

A Monte Carlo study for the temporal aggregation problem using one factor continuous time short rate models

Y. C. Lin

Queen Mary College, University of London, UK

Abstract

For most continuous time models formulated in finance, there is no closed form for the likelihood function and estimation of the parameters on the basis of discrete data will be based on an approximation rather than an exact discretization. For example, the Euler method introduces discretization bias because it ignores the internal dynamics that can be excessively erratic. We view the approximation as a difference equation and note that the solution of the continuous time model does not satisfy this difference equation. The effectiveness of the approximation will depend on the rate at which the underlying process is sampled. We investigate how much it matters: can we get significantly different estimates of the same structural parameter when we use say hourly data as compared with using monthly data under given discretization? If yes, then that discretization when applied to a data set in hand, as is done in practice, cannot be said to give robust results. We compare numerically the application of methods by Yu and Phillips (2001), Shoji and Ozaki (1998) and Ait-Sahalia (2002) in the maximum likelihood estimation of the unrestricted interest rate model proposed by Chan et al. (1992). We find that reducing the sampling rate yield large biases in the estimation of the parameters. The Ait-Sahalia method is shown to offer a good approximation and has the advantage of reducing some of the temporal aggregation bias.

Keywords: the discretization method.

1 Introduction

The purpose of the paper is to evaluate the performance of different discretization approximation to a structure continuous time model formulated as



a stochastic differential equation and show that the fact that the discretization approximation depends on the time interval. For most models formulated in continuous time, there is no closed form for the likelihood function and estimation of the parameters of the model on the basis of discrete data needs to be based on an approximate rather than an exact discretization. This has been one of the main issues in formulating and estimating interest rate diffusion models. For example, the discretization method used by Chan et al. [6] (CKLS, hereafter) is based on the Euler method. However, the Euler method introduces discretization bias because it ignores the internal dynamics that can be excessively erratic. It therefore motivates the main emphasis will be on how to use the accurate restrictions to the data (the solution of the stochastic models) to study the econometric properties. Our model is specified as a simple first order stochastic differential equation system but we allow this system to be driven by a constant elasticity of volatility. This model is called the CKLS model in the literature of interest rates. The model considered represented some of the well known and most frequently used models in practice (Merton, 1973; Vaslek, 1977; CIR SD, 1985, the geometric Brownian motion (GBM) process of Black and Scholes, 1973). Our starting point is to view the discretization as a difference equation and to note that the solution of the continuous time model does *not* satisfy this difference equation when the discretization is not exact. This has major implications for estimation. With discrete time sampling, we must simulate a large number of sample paths along which the process is sampled very finely; otherwise, ignoring the difference generally results in inconsistent estimates, unless the discretization happens to be an exact one. This is the time aggregation problem inherent in the dichotomy between the time scale of the continuous time model and that of the observed data. As a result, the effectiveness of the discretization will vary depending on the rate at which the underlying process is sampled. Since the rate at which we sample the data matters when the discretization is approximated, we investigate how much it matters: can we get significantly different estimates of the same structural parameter when we use say hourly data as compared with using say monthly data under given discretization? If the answer is “yes”, then that discretization when applied to a data set in hand, as is done in practice, cannot be said to give robust results. By Monte Carlo simulations and empirical study our aim is to investigate which approximation discretization is most robust to temporal aggregation for the interest rates we usually consider. We compare numerically the application of methods by Yu and Phillips [12], Shoji and Ozaki [11] and Ait-Sahalia [1] in the maximum likelihood estimation of the unrestricted interest rate model proposed by Chan et al. [6]. In this paper we look at the effects of systematic sampling, where we use observations picked out every n periods. For all estimation methods considered in this paper, we find that reducing the sampling rate will yield large biases in the estimation of the parameters. The Ait-Sahalia method is shown to offer a good approximation and has the advantage of reducing some of the temporal aggregation bias.



2 The model and the estimation methods

Following Chan et al. [6] (hereafter CKLS) a one dimensional continuous time specification of the interest rate is considered:

$$dx(t) = (\alpha + \beta x(t))dt + \sigma x(t)^{\gamma} dB(t). \quad (2.1)$$

where $\{x(t), t > 0\}$ is the interest rate variable, α, β, σ , and γ are unknown structural parameters, $\{B_t, t \geq 0\}$ is a standard Brownian motion.

