Econometrics Journal (2001), volume **4**, pp. 210–224.

A Gaussian approach for continuous time models of the short-term interest rate

JUN YU[†] AND PETER C. B. PHILLIPS[‡]

[†]*Department of Economics, University of Auckland, Private Bag 92019, Auckland, New Zealand*
E-mail: j.yu@auckland.ac.nz

[‡]*Cowles Foundation for Research in Economics, Yale University, University of Auckland, New Zealand and University of York, UK*
E-mail: peter.phillips@yale.edu

Received: May 2001

Summary This paper proposes a Gaussian estimator for nonlinear continuous time models of the short-term interest rate. 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 US and British interest rates is given.

Keywords: *Gaussian Estimation, Continuous Time Models, Stochastic Differential Equation, Nonlinear Diffusion, Short-term Interest Rate, Normalizing Transformation, Maximum Likelihood, Level Effect.*

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 *et al.* (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), Gallant and Tauchen (1996), Eraker (2001), Elerian *et al.* (2001), Durham and Gallant (2002)), some are based on numerical approximations (such as Lo (1988)), while others are based on closed-form approximations (such as Aït-Sahalia (1999, 2000)). A third approach seeks to estimate the drift and diffusion functions directly by nonparametric kernel techniques (Florens-Zmirou (1993), Jiang and Knight (1997), 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 of interest rates. 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 general difficult to gauge the accuracy of the approximations.

This paper proposes a different approach to forming a discrete time model. It has the interesting feature that it produces a Gaussian approach to estimating 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 discrete observations of the process that satisfy this model are no longer equally spaced. The proposed estimator uses this new discrete time model and is a Gaussian estimator in the sense that it maximizes the Gaussian likelihood. The procedure exploits the martingale property of the process driving the diffusion and uses a time-change technique as a normalizing transformation to convert the process to a Gaussian one. The time-change transformation is itself of empirical interest because it depends on the properties of the process and, upon estimation, reveals the extent of the departure from Gaussianity during the observation period. Our method is also related to the idea of subordination in the sense that Gaussianity can be induced. Many types of subordinators have been used in the literature. Clark (1973) suggests using a Brownian motion as a subordinator while others use more complicated processes or exogenous variables as subordinators (see, for example, Ané and Geman (2000), Barndorff-Nielsen and Shephard (2001), Conley *et al.* (1997), Geman *et al.* (1998), Madan and Senata (1990)). Rather than compounding two processes, however, our method is more on the normalization transformation and the normality is endogenously determined.

The paper is organized as follows. Section 2 reviews various continuous time models of the short-term interest rate and Nowman's estimation method. Section 3 develops the alternative approach of this paper. Section 4 reports a simulation study of the performance of the proposed approach in comparison with the Nowman method. Section 5 illustrates the procedure in an empirical application. Section 6 concludes.

2. CONTINUOUS TIME INTEREST RATE MODELS

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

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

Table 1. Alternative one-factor short-term interest rate 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 *et al.* (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 *et al.* (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)	$dr(t) = (\alpha + \beta r(t))dt + \sigma r^{\gamma} dB$			

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 $(0, \Delta, 2\Delta, \ldots, T\Delta)$ in the time interval $[0, T\Delta]$, where Δ is a discrete time step in a sequence of observations on $r(t)$. 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 models of the short-term interest rate. Their specifications and the parameter restrictions are summarized in Table 1 which can be also found in Jiang and Knight (1997). 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 *et al.* (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

[1] Although we focus on the 1-factor model in this paper, there are many multi-factor models that have been studied in the short-term interest rate literature. Examples include Andersen and Lund (1997), Babbs and Nowman (1999), Brennan and Schwartz (1979), Brenner *et al.* (1996), Chen and Scott (1992), Longstaff and Schwartz (1992), Duffie and Kan (1996). These extensions are not considered in the present paper and the simple but popular model (2.1) is used to illustrate our new approach.

[2] The specification of a linear drift has been criticized in the recent literature. For example, using a nonparametric test, Aït-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 Aït-Sahalia (1996) and Stanton (1997). Pritsker (1998) found that the Aït-Sahalia (1996) test rejects the true model too often. Some other recent work by Bandi and Phillips (1999) proposed nonparametric estimators of the drift and diffusion that are applicable in nonstationary cases.

remains unchanged over the unit intervals, $[s\Delta, (s+1)\Delta)$, $s = 0, 1, \ldots,$ and then approximates the stochastic equation (2.1) over these intervals by the equation

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

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 method where the Euler method is applied to both the drift and diffusion terms in the diffusion process, 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.

