $(y_{2,t}, y_{10,t})$. The rest of the yields, denoted as $y_{\tau,t}^c$, are explained with measurement errors, $\epsilon_{\tau,t}$:

$$y_{\tau,t}^c = \tilde{Y}(X_t, \tau) + \epsilon_{\tau,t} \quad \epsilon_{\tau,t} | X_t \sim i.i.d.N(0, \xi^2), \quad (25)$$

where $X_t = (r_t, v_t)$ for the SV model and $X_t = (r_t, \theta_{2,t}, \theta_{3,t})$ for the SCT model. The

⁸When one or more of the Wednesday observations are missing, we choose another day of the week on which all rates are available in the following order: Tuesday, Thursday, Friday, and Monday.

distributional assumptions for $\epsilon_{.,t}$ conditioned on X_t are summarized as: all errors have the same variance, ξ^2 , and they are contemporaneously and serially independent.

For the SV model, the density function at time t conditioned on time $t - \Delta$ is

$$\begin{aligned} p(r_t, y_{5,t}, \{y_{\tau_i,t}^c\}; r_{t-\Delta}, y_{5,t-\Delta}) &= p(X_t, \{\epsilon_{\tau_i,t}\}; X_{t-\Delta}) \left| \frac{\partial \tilde{Y}(X_t, 5)}{\partial v_t} \right|^{-1} \\ &= p_T(X_t; X_{t-\Delta}) \left| \frac{\partial \tilde{Y}(X_t, 5)}{\partial v_t} \right|^{-1} \times \prod_i p_C(\epsilon_{\tau_i,t}; X_t). \end{aligned} \quad (26)$$

The first equality follows from the change of variables. The second equality follows from the decomposition into the time-series (marginal) and cross-sectional (conditional) components with noticing that X_t is Markovian. Note that the Jacobian, $\partial \tilde{Y}(X_t, 5)/\partial v_t$, is also based on the second-order approximation of the yield function, which is already obtained in the iteration procedure for backing out v_t . The density function for the SCT model is similar but simpler, as X_t is proxied by the observed yield vector so the Jacobian does not appear.

In computing p_T , the SDEs for X_t in the actual measure are required. To obtain them, we need to specify the market prices of risk. Here, a simple specification will suffice, as the physical dynamics are actually irreverent to the assessment of the accuracy. For the SV model, the prices of risk for uncertain variations in r_t and v_t are λ_1/v_t and λ_2 (constant), respectively.⁹ For the SCT model, the prices of risk are all assumed to be constant. Then, the resulting drift terms in the physical measure are also linear for both models. Still, since analytical expressions for p_T are unknown, it is assumed to be the normal density function, where the first and second conditional moments of X_t in the actual measure are substituted. Using the proposed method, this computation is straightforward.

⁹The former specification is the so-called extended market price of risk, proposed by Cheridito et al. (2007), which is consistent with non-arbitrage if v_t , given $v_0 > 0$, does not reach zero in finite time under both measures. For the SV model considered here, the condition is $b_0 > 0$, which indeed is implied by the data.

References

- Ahn, D., and B. Gao, 1999, "A parametric nonlinear model of term structure dynamics," *Review of Financial Studies*, 12, 721-762.
- 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.
- Bakshi, G., and D. Madan, 2000, "Spanning and derivative-security valuation," *Journal of Financial Economics* 55, 205-238.
- 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.
- Ball, C. A., and W. N. Torous, 1999, "The stochastic volatility of short-term interest rates: some international evidence," *Journal of Finance*, 54, 2339-2359.
- Brenner, R. J., R. H. Harjes, and K. F. Kroner, 1996, "Another look at models of the short-term interest rate," *Journal of Financial and Quantitative Analysis*, 31, 85-107.
- 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," *Journal of Finance*, 47, 1209-1227.
- Chen, R., and L. Scott, 1993, "Maximum likelihood estimation for a multifactor equilibrium model of the term structure of interest rates," *Journal of Fixed Income*, 3, 14-31.
- Cheridito, P., D. Filipović, and R. L. Kimmel, 2007, "Market price of risk specifications for affine models: theory and evidence," *Journal of Financial Economics*, 83, 123-170.

