Using Proxies for the Short Rate: When are Three Months Like an Instant?

David A. Chapman*† John B. Long, Jr.‡ Neil D. Pearson§

September 23, 1998

Abstract

The dynamics of the unobservable "short" or "instantaneous" rate of interest are frequently estimated using a proxy. We show how the biases resulting from this practice (the "proxy problem") are related to the derivatives of the proxy with respect to the short rate and the (inverse) function from the proxy to the short rate. Analytic results show that the proxy problem is not economically significant for single-factor affine models, for parameter values consistent with US data. In addition, for the two-factor affine model of Longstaff and Schwartz (1992), the proxy problem is only economically significant for pricing discount bonds with maturities of more than five years. We also describe two different procedures which can be used to assess the magnitude of the proxy problem in more general interest rate models. Numerical evaluation of a nonlinear single-factor model suggests that the proxy problem can significantly affect both estimates of the diffusion function and discount bond prices.

*We thank Murray Carlson, Narasihman Jegadeesh, an anonymous referee, and the editor (Bernard Dumas) for extremely helpful comments and Xiongwei Ju for checking some of the computations. Neil Pearson was partially supported by the Office for Futures and Options Research at the University of Illinois.

†Finance Department, Graduate School of Business, The University of Texas at Austin, Austin, Texas 78712-1179. Phone: (512) 471-6621. e-mail: chapman@eco.utexas.edu. The most recent version of this paper is available online at http://www.bus.utexas.edu/~chapmand/clp.html

‡William E. Simon Graduate School of Business Admininstration, University of Rochester, Rochester, New York 14627. Phone: (716) 275-3358. e-mail: long@mail.ssb.rochester.edu

§Finance Department, University of Illinois at Urbana-Champaign, 340 Commerce West, 1206 South Sixth Street, Champaign, Illinois 61820. Phone: (217) 244-0490. e-mail: pearson2@uiuc.edu.

1 Introduction

There is a large recent empirical literature devoted to examining the time series properties of short-term interest rate data. Examples of this research include Chan, Karolyi, Longstaff, and Sanders (1992), Aït-Sahalia (1996a,b), Conley, Hansen, Luttmer, and Scheinkman (1997) (hereafter CHLS), Andersen and Lund (1997), and Stanton (1997). These authors use a variety of econometric techniques including simple method of moments, efficient method of moments, and nonparametric estimators. Regardless of technique, this research is often motivated, explicitly or implicitly, by a desire to understand the time series properties of the (unobservable) rate of return on a default-free bond maturing in the next instant of time, the so-called "instantaneous" or "short" rate. This rate often plays a critical role in simple models of the term structure of interest rates. See, for example, Vasicek (1977), Cox, Ingersoll and Ross (1985) (hereafter CIR), or Brennan and Schwartz (1979).

Empirical research on the time-series properties of the short rate typically requires specifying a proxy for this unobservable variable. The papers cited above make a variety of different choices. For example, Anderson and Lund (1997) and Stanton (1997) use the yield on a 3-month Treasury bill as a proxy for the short rate, while Chan et al (1992) use the 1-month Treasury bill yield. In contrast, Aït-Sahalia (1996a,b) uses the 1-week Eurodollar rate, and CHLS use the Federal Funds rate. In making these choices, researchers face a clear trade-off. Longer term rates, like the yield on a 3-month Treasury bill, may deviate in important ways from the unobservable short rate. For example, the CIR model provides an explicit formula for the yield of a three-month discount bond, and it clearly differs from the instantaneous rate. On the other hand, there are potentially serious microstructure or institutional features of the markets for very short-term instruments that make the use of these data problematic in estimating and testing simple term structure models. See, for example, the bi-weekly settlement effect in the Federal Funds market documented in Hamilton (1996) or the "cash equivalence" effect on the prices of 1-month Treasury bills discussed in Knez, Litterman, and Scheinkman (1994).

In this paper, we examine the magnitude of the biases (the "proxy problem") associated with the use of a 1- or 3-month discount bond yield in place of the unobservable short rate. When a proxy is used, the drift and diffusion functions estimated by an econometrician are those of the proxy process, not the short rate process. We first show how the drift and diffusion functions of the proxy process are related to the derivatives of the proxy with respect to the short rate and the (inverse) function from the proxy to the short rate. Using this framework, we analyze the class of single-factor affine bond pricing models, which includes the Vasicek (1977) and CIR (1985) models as special cases, and the two-factor affine model of Longstaff and Schwartz (1992) (hereafter LS). In addition, we develop two procedures for analyzing more general (i.e., nonlinear) interest rate models, and apply them to a simple nonlinear single-factor model based on the empirical work of Aït-Sahalia (1996b), CHLS (1997), and Stanton (1997).

Our first result is that the proxy problem does not appear to be economically significant in the case of single-factor affine bond price models, for parameter values that are consistent with the US data. Analytic results show that the errors in the estimation of the drift and diffusion functions are trivial, even when using a 3-month bill as a proxy. Simple method of moments parameter estimates based on short-maturity proxies in the Vasicek (1977) and CIR cases and the bond prices based on these parameter estimates are virtually indistinguishable from the true parameters and bond prices, for maturities from one to twenty years. In the case of the two-factor affine model of LS, the proxy problem (using a 3-month Treasury bill) is generally small when pricing discount bonds of maturities of five years or less, although there are some significant pricing differences when the short rate is high while the variance is simultaneously low.

In the case of the nonlinear model, analytic results are not possible, since there is no closed-form solution to the fundamental bond pricing partial differential equation (PDE). However, we describe two separate procedures for evaluating the magnitude of the proxy problem. The first procedure is based on a finite-difference approximation to the bond pricing PDE and a numerical approximation to the derivatives of the yield function. The second approximation approach is to use a truncated Taylor series expansion of the yield function and then define the derivatives in terms of the approximate yield function. Both approaches can be implemented with only knowledge of the drift and diffusion functions of the short rate and the market price of risk process. When applied to the nonlinear single-factor model, both approaches provide very similar results.

A comparison of the drift functions of the true short rate and the 1- and 3-month proxies in the nonlinear model shows that the proxies display approximately the same amount of mean reversion as the estimates based on the true short rate, which is consistent with the results in the affine case.

The diffusion functions of the proxies, however, exhibit some important differences from the true diffusion. In fact, the diffusion function for the 3-month yield appears to be a *concave* function of the interest rate for a substantial portion of the support of the short rate even though the true diffusion is a *convex* function. The shape of the diffusion function has been a focus of the empirical literature, and it remains an unresolved question. It is obvious that the shape of this function has implications for pricing interest rate options and other fixed-income instruments with "option-like" features. In fact, it turns out that even for discount bonds with maturities from one to ten years there are differences of roughly one percent per year between the true bond prices and the prices implied by the functions estimated from the 3-month proxy.

These results illustrate that empirical analyses and conclusions about the short rate process can depend on the choice of proxy. They also provide information about when it is reasonable to interpret the results of the empirical literature as saying something substantive about the nature of the short rate process and when such an interpretation might be misleading. The results suggest that the trade-off between maturity misspecification and potential microstructure problems must be evaluated on a model by model basis. While these results should be treated with some caution, since they only apply directly to a specific set of term structure models and for parameters consistent with US data, the two numerical procedures described in the paper provide general tools for identifying situations where the proxy problem is severe. In addition, although we do not emphasize it in this paper, the Taylor series approximation can be used to construct moment-based estimators of the true short rate (or multiple state variables) from the moments of observed yields of short term bonds.

The remainder of the paper is organized as follows. Section 2 introduces some basic definitions, including the fundamental bond pricing equation. It also shows how the drift and diffusion functions of the process for the proxy are related to the derivatives of the proxy with respect to the short rate and the function mapping the proxy to the short rate. Section 3 contains the results for the affine bond pricing models. A nonlinear model consistent with some of the recent empirical literature is examined in Section 4, and Section 5 contains the results on the two-factor LS model. The conclusions are in Section 6. Appendix A provides some evidence on the robustness and accuracy of the finite-difference approach, and Appendix B contains some calculations used in Section 5.

2 Bond Pricing and the Proxy Problem

The short rate is assumed to be the solution to a stochastic differential equation (SDE) of the form

$$dr_t = \mu(r_t)\,dt + \sigma(r_t)\,dW_t, \tag{1}$$

where μ and σ are the "drift" and "diffusion" functions, respectively, and W is a standard Brownian motion. A unique solution to (1), for any initial condition, is assumed to exist, and given that μ and σ do not depend on time, the solution is time-homogeneous.

The essence of a single-factor bond pricing model is the assumption that the prices of default-free bonds of all maturities are given by a function of a single state variable and the term to maturity. In particular, let the price at time t of a (default-free) bond making a fixed payment at date $s < T$ be a function of the short rate, and denote this price by $B(r,\tau)$, where $\tau \equiv s - t$. Assuming that the function $B(r,\tau)$ is differentiable as required, Itô's formula implies

$$dB(r_t,\tau) = \alpha(r_t,\tau)\,B(r_t,\tau)\,dt + \delta(r_t,\tau)\,B(r_t,\tau)\,dW_t, \tag{2}$$

where the expected rate of return $\alpha(r,\tau)$ and volatility $\delta(r,\tau)$ are

$$\alpha(r,\tau)\,B = \frac{1}{2}\sigma^2(r)\,B_{rr} + \mu(r)\,B_r - B_\tau \tag{3}$$

and

$$\delta(r,\tau)\,B = \sigma(r)\,B_r, \tag{4}$$

respectively. B_r denotes the partial derivative $\partial B(r,\tau)/\partial r$, and B_τ and B_{rr} are defined analogously.

Since bond prices are functions solely of the short rate, price changes on bonds of different maturities must be perfectly correlated. It follows that, in order to preclude arbitrage, the risk premiums on bonds of different maturities must be proportional to the standard deviations of their returns.[1] Therefore, the expected rate of return in (2) must satisfy

$$\alpha(r,\tau) = r + \lambda(r)\,\frac{B_r(r,\tau)}{B(r,\tau)}, \tag{5}$$

[1] See Vasicek (1977), pages 180-181, or Ingersoll (1987), pages 381-382 and page 394.

where λ, the risk premium process, is a function of the short rate alone.[2] Substituting (5) into (3) yields the PDE

$$\frac{1}{2}\sigma^2(r) B_{rr} + [\mu(r) - \lambda(r)] B_r - B_\tau - rB = 0 \tag{6}$$

with the associated boundary condition

$$B(r, 0) = 1, \tag{7}$$

where the bond's payment has been normalized to one unit of account.

The potential bias in using the yield on a finite-maturity bond as a proxy for the short rate (the "proxy problem") stems from the obvious fact that the yield-to-maturity on a τ-period bond, defined as

$$y(r, \tau) = -\frac{1}{\tau} \log(B(r, \tau)), \tag{8}$$

is, in general, not identical to the short rate. As a result, the stochastic process for the yield-to-maturity,

$$dy_t = \mu^y(y_t) dt + \sigma^y(y_t) dW_t, \tag{9}$$

is not identical to the short rate process (1). If y is used as a proxy for r, the functions being estimated are μ^y and σ^y of (9) and not the functions μ and σ that appear in (1). The "biases" $\mu^y - \mu$ and $\sigma^y - \sigma$ (note that they are functions) may be due to μ^y and σ^y having different functional forms than μ and σ. Even when μ^y and σ^y have the same functional forms as μ and σ, the biases may also be due to the values of the parameters of μ^y and σ^y differing from the values of the corresponding parameters of μ and σ.

Applying Itô's lemma, it is possible to identify the specific nature and magnitude of the bias. Holding the maturity τ constant, $y_t = y(r_t)$ according to (8), and

$$dy_t = \left[\frac{\partial y(r_t)}{\partial r}\mu(r_t) + \frac{1}{2}\frac{\partial^2 y(r_t)}{\partial r^2}\sigma^2(r_t)\right] dt + \frac{\partial y(r_t)}{\partial r}\sigma(r_t) dW_t,$$

or

[2] According to (5), $\lambda(r)$ determines (in part) the instantaneous excess holding period return of a bond of maturity τ over r. If $\lambda(r) = 0$, then the "local expectations hypothesis" holds, while $\lambda < 0$ implies that all maturities offer an expected return premium above the short rate.

$$
\begin{aligned}
dy_t = & \left[\frac{\partial y(r(y_t))}{\partial r}\mu(r(y_t)) + \frac{1}{2}\frac{\partial^2 y(r(y_t))}{\partial r^2}\sigma^2(r(y_t))\right]dt \\
& + \frac{\partial y(r(y_t))}{\partial r}\sigma(r(y_t))\,dW_t
\end{aligned} \tag{10}
$$

where $r_t = r(y_t)$ and $r : \mathbb{R} \to \mathbb{R}$ is the inverse of the yield function y defined by (8). Equation (10) reveals that μ^y and σ^y are given by

$$
\mu^y(y) = \frac{\partial y(r(y_t))}{\partial r}\mu(r(y_t)) + \frac{1}{2}\frac{\partial^2 y(r(y_t))}{\partial r^2}\sigma^2(r(y_t)) \tag{11}
$$

and

$$
\sigma^y(y) = \frac{\partial y(r(y_t))}{\partial r}\sigma(r(y_t)). \tag{12}
$$

According to (11) and (12), the magnitude of the biases are determined by the partial derivatives $\partial y/\partial r$ and $\partial^2 y/\partial r^2$, and the inverse function r.

The magnitude of the biases can be determined only for specific bond pricing models; i.e., for specific choices of the functions μ, σ, and λ. Given a specific model, the magnitude of the biases can be examined by evaluating the right-hand sides of (11) and (12) and comparing the results to the true drift and diffusion functions, μ and σ. While this can be done explicitly for only a limited class of models, it is straightforward to compute numerical approximations to the relevant derivatives and functions. Thus (11) and (12) provide a general procedure for evaluating the magnitude of the bias.

The following sections first consider the class of affine bond pricing models, which include the well-known Vasicek (1977) and CIR models as special cases, and then a nonlinear single-factor model inspired by recent empirical research. In both cases, we focus on the magnitude of the biases introduced by substituting an observable short-maturity bond for the true short rate, and whether these biases are likely to be economically significant.

3 Affine Single-Factor Economies

3.1 General Results

Bond prices are said to be *affine* if they can be written as

$$
B(r,\tau) = \exp\left[a(\tau) + b(\tau)\,r\right]. \tag{13}
$$

The following proposition provides the connection between the short rate process, the risk premium, and affine bond prices:

Proposition 1 *There exist functions $a(\tau)$ and $b(\tau)$ such that bond prices are given by (13) if and only if the diffusion coefficient in (1) is*

$$\sigma(r) = \sqrt{\beta_0 + \beta_1 r} \tag{14}$$

and the drift adjusted by the risk premium function is

$$\mu(r) - \lambda(r) = \rho_0 + \rho_1 r, \tag{15}$$

where ρ_0, ρ_1, β_0, and β_1 are constants.

Proof. See Duffie (1996), Section $7E$. ■

Since bond prices are given by (13), it follows that at time t the yield-to-maturity on a bond maturing at $t+\tau$ is

$$y(r_t, \tau) = -\frac{1}{\tau}\left[a(\tau) + b(\tau) r_t\right]. \tag{16}$$

The dynamics of $y(r_t, \tau)$ (for a constant value of τ) then follow from (16) and the dynamics of r. For example, assume that the actual (not risk-adjusted) drift of r is affine[3]

$$\mu(r) = \alpha_0 + \alpha_1 r.$$

Then, equations (11) and (12) imply that the drift and diffusion functions of y for a fixed τ are

$$\mu^y(y) = \alpha_0^y + \alpha_1^y y \tag{17}$$

$$\sigma^y(y) = \sqrt{\beta_0^y + \beta_1^y y} \tag{18}$$

where

$$\alpha_0^y = [-b(\tau)/\tau]\,\alpha_0 - [-a(\tau)/\tau]\,\alpha_1, \tag{19}$$

$$\alpha_1^y = \alpha_1, \tag{20}$$

$$\beta_0^y = [-b(\tau)/\tau]^2\,\beta_0 - [-a(\tau)/\tau]\,[-b(\tau)/\tau]\,\beta_1, \tag{21}$$

and

$$\beta_1^y = [-b(\tau)/\tau]\,\beta_1. \tag{22}$$

If y is used as a proxy for r, then α_0^y will be estimated instead of α_0, β_0^y instead of β_0, and β_1^y instead of β_1. Equation (20) implies that α_1, can be estimated from the finite-maturity proxy. The following proposition examines the magnitude of the bias.

