Exact Gaussian Estimation of Continuous Time Models of The Term Structure of Interest Rates*

Peter C.B. Phillips† and Jun Yu‡

December 31, 2000

*We thank The Center for Research in Security Prices (CRSP), Graduate School of Business, University of Chicago, for providing the US Treasury bill data, Andrew Karolyi for confirming the data with us, and Ben Nowman for providing the UK sterling interbank rate and for his comments. The second author gratefully acknowledges financial support from the University of Auckland Research Committee and the University of Auckland Business School.

†Cowles Foundation for Research in Economics, Yale University & Department of Economics, The University of Auckland, Auckland, New Zealand, email: peter.phillips@yale.edu

‡Department of Economics, The University of Auckland, Auckland, New Zealand, email: j.yu@auckland.ac.nz

Abstract

This paper proposes an exact Gaussian estimator for nonlinear continuous time models of the term structure of interest rates. The approach is based on a stopping time argument that produces a normalizing transformation facilitating the use of a Gaussian likelihood. A Monte Carlo study shows that the finite sample performance of the proposed procedure offers an improvement over the discrete approximation method proposed by Nowman (1997). An empirical application to U.S. and British interest rates is given.

JEL: C14, C22, G12

1 Introduction

Continuous time models of the interest rate are now frequently formulated in terms of nonlinear stochastic differential equations. Econometric estimation of such models has been intensively studied in the recent literature. Broadly speaking, three methods have been proposed to estimate the parameters of such systems. The first method employs a discrete time approximation to the continuous system and estimation of the discrete time model is conducted by nonlinear regression or maximum likelihood. This is the approach used by Chan, Karolyi, Longstaff, and Sanders (1992) (CKLS, hereafter) and Nowman (1997). The second method exploits the martingale property of the diffusion process and approximates the transition function, the likelihood or the moment conditions. Some of these approximations are based on simulations (e.g. Duffie and Singleton, 1993), some are based on numerical approximations (such as Lo, 1988), while others are based on closed-form approximations (such as Ait-Sahalia, 1999, 2000)). A third approach seeks to estimate the drift and diffusion functions directly by nonparametric kernel techniques (Florens-Zmirou, 1993, and Bandi and Phillips, 1999).

The approximation scheme used in the discretization method proposed by CKLS is based on the Euler method. In comparison to the continuous time model, the discrete time model is relatively easy to estimate. As a linear approximation, however, the

Euler method introduces a discretization bias since it ignores the internal dynamics which can be excessively erratic. It is well known that ignoring such a bias can result in inconsistent estimators (see Melino (1994)). The discrete approximation method proposed by Nowman (1997) presents the first application of Gaussian methods of estimation for nonlinear continuous time models. It is based on the Gaussian estimation method developed by Bergstrom (1983, 1984, 1985, 1986, 1990) for linear systems. Since the general form of continuous time models of interest rates involve conditional heteroscedasticity, however, the process is not Gaussian. So, in order to use Gaussian estimation, Nowman (1997) assumes the volatility of the interest rate is constant over each unit observation period, thereby facilitating the construction of a discrete time version of the model. In essence, this procedure uses the Euler method to approximate the diffusion term over the unit interval. In so doing, the method replaces a non-Gaussian process by an approximate Gaussian one. Since only the diffusion term is approximated, the Nowman method has the advantage of reducing some of the aggregation bias relative to full discretization. Strictly speaking, the Nowman procedure is a form of quasi maximum likelihood. While simulations or approximations can overcome the difficulties involved in calculating the likelihood function or the moments of the diffusion process, it is in generaly difficult to gauge the accuracy of the approximations.

The present paper proposes a different approach to forming a discrete time model. It has the interesting feature that it produces an exact Gaussian approach to estimating a non-Gaussian diffusion processes. It is related to the Nowman discrete approximation method in the sense that a discrete model is derived and used for estimation. However, we use a very different mechanism to obtain an exact discrete model with Gaussian errors and the proposed estimator is, in fact, an exact Gaussian estimator. The procedure exploits the martingale property of the process driving the diffusion and uses a time-change technique as a normalizing transformation to convert this process to a Gaussian one.

The paper is organized as follows. Section I reviews various continuous time models of term structure of interest rates and Nowman's estimation method. Section II develops the alternate approach of the present paper. Section III reports a simulation study of the performance of the proposed approach in comparision with the Nowman method. Section IV illustrates the procedure in an empirical application. Section V concludes.

