10-2010

Asymptotic Distributions of the Least Squares Estimator for Diffusion Processes

Qiankun ZHOU
Singapore Management University, qkzhou@smu.edu.sg

Jun YU
Singapore Management University, yujun@smu.edu.sg

SMU ECONOMICS & STATISTICS
WORKING PAPER SERIES

Asymptotic Distributions of the Least Squares Estimator for Diffusion Processes

Qiankun Zhou & Jun Yu

October 2010

Paper No. 20-2010

Asymptotic Distributions of the Least Squares Estimator for Diffusion Processes[1]

Qiankun Zhou
School of Economics
Singapore Management University

Jun Yu
School of Economics and Sim Kee Boon Institute for Financial Economics
Singapore Management University

January 21, 2010

[1] We gratefully acknowledge financial support from the Ministry of Education AcRF Tier 2 fund under Grant No. T206B4301-RS. Qiankun Zhou, School of Economics, Singapore Management University, 90 Stamford Road, Singapore 178903. Email: qkzhou@smu.edu.sg. Jun Yu, School of Economics and Sim Kee Boon Institute for Financial Economics, Singapore Management University, 90 Stamford Road, Singapore 178903. Email: yujun@smu.edu.sg.

Abstract

The asymptotic distributions of the least squares estimator of the mean reversion parameter (κ) are developed in a general class of diffusion models under three sampling schemes, namely, long-span, in-fill and the combination of long-span and in-fill. The models have an affine structure in the drift function, but allow for nonlinearity in the diffusion function. The limiting distributions are quite different under the alternative sampling schemes. In particular, the in-fill limiting distribution is non-standard and depends on the initial condition and the time span whereas the other two are Gaussian. Moreover, while the other two distributions are discontinuous at $\kappa = 0$, the in-fill distribution is continuous in κ. This property provides an answer to the Bayesian criticism to the unit root asymptotics. Monte Carlo simulations suggest that the in-fill asymptotic distribution provides a more accurate approximation to the finite sample distribution than the other two distributions in empirically realistic settings. The empirical application using the U.S. Federal fund rates highlights the difference in statistical inference based on the alternative asymptotic distributions and suggests strong evidence of a unit root in the data.

Keywords: Vasicek Model, One-factor Model, Mean Reversion, In-fill Asymptotics, Long-span Asymptotics, Unit Root Test

JEL classification: C12, C22, G12

1 Introduction

Consider a stochastic process that is specified in terms of a stochastic differential equation (SDE):

$$dX(t) = \kappa(\mu - X(t))dt + \sigma_X(X(t))dW(t) \tag{1}$$

where $W(t)$ is a standard Brownian motion, μ is the long term mean of $X(t)$ and κ captures the speed of mean reversion of $X(t)$ towards μ if $\kappa > 0$. This one factor model includes as a special case many important models used in financial economics and econometrics.

As an earlier contribution to the continuous time finance literature, Vasicek (1977) proposed to use the Ornstein-Uhlenbeck (OU) diffusion process to describe the evolution of interest rates. In this case, the stochastic process $X(t)$ is given by the following SDE:

$$dX(t) = \kappa(\mu - X(t))dt + \sigma dW(t) \tag{2}$$

where σ is the instantaneous volatility. If $\sigma_X(X(t)) = \sigma\sqrt{X(t)}$, the model is the well-known square root model proposed by Cox, Ingersoll and Ross (1985, CIR here after). Chan et al (1992, CKLS hereafter) proposed a model with $\sigma_X(X(t)) = \sigma X^{\gamma}(t)$. Aït-Sahalia (1996a) introduced a semiparametric model with $\sigma_X(X(t))$ being nonparametrically specified.

In practice, $X(t)$ are directly observable but only at discrete points in time, say $t = 0, \delta, 2\delta, \ldots, n\delta (:= T)$, where n is the sample size, δ the sampling interval, and T the time span of the data. Econometric analysis aims to bring the continuous time model (1) to the discrete data. A recent literature on realized volatility has focused on the diffusion function, based on the assumption that T is fixed (usually set at 1) but $\delta \to 0$; see, for example, Andersen et al (2001) and Barndorff-Nielsen and Shephard (2002). In this paper, we shift this attention to the drift function because the drift function determines the dynamic property and is important for pricing and forecasting.

Many estimation methods have been proposed to estimate parameters in (2) from the discrete observations on $X(t)$. Examples include GMM, maximum likelihood (ML), Gaussian methods, quasi-ML, simulation-based methods such as simulated ML, indirect inference, EMM and Bayesian MCMC, and nonparametric methods. It has been argued that when the model is correctly specified, the preferred choice of estimator should be ML (Durham and Gallant, 2002).

One reason for this choice is that under general regularity conditions, the maximum likelihood estimator (MLE) is asymptotically efficient as $n \to \infty$. The other reason for this choice is that MLE is asymptotically normal as $n \to \infty$, facilitating statistical inferences (Aït-Sahalia, 2002, and Tang and Chen, 2009).

It is now known that ML methods, both the exact and the approximate ML methods, have a serious *finite sample estimation bias* in the mean reversion parameter κ. This bias is related to but much more serious than the finite sample bias in the correlation coefficient estimator (Phillips and Yu, 2005). The bias is shown to have important implications for financial decisions (Phillips and Yu, 2005 and 2009b). Various methods have been introduced to reduce the bias in κ, including the jackknife method (Phillips and Yu, 2005), indirect inference (Phillips and Yu, 2009a) and the bootstrap method (Tang and Chen, 2009). Various authors have obtained analytic forms to approximate the bias under various one-factor models (Tang and Chen, 2009, Yu, 2009b, Ullah, Wang and Yu, 2009).

In addition to the finite sample bias problem, when the true value of κ is small, evidence has been reported on substantial deviations of the finite sample distribution of the MLE of κ from its classical asymptotic distribution developed under the assumption of $n \to \infty$. For example, in the context of Vasicek model with a known μ, Yu (2009a) showed that the finite sample distribution of the MLE of κ and the classical asymptotic distribution behave quite differently. The former is skewed to the right even when n is very large (for example, even when 25,000 daily observations are used!). Similar evidence is documented for other statistics used in the literature. For example, Pritsker (1998) found that the asymptotic distribution of the nonparametric test of Aït-Sahalia (1996b) and that of the kernel density estimator of the marginal distribution do not provide good approximations to their finite sample distributions unless several thousands years of data become available. Similar evidence can be also found in Chapman and Pearson (2001). These pieces of evidence naturally raise the concern of making statistical inferences based on the classical asymptotic theory developed under the assumption that $n \to \infty$.

This problem is related to the unit root literature where it is found that when the root is near unity, the finite sample distribution of the AR coefficient is closer to the Dickey-Fuller distribution than to the asymptotic distribution under the stationary assumption (Ahtola and Tiao, 1984). To address this problem, Phillips (1987b) provided an asymptotic theory for a first-order autoregression with a root near unity. Perron (1991) extended the study by allowing for a more flexible initial condition and a general but finite value for time span (T). Both Phillips and Perron suggested using a SDE model to approximate the discrete time model with a root local to unity and developed the asymptotic theory by assuming $\delta \to 0$ instead of letting $T \to \infty$.[1] Recently, Aït-Sahalia and Park (2009) used the local time approach to develop the asymptotic theory for the kernel estimate of the marginal distribution for diffusions, with the hope to better approximate its finite sample distribution.

The main purpose of the present paper is to develop the asymptotic distribution of the least squares (LS) estimator of κ in Model (2) under three different sampling schemes. The three alternative sampling schemes are listed below:

$$T \to \infty, \quad \delta \ \text{ is fixed, hence } n(:= T/\delta) \to \infty \qquad (A1)$$

$$T \to \infty, \ \delta \to 0 \text{ and hence } n \to \infty \qquad (A2)$$

$$\delta \to 0, \quad T \text{ is fixed and hence } n \to \infty \qquad (A3)$$

where δ is the sampling interval, n the sample size and T the time span.

Scheme ($A1$) assumes that the sampling interval is fixed and the sample size increases as the time span increases. This scheme corresponds to the classical approach to establishing the asymptotic theory. It is widely used in the literature and referred to as the *long-span* asymptotics in the present paper. Tang and Chen (2009) developed the asymptotic distribution of the MLE of κ (and other parameters) in the context of the Vasicek model and the CIR model under this scheme. Aït-Sahalia (2002) made use of this scheme to develop the asymptotic distribution of his approximate MLE. In practical applications in economics and finance, T measures the number of years from which the sample is collected. Typical values for T is not very large (between 1 and 50). In some cases, even if T may be large, a smaller T may be used to avoid possible structural breaks in Model (2). The long-span asymptotic distribution of the MLE of κ is Gaussian for $\kappa > 0$ (stationary) but is skewed for $\kappa = 0$ (unit root). The later result corresponds to the important finding in the unit root literature (Phillips, 1987a). On the other hand, the finite sample distribution is continuous for all values of κ. This observation suggests that the long-span asymptotics fail to provide an accurate approximation to the finite

[1] See Phillips and Magdalinos (2007), Phillips and Han (2008), and Han, Phillips and Sul (2009) for further contributions to bridge the asymptotic distribution of the unit root case and that of the stationary case.

sample distribution when κ is close to 0. The discontinuity in the asymptotic distributions has led to severe criticisms of the use of unit root limit theory in the Bayesian literature; see, for example, Sim (1988) and Sim and Uhlig (1991).

Like Scheme ($A1$), Scheme ($A3$) also allows the sample size to go to infinity. However, this is achieved by decreasing the sampling interval but fixing the time span. In this paper this scheme is referred to as the *in-fill* asymptotics. Under this scheme, Phillips (1987b) and Perron (1991) developed the asymptotic distribution of the LS estimator of the AR coefficient (ϕ) in the discrete time models. Yu (2009a) developed the in-fill asymptotic distribution of κ in the context of the Vasicek model with a known intercept. He found that the in-fill asymptotic distribution is much closer to the finite sample distribution than the long-span asymptotic distribution in the empirically realistic cases. It is important to investigate the robustness of this result under a more general set-up. In practical applications in economics and finance, data are often measured in the annualized term. As a result, $\delta = 1/252$ ($1/52$, $1/12$), corresponding to the daily (weekly, monthly) data. For intra-day data, δ is even smaller than and $1/252$.

Scheme ($A2$) combines both the long-span scheme and the in-fill scheme and is referred to as the *double* asymptotics in this paper. Not surprisingly, this set of assumptions is strongest. Under this scheme, Brown and Hewitt (1975) developed the asymptotic distribution for the MLE of κ in the Vasicek model when μ is known. Bandi and Phillips (2003, 2007) developed the asymptotic distribution for both the non-parametric and the parametric estimators of a continuous time model. Phillips and Yu (2009b) employed this scheme to develop the asymptotic distribution for a two-stage ML estimator.

The present paper contributes to the literature in three aspects. First, the limit theory is developed for the LS estimator of κ in the context of a general class of continuous time models under the three schemes. Under Schemes ($A1$) and ($A2$) the limiting distribution is Gaussian that is independent on the initial condition as well as the parameters in the diffusion function. However, under Scheme ($A3$) the limiting distribution is no-Gaussian and skewed to the right. It depends on both the initial condition and the parameters in the diffusion function. Our results differs from Perron (1991) in that he was primarily concerned about the distribution of the AR coefficient. Our result significantly extend the work of Yu (2009a) in that his model specification is much more restrictive (namely $\mu = 0$ and $\sigma_X(X(t)) = \sigma$). Our asymptotic results under Scheme ($A3$) generalize those of Phillips (1987b) because we allow a general initial condition and a general value for the time span. We extend the asymptotic results of Tang and Chen (2009) in two important ways: (1) the model is more general (the diffusion function is more flexible); (2) different sampling schemes are considered.

Second, we compare the performance of the three alternative distributions. To the best of our knowledge, this is the first time in the literature that the relative performance of all three alternative distributions is examined. Our results suggest that for empirically realistic cases, Schemes ($A1$) and ($A2$) fail to provide accurate approximations to the finite sample distribution, whereas the distribution under Scheme ($A3$) is very accurate, even under the monthly frequency.

Third, we provide an answer to the Bayesian criticisms to unit root econometrics. Since the limiting distribution under Scheme ($A3$) is continuous in κ, the same distribution is used to construct the confidence interval, regardless of the true value of κ. Consequently, the confidence regions based on our asymptotic distribution is connected. Our results show that it is the limiting distribution developed under Scheme ($A1$) or ($A2$) but not the unit root limiting distribution that fails to provide a satisfactory approximation to the finite sample distribution of κ when κ is close to 0. Our answer to the Bayesian criticisms is to use the limiting distribution under Scheme ($A3$) to construct the confidence interval.