Nowman's method is related to the local linearization method proposed by Shoji and Ozaki (1997, 1998) for estimating diffusion processes with a constant diffusion term and a possible nonlinear drift term, that is[3]

$$dr(t) = \mu(r)dt + \sigma dB(t). \tag{2.4}$$

While Nowman approximates the nonlinear diffusion term by a locally linear function, Shoji and Ozaki (1998) approximate the drift term. In essence, Nowman's procedure can be regarded as a local linearization method.

3. GAUSSIAN ESTIMATION USING RANDOM TIME CHANGES

In this section a 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 theorem)—see Revuz and Yor (1999)—we have the following result which gives a normalizing transformation for an arbitrary continuous martingale.

Lemma 3.1 (DDB Theorem). *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*

$$T_t = \inf\{s | [M]_s > t\}. \tag{3.5}$$

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

[3]The local linearization method by Shoji and Ozaki is more generally applicable to diffusion processes $dr(t) = \mu(r)dt + \sigma(r)dB(t)$ with a nonlinear drift and a nonlinear diffusion. This is because a nonlinear transformation can be made to transform the above diffusion process to equation (2.4); see Ozaki (1985) for details about the nonlinear transformation.

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 (3.5), 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:

$$r(t) = \left[r(0) + \frac{\alpha}{\beta}\right]e^{\beta t} - \frac{\alpha}{\beta} + \int_0^t e^{\beta(t-s)}\sigma r^{\gamma}(s)dB(s), \tag{3.6}$$

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

$$r(t+h) = \frac{\alpha}{\beta}(e^{\beta h} - 1) + e^{\beta h}r(t) + \int_0^h \sigma e^{\beta(h-\tau)}r^{\gamma}(t+\tau)dB(\tau). \tag{3.7}$$

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

$$[M]_h = \sigma^2 \int_0^h e^{2\beta(h-\tau)}r^{2\gamma}(t+\tau)d\tau. \tag{3.8}$$

We now use the time transformation (3.5) 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

$$h_{j+1} = \inf\{s|[M_j]_s \geq a\} = \inf\left\{s\,\middle|\,\sigma^2 \int_0^s e^{2\beta(s-\tau)}r^{2\gamma}(t_j+\tau)d\tau \geq a\right\}. \tag{3.9}$$

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 (3.7) at $\{t_j\}$, we have

$$r(t_{j+1}) = \frac{\alpha}{\beta}(e^{\beta h_{j+1}} - 1) + e^{\beta h_{j+1}}r(t_j) + M(h_{j+1}). \tag{3.10}$$

According to the lemma, $M(h_{j+1}) = B(a) \sim N(0, a)$. Hence, equation (3.10) is an exact discrete model with Gaussian disturbances and can be estimated directly by maximum likelihood. Although both (2.3) and (3.10) are exact discrete models, only (3.10) is the exact discrete model with Gaussian disturbances. The time-transformed model (3.10) has both theoretical and practical significance. An interesting feature of (3.10) is that the discrete time model does not have equally spaced 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

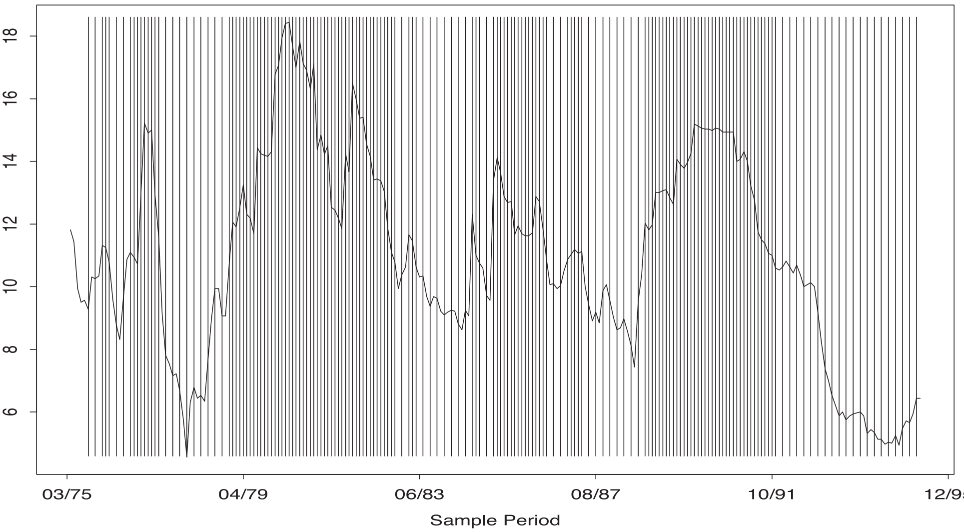

Figure 1. Time transformations for the UK interest rate.