[3] If $\mu(r)$ is affine, then (15) implies that $\lambda(r)$ is also affine, i.e., $\lambda(r) = \lambda_0 + \lambda_1 r$. Moreover, since $\lambda(r)$ should be zero whenever $\sigma(r)$ is zero, (14) implies $\lambda_0 = \phi\beta_0$ and $\lambda_1 = \phi\beta_1$ for some ϕ. Thus the risk adjusted drift parameters in (15) are $\rho_0 = \alpha_0 - \phi\beta_0$ and $\rho_1 = \alpha_1 - \phi\beta_1$.

Proposition 2 *Let bond prices be defined by (13). The functions $a(\tau)$ and $b(\tau)$ are the solutions to*

$$b'(\tau) = \rho_1 b(\tau) + \frac{1}{2}\beta_1 b^2(\tau) - 1, \tag{23}$$

$$a'(\tau) = \rho_0 b(\tau) + \frac{1}{2}\beta_0 b^2(\tau), \tag{24}$$

with the initial conditions $b(0) = 0$ and $a(0) = 0$, where β_0 and β_1 are defined in (14), and ρ_0 and ρ_1 are defined in (15). The second-order Taylor series expansions of $-b(\tau)/\tau$ and $-a(\tau)/\tau$ around $\tau = 0$ are

$$-\frac{b(\tau)}{\tau} \cong 1 + \frac{1}{2}\rho_1\tau + \frac{1}{6}\left(\rho_1^2 - \beta_1\right)\tau^2, \tag{25}$$

$$-\frac{a(\tau)}{\tau} \cong \frac{1}{2}\rho_0\tau + \frac{1}{6}\left(\rho_0\rho_1 - \beta_0\right)\tau^2. \tag{26}$$

Proof. For (23) and (24), see Duffie (1996), Section 7*E*. The Taylor series expansions follow by direct calculation. ■

Together with (19) through (22), Proposition 2 suggests that the misspecification induced by substituting an observed bond yield for the true short rate is small for short maturities and for cases when ρ_1 and β_1 are close to zero. Proposition 1 and (16) imply that the unconditional mean of the yield on a τ-period bond is

$$E\left[y(r,\tau)\right] = -\frac{1}{\tau}\left[a(\tau) + b(\tau)E(r)\right], \tag{27}$$

and the unconditional variance is

$$\operatorname{Var}\left[y(r,\tau)\right] = \frac{1}{\tau^2}b^2(\tau)\operatorname{Var}(r). \tag{28}$$

Given a sequence of observations of the yield process at an arbitrary discrete interval Δ, the unconditional k-th order autocorrelation of the yields is

$$\operatorname{Corr}\left[y_{t+k\Delta}(r,\tau), y_t(r,\tau)\right] = \operatorname{Corr}\left[r_{t+k\Delta}, r_t\right]. \tag{29}$$

If the observations on a yield with τ-periods until maturity were treated as observations of r itself, (27) through (29) indicate the magnitude of the misspecification error that would result from a simple method-of-moments estimator. In particular, (29) demonstrates that the autocorrelation function can be inferred directly without a problem, while (27) and (28) show that estimates of the long-run mean and the variance of the short rate will have biases related to the levels of $-a(\tau)/\tau$ and $-b(\tau)/\tau$. The magnitude of these effects will depend on the exact forms of the short rate and risk premium functions and on the relevant parameter values.[4]

3.2 Vasicek (1977)

Consider a short rate process defined by the SDE

$$dr_t = \kappa(\theta - r_t)\,dt + \sigma dW_t, \tag{30}$$

where $\kappa > 0$. This is a Gaussian process. If $\lambda(r) = \lambda_0$ (< 0), then $\alpha_0 = \kappa\theta$, $\alpha_1 = -\kappa$, $\beta_0 = \sigma^2$, $\beta_1 = 0$, and r and $\lambda(r)$ satisfy the conditions for an affine bond price model.[5] Vasicek (1977) shows that

$$a(\tau) = \left(\theta - \frac{\lambda_0}{\kappa} - \frac{1}{2}\frac{\sigma^2}{\kappa^2}\right)\left[\frac{1}{\kappa}(1 - \exp(-\kappa\tau)) - \tau\right] - \frac{\sigma^2}{4\kappa^3}\left[1 - \exp(-\kappa\tau)\right]^2, \tag{31}$$

$$b(\tau) = -\frac{1}{\kappa}\left[1 - \exp(-\kappa\tau)\right]. \tag{32}$$

Equations (19) through (22) then give the dynamics of y for fixed τ as

$$dy_t = \left[-\frac{b(\tau)}{\tau}\kappa\theta - \frac{a(\tau)}{\tau}\kappa - \kappa y_t\right]dt - \frac{b(\tau)}{\tau}\sigma dW_t. \tag{33}$$

Figures 1 and 2 contain plots of the true drift and diffusion functions and the functions implied by 1- and 3-month proxies for three sets of parameter choices that embody different levels of persistence. The first two sets of parameter choices are generally consistent with the short-term interest rate data used in earlier research. The parameter θ is set close to the value of the long-run mean of the Eurodollar rate used in Aït-Sahalia (1996a,b), and the κ and σ combinations are chosen to imply the same steady-state variance of the short rate process. κ is allowed to assume values that imply a first-order autocorrelation of the discrete observations as high as 0.98 ($\kappa = 0.22$) or as

[4] Equations (27) through (29) can be used, in conjunction with the sample moments of (at least) two yields and an explicit specification for r and λ (and, therefore, $a(\tau)$ and $b(\tau)$), in a standard generalized method of moments estimation of the short rate parameters. This estimation is not the focus of the questions that we examine in this paper. However, if the proxy problem is quantitatively small, then the corrections to a moment-based estimation strategy implied by (27) through (29) will also be small.

[5] See Goldstein and Zapatero (1996) for an example of an exchange-economy equilibrium model in which this bond pricing model holds.

low as 0.87 ($\kappa = 1.72$). The small value for the risk premium parameter, $\lambda_0 = -0.02$, is consistent with the estimate in Stanton (1997). The drift and diffusion functions plotted in Figures 1 and 2 are quite close (essentially *identical* for the drifts) to the true functions, which suggests that the proxy problem is small.

Equation (30) is an Ornstein-Uhlenbeck process, which is a first-order autoregressive process when observed at discrete intervals. The parameter κ reflects the speed of adjustment of r back towards its unconditional (or long-run) mean of θ. The unconditional moments of the short rate process in this case are well known. In particular:

$$E\left[r_t\right] = \theta, \tag{34}$$

$$\mathrm{Var}\left[r_t\right] = \frac{\sigma^2}{2\kappa}, \tag{35}$$

and

$$\mathrm{Corr}\left[r_{t+\Delta}, r_t\right] = \exp\left(-\kappa\Delta\right). \tag{36}$$

In this case (27) through (28) become:

$$E\left[y\left(r,\tau\right)\right] = -\frac{1}{\tau}\left[a\left(\tau\right) + b\left(\tau\right)\theta\right], \tag{37}$$

$$\mathrm{Var}\left[y\left(r,\tau\right)\right] = \frac{1}{\tau^2} b^2\left(\tau\right)\frac{\sigma^2}{2\kappa}, \tag{38}$$

and

$$\mathrm{Corr}\left[y_{t+\Delta}\left(r,\tau\right), y_t\left(r,\tau\right)\right] = \exp\left(-\kappa\Delta\right), \tag{39}$$

where $a\left(\tau\right)$ and $b\left(\tau\right)$ are defined in (31) and (32), respectively.[6]

The economic significance of the proxy problem is evaluated by assuming that an econometrician has constructed the sample analogs of (37) through (39), but proceeds under the assumption that they are the sample analogs of (34) through (36). This results in (potentially) biased estimates of the parameters. We then compare the yields on discount bonds computed using the true and biased parameters. Comparing (36) and (39), it is clear that the time series of observations on the yield of any bill allows κ to be estimated, and that the use of a proxy does not result in any bias for this parameter.

[6] Using the yield moments to recover estimates of the short rate parameters is particularly simple in the Vasicek case. Equation (39) and the sample autocorrelation coefficient on any yield can be used to estimate κ. Given an estimate of κ, (38) and the sample variance on any yield can be used to construct an estimate of σ, and (37) and the sample means of any two yields can be used to construct estimates of θ and λ_0.

The estimates of parameters θ and σ are obtained by equating (34) and (35) to the sample analogs of (37) and (38), and will be biased. The risk premium λ_0 affects the results of this exercise through the form of (37) and (38), and we assume that λ_0 is known.

Table 1 shows the results of this exercise in the Vasicek case. For each of the three parameterizations used in Figures 1 and 2, the true yields on discount bonds of maturities of up to twenty years are compared with the yields that would be calculated under the following assumptions: (1) A 3-month bill is used in place of the true r; (2) A long enough sample of observations of the proxy is used so that the sample moments are equal to the population moments (37) through (39); and (3) The market price of risk parameter is known. Consistent with the results in Figures 1 and 2, the pricing differences implied by using a proxy in place of the true short rate are quite small, with the largest pricing error being 25 basis points on a 20-year discount bond.[7]

3.3 Cox, Ingersoll, and Ross (1985)

In this case, the short rate follows a square-root diffusion

$$dr_t = \kappa(\theta - r_t)\,dt + \sigma\sqrt{r_t}dW_t. \tag{40}$$

Let $\lambda(r) = \lambda_1 r$, where $\lambda_1 < 0$. This structure fits into the general affine class with $\alpha_0 = \kappa\theta$, $\alpha_1 = -\kappa$, $\beta_0 = 0$, and $\beta_1 = \sigma^2$. It is further assumed that κ, $\theta > 0$ and $2\kappa\theta \geq \sigma^2$. Bond prices in this model satisfy (13) with

$$a(\tau) = \frac{2\kappa\theta}{\sigma^2}\log\left[\frac{2\gamma\exp\left(\frac{1}{2}\tau(\kappa+\lambda_1+\gamma)\right)}{(\kappa+\lambda_1+\gamma)\left[\exp(\gamma\tau)-1\right]+2\gamma}\right] \tag{41}$$

and

$$b(\tau) = \frac{-2\left[\exp(\gamma\tau)-1\right]}{(\kappa+\lambda_1+\gamma)\left[\exp(\gamma\tau)-1\right]+2\gamma}, \tag{42}$$

where $\gamma \equiv \sqrt{(\kappa+\lambda_1)^2 + 2\sigma^2}$. Equations (19) through (22) then give the dynamics of y for fixed τ as

$$dy_t = \left[-\frac{b(\tau)}{\tau}\kappa\theta - \frac{a(\tau)}{\tau}\kappa - \kappa y_t\right]dt - \sigma\sqrt{-\frac{a(\tau)\,b(\tau)}{\tau^2} - \frac{b(\tau)}{\tau}y_t}dW_t. \tag{43}$$

[7]Note that if the market price of risk had also been estimated, the proxy-generated prices would have been even more accurate, since the estimated market prices of risk reported in the previous footnote are all smaller in absolute value than the true λ_0.

Figures 3 and 4 contain plots of the drift and diffusion functions of the short rate in the CIR case for three parameterizations that imply the same unconditional mean, variance, and first-order autocorrelation coefficients as the parameterizations used in the Vasicek (1977) case. Again, as in the Vasicek case, these plots suggest that the magnitude of any misspecification due to substituting a finite maturity bond in place of the true short rate is quite small.

In the case of the CIR short rate model, the unconditional moments of the true short rate process are:

$$E\left[r_t\right] = \theta, \tag{44}$$

$$\text{Var}\left[r_t\right] = \frac{\theta\sigma^2}{2\kappa}, \tag{45}$$

and

$$\text{Corr}\left[r_{t+\Delta}, r_t\right] = \exp\left(-\kappa\Delta\right) \tag{46}$$

in this case. The unconditional moments of a proxy for the short rate with τ periods until maturity are

$$E\left(y_t\left(r,\tau\right)\right) = -\frac{1}{\tau}\left[a\left(\tau\right) + b\left(\tau\right)\theta\right], \tag{47}$$

$$\text{Var}\left(y_t\left(r,\tau\right)\right) = \frac{1}{\tau^2}b^2\left(\tau\right)\frac{\theta\sigma^2}{2\kappa}, \tag{48}$$

and

$$\text{Corr}\left[y_{t+\Delta}\left(r,\tau\right), y_t\left(r,\tau\right)\right] = \exp\left(-\kappa\Delta\right), \tag{49}$$

where $a\left(\tau\right)$ and $b\left(\tau\right)$ are defined by (41) and (42), respectively.

Table 2 shows the economic significance of differences in the estimates of the parameters of the short rate associated with the use of a 3-month bill as a proxy, again by examining the yields-to-maturity on zero-coupon bonds and using the same procedure used in Table 1. As in the Vasicek case, the implied pricing differences are small across all three parameterizations and for all maturities from three months to twenty years.[8]

[8] The calculations in Table 3 are based on the same assumptions used to generate the results in Table 2.

3.4 Summary of the Affine Examples

These two examples demonstrate, *in an affine bond price model*, that accurate estimates of the drift and diffusion of the short rate process can be obtained, even using proxies with maturities as long as three months. However, CHLS, Aït-Sahalia (1996b) and Stanton (1997) all find evidence that is inconsistent with the assumption of a linear drift in the short rate process, and Stanton (1997) estimates a nonlinear risk premium function.[9] The following section examines a simple example that is consistent with these findings.

4 A Nonlinear Single-Factor Model

4.1 The Structure of the Model

The short rate process is defined by the SDE

$$dr_t = \mu(r_t)\,dt + \sigma(r_t)\,dW_t, \tag{50}$$

where

$$\mu(r) = \psi_0 r^{-1} + \psi_1 + \psi_2 r + \psi_3 r^2 \tag{51}$$

and

$$\sigma(r) = \sigma r^{3/2}. \tag{52}$$

The drift specification in (51) is the flexible functional form used in Aït-Sahalia (1996b) and CHLS, and (52) is the diffusion specification suggested by Chan et al (1992) and further examined in CHLS. The diffusion function (52) is also a special case of the flexible diffusion specification used by Aït-Sahalia (1996b). The risk premium function is

$$\lambda(r) = \sigma(r)\left[\lambda_0 + \lambda_1 r + \lambda_2 r^2\right], \tag{53}$$

where $\sigma(r)$ is the diffusion function in (52). The arbitrage-free bond prices associated with (50) and (53) follow from the PDE (6) and the boundary condition (7). As Stanton (1997) notes, this specification of $\lambda(r)$ is consistent with the absence of arbitrage because it satisfies the condition that

$$\lambda(r) = 0 \qquad \text{if} \qquad \sigma(r) = 0. \tag{54}$$

[9] Chapman and Pearson (1998) examine the finite-sample properties of the estimators used in Stanton (1997) and Aït-Sahalia (1996b), and they conclude that these papers do not provide compelling evidence that the drift of the short rate is actually nonlinear. Chapman and Pearson (1998), however, do not examine the estimator in CHLS. Nor do they conclude that the drift function in the actual data is linear.

Violation of (54) results in an arbitrage opportunity in an "arbitrage-free" term structure model (see CIR page 398 or Ingersoll (1987), pages 400-401).

The parameter values for the drift and diffusion functions are taken from CHLS. In particular, $\psi_0 = 0.0073$, $\psi_1 = -0.4446$, $\psi_2 = 9.5178$, $\psi_3 = -56.9038$, and $\sigma = 1.00$.[10] The risk premium function (53) can approximate the nonparametric risk premium function estimated in Stanton (1997) for the parameter choices $\lambda_0 = -1.0$, $\lambda_1 = 15$, and $\lambda_2 = -115$.

4.2 Approximating the Proxy Drift and Diffusion Functions

As in the affine case, the proxy problem can be analyzed using equations (11) and (12). In this nonlinear case, however, there are no closed-form expressions for bond prices and yields, which makes it impossible to compute explicitly the partial derivatives $\partial y/\partial r$ and $\partial^2 y/\partial r^2$. Nonetheless, there are approximations to the price and yield functions that can be used to construct approximations to μ^y and σ^y. In this sub-section we describe two different approaches that can be used to extend an analysis of the proxy problem to any single-factor term structure model in which the drift and diffusion functions of the short rate and the market price of interest rate risk function can be specified.[11]

The first approach is based on a finite-difference approximation to the solution of the bond pricing PDE, equation (6), subject to the boundary condition (7). It consists of the following steps. First, use a finite-difference algorithm to find an approximate solution to (6).[12] Given the approximation to the bond price function, compute the yield function at the grid points of the bond price function. Interpolate between the grid points using cubic splines to construct an approximation to the continuous yield function. Given the yield function, at a maturity τ, construct a grid of size N on a bounded interval of the short rate $[\underline{r}, \overline{r}]$ and approximate the required derivatives as

$$\frac{\partial \widehat{y}\,(r_i, \tau)}{\partial r} \approx \frac{\widehat{y}\,(r_{i+1}, \tau) - \widehat{y}\,(r_{i-1}, \tau)}{2\delta} \tag{55}$$

[10] See CHLS, Appendix G, for the drift parameter estimates.