The paper is organized as follows. Section 2 reviews and extends the results for the Vasicek model with a known mean. Section 3 derives the results for Vasicek model with a unknown mean. In Section 4, the results are generalized to the model with a flexible diffusion function. Section 5 reports Monte Carlo results and compares the performance of the alternative schemes. Section 6 examines the practical effects of the alternative asymptotic distributions using monthly Federal fund data and tests for unit root in the data. Section 7 concludes. Proofs of the main results in the paper are given in Appendix

2 Vasicek Model with a Known Mean

The Vasicek model with a known mean (without the loss of generality, it is assumed to be zero) is given by:

$$dX(t) = -\kappa X(t)dt + \sigma dW(t), X(0) = X_0. \tag{3}$$

The exact discrete time model corresponding to (3) has the AR(1) structure:

$$X_{t\delta} = \phi X_{(t-1)\delta} + \sigma \sqrt{\frac{1 - e^{-2\kappa\delta}}{2\kappa}} \epsilon_t, \tag{4}$$

where $\phi = e^{-\kappa\delta}$, $\epsilon_t \overset{i.i.d}{\sim} N(0, 1)$.

When there is no confusion, we will simply write $X_{t\delta}$ by X_t. When the discrete data $\{X_{0\delta}, X_{1\delta}, \ldots, X_{n\delta}\}$ $(n\delta = T)$ are available, the LS estimator of ϕ is:

$$\hat{\phi}_n = \frac{\sum X_{t-1} X_t}{\sum X_{t-1}^2},$$

where $\sum := \sum_{t=1}^{n}$. If $\kappa > 0$, the model is strictly stationary. In this case, under Scheme $(A1)$, by the central limit theory of the martingale difference sequences, we have $\sqrt{n}(\hat{\phi}_n - \phi) \xrightarrow{d} N(0, 1 - \phi^2)$ as $n \to \infty$. Since $\hat{\kappa} = -\ln \hat{\phi}_n / \delta$, by the *Delta* method, we have for $\kappa > 0$, as $T \to \infty$

$$\sqrt{T}(\hat{\kappa} - \kappa) \xrightarrow{d} N\left(0, \frac{e^{2\kappa\delta} - 1}{\delta}\right). \tag{5}$$

The asymptotic distribution of $\hat{\kappa}$ was developed in Tang and Chen (2009). It can be seen that the limiting distribution of $\hat{\kappa}$ is independent on the diffusion parameter of the model as well as the initial condition, greatly facilitating statistical inference of κ.

If $\kappa = 0$, then $\phi = 1$ and the model has a unit root. Phillips (1987a) showed that under Scheme $(A1)$:

$$n(\hat{\phi}_n - \phi) \xrightarrow{d} \frac{\int_0^1 W dW}{\int_0^1 W^2 dr}, \tag{6}$$

as $n \to \infty$. By the generalized *Delta* method (Shao, 2003), as $T \to \infty$, we have for $\kappa = 0$, as $T \to \infty$,

$$T(\hat{\kappa} - \kappa) \xrightarrow{d} -\frac{\int_0^1 W dW}{\int_0^1 W^2 dr}. \tag{7}$$

Similarly, under Scheme $(A2)$ with $T \to \infty$ and $\delta \to 0$, the asymptotic distribution is:

$$\sqrt{T}(\hat{\kappa} - \kappa) \xrightarrow{d} N(0, 2\kappa), \tag{8}$$

for $\kappa > 0$ and

$$T(\hat{\kappa} - \kappa) \xrightarrow{d} -\frac{\int_0^1 W dW}{\int_0^1 W^2 dr}, \tag{9}$$

for $\kappa = 0$.

To review the asymptotic results under Scheme ($A3$), we follow Perron (1991) and introduce a few new notations. Denote $J_c(r) = \int_0^r e^{c(r-s)} dW(s)$, $\gamma_0 = X_0 / \left(\sigma\sqrt{T}\right)$, $c = -\kappa T$ and

$$A_1(\gamma_0, c) = \gamma_0 \int_0^1 e^{cr} dW(r) + \int_0^1 J_c(r) dW(r),$$

$$B_1(\gamma_0, c) = \gamma_0^2 (e^{2c} - 1)/2c + 2\gamma_0 \int_0^1 e^{cr} J_c(r) dW(r) + \int_0^1 J_c^2(r) dr.$$

Phillips (1987b) derived the in-fill asymptotic distribution of $\widehat{\phi}_n$ when $T = 1$ and $X(0) = 0$,

$$n(\hat{\phi}_n - \phi) \xrightarrow{d} \frac{\int_0^1 J_{-\kappa}(r) dW(r)}{\int_0^1 J_{-\kappa}^2(r) dr}. \tag{10}$$

For a general T and a general initial condition $X(0) = X_0$, Perron (1991) extended the results of Phillips and showed that:

$$n(\hat{\phi}_n - \phi) \xrightarrow{d} \frac{A_1(\gamma_0, c)}{B_1(\gamma_0, c)}. \tag{11}$$

He further derived the analytical expression for the moment generating function (MGF) of the limiting distribution, facilitating the calculation of its distribution. The asymptotic distribution of $\hat{\kappa}$ under ($A3$) can be easily obtained by applying the generalized *Delta* method to (11):

$$T(\hat{\kappa} - \kappa) \xrightarrow{d} -\frac{A_1(\gamma_0, c)}{B_1(\gamma_0, c)}. \tag{12}$$

This result is closely related to that obtained in Yu (2009a) who showed that, under Scheme ($A3$),

$$\hat{\kappa} \xrightarrow{d} -\frac{\int_0^T X_t dX_t}{\int_0^T X_t^2 dt}, \tag{13}$$

where $X_t = e^{-\kappa t} X_0 + \sigma \int_0^t e^{-\kappa(t-s)} dW(s)$. To simplify the calculation, Yu obtained an alternative form of the limiting distribution by replacing the stochastic integral with the Riemann integral, i.e.,

$$\hat{\kappa} \xrightarrow{d} \frac{T - X(T)^2}{2\int_0^T X(t)^2 dt} \tag{14}$$

Using simulations, Yu demonstrated the superiority of this in-fill asymptotic distribution over the long-span asymptotic distribution (5). It can be verified that the limiting distribution given in (12) is the same as that given in (14). In (12) the initial condition and the parameter in the diffusion function are explicit whereas they are implicit in (14). Interestingly, the in-fill asymptotic theory is the same for $\kappa < 0$ as for $\kappa = 0$. This is in sharp contrast to the long-span asymptotic theory and the double asymptotic theory reviewed earlier.

There is an extensive literature on unit root testing. Nearly all unit root tests are formulated from the discrete time models. In Equation (4) the unit root hypothesis is equivalent to $\phi = 1$. However, the unit root tests can be also performed in continuous time. For example, the unit root hypothesis is equivalent to $\kappa = 0$ in Equation (3). The asymptotic distribution of $\hat{\kappa}$ under Scheme $(A1)$ and $\kappa = 0$ is different from that under Scheme $(A1)$ and $\kappa = \kappa_0 > 0$. This discontinuity is the same as the well-known discontinuity in the asymptotic theory in $\widehat{\phi}$ and suggests that the confidence intervals obtained from (5) and (7) may be two disjoint pieces (Sim, 1988). On the other hand, the confidence intervals obtained from the finite sample distributions must be connected because the finite sample distribution is continuous in κ. This observation has generated some severe criticisms in the Bayesian literature to the use of the nonstationary asymptotic theory (Sim and Uhlig, 1991). See also the critique of the criticisms (Phillips, 1991). Since the in-fill asymptotic distribution is continuous in κ, it provides a unified framework to make statistical inference about κ. In particular, the limiting distribution in (12) is skewed and behaves similar to the unit root limiting distribution when κ is positive and close to 0. Consequently, our answer to the Bayesian criticisms is that the disconnecting confidence intervals are caused by the poor approximation of (5) and (8) to the finite sample distribution, but not by the use of the nonstationary asymptotic theory. Extensive simulations will be carried out later to verify the validity of this claim.

3 Vasicek Model with Unknown Mean

In this section, we consider the Vasicek model with an unknown mean:

$$dX(t) = \kappa(\mu - X(t))dt + \sigma dW(t), X(0) = X_0. \tag{15}$$

The exact discrete time model corresponding to (15) is an AR(1) model with intercept:

$$X_{i\delta} = \mu(1 - e^{-\kappa\delta}) + \phi X_{(i-1)\delta} + \sigma\sqrt{\frac{1 - e^{-2\kappa\delta}}{2\kappa}}\epsilon_i, \tag{16}$$

where $\phi = e^{-\kappa\delta}$, $\epsilon_i \overset{i.i.d}{\sim} N(0, 1)$.

The LS estimator of ϕ is:

$$\hat{\phi}_n = \frac{\sum (X_{t-1} - \overline{X}_-)(X_t - \overline{X})}{\sum (X_{t-1} - \overline{X})^2},$$

where $\overline{X}_- = \frac{1}{n}\sum X_{t-1}$ and $\overline{X} = \frac{1}{n}\sum X_t$.

Under Scheme $(A1)$, Tang and Chen (2009) derived the long-span asymptotic distribution of $\hat{\kappa}$ when $\kappa > 0$:

$$\sqrt{T}(\hat{\kappa} - \kappa) \xrightarrow{d} N\left(0, \frac{e^{2\kappa\delta} - 1}{\delta}\right), \tag{17}$$

as $T \to \infty$. Letting $\delta \to 0$, when $\kappa > 0$, the asymptotic distribution of $\hat{\kappa}$ under $(A2)$ is

$$\sqrt{T}(\hat{\kappa} - \kappa) \xrightarrow{d} N(0, 2\kappa). \tag{18}$$

Asymptotic distributions given in (17) and (18) are the same as those in (5) and (8), respectively.

The in-fill asymptotic distribution has not been derived in the literature and it is more complicated than that in the known mean case. Theorem 3.1 presents the result.

Theorem 3.1 *For Model (15), under Scheme* $(A3)$, *the in-fill asymptotic distribution of* $\hat{\kappa}$ *is*

$$T(\hat{\kappa} - \kappa) \xrightarrow{d} -\frac{A_2(\gamma_0, c)}{B_2(\gamma_0, c)}, \tag{19}$$

where

$$A_2(\gamma_0, c) = \frac{b}{c}\int_0^1 c_1 dW(r) + \int_0^1 J_c(r)dW(r) + \gamma_0\int_0^1 e^{rc}dW(r) - \int_0^1 dW(r)\left(c_2 b + \int_0^1 J_c(r)dr + c_4\gamma_0\right),$$

$$\begin{aligned} B_2(\gamma_0, c) &= c_3 b^2 + \frac{2b}{c}\int_0^1 c_1 J_c(r)dr + \int_0^1 J_c^2(r)dr + c_4^2 b\gamma_0 + 2\gamma_0\int_0^1 e^{rc}J_c(r)dr \\ &\quad + \gamma_0^2\frac{e^{2c} - 1}{2c} - \left(c_2 b + \int_0^1 J_c(r)dr + c_4\gamma_0\right)^2. \end{aligned}$$

and $c = -\kappa T$, $c_1 = e^{rc} - 1$, $c_2 = \frac{e^c - c - 1}{c^2}$, $c_3 = \frac{e^{2c} - 4e^c + 2c + 3}{2c^3}$, $c_4 = \frac{e^c - 1}{c}$, $J_c(r) = \int_0^r e^{c(r-s)}dW(s)$, $b = \mu\sqrt{-c\kappa}/\sigma$, $\gamma_0 = X_0/\left(\sigma\sqrt{T}\right)$.

Remark 3.1 *The in-fill asymptotic theory in (19) is analogous to that of (12) in the Vasicek model with a known mean. It holds true for all values of κ, whether $\kappa < 0$ or $\kappa = 0$.*

Remark 3.2 *In the Vasicek model with a known mean, Perron (1991) derived the expression for the MGF of $-A_1(\gamma_0, c)/B_1(\gamma_0, c)$. Unfortunately, it does not seem that the MGF has an analytic expression for $-A_2(\gamma_0, c)/B_2(\gamma_0, c)$.*

Remark 3.3 *If the mean μ in model (15) is known (and assumed to be 0) and $X_0 = 0$, then model (15) reduces to model (3) with $X_0 = 0$. In this case, by letting $b = 0$ and $X_0 = 0$, we get:*

$$T(\hat{\kappa} - \kappa) \xrightarrow{d} -\frac{\int_0^1 J_c(r)dW(r) - \int_0^1 dW(r)\int_0^1 J_c(r)dr}{\int_0^1 J_c^2(r)dr - \left(\int_0^1 J_c(r)dr\right)^2}.$$

This asymptotic distribution coincides with that in Phillips (1987b).