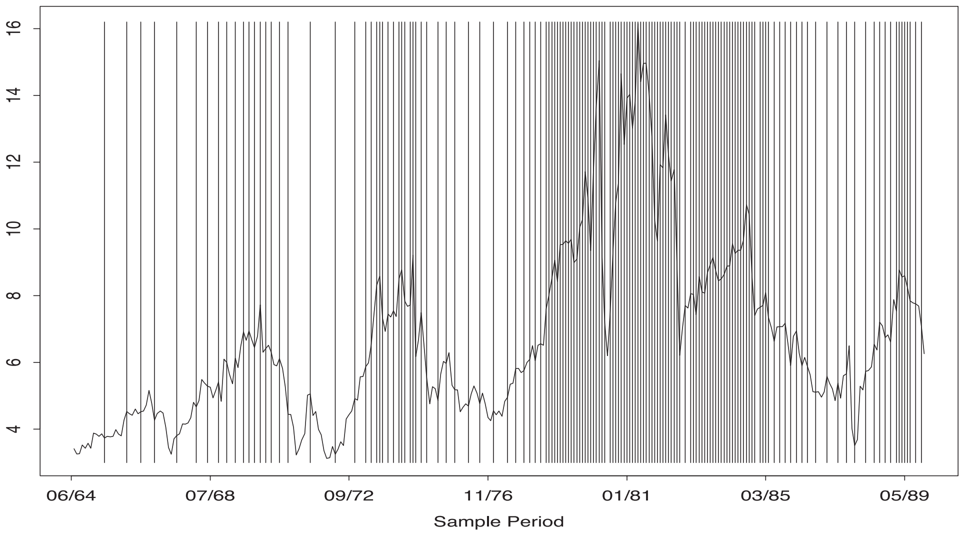

Figure 2. Time transformations for the US interest rate.

experienced high interest rates at the beginning of 1980s but sample much less frequently when the market experienced lower interest rates in the 1960s. Also, from equation (3.9) we note that the sampling points $\{t_j\}$ are more sensitive when γ is larger. This is confirmed by Figures 1 and 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 the model

$$L(\theta) = \frac{1}{N} \sum_j \left[2 \log a + \frac{(r(t_{j+1}) - \frac{\alpha}{\beta}(e^{\beta h_{j+1}} - 1) - e^{\beta h_{j+1}} r(t_j))^2}{a^2} \right], \tag{3.11}$$

where N is the number of sample points resulting from the transformation. Minimization of equation (3.11) 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.

$$\min_{\alpha,\beta} \frac{1}{N} \sum_j \left(r(t_{j+1}) - \frac{\alpha}{\beta}(e^{\beta h_{j+1}} - 1) - e^{\beta h_{j+1}} r(t_j) \right)^2. \tag{3.12}$$

The autocorrelation properties of the sequence $\{r(t_j)\}$ are determined by the parameter β. It is well known that the ML estimate of the autocorrelation 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, although this would be possible by taking into account the further nonlinearities involved in the dependence of the time interval h_j on these parameters. In our implementation, we therefore take Nowman's estimates of σ and γ and fix them in our algorithm.

Subject to a record fine enough to enable choice of the time intervals $h_{j,}$ the estimates of θ are maximum likelihood estimates and standard asymptotic theory applies. For example, the asymptotic distribution is given by

$$\sqrt{N}(\hat{\theta} - \theta) \overset{a}{\sim} N(0, I^{-1}(\theta_0)), \tag{3.13}$$

where I is the Fisher information matrix

$$I(\theta) \equiv \lim_{N \to \infty} -E\left[\frac{1}{N} \frac{\partial^2 L(\theta)}{\partial\theta\partial\theta'} \right], \tag{3.14}$$

and θ_0 represents the true values of the parameters. The information matrix can be estimated by its sample counterpart

$$I_N(\theta) = -\frac{1}{N} \sum_j \frac{\partial^2 l_j(\hat{\theta})}{\partial\theta\partial\theta'}. \tag{3.15}$$