[11] The following section discusses how to extend these approaches to a specific multi-factor term structure example.

[12] In the following sub-section (and in the appendix), the numerical approximation is conducted using a Crank-Nicholson finite-difference algorithm, as described in Sections H and I of Chapter 11 of Duffie (1996). The backward iterations through the time steps of the problem are computed by solving a tridiagonal system of linear equations. This was accomplished using the **tridag(•)** function defined in Section 2.4 (page 51) of Press et al. (1992) (as implemented in MathCad's *Numerical Recipes Function Pack*).

and

$$\frac{\partial \widehat{y}\left(r_i,\tau\right)}{\partial r} \approx \frac{\widehat{y}\left(r_{i+1},\tau\right)-2\widehat{y}\left(r_i,\tau\right)+\widehat{y}\left(r_{i-1},\tau\right)}{\delta^2}, \tag{56}$$

where $\widehat{y}$ is the approximation to the yield function, $i-1$, i, and $i+1$ refer to points on the partition of $[\underline{r},\overline{r}]$, and δ is the (constant) width of the partition intervals. Finally, the partial derivatives (55) and (56) and the approximate yield function (and its inverse) from the finite-difference solution, it is possible to construct approximations to (11) and (12).

There are three sources of approximation error in this "finite-difference approach." The first is introduced by the computation of the finite-difference approximation. The second comes from the interpolation of the finite-difference solution, and the final source of error is due to the approximations to the true derivatives given in (55) and (56). The error order of magnitude of the finite-difference approximation depends on the particular algorithm used to solve (6).[13] The derivative approximations, (55) and (56), both have error order of magnitudes of δ^2, where $\delta \equiv \left(\overline{r}-\underline{r}\right)/N$. The size of these errors, and the associated quality of the drift and diffusion approximations from the finite-difference approach is examined in detail in Appendix A, where it is shown (for the affine case) that the errors are small and economically insignificant. In all of the results reported in the next sub-section, the finite-difference grid uses 720 time points and 240 space points, and the numerical derivatives, (55) and (56), are constructed using a grid of 76 points.

A second approximation, the "Taylor series approach," can be implemented using only knowledge of the drift and diffusion functions of the short rate and the market price of risk function. It is defined in the following proposition and corollary.

Proposition 3 *Let $\mu(r)$, $\lambda(r)$, and $\sigma^2(r)$ be continuous and have $2(N-2)$ continuous derivatives, for $N \geq 2$ and $r \in D$, where D is the support of the stochastic process r. The N^{th}-order Taylor series approximation of $B(r,\tau)$ is*

$$B(r,\tau,N)=\sum_{n=0}^{N}\frac{1}{n!}f_n(r)\,\tau^n,$$

[13] The Crank-Nicholson scheme used below has an error that is $O\left((\Delta\tau)^2\right)$, where $\Delta\tau$ is the size of the grid in the time dimension. $O(h)$ is the asymptotic order symbol, meaning that if f is a function of Δ and f is $O(\Delta)$ then

$$\lim_{\Delta\downarrow 0} f(\Delta)/\Delta = K,$$

where K is a constant.

where

$$f_{n+1}(r) = [\mu(r) - \lambda(r)] \frac{df_n(r)}{dr} + \frac{1}{2}\sigma^2(r) \frac{d^2 f_n(r)}{dr^2} - r f_n(r),$$

$n \geq 0$, *and*

$$f_0(r) = 1.$$

Proof. Under the assumptions of the proposition, $f_N(r)$ is continuous on D. For $0 \leq n \leq N$, $f_n(r)$ is continuous and has two continuous derivatives on D. Itô's formula then implies

$$E_0^Q \left[f_n(r_\tau) \exp\left(- \int_0^\tau r_s ds \right) \right] =$$

$$f_n(r_0) + E_0^Q \int_0^\tau f_{n+1}(r_s) \exp\left(- \int_0^s r_u du \right) ds$$

for $0 \leq n \leq N$, where $E_0^Q[\bullet]$ denotes expectation computed under the risk-neutral measure. Thus

$$\frac{\partial}{\partial \tau} E_0^Q \left[f_n(r_\tau) \exp\left(- \int_0^\tau r_s ds \right) \right] = E_0^Q \left[f_{n+1}(r_\tau) \exp\left(- \int_0^\tau r_s ds \right) \right]$$

for $0 \leq n \leq N$. Bond prices are given by

$$B(r_0, \tau) = E_0^Q \left[f_0(r_\tau) \exp\left(- \int_0^\tau r_s ds \right) \right].$$

Thus

$$\frac{\partial^n}{\partial \tau^n} B(r_0, \tau) = E_0^Q \left[f_n(r_\tau) \exp\left(- \int_0^\tau r_s ds \right) \right] \tag{57}$$

for $0 \leq n \leq N$. $B(r_0, \tau)$ and its first N derivatives with respect to τ are continuous with respect to τ for $0 \leq \tau < \infty$. The approximation in the proposition is the truncated Taylor series expansion of $B(r_0, \tau)$ around $\tau = 0$. It is based on

$$\left. \frac{\partial^n}{\partial \tau^n} B(r_0, \tau) \right|_{\tau = 0} = f_n(r_0)$$

for $1 \leq n \leq N$, which is an implication of (57). ■

Corollary 1 *Let* $y(r,\tau)$, $\mu^y(r,\tau)$, *and* $\sigma^y(r,\tau)$ *be the bond yield, the drift of the yield (for a constant* τ*), and the volatility of the yield (for a constant* τ*), as functions of* r *and* τ*. Approximations of these functions based on* $B(r,\tau,N)$ *are*

$$y(r,\tau,N) = -\frac{1}{\tau}\log\left(B(r,\tau,N)\right),$$

$$\mu^y(r,\tau,N) = \mu(r)\frac{\partial y(r,\tau,N)}{\partial r} + \frac{1}{2}\sigma^2(r)\frac{\partial^2 y(r,\tau,N)}{\partial r^2},$$

and

$$\sigma^y(r,\tau,N) = \sigma(r)\frac{\partial y(r,\tau,N)}{\partial r}.$$

Proof. This follows immediately from Itô's lemma, and the derivatives can be calculated from the results in Proposition 3. ■

In the results reported in the next sub-section, the Taylor series approach is implemented using a third order ($N = 3$) approximation.[14]

4.3 The Results

As a first step in examining the nonlinear model, we construct estimates of the stationary density of the true short rate and the 1–month and 3-month yields. Following Karlin and Taylor (1981), the stationary density, denoted $\pi(r)$, is the solution to the stationary version of the Kolmogorov forward equation:

$$0 = \frac{1}{2}\frac{\partial^2}{\partial r^2}\left[\sigma^2(r)\pi(r)\right] - \frac{\partial}{\partial r}\left[\mu(r)\pi(r)\right]. \tag{58}$$

The density $\pi(r)$ must, of course, also satisfy the conditions: $\int_{\underline{b}}^{\overline{b}}\pi(r) = 1$ and $\pi(r) \geq 0$ for $r \in \left[\underline{b},\overline{b}\right]$, where $\left[\underline{b},\overline{b}\right]$ is the support of the stationary density. Karlin and Taylor (1981) show that the stationary density is of the form

$$\pi(r) = C_1 m(r) S(r) + C_2 m(r), \tag{59}$$

where

$$S(x) = \int^x s(r)\,dr$$

[14] The process of "inverting" the moments of finite-maturity yields to recover the short rate (and market price of risk) parameters is more complicated in this nonlinear problem because there is no closed-form solution for yields. However, the yield moments based on the Taylor series approximation in Proposition 3 are nonlinear functions that relate the moments of observed yields to the parameters of short rate process, and they could be used in a generalized method of moments estimation of the short rate parameters.

and

$$m(x) = \frac{1}{s(x)\,\sigma^2(x)},$$

with

$$s(x) = \exp\left\{-\int^x \frac{2\mu(r)}{\sigma^2(r)}dr\right\},$$

and C_1 and C_2 are constants of integration chosen to ensure that $\pi(r)$ is a density. In this case, the conditions above imply that $C_1 = 0$ and

$$C_2 = \int_{\underline{b}}^{\bar{b}} m(r)\,dr.$$

In computing this constant, we use the interval $[0.03, 0.18]$ as an approximation to the support of the stationary density.[15]

The stationary density of the true short rate is computed using the true drift and diffusion functions, and the stationary densities of the 1- and 3-month yield series are computed using the numerical estimates of the drift and diffusion functions. Figure 5 shows the resulting densities based on the finite-difference approximation approach, and Figure 6 shows the comparable results based on the Taylor series approximation. In both pictures, the 3-month yield has the highest peak and the smallest tails. The means of the 1-month yield and the 3-month yield are both higher than the mean of the true short rate. All of the densities in both figures suggest that both the true short rate and proxies are generally between 0.05 and 0.16.

Figures 7 plots the drift function of the true short rate, (51), along with the numerical approximations to $\mu^y(y)$ based on both the 1- and 3-month yields using the finite-difference approximation. Figure 8 shows the same functions based on the Taylor series approximation. A researcher using a 1-month proxy would observe the second panel in each figure instead of the first panel, and a researcher using a 3-month proxy would observe the third panel. All three functions in Figure 7 are essentially identical, which suggests that the proxy problem is not a substantial issue in the estimation of the drift function. However, using the Taylor series approximation, Figure 8 suggests that the drift of the 1-month rate exhibits slightly more curvature (i.e., mean reversion at higher rate levels) than the true short rate drift, and the 3-month rate exhibits even more curvature than the 1-month rate. The economic significance of these apparently minor differences depends on their

[15] The numerical calculations of the integrals involved in solving for $\pi(r)$ is done using MathCad7, Professional. MathCad uses the Romberg method to compute the integrals. See Press, et al. (1992), Section 4.3.

implications (along with the effect of the proxy on the diffusion function) for bond prices and interest rate derivative prices.

The diffusion functions for the true short rate and the 1- and 3-month proxies, using the two different approximation approaches, are shown in Figures 9 and 10. These figures have the same general structure as Figures 7 and 8, with the true function in the top panel and what a researcher using a 1- or 3-month proxy would observe in the middle and bottom panels, respectively. One important issue addressed in Chan et al (1992), Aït-Sahalia (1996a,b), and CHLS is the shape of the diffusion function. It is related to potentially delicate issues of stationarity as well as consistency with existing single-factor models. Figures 9 and 10 show that the shape of the diffusion can depend on the choice of proxy. The true diffusion is a convex function of the short rate, whereas the 1-month diffusion is essentially linear. Strikingly, in the bottom panels of Figures 9 and 10, the diffusion functions based on the 3-month proxy are concave for a large part of the range of interest rates, and it appears that the diffusion functions have inflection points at a yield level of approximately 11 percent.[16] Given the importance of the diffusion function in pricing options and other interest rate derivatives, the result that the shape of the diffusion function can be sensitive to the choice of proxy suggests that proxy problem can be significant.

The partial differential equation (6) demonstrates that it is the diffusion and the drift under the equivalent martingale measure, $\mu(r) - \lambda(r)$, that are important in pricing discount bonds. Figures 11 and 12 show plots of the drift under the equivalent martingale measure constructed from the actual drifts of the true rate and the 1- and 3-month proxies and the true risk premium function. Again, these figures have the same general structures as Figures 7 and 8 and Figures 9 and 10. In both Figures 11 and 12, the adjusted drifts of the 1-month proxy (plotted against the 1-month yield) looks very similar to the shape of the true risk adjusted drift. However, the drifts of the 3-month proxy differ for large yields.[17]

[16] The bottom panels of Figures 9 and 10 indicate that the two different approximation schemes produce different estimates of the diffusion function for large values of the 3-month yield. Figures 5 and 6 indicate that such large yield values are realized infrequently, suggesting that these differences due to the different approximation methods may not be important for pricing bonds. Tables 3 and 4 show that this is in fact the case. The estimates of the drift and diffusion functions from the different approximation schemes result in bond prices and yields that are very similar.

[17] The shape of the risk neutral drift for the 3-month yield based on the finite-difference approximation is different from the shape of the same function estimated using the Taylor series approximation to the yield function, for large interest rates. In particular, the finite-difference based approximation in Figure 11 implies yield processes with substantially less

Tables 3 and 4 examine the economic significance of Figures 7 through 12. They show the yields on discount bonds of maturities from one to ten years using the true risk premium function combined with the true drift and diffusion functions and with the drift and diffusion functions obtained using the 1- and 3-month yields as proxies. The first thing to note about these tables is that there is virtually no difference between the prices based on the functions from the finite-difference approximation when compared to the prices from the Taylor series approximation. Accordingly, the analysis will focus on the finite-difference results in Table 3.

Panel A of Table 3 presents the results using the true functions. The model implies upward sloping term structures for low levels of r and inverted term structures for high levels of r. Furthermore, comparing down the columns of Panel A shows that the yields on shorter maturity bonds are more volatile than the yields on longer maturity bonds. The results in Panels B and C are quite similar to each other. The entries for corresponding maturities and levels of the relevant proxy are, typically within three basis points of each other. Panels B and C demonstrate that the proxies produce discount yields that are both higher and more variable than the true bond prices. The increased variability is particularly true of maturities at the short-end of the yield curve.

These results indicate that the proxy problem can be important for non-linear interest rate models. moreover, there are two reasons to suspect that Tables 3 and 4 may understate the extent of the proxy problem. First, by using the true functions for the drifts and diffusions of the true short rate and the short rate proxies and using the true market price of risk function, these tables abstract from any estimation error. Second, the differences in the shape of the diffusion functions documented in Figures 9 and 10 will have a greater impact on the pricing of interest rate and bond options than on the pricing of discount bonds.

5 The Longstaff and Schwartz (1992) Model

The preceding analysis of the proxy problem can be extended to multi-factor term structure models. While there is no unifying framework analogous to the affine structure of (13), (14), and (15) in a multi-factor setting[18], the

mean reversion than the comparable estimate based on the Taylor series approach, shown in Figure 12. However, Tables 3 and 4 show that the differences for large yields due to the different approximation schemes have little impact on bond prices.

[18] A few examples of alternate formulations of multi-factor term structure models are Brennan and Schwartz (1979), Brown and Schaefer (1993), Pearson and Sun (1994), and

general procedures outlined in the previous sections can still be applied to specific multi-factor models. As an example, we examine the two-factor model described in Longstaff and Schwartz (1992). This model begins by specifying the realized returns on physical capital:

$$\frac{dQ}{Q} = (\mu X + \theta Y)\, dt + \sigma \sqrt{Y} dW_1, \tag{60}$$

where μ, θ, and σ are positive constants and W_1 is a Brownian motion. The two state variables in this model are

$$dX = (a - bX)\, dt + c\sqrt{X} dW_2 \tag{61}$$

and

$$dY = (d - eY)\, dt + f\sqrt{Y} dW_3, \tag{62}$$

where $a, b,\ c,\ d,\ e,\ f\ >\ 0$, W_2 and W_3 are Brownian motions, and W_2 is uncorrelated with W_1 and W_3.[19] Both (61) and (62) are square-root diffusions, which implies that their stationary and transition densities are known, as are their conditional and unconditional moments. X is interpreted as an exogenous shock to the expected return on a constant returns to scale production technology, and Y is a shock to both the expected return and volatility of production.

Under the additional assumption that the economy consists of a large number of identical individuals with constant rate of time preference and logarithmic preferences over real consumption, LS demonstrate that the equilibrium term structure can be written as a function of two state variables: the instantaneous interest rate, r, and the variance of changes in the instantaneous rate, V. The dynamics of these state variables are

$$\begin{aligned} dr = &\left(\alpha\gamma + \beta\eta - \frac{\beta\delta - \alpha\xi}{\beta - \alpha} r - \frac{\xi - \delta}{\beta - \alpha} V\right) dt \\ &+ \alpha\sqrt{\frac{\beta r - V}{\alpha\,(\beta - \alpha)}} dW_2 + \beta\sqrt{\frac{V - \alpha r}{\beta\,(\beta - \alpha)}} dW_3 \end{aligned} \tag{63}$$

and

Duffie and Kan (1996).