Remark 3.4 *If $\kappa \to 0$ (so $c \to 0$) and $X_0 = 0$, there is a unit root in the model in the limit. The numerator in (19) becomes*

$$\begin{aligned} & \lim_{c\to 0}\frac{b}{c}\int_0^1 c_1 dW(r) + \int_0^1 J_c(r)dW(r) - \int_0^1 dW(r)\left(c_2 b + \int_0^1 J_c(r)dr\right) \\ = \; & b\int_0^1\left(r - \frac{1}{2}\right)dW(r) + \int_0^1 W(r)dW(r) - \int_0^1 dW(r)\int_0^1 W(r)dr, \end{aligned}$$

and the denominator becomes

$$\begin{aligned} & \lim_{c\to 0} c_3 b^2 + \frac{2b}{c}\int_0^1 c_1 J_c(r)dr - \left(c_2 b + \int_0^1 J_c(r)dr\right)^2 + \int_0^1 J_c^2(r)dr \\ = \; & \frac{b^2}{12} + 2b\int_0^1\left(r - \frac{1}{2}\right)W(r)dr + \int_0^1 W^2(r)dr - \left(\int_0^1 W(r)dr\right)^2. \end{aligned}$$

Hence, the in-fill asymptotic distribution of $\hat{\phi}_n$ in this case is (see Appendix)

$$n(\widehat{\phi}_n - \phi) \xrightarrow{d} \frac{b\int_0^1\left(r - \frac{1}{2}\right)dW(r) + \int_0^1 W(r)dW(r) - \int_0^1 dW(r)\int_0^1 W(r)dr}{\frac{b^2}{12} + 2b\int_0^1\left(r - \frac{1}{2}\right)W(r)dr + \int_0^1 W^2(r)dr - \left(\int_0^1 W(r)dr\right)^2}.$$

This distribution is the same as that obtained in Haldrup and Hylleberg (1995). Haldrup and Hylleberg considered the asymptotic distribution of the LS estimator for a random walk with a drift. Obviously, the results of Haldrup and Hylleberg is a special case of ours. We must note that $c = 0$ means $b = 0$, but here we keep b in the distribution for the purpose of comparison.

Remark 3.5 *If the initial value X_0 is set to zero, we get the asymptotic distribution of κ:*

$$T(\hat{\kappa}-\kappa) \xrightarrow{d} -\frac{\frac{b}{c}\int_0^1 c_1 J_c(r)dr + \int_0^1 J_c(r)dW(r) - \int_0^1 dW(r)\left(c_2 b + \int_0^1 J_c(r)dr\right)}{c_3 b^2 + \frac{2b}{c}\int_0^1 c_1 J_c(r)dr + \int_0^1 J_c^2(r)dr - \left(c_2 b + \int_0^1 J_c(r)dr\right)^2}. \tag{20}$$

Remark 3.6 *We have obtained the double asymptotic distribution of $\hat{\kappa}$ in (18) as a limit case of the long-span asymptotic distribution in (17). The double asymptotic distribution can be also obtained as the limit of the in-fill asymptotic distribution (19). To see it, let the time span $T \to \infty$, i,e. $c \to -\infty$, and we have $(-2c)\int_0^1 J_c^2(r)dr \xrightarrow{p} 1$, $(-2c)^{1/2}\int_0^1 J_c(r)dW(r) \xrightarrow{d} N(0,1)$, $(-2c)^{3/2}\int_0^1 e^{rc}J_c(r)dr \xrightarrow{d} N(0,1)$ and $(-2c)^{1/2}\int_0^1 e^{rc}dW(r) \xrightarrow{d} N(0,1)$. Therefore, the limit of the numerator is*

$$\begin{aligned}
& \frac{b}{c}\int_0^1 c_1 dW(r) + \int_0^1 J_c(r)dW(r) + \gamma_0\int_0^1 e^{rc}dW(r) - \int_0^1 dW(r)\left(c_2 b + \int_0^1 J_c(r)dr + c_4\gamma_0\right) \\
\sim\ & \frac{b}{c}(-2c)^{-1/2}N(0,1) - \frac{b}{c}\left(1 + \frac{e^c - c - 1}{c}\right)N(0,1) + (-2c)^{-1/2}N(0,1) - (-c)^{-1}\chi^2(1) \\
\sim\ & (-2c)^{-1/2}N(0,1) + o_p(c^{-1/2}),
\end{aligned}$$

and the limit of the denominator is

$$\begin{aligned}
& c_3 b^2 + \frac{2b}{c}c_1 J_c(r)dr + \int_0^1 J_c^2(r)dr + c_4^2 b\nu + 2\gamma_0\int_0^1 e^{rc}J_c(r)dr + \gamma_0^2\frac{e^{2c}-1}{2c} \\
& -\left(c_2 b + \int_0^1 J_c(r)dr + c_4\gamma_0\right)^2 \\
\sim\ & \frac{b^2}{c^2} + \frac{2b}{c}(-2c)^{-3/2}N(0,1) - \frac{2b}{c}(-c)^{-1}N(0,1) - \left(-\frac{b}{c} + (-c)^{-1}N(0,1)\right)^2 + (-2c)^{-1} \\
=\ & (-2c)^{-1} + o_p(c^{-1}).
\end{aligned}$$

Consequently,

$$T(\hat{\kappa}-\kappa) \sim -\frac{(-2c)^{-1/2}N(0,1) + o_p(c^{-1/2})}{(-2c)^{-1} + o_p(c^{-1})},$$

and

$$\sqrt{T}(\hat{\kappa}-\kappa) \xrightarrow{d} N(0, 2\kappa). \tag{21}$$

4 General One-factor Model

The model considered in this section has the following expression:

$$dX(t) = \kappa(\mu - X(t))dt + \sigma(X(t))dW(t), X(0) = X_0. \tag{22}$$

Obviously, the standard Lipschitz condition is needed for $\sigma(X(t))$ to ensure that the solution to this SDE exists and is unique. Moreover, we need $X(t)$ to be a positive recurrent and strictly stationary time-reversible process which satisfies strong mixing properties. In particular, following Genon-Catalot, et al (2000), we make the following standard assumptions.

Assumption 1: The function $\sigma(X(t))$ is defined on $(0, +\infty)$ and satisfies

$$\sigma^2(x) \in C^2 \text{ and } 0 < \sigma(x) < +\infty, \forall x \in (0, +\infty),$$

and

$$\exists K > 0, \quad \forall x \in (0, +\infty), \quad |\sigma^2(x)| \leq K(1 + x^2).$$

For $u_0 \in (0, +\infty)$, denote the scale and the speed densities of $X(t)$, respectively, by,

$$s(x) = \exp\left\{-2\int_{u_0}^{x} \frac{\kappa(\mu - u)}{\sigma^2(u)} du\right\} \text{ and } m(x) = \frac{1}{\sigma^2(x)s(x)}.$$

Assumption 2: $\int_0^\infty s(x)dx = +\infty$, $\int_0^\infty m(x)dx = M < +\infty$.

Define the stationary probability density by

$$\pi(x) = \frac{1}{M} m(x) I_{[x \in (0, +\infty)]},$$

where $I_{[\cdot]}$ is the indicator function.

Assumption 3: As $x \to 0$ or $x \to +\infty$, $\lim \sigma(x)m(x) = 0$.

Assumption 4: Define $\gamma(x) = \sigma'(x) - 2\kappa(\mu - x)/\sigma(x)$. If $x \to 0$ or $x \to +\infty$, $\lim 1/\gamma(x) = \tilde{\gamma}_0 < \infty$.

Assumption 5: $E(|X(t)|^p) < \infty$ for some $p > 2$.

Remark 4.1 *Assumption 1 is the global Lipschitz and growth condition. It is typically used in the literature to ensure the existence and uniqueness of a strong solution to SDE (22). Together with Assumption 2, it guarantees the positive recurrence (Genon-Catalot et al, 2000). However, the global Lipschitz may be replaced by the local Lipschitz and growth condition in the one-factor model, as explained in Aït-Sahalia (2002).*

Remark 4.2 *The ρ-mixing property is ensured by Assumptions 3-4 as shown in Genon-Catalot et al (2000) where the mixing rate is also provided; see Appendix.*

Remark 4.3 *Assumption 5 is not as primitive as other assumptions. However, it has been widely used in the literature; see, for example, Yoshida (1992) and Phillips and Yu (2009b).*

We now develop the in-fill asymptotic distribution of the LS estimator of κ. First, note that the exact discrete time model of (22) is given by

$$X_{t\delta} = \mu(1 - e^{-\kappa\delta}) + \phi X_{(t-1)\delta} + \int_0^\delta e^{-\kappa(\delta-\tau)}\sigma(X_{(t-1)\delta+\tau})dW(\tau). \tag{23}$$

Define $Y_{t\delta} = X_{t\delta}/\sqrt{\delta}$ and we can rewrite (23) as

$$Y_{t\delta} = \mu(1 - e^{-\kappa\delta})/\sqrt{\delta} + \phi Y_{(t-1)\delta} + \frac{1}{\sqrt{\delta}}\int_0^\delta e^{-\kappa(\delta-\tau)}\sigma(\sqrt{\delta}Y_{(t-1)\delta+\tau})dW(\tau).$$

Letting

$$\begin{aligned} u_{th} &= \frac{1}{\sqrt{\delta}}\int_0^\delta e^{-\kappa(\delta-\tau)}\sigma(\sqrt{\delta}Y_{(t-1)\delta+\tau})dW(\tau) \\ &= \frac{e^{-\kappa\delta}}{\sqrt{\delta}}\int_0^\delta e^{-\kappa\tau}\sigma(\sqrt{\delta}Y_{(t-1)\delta+\tau})dW(\tau) := \frac{e^{-\kappa\delta}}{\sqrt{\delta}}v_{th}, \end{aligned}$$

we have

$$Y_{th} = \mu(1 - e^{-\kappa\delta})/\sqrt{\delta} + \phi Y_{(t-1)h} + u_{th}. \tag{24}$$

Note that, in general, u_{th} is conditionally heteroskedastic. The LS estimator of ϕ is

$$\hat{\phi}_n = \frac{\sum (Y_{(t-1)h} - \overline{Y}_-)(Y_{th} - \overline{Y})}{\sum (Y_{(t-1)h} - \overline{Y}_-)^2}$$

where $\overline{Y}_- = \frac{1}{n}\sum Y_{(t-1)h}$ and $\overline{Y} = \frac{1}{n}\sum Y_{th}$. The LS estimator of κ is $\hat{\kappa} = -\ln(\hat{\phi}_n)/\delta$. Theorem 4.1 establishes the in-fill asymptotic theory of $\hat{\kappa}$ under Scheme $(A3)$.

Theorem 4.1 *For Model (22), under Scheme $(A3)$ and Assumptions 1-5, the in-fill asymptotic distribution of $\hat{\kappa}$ is*

$$T(\hat{\kappa} - \kappa) \xrightarrow{d} -\frac{A_3(\gamma_0', c)}{B_3(\gamma_0', c)}. \tag{25}$$

where

$$A_3(\gamma_0', c) = \frac{b'}{c}\int_0^1 c_1 dW(r) + \int_0^1 J_c(r)dW(r) + \gamma_0'\int_0^1 e^{rc}dW(r) - \int_0^1 dW(r)\left(c_2 b' + \int_0^1 J_c(r)dr + c_4\gamma_0'\right),$$

$$\begin{aligned} B_3(\gamma_0', c) &= c_3^2 b'^2 + \frac{2b'}{c}\int_0^1 c_1 J(r)dr + \int_0^1 J_c^2(r)dr + c_4^2 b'\gamma_0' + 2\gamma_0'\int_0^1 e^{rc}J_c(r)dr \\ &\quad + \gamma_0'^2\frac{e^{2c}-1}{2c} - \left(c_2 b' + \int_0^1 J_c(r)dr + c_4\gamma_0'\right)^2, \end{aligned}$$

and $c = -\kappa T$, $c_1 = e^{rc} - 1$, $c_2 = \frac{e^c - c - 1}{c^2}$, $c_3 = \frac{e^{2c} - 4e^c + 2c + 3}{2c^3}$, $c_4 = \frac{e^c - 1}{c}$, $\tilde{\sigma}^2 = \lim_{n\to\infty} E(n^{-1}\sum u_t^2)$, $b' = \mu\kappa\sqrt{T}/\tilde{\sigma}$, $\gamma_0' = X_0/(\tilde{\sigma}\sqrt{T})$, $J_c(r) = \int_0^r e^{c(r-s)}dW(s)$.