2 Continuous Time Interest Rate Models

Consider an interest rate diffusion process $\{r(t) : t \geq 0\}$ generated by

$$dr(t) = (\alpha + \beta r(t))dt + \sigma r^{\gamma}(t)dB(t), \tag{2.1}$$

where $B(t)$ is a standard Brownian motion defined on the probability space $(\Omega, \Im^B, (\Im_t^B)_{t\geq 0}, P)$, and α, β, σ, and γ are unknown system parameters.[1] In this model, $r(t)$ mean-reverts towards the unconditional mean $-\frac{\alpha}{\beta}$, $-\beta$ measures the speed of the reversion, and γ determines the sensitivity of the variance with respect to the level of $r(t)$. Assume the data $r(t)$ is recorded discretely at $(0, \Delta, 2\Delta, \cdots, T\Delta)$ in the time interval $[0, T\Delta]$, where Δ is a discrete time step in a sequence of observations $r(t)$. If $r(t)$ is the annualized interest rate observed monthly (weekly or daily), then $\Delta = 1/12$ (1/52 or 1/250).

The specification of equation (2.1) allows a possible nonlinear diffusion term but only a linear drift.[2] Equation (2.1) nests some well-known term structure models of

[1]Although we focus on the 1-factor model in this paper, there are many multi-factor models that have been studied in the term structure literature. Examples include Babbs and Nowman (1999), Brennan and Schwartz (1979), Chen and Scott (1992), Duffie and Kan (1996), and Longstaff and Schwartz (1992).

[2]The specification of a linear drift has been criticized in the recent literature. For example, using a nonparametric test, Ait-Sahalia (1996) rejected all parametric models and argues that the linearity in the drift is a major source of misspecification. Stanton (1997) proposed nonparametric estimators of the drift and diffusion functions and found that the estimated drift is highly nonlinear, especially when the interest rate is more than 14%. However, a Monte Carlo study performed by Chapman and Pearson (2000) indicates poor finite sample properties of the nonparametric estimators of Ait-

interest rates. Their specifications and the parameter restrictions are summarized in Table 1.

Except for a few special cases, maximum likelihood is difficult to use since the likelihood function does not have a closed form expression. Also, in almost all practical contexts the diffusion process is not Gaussian. For example, Cox, Ingersoll and Ross (1985) show that when $\gamma = 0.5$ the distribution of $r(t+1)$ conditional on $r(t)$ is non-central $\chi^2[2cr(t), 2q+2, 2\lambda(t)]$, where $c = -2\beta/(\sigma^2(1-e^{\beta})), \lambda(t) = cr(t)e^{\beta}, q = 2\alpha/\sigma^2 - 1$, and the second and third arguments are the degrees of freedom and non-centrality parameters, respectively.

When $\gamma > 0$, the conditional volatility of the model increases with the level of the interest rate. This is the so-called "level effect". Since the conditional variance is not constant for $\gamma \neq 0$, the Gaussian estimation method proposed by Bergstrom (1983, 1985, 1986, 1990) is not directly applicable. To use Bergstrom's procedure, Nowman (1997) assumes that the conditional volatility remains unchanged over the unit intervals, $[s\Delta, (s+1)\Delta)$, $s = 0, 1, ...,$ and then approximates the stochastic equation (2.1) over these intervals by the equation:

$$dr(t) = (\alpha + \beta r(t))dt + \sigma r^{\gamma}(s\Delta)dB(t), \quad s\Delta \leq t < (s+1)\Delta. \tag{2.2}$$

The corresponding exact discrete model of (2.2) then has the form (e.g., Bergstrom, 1984)

$$r(t) = e^{\Delta\beta}r(t-\Delta) + \frac{\alpha}{\beta}(e^{\Delta\beta} - 1) + \eta(t), \tag{2.3}$$

where the conditional distribution $\eta(t)|\Im^B_{t-1} \sim N(0, \frac{\sigma^2}{2\beta}(e^{2\Delta\beta} - 1)(r^{2\gamma}(t-1)))$. With this approximation, the Gaussian method can be used to estimate equation (2.3).

The Nowman procedure can be understood as using the Euler method to approximate the diffusion term over the unit interval. Compared with the discretization

Sahalia (1996) and Stanton (1997). Pritsker (1998) found that the Ait-Sahalia (1996) test rejects the true model too often. Some other recent work by Bandi and Phillips (2000) proposed nonparametric estimators of the drift and diffusion that are applicable in nonstationary cases.

method where the Euler method is applied to both the drift and diffusion terms in the diffusion process, the Nowman's method can be expected to reduce some of the temporal aggregation bias. Strictly speaking, however, the method is a form of quasi-maximum method since (2.3) is not the true discrete model corresponding to equation (2.1) but is merely a conditional Gaussian approximation.