[19] For this section of the paper, we have (generally) adopted the notation used in LS.

$$dV = \left(\alpha^2\gamma + \beta^2\eta - \frac{\alpha\beta(\delta-\xi)}{\beta-\alpha} r - \frac{\beta\xi - \alpha\delta}{\beta-\alpha} V\right) dt$$

$$+ \alpha^2 \sqrt{\frac{\beta r - V}{\alpha(\beta-\alpha)}} dW_2 + \beta^2 \sqrt{\frac{V - \alpha r}{\beta(\beta-\alpha)}} dW_3, \quad (64)$$

where $\alpha = \mu c^2$, $\gamma = a/c^2$, $\beta = \left(\theta - \sigma^2\right) f^2$, $\eta = d/f^2$, $\delta = b$, and $\xi = e$. The parameters α, γ, β, η, δ, ξ are also all positive. Furthermore, (63) and (64) show that these six parameters determine the joint distribution of the state variables r and V.

In particular – analogous to (34) and (35) or (44) and (45) in the Vasicek and CIR examples – the unconditional means and variances of the state variables (from LS) are

$$E[r] = \frac{\alpha\gamma}{\delta} + \frac{\beta\eta}{\xi}, \quad (65)$$

$$\mathrm{Var}[r] = \frac{\alpha^2\gamma}{2\delta^2} + \frac{\beta^2\eta}{2\xi^2}, \quad (66)$$

$$E[V] = \frac{\alpha^2\gamma}{\delta} + \frac{\beta^2\eta}{\xi}, \quad (67)$$

and

$$\mathrm{Var}[V] = \frac{\alpha^4\gamma}{2\delta^2} + \frac{\beta^4\eta}{2\xi^2}. \quad (68)$$

The unconditional autocovariances of r and V, for an arbitrary time interval Δ, are

$$\mathrm{Cov}[r_{t+\Delta}, r_t] = \frac{\alpha^2\gamma}{2\delta^2} \exp(-\delta\Delta) + \frac{\beta^2\eta}{2\xi^2} \exp(-\xi\Delta), \quad (69)$$

and

$$\mathrm{Cov}[V_{t+\Delta}, V_t] = \frac{\alpha^4\gamma}{2\delta^2} \exp(-\delta\Delta) + \frac{\beta^4\eta}{2\xi^2} \exp(-\xi\Delta). \quad (70)$$

We use the means, variances, and covariances in assessing the importance of the proxy problem. Clearly the choice of these moments is somewhat arbitrary. For example, an alternative estimation strategy is to use both short and long-term yields, as Boudoukh, Richardson, Stanton, and Whitelaw (1998) do in their estimation of a nonlinear two-factor model. We use these six moments, and construct all of them using a proxy for the short rate, because it seems likely that this choice of moments is the case when the proxy problem will be most severe.

In the single-factor Vasicek and CIR cases, it is straightforward to invert the mean, variance, and autocorrelation definitions to recover the three model parameters. In the LS model, however, the six moment conditions (65) through (70) each involve all six parameters. Given estimates of the population moments of r and V (assuming they are observable), we can recover the six parameters of the model by solving for α and β numerically and using these values to solve for δ, ξ, γ and η. Appendix B provides the calculations. Of course, we are interested in assessing the magnitude of the proxy problem in this two-factor model. This first requires computing the unconditional moments of the short-maturity proxies.

LS prove that the yield-to-maturity on a τ-period discount bond can be written as

$$y(r,V,\tau) = -\frac{\kappa\tau + 2\gamma\log(A(\tau)) + 2\eta\log(B(\tau)) + C(\tau)r + D(\tau)V}{\tau} \quad (71)$$

where

$$A(\tau) = \frac{2\phi}{(\delta+\phi)[\exp(\phi\tau)-1]+2\phi},$$

$$B(\tau) = \frac{2\psi}{(\nu+\psi)[\exp(\psi\tau)-1]+2\psi},$$

$$C(\tau) = \frac{\alpha\phi[\exp(\psi\tau)-1]B(\tau) - \beta\psi[\exp(\phi\tau)-1]A(\tau)}{\phi\psi(\beta-\alpha)},$$

$$D(\tau) = \frac{\psi[\exp(\phi\tau)-1]A(\tau) - \phi[\exp(\psi\tau)-1]B(\tau)}{\phi\psi(\beta-\alpha)},$$

and

$$\nu = \xi + \lambda,$$

$$\phi = \sqrt{2\alpha+\delta^2},$$

$$\psi = \sqrt{2\beta+\nu^2},$$

$$\kappa = \gamma(\delta+\phi) + \eta(\nu+\psi),$$

where λ is the market price of risk parameter. That is, LS show that the instantaneous expected return on a τ-period bond is

$$r + \lambda\frac{[\exp(\psi\tau)-1]B(\tau)}{\psi(\beta-\alpha)}(\alpha r - V).$$

Given a proxy yield with τ-periods until maturity, the moment conditions analogous to the short-rate moment conditions (65), (66), and (69) are

$$E[y(r,V,\tau)] = -\kappa - \frac{2\gamma}{\tau}\log(A(\tau)) - \frac{2\eta}{\tau}\log(B(\tau))$$

$$-\frac{1}{\tau}C(\tau)E(r)-\frac{1}{\tau}D(\tau)E(V), \tag{72}$$

$$\operatorname{Var}[y(r,V,\tau)]=\frac{1}{\tau^2}C(\tau)^2\operatorname{Var}[r]+\frac{1}{\tau^2}D(\tau)^2\operatorname{Var}[V]$$
$$+2\frac{1}{\tau^2}C(\tau)D(\tau)\operatorname{Cov}[r,V], \tag{73}$$

and

$$\operatorname{Cov}[y_{t+\Delta}(r,V,\tau),y_t(r,V,\tau)]=\frac{C(\tau)^2}{\tau^2}\operatorname{Cov}[r_{t+\Delta},r_t]$$

$$+\frac{D(\tau)^2}{\tau^2}\operatorname{Cov}[V_{t+\Delta},V_t]+\frac{C(\tau)D(\tau)}{\tau^2}\left\{\operatorname{Cov}[r_{t+\Delta},V_t]+\operatorname{Cov}[V_{t+\Delta},r_t]\right\}, \tag{74}$$

where

$$\operatorname{Cov}[r,V]=\frac{\alpha^3\gamma}{2\delta^2}+\frac{\beta^3\eta}{2\xi^2},$$

and

$$\operatorname{Cov}[r_{t+\Delta},V_t]=\frac{\alpha^3\gamma}{2\delta^2}\exp(-\delta\Delta)+\frac{\beta^3\eta}{2\xi^2}\exp(-\xi\Delta)=\operatorname{Cov}[V_{t+\Delta},r_t].$$

Applying Itô's lemma to (63) and (64) and taking expectations, the volatility of the yield proxy is

$$V^y=\left[-\frac{\alpha C(\tau)}{\tau}-\frac{\alpha^2D(\tau)}{\tau}\right]^2\frac{\beta r-V}{\alpha(\beta-\alpha)}$$
$$+\left[-\frac{\beta C(\tau)}{\tau}-\frac{\beta^2D(\tau)}{\tau}\right]^2\frac{V-\alpha r}{\beta(\beta-\alpha)}.$$

The moment conditions analogous to the volatility moment conditions (67), (68), and (70) are

$$E[V^y]=\left[-\frac{\alpha C(\tau)}{\tau}-\frac{\alpha^2D(\tau)}{\tau}\right]^2\frac{\gamma}{\delta}+\left[-\frac{\beta C(\tau)}{\tau}-\frac{\beta^2D(\tau)}{\tau}\right]^2\frac{\eta}{\xi}, \tag{75}$$

$$\operatorname{Var}[V^y]=\left[-\frac{\alpha C(\tau)}{\tau}-\frac{\alpha^2D(\tau)}{\tau}\right]^4\frac{\gamma}{2\delta^2}+\left[-\frac{\beta C(\tau)}{\tau}-\frac{\beta^2D(\tau)}{\tau}\right]^4\frac{\eta}{2\xi^2}, \tag{76}$$

and

$$\operatorname{Cov}[V^y_{t+\Delta},V^y_t]=\left[-\frac{\alpha C(\tau)}{\tau}-\frac{\alpha^2D(\tau)}{\tau}\right]^4\frac{\gamma}{2\delta^2}\exp(-\delta\Delta)$$

$$+\left[-\frac{\beta C(\tau)}{\tau}-\frac{\beta^2 D(\tau)}{\tau}\right]^4 \frac{\eta}{2\xi^2}\exp(-\xi\Delta). \qquad (77)$$

Similar to our approach with the one-factor affine models, we evaluate the proxy problem by assuming that an econometrician has constructed the sample analogs of (72) through (77), but proceeds as if he has constructed the sample analogs of (65) through (70).

The first step in examining the extent of any possible proxy problem in the LS model is to obtain reasonable estimates of the model parameters. To do this, we assume that the one-week Eurodollar data used in Aït-Sahalia (1996a,b) is effectively the short-rate, construct estimates of the unconditional moments of r and V, and use (65) through (70) to obtain estimates of the parameters.[20] Figure 13 contains plots of the end-of-month value of the Eurodollar rate and the square-root of the monthly average of the volatility of the daily change in the Eurodollar rate, an estimate of V. The point estimates of the unconditional moments used to identify the model parameters are shown in Table 5.[21]

Given these moment conditions, the nonlinear equations defined in Appendix B can be solved numerically for α and β to yield the values $\alpha = 0.0447$ and $\beta = 0.0030$.[22] These values imply that $\delta = 0.1679$, and $\xi = 0.1294$, which (combined with α and β) imply that $\gamma = 0.0121$ and $\eta = 3.5359$. These are the "true" parameter values implied by the "true" state variables r and V. These six parameter values, along with the risk premium parameter $\lambda = -0.235$, imply values for the moments of finite-maturity yields and the volatility of yield changes. Calculating these moments for a three-month bond and using these moments in the calculation of the model parameters, yields the following proxy-based parameters: $\alpha^p = 0.1770$, $\beta^p = 0.0044$, $\delta^p = 0.1455$, $\xi^p = 0.1407$, $\gamma^p = 0.0000335$, and $\eta = 2.8202$. A comparison of the true and the proxy-based parameters indicates that the largest differences are in the calculation of α and γ.

These parameter differences imply different estimates of the drift and the diffusion functions. In order to assess the economic significance of these differences, Table 6 presents calculations of the differences in the yields on discount bonds of different maturities implied by the Longstaff and Schwartz

[20] We thank Yacine Aït-Sahalia for generously making his data available.

[21] Figure 13 suggests that the volatility of Eurodollar rate changes over this period experienced two distinct "regimes," a high volatility regime in the first half of the sample and a low volatility regime in the second half of the sample. Regime switching is inconsistent with the model's assumptions about the evolution of V, but our goal here is to provide a rough calibration of the model's parameters.

[22] These are not necessarily the *unique* solution to the pair of nonlinear equations.

model evaluated at the two sets of parameters. When the difference has a negative sign, it indicates that the proxy-based parameters understate the yield levels relative to the true parameter values. Consistent with the results from the single-factor affine cases, these differences are generally small for maturities of five years or less, although extreme values of volatility and the short rate can generate pricing differences as large as 1 percent. At the ten and twenty year horizons, the differences in the yields can be larger. At high short rate levels, the proxy-based estimates understate the true ten-year yields by as much as 2.9 percent for low volatility, and for low levels of the short rate, the proxy-based estimates overstate the true yield by 1.8 percent for high volatility. At the twenty year horizon, the proxy-based estimates always undervalue the bonds (relative to the true price), with the largest difference being 8.8 percent. This implies that using a three-month proxy to construct bond price estimates will result in flatter yield curve than would be implied by the true state variables.

As in the single-factor affine models, it is possible to invert the moments of the measured yields and yield volatilities to recover estimates of the parameters of the true state variables r and V. However, as in the single factor case, if the proxy problem is unimportant, then the corrections to a moment-based estimation strategy will also be unimportant.

An examination of the proxy problem in multi-factor term structure models can be generalized in at least two directions. First, there are two-factor affine models that identify the state variables differently from LS, and these formulations will generate moment conditions that can be evaluated in a manner that is qualitatively similar to the LS case.[23] In fact, because it relies on a proxy for the short rate to construct the moments of both factors, it may well be the case that when estimation is based on moments of the short rate and its volatility the LS formulation of a two-factor affine model has a larger proxy problem than other formulations. Second, the Taylor series approximation introduced in Proposition 3 can be extended in a straightforward manner to evaluate a nonlinear multivariate model.[24]

[23] For example, CIR and Pearson and Sun (1994) identify the factors as a "real rate" process and an "expected inflation" process, and Brennan and Schwartz (1979) and Boudoukh, Richardson, Stanton, and Whitelaw (1998) use a short rate and a long rate.

[24] This multivariate Taylor series expansion could also serve as the basis for method of moments estimation of the parameters of r, V, and λ.

6 Conclusions

In this paper, we have provided a detailed analysis of the impact of using a 1- or 3-month yield as a proxy for the unobservable short rate. The results are encouraging for a large part of the existing literature. For single factor affine models in general and for the two-factor affine model we examined, the economic significance of the parameter estimate errors and the resulting bond price errors is generally negligible, although for long-term bonds there can be noticeable differences between the true bond prices and the prices based on the use of a short maturity proxy in the two-factor LS model. However, the use of a proxy results in economically significant errors in the nonlinear model we examined. In particular, one of the most basic properties of the diffusion function, whether it is convex or concave, can be sensitive to the choice of proxy. Of course, these conclusions apply only to the specific models that we examined. In addition, the economic significance of the deviations in the estimates of the drift and diffusion functions was only evaluated using the prices of discount bonds. Nonetheless, the results indicate that the proxy problem can be important for nonlinear models and suggest that researchers who estimate such models may need to evaluate it on a case-by-case basis.

One of the principal contributions of the paper is to provide two separate methods, based on the numerical solution to the fundamental bond pricing partial differential equation and truncated Taylor series expansions, to assess the magnitude of the proxy problem. For models in which the proxy problem is not severe, empirical researchers can feel confident in examining the properties of the model using the yields of bills with maturities as long as three months. These data are more likely to be free from some of the microstructure problems or institutional features of the market that might affect the observed prices of very short-term interest rate series like the Federal Funds rate, overnight repo rates or even the 1-month Treasury bill rate. However, for some single factor models, an explicit calculation of the magnitude of the proxy problem might dictate devoting more attention to explicit modelling of some of the data issues associated with the use of very short-term interest rate series. Furthermore, the Taylor series approximation could serve as the basis for a direct method of moments estimation of the short rate parameters from the moments of finite-maturity bonds, even in settings where there are no closed-form bond pricing solutions.

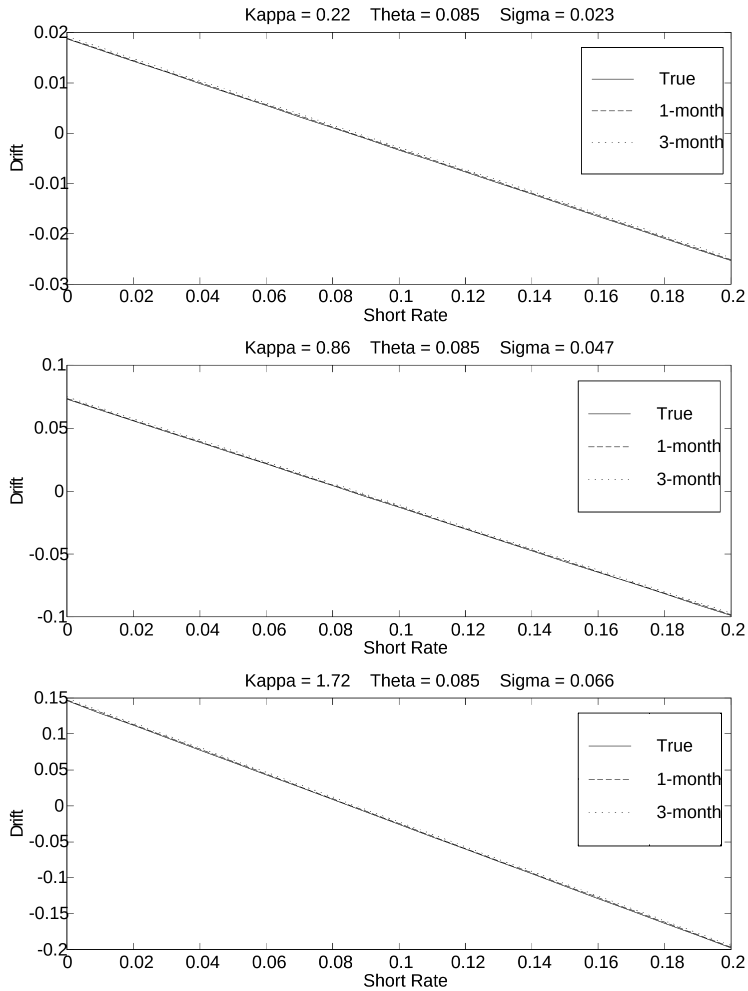

Figure 1: The Drift Functions in the Vasicek (1977) Case.