Remark 4.4 *In the Vasicek model, since $\sigma(X(t)) = \sigma$, $\tilde{\sigma} = \sigma$ and the result in Theorem 4.1 reduces to that in Theorem 3.1.*

Remark 4.5 *Using the standard limit theory of martingale difference sequence, under Scheme (A1), we get*

$$\sqrt{n}(\hat{\phi}_n - \phi) \xrightarrow{d} N\left(0, \frac{(1-\phi^2)^2}{\tilde{\sigma}^4}\bar{\sigma}^2\right)$$

where $\bar{\sigma}^2 = \lim_{n\to\infty} E(n^{-1}\sum Y_{t-1}^2 u_t^2)$. By the Delta method, we can easily get

$$\sqrt{T}(\hat{\kappa} - \kappa) \xrightarrow{d} N\left(0, \frac{(e^{\kappa\delta} - e^{-\kappa\delta})^2}{\delta}\frac{\bar{\sigma}^2}{\tilde{\sigma}^4}\right). \tag{26}$$

Remark 4.6 *Using the same argument as for the Vasicek model, we get the double asymptotics for the one-factor model under Scheme (A2):*

$$\sqrt{T}(\hat{\kappa} - \kappa) \xrightarrow{d} N(0, 2\kappa). \tag{27}$$

Interestingly, this is the same as that under the homoskedastic model. Under the CIR model, Tang and Chen (2009, Theorem 3.2.4) obtained the same doubt asymptotic distribution of a quasi ML estimator.

Remark 4.7 *If the initial value $X_0 = 0$, then the distribution (25) reduces to*

$$T(\hat{\kappa} - \kappa) \xrightarrow{d} -\frac{\frac{b'}{c}\int_0^1 c_1 dW(r) + \int_0^1 J_c(r)dW(r) - \int_0^1 dW(r)\left(c_2 b' + \int_0^1 J_c(r)dr\right)}{c_3^2 b'^2 + \frac{2b'}{c}\int_0^1 c_1 J(r)dr + \int_0^1 J_c^2(r)dr - \left(c_2 b' + \int_0^1 J_c(r)dr\right)^2}.$$

5 Monte Carlo Simulations

Perron (1991) obtained the MGF of $-A_1(\gamma_0, c)/B_1(\gamma_0, c)$ in Equation (12) and used it to tabulate the distribution and the density function. Unfortunately, the in-fill asymptotic distributions in (19) and (25) do not have a closed-form expression for the MGF, nor for the density. In the present paper, we use the method proposed by Chan (1988) to obtain the density of the limiting distributions. As suggested by Chan, the in-fill asymptotic distributions expressed in (19) and (25) may be approximated by Riemann sums and $dW(r)$ by $\epsilon_i/\sqrt{n}$, where $\{\epsilon_i\}$ is a sequence of the standard normal random variables and n the sample size. Consequently, the limiting distribution $\int_0^1 J_c(r)dW(r)/\int_0^1 J_c^2(r)dr$ may be approximated by $n\left(\sum_{i=1}^n \sum_{k=1}^i e^{c(i-k)/n}\epsilon_k\epsilon_{i+1}\right) / \left(\sum_{i=1}^n (\sum_{k=1}^i e^{c(i-k)/n}\epsilon_k)^2\right)$. Chan compared several approximation methods and concluded that the above approximation performs better in the sense that it generate smaller approximation errors, converges faster and is easy to implement.

We design several Monte Carlo experiments to compare the accuracy of the alternative asymptotic distributions of $\widehat{\kappa}$ to the true distribution, all in the context of the following Vasicek model with μ being a unknown parameter:

$$dX(t) = \kappa(\mu - X(t))dt + \sigma dW(t), X(0) = X_0.$$

The true value of κ is set at 0.01, 0.1 and 1, respectively. The first two values are empirically realistic for interest rate data while the last value is empirically realistic for volatility. The true value of μ is set to 0.1, σ to 0.1 and $X_0 = 0$ or $X_0 \sim N(\mu, \sigma^2/2\kappa)$. The value of the sampling interval δ is set at 1/12, 1/52 and 1/252. The time span T is set at 10, so the sample size is 120, 520 and 2520 for monthly, weekly and daily frequencies, respectively.

The percentiles of the statistic $T(\hat{\kappa} - \kappa)$ and the in-fill asymptotic distribution are obtained from 10,000 replications. The Monte Carlo simulation results are reported in Tables 1-6 where the 0.5%, 1%, 5%, 10%, 90%, 95%, 99%, and 99.5% quantiles of the four distributions (i.e., the true distribution, the asymptotic distributions developed under Schemes ($A1$), ($A2$) and ($A3$)), for $\kappa = 0.01, 0.1, 1$, respectively. Tables 1-3 report the results when $X_0 = 0$ and Tables 4-6 report the results when $X_0 \sim N(\mu, \sigma^2/2\kappa)$.

Several features are apparent in the Tables. First, in all cases, the percentiles are not sensitive to the frequency. This observation suggests that the precision of the estimation and the power of a unit root test cannot be increased by using data in a higher frequency but with a fixed time span, even though the sample size increases in this case. On the other hand, the percentiles are sensitive to the value of κ and to the initial condition. The smaller the value of κ, the more sensitive the percentiles to the initial condition. This feature is related to the role that the initial condition plays in the unit root tests; see, for example, Phillips (1987), Müller and Elliott (2003), and Harvey et al (2009).

Second, normality always provides inaccurate approximations of the finite sample distribution, suggesting that when κ is in the range, $(A1)$ and $(A2)$ should not be used in practice as far as statistical inference of κ is concerned. The percentiles from the limiting distribution under Schemes $(A1)$ and $(A2)$ are very different from those obtained from the true distribution, even when $\kappa = 1$. It is obvious that the true distribution of $\hat{\kappa}$ is highly skewed to the right. The long-span asymptotic distribution and the doubt asymptotic distribution perform particularly poorly in the right tail. Interestingly, in all cases, the percentiles of the long-span asymptotic distribution match well to those of the double asymptotic distribution, even when $\delta = 1/12$, suggesting that $\delta \to 0$ is not a too strong assumption.

Third, the in-fill asymptotic distribution provides much more adequate approximations to the finite sample distribution. The smaller the δ is, the better the performance of the in-fill distribution, consistent with our expectation.

Fourth, in all cases, the median of $T(\hat{\kappa} - \kappa)$ is substantially bigger than zero, suggesting a severe positive bias in $\hat{\kappa}$. The bias cannot be reduced by using data in a higher frequency but with a fixed time span. All these results are consistent with those in Phillips and Yu (2005) and Tang and Chen (2009). The bias also manifests in the in-fill asymptotic distribution but not in the long-span and the doubt asymptotic distributions.

Finally, the in-fill asymptotic distribution is less accurate when κ and δ become larger and hence a root is further away from unity. However, the in-fill asymptotic distribution continues to perform much better than the long-span and the doubt asymptotic distributions.

6 An Empirical Application

In this section, we apply the alternative asymptotic theory to the Vasicek model based on real monthly time series data on a short term interest rate series. The data involve the Federal funds rate and are available from the H-15 Federal Reserve Statistical Release. It is sampled monthly and has 432 observations covering the period from July 1954 to June 2002. Since all yields are expressed in annualized form, we have $\delta = 1/12$ for the monthly data. The same data were used in Aït-Sahalia (1999).

Table 7 shows the sample sizes, means, standard deviations, first seven autocorrelations, and Phillips-Perron $Z(t)$ unit root test statistic (with a fitted intercept in the regression) for the series. The presence of a unit root cannot be rejected at the 10% level. These results, together with the form of the sample autocorrelogram, suggest that the interest rate is highly persistent.

Assuming X_0 is the same as the first observation, the ML/LS estimates of the three parameters κ, μ and σ are: $\hat{\kappa} = 0.2613$, $\hat{\mu} = 0.0717$ and $\hat{\sigma} = 0.0223$. Consequently, we can get the 90% and 95% confidence intervals for κ under the three schemes, which are reported in Table 8. Under Schemes $(A1)$ and $(A2)$, the limit distribution is different when $\kappa > 0$ from that when $\kappa = 0$. So two sets of confidence intervals are reported in the two cases. As found in the Monte Carlo study, the confidence intervals obtained from $(A1)$ and $(A2)$ are nearly identical since $\delta = 1/12$ is small.

It is well documented in the term structure literature that the short term interest rates are highly persistent. However, no agreement has reached among economists whether or not the short term interest rates have a unit root. For example, Aït-Sahalia (1996b) argued that the short term interest rate is stationary while Stock and Watson (1988) reported evidence of a unit root in the Federal fund rate. Using the confidence intervals (either 90% or 95%) constructed under Schemes $(A1)$ and $(A2)$ and $\kappa = \kappa_0 > 0$, one would conclude that there is no unit root in the data. However, the confidence intervals (both 90% and 95%) constructed under Schemes $(A1)$ and $(A2)$ and $\kappa = 0$ suggest that there is a unit root in the data. This discrepancy is, of course, due to the discontinuity in the asymptotic distributions at unity.

Under Scheme $(A3)$ the confidence interval does not depend on the true value of κ and hence only one confidence interval is needed. In this case, both the 90% and the 95% confidence intervals contain zero, suggesting that there is a unit root in the data. Interestingly, the confidence intervals are very similar to those obtained from the unit root asymptotic distribution. We conclude that it is the asymptotic normality but not the unit root asymptotic distribution that causes the problem of the disconnected confidence interval. As we showed earlier, the asymptotic distribution under Scheme ($A3$) is more accurate and robust to the hypothesized value of κ. Consequently, we believe the empirical result based on Scheme ($A3$) and hence the unit root hypothesis are more reliable.

7 Conclusion

In this paper, we have developed the asymptotic distributions of the LS estimator of the mean reversion parameter (κ) in a general class of continuous time models under three schemes, namely, long-span, in-fill and the combination of long-span and in-fill. While the drift has an affine structure in our model, nonlinearity is allowed in the diffusion function. The limiting distributions are quite different under the alternative schemes. In particular, the in-fill limiting distribution is non-standard and depend on the time span and the initial value. However, it is applicable to all values of κ, including the unit root case. Consequently, the confidence intervals obtained from the in-fill limiting distribution are not disconnected. Monte Carlo simulations suggest that the in-fill asymptotic distribution provides more accurate approximations to the finite sample distribution than the other two asymptotic distributions in empirically realistic cases. Empirical applications to U.S. Federal fund rates suggest an importance difference in statistical inference based on the alternative asymptotic distributions. While the long-span and the double asymptotic distributions reject the hypothesis of unit root in the model, the in-fill asymptotic distribution does not reject the hypothesis of unit root in the model.

A more general continuous time model may be specified by the following system of SDEs:

$$d\begin{pmatrix} X(t) \\ V(t) \end{pmatrix} = \begin{pmatrix} \kappa(\mu - X(t)) \\ \mu_V(V(t)) \end{pmatrix} dt + \begin{pmatrix} \sigma_X(X(t), V(t)) \\ \sigma_V(V(t)) \end{pmatrix} dW(t), \tag{28}$$

Here $X(t)$ is observed by the econometrician and $V(t)$ helps to determine the volatility of $X(t)$ that is latent and evolves randomly. This stochastic volatility model has been widely used in finance to price contingent-claims. It is useful to generalize the in-fill limit theory to cover the stochastic volatility model. We plan to report the results in future work.

8 Appendix

8.1 Proof of Theorem 3.1

To prove this theorem, we follow Phillips (1987b), Perron (1991) and Haldrup and Hylleberg (1995). Define $a(\delta) = \sigma\sqrt{(1-e^{-2\kappa\delta})/2\kappa}$ and $Y_t = X_t/a(\delta)$. Dividing Equation (16) by $a(\delta)$, we get:

$$Y_t = \mu(1-e^{-\kappa\delta})/a(\delta) + \phi Y_{t-1} + \epsilon_t. \tag{29}$$

In Equation (29), the drift has the order of $O_p(n^{-1/2})$. When $n \to \infty$, we can define the drift as $\mu^* = b/\sqrt{n}$, where $b = \mu\sqrt{-c\kappa}/\sigma$.

Expanding (29), we have:

$$\begin{aligned} Y_t &= \mu^* \frac{\phi^t - 1}{\phi - 1} + \sum_{j=0}^{t} e^{(t-j)c/n}\epsilon_j + e^{tc/n}Y_0 + o_p(n^{-1/2}) \\ &= \frac{b}{\sqrt{n}} \frac{e^{tc/n} - 1}{e^{c/n} - 1} + \sum_{j=0}^{t} e^{(t-j)c/n}\epsilon_j + e^{tc/n}Y_0 + o_p(n^{-1/2}), \end{aligned}$$

where $c = -\kappa T$, T is the time span, and $Y_0 = X_0/a(\delta)$ is the initial condition. To simplify the expressions, denote $X_0/(\sigma\sqrt{T})$ by γ_0. Obviously, $n^{-1/2}Y_0 \to \gamma_0$ (as $n \to \infty$).