Although we illustrate our approach using the diffusion process with a linear drift term, one can apply our procedure to estimate a process with a nonlinear drift, provided the model can be transformed into a process with a linear drift. In general, the transformation needed is nonlinear and dependent on the specification of the drift term as well as the diffusion term. Consider the following diffusion process:

$$dr(t) = \mu(r)dt + \sigma(r)dB(t). \tag{3.16}$$

Using Ito's formula with $x = G(r)$, equation (3.16) is equivalent to

$$\begin{aligned}
dx(t) &= \frac{\partial G}{\partial r} dr(t) + \frac{1}{2}\frac{\partial^2 G}{\partial r^2}(dr(t))^2 \\
&= \frac{\partial G}{\partial r} dr(t) + \frac{1}{2}\frac{\partial^2 G}{\partial r^2}\sigma^2(r)dt \\
&= \frac{\partial G}{\partial r}\mu(r)dt + \frac{\partial G}{\partial r}\sigma(r)dB(t) + \frac{1}{2}\frac{\partial^2 G}{\partial r^2}\sigma^2(r)dt \\
&= \left(\frac{\partial G}{\partial r}\mu(r) + \frac{1}{2}\frac{\partial^2 G}{\partial r^2}\sigma^2(r)\right)dt + \frac{\partial G}{\partial r}\sigma(r)dB(t).
\end{aligned}$$

If one can choose $G(r)$ to satisfy the second-order differential equation $\frac{\partial G}{\partial r}\mu(r) + \frac{1}{2}\frac{\partial^2 G}{\partial r^2}\sigma^2(r) = \alpha + \beta x$, equation (3.16) is transformed into a model with a linear drift term and hence our procedure can be applied.

4. IMPLEMENTATION AND SIMULATION

In practice, interest rates are observed at discrete, albeit short, time intervals. In consequence, the time-change formula (3.9) is not directly applicable. Instead, we use the discrete time approximation

$$h_{j+1} = \Delta \min\left\{ s \,\middle|\, \sum_{i=1}^{s} \sigma^2 e^{2\beta(s-i)\Delta} r^{2\gamma}(t_j + i\Delta) \geq a \right\}. \tag{4.17}$$

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 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 (i.e. the Vasicek model):

$$r(t + \Delta) = \frac{\alpha}{\beta}(e^{\Delta\beta} - 1) + e^{\Delta\beta} r(t) + \varepsilon, \tag{4.18}$$

with $\varepsilon \sim N(0, a)$. Thus, a is the unconditional volatility of the error term in (4.18).

To demonstrate the effectiveness of the above procedure, we generate 72 000 hourly observations from a square-root process with $\alpha = 12, \beta = -2, \sigma = 0.2$. Instead of using all the 72 000 observations, we pretend that we only have daily observations and hence only a subset of the simulated sequence is used. Based on the daily data and the transformation scheme (4.17), we obtain a sequence of $\{h_j\}$ and time points $\{t_j\}$ which are then used to calculate the residuals in (4.18). Figure 3 shows the Quantile–Quantile plot of the residuals. It can be seen that the normality is well induced.

Implementation of the proposed method then proceeds as follows: (1) estimate equation (4.18) 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 $\hat{\gamma}$, i.e. obtain the Nowman 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

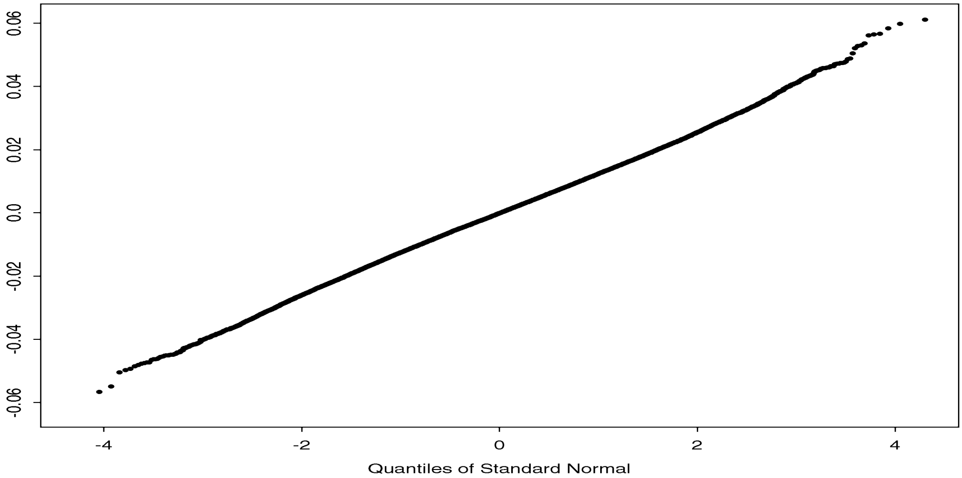

Figure 3. QQ plot of the normalized sequence of a simulated series.