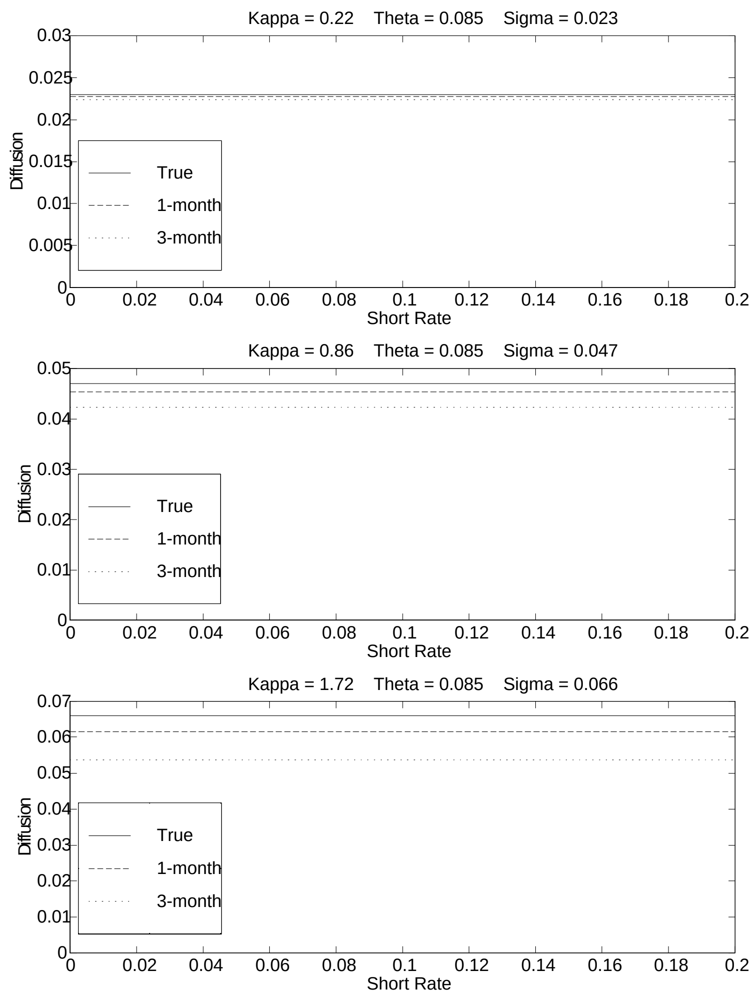

Figure 2: The Diffusion Functions in the Vasicek (1977) Case.

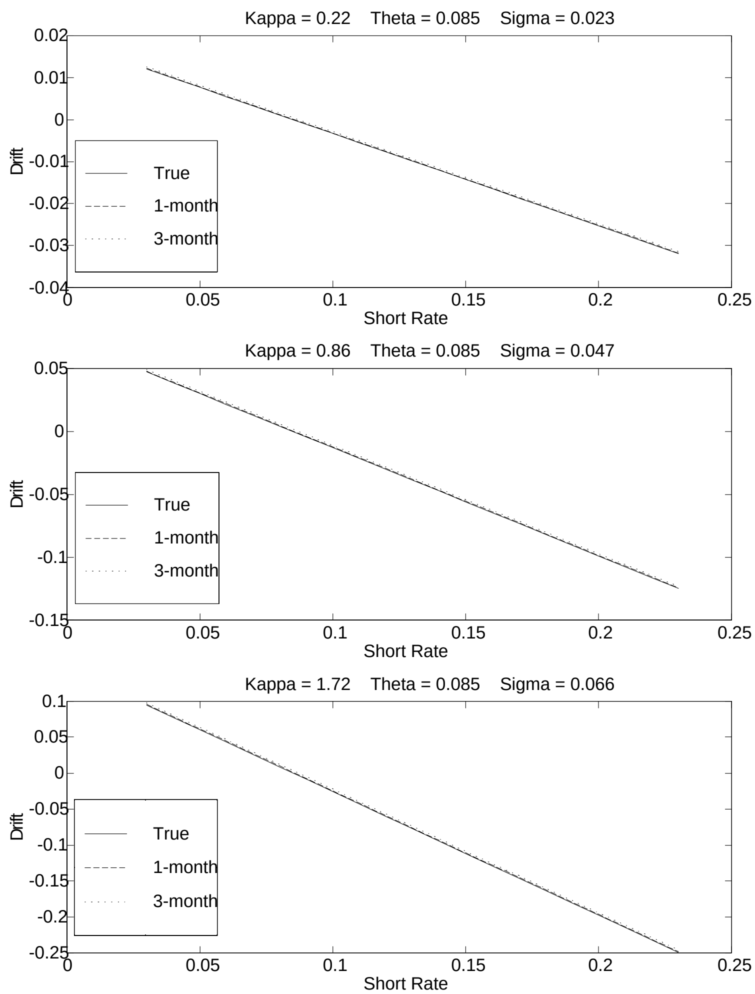

Figure 3: The Drift Functions in the CIR Case.

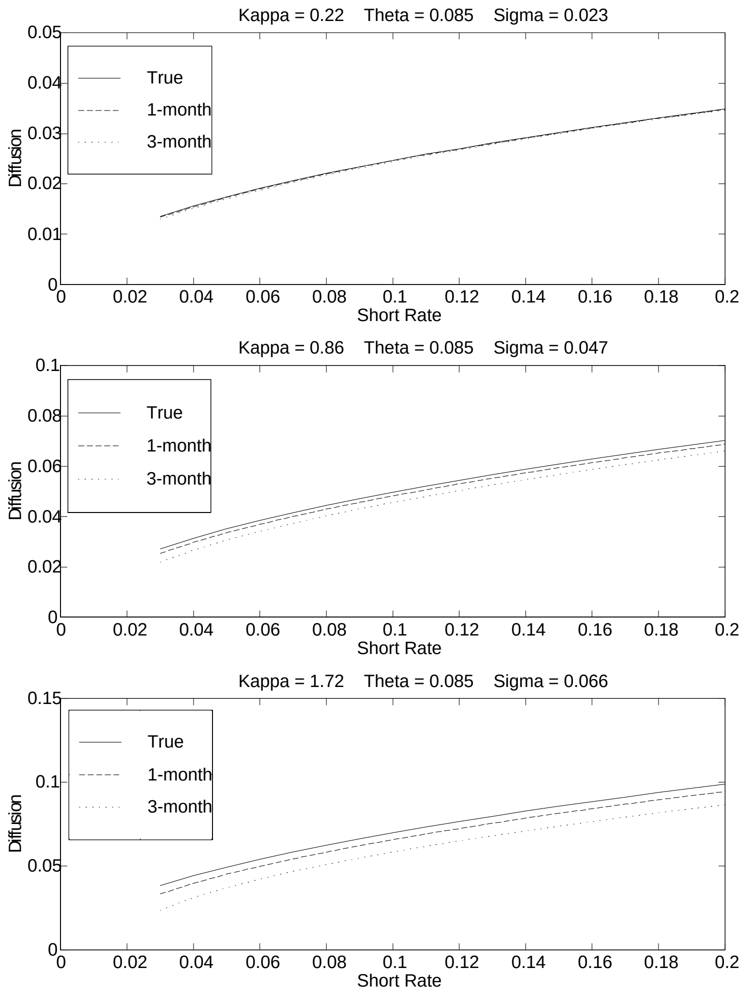

Figure 4: The Diffusion Functions in the CIR Case.

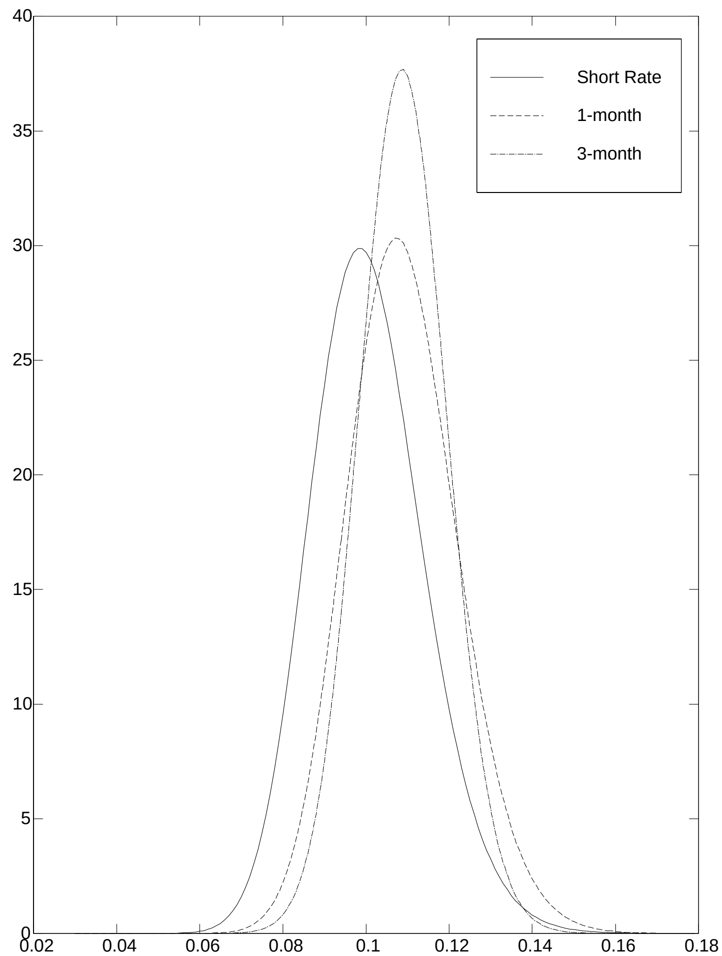

Figure 5: The Stationary Densities of the True short rate, the 1-month Yield, and the 3-Month Yield in the Nonlinear Model Approximated Using the Finite-Difference Approach.

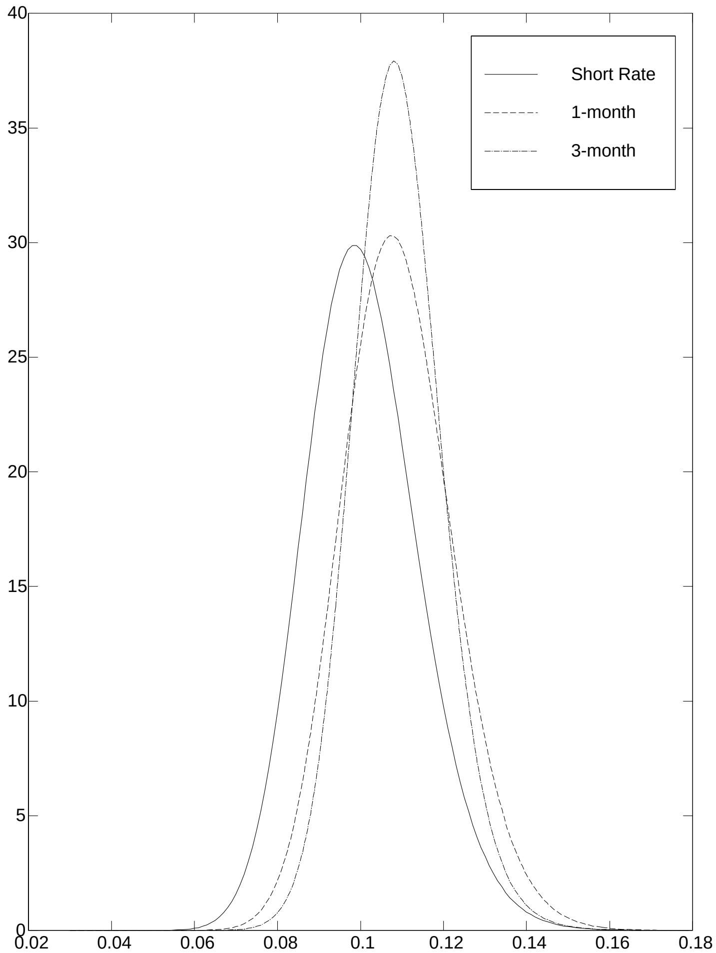

Figure 6: The Stationary Densities of the True short rate, the 1-month Yield, and the 3-Month Yield in the Nonlinear Model Approximated Using the Taylor Series Approach.

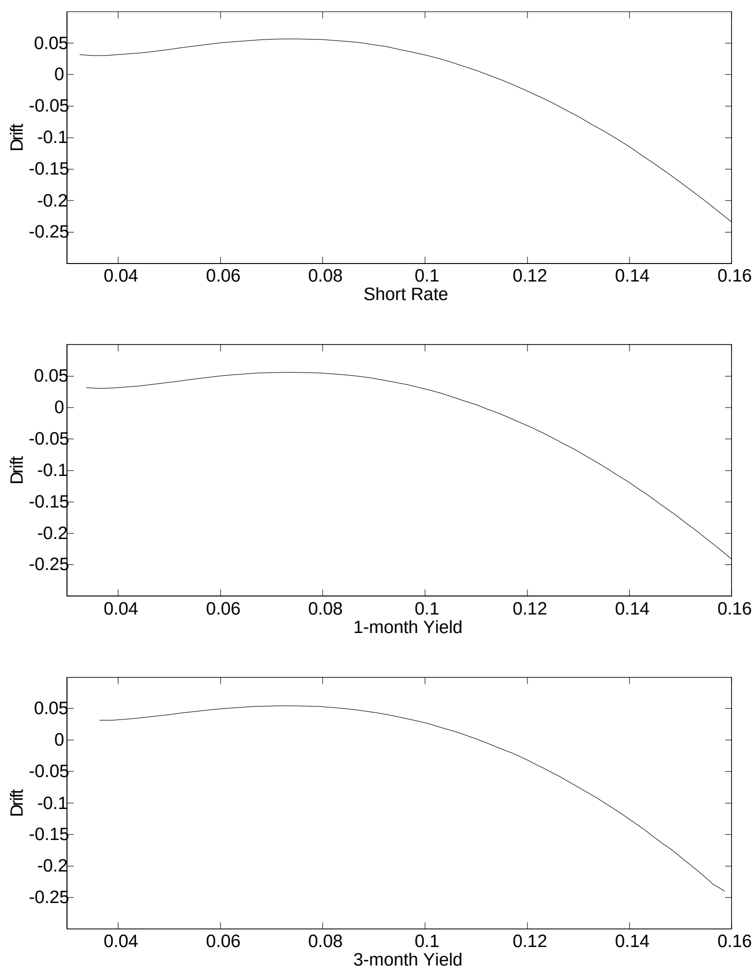

Figure 7: The Drift Functions in the Nonlinear Model Approximated Using the Finite-Difference Approach.

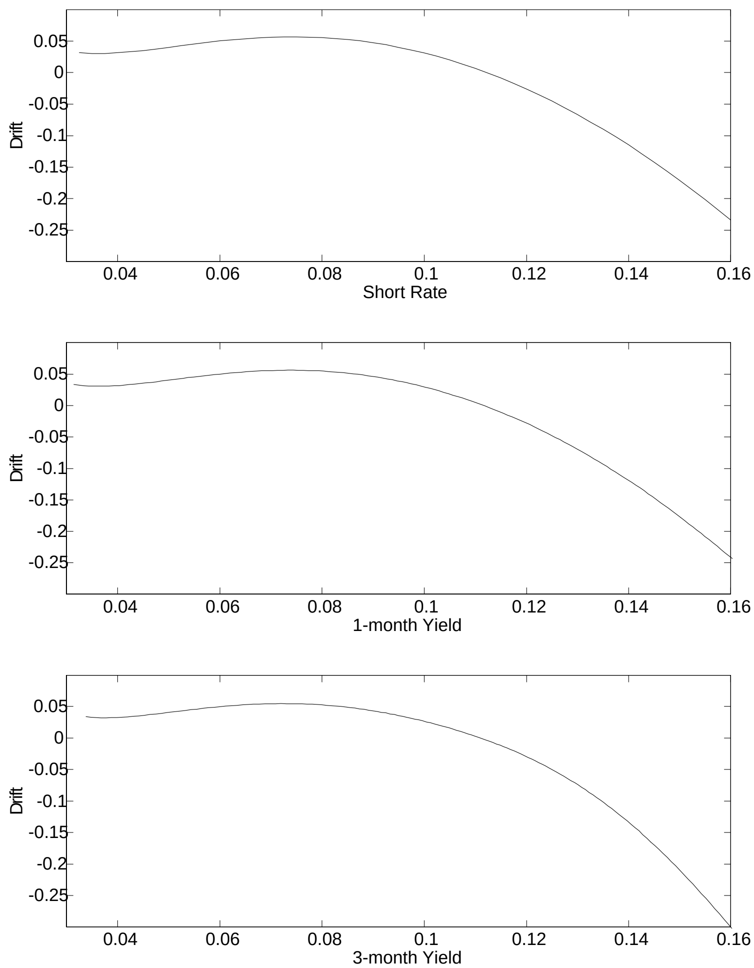

Figure 8: The Drift Functions in the Nonlinear Model Approximated Using the Taylor-Series Approach.