Define the partial sum of ϵ_t as $Z_n(r) = n^{-1/2}S_{[nr]} = n^{-1/2}\sum_{t=1}^{[nr]}\epsilon_t$ $(0 \le r \le 1)$. We have, as $n \to \infty$,

$$Z_n(r) \xrightarrow{d} W(r).$$

Before proving Theorem 3.1 we first establish the following lemma.

Lemma 8.1 *If Y_t is generated according to (29), then as $n \to \infty$*

$$n^{-1/2}Y_{[nr]} \xrightarrow{d} \frac{b(e^{cr}-1)}{c} + J_c(r) + \gamma_0 e^{rc}, \text{ for } 0 \le r \le 1; \tag{a}$$

$$n^{-3/2}\sum_{t=1}^{n} Y_t \xrightarrow{d} \frac{e^c - c - 1}{c^2} b + \int_0^1 J_c(r)dr + \frac{e^c-1}{c}\gamma_0; \tag{b}$$

$$\begin{aligned} n^{-2}\sum_{t=1}^{n} Y_t^2 \xrightarrow{d} & \frac{e^{2c} - 4e^c + 2c + 3}{2c^3} b^2 + \frac{2b}{c}\int_0^1 (e^{rc}-1)J_c(r)dr + \int_0^1 J_c^2(r)dr \\ & + \frac{e^{2c} - 2e^c + 1}{c^2} b\gamma_0 + 2\gamma_0 \int_0^1 e^{rc} J_c(r)dr + \gamma_0^2 \frac{e^{2c}-1}{2c}; \end{aligned} \tag{c}$$

$$n^{-1}\sum_{t=1}^{n} Y_{t-1}\epsilon_t \xrightarrow{d} \frac{2b}{c}\int_0^1 (e^{cr}-1)J_c(r)dr + \int_0^1 J_c(r)dW(r) + \gamma_0 \int_0^1 e^{rc} dW(r). \tag{d}$$

Proof of Lemma 8.1: (a)

$$\begin{aligned} n^{-1/2}Y_{[nr]} &= n^{-1/2}\left(\frac{b}{\sqrt{n}}\frac{e^{[nr]c/n}-1}{e^{c/n}-1} + \sum_{j=0}^{[nr]} e^{([nr]-j)c/n}\epsilon_j + e^{[nr]c/n}Y_0 + o_p(n^{-1/2})\right) \\ &= \frac{b(e^{[nr]c/n}-1)}{n(e^{c/n}-1)} + n^{-1/2}\sum_{j=0}^{[nr]} e^{([nr]-j)c/n}\epsilon_j + n^{-1/2}e^{[nr]c/n}Y_0 + o_p(n^{-1/2}) \\ &\xrightarrow{d} \frac{b(e^{rc}-1)}{c} + J_c(r) + e^{rc}\gamma_0. \end{aligned}$$

(b)

$$\begin{aligned} n^{-3/2}\sum_{t=1}^{n} Y_t &= \frac{n^{-2}b}{e^{c/n}-1}\left(\sum_{t=1}^{n} e^{tc/n} - n\right) + n^{-3/2}\sum_{t=1}^{n}\sum_{j=1}^{t} e^{(t-j)c/n}\epsilon_j + n^{-3/2}\sum_{t=1}^{n} e^{tc/n}Y_0 + o_p(1) \\ &= n^{-2}\frac{b}{e^{c/n}-1}\left(\frac{e^{c(n+1)/n} - e^{c/n}}{e^{c/n}-1} - n\right) + n^{-1}\sum_{t=1}^{n} n^{-1/2}\sum_{j=1}^{t} e^{(t-j)c/n}\epsilon_j \\ &\quad + n^{-3/2}Y_0\frac{e^{c(n+1)/n} - e^{c/n}}{e^{c/n}-1} + o_p(1) \\ &= \frac{b(e^{c(n+1)/n} - e^{c/n})}{n^2(e^{c/n}-1)^2} - \frac{b}{n(e^{c/n}-1)} + n^{-1}\sum_{t=1}^{n} J_c\left(\frac{t}{n}\right) + \frac{e^c-1}{c}\gamma_0 + o_p(1) \\ &\xrightarrow{d} \frac{e^c - c - 1}{c^2} b + \int_0^1 J_c(r)dr + \frac{e^c-1}{c}\gamma_0. \end{aligned}$$

(c)

$$
\begin{aligned}
n^{-2}\sum_{t=1}^{n} Y_t^2 &= n^{-2}\sum_{t=1}^{n}\left(\frac{b}{n^{1/2}}\frac{e^{tc/n}-1}{e^{c/n}-1}+\sum_{j=0}^{t} e^{(t-j)c/n}\epsilon_j + e^{tc/n}Y_0\right)^2 \\
&= n^{-2}\sum_{t=1}^{n}\Bigg\{\left(\frac{b}{n^{1/2}}\frac{e^{tc/n}-1}{e^{c/n}-1}+\sum_{j=0}^{t} e^{(t-j)c/n}\epsilon_j\right)^2 + \\
&\quad 2\left(\frac{b}{n^{1/2}}\frac{e^{tc/n}-1}{e^{c/n}-1}+\sum_{j=0}^{t} e^{(t-j)c/n}\epsilon_j\right) e^{tc/n}Y_0 + e^{2tc/n}Y_0^2\Bigg\} \\
&= n^{-2}\sum_{t=1}^{n}\Bigg\{\frac{b^2}{n}\frac{(e^{tc/n}-1)^2}{(e^{c/n}-1)^2}+\left(\sum_{j=0}^{t} e^{(t-j)c/n}\epsilon_j\right)^2+\frac{2b}{n^{1/2}}\frac{e^{tc/n}-1}{e^{c/n}-1}\sum_{j=0}^{t} e^{(t-j)c/n}\epsilon_j \\
&\quad +\frac{2b}{n^{1/2}}\frac{e^{tc/n}-1}{e^{c/n}-1}e^{tc/n}Y_0 + 2e^{tc/n}Y_0\sum_{j=0}^{t} e^{(t-j)c/n}\epsilon_j + e^{2tc/n}Y_0^2\Bigg\}.
\end{aligned}
$$

The first term of the sum is:

$$
\begin{aligned}
&\frac{1}{n^2}\sum_{t=1}^{n}\frac{b^2}{n}\frac{(e^{tc/n}-1)^2}{(e^{c/n}-1)^2} = \frac{b^2}{n^3(e^{c/n}-1)^2}\sum_{t=1}^{n}(e^{2ct/n}-2e^{ct/n}+1) \\
= &\ \frac{b^2}{n^3(e^{c/n}-1)^2}\left(\frac{e^{2c/n}-e^{2c(n+1)/n}}{1-e^{2c/n}} - 2\frac{e^{c/n}-e^{c(n+1)/n}}{1-e^{c/n}} + n\right) \\
= &\ b^2\left(\frac{e^{2c/n}-e^{2c(1+1/n)}}{(e^{c/n}-1)^2 n^2(1-e^{2c/n})n} - 2\frac{e^{c/n}-e^{(1+1/n)c}}{(e^{c/n}-1)^2 n^2(1-e^{c/n})n} + \frac{1}{n^2(e^{c/n}-1)^2}\right) \\
\to &\ \frac{e^{2c}-4e^{c}+2c+3}{2c^3}b^2.
\end{aligned}
$$

The second term of the sum is:

$$
\begin{aligned}
\frac{1}{n^2}\sum_{t=1}^{n}\left(\sum_{j=1}^{t} e^{(t-j)c/n}\epsilon_j\right)^2 &= \frac{1}{n}\sum_{t=1}^{n}\left(n^{-1/2}\sum_{j=1}^{t} e^{(t-j)c/n}\epsilon_j\right)^2 \\
= \frac{1}{n}\sum_{t=1}^{n} J_c^2\left(\frac{t}{n}\right) + O_p(n^{-1}) &\xrightarrow{d} \int_0^1 J_c^2(r)dr.
\end{aligned}
$$

The third term of the sum is:

$$
\begin{aligned}
&\frac{1}{n^2}\frac{2b}{n^{1/2}(e^{c/n}-1)}\sum_{t=1}^{n}\left(e^{tc/n}-1\right)\sum_{j=1}^{t}e^{(t-j)c/n}\epsilon_j \\
&= \frac{2b}{n(e^{c/n}-1)}\frac{1}{n}\sum_{t=1}^{n}\left(e^{tc/n}-1\right)\frac{1}{\sqrt{n}}\sum_{j=1}^{t}e^{(t-j)c/n}\epsilon_j \\
&= \frac{2b}{n(e^{c/n}-1)}\left(\frac{1}{n}\sum_{t=1}^{n}e^{tc/n}\frac{1}{\sqrt{n}}\sum_{j=1}^{t}e^{(t-j)c/n}\epsilon_j-\frac{1}{n}\frac{1}{\sqrt{n}}\sum_{j=1}^{t}e^{(t-j)c/n}\epsilon_j\right) \\
&= \frac{2b}{c}\left(\frac{1}{n}\sum_{t=1}^{n}e^{ct/n}J_c\left(\frac{t}{n}\right)-\frac{1}{n}\sum_{t=1}^{n}J_c\left(\frac{t}{n}\right)\right)+O_p(n^{-1}) \\
&\xrightarrow{d} \frac{2b}{c}\int_0^1(e^{cr}-1)J_c(r)dr.
\end{aligned}
$$

The fourth term of the sum is:

$$
\begin{aligned}
\frac{2b}{n^{5/2}}\sum_{t=1}^{n}\frac{e^{tc/n}-1}{e^{c/n}-1}e^{tc/n}Y_0 &= \frac{2bY_0}{n^{5/2}(e^{c/n}-1)}\sum_{t=1}^{n}\left(e^{2tc/n}-e^{tc/n}\right) \\
&= \frac{2bY_0}{n(e^{c/n}-1)}\frac{1}{n^{3/2}}\left(\frac{e^{2c/n}-e^{2c(n+1)/n}}{1-e^{2c/n}}-\frac{e^{c/n}-e^{c(n+1)/n}}{1-e^{c/n}}\right) \\
&= \frac{2bY_0}{n(e^{c/n}-1)}\frac{1}{n^{1/2}}\left(\frac{e^{2c}-1}{2c}-\frac{e^{c}-1}{c}\right)\to\frac{e^{2c}-2e^{c}+1}{c^2}b\gamma_0.
\end{aligned}
$$

The fifth term of the sum is:

$$
\begin{aligned}
2n^{-2}\sum_{t=1}^{n}\left(e^{tc/n}Y_0\sum_{j=1}^{t}e^{(t-j)c/n}\epsilon_j\right) &= 2n^{-3/2}Y_0\sum_{t=1}^{n}\left(e^{tc/n}n^{-1/2}\sum_{t=1}^{n}e^{(t-j)c/n}\epsilon_j\right) \\
&= 2n^{-3/2}Y_0\sum_{t=1}^{n}e^{tc/n}J_c\left(\frac{t}{n}\right)+O_p(n^{-3/2}) \\
&= 2n^{-1/2}Y_0\int_0^1e^{rc}J_c(r)dr+O_p(n^{-3/2}) \\
&\xrightarrow{d} 2\gamma_0\int_0^1e^{rc}J_c(r)dr.
\end{aligned}
$$

Obviously, the last term of the sum converges to $\gamma_0^2\frac{e^{2c}-1}{2c}$. Combing the above equations we can easily get the results of Lemma 1 (c).