3 Exact Gaussian Estimation

In this section an exact Gaussian method is developed to estimate the equation (2.1). The approach is based on the idea that any continuous time martingale can be written as a Brownian motion after a suitable time change. In particular, by the Dambis, Dubins-Schwarz theorem (hereafter DDB theorm) - see Revuz and Yor (1999) - we have the following result which gives a normalizing transformation for an arbitrary continuous martingale.

Lemma (DDB Theorm) 3.1 *Let M be a $(\Im_t, P)$-continuous local martingale vanishing at 0 with quadratic variation process $[M]_t$ such that $[M]_\infty = \infty$. Set*

$$T_t = \inf\{s | [M]_s > t\}. \tag{3.4}$$

Then, $B_t = M_{T_t}$ is a $(\Im_{T_t})$-Brownian motion and $M_t = B_{[M]_t}$.

The process B_t is referred to as the DDB Brownian motion of M. According to this result, when we adjust from chronological time in the local martingale M to time T_t we transform the process to a Brownian motion. As we move along the new path in the resulting Gaussian process, sampling speed needs to be varied in order to accomplish the transformation. But this is something that can be done when we have finely spaced data. The required time changes are given by equation (3.4), so they depend on the quadratic variation of the process M_t. Since this process is path dependent, the time adjustment will be made according to the observed path of the process.

We can use the lemma to extract an exact discrete Gaussian model for (2.1). First, note that model (2.1) for $r(t)$ has for any given $r(0)$ the following solution

$$r(t) = [r(0) + \frac{\alpha}{\beta}]e^{\beta t} - \frac{\alpha}{\beta} + \int_0^t e^{\beta(t-s)}\sigma r^{\gamma}(s)dB(s), \tag{3.5}$$

so that we can write for any $h > 0$

$$r(t+h) = \frac{\alpha}{\beta}(e^{\beta h} - 1) + e^{\beta h} r(t) + \int_0^h \sigma e^{\beta(h-\tau)} r^{\gamma}(t+\tau)dB(\tau). \tag{3.6}$$

Let $M(h) = \sigma \int_0^h e^{\beta(h-\tau)} r^{\gamma}(t+\tau)dB(\tau)$. $M(h)$ is a continuous martingale with quadratic variation

$$[M]_h = \sigma^2 \int_0^h e^{2\beta(h-\tau)} r^{2\gamma}(t+\tau)d\tau. \tag{3.7}$$

We now use the time transformation (3.4) in the lemma to construct a DDB Brownian motion to represent the process $M(h)$. To do so, we introduce a sequence of positive numbers $\{h_j\}$ which deliver the required time changes. For any fixed constant $a > 0$, let

$$h_{j+1} = \inf\{s|[M_j]_s \geq a\} = \inf\{s|\sigma^2 \int_0^s e^{2\beta(s-\tau)} r^{2\gamma}(t_j+\tau)d\tau \geq a\}. \tag{3.8}$$

Next, construct a sequence of time points $\{t_j\}$ using the iterations $t_{j+1} = t_j + h_{j+1}$ with t_1 assumed to be 0. Evaluating equation (3.6) at $\{t_j\}$, we have

$$r(t_{j+1}) = \frac{\alpha}{\beta}(e^{\beta h_{j+1}} - 1) + e^{\beta h_{j+1}} r(t_j) + M(h_{j+1}). \tag{3.9}$$

According to the lemma, $M(h_{j+1}) = B(a) \sim N(0, a)$. Hence, equation (3.9) is an exact discrete model with Gaussian disturbances and can be estimated directly by maximum likelihood. Although both (2.3) and (3.9) are exact discrete models, only (3.9) is the exact discrete model with Gaussian disturbances. The time transformed model (3.9) has both theoretical and practical significance. An interesting feature of (3.9) is that the discrete time model does not have equispaced observations. One needs to sample the process more frequently when the level of interest rates, and hence the conditional volatility, is higher. Thus, the sampling process is endogenous. Figures 1 and 2 illustrate how the time transformation varies according to the generating process and the sample path using the two real data sets from Section 5. In both figures the vertical lines represent the sequence of sampling points $\{t_j\}$. The finer they are, the higher the sampling speed is. Obviously the sampling speed varies in both cases. For example, for the US treasury bill rate, we have to sample all the observations available to us when the market experienced high interest rates at the beginning of 1980s but sample much less frequently when the market experienced lower interest rates in 1960s. Also, from equation (3.8) we note that the sampling points $\{t_j\}$ are more sensitive when γ is larger. This is confirmed by Figure 1 and Figure 2 since γ is estimated to be 1.3610 in the US market and 0.2898 in the UK market.