the Nowman estimates and perform a numerical optimization on (3.12) with h_{j+1} chosen according to the time-change formula (4.17). The numerical solutions of this extremum estimation problem are then the desired estimates. This algorithm has the advantage of being simple and convenient for practical implementation. It has the disadvantage that it depends (and conditions) on first-stage estimates of volatility parameters obtained from Nowman's approximate model. The simulations reported below indicate that this procedure works well in practice.

The objective function (3.12) 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

$$dr(t) = (\alpha + \beta r(t))dt + \sigma r^{\gamma}(t)dB(t), \tag{4.19}$$

with $\gamma = 0.5$.

For any given parameter setting, a sample path for the square-root diffusion is simulated according to the two-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 sizes 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 those 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

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
γ	0.5	0.5	0.5

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
BIAS	0.624	−0.1132	0.0173	−0.0081	0.5130	−0.1075	0.0173	−0.0081
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 1000 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
BIAS	1.409	−0.2320	0.0262	−0.0075	1.165	−0.2011	0.0262	−0.0075
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 1000 weekly observations for each of the 1000 replications.

model in order to obtain the ML estimate of a. We repeat the experiment using 1000 replications.[4] 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

[4]We also tried to estimate the model using a simulation-based method: efficient method of moments (EMM) proposed by Gallant and Tauchen (1996). EMM is shown to be asymptotically efficient if the auxiliary model encompasses the data generating process (see Gallant and Tauchen (1996)) its efficiency is close to that of maximum likelihood if the auxiliary model approximates well the data generating process (see Gallant and Long (1997)). Since it is based on simulations, however, computationally it is less efficient. We implemented EMM in the Monte Carlo study with the first parameter setting. The same Gauss computer language as in Pagan *et al.* (1996) was run on a Pentium III PC. After two weeks of calculations, we are still unable to obtain results for 1000 replications. Moreover, we have found that often the objective function for the EMM procedure is numerically unstable, resulting in singularity crashes and/or long convergence time. The same observations have been also documented in Andersen *et al.* (1999) in the context of estimating the discrete time stochastic volatility models. For these obvious reasons, we do not compare our estimates with the EMM estimates.

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
BIAS	3.8250	−0.5360	0.0021	−0.0030	2.8440	−0.4750	0.0021	−0.0030
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 2000 daily observations for each of the 1000 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	3.3283	−0.3389	2.1111	0.2898
		(1.1693)	(0.1122)		

Note: The data used are 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.

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 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 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. However, the reductions of the MSEs are largely due to the decreases in bias.

We should point out that both Nowman's method and our method are computationally highly efficient although they may not be asymptotically most efficient. Using FORTRAN code in an alpha-digital Unix system, for example, it takes Nowman's method and our method no more than 1 and 3 minutes respectively to do all 1000 replications for each of the three parameter settings.

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	−0.6072	0.6554	
CKLS	Nowman	2.4272	−0.3277	0.0303	1.3610
CKLS	Exact Gaussian	2.0069	−0.3330	0.0303	1.3610
		(0.5216)	(0.0677)		

Note: The data used are 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.

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.[5] 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 (Chan *et al.* (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. (We should stress that the asymptotic standard errors given are conditional on the Nowman estimates and they may understate the unconditional asymptotic standard errors.) Our method produces estimates that are similar to Nowman's, but leads to a 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 9.821 percent, with implied estimates of the speed of the reversion by our method of 0.3389, 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 very close to our estimates. Our method results in a smaller estimate of α, once again consistent with the findings from the Monte Carlo study. However, contrary to the findings in the Monte Carlo study, it results in a smaller estimate of β. The Nowman estimate of the unconditional mean is 7.41 percent while our estimate is 6.03 percent. The implied estimates of the speed of the reversion are 0.3330 for our method and 0.3277 for Nowman's method.