In practice, one could simulate a discretized process with a discretization step Δ . Then one might consider the estimator based on this approximated process. The Euler approximation to (2.1) is given by

$$x(t + \Delta) = x(t) + [\alpha + \beta x(t)]\Delta + u_{t+\Delta}^e \quad (2.2)$$

where $u_{t+\Delta}^e = \sigma x_t^{\gamma} \Delta B(t) = \sigma x_t^{\gamma} (B(t + \Delta) - B(t))$ is the disturbance term.

In principle, we can obtain more and more accurate discretization scheme including further stochastic terms from the stochastic Taylor expansion to the approximation scheme (2.2). This is because these stochastic terms contain additional information about the sample path of the Brownian motion. Despite this possibility, we need to stress the importance of the discretization scheme because neglect errors introduced as a result of time aggregation. Moreover, the approximation scheme (2.2) will not allow us to derive the exact maximum likelihood estimator. The Gaussian estimators will be consistent and asymptotically normal provided $\Delta \rightarrow 0$ or $N \rightarrow \infty$. The size of the approximation error in the discretized process is a function of the length of the discrete time interval. In other words, the approximation error is smaller for shorter time intervals. It is well known that ignoring this bias in the estimation process would give rise to inconsistent estimates of the model's parameters.

On the other hand, (2.1) could be interpreted as representing the integral equation:

$$x(t + \Delta) - x(0) = \int_0^{t+\Delta} [\alpha + \beta x(s)]ds + \int_0^{t+\Delta} \sigma x^{\gamma} dB(s), \quad (t > 0). \quad (2.3)$$

For any initial value $x(0)$, the solution to model (2.1) is thus given by

$$x(t + \Delta) = \frac{\alpha}{\beta} (e^{\beta\Delta} - 1) + e^{\beta\Delta} x(t) + \int_0^{\Delta} e^{\beta(\Delta-\tau)} \sigma x^{\gamma}(t + \tau) dB(\tau). \quad (2.4)$$

Equation (2.4) is the exact discrete model. But, (2.4) cannot be used for estimation because the last term on the right hand involves the level of the process. Along the line of Bergstrom's method [3], Nowman [10] to assume that the volatility of the interest rate change at the beginning of the unit observation period and then remains constant and then apply the Bergstrom's method to estimate the parameters of interest. Let t' be the smallest integer greater than or equal to t , Nowman considers the following SDE:



$$dx(t) = (\alpha + \beta x(t))dt + \sigma x(t' - 1)^r dB(t), t' \leq t < t' + 1. \quad (2.5)$$

Then, following Bergstrom ([3], Theorem 2) the form of the corresponding exact discrete model of (2.1) can be expressed as:

$$x(t+1) = \frac{\alpha}{\beta}(e^\beta - 1) + e^\beta x(t) + \eta_t, t = 1, \dots, T, \quad (2.6)$$

where $\eta_t (t = 1, \dots, T)$ is assumed to follow a normal and satisfies the conditions:

$$E(\eta_s \eta_t) = 0 \text{ } s \neq t, \quad (2.7)$$

$$E(\eta_t^2) = \int_{t-1}^t e^{2(t-\tau)\beta} \sigma^2 x^{2\gamma}(t-1) d\tau = \frac{\sigma^2}{2\beta} (e^{2\beta} - 1) x^{2\gamma}(t-1). \quad (2.8)$$

Comparing to the approximation scheme (2.4), equation (2.6) allows us to use the exact maximum likelihood estimator. This should be help to reduce some of the temporal aggregation bias.

Also along the line of Bergstrom's method [3], Yu and Phillips employ the Dsmbis, Dubins-Schwarz (DDS) theorem and apply the time change formula to cover the residual processes to follow a Normal density. Let the last term in (2.5) be $M(\Delta)$ and it will be a continuous martingale with quadratic variation:

$$[M]_\Delta = \sigma^2 \int_0^\Delta e^{2\beta(\Delta-\tau)} x^{2\gamma}(t+\tau) d\tau.$$

Applying DDS theorem, Yu and Phillips transform $M(\Delta)$ to DDS Brownian motion. This method produces an exact discrete Gaussian model. Comparing to the Nowman's method, which is to equate the observation interval with the unit interval and to consider the exact discrete model on the sequence of the equispaced observations, the Yu and Phillips's method will cause a sequence to be non-equispaced observations.