Letting $\theta = (\alpha, \beta, \sigma, \gamma)$ and defining $L(\theta)$ as minus twice the averaged logarithm of the likelihood function of model

$$L(\theta) = \frac{1}{N}\sum_j [2\log a + \frac{(r(t_{j+1}) - \frac{\alpha}{\beta}(e^{\beta h_{j+1}} - 1) - e^{\beta h_{j+1}} r(t_j))^2}{a^2}], \tag{3.10}$$

where N is the number of sample points resluting from the transformation. Minimization of equation (3.10) leads to the ML estimators of θ. It can be seen that in terms of the estimation of α and β the above maximum likelihood procedure is equivalent to least squares, i.e.

$$\min_{\alpha,\beta} \frac{1}{N}\sum_j (r(t_{j+1}) - \frac{\alpha}{\beta}(e^{\beta h_{j+1}} - 1) - e^{\beta h_{j+1}} r(t_j))^2. \tag{3.11}$$

The autorrelation properties of the sequence $\{r(t_j)\}$ are determined by the parameter β. It is well known that the ML estimate of the autorrelation parameter for a sequence that almost has a "unit root" is downward biased (cf Andrews (1993)). Since interest rates, when observed at the daily, weekly and even monthly frequencies, tend to have large autoregressive coefficients, the ML estimate of β has a downward bias which results in upward bias in the estimate of α. On the other hand, simulations we have performed and which will be discussed below show that the Nowman estimates of σ and γ are quite good in finite samples. In consequence, we propose to use the new discrete time model to improve estimation of α and β but make no attempt to improve estimation of σ and γ. To do so we take Nowman's estimates of σ and γ and fix them in our algorithm.

4 Implementation and Simulation

In practice interest rates are observed at discrete, albeit short, time intervals. In consequence, the time-change formula (3.8) is not directly applicable. Instead, we use the discrete time approximation

$$h_{j+1} = \Delta \min\{s | \sum_{i=1}^{s} \sigma^2 e^{2\beta(s-i)\Delta} r^{2\gamma}(t_j + i\Delta) \geq a\}. \tag{4.12}$$

To use the proposed procedure, a value for a must be selected. Asymptotically, the choice of a should not matter as long as a is finite, but the same is not be true in finite samples. If a is chosen too large, then the effective sample size is too small and we cannot collect a sample with enough information. If a is too small, then we lose the opportunity to adjust the sampling interval to transform the process to Gaussianity. For practical implementation, we therefore propose to choose a in a data based fashion to reflect the average volatility in the data. To do so, we select a as the ML estimate, say $\hat{a}$, in the following constant volatility model (ie the Vasicek model)

$$r(t+\Delta) = \frac{\alpha}{\beta}(e^{\Delta\beta} - 1) + e^{\Delta\beta} r(t) + \varepsilon, \tag{4.13}$$

with $\varepsilon \sim N(0, a)$. Thus, a is the unconditional volatility of the error term in (4.13).

Implementation of the proposed method then proceeds as follows: (1) estimate equation (4.13) using the ML method and obtain $\hat{a}$; (2) estimate equation (2.3) using the ML method and obtain $\hat{\alpha}, \hat{\beta}, \hat{\sigma}$ and $\hat{\gamma}$, ie, obtain the Nowman's estimates of model

(2.1); (3) set a, σ, γ as $\hat{a}, \hat{\sigma}, \hat{\gamma}$ respectively and condition on them in the subsequent step; (4) choose initial values of α, β to be the Nowman estimates and perform a numerical optimization on (3.11) with h_{j+1} chosen according to the time change formula (4.12). The numerical solutions of this extremum estimation problem are then the desired estimates.

The objective function (3.11) has no direct analytic expression for its derivatives with respect to β since both the sampling frequency and the total number of sample observations depend on β. Consequently, the numerical optimization is carried out using Powell's conjugate direction algorithm (Powell (1964)).

To evaluate the finite sample performance of our method, we conduct a small Monte Carlo study. Suppose that the interest rate $r(t)$ follows the square-root process

$$dr(t) = (\alpha + \beta r(t))dt + \sigma r^{\gamma}(t)dB(t), \tag{4.14}$$

with $\gamma = 0.5$.