- Duffee, G. R., 2002, "Term premia and interest rate forecasts in affine models," *Journal of Finance*, 57, 405-443.
- Duffie, D., J. Pan, and K. Singleton, 2000, "Transform analysis and asset pricing for affine jump-diffusions," *Econometrica*, 68, 1343-1376.
- Durham, G. B., 2003, "Likelihood-based specification analysis of continuous-time models of the short-term interest rate," *Journal of Financial Economics*, 70, 463-487.
- 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.
- Karatzas, I., and S. E. Shreve, 1991, *Brownian motion and stochastic calculus*, 2nd ed., Springer.
- Karlin, S., and H. M. Taylor, 1981, *A second course in stochastic processes*, Academic Press.
- Knez, P. J., R. Litterman, and J. Scheinkman, 1994, "Explorations into factors explaining money market returns," *Journal of Finance*, 49, 1861-1882.
- Litterman, R., and J. Scheinkman, 1991, "Common factors affecting bond returns," *Journal of Fixed Income*, June, 54-61.
- 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.
- Pearson, N. D., and T. Sun, 1994, "Exploiting the conditional density in estimating the term structure: an application to the Cox, Ingersoll, and Ross model," *Journal of Finance*, 49, 1279-1304.
- Rogers, L. C. G., and D. Williams, 1994, *Diffusions, Markov processes, and martingales, Vol.1*, 2nd ed., Wiley.

Shoji, I., 2002, "Approximation of conditional moments of diffusion processes," *International Journal of Computational and Numerical Analysis and Applications*, 1, 163-190.

Takamizawa, H., and I. Shoji, 2003, "Modeling the term structure of interest rates with general short-rate models," *Finance and Stochastics*, 7, 323-335.

Takamizawa, H., and I. Shoji, 2004, "On accuracy of local linear approximation for the term structure of interest rates," *Quantitative Finance*, 4, 151-157.

$[A]_{i,j}$	1	2	3	4	5
1	$f_1^{(1,0,0)}$	$f_1^{(0,1,0)}$	$f_1^{(0,0,1)}$	$f_1^{(2,0,0)}/2$	$f_1^{(0,2,0)}/2$
2	$f_2^{(1,0,0)}$	$f_2^{(0,1,0)}$	$f_2^{(0,0,1)}$	$f_2^{(2,0,0)}/2$	$f_2^{(0,2,0)}/2$
3	$f_3^{(1,0,0)}$	$f_3^{(0,1,0)}$	$f_3^{(0,0,1)}$	$f_3^{(2,0,0)}/2$	$f_3^{(0,2,0)}/2$
4	$2f_1 + g_{11}^{(1,0,0)}$	$g_{11}^{(0,1,0)}$	$g_{11}^{(0,0,1)}$	$2f_1^{(1,0,0)} + g_{11}^{(2,0,0)}/2$	$g_{11}^{(0,2,0)}/2$
5	$g_{22}^{(1,0,0)}$	$2f_2 + g_{22}^{(0,1,0)}$	$g_{22}^{(0,0,1)}$	$g_{22}^{(2,0,0)}/2$	$2f_2^{(0,1,0)} + g_{22}^{(0,2,0)}/2$
6	$g_{33}^{(1,0,0)}$	$g_{33}^{(0,1,0)}$	$2f_3 + g_{33}^{(0,0,1)}$	$g_{33}^{(2,0,0)}/2$	$g_{33}^{(0,2,0)}/2$
7	$f_2 + g_{12}^{(1,0,0)}$	$f_1 + g_{12}^{(0,1,0)}$	$g_{12}^{(0,0,1)}$	$f_2^{(1,0,0)} + g_{12}^{(2,0,0)}/2$	$f_1^{(0,1,0)} + g_{12}^{(0,2,0)}/2$
8	$f_3 + g_{13}^{(1,0,0)}$	$g_{13}^{(0,1,0)}$	$f_1 + g_{13}^{(0,0,1)}$	$f_3^{(1,0,0)} + g_{13}^{(2,0,0)}/2$	$g_{13}^{(0,2,0)}/2$
9	$g_{23}^{(1,0,0)}$	$f_3 + g_{23}^{(0,1,0)}$	$f_2 + g_{23}^{(0,0,1)}$	$g_{23}^{(2,0,0)}/2$	$f_3^{(0,1,0)} + g_{23}^{(0,2,0)}/2$
$[b]_j$	f_1	f_2	f_3	g_{11}	g_{22}

Table 1:

6	7	8	9
$f_1^{(0,0,2)}/2$	$f_1^{(1,1,0)}$	$f_1^{(1,0,1)}$	$f_1^{(0,1,1)}$
$f_2^{(0,0,2)}/2$	$f_2^{(1,1,0)}$	$f_2^{(1,0,1)}$	$f_2^{(0,1,1)}$
$f_3^{(0,0,2)}/2$	$f_3^{(1,1,0)}$	$f_3^{(1,0,1)}$	$f_3^{(0,1,1)}$
$g_{11}^{(0,0,2)}/2$	$2f_1^{(0,1,0)} + g_{11}^{(1,1,0)}$	$2f_1^{(0,0,1)} + g_{11}^{(1,0,1)}$	$g_{11}^{(0,1,1)}$
$g_{22}^{(0,0,2)}/2$	$2f_2^{(1,0,0)} + g_{22}^{(1,1,0)}$	$g_{22}^{(1,0,1)}$	$2f_2^{(0,0,1)} + g_{22}^{(0,1,1)}$
$2f_3^{(0,0,1)} + g_{33}^{(0,0,2)}/2$	$g_{33}^{(1,1,0)}$	$2f_3^{(1,0,0)} + g_{33}^{(1,0,1)}$	$2f_3^{(0,1,0)} + g_{33}^{(0,1,1)}$
$g_{12}^{(0,0,2)}/2$	$f_1^{(1,0,0)} + f_2^{(0,1,0)} + g_{12}^{(1,1,0)}$	$f_2^{(0,0,1)} + g_{12}^{(1,0,1)}$	$f_1^{(0,0,1)} + g_{12}^{(0,1,1)}$
$f_1^{(0,0,1)} + g_{13}^{(0,0,2)}/2$	$f_3^{(0,1,0)} + g_{13}^{(1,1,0)}$	$f_1^{(1,0,0)} + f_3^{(0,0,1)} + g_{13}^{(1,0,1)}$	$f_1^{(0,1,0)} + g_{13}^{(0,1,1)}$
$f_2^{(0,0,1)} + g_{23}^{(0,0,2)}/2$	$f_3^{(1,0,0)} + g_{23}^{(1,1,0)}$	$f_2^{(1,0,0)} + g_{23}^{(1,0,1)}$	$f_2^{(0,1,0)} + f_3^{(0,0,1)} + g_{23}^{(0,1,1)}$
g_{33}	g_{12}	g_{13}	g_{23}

Table 1 (continued): The elements of $A(X_s)$ and $b(X_s)$ in the approximation formula

A moment vector $\Psi_s(t)$ contains up to conditional second moments of a three-dimensional process, $X'_t = (x_{1,t} \ x_{2,t} \ x_{3,t})$: the first three elements are $E_s[x_{i,t} - x_{i,s}]$ ($i = 1, 2, 3$), the next three elements are $E_s[(x_{i,t} - x_{i,s})^2]$ ($i = 1, 2, 3$), and the last three elements are $E_s[(x_{i,t} - x_{i,s})(x_{j,t} - x_{j,s})]$ ($(i, j) = (1, 2), (1, 3), (2, 3)$). $x_{i,t}$ follows $dx_{i,t} = f_i(X_t)dt + \sigma_i(X_t)'dW_t$, where W_t is three-dimensional Brownian motion. The notations in the table are as follows: $g_{ij} = \sigma_i'\sigma_j$,

$$f_i^{(k,l,m)} = \frac{\partial f_i^{k+l+m}}{\partial x_1^k \partial x_2^l \partial x_3^m}, \quad \text{and} \quad g_{ij}^{(k,l,m)} = \frac{\partial g_{ij}^{k+l+m}}{\partial x_1^k \partial x_2^l \partial x_3^m}.$$

Panel A: Parameter values for the SV model				
	SV-P1	SV-P2	SV-P3	SV-P4
a_0	-0.034	-0.028	-0.034	-0.034
a_1	-0.048	-0.096	-0.048	-0.048
a_2	0.291	0.291	0.291	0.291
b_0	0.032	0.032	0.064	0.032
b_2	-0.229	-0.229	-0.458	-0.229
σ	0.125	0.125	0.125	0.063
ρ_{12}	-0.143	-0.143	-0.143	-0.143

Panel B: Parameter values for the SCT model			
	SCT-P1	SCT-P2	SCT-P3
k_1	1.133	2.266	1.133
k_2	0.712	1.424	0.712
k_3	0.010	0.020	0.010
$\bar{\theta}$	0.263	0.263	0.263
σ_1	0.157	0.157	0.078
σ_2	0.232	0.232	0.116
σ_3	0.148	0.148	0.074
ρ_{12}	0.533	0.533	0.533
ρ_{13}	0.314	0.314	0.314
ρ_{23}	0.756	0.756	0.756