(d) For the sum

$$\begin{aligned}
n^{-1}\sum_{t=1}^{n} Y_t\epsilon_{t+1} &= n^{-1}\sum_{t=1}^{n}\left(\frac{b}{\sqrt{n}}\frac{e^{tc/n}-1}{e^{c/n}-1}+\sum_{j=1}^{t}e^{(t-j)c/n}\epsilon_j+e^{tc/n}Y_0\right)\epsilon_{t+1}\\
&= \frac{1}{n}\sum_{t=1}^{n}\frac{b}{\sqrt{n}}\frac{e^{tc/n}-1}{e^{c/n}-1}\epsilon_{t+1}+\frac{1}{n}\sum_{t=1}^{n}\epsilon_{t+1}\sum_{j=1}^{t}e^{(t-j)c/n}\epsilon_j+\frac{Y_0}{n}\sum_{t=1}^{n}e^{tc/n}\epsilon_{t+1},
\end{aligned}$$

the first term is:

$$\begin{aligned}
\frac{1}{n}\sum_{t=1}^{n}\frac{b}{\sqrt{n}}\frac{e^{tc/n}-1}{e^{c/n}-1}\epsilon_{t+1} &= \frac{b}{n(e^{c/n}-1)}\frac{1}{\sqrt{n}}\sum_{t=1}^{n}(e^{tc/n}-1)\epsilon_{t+1}\\
&= \frac{b}{c}\sum_{t=1}^{n}(e^{tc/n}-1)\int_{t/n}^{(t+1)/n}dZ_n(r)\\
&= \frac{b}{c}\sum_{t=1}^{n}\int_{j/n}^{(j+1)/n}(e^{rc}-1)dZ_n(r)+O_p(n^{-1})\\
&\overset{d}{\to} \frac{b}{c}\int_0^1(e^{rc}-1)dW(r).
\end{aligned}$$

It is easy to see that the second term is the same as the first term except for the coefficient, i.e.,

$$\frac{1}{n}\sum_{t=1}^{n}\epsilon_{t+1}\sum_{j=1}^{t}e^{(t-j)c/n}\epsilon_j \overset{d}{\to} \int_0^1 J_c(r)dW(r).$$

And the last term is:

$$\frac{Y_0}{n}\sum_{t=1}^{n}e^{tc/n}\epsilon_{t+1}\overset{d}{\to}\frac{Y_0}{\sqrt{n}}\int_0^1 e^{rc}dW(r)=\gamma_0\int_0^1 e^{rc}dW(r)$$

Therefore,

$$\frac{1}{n}\sum Y_t\epsilon_{t+1}\overset{d}{\to}\frac{2b}{c}\int_0^1(e^{cr}-1)J_c(r)dr+\int_0^1 J_c(r)dW(r)+\nu\int_0^1 e^{rc}dW(r)$$

To prove Theorem 3.1, we note that

$$\hat{\phi}_n=\frac{\sum(Y_{t-1}-\overline{Y}_-)(Y_t-\overline{Y})}{\sum(Y_{t-1}-\overline{Y}_-)^2}=\phi+\frac{\sum(Y_{t-1}-\overline{Y}_-)\epsilon_t}{\sum\ (Y_{t-1}-\overline{Y}_-)^2}.$$

Hence,

$$n(\hat{\phi}_n - \phi) = \frac{n^{-1}\sum (Y_{t-1} - \overline{Y}_{-})\epsilon_t}{n^{-2}\sum (Y_{t-1} - \overline{Y}_{-})^2} = \frac{n^{-1}\sum Y_{t-1}\epsilon_t - n^{-1/2}\sum \epsilon_t n^{-3/2}\sum Y_t}{n^{-2}\sum Y_{t-1}^2 - (n^{-3/2}\sum Y_{t-1})^2} \xrightarrow{d}$$

$$-\frac{\frac{b}{c}\int_0^1 c_1 dW(r) + \int_0^1 J_c(r)dW(r) + \gamma_0\int_0^1 e^{rc}dW(r) - \int_0^1 dW(r)\left(c_2 b + \int_0^1 J_c(r)dr + c_4\gamma_0\right)}{c_3 b^2 + \frac{2b}{c}c_1\int_0^1 J_c(r)dr + \int_0^1 J_c^2(r)dr + c_4^2 b\gamma_0 + 2\gamma_0\int_0^1 e^{rc}J_c(r)dr + \gamma_0^2\frac{e^{2c}-1}{2c} - \left(c_2 b + \int_0^1 J_c(r)dr + c_4\gamma_0\right)^2}$$

where $c = -\kappa T, c_1 = e^{rc} - 1, c_2 = \frac{e^c - c - 1}{c^2}, c_3 = \frac{e^{2c} - 4e^c + 2c + 3}{2c^3}, c_4 = \frac{e^c - 1}{c}$ and $J_c(r) = \int_0^r e^{c(r-s)}dW(s)$, $b = \mu\sqrt{-c\kappa}/\sigma$, $\gamma_0 = X_0/\sigma\sqrt{T}$. Since $\hat{\kappa} = -\ln(\hat{\phi}_n)/\delta$, by the generalized *Delta* method (Theorem 1.12, Shao, 2003), we can get the result of the theorem.

Before we prove Theorem 4.1, we need a lemma. Its proof can be found in Genon-Catalot et al (2000).

Lemma 8.2 *(Genon-Catalot et al, 2000): (1) Under Assumptions 1-4, X_t is time reversible, and X_t as well as $X_{t\delta}$, for all δ, are ergodic and β-mixing. (2) Under Assumptions 1-4, X_t is ρ-mixing if and only if the limits in Assumption 4 are finite. (3) Under Assumptions 1-4 and assume that the limits in Assumption 4 are finite, there exists a positive λ such that $\alpha_X(t) \le e^{-\lambda t}/4$.*

8.2 Proof of Theorem 4.1

For Model (24), we need to show that u_{th} in the following local-to-unity model

$$Y_{th} = e^{-\kappa\delta}Y_{(t-1)h} + u_{th}, t = 0, h, 2h, \cdots nh(:= T)$$

satisfies the four conditions imposed by Phillips (1987b, page 537):

(i) $E(u_{th}) = 0$ for all t;

(ii) $\sup_t E|u_{th}|^p < \infty$ for some $p > 2$;

(iii) As $n \to \infty$, $\tilde{\sigma}^2 = \lim E(n^{-1}S_n^2)$ exists and $\tilde{\sigma}^2 > 0$, where $S_{th} = u_{1h} + ... + u_{th}$;

(iv) u_t is strong mixing with mixing coefficients α_m that satisfy

$$\sum_{m=1}^{\infty} \alpha_m^{1-2/p} < \infty.$$

It is easily to see that (i) is satisfied using conditioning argument. To verify condition (ii), first note that the exact discrete time model of (22) is given by

$$X_{t\delta} = \mu(1 - e^{-\kappa\delta}) + \phi X_{(t-1)\delta} + \int_0^\delta e^{-\kappa(\delta-\tau)}\sigma(X_{(t-1)\delta+\tau})dW(\tau).$$

Defining $Y_{t\delta} = X_{t\delta}/\sqrt{\delta}$ and

$$\begin{aligned} u_{th} &= \frac{1}{\sqrt{\delta}} \int_0^\delta e^{-\kappa(\delta-\tau)} \sigma(\sqrt{\delta} Y_{(t-1)\delta+\tau}) dW(\tau) \\ &= \frac{e^{-\kappa\delta}}{\sqrt{\delta}} \int_0^\delta e^{-\kappa\tau} \sigma(\sqrt{\delta} Y_{(t-1)\delta+\tau}) dW(\tau) := \frac{e^{-\kappa\delta}}{\sqrt{\delta}} v_{th}, \end{aligned}$$

where $v_{th} = \int_0^\delta e^{-\kappa\tau} \sigma(\sqrt{\delta} Y_{(t-1)\delta+\tau}) dW(\tau)$, we get

$$Y_{th} = \mu(1 - e^{-\kappa\delta})/\sqrt{\delta} + \phi Y_{(t-1)h} + u_{th}. \tag{30}$$

Obviously, v_{th} is a martingale. Suppose M is an positive integer, we now introduce M martingale increments, $\{\zeta_m\}_{m=1}^M$, where

$$\zeta_1 = \int_0^{\delta/M} e^{-\kappa\tau} \sigma(\sqrt{\delta} Y_{(t-1)\delta+\tau}) dW(\tau), \cdots, \zeta_M = \int_{\delta(M-1)/M}^\delta e^{-\kappa\tau} \sigma(\sqrt{\delta} Y_{(t-1)\delta+\tau}) dW(\tau).$$

The quadratic variation of each ζ_m is given by

$$\zeta_m^2 = \int_{\delta(m-1)/M}^{\delta m/M} e^{-2\kappa\tau} \sigma^2(\sqrt{\delta} Y_{(t-1)\delta+\tau}) d\tau.$$

By the Burkholder inequality (Burkholder, 1966), for any $\alpha > 1$, $\exists C_\alpha > 0$ such that

$$(E|v_{th}|)^\alpha \le C_\alpha \left(E\left|\zeta_1^2 + \cdots + \zeta_m^2\right|\right)^{\alpha/2} = C_\alpha \left(E\left|\int_0^\delta e^{-2\kappa\tau} \sigma^2(\sqrt{\delta} Y_{(t-1)\delta+\tau}) d\tau\right|\right)^{\alpha/2}.$$

Now if we choose $\alpha = p$, then by Assumption 5 we have:

$$\begin{aligned} \sup_t E|u_{th}|^p &= \left(\frac{e^{-\kappa\delta}}{\sqrt{\delta}}\right)^p \sup_t (E|v_{th}|)^p \\ &\le C_p \left(\frac{e^{-\kappa\delta}}{\sqrt{\delta}}\right)^p \left(E\left|\int_0^\delta e^{-2\kappa\tau} \sigma^2(\sqrt{\delta} Y_{(t-1)\delta+\tau}) d\tau\right|\right)^{p/2} \\ &\le C_p' \left(\frac{e^{-\kappa\delta}}{\sqrt{\delta}}\right)^p \left(E\left|\int_0^\delta e^{-2\kappa\tau} d\tau\right|\right)^{p/2} \\ &= C_p' \left(\frac{1 - e^{-2\kappa\delta}}{2\kappa\delta}\right)^{p/2}. \end{aligned}$$

This quantity converges to C_p' as $\delta \to 0$ and hence verifies condition (ii).

For condition (iii), since $\{u_t\}$ is α-mixing (strong mixing), by Corollary 5.1 of Hall and Heyde (1980), we obtain that $\lim_{n\to\infty} n^{-1} E S_n^2 = \tilde{\sigma}^2$, where $0 < \tilde{\sigma}^2 < \infty$. We must note that

$\tilde{\sigma}^2$ cannot be zero due to the fact that $ES_n^2 = \sum_{t=1}^n E(u_t^2) + 2\sum_{i>j} E(u_i u_j) = \sum_{t=1}^n E(u_t^2)$ and $E(u_t^2)$ is some certain constant.

For (iv), we note that $u_t = g(X_{t-1})$, where $g(\cdot)$ is a measurable function. By Theorem 3.49 of White (2001), u_t is also α-mixing with $\alpha_U(t) \leq e^{-\lambda t}/4$ under Assumptions 1-5. Thus, $\sum_{m=1}^\infty \alpha_{m,U}^{1-2/\beta} \leq \sum_{m=1}^\infty e^{-\lambda m(1-2/\beta)}/4 < \infty$ for some $\beta > 2$ and positive λ.

Define $\tilde{\sigma}^2 = \lim_{n\to\infty} E(n^{-1}S_n^2)$. Under Assumptions 1-5, the partial sum $\{S_t\}$ obeys a central limit theory on the functional space D, i.e, as $n \to \infty$,

$$\tilde{Z}_n(r) = n^{-1/2}\tilde{\sigma}^{-1}S_{[nr]} \xrightarrow{d} W(r) \qquad (0 \leq r \leq 1)$$

where $[nr]$ denotes the integer part of nr. This result can be found in Phillips (1987a, 1987b).