For any given parameter setting, a sample path for the square root diffusion is simulated according to the 2-step method used by Chapman and Pearson (2000). To ensure the validity of our method for the frequencies commonly used in practice, we choose $\Delta = 1/12, 1/52, 1/250$ which correspond to monthly, weekly, and daily frequencies, respectively.

Table 2 shows the parameter settings and the sample size for all three frequencies. The parameter values are close to what would be obtained from empirical applications when a square-root diffusion model is fitted. For example, the parameter setting implies that the long term mean for annualized interest rates is 6.0 percent for all three frequencies. Daily interest rates revert more quickly to the long term mean than weekly and monthly rates. Moreover, we try to choose the sample sizes close to what have been used in actual empirical studies in the literature.

The model is fitted to the simulated sequence by both Nowman's method and the proposed method with γ treated as additional unknown parameter. We also fit the sequence to the Vasicek model in order to obtain the ML estimate of a. We repeat the experiment in 1,000 replications. The means, variances and mean square errors (MSE) of the resulting estimates are displayed in Tables 3-5.

One result emerging from these tables is that Nowman's method provides very good estimates of σ and γ in terms of both bias and MSE. The sample bias for σ is $3\%, 7\%, 1\%$ with monthly, weekly and daily data respectively, and $2\%, 2\%, 1\%$ for γ and hence is negligible. The result justifies the choice of Nowman's procedure to estimate σ and γ. On the other hand, the finite sample performance of Nowman's estimates of α and β method are nowhere near as good. For example, the sample bias for α is $86.7\%, 47.0\%, 63.8\%$ with monthly, weekly and daily data respectively, and $94.3\%, 46.4\%, 53.6\%$ for β. Moreover, the sampling distribution of β is biased downward for all three frequencies. The bias is still substantial even when the sample size is reasonably large. This is consistent with the well known problems with estimation of first-order autoregressive/unit root models, especially when the AR parameter is large. The downward bias for β implies that the sampling distribution of α is biased upward for all three frequencies. This bias is still present in our exact Gaussian estimates. However, it is smaller than that of Nowman's method . For example, our method produces $15\%, 8\%, 16\%$ less bias than the Nowman's method when estimating α with monthly, weekly and daily data, respectively, and $5\%, 6\%, 6\%$ when estimating β. Furthermore, our method appears to be more efficient than Nowman's method. For example, in terms of the MSE, the efficiency gain is $3\%, 7\%, 7\%$ when estimating α with monthly, weekly and daily data respectively, and $7\%, 7\%, 8\%$ when estimating β.

5 Empirical Results

Two series of interest rates are used in the empirical study, including one British rate obtained from *Datastream* and one US rate obtained from the Center for Research in

Security Prices (CRSP).[3] The British rate was used also in Nowman's (1997) study and is the one-month sterling interbank middle rate over the period from 03/1975 to 03/1995 (see Nowman for details). It contains 242 observations. The US rate is the US Treasury bill one-month yield data over the period from 06/1964 to 12/1989. It has 307 observations. The same dataset is used also by CKLS (1992) and Nowman (1997) (see CKLS for details).

In Table 6 we present the ML estimates of the Vasicek model, the Nowman estimates in the CKLS model and our exact Gaussian estimates for the UK interest rate. We also provide asymptotic standard errors of our exact Gaussian estimates.[4] Our method produces estimates that are similar to Nowman's, but leads to smaller estimate of α and a larger estimate of β, consistent with the findings from the Monte Carlo study. The Nowman method provides an estimate of the unconditional mean of 10.20 percent while our method leads to 9.977 percent, with implied estimates of the speed of the reversion by our method of 0.3279, which is smaller than the Nowman estimate of 0.3490.

In Table 7 we present the ML estimates in the Vasicek model, the Nowman estimates in the CKLS model, and our exact Gaussian estimates for the US interest rate. We also provide asymptotic standard errors of our exact Gaussian estimates. In this case, Nowman's estimates are not close to our estimates. Our method results in a larger estimate of α and a smaller estimate of β, contrary to the findings in the Monte Carlo study. The Nowman estimate of the unconditional mean is 7.41 percent while our estimate is 6.33 percent. The implied estimates of the speed of the reversion are 0.51 for our method and 0.3277 for Nowman's method.

[3]Source: CRSP, Center for Research in Security Prices. Graduate School of Business, The University of Chcago. Used with Permission. All right reserved. www.crsp.com.

[4]We should stress that the asymptotic standard errors given are conditional on the Nowman estimates and they may understate the unconditional asymptotic standard errors.