Table 2: Parameter values for the SV and SCT models

The SV and SCT models are given by

SV model:

$$dr_t = (a_0 + a_1 r_t + a_2 v_t)dt + v_t r_t dW_{1,t},$$

$$dv_t = (b_0 + b_2 v_t)dt + \sigma v_t dW_{2,t},$$

SCT model:

$$dr_t = k_1(\theta_{2,t} - r_t)dt + \sigma_1 r_t dW_{1,t},$$

$$d\theta_{2,t} = k_2(\theta_{3,t} - \theta_{2,t})dt + \sigma_2 \theta_{2,t} dW_{2,t},$$

$$d\theta_{3,t} = k_3(\bar{\theta} - \theta_{3,t})dt + \sigma_3 \theta_{3,t} dW_{3,t},$$

and the correlation coefficients between Brownian motions are ρ_{ij} . The columns of SV-P1 and SCT-P1 present the actual estimates using data on U.S. LIBOR and Swap rates over 1990–2005. Alternative sets of parameter values are obtained based on them.

τ	0.5	1	2	3	5	7	10
Panel A: MAE of the second-order approximation for the SV model							
SV-P1	0.15	0.16	0.20	0.37	1.91	7.08	29.10
SV-P2	0.14	0.14	0.17	0.32	1.61	6.01	25.39
SV-P3	0.18	0.18	0.23	0.39	1.68	5.44	19.67
SV-P4	0.15	0.16	0.19	0.34	1.55	5.24	20.08
Panel B: MAE of the second-order approximation for the SCT model							
SCT-P1	0.17	0.18	0.24	0.42	1.69	5.03	17.06
SCT-P2	0.35	0.36	0.43	0.70	2.60	7.86	28.59
SCT-P3	0.17	0.17	0.20	0.31	0.96	2.64	8.47

Table 3: Mean absolute errors (MAE) of the second-order approximation

The MAE (expressed in basis point) is given by $10^4 \times \frac{1}{T} \sum_t |e_{\tau,t}|$ ($T = 167$), where $e_{\tau,t} = y_{\tau,t}^{ap} - y_{\tau,t}^{mc}$, and where $y_{\tau,t}^{ap}$ and $y_{\tau,t}^{mc}$ are the τ -maturity yields computed by the approximation and MC methods, respectively.

τ	0.5	1	2	3	5	7	10
Panel A: MAE of the third-order approximation for the SV model							
SV-P1	0.15	0.15	0.15	0.15	0.14	0.47	4.35
SV-P2	0.14	0.14	0.14	0.14	0.12	0.29	2.91
SV-P3	0.18	0.18	0.18	0.16	0.14	0.34	2.52
SV-P4	0.15	0.15	0.15	0.16	0.15	0.27	2.26
Panel B: MAE of the third-order approximation for the SCT model							
SCT-P1	0.17	0.18	0.18	0.18	0.15	0.28	1.95
SCT-P2	0.35	0.35	0.32	0.26	0.15	0.39	3.29
SCT-P3	0.17	0.17	0.17	0.17	0.12	0.06	0.38
Panel C: MAE of the extended second-order approximation for the SV model							
SV-P1	0.15	0.15	0.16	0.18	0.46	1.59	6.64
SV-P2	0.14	0.14	0.14	0.17	0.51	1.95	8.85
SV-P3	0.18	0.18	0.18	0.17	0.29	1.33	6.86
SV-P4	0.15	0.15	0.16	0.17	0.23	0.62	3.56
Panel D: MAE of the extended second-order approximation for the SCT model							
SCT-P1	0.17	0.18	0.22	0.36	1.45	4.52	15.93
SCT-P2	0.35	0.36	0.41	0.64	2.41	7.47	27.72
SCT-P3	0.17	0.17	0.19	0.26	0.80	2.27	7.65

Table 4: Mean absolute errors (MAE) of the third-order and extended second-order approximations

The MAE (expressed in basis point) is given by $10^4 \times \frac{1}{T} \sum_t |e_{\tau,t}|$ ($T = 167$), where $e_{\tau,t} = y_{\tau,t}^{ap} - y_{\tau,t}^{mc}$, and where $y_{\tau,t}^{ap}$ and $y_{\tau,t}^{mc}$ are the τ -maturity yields computed by the approximation and MC methods, respectively. The moment vector for the third-order approximation contains up to conditional third moments, whereas the moment vector for the extended second-order approximation includes only $E_t[(r_T - r_t)^2(z_{t,T} - z_{t,t})]$ in addition to conditional first and second moments.