The remaining part of the proof is the same as in the Vasicek model. To save space, we just list the main results here. Defining $\mu' = \frac{b^\star}{\sqrt{n}}$, $b^\star = \mu\kappa\sqrt{T}$, $b' = b^\star/\tilde{\sigma} = \mu\kappa\sqrt{T}/\tilde{\sigma}$ and $\gamma_0' = \frac{X_0}{\tilde{\sigma}\sqrt{T}}$, where X_0 is the initial value, we get:

$$\begin{aligned} Y_t &= \mu(1-e^{-\kappa\delta})/\sqrt{\delta} + e^{-\kappa\delta}Y_{t-1} + u_t \\ &= \mu'\frac{\phi^t - 1}{\phi - 1} + \sum_{j=0}^t e^{(t-j)c/n}u_j + e^{tc/n}Y_0 + o_p(n^{-1/2}) \\ &= \frac{b^\star}{\sqrt{n}}\frac{e^{tc/n}-1}{e^{c/n}-1} + \sum_{j=0}^t e^{(t-j)c/n}u_j + e^{tc/n}Y_0 + o_p(n^{-1/2}) \end{aligned}$$

The following lemma is important to prove Theorem (4.1):

Lemma 8.3 *If Y_t is generated according to (24), then as $n \to \infty$*

$$n^{-1/2}\tilde{\sigma}^{-1}Y_{[nr]} \xrightarrow{d} \frac{b'(e^{cr}-1)}{c} + J_c(r) + \gamma_0' e^{rc} \quad \text{for} \quad 0 \leq r \leq 1; \tag{a'}$$

$$n^{-3/2}\tilde{\sigma}^{-1}\sum_{t=1}^n Y_t \xrightarrow{d} \frac{e^c - c - 1}{c^2}b' + \int_0^1 J_c(r)dr + \frac{e^c - 1}{c}\gamma_0'; \tag{b'}$$

$$\begin{aligned} n^{-2}\tilde{\sigma}^{-2}\sum_{t=1}^n Y_t^2 \xrightarrow{d}\ & \frac{e^{2c} - 4e^c + 2c + 3}{2c^3}b'^2 + \frac{2b'}{c}\int_0^1 (e^{rc}-1)J_c(r)dr + \int_0^1 J_c^2(r)dr \\ & + \frac{e^{2c} - 2e^c + 1}{c^2}b'\gamma_0' + 2\gamma_0'\int_0^1 e^{rc}J_c(r)dr + \gamma_0'^2\frac{e^{2c}-1}{2c}; \end{aligned} \tag{c'}$$

$$n^{-1}\tilde{\sigma}^{-2}\sum_{t=1}^n Y_{t-1}u_t \xrightarrow{d} \frac{2b'}{c}\int_0^1 (e^{cr}-1)J_c(r)dr + \int_0^1 J_c(r)dW(r) + \gamma_0'\int_0^1 e^{rc}dW(r). \tag{d'}$$

The proof of Lemma 8.3 is the same as that of Lemma 8.1. By using the results in Lemma 8.2, one can get easily get the results of Theorem 4.1. The proof is omitted.

References

[1] Ahtola, J. and G. C. Tiao, 1984, Parameter inference for a nearly nonstationary first order autoregressive model. *Biometrika*, 71, 263-272.

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

[3] Aït-Sahalia, Y., 1996b, Testing Continuous-time Models of Spot Interest Rate Derivative Securities, *Review of Financial Studies*, 9, 385–426.

[4] Aït-Sahalia, Y., 1999, Transition Densities for Interest Rate and Other Non-linear Diffusions. *Journal of Finance*, 54, 1361–1395.

[5] Aït-Sahalia, Y., 2002, Maximum likelihood estimation of discretely sampled diffusion: A closed-form approximation approach. *Econometrica*, 70, 223–262.

[6] Aït-Sahalia, Y., and J. Park, 2009, Stationarity-Based Specification Tests for Diffusions when the Process is Nonstationary, Working Paper, Princeton University.

[7] Andersen, T., T. Bollerslev, F.X. Diebold, and P. Labys 2001, The Distribution of Realized Exchange Rate Volatility. *Journal of the American Statistical Association* 96, 42–55.

[8] Bandi, F. M. and P.C.B. Phillips, 2003, Fully nonparametric estimation of scalar diffusion models. *Econometrica*, 71, 241–283.

[9] Bandi, F. M. and P.C.B. Phillips, 2007, A simple approach to the parametric estimation of potentially nonstationary diffusions. *Journal of Econometrics*, 137, 354-395.

[10] Barndorff-Nielsen, O. and N. Shephard 2002, Econometric analysis of realized volatility and its use in estimating stochastic volatility models, *Journal of the Royal Statistical Society, Series B*, 64, 253-280.

[11] Brown, B.W. and Hewitt, J.T., 1975, Asymptotic likelihood theory for diffusion prcesses, *Journal of Applied Probability*, 12, 228-238.

[12] Burkholder, D. L., 1966, Martingale transforms, *Annals of Mathematical Statistics*, 37, 1494-1505.

[13] Chapman, D. A., and N. Pearson, 2000, Is the Short Rate Drift Actually Nonlinear? *Journal of Finance*, 55, 355–388.

[14] Chan, K. C., G. A. Karolyi, F. A. Longstaff, and A. B. Sanders 1992, An Empirical Comparison of Alternative Models of Short Term Interest Rates, *Journal of Finance*, 47, 1209–1227.

[15] Chan, N.H., 1988. The parameter inference for nearly nonstationary time series. *Journal of the American Statistical Association*. 83,857-862.

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

[17] Durham, G., and A. R. Gallant, 2002, Numerical Techniques for Maximum Likelihood Estimation of Continuous-time Diffusion Processes, *Journal of Business and Economic Statistics*, 20, 297–316.

[18] Genon-Catalot, V., T. Jeantheau and C. Laredo 2000. Stochastic Volatility Models as Hidden Markov Models and Statistical Applications. *Bernoulli*. 6(6),1051-1079.

[19] Haldrup N., and S. Hyllenberg 1995. A note on the distribution of the least squares estimator of a random walk with drift: some analytical evidence. *Economics Letters*. 48,221-228.

[20] Hall, P., and C. Heyde 1980. *Martingale Limit Theory and Its Application*. Academic Press.

[21] Han, C., Phillips, P.C.B., and D. Sul, 2009, Uniform Asymptotic Normality in Stationary and Unit Root Autoregression. Working Paper, Yale University.

[22] Harvey, D.I., Leybourne, S.J. and Taylor, A.M.R., 2009, Unit root testing in practice: Dealing with uncertainty over the trend and initial condition (with commentaries and rejoinder). *Econometric Theory*, 25, 587-636.

[23] Müller, U.K. and G. Elliott, 2003, Tests for unit roots and the initial condition. *Econometrica,* 71, 1269–86.

[24] Perron P., 1991. A continuous time approximation to the unstable first order autoregressive processes: the case without an intercept. *Econometrica.* 59,211-236.

[25] Phillips, P.C.B., 1987a. Time series regression with a unit root. *Econometrica.* 2,277-301.

[26] Phillips, P.C.B., 1987b. Toward a unified asymptotic theory for autoregression. *Biometrika.* 74,533-547.

[27] Phillips, P.C.B., 1991. To critize the critics: An objective Bayesian analysis of stochastic trends, *Journal of Applied Econometrics,* 6, 333-364.

[28] Phillips, P.C.B., and C. Han, 2008, Gaussian Inference in AR(1) Time Series with or without a Unit Root, *Econometric Theory,* 24, 631-650.

[29] Phillips, Peter C. B., and T. Magdalinos, 2007, Limit theory for moderate deviations from unity, *Journal of Econometrics,* 136, 115-130.

[30] Phillips, P.C.B. and J. Yu, 2005, Jackknifing Bond Option Prices. *Review of Financial Studies*, 18, 707-742.

[31] Phillips, P.C.B. and J. Yu, 2009a, Simulation-based Estimation of Contingent-claims Prices, *Review of Financial Studies*, 22, 3669-3705.

[32] Phillips, P.C.B. and J. Yu, 2009b, A Two-Stage Realized Volatility Approach to Estimation of Diffusion Processes with Discrete Data. *Journal of Econometrics*, 150, 139-150

[33] Pritsker, M., 1998, Nonparametric Density Estimation and Tests of Continuous Time Interest Rate Models. *Review of Financial Studies*, 11, 449-487.

[34] Sim, C. A., 1988, Bayesian skepticism on unit root econometrics. *Journal of Economic Dynamics and Control*, 12, 463-474.

[35] Sim, C. A. and H. Uhlig, 1991, Understanding unit rooters: A helicopter tour. *Econometrica, 59, 1591-1599.*

[36] Stock, J., and M. Watson, 1988. Testing for common trends. *Journal of the American Statistical Association.* 83, 1097-1107.

[37] Shao, J., 2003. *Mathematical Statistics*. Springer.

[38] Sundaresan, S.M., 2000. Continuous-time methods in finance: a review and assessment. *Journal of Finance*. 4,1569-1622.

[39] Tang, C.Y., and S.X. Chen., 2009. Parameter estimation and bias correction for diffusion processes. *Journal of Econometrics*. 149,65-81

[40] Ullah, A., Y. Wang, and J, Yu 2009, Bias in the Mean Reversion Estimator in the Continuous Time Gaussian and Levy Processes, Working Paper, Singapore Management University.

[41] Vasicek, O., 1977, An equilibrium characterization of the term structure. *Journal of Financial Economics*. 5,177-186.

[42] White, H., 2001, *Asymptotic Theory for Econometricians*. Academic Press.

[43] Yoshida, N., 1992, Estimation for diffusion processes from discrete observation. *Journal of Multivariate Analysis*. 41, 220-242.

[44] Yu, J., 2009a, Econometric Analysis of Continuous Time Models: A Survey of Peter Phillips' Work and Some New Results, Working Paper, Singapore Management University.

[45] Yu, J., 2009b, Bias in the Estimation of the Mean Reversion Parameter in Continuous Time Models, Working Paper, Singapore Management University.

Table 1: Percentiles of $T(\hat{\kappa} - \kappa)$ when κ=0.01, $X_0 = 0$

Percentile	0.5%	1%	2.5%	5%	10%	50%	90%	95%	97.5%	99%	99.5%
exact^M	-1.5144	-1.1099	-0.4235	0.1666	0.8846	4.4953	11.6968	15.0534	17.7641	22.4038	24.7566
$(A3)^M$	-1.5187	-1.1365	-0.5000	0.1091	0.8524	4.3583	11.2249	13.9628	16.5831	19.7878	22.1882
$(A1)^M$	-1.1521	-1.0402	-0.8769	-0.7360	-0.5749	0	0.5749	0.7360	0.8769	1.0402	1.1521
$(A2)^M$	-1.1508	-1.0398	-0.8762	-0.7347	-0.5722	0	0.5722	0.7347	0.8762	1.0398	1.1508
exact^W	-1.4801	-1.0809	-0.4250	0.1522	0.8785	4.3577	11.3140	14.2032	16.9497	20.6093	23.8075
$(A3)^W$	-1.5282	-1.0974	-0.3888	0.1401	0.8303	4.3243	11.1558	13.9867	16.6687	20.2992	23.1902
$(A1)^W$	-1.1517	-1.0399	-0.8766	-0.7357	-0.5747	0	0.5747	0.7357	0.8766	1.0399	1.1517
$(A2)^W$	-1.1508	-1.0398	-0.8762	-0.7347	-0.5722	0	0.5722	0.7347	0.8762	1.0398	1.1508
exact^D	-1.4340	-1.0316	-0.3965	0.1143	0.8423	4.3880	11.2112	13.9702	16.5959	20.2595	23.3660
$(A3)^D$	-1.4910	-1.0633	-0.4178	0.1468	0.8681	4.4808	11.6358	14.5646	17.4274	20.7899	23.5020
$(A1)^D$	-1.1516	-1.0398	-0.8766	-0.7357	-0.5747	0	0.5747	0.7357	0.8766	1.0398	1.1516
$(A2)^D$	-1.1508	-1.0398	-0.8762	-0.7347	-0.5722	0	0.5722	0.7347	0.8762	1.0398	1.1508

Note:
1. Data, simulated from the Vasicek model with $\sigma = 0.1, \mu = 0.1$, $\kappa = 0.01$ and $X_0 = 0$, are used to estimate the Vasicek model using LS.
2. The powers, ^{M}, ^{W} and ^{D}, denote statistics calculated from the monthly, weekly and daily data, respectively. $(A1)$, $(A3)$ and $(A2)$ correspond to the long-span, the in-fill and the doubt asymptotics, respectively.
3. Percentiles for the exact distribution and the in-fill asymptotic distribution are based on 10000 times of replications.