6 Conclusion

This paper gives an exact discrete time Gaussian model of a nonlinear continuous time diffusion. The discrete model is suitable for Gaussian estimation of the term structure of interest rates even when there are nonlinear volatility effects. Monte Carlo simulations show that the finite sample performance of the proposed method compares well with estimates based on the alternate discrete approximation of Nowman (1997). Nowman's method provides very good estimates of the two parameters in the diffusion term, but is less accurate in estimating the parameters of the drift. The new procedure reduces the finite sample bias and improves the finite sample efficiency of Nowman's method in our simulations for all frequencies that are common used in empirical work. In an empirical application of both procedures to British and US interest rates, it is found that the two procedures produce similar estimates for British interest rates but different estimates for US interest rates, where the speed of reversion is estimated to be 60% faster by our procedure.

References

[1] Ait-Sahalia, Yacine. Testing continuous-time models of the spot interest rate. *Review of Financial Studies*, 9:385–426, 1996.

[2] Ait-Sahalia, Yacine. Transition densities for interest rate and other nonlinear diffusions. *Journal of Finance*, 54:1361–1395, 1999.

[3] Ait-Sahalia, Yacine. Maximum likelihood estimation of discretely sampled diffusion: A closed-form approximation approach. Forthcoming, *Econometrica*, 2000.

[4] Andrews, Donald W. K. Exactly median-unbiased estimation of first order autoregressive/unit root models. *Econometrica*, 61:139–166, 1993.

[5] Babbs, Simon H. and K.B. Nowman. Kalman filtering of generalized vasicek term structure models. *Journal of Financial and Quantitative Analysis*, 34:115–130, 1999.

[6] Bandi, F.M. and P.C.B. Phillips. Econometric estimation of diffusion models. Working Paper, Yale University, 1999.

[7] Bergstrom, Albert R. Gaussian estimation of structural parameters in higher order continuous time dynamic models. *Econometrica*, 51:117–152, 1983.

[8] Bergstrom, Albert R. Continuous time stochastic models and issues of aggregation over time. In Z. Griliches and M.D. Intriligator, editors, *Handbook of Econometrics*. Vol. II (Elsevier Science, Amsterdam), 1984.

[9] Bergstrom, Albert R. The estimation of parameters in nonstationary higher-order continuous-time dynamic models. *Econometric Theory*, 1:369–385, 1985.

[10] Bergstrom, Albert R. The estimation of open higher-order continuous time dynamic models with mixed stock and flow data. *Econometric Theory*, 2:350–373, 1986.

[11] Bergstrom, Albert R. *Continuous Time Econometric Modelling*. Oxford University Press, 1990.

[12] Brennan, M. J. and E. S. Schwartz. A continuous time approach to the pricing of bonds. *Journal of Banking and Finance*, 3:133–155, 1979.

[13] Chan, K., Karolyi, F, Longstaff, F., and A. Sanders. A empirical comparison of alternative models of short term interest rates. *Journal of Finance*, 47:1209–1227, 1992.

[14] Chapman, David A. and Neil Pearson. Is the short rate drift actually nonlinear? *Journal of Finance*, 55:355–388, 2000.

[15] Chen, R. R. and L. Scott. Pricing interest rate options in a two-factor cox-ingersoll-ross model of the term structure. *Review of Financial Studies*, 5:613–636, 1992.

[16] Cox, J., Ingersoll, J., and S. Ross. A theory of the term structure of interest rates. *Econometrica*, 53:385–407, 1985.

[17] Duffie, D. and K.J. Singleton. Simulated moments estimation of Markov models of asset prices. *Econometrica*, 61:929–952, 1993.

[18] Duffie, D. and R. Kan. A yield-factor model of interest rates. *Mathematical Finance*, 6:379–406, 1993.

[19] Florens-Zmirou, D. On estimating the diffusion coefficient from discrete observations. *Journal of Applied Probability*, 30:790–804, 1999.

[20] Lo, Andrew W. Maximum likelihood estimation of generalized Itô processes with discretely sampled data. *Econometric Theory*, 4:231–247, 1988.

[21] Longstaff, F. and E. S. Schwartz. Interest rate volatility and the term structure: A two-factor general equilibrium model. *Journal of Finance*, 47:1259–1282, 1992.

[22] Melino, Angelo. Estimation of continuous-time models in finance. In Sims, C.A., editor, *Advances in Econometrics: Sixth World Congress*. Vol. II (Cambridge University Press, Cambridge), 1994.

[23] Nowman, K. Gaussian estimation of single-factor continuous time models of the term structure of interest rates. *Journal of Finance*, 52:1695–1703, 1997.