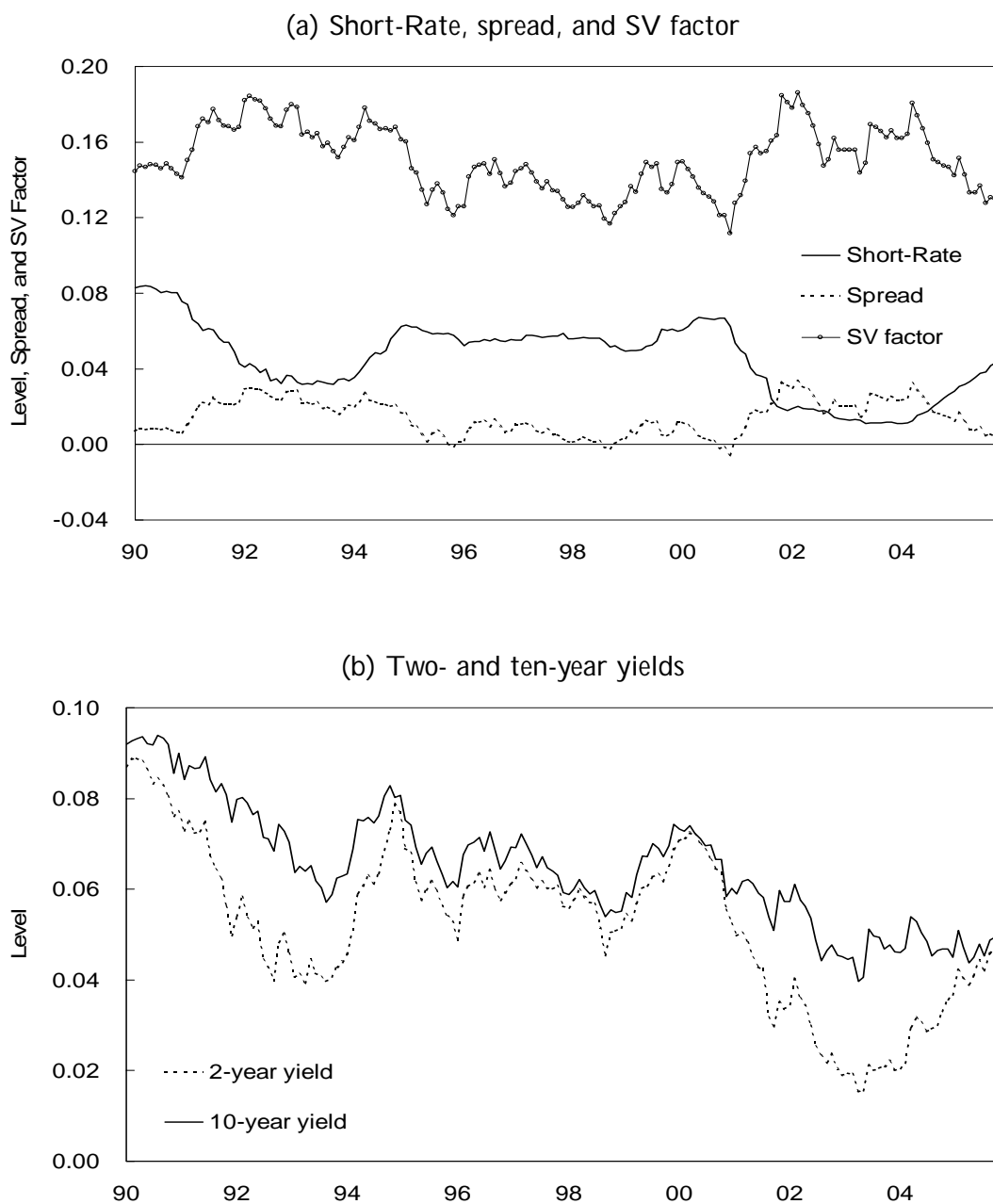


Figure 1: Time-series of the state variables over 1990–2005

The data consist of discount bond yields implicit in the U.S. LIBOR and Swap rates. Panel (a) displays the short-rate proxied by the 3-month yield, the spread between the 5-year and 3-month yields, and the SV factor calculated from the SV model. Panel (b) displays the 2- and 10-year yields, which are used as proxies for $(\theta_{2,t}, \theta_{3,t})$ in the SCT model.