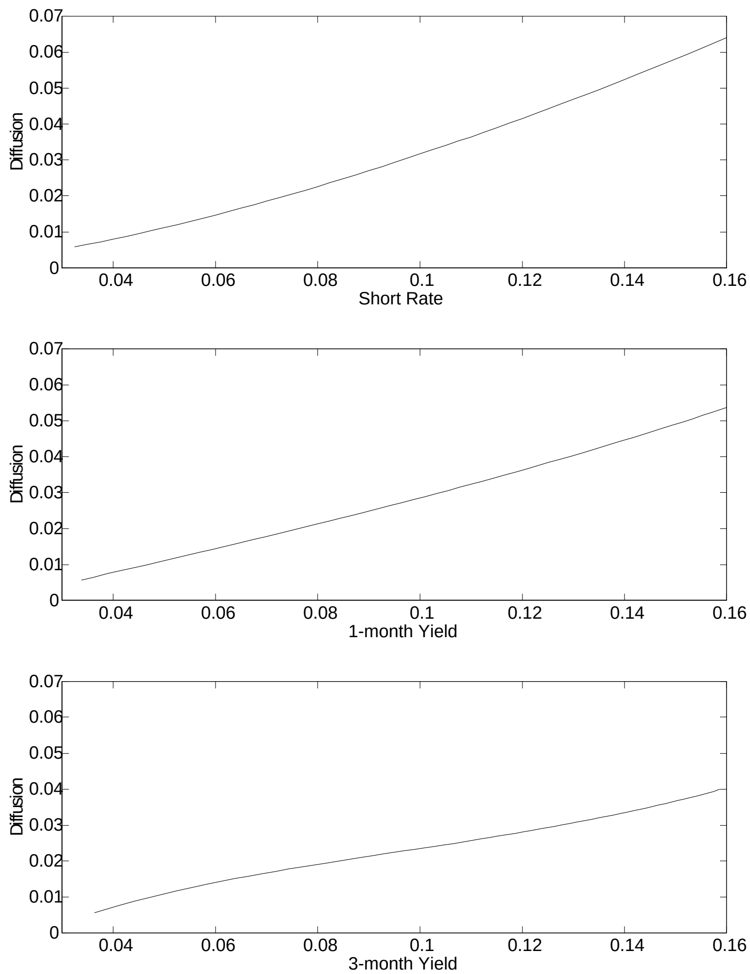

Figure 9: The Diffusion Functions in the Nonlinear Model Approximated Using the Finite-Difference Approach.

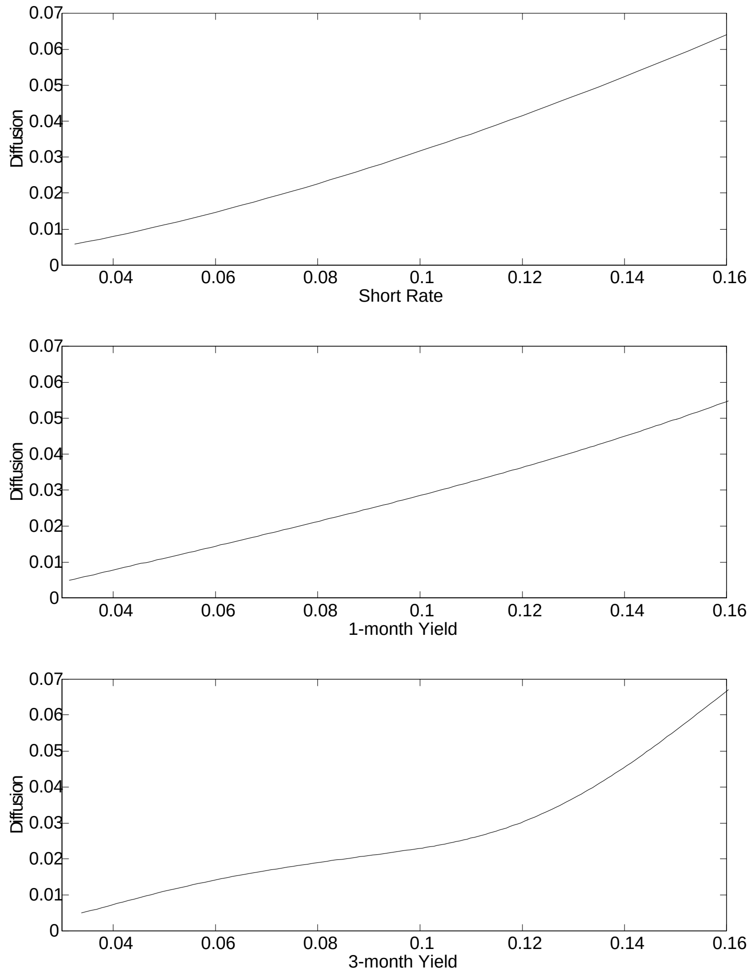

Figure 10: The Diffusion Functions in the Nonlinear Model Approximated Using the Taylor-Series Approach.

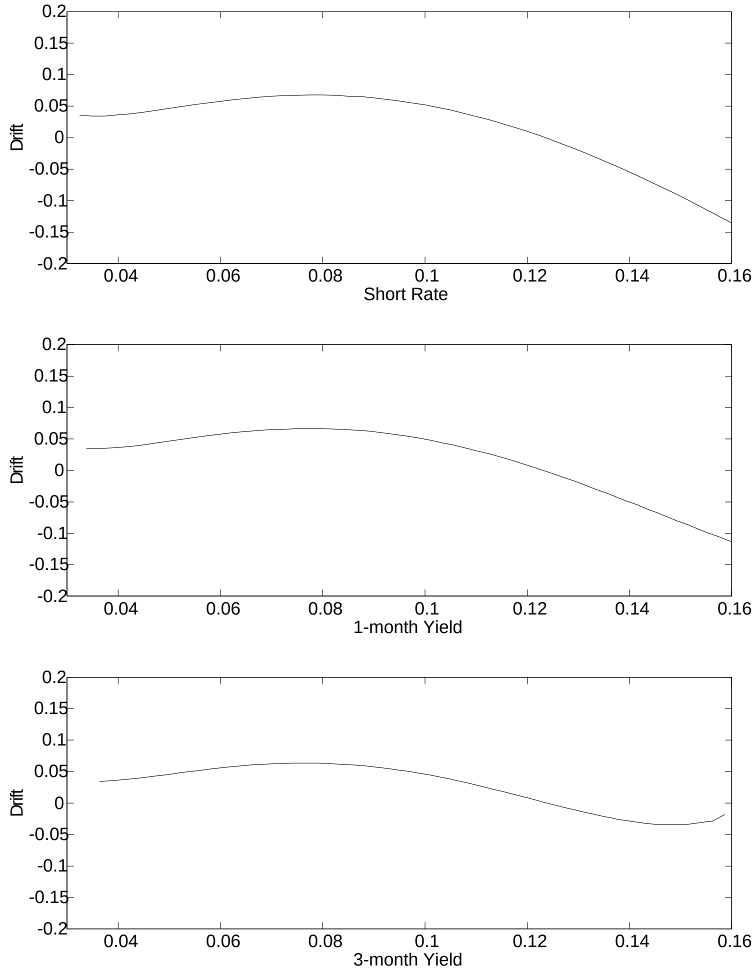

Figure 11: The Drift Functions in the Nonlinear Model Under the Equivalent Martingale Measure Approximated Using the Finite-Difference Approach.

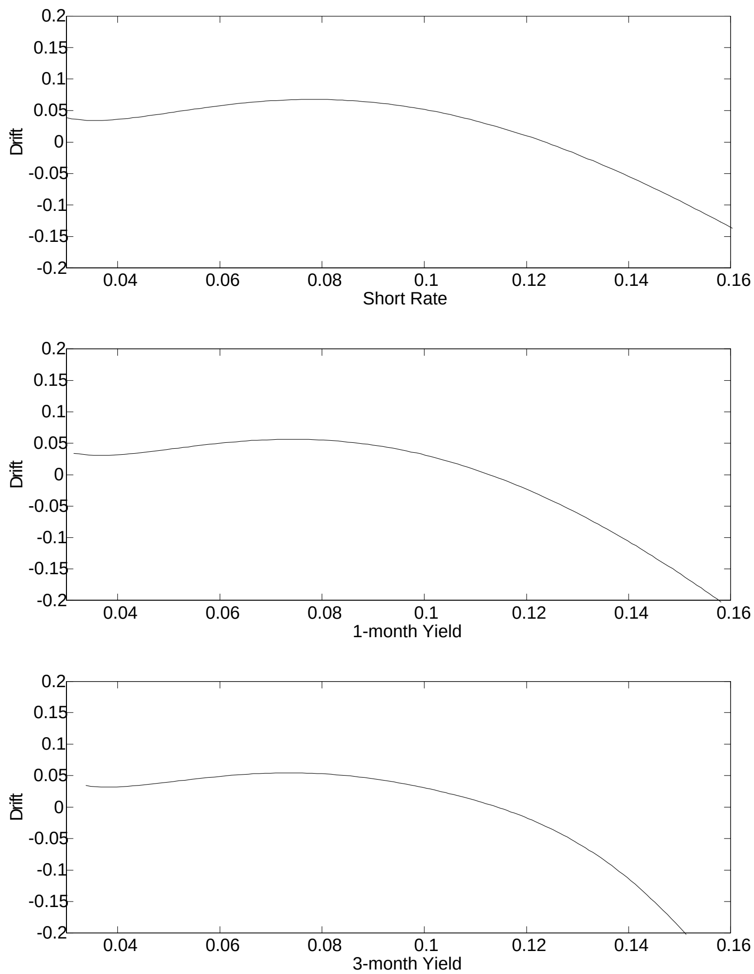

Figure 12: The Drift Functions in the Nonlinear Model Under the Equivalent Martingale Measure Approximated Using the Taylor-Series Approach.

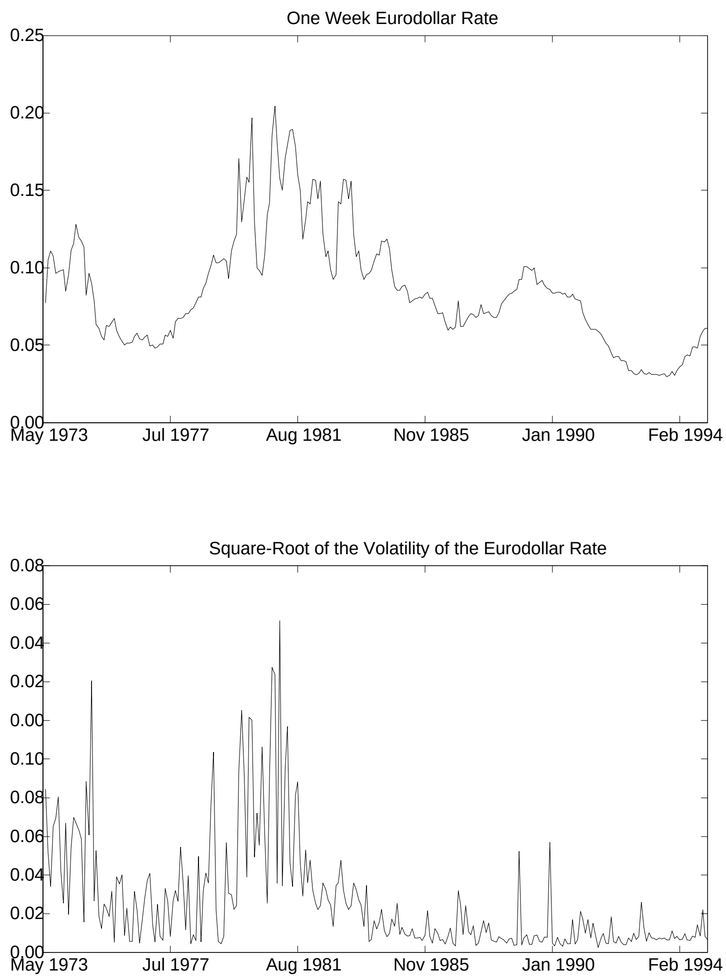

Figure 13: Monthly Observations of the Level of the One-Week Eurodollar Rate and the Square-Root of the Monthly Average Volatility of the Daily Change in the One Week Eurodollar Rate.

Table 1: The Effect of Maturity-Induced Specification Error on Measured Yields in the Vasicek (1977) Model.

Panel A: $\kappa = 0.22$, $\theta = 0.0850$, $\sigma = 0.023$, and $\lambda_0 = -0.02$.

	Maturities				
Estimates	3-month	1-year	5-year	10-year	20-year
True	8.74	9.42	11.97	13.69	15.19
Simple MoM	8.75	9.45	12.08	13.85	15.40

Panel B: $\kappa = 0.86$, $\theta = 0.0850$, $\sigma = 0.047$, and $\lambda_0 = -0.02$.

	Maturities				
Estimates	3-month	1-year	5-year	10-year	20-year
True	8.73	9.25	10.19	10.43	10.55
Simple MoM	8.76	9.33	10.39	10.66	10.80

Panel C: $\kappa = 1.72$, $\theta = 0.0850$, $\sigma = 0.066$, and $\lambda_0 = -0.02$.

	Maturities				
Estimates	3-Month	1-year	5-year	10-year	20-year
True	8.71	9.08	9.47	9.53	9.56
Simple MoM	8.76	9.20	9.68	9.75	9.79

"True" refers to the bond yields implied by the assumed parameter values evaluated at the true long-run mean. "Simple MoM" are the bond yields implied by using the simple method-of-moments estimator described in the text applied to a sample of observations of a proxy bill with three months to maturity; i.e., $\tau = 0.25$, evaluated at the long-run mean estimated from the method-of-moments. The sample is assumed to be long enough to recover the population moments of the proxy, and the market price of risk parameter is assumed to be known in constructing the estimated yields.

Table 2: The Effect of Maturity-Induced Specification Error on Measured Yields in the CIR Model.

Panel A: $\kappa = 0.22$, $\theta = 0.085$, $\sigma = 0.078$, and $\lambda_1 = -0.235$.

	Maturities				
Estimates	3-month	1-year	5-year	10-year	20-year
True	8.75	9.49	13.33	17.63	24.04
Simple MoM	8.76	9.52	13.47	17.90	24.53

Panel B: $\kappa = 0.86$, $\theta = 0.085$, $\sigma = 0.157$, and $\lambda_1 = -0.235$.

	Maturities				
Estimates	3-month	1-year	5-year	10-year	20-year
True	8.74	9.30	10.54	10.93	11.14
Simple MoM	8.76	9.38	10.79	11.23	11.47

Panel C: $\kappa = 1.72$, $\theta = 0.085$, $\sigma = 0.221$, and $\lambda_1 = -0.235$.

	Maturities				
Estimates	3-Month	1-year	5-year	10-year	20-year
True	8.72	9.12	9.58	9.66	9.70
Simple MoM	8.76	9.25	9.83	9.92	9.97

"True" refers to the bond yields implied by the assumed parameter values evaluated at the true long-run mean. "Simple MoM" are the bond yields implied by using the simple method-of-moments estimator described in the text applied to a sample of observations of a proxy bill with three months to maturity; i.e., $\tau = 0.25$, evaluated at the long-run mean estimated from the method-of-moments. The sample is assumed to be long enough to recover the population moments of the proxy, and the market price of risk parameter is assumed to be known in constructing the estimated yields.

Table 3: Discount Bond Yields in the Nonlinear Model Using the Drift and Diffusion Functions Constructed Using the Finite-Difference Approach.