[24] Powell, M.J.D. An efficient method for finding the minimum of a function several variables without calculating derivatives. *The Computer Journal*, 7:155–162, 1964.

[25] Pritsker, M. Nonparametric density estimation and tests of continuous time interest rate models. *Review of Financial Studies*, 11:449–487, 1998.

[26] Revuz, Daniel and Marc Yor. *Continuous Martingales and Brownian Motion.* Springer-Verlag, 1999.

[27] Stanton, Richard. A nonparametric model of term structure dynamics and the market prices of interest rate risk. *Journal of Finance*, 52:1973–2002, 1997.

Table 1: Alternative One-factor Term Structure Models and Parameter Relationship

	Model	α	β	γ
Merton (1973)	$dr(t) = \alpha dt + \sigma dB$		0	0
Vasicek (1977)	$dr(t) = (\alpha + \beta r(t))dt + \sigma dB$			0
Cox, Ingersoll and Ross (1985)	$dr(t) = (\alpha + \beta r(t))dt + \sigma r^{1/2} dB$			1/2
Dothan (1978)	$dr(t) = \sigma r dB$	0	0	1
Geometric Brownian Motion	$dr(t) = \beta r(t)dt + \sigma r dB$	0		1
Brennan and Schwartz (1980)	$dr(t) = (\alpha + \beta r(t))dt + \sigma r dB$			1
Cox, Ingersoll and Ross (1980)	$dr(t) = \sigma r^{3/2} dB$	0	0	3/2
Constant Elasticity of Variance	$dr(t) = \beta r(t)dt + \sigma r^{\gamma} dB$	0		
CKLS (1992)	$dr(t) = (\alpha + \beta r(t))dt + \sigma r^{\gamma} dB$			

Table 2: Parameter Setting and Sample Size in the Monte Carlo Study

	Monthly	Weekly	Daily
Sample Size	500	1000	2000
α	0.72	3.0	6.0
β	-0.12	-0.5	-1.0
σ	0.6	0.35	0.25

Table 3: Monte Carlo Study Comparing Nowman's Method and Proposed Method for Monthly Data

	Nowman's Method				Our Method			
	α	β	σ	γ	α	β	σ	γ
MEAN	1.344	-0.2332	0.6173	0.4919	1.2330	-0.2275	0.6173	0.4919
VAR	0.5897	0.1713	0.0099	0.0071	0.5732	0.1589	0.0099	0.0071
MSE	0.9791	0.1841	0.0102	0.0071	0.8363	0.1705	0.0102	0.0071

Note: A square-root model with $\alpha = 0.72, \beta = -0.12, \sigma = 0.6, \gamma = 0.5$ is used to simulate 500 monthly observations for each of the 1,000 replications.

Table 4: Monte Carlo Study Comparing Nowman's Method and Proposed Method for Weekly Data

	Nowman's Method				Our Method			
	α	β	σ	γ	α	β	σ	γ
MEAN	4.409	-0.7320	0.3762	0.4925	4.1650	-0.7011	0.3762	0.4925
VAR	4.0663	0.1109	0.0172	0.0357	3.8608	0.1030	0.0172	0.0357
MSE	6.0855	0.1647	0.0179	0.0358	5.2180	0.1435	0.0179	0.0358

Note: A square-root model with $\alpha = 3.0, \beta = -0.5, \sigma = 0.35, \gamma = 0.5$ is used to simulate 1,000 weekly observations for each of the 1,000 replications.

Table 5: Monte Carlo Study Comparing Nowman's Method and Proposed Method for Daily Data

	Nowman's Method				Our Method			
	α	β	σ	γ	α	β	σ	γ
MEAN	9.8250	-1.5360	0.2521	0.4970	8.8440	-1.4750	0.2521	0.4970
VAR	19.1959	0.5219	0.0244	0.0746	17.822	0.4798	0.2521	0.4970
MSE	33.8266	0.8092	0.0244	0.0746	25.910	0.7054	0.0244	0.0746

Note: A square-root model with $\alpha = 6.0, \beta = -1.0, \sigma = 0.25, \gamma = 0.5$ is used to simulate 2,000 daily observations for each of the 1,000 replications.