Shoji and Ozaki [11] use the Ito formula to transform (2.1) as a diffusion process with a nonlinear drift term but a constant diffusion term. They use the local linear technique to approximate that new process. Basically, by the method of Shoji and Ozaki we will have a linear SDE as an approximation to any continuous diffusion, which allows us to derive the exact discretization of the continuous diffusion. The exact representation allows us to use the Bergstrom methods to estimate the parameters of a continuous time systems from discrete data.

Alternative estimation method that efficiently takes account of the time aggregation bias is Ait-Sahalia's method [1]. Comparing to the Shoji and Ozaki method, to simulate the discretized time observations of the process that is the solution of the locally linearization, Ait-Shalia approximates the unknown transition density function by Gaussian. Let $\theta = [\alpha, \beta, \sigma, \gamma]$. Also based on



the Ito formula, Ait-Sahalia considers the new process $y(t)$ is observed at the time points $\{t = i\Delta, 0 \leq i \leq n\}$ for Δ is fixed and defines the increment of $y(t)$ as

$$z(t) = \Delta^{-1/2}(y(t) - y(0) - h(y(0); \theta)), \quad (2.9)$$

where

$$h(y(t); \theta) = \alpha \sigma^{-1} (\sigma(1 - \gamma)y(t))^{\frac{-\gamma}{1-\gamma}} + \beta(1 - \gamma)y(t) - \frac{1}{2}\gamma(1 - \gamma)y^{-1}(t).$$

Then, Ait-Sahalia [1] constructs the random variable $z(t)$ so that its density p_z can be close to a standard normal density. Following Ait-Sahalia [1] one can use the Hermite series expansion up to the J th term to approximate the density function p_z for fixed $\Delta, y(0), \theta$. One then can construct the approximation to the unknown density function for the diffusion process $x(t)$. Ait-Sahalia [1] proves that the density of the random variable $z(t)$ is close to the standard normal density and the approximation is close to the true density function of $x(t)$ when $J \rightarrow \infty$ but the sampling interval Δ remains fixed. Further, more and more accurate approximation to the true density can be obtained provided the order of approximation J gets larger and larger in this scheme. Comparing to the Euler scheme, we note that the sampling interval is not assumed that $\Delta \rightarrow 0$ in order to calculate the parameters explicitly.

In conclusion, when the sampling time interval is sufficiently small, one could expect that the approximation path for (2.1) by the Euler scheme would be close to the true trajectories such that these estimates of the parameters could converge to the true one. However, when the discretization step is observed equidistantly, then the estimates will show different performance depending on the frequency of the data. This is the problem of temporal aggregation in continuous time econometrics. To overcome this problem we would like to derive a discrete time model that will correspond exactly to the underlying continuous time process, in the sense that it generates exactly the same data at discrete points as does the continuous time model. We thus examine this problem of temporal aggregation by discussing several discretization schemes for the stochastic process (2.1) and estimation of the parameters of these discretized models. Basically, we extend the Monte Carlo results in Shoji and Ozaki [11] and Cleur [5]. Both studies only consider the effect of varying the frequency of the data on the estimation of parameters. But, Cleur [7] does not discuss the existence of the exact discretization of the diffusion equation in (2.1) that takes into account time aggregation bias. It is well known that ignoring this bias in the estimation process would give rise to inconsistent estimates of the model's parameters. In our empirical studies we will focus on the strategy of discretizing equation (2.1), which is the correct representation of the diffusion equation (2.1), by solving the stochastic differential equation and then discretizing the solution to this stochastic differential equation. See Nowman [10] and Yu and Phillips [12].



3 Monte Carlo results

Our results are follows.