Panel A: Bond Yields Using the True Drift and Diffusion Functions.

	Maturities			
Short Rate	1-Year	3-Year	5-Year	10-Year
0.08	9.75	10.48	10.65	10.77
0.10	10.63	10.80	10.84	10.87
0.12	11.28	11.03	10.98	10.94
0.14	11.78	11.21	11.08	10.99

Panel B: Bond Yields Using the Drift and Diffusion Functions from the 1-month Yield.

	Maturities			
1-month Proxy	1-Year	3-Year	5-Year	10-Year
0.08	10.31	11.37	11.62	11.80
0.10	11.28	11.73	11.83	11.91
0.12	12.02	12.00	11.99	11.99
0.14	12.63	12.21	12.12	12.05

Panel C: Bond Yields Using the Drift and Diffusion Functions from the 3-month Yield.

	Maturities			
3-month Proxy	1-Year	3-Year	5-Year	10-Year
0.08	10.27	11.35	11.59	11.77
0.10	11.25	11.71	11.81	11.88
0.12	11.99	11.97	11.96	11.96
0.14	12.59	12.18	12.09	12.02

Discount bond prices are found by solving (6) numerically, using the drift and diffusion functions computed from the finite-difference approximation and the true risk premium function (53). All yields are expressed in percent at an annual rate.

Table 4: Discount Bond Yields in the Nonlinear Model Using the Drift and Diffusion Functions Constructed Using the Taylor-Series Approach.

Panel A: Bond Yields Using the True Drift and Diffusion Functions.

	Maturities			
Short Rate	1-Year	3-Year	5-Year	10-Year
0.08	9.75	10.48	10.65	10.77
0.10	10.63	10.80	10.84	10.87
0.12	11.28	11.03	10.98	10.94
0.14	11.78	11.21	11.08	10.99

Panel B: Bond Yields Using the Drift and Diffusion Functions from the 1-month Yield.

	Maturities			
1-month Proxy	1-Year	3-Year	5-Year	10-Year
0.08	10.31	11.39	11.63	11.81
0.10	11.29	11.75	11.85	11.92
0.12	12.03	12.01	12.01	12.00
0.14	12.64	12.23	12.14	12.07

Panel C: Bond Yields Using the Drift and Diffusion Functions from the 3-month Yield.

	Maturities			
3-month Proxy	1-Year	3-Year	5-Year	10-Year
0.08	10.27	11.33	11.57	11.75
0.10	11.26	11.69	11.78	11.86
0.12	11.98	11.95	11.94	11.93
0.14	12.51	12.13	12.05	11.99

Discount bond prices are found by solving (6) numerically, using the drift and diffusion functions computed from the Taylor-series approximation and the true risk premium function (53). All yields are expressed in percent at an annual rate.

Table 5: Unconditional Sample Moments of the Short-Rate and Volatility, Based on the One-Week Eurodollar Rate. June 1973 to February 1995.

Moment	Short-Rate (r)	Volatility (V)
Mean	0.0853	0.00039
Variance	0.00138	8.61×10^{-7}
Autocovariance	0.00132	3.46×10^{-7}

The short-rate is the one week Eurodollar rate, from Aït-Sahalia (1996a,b), recorded on the last trading day of each month. The volatility series is estimated as the monthly variance of the daily changes in the Eurodollar rate, calculated as $V_t = \sum_{i=1}^{N_t-1} \Delta r_{t,i}$, where N_t is the number of days in month t, $\Delta r_{t,i} = r_{t,i} - r_{t,i-1}$, and i indexes days within month t. 'Autocovariance' is the first-order autocovariance, defined as $\mathrm{Cov}[r_{t+\Delta}, r_t]$ and $\mathrm{Cov}[V_{t+\Delta}, V_t]$ with Δ corresponding to one month.

Table 6: Differences in the Yields on Discount Bonds in the Longstaff and Schwartz Model: Yield Implied by.the Proxies Minus Yields Implied by the True State Variables.

Panel A: One Year Maturity.

	r			
$\sqrt{V}$	2.0	6.0	10.0	14.0
0.316	−0.0144	−0.0694	−0.1240	−0.1800
2.000	0.0853	0.0302	−0.0248	−0.0799
3.160	0.2390	0.1840	0.1280	0.0734

Panel B: Five Years Maturity.

	r			
$\sqrt{V}$	2.0	6.0	10.0	14.0
0.316	−0.1110	−0.4350	−0.7590	−1.1000
2.000	0.3670	0.0430	−0.2810	−0.605
3.160	1.1000	0.7780	0.4540	0.1300

Panel C: Ten Years Maturity.

	r			
$\sqrt{V}$	2.0	6.0	10.0	14.0
0.316	−0.4740	−1.3000	−2.1000	−2.9000
2.000	−0.4270	−0.3890	−1.2000	−2.0000
3.160	1.8000	0.9960	0.1810	−0.6350

Panel D: Twenty Years Maturity.

	r			
$\sqrt{V}$	0.02	0.06	0.10	0.14
0.316	−2.7000	−4.7000	−6.8000	−8.8000
2.000	−1.2000	−3.3000	−5.3000	−7.4000
3.160	−0.9750	−1.1000	−3.1000	−5.2000

The yield, at each maturity, under the proxies is calculated using the parameter values constructed from the moment conditions (72) through (77) and the yield equation (71). The yields under the true state variables is computed analogously, using the moment conditions (65) through (70). All differences are reported in percent at an annual rate.

Appendix A: Evaluating the Finite-Difference Approach.

This appendix provides information on the magnitude of the approximation errors associated with the finite-difference approach, when applied to the Vasicek and CIR models. Tables A-1 through A-4 show the maximum absolute error in approximating the 1- and 3-month yield functions for the three parameterizations examined in Figures 1 through 4, for a variety of grid sizes in the time and space dimensions.[25] For all parameterization, both maturities, and both models, these tables show that the approximation error declines uniformly in the number of *time* points, but the error is insensitive to increases in the number of *space* points. When the number of time points equals 480, the maximum absolute error in approximation the yield over the range of 0.01 to 0.20 is roughly four basis points. When the number of times points is increased to 720, the maximum absolute error reduces to three basis points. These results are consistent across both models and all of the parameterizations examined in Tables A-1 through A-4.

Tables A-5 and A-6 evaluate the maximum absolute error in the numerical approximation of the first and second derivative of the yield in the Vasicek case. Tables A-7 and A-8 report the same quantities for the CIR case. Each table reports the results from grids of size $N = 19$ to $N = 152$ defined on the interpolated yield function for the interval $r \in [0.03, 0.20]$.[26] The magnitudes of the errors for the first derivatives in both models are roughly the same, and they appear to be small. The magnitudes of the errors of the second derivative estimates are larger and they are increasing in κ, the mean reversion parameter in each model. The maximum absolute error for the Vasicek case with a 3-month maturity and $\kappa = 1.72$ are an order of magnitude larger than the errors for other parameterizations of either model. However, graphs of the approximation (not reported here) indicate that the error is only large for very small levels of the short rate (approximately 0.03 to 0.035).

The previous tables provide important information on the overall per-

[25] As Wilmot, Howison, and Dewynne (1995) note: "The Crank-Nicholson implicit finite-difference scheme is essentially an average of . . . implicit and explicit methods." [pages 156-157]. It is both stable and unconditionally convergent. See Wilmot, Howison, and Dewynne (1995) Exercise 17 on page 163. Press et al (1992) Section 19.2 also discusses the unconditional convergence properties of the Crank-Nicholson scheme. This means that there is no need to constrain the ratio of the size of the time steps and the state steps to any specific magnitude.

[26] Using the results from Tables A-1 through A-4, all of the results in Tables A-5 through A-8 are based on a finite difference approximation with 720 time points and 240 space points.

formance of the approximation, but they do not answer the most important question: How close are the approximate drift and diffusion functions to the true drift and diffusion functions? This question is answered in Table A-9 (for the Vasicek case) and Table A-10 (for the CIR case). In both cases, the approximations are quite precise. The maximum error is typically just a few basis points, and the approximations are accurate for both the drift and diffusion functions. On the basis of these results, we conclude that the finite-difference approach can provide valuable information on the extent of the proxy problem in cases where there is no explicit solution for the yield function.

This appendix has evaluated the finite-difference approach only. However, the absolute value of the maximum error – over the range $[0.03, 0.18]$ – in constructing the drift and diffusion functions using the Taylor series approximations of order 3 was less than 1/100 of a basis point. These results are available upon request.

Table A-1: Maximum Absolute Error in the Finite-Difference Calculation of $y\,(r, \tau = 1 \text{ month})$ in theVasicek Case.

Panel A: $\kappa = 0.22$, $\theta = 0.085$, $\sigma = 0.023$, and $\lambda_0 = -0.02$.

	No. of Space Points			
No. of Time Points	120	240	480	720
120	0.00166	0.00164	0.00166	0.00166
240	0.00083	0.00082	0.00083	0.00083
480	0.00042	0.00041	0.00042	0.00042
720	0.00028	0.00027	0.00028	0.00028

Panel B: $\kappa = 0.86$, $\theta = 0.085$, $\sigma = 0.047$, and $\lambda_0 = -0.02$.

	No. of Space Points			
No. of Time Points	120	240	480	720
120	0.00164	0.00160	0.00162	0.00161
240	0.00083	0.00081	0.00081	0.00081
480	0.00043	0.00041	0.00041	0.00041
720	0.00030	0.00028	0.00027	0.00027

Panel C: $\kappa = 1.72$, $\theta = 0.085$, $\sigma = 0.066$, and $\lambda_0 = -0.02$.

	No. of Space Points			
No. of Time Points	120	240	480	720
120	0.00161	0.00156	0.00156	0.00155
240	0.00083	0.00079	0.00079	0.00078
480	0.00044	0.00040	0.00040	0.00040
720	0.00032	0.00031	0.00032	0.00032

The bond pricing function is approximated using a Crank-Nicholson finite difference algorithm, as described in Duffie (1996). The maximum error is reported over the range $r \in [0.01, 0.20]$.

Table A-2: Maximum Absolute Error in the Finite-Difference Calculation of $y(r, \tau = 3 \text{ months})$ in theVasicek Case.

Panel A: $\kappa = 0.22$, $\theta = 0.085$, $\sigma = 0.023$, and $\lambda_0 = -0.02$.

	No. of Space Points			
No. of Time Points	120	240	480	720
120	0.00166	0.00163	0.00165	0.00165
240	0.00083	0.00082	0.00083	0.00083
480	0.00042	0.00041	0.00041	0.00041
720	0.00028	0.00027	0.00028	0.00028

Panel B: $\kappa = 0.86$, $\theta = 0.085$, $\sigma = 0.047$, and $\lambda_0 = -0.02$.

	No. of Space Points			
No. of Time Points	120	240	480	720
120	0.00159	0.00153	0.00153	0.00152
240	0.00083	0.00078	0.00077	0.00077
480	0.00045	0.00046	0.00048	0.00049
720	0.00040	0.00045	0.00046	0.00047

Panel C: $\kappa = 1.72$, $\theta = 0.085$, $\sigma = 0.066$, and $\lambda_0 = -0.02$.

	No. of Space Points			
No. of Time Points	120	240	480	720
120	0.00150	0.00141	0.00140	0.00138
240	0.00098	0.00105	0.00107	0.00109
480	0.00091	0.00098	0.00100	0.00102
720	0.00088	0.00095	0.00097	0.00099

The bond pricing function is approximated using a Crank-Nicholson finite difference algorithm, as described in Duffie (1996). The maximum error is reported over the range $r \in [0.01, 0.20]$.

Table A-3: Maximum Absolute Error in the Finite-Difference Calculation of $y\,(r, \tau = 1 \text{ month})$ in the CIR Case.

Panel A: $\kappa = 0.22$, $\theta = 0.085$, $\sigma = 0.078$, and $\lambda_1 = -0.235$.

	No. of Space Points			
No. of Time Points	120	240	480	720
120	0.00167	0.00165	0.00168	0.00167
240	0.00083	0.00082	0.00084	0.00084
480	0.00041	0.00041	0.00042	0.00042
720	0.00027	0.00027	0.00028	0.00028

Panel B: $\kappa = 0.86$, $\theta = 0.085$, $\sigma = 0.157$, and $\lambda_1 = -0.235$.

	No. of Space Points			
No. of Time Points	120	240	480	720
120	0.00164	0.00161	0.00163	0.00163
240	0.00083	0.00081	0.00082	0.00081
480	0.00042	0.00041	0.00041	0.00041
720	0.00029	0.00027	0.00027	0.00027

Panel C: $\kappa = 1.72$, $\theta = 0.085$, $\sigma = 0.221$, and $\lambda_1 = -0.235$.

	No. of Space Points			
No. of Time Points	120	240	480	720
120	0.00161	0.00157	0.00158	0.00157
240	0.00083	0.00079	0.00079	0.00079
480	0.00043	0.00041	0.00040	0.00040
720	0.00030	0.00028	0.00027	0.00027

The bond pricing function is approximated using a Crank-Nicholson finite difference algorithm, as described in Duffie (1996). The maximum error is reported over the range $r \in [0.01, 0.20]$.

Table A-4: Maximum Absolute Error in the Finite-Difference Calculation of $y(r, \tau = 3 \text{ month})$ in the CIR Case.

Panel A: $\kappa = 0.22$, $\theta = 0.085$, $\sigma = 0.078$, and $\lambda_1 = -0.235$.

	No. of Space Points			
No. of Time Points	120	240	480	720
120	0.00168	0.00167	0.00170	0.00170
240	0.00083	0.00083	0.00085	0.00085
480	0.00040	0.00041	0.00042	0.00042
720	0.00026	0.00027	0.00028	0.00028

Panel B: $\kappa = 0.86$, $\theta = 0.085$, $\sigma = 0.157$, and $\lambda_1 = -0.235$.

	No. of Space Points			
No. of Time Points	120	240	480	720
120	0.00160	0.00156	0.00157	0.00157
240	0.00082	0.00079	0.00079	0.00079
480	0.00043	0.00040	0.00040	0.00040
720	0.00030	0.00027	0.00027	0.00027

Panel C: $\kappa = 1.72$, $\theta = 0.085$, $\sigma = 0.221$, and $\lambda_1 = -0.235$.

	No. of Space Points			
No. of Time Points	120	240	480	720
120	0.00151	0.00143	0.00143	0.00141
240	0.00081	0.00074	0.00072	0.00071
480	0.00046	0.00039	0.00037	0.00037
720	0.00034	0.00028	0.00026	0.00025

The bond pricing function is approximated using a Crank-Nicholson finite difference algorithm, as described in Duffie (1996). The maximum error is reported over the range $r \in [0.01, 0.20]$.

Table A-5: Maximum Absolute Error in the Estimates of $\partial y/\partial r$ for the Vasicek Case.

Panel A: $\kappa = 0.22$, $\theta = 0.085$, $\sigma = 0.023$, and $\lambda_0 = -0.02$.

	$N = \#$ of Grid Points			
Maturity	$N = 19$	$N = 38$	$N = 76$	$N = 152$
1-month	0.00141	0.00141	0.00141	0.00141
3-months	0.00144	0.00144	0.00144	0.00144

Panel B: $\kappa = 0.86$, $\theta = 0.085$, $\sigma = 0.047$, and $\lambda_0 = -0.02$.

	$N = \#$ of Grid Points			
Maturity	$N = 19$	$N = 38$	$N = 76$	$N = 152$
1-month	0.00146	0.00146	0.00146	0.00146
3-months	0.00369	0.00262	0.00155	0.00232

Panel C: $\kappa = 1.72$, $\theta = 0.085$, $\sigma = 0.066$, and $\lambda_0 = -0.02$.

	$N = \#$ of Grid Points			
Maturity	$N = 19$	$N = 38$	$N = 76$	$N = 152$
1-month	0.00151	0.00151	0.00151	0.00151
3-months	0.01356	0.01185	0.00854	0.01134

The approximation to the partial derivative of the yield function with respect to r is calculated by interpolating the finite-difference solution using, 720 time points and 240 space points, over the range of r from $[0.01, 0.20]$ using cubic splines. The interpolated function is then evaluated at a grid of size N, and the first derivative is approximated using the formula

$$\frac{\partial \widehat{y}\left(r_i, \tau\right)}{\partial r} \approx \frac{\widehat{y}\left(r_{i+1}, \tau\right) - \widehat{y}\left(r_{i-1}, \tau\right)}{2\delta}$$

where τ is held constant, and $\widehat{y}\left(r_i, \tau\right)$ refers to the approximate yield function evaluated at the grid point i. δ is the constant width of the interval between grid points. The maximum errors are reported over the range of $r \in [0.03, 0.20]$. There were large approximation errors for extremely small values of r.