Table 6: Empirical Study Comparing Nowman's Method and Proposed Method Using UK Short-Term Interest Rates

Model	Estimation Method	α	β	$\sigma^2(a)$	γ
Vasicek	ML	3.8305	-0.3730	0.6767	
CKLS	Nowman	3.5615	-0.3490	2.1111	0.2898
CKLS	Exact Gaussian	3.2715 (0.6583)	-0.3279 (0.0769)	2.1111	0.2898

Note: The data used is the one-month sterling interbank rate from March 1975 to March 1995 (242 observations). The Vasicek model estimated by ML is given by

$$dr(t) = (\alpha + \beta r(t)) + \sigma dB(t),$$

and the CKLS model estimated by Nowman's method and our exact Gaussian method is given by

$$dr(t) = (\alpha + \beta r(t)) + \sigma r^{\gamma}(t) dB(t).$$

Asymptotic standard errors are in brackets.

Table 7: Empirical Study Comparing Nowman's Method and Proposed Method Using US Short-Term Interest Rates

Model	Estimation Method	α	β	$\sigma^2(a)$	γ
Vasicek	ML	4.1889	-0.6672	0.6554	
CKLS	Nowman	2.4272	-0.3277	0.0303	1.3610
CKLS	Exact Gaussian	3.2298	-0.5100	0.0303	1.3610
		(1.8448)	(0.1582)		

Note: The data used is the one-month sterling interbank rate from June 1964 to December 1989 (307 observations). The Vasicek model estimated by ML is given by

$$dr(t) = (\alpha + \beta r(t)) + \sigma dB(t),$$

and the CKLS model estimated by Nowman's method and our proposed Gaussian method is given by

$$dr(t) = (\alpha + \beta r(t)) + \sigma r^{\gamma}(t) dB(t).$$

Asymptotic standard errors are in brackets.

Figure 1: time transformations for the UK interest rate

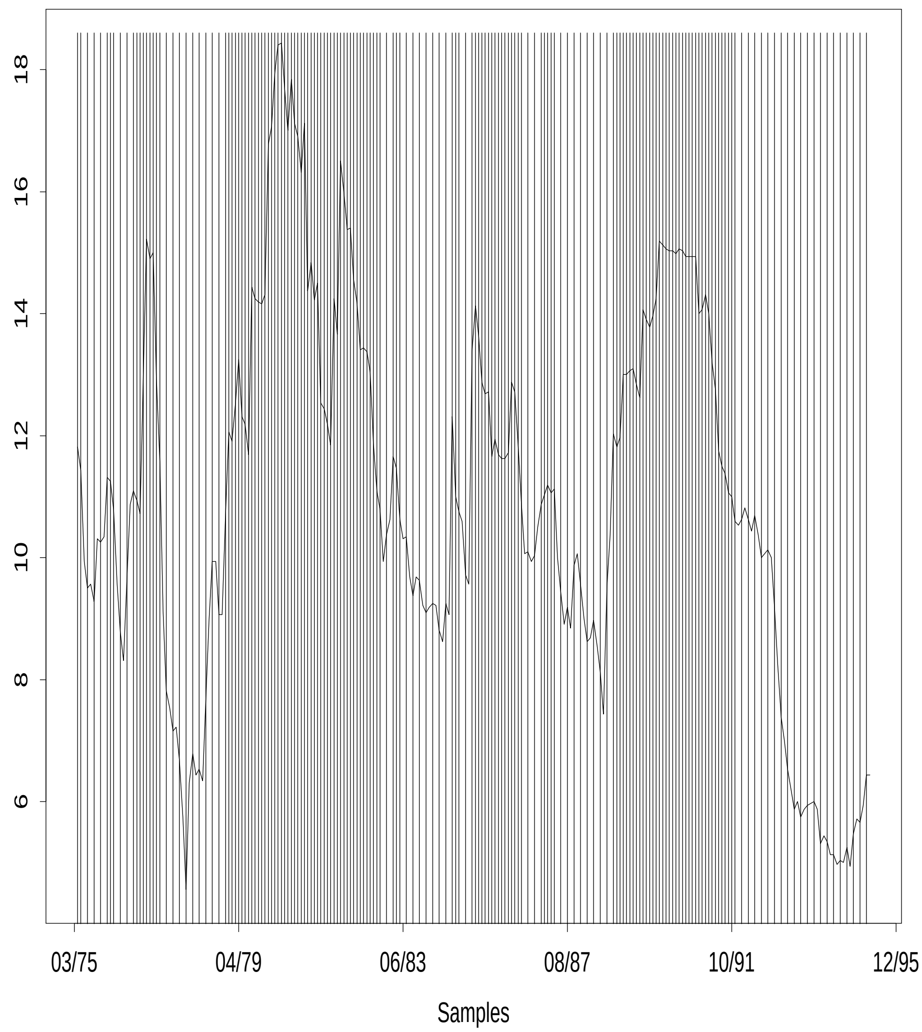

Figure 2: time transformations for the US interest rate