1. When the frequency of the data is lower, for example using the monthly data, the estimates appear to converge toward a value away from the corresponding true value, particularly, inaccuracy of the estimates of α and β are quite impressive. This asymptotic bias is becoming increasing evident for all methods.
2. For the low frequency data (monthly or weekly data), the estimate of the parameters is biased and the rise in the frequency of the data will lead to an increase in the bias. In the simulation of daily data, the discretization bias is small by the use of Ait-Sahalia's $J=3$ method. This implies that discretization bias may not be very important as expected for Ait-Sahalia's $J=3$ method. This provides some evidences that high frequency data may not be particularly important.
3. In all cases, the biases are serious for empirically relevant of α . We also find that the bias in for the estimator of the parameters α and β will translate into a serious bias for the diffusion parameters σ and γ . Instead of the CIR model, we use the CKLS model to estimate parameters $[\alpha, \beta, \sigma, \gamma]$. We still use the CIR SR type process to generate the hourly data. Our outcome shows the estimates of σ and γ are sensitive to changes in α . For example, using Ait-Sahalia's method, γ is always downward biased and this is consistent with the upward bias in estimated α . In magnitude, the downward bias for γ stays within the 2%. By contrast, σ is substantially upward biased. For the $\alpha=6.0$ case, the percentage bias for σ in the worse case is large than 40% (using Ait-Sahalia's ($J=2$) method). To examine whether the bias of the estimator of γ is affected by other parameters, we show that the bias in the estimator of γ is indeed affected by the parameter α .
4. We compare the MSEs between these three estimation methods by using 36000 simulated data. Ait-Sahalia $J=3$ method appears to be more efficient than other two methods. Hence, for a small sample size, the Ait-Sahalia's method would have efficiency gain because that method will produce a less bias and a less increase in standard errors.
5. After 1000 replications of the estimation procedure, we perform the Kolmogorov-Smirnov test to compare the distribution of the 1000 estimates. Our aim is to examine if these 1000 estimates come from the same distribution for two different sampling frequencies. The null hypothesis is that two samples come from the same distribution. We compare the distributions for hourly / daily, hourly / weekly, and hourly / monthly. Because hourly data can provide much precise confidence intervals, we can investigate the distorted effects by comparing if the sampling distribution of estimates for other sampling frequencies is far from that of estimates for hourly data. Hence, 84.8% for the Yu and Phillips method for hourly / daily data should be compared to one. Obviously, for the Yu and Phillips method and Ait-Sahalia's $J=2$ method, the rejection rates are too large. For example, for hourly / daily data under both methods the empirical rejection rates are



one almost for all cases. This means that the distributions for hourly and monthly data are not the same due to the effects of systematic sampling. We expect the distorted effects should increase as the extent of the data frequency decreases. Hence, for the shoji and Ozaki method and Ait-Sahalia's $J=3$ method, rejection rates are reasonable because the rates increase to one as the data move from hourly /daily to hourly / monthly. However, the test results reflect that fact that for the Yu and Phillips method and Ait-Sahalia's $J=2$ method, the serious distorted effects will occur even using high frequency data and therefore these two methods cannot effectively eliminate the biases.

In addition, as expected in the parameter estimation, our test results also show that, for Ait-Sahalia's $J=3$ method, the distorted effects are not as strong as the Yu and Phillips method. Hence, although Ait-Sahalia's $J=3$ method does not completely eliminate this sort bias it still can be expected relatively powerful on reduction of bias. Although we do not report here, it will be easy to find the reduction is not so obvious when using lower frequency monthly data, and the reduction will be much small the smaller the sample size and the greater the frequency of sampling. Also we show that there is little reduction in bias in using the higher frequency weekly data over and above monthly data, and there could be a substantial reduction in bias from using daily data for Ait-Sahalia's $J=3$ method.

The results by using the Kolmogorov-Smirnov test are consistent with the results using the Mann Whitney rank sum test to examine if the variances for two sampling frequencies are equivalent and the usual F test to examine if the means are equivalent. Also we report the CDF value for the Mann Whitney rank sum test and the usual F test. All of our cases in Tables are one, which means that we reject the null hypothesis that two samples come from the same distribution.