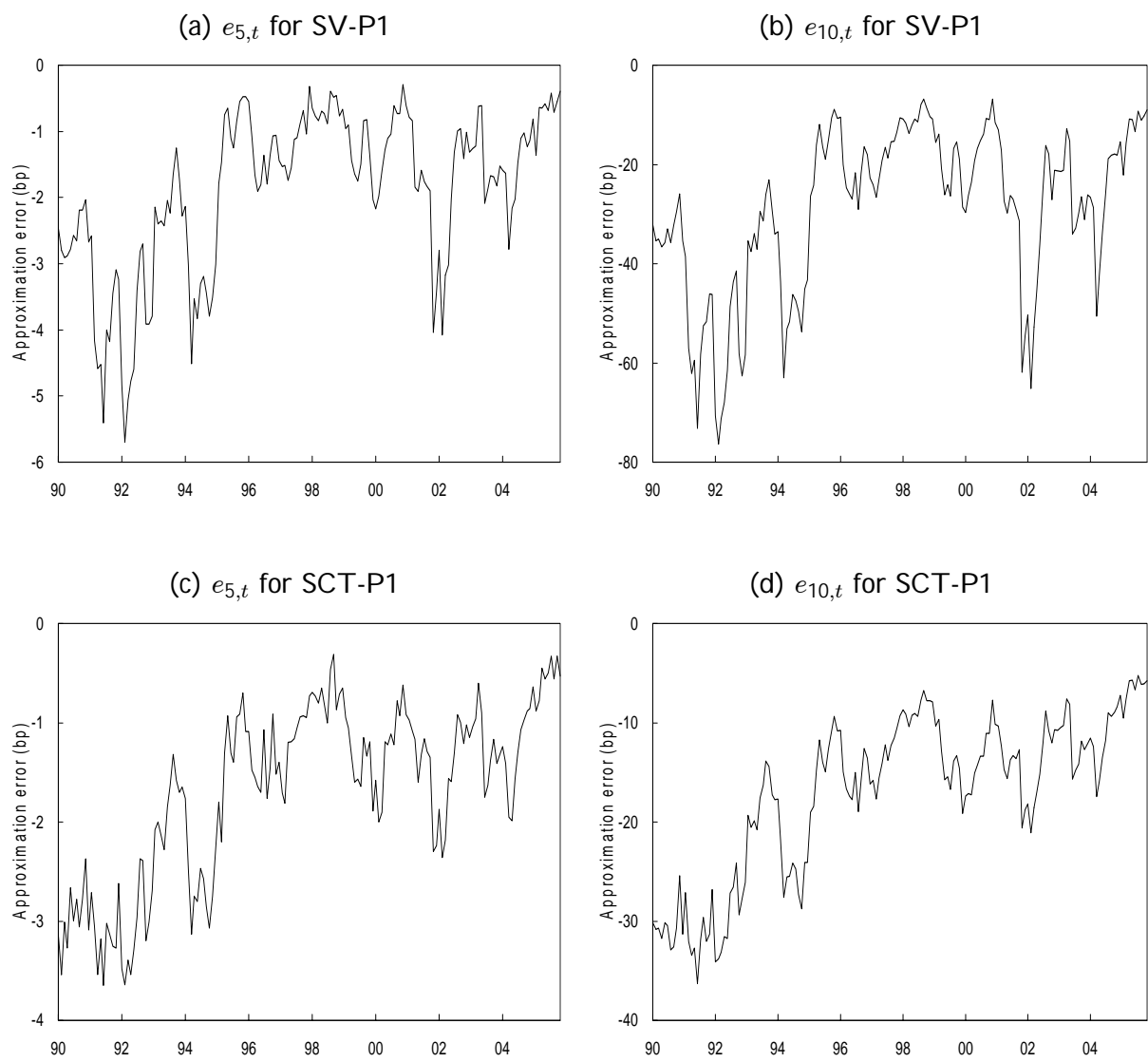


Figure 2: Time-series of errors of the second-order approximation over 1990–2005. The approximation error is defined as $e_{\tau,t} = y_{\tau,t}^{ap} - y_{\tau,t}^{mc}$, where $y_{\tau,t}^{ap}$ and $y_{\tau,t}^{mc}$ are the τ -maturity yields computed by the approximation and MC methods, respectively.