Table A-6: Maximum Absolute Error in the Estimates of $\partial^2 y/\partial r^2$ for the Vasicek Case.

Panel A: $\kappa = 0.22$, $\theta = 0.085$, $\sigma = 0.023$, and $\lambda_0 = -0.02$.

	$N = \#$ of Grid Points			
Maturity	$N = 19$	$N = 38$	$N = 76$	$N = 152$
1-month	0.00010	0.00012	0.00012	0.00013
3-months	0.01095	0.00556	0.00164	0.00433

Panel B: $\kappa = 0.86$, $\theta = 0.085$, $\sigma = 0.047$, and $\lambda_0 = -0.02$.

	$N = \#$ of Grid Points			
Maturity	$N = 19$	$N = 38$	$N = 76$	$N = 152$
1-month	0.07889	0.05498	0.02549	0.04858
3-months	0.64403	0.57596	0.39545	0.55577

Panel C: $\kappa = 1.72$, $\theta = 0.085$, $\sigma = 0.066$, and $\lambda_0 = -0.02$.

	$N = \#$ of Grid Points			
Maturity	$N = 19$	$N = 38$	$N = 76$	$N = 152$
1-month	0.40229	0.34343	0.21587	0.32624
3-months	1.40602	1.32290	1.01671	1.29778

The approximation to the partial derivative of the yield function with respect to r is calculated by interpolating the finite-difference solution using, 720 time points and 240 space points, over the range of r from $[0.01, 0.20]$ using cubic splines. The interpolated function is then evaluated at a grid of size N, and the second derivative is approximated using the formula

$$\frac{\partial^2 \widehat{y}\left(r_i, \tau\right)}{\partial r^2} \approx \frac{\widehat{y}\left(r_{i+1}, \tau\right) - 2\widehat{y}\left(r_i, \tau\right) + \widehat{y}\left(r_{i-1}, \tau\right)}{\delta^2}$$

where τ is held constant, and $\widehat{y}\left(r_i, \tau\right)$ refers to the approximate yield function evaluated at the grid point i. δ is the constant width of the interval between grid points. The maximum errors are reported over the range of $r \in [0.03, 0.20]$. There were large approximation errors for extremely small values of r.

Table A-7: Maximum Absolute Error in the Estimates of $\partial y / \partial r$ for the CIR Case.

Panel A: $\kappa = 0.22$, $\theta = 0.085$, $\sigma = 0.023$, and $\lambda_0 = -0.02$.

	$N = \#$ of Grid Points			
Maturity	$N = 19$	$N = 38$	$N = 76$	$N = 152$
1-month	0.00139	0.00139	0.00139	0.00139
3-months	0.00138	0.00138	0.00138	0.00138

Panel B: $\kappa = 0.86$, $\theta = 0.085$, $\sigma = 0.047$, and $\lambda_0 = -0.02$.

	$N = \#$ of Grid Points			
Maturity	$N = 19$	$N = 38$	$N = 76$	$N = 152$
1-month	0.00143	0.00143	0.00143	0.00143
3-months	0.00149	0.00149	0.00149	0.00149

Panel C: $\kappa = 1.72$, $\theta = 0.085$, $\sigma = 0.066$, and $\lambda_0 = -0.02$.

	$N = \#$ of Grid Points			
Maturity	$N = 19$	$N = 38$	$N = 76$	$N = 152$
1-month	0.00147	0.00148	0.00148	0.00148
3-months	0.00158	0.00159	0.00159	0.00159

The approximation to the partial derivative of the yield function with respect to r is calculated by interpolating the finite-difference solution using, 720 time points and 240 space points, over the range of r from $[0.01, 0.20]$ using cubic splines. The interpolated function is then evaluated at a grid of size N, and the first derivative is approximated using the formula

$$\frac{\partial \widehat{y}\left(r_i, \tau\right)}{\partial r} \approx \frac{\widehat{y}\left(r_{i+1}, \tau\right) - \widehat{y}\left(r_{i-1}, \tau\right)}{2\delta}$$

where τ is held constant, and $\widehat{y}\left(r_i, \tau\right)$ refers to the approximate yield function evaluated at the grid point i. δ is the constant width of the interval between grid points. The maximum errors are reported over the range of $r \in [0.03, 0.20]$. There were large approximation errors for extremely small values of r.

Table A-8: Maximum Absolute Error in the Estimates of $\partial^2 y/\partial r^2$ for the CIR Case.

Panel A: $\kappa = 0.22$, $\theta = 0.085$, $\sigma = 0.023$, and $\lambda_0 = -0.02$.

	$N = \#$ of Grid Points			
Maturity	$N = 19$	$N = 38$	$N = 76$	$N = 152$
1-month	0.00012	0.00012	0.00011	0.00012
3-months	0.00035	0.00035	0.00033	0.00035

Panel B: $\kappa = 0.86$, $\theta = 0.085$, $\sigma = 0.047$, and $\lambda_0 = -0.02$.

	$N = \#$ of Grid Points			
Maturity	$N = 19$	$N = 38$	$N = 76$	$N = 152$
1-month	0.00049	0.00049	0.00046	0.00049
3-months	0.00123	0.00122	0.00116	0.00122

Panel C: $\kappa = 1.72$, $\theta = 0.085$, $\sigma = 0.066$, and $\lambda_0 = -0.02$.

	$N = \#$ of Grid Points			
Maturity	$N = 19$	$N = 38$	$N = 76$	$N = 152$
1-month	0.00089	0.00088	0.00083	0.00088
3-months	0.00187	0.00186	0.00177	0.00186

The approximation to the partial derivative of the yield function with respect to r is calculated by interpolating the finite-difference solution using, 720 time points and 240 space points, over the range of r from $[0.01, 0.20]$ using cubic splines. The interpolated function is then evaluated at a grid of size N, and the second derivative is approximated using the formula

$$\frac{\partial^2 \widehat{y}(r_i, \tau)}{\partial r^2} \approx \frac{\widehat{y}(r_{i+1}, \tau) - 2\widehat{y}(r_i, \tau) + \widehat{y}(r_{i-1}, \tau)}{\delta^2}$$

where τ is held constant, and $\widehat{y}(r_i, \tau)$ refers to the approximate yield function evaluated at the grid point i. δ is the constant width of the interval between grid points. The maximum errors are reported over the range of $r \in [0.03, 0.20]$. There were large approximation errors for extremely small values of r.

Table A-9: Maximum Absolute Error in the Estimates of the Drift and Diffusion Functions in the Vasicek Case Using the Finite-Difference Approach.

Panel A: $\kappa = 0.22$, $\theta = 0.085$, $\sigma = 0.023$, and $\lambda_0 = -0.02$.

	1-month Proxy	3-month Proxy
Drift Function	0.00003	0.00003
Diffusion Function	0.00003	0.00003

Panel B: $\kappa = 0.86$, $\theta = 0.085$, $\sigma = 0.047$, and $\lambda_0 = -0.02$.

	1-month Proxy	3-month Proxy
Drift Function	0.00012	0.00042
Diffusion Function	0.00007	0.00007

Panel C: $\kappa = 1.72$, $\theta = 0.085$, $\sigma = 0.066$, and $\lambda_0 = -0.02$.

	1-month Proxy	3-month Proxy
Drift Function	0.00057	0.00164
Diffusion Function	0.00010	0.00056

The approximation is on the interval $r \in [0.03, 0.18]$. The finite-difference algorithm is a Crank-Nicholson scheme with 720 time points and 240 space points. The numerical derivatives are constructed using a grid of size $N = 76$ applied to the interpolated yield function.

Table A-10: Maximum Absolute Error in the Estimates of the Drift and Diffusion Functions in the CIR Case Using the Finite-Difference Approach.

Panel A: $\kappa = 0.22$, $\theta = 0.085$, $\sigma = 0.078$, and $\lambda_1 = -0.235$.

	1-month Proxy	3-month Proxy
Drift Function	0.00003	0.00002
Diffusion Function	0.00002	0.00002

Panel B: $\kappa = 0.86$, $\theta = 0.085$, $\sigma = 0.157$, and $\lambda_1 = -0.235$.

	1-month Proxy	3-month Proxy
Drift Function	0.00010	0.00009
Diffusion Function	0.00005	0.00006

Panel C: $\kappa = 1.72$, $\theta = 0.085$, $\sigma = 0.221$, and $\lambda_1 = -0.235$.

	1-month Proxy	3-month Proxy
Drift Function	0.00019	0.00018
Diffusion Function	0.00008	0.00009

The approximation is on the interval $r \in [0.03, 0.18]$. The finite-difference algorithm is a Crank-Nicholson scheme with 720 time points and 240 space points. The numerical derivatives are constructed using a grid of size $N = 76$ applied to the interpolated yield function.

Appendix B: Inverting the Longstaff-Schwartz Model.

In the LS model, the six parameters are recovered from the six moment conditions by reducing the six nonlinear equations to three pairs of nonlinear equations. The first step in this process is to re-write (66) and (68) as the following "linear" system

$$\begin{bmatrix} 1 & 1 \\ \alpha^2 & \beta^2 \end{bmatrix} \begin{bmatrix} \frac{\alpha^2\gamma}{2\delta^2} \\ \frac{\beta^2\eta}{2\xi^2} \end{bmatrix} = \begin{bmatrix} \mathrm{Var}(r) \\ \mathrm{Var}(V) \end{bmatrix}.$$

Solving (and re-arranging) yields

$$\frac{\gamma}{\delta^2} = \frac{2\beta^2\mathrm{Var}(r) - 2\mathrm{Var}(V)}{\alpha^2\left(\beta^2 - \alpha^2\right)} \tag{78}$$

and

$$\frac{\eta}{\xi^2} = \frac{2\mathrm{Var}(V) - 2\alpha^2\mathrm{Var}(r)}{\beta^2\left(\beta^2 - \alpha^2\right)}. \tag{79}$$

The two expected return conditions, (65) and (67), can be written as

$$\begin{bmatrix} \alpha & \beta \\ \alpha^2 & \beta^2 \end{bmatrix} \begin{bmatrix} \frac{\gamma}{\delta} \\ \frac{\eta}{\xi} \end{bmatrix} = \begin{bmatrix} E(r) \\ E(V) \end{bmatrix}.$$

Solving (and re-arranging) yields

$$\frac{\gamma}{\delta} = \frac{\beta E(r) - E(V)}{\alpha(\beta - \alpha)} \tag{80}$$

and

$$\frac{\eta}{\xi} = \frac{E(V) - \alpha E(r)}{\beta(\beta - \alpha)}. \tag{81}$$

Taking the ratio of (80) to (78) and (81) to (79) produces the following expressions for δ and ξ as functions of the moments of r and V and the parameters α and β:

$$\delta(\alpha, \beta) = \frac{\alpha(\alpha + \beta)(\beta E(R) - E(V))}{2\beta^2\mathrm{Var}(r) - 2\mathrm{Var}(V)}$$

and

$$\xi(\alpha, \beta) = \frac{\beta(\alpha + \beta)(E(V) - \alpha E(r))}{2\mathrm{Var}(V) - 2\alpha^2\mathrm{Var}(r)}.$$

The two equation system for α and β (the key to the entire inversion process) follows by substituting $\delta(\alpha, \beta)$ and $\xi(\alpha, \beta)$ into (69) and (70) and simplifying to get

$$\frac{\beta^2 \text{Var}(r) - \text{Var}(V)}{\beta^2 - \alpha^2} \exp(-\delta(\alpha, \beta)\Delta) + \frac{\text{Var}(V) - \alpha^2 \text{Var}(r)}{\beta^2 - \alpha^2} \exp(-\xi(\alpha, \beta)\Delta) - \text{Cov}[r_{t+\Delta}, r_t] = 0 \tag{82}$$

and

$$\frac{\alpha^2 \beta^2 \text{Var}(r) - \alpha^2 \text{Var}(V)}{\beta^2 - \alpha^2} \exp(-\delta(\alpha, \beta)\Delta) + \frac{\beta^2 \text{Var}(V) - \alpha^2 \beta^2 \text{Var}(r)}{\beta^2 - \alpha^2} \exp(-\xi(\alpha, \beta)\Delta) - \text{Cov}[V_{t+\Delta}, V_t] = 0. \tag{83}$$

Given α, β, δ, and ξ, the remaining two parameters can be solved as

$$\gamma(\alpha, \beta) = \frac{\beta E(r) - E(V)}{\alpha(\beta - \alpha)} \cdot \delta(\alpha, \beta) \tag{84}$$

and

$$\eta(\alpha, \beta) = \frac{E(V) - \alpha E(r)}{\beta(\beta - \alpha)} \cdot \xi(\alpha, \beta). \tag{85}$$

References

[1] Aït-Sahalia, Y., 1996a, "Nonparametric Pricing of Interest Rate Derivative Securities," *Econometrica* 64, pages 527-560.

[2] Aït-Sahalia, Y., 1996b, "Testing Continuous-Time Models of the Spot Interest Rate," *The Review of Financial Studies* 9, pages 385-426.

[3] Andersen, T. G., and J. Lund, 1997, "Estimating Continuous Time Stochastic Volatility Models of the Short-Term Interest Rate," *Journal of Econometrics* 77, pages 343-377.

[4] Boudoukh, J., M. Richardson, R. Stanton, and R. F. Whitelaw, 1998, "The Stochastic Behavior of Interest Rates: Implications from a Multifactor, Nonlinear, Continuous-Time Model," Manuscript.

[5] Brennan, M. J., and E. S. Schwartz, 1979, "A Continuous-Time Approach to the Pricing of Bonds," *Journal of Banking and Finance* 3, pages 133-155.

[6] Brown, R., and Schaefer, 1993, "Interest Rate Volatility and the Shape of the Term Structure," *Philosophical transactions of the Royal Society: Physical Sciences and Engineering* 347, pages 449-598.

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

[8] Chapman, D. A., and N. D. Pearson, 1998, "Is the Short Rate Drift Actually Nonlinear?" Manuscript.

[9] Conley, T. G., L. P. Hansen, E. G. J. Luttmer, and J. A. Scheinkman, 1997, "Short-Term Interest Rates as Subordinated Diffusions," *The Review of Financial Studies* 10, pages 525-577.

[10] Cox, J. C., J. E. Ingersoll, Jr., and S. A. Ross, 1985, "A Theory of the Term Structure of Interest Rates," *Econometrica* 53, pages 321-346.

[11] Duffie, D., 1996, *Dynamic Asset Pricing Theory,* 2^{nd} *Edition.* (Princeton: Princeton University Press).

[12] Duffie, D., and R. Kan, 1996, "A Yield Factor Model of Interest Rates," *Mathematical Finance* 6, pages 379-406.

[13] Goldstein, R., and F. Zapatero, 1996, "General Equilibrium with Constant Relative Risk Aversion and Vasicek Interest Rates," *Mathematical Finance* 6, page 331-340.

[14] Hamilton, J. D., 1996, "The Daily Market for Federal Funds," *Journal of Political Economy* 104, pages 26-56.

[15] Ingersoll, J. E., Jr., 1987, *Theory of Financial Decision Making.* (Totowa, NJ: Rowman & Littlefield).

[16] Karlin, S., and H. M. Taylor, 1981, *A Second Course in Stochastic Processes.* (San Diego: Academic Press, Inc.)

[17] Knez, P. J., R. Litterman, and J. Scheinkman, 1994, "Explorations Into Factors Explaining Money Market Returns," *The Journal of Finance* 49, pages 1861-1882.

[18] Longstaff, F. A., and E. S. Schwartz, 1992, "Interest Rate Volatility and the Term Structure: A Two-Factor General Equilibrium Model," *The Journal of Finance* 47, pages 1259-1282

[19] 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, pages 1279-1304.

[20] Press, W. H., S. A. Teukovsky, W. T. Vetterling, and B. P. Flannery, 1992, *Numerical Recipes in C,* 2^{nd} *Edition.* (New York: Cambridge University Press).

[21] Stanton, R., 1997, "A Nonparametric Model of Term Structure Dynamics and the Market Price of Interest Rate Risk," *The Journal of Finance* 52, pages 1973-2002.

[22] Vasicek, O. A., 1977, "An Equilibrium Characterization of the Term Structure," *Journal of Financial Economics* 5, pages 177-188.

[23] Wilmot, P., S. Howison, and J. Dewynne, 1995, *The Mathematics of Financial Derivatives.* (Cambridge; Cambridge University Press).