4 Empirical results

Six series of daily and monthly interest rates are used in the empirical study, including the Canada rates, the Germany rates and the US rates. Our goal is to determine the robustness of discretization methods to different sampling intervals. In addition to estimating the models using the entire daily and monthly samples, we also use the sampling scheme in the simulation to augment weekly and monthly observations with daily data. Then, we repeat the estimations using these observations. We estimate the real daily rates and real monthly rate. However, we take every 5 daily observations to be the weekly data and every 4 weekly data to be the monthly data, which forms our augmented monthly data in our Monte Carlo study. By using augmented monthly data, we show the Ait-Sahalia $J=3$ method produces estimates that are similar to the ones by using the real monthly data. But, this is not the case for the Yu and Phillips method. The Yu and Phillips method will produce seriously biased estimates when estimating α . For example, using sampling scheme in our Monte Carlo study, the Yu and Phillips method will provide an estimate of α of 49.4331 for the Germany case, while it is 18.4542 for real monthly data. However, Ait-Sahalia's



J=3 method will provide a small estimate for α of 4.4694, which is more consistent with the estimate, 4.0620, for the real monthly data. For σ and γ , the performances of the Yu and Phillips method for the augmented monthly data and the real monthly data are similar to each other. This is because in the Yu and Phillips method the Nowman's procedure is used to estimate σ and γ . This result shows that Ait-Sahalia's J=3 method has better performances than the Yu and Phillips method, consistent with the findings from the Monte Carlo study. Furthermore, by using augmented monthly data, Ait-Sahalia's J=3 method will produce a small estimate of α and a larger estimate of β comparing to the real monthly data. Also all methods show that there is a more distorted effect in the estimate of α comparing to the estimates of β , once again consistent with the findings from the Monte Carlo study. However, contrary to the findings in the Monte Carlo study, Ait-Sahalia's J=2 method does not result in more distorted effects comparing to the Shoji and Ozaki method and the Yu and Phillips method.

5 Conclusions

In this paper we compare the estimation performances for the continuous time short rate models. We investigate which approximation discretization is most robust to temporal aggregation for the interest rates we usually consider. We compare numerically the application of methods by Yu and Phillips [12], Shoji and Ozaki [11] and Ait-Sahalia [1] in the maximum likelihood estimation of the unrestricted interest rate model proposed by Chan et al. [6]. We find that reducing the sampling rate yield large biases in the estimation of the parameters. The Ait-Sahalia method is shown to offer a good approximation and has the advantage of reducing some of the temporal aggregation bias.

References

- [1] Ait-Sahalia, Y. (2002) Maximum likelihood estimation of discrete sampled diffusion: A closed form approximation approach. *Econometrica* 70, 223-262.
- [2] Bergstrom, A.R. (1983) Gaussian estimation of structure parameters in high – order continuous time dynamic models. *Econometrica* 51, 117-151.
- [3] Bergstrom, A.R. (1984) Continuous time stochastic models and issues of aggregation over time. In: Z. Griliches & M.D. Intriligator (eds), *Handbook of Econometrics*, pp1145-1212. Amsterdam: North-Holland.
- [4] Bergstrom, A.R. (1985) The estimation of parameters in non-stationary higher-order continuous time dynamic models, *Econometric Theory*, 1 369-385.
- [5] Bergstrom, A.R. (1986) The estimation of open higher-order continuous time dynamic models with mixed stock and flow data, *Econometric Theory*, 2 350-373.
- [6] Chan, K.C.G., G. Andrew Karolyi, Francis A Longstaff, and Anthony B Sanders (1992) An empirical comparison of alternative models of the short- term interest rate, *Journal of finance* 47, 1209-1227.



- [7] Cleur, E. M. (2001) Maximum likelihood estimates of a class of one-dimensional stochastic differential equation models from discrete data, *Journal of Time Series Analysis*, Vol. 22, No. 5, 505-516.
- [8] Cox, John C., Jonathan E. Ingersoll, and Stephen A. Ross (1985a), A Theory of the Term Structure of Interest Rates, *Econometrica*, vol. 53, 385-407.
- [9] McCrorie, J.R. (2000a) Deriving the Exact Discrete Analog of a Continuous Time System. *Econometric Theory* 16, 998-1015.
- [10] Nowman, K. (1997) Gaussian estimation of a single – factor continuous time models of the term structure of interest rate, *Journal of Finance*, 52, 1695 -1703.
- [11] Shoji, I. and Ozaki, T. (1997) Comparative study of estimation methods for continuous time stochastic processes, *Journal of Time Series Analysis*, 18, 485-506.
- [12] Yu, J. and P.C.B. Phillips (2001) Gaussian estimation of continuous time models of the short interest rate, *Econometrics Journal*, vol. 4, 2, 210-224.