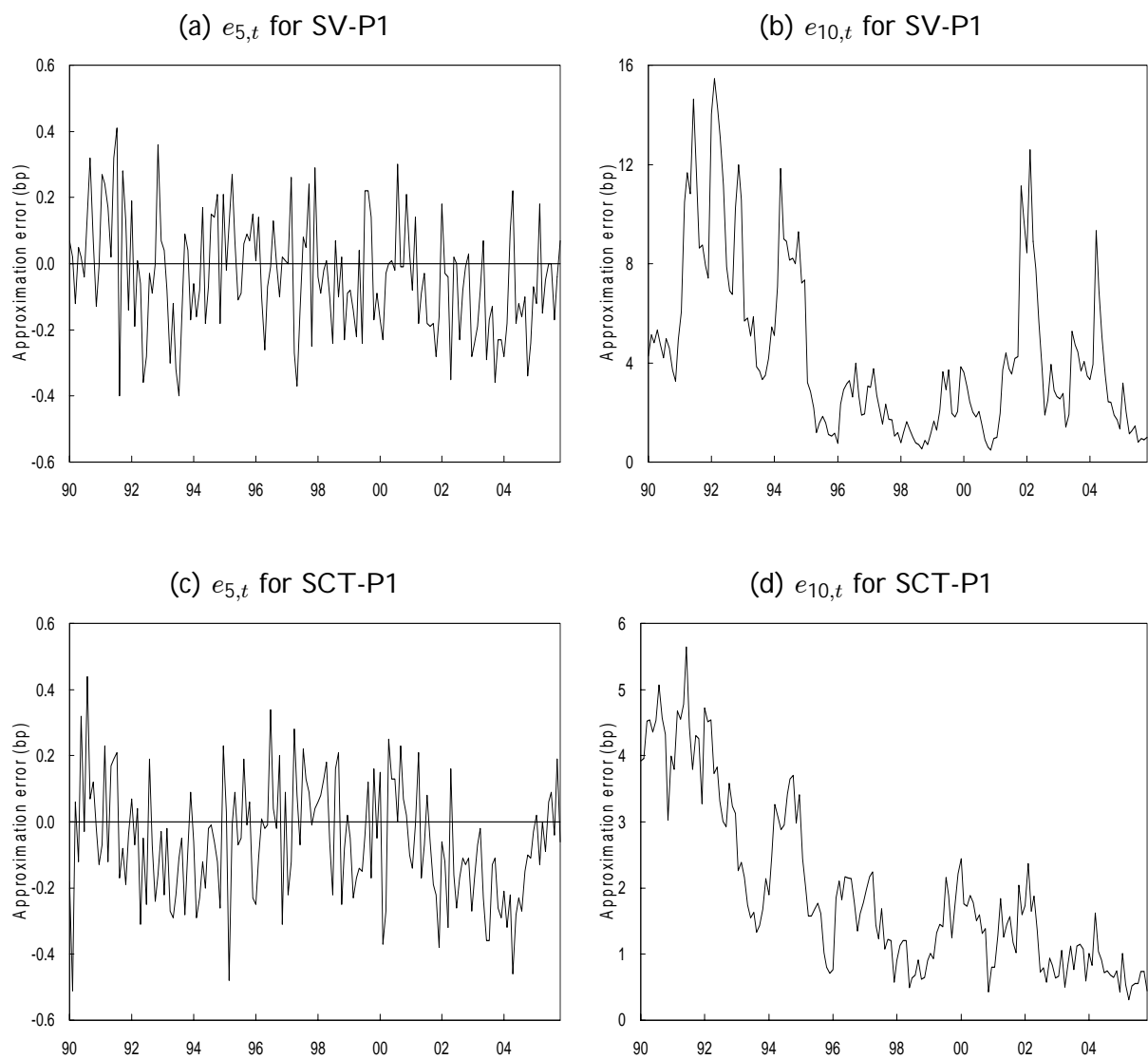


Figure 3: Time-series of errors of the third-order approximation over 1990–2005. The approximation error is defined as $e_{\tau,t} = y_{\tau,t}^{ap} - y_{\tau,t}^{mc}$, where $y_{\tau,t}^{ap}$ and $y_{\tau,t}^{mc}$ are the τ -maturity yields computed by the approximation and MC methods, respectively. The moment vector for the third-order approximation contains up to conditional third moments.