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 short-term interest rate even

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

when there are nonlinear volatility effects. Implementation of the model involves the use of non-equispaced observations and the time-change transformation shows how the process needs to be sampled more frequently when conditional volatility is higher. Monte Carlo simulations show that the finite-sample performance of the proposed method compares well with estimates based on the alternative 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 commonly 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 unconditional mean is estimated to be 19% lower using our procedure.

ACKNOWLEDGEMENTS

We thank The Center for Research in Security Prices, 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 data. We acknowledge comments from Karim Abadir, John Knight, Ben Nowman, Joon Park, Neil Shephard and seminar participants at the New Zealand Econometric Study Group Meeting in Auckland, University of Auckland, University of Melbourne, Monash University, National University of Singapore (Department of Finance and Accounting), and 2001 Far Eastern Meeting of the Econometric Society. We also wish to thank Vance Martin for allowing us to use his Gauss code to compute EMM estimates. Yu gratefully acknowledges financial support from the University of Auckland Research Committee and the University of Auckland Business School. Phillips thanks the NSF for support under Grant No. SBR 97-30295. This work was supported by the Marsden Fund under No. 01-VOA-015.

REFERENCES

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

Aït-Sahalia, Y. (1999). Transition densities for interest rate and other nonlinear diffusions. *Journal of Finance 54*, 1361–95.

Aït-Sahalia, Y. (2000). Maximum likelihood estimation of discretely sampled diffusion: A closed-form approximation approach. *Econometrica*, forthcoming.

Andersen, T. G., C. Hyung-Jun and B. E. Sorensen (1999). Efficient method of moments estimation of a stochastic volatility model: A Monte Carlo study. *Journal of Econometrics 91*, 61–87.

Andersen, T. G. and J. Lund (1997). Estimation continuous-time stochastic volatility models of the short-term interest rate. *Journal of Econometrics 77*, 343–77.

Andrews, D. W. K. (1993). Exactly median-unbiased estimation of first order autoregressive/unit root models. *Econometrica 61*, 139–66.

Ané, T. and H. Geman (2000). Order flow, transaction clock and normality of asset returns. *Journal of Finance 55(5)*, 2259–84.

Babbs, S. H. and K. B. Nowman (1999). Kalman filtering of generalized Vasicek term structure models. *Journal of Financial and Quantitative Analysis 34*, 115–30.

Bandi, F. M. and P. C. B. Phillips (1999). *Econometric Estimation of Diffusion Models*, Working paper, Yale University.

Barndorff-Nielsen, O. and N. Shephard (2001). Non-Gaussian Ornstein–Uhlenbeck-based models and some of their uses in financial economics (with discussion). *Journal of the Royal Statistical Society, Series B 63*, 167–241.

Bergstrom, A. R. (1983). Gaussian estimation of structural parameters in higher order continuous time dynamic models. *Econometrica 51*, 117–52.

Bergstrom, A. R. (1984). Continuous time stochastic models and issues of aggregation over time. In Z. Griliches and M. D. Intriligator (eds), *Handbook of Econometrics*, vol. II, Amsterdam: Elsevier.

Bergstrom, A. R. (1985). The estimation of parameters in nonstationary higher-order continuous-time dynamic models. *Econometric Theory 1*, 369–85.

Bergstrom, A. R. (1986). The estimation of open higher-order continuous time dynamic models with mixed stock and flow data. *Econometric Theory 2*, 350–73.

Bergstrom, A. R. (1990). *Continuous Time Econometric Modelling*. Oxford: Oxford University Press.

Brennan, M. J. and E. S. Schwartz (1979). A continuous time approach to the pricing of bonds. *Journal of Banking and Finance 3*, 133–55.

Brennan, M. J. and E. S. Schwartz (1980). Analyzing convertible bonds. *Journal of Financial and Quantitative Analysis 15*, 907–29.

Brenner, R. J., R. H. Harjes and F. K. Kenneth (1996). Another look at models of the short-term interest rate. *Journal of Financial and Quantitative Analysis 31*, 85–107.

Chan, K., F. Karolyi, F. Longstaff and A. Sanders (1992). An empirical comparison of alternative models of short term interest rates. *Journal of Finance 47*, 1209–27.

Chapman, D. A. and N. Pearson (2000). Is the short rate drift actually nonlinear? *Journal of Finance 55*, 355–88.