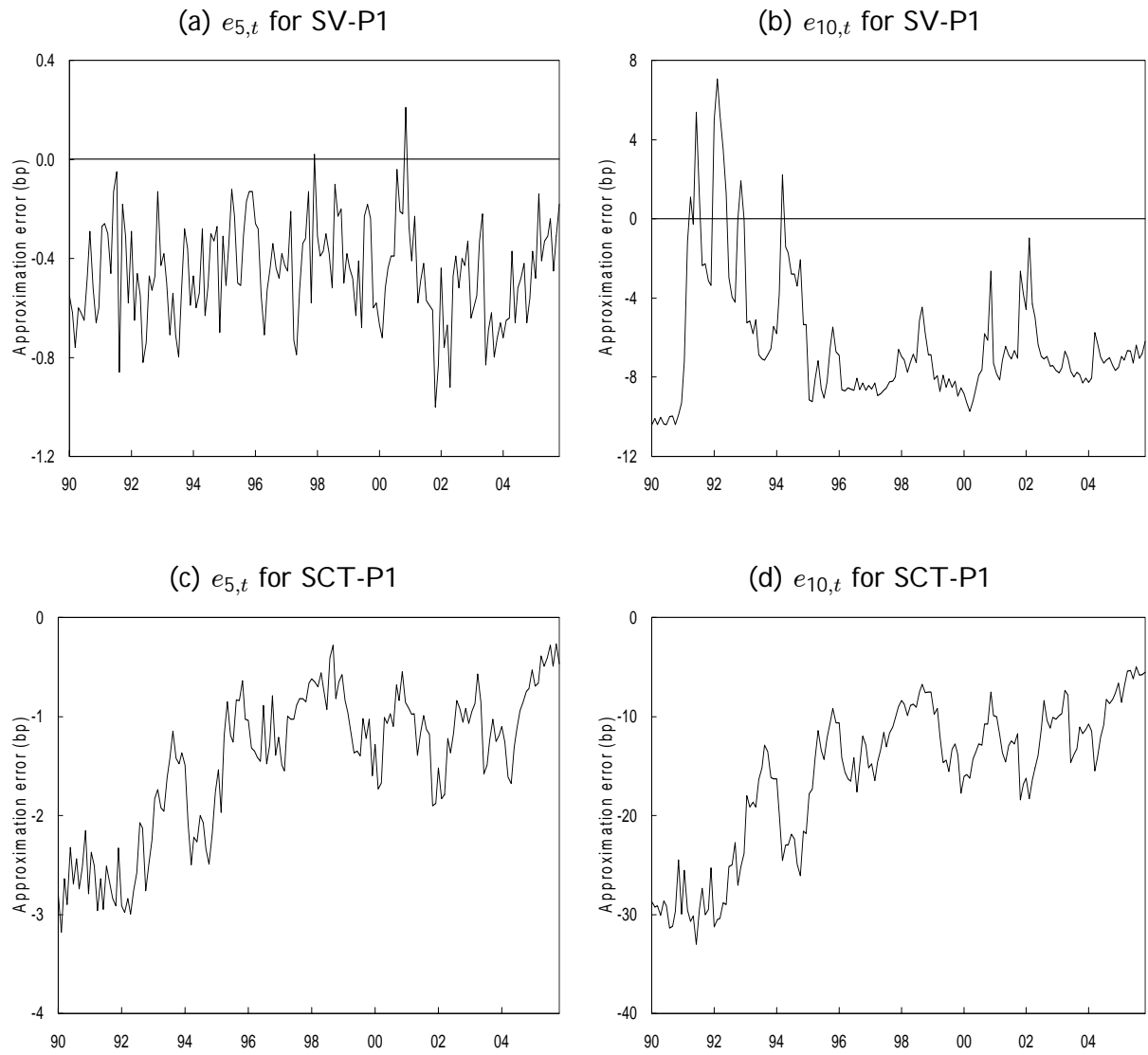


Figure 4: Time-series of errors of the extended second-order approximation over 1990–2005

The approximation error is defined as $e_{\tau,t} = y_{\tau,t}^{ap} - y_{\tau,t}^{mc}$, where $y_{\tau,t}^{ap}$ and $y_{\tau,t}^{mc}$ are the τ -maturity yields computed by the approximation and MC methods, respectively. The moment vector for the extended second-order approximation includes only $E_t[(r_T - r_t)^2(z_{t,T} - z_{t,t})]$ in addition to conditional first and second moments.