Chen, R. R. and L. Scott (1992). Pricing interest rate options in a two-factor Cox–Ingersoll–Ross model of the term structure. *Review of Financial Studies 5*, 613–36.

Clark, P. K. (1973). A subordinated stochastic process model with finite variance for speculative prices. *Econometrica 41*, 135–55.

Conley, T. G., L. P. Hansen, E. G. J. Lutterm and J. A. Scheinkman (1997). Short-term interest rate as subordinated diffusion? *Review of Financial Studies 10*, 525–77.

Cox, J., J. Ingersoll and S. Ross (1980). An analysis of variable rate loan contracts. *Journal of Finance 35*, 389–403.

Cox, J., J. Ingersoll and S. Ross (1985). A theory of the term structure of interest rates. *Econometrica 53*, 385–407.

Dothan, U. L. (1978). On the term structure of interest rates. *Journal Financial Economics 6*, 59–69.

Duffie, D. and K. J. Singleton (1993). Simulated moments estimation of Markov models of asset prices. *Econometrica 61*, 929–52.

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

Durham, G. and A. R. Gallant (2002). Numerical techniques for maximum likelihood estimation of continuous-time diffusion processes. *Journal of Business and Economic Statistics*, forthcoming.

Elerian, O., S. Chib and N. Shephard (2001). Likelihood inference for discretely observed non-linear diffusions. *Econometrica 69*, 959–93.

Eraker, B. (2001). MCMC analysis of diffusion models with application to finance. *Journal of Business and Economic Statistics 19*, 177–91.

Florens-Zmirou, D. (1993). On estimating the diffusion coefficient from discrete observations. *Journal of Applied Probability 30*, 790–804.

Gallant, A. R. and J. Long (1997). Estimating stochastic differential equations efficiently by minimum chi-squared. *Biometrika 84*, 125–41.

Gallant, A. R. and G. Tauchen (1996). Which moments to match. *Econometric Theory 12*, 657–81.

Geman, H., D. Madan and M. Yor (1998). *Asset Prices are Brownian Motion: Only in Business Time*, Working paper, University of Maryland.

Jiang, G. and J. Knight (1997). A nonparametric approach to the estimation of diffusion processes—with an application to a short-term interest rate model. *Econometric Theory 123*, 615–45.

Lo, A. W. (1988). Maximum likelihood estimation of generalized Itô processes with discretely sampled data. *Econometric Theory 4*, 231–47.

Longstaff, F. and E. S. Schwartz (1992). Interest rate volatility and the term structure: A two-factor general equilibrium model. *Journal of Finance 47*, 1259–82.

Madan, D. and E. Senata (1990). The variance gamma (V.G.) model from share market returns. *Journal of Business 63*, 511–24.

Melino, A. (1994). Estimation of continuous-time models in finance. In C. A. Sims (ed.), *Advances in Econometrics: Sixth World Congress*, vol. II, Cambridge: Cambridge University Press.

Merton, R. C. (1973). Theory of rational option pricing. *Bell Journal of Economics and Management Science 4*, 141–83.

Nowman, K. B. (1997). Gaussian estimation of single-factor continuous time models of the term structure of interest rates. *Journal of Finance 52*, 1695–703.

Ozaki, T. (1985). Nonlinear time series models and dynamic systems. In E. J. Hannan (ed.), *Handbook of Statistics*, vol. V, pp. 25–83. Amsterdam: North-Holland.

Pagan, A. R., A. D. Hall and V. Martin (1996). Modeling the term structure. In G. S. Maddala and C. R. Rao (eds), *Handbook of Statistics*, vol. 14, pp. 91–118. Amsterdam: Elsevier.

Powell, M. J. D. (1964). An efficient method for finding the minimum of a function several variables without calculating derivatives. *The Computer Journal 7*, 155–62.

Pritsker, M. (1998). Nonparametric density estimation and tests of continuous time interest rate models. *Review of Financial Studies 11*, 449–87.

Revuz, D. and M. Yor (1999). *Continuous Martingales and Brownian Motion*. New York: Springer.

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

Shoji, I. and T. Ozaki (1998). Estimation for nonlinear stochastic differential equations by a local linearization method. *Stochastic Analysis and Applications 16*, 733–752.

Stanton, R. (1997). A nonparametric model of term structure dynamics and the market prices of interest rate risk. *Journal of Finance 52*, 1973–2002.

Vasicek, O. (1977). An equilibrium characterization of the term structure. *Journal of Financial Economics 5*, 177–86.