Table 2: Percentiles of $T(\hat{\kappa}-\kappa)$ when κ=0.1 and $X_0 = 0$

Percentile	0.5%	1%	2.5%	5%	10%	50%	90%	95%	97.5%	99%	99.5%
exact^M	-1.8713	-1.4383	-0.7649	-0.1663	0.6379	4.5681	12.1090	15.3824	18.7083	23.1338	26.2382
$(A3)^M$	-1.8099	-1.4989	-0.7948	-0.1341	0.5894	4.3725	11.5566	14.3178	17.0514	20.9325	23.4568
$(A1)^M$	-3.6568	-3.3018	-2.7834	-2.3361	-1.8249	0	1.8249	2.3361	2.7834	3.3018	3.6568
$(A2)^M$	-3.6416	-3.288	-2.7719	-2.3264	-1.8173	0	1.8173	2.3264	2.7719	3.288	3.6416
exact^W	-1.9522	-1.4744	-0.8325	-0.2564	0.5681	4.3405	11.4977	14.4374	17.3876	20.9383	23.3842
$(A3)^W$	-1.8813	-1.4065	-0.7990	0.1816	0.5594	4.3148	11.3819	14.2886	17.2340	20.7636	23.6784
$(A1)^W$	-3.6451	-3.2912	-2.7745	-2.3286	-1.8190	0	1.8190	2.3286	2.7745	3.2912	3.6451
$(A2)^W$	-3.6416	-3.2880	-2.7719	-2.3264	-1.8173	0	1.8173	2.3264	2.7719	3.2880	3.6416
exact^D	-1.9257	-1.3799	-0.6801	-0.1366	0.6064	4.3718	11.3478	13.9965	16.9063	20.7398	24.4426
$(A3)^D$	-1.7994	-1.4271	-0.7142	-0.1086	0.6401	4.4752	11.9249	15.0384	17.8352	21.9500	24.5311
$(A1)^D$	-3.6423	-3.2887	-2.7724	-2.3268	-1.8176	0	1.8176	2.3268	2.7724	3.2887	3.6423
$(A2)^D$	-3.6416	-3.2880	-2.7719	-2.3264	-1.8173	0	1.8173	2.3264	2.7719	3.2880	3.6416

Note:
1. Data, simulated from the Vasicek model with $\sigma = 0.1, \mu = 0.1$, $\kappa = 0.1$ and $X_0 = 0$, are used to estimate the Vasicek model using LS.
2. The powers, ^{M}, ^{W} and ^{D}, denote statistics calculated from the monthly, weekly and daily data, respectively. $(A1)$, $(A3)$ and $(A2)$ correspond to the long-span, the in-fill and the doubt asymptotics, respectively.
3. Percentiles for the exact distribution and the infill asymptotic distribution are based on 10000 times of replications.

Table 3: Percentiles of $T(\hat{\kappa}-\kappa)$ when κ=1 and $X_0 = 0$

Percentile	0.5%	1%	2.5%	5%	10%	50%	90%	95%	97.5%	99%	99.5%
exact^M	-5.9833	-5.4156	-4.5158	-3.6159	-2.4666	3.4199	13.1728	16.9335	20.4165	25.0462	29.0258
$(A3)^M$	-5.8584	-5.1766	-4.3099	-3.4872	-2.3805	3.0166	11.4997	14.3912	16.9620	20.8324	23.0600
$(A1)^M$	-12.0127	-10.8464	-9.1436	-7.6741	-5.9947	0	5.9947	7.6741	9.1436	10.8464	12.0127
$(A2)^M$	-11.5158	-10.3977	-8.7654	-7.3567	-5.7467	0	5.7467	7.3567	8.7654	10.3977	11.5158
exact^W	-5.8366	-5.3298	-4.5118	-3.6172	-2.4924	3.0410	12.2778	15.7920	18.8424	22.2474	25.2853
$(A3)^W$	-5.7605	-5.2849	-4.4108	-3.5682	-2.4174	3.1770	11.8582	15.1657	18.2719	21.8524	24.3581
$(A1)^W$	-11.6274	-10.4985	-8.8503	-7.428	-5.8024	0	5.8024	7.428	8.8503	10.4985	11.6274
$(A2)^W$	-11.5158	-10.3977	-8.7654	-7.3567	-5.7467	0	5.7467	7.3567	8.7654	10.3977	11.5158
exact^D	-5.7852	-5.1738	-4.3021	-3.3933	-2.2925	3.1251	11.8937	15.2208	18.3826	22.2328	26.2669
$(A3)^D$	-5.8282	-5.3738	-4.3753	-3.4309	-2.2804	3.2532	12.3750	15.4324	18.9931	23.3231	25.8484
$(A1)^D$	-11.5386	-10.4184	-8.7828	-7.3713	-5.7581	0	5.7581	7.3713	8.7828	10.4184	11.5386
$(A2)^D$	-11.5158	-10.3977	-8.7654	-7.3567	-5.7467	0	5.7467	7.3567	8.7654	10.3977	11.5158

Note:
1. Data, simulated from the Vasicek model with $\sigma = 0.1, \mu = 0.1$, $\kappa = 1$ and $X_0 = 0$, are used to estimate the Vasicek model using LS.
2. The powers, ^{M}, ^{W} and ^{D}, denote statistics calculated from the monthly, weekly and daily data, respectively. $(A1)$, $(A3)$ and $(A2)$ correspond to the long-span, the in-fill and the doubt asymptotics, respectively.
3. Percentiles for the exact distribution and the in-fill asymptotic distribution are based on 10000 times of replications.

Table 4: Percentiles of $T(\hat{\kappa}-\kappa)$ when κ=0.01, $\mu = 0.1$, $\sigma = 0.1$, and $X_0 \sim N(\mu, \sigma^2/2\kappa)$

Percentile	0.5%	1%	2.5%	5%	10%	50%	90%	95%	97.5%	99%	99.5%
exact^M	-1.3763	-1.0314	-0.4091	0.1743	0.8979	4.4736	11.6923	14.7856	17.6383	21.2153	24.4151
$(A3)^M$	-1.5073	-1.1364	-0.5154	0.1156	0.8443	4.3600	11.1829	14.0644	16.5516	19.6562	22.3832
$(A1)^M$	-1.1521	-1.0402	-0.8769	-0.736	-0.5749	0	0.5749	0.736	0.8769	1.0402	1.1521
$(A2)^M$	-1.1508	-1.0398	-0.8762	-0.7347	-0.5722	0	0.5722	0.7347	0.8762	1.0398	1.1508
exact^W	-1.4790	-1.0937	-0.3799	0.1677	0.8355	4.3711	11.3132	14.2141	16.8817	20.6378	23.4944
$(A3)^W$	-1.5183	-1.1096	-0.3889	0.1367	0.8197	4.3274	11.1589	13.9485	16.6963	20.2543	23.2539
$(A1)^W$	-1.1517	-1.0399	-0.8766	-0.7357	-0.5747	0	0.5747	0.7357	0.8766	1.0399	1.1517
$(A2)^W$	-1.1508	-1.0398	-0.8762	-0.7347	-0.5722	0	0.5722	0.7347	0.8762	1.0398	1.1508
exact^D	-1.4952	-1.0347	-0.4143	0.1559	0.8795	4.4741	11.6296	14.6358	17.5448	20.8621	24.0828
$(A3)^D$	-1.4950	-1.0478	-0.4064	0.1531	0.8725	4.4805	11.6052	14.6322	17.3911	20.7465	23.9938
$(A1)^D$	-1.1516	-1.0398	-0.8766	-0.7357	-0.5747	0	0.5747	0.7357	0.8766	1.0398	1.1516
$(A2)^D$	-1.1508	-1.0398	-0.8762	-0.7347	-0.5722	0	0.5722	0.7347	0.8762	1.0398	1.1508

Note:
1. Data, simulated from the Vasicek model with $\sigma = 0.1, \mu = 0.1$, $\kappa = 0.01$ and $X_0 \sim N(\mu, \sigma^2/2\kappa)$, are used to estimate the Vasicek model using LS.
2. The powers, ^{M}, ^{W} and ^{D}, denote statistics calculated from the monthly, weekly and daily data, respectively. $(A1)$, $(A3)$ and $(A2)$ correspond to the long-span, the in-fill and the doubt asymptotics, respectively.
3. Percentiles for the exact distribution and the in-fill asymptotic distribution are based on 10000 times of replications.

Table 5: Percentiles of $T(\hat{\kappa}-\kappa)$ when κ=0.1, $\mu = 0.1$, $\sigma = 0.1$, and $X_0 \sim N(\mu, \sigma^2/2\kappa)$

Percentile	0.5%	1%	2.5%	5%	10%	50%	90%	95%	97.5%	99%	99.5%
exact^M	-1.7676	-1.4004	-0.6954	-0.0581	0.7681	4.5518	12.2125	15.3651	18.2946	22.4320	25.1627
$(A3)^M$	-1.8428	-1.4586	-0.7960	-0.1221	0.6648	4.3950	11.6004	14.2922	17.1580	20.8494	23.0308
$(A1)^M$	-3.6568	-3.3018	-2.7834	-2.3361	-1.8249	0	1.8249	2.3361	2.7834	3.3018	3.6568
$(A2)^M$	-3.6416	-3.288	-2.7719	-2.3264	-1.8173	0	1.8173	2.3264	2.7719	3.288	3.6416
exact^W	-1.8282	-1.3872	-0.6835	-0.1256	0.6906	4.3916	11.6948	14.3183	17.1894	21.4618	24.6075
$(A3)^W$	-1.8881	-1.3992	-0.7037	-0.1412	0.6351	4.3582	11.5612	14.2489	16.9576	21.2970	24.2865
$(A1)^W$	-3.6451	-3.2912	-2.7745	-2.3286	-1.8190	0	1.8190	2.3286	2.7745	3.2912	3.6451
$(A2)^W$	-3.6416	-3.2880	-2.7719	-2.3264	-1.8173	0	1.8173	2.3264	2.7719	3.2880	3.6416
exact^D	-1.8415	-1.3762	-0.6445	-0.0644	0.6855	4.5450	12.0387	15.1652	17.8393	22.0280	24.9007
$(A3)^D$	-1.8435	-1.3763	-0.6611	-0.0700	0.6911	4.5500	11.9984	15.1573	17.7976	22.0178	24.7715
$(A1)^D$	-3.6423	-3.2887	-2.7724	-2.3268	-1.8176	0	1.8176	2.3268	2.7724	3.2887	3.6423
$(A2)^D$	-3.6416	-3.288	-2.7719	-2.3264	-1.8173	0	1.8173	2.3264	2.7719	3.2880	3.6416

Note:
1. Data, simulated from the Vasicek model with $\sigma = 0.1, \mu = 0.1$, $\kappa = 0.1$ and $X_0 \sim N(\mu, \sigma^2/2\kappa)$, are used to estimate the Vasicek model using LS.
2. The powers, ^{M}, ^{W} and ^{D}, denote statistics calculated from the monthly, weekly and daily data, respectively. $(A1)$, $(A3)$ and $(A2)$ correspond to the long-span, the in-fill and the doubt asymptotics, respectively.
3. Percentiles for the exact distribution and the infill asymptotic distribution are based on 10000 times of replications.

Table 6: Percentiles of $T(\hat{\kappa}-\kappa)$ when κ=1, $\mu = 0.1$, $\sigma = 0.1$, and $X_0 \sim N(\mu, \sigma^2/2\kappa)$

Percentile	0.5%	1%	2.5%	5%	10%	50%	90%	95%	97.5%	99%	99.5%
exact^M	-6.2242	-5.5086	-4.5765	-3.6693	-2.4706	3.6723	14.0061	18.0038	22.0170	26.5343	30.8223
$(A3)^M$	-5.8935	-5.3139	-4.4239	-3.5500	-2.4283	3.2503	12.2221	15.5332	18.5478	22.0916	25.2350
$(A1)^M$	-12.0127	-10.8464	-9.1436	-7.6741	-5.9947	0	5.9947	7.6741	9.1436	10.8464	12.0127
$(A2)^M$	-11.5158	-10.3977	-8.7654	-7.3567	-5.7467	0	5.7467	7.3567	8.7654	10.3977	11.5158
exact^W	-6.0139	-5.3174	-4.5273	-3.6070	-2.3606	3.5724	13.0060	16.4764	20.2147	23.8976	27.2821
$(A3)^W$	-5.8653	-5.2880	-4.5187	-3.6009	-2.3654	3.4840	12.6226	16.0858	19.5128	23.0327	26.0710
$(A1)^W$	-11.6274	-10.4985	-8.8503	-7.4280	-5.8024	0	5.8024	7.4280	8.8503	10.4985	11.6274
$(A2)^W$	-11.5158	-10.3977	-8.7654	-7.3567	-5.7467	0	5.7467	7.3567	8.7654	10.3977	11.5158
exact^D	-5.9303	-5.3399	-4.4215	-3.5354	-2.2616	3.5449	13.2887	16.9237	20.3982	24.6560	27.8147
$(A3)^D$	-5.8146	-5.3567	-4.3562	-3.4135	-2.3017	3.2322	12.3168	15.3456	18.8294	23.2107	25.7605
$(A1)^D$	-11.5386	-10.4184	-8.7828	-7.3713	-5.7581	0	5.7581	7.3713	8.7828	10.4184	11.5386
$(A2)^D$	-11.5158	-10.3977	-8.7654	-7.3567	-5.7467	0	5.7467	7.3567	8.7654	10.3977	11.5158

Note:

1. Data, simulated from the Vasicek model with $\sigma = 0.1, \mu = 0.1$, $\kappa = 1$ and $X_0 \sim N(\mu, \sigma^2/2\kappa)$, are used to estimate the Vasicek model using LS.
2. The powers, ^{M}, ^{W} and ^{D}, denote statistics calculated from the monthly, weekly and daily data, respectively. $(A1)$, $(A3)$ and $(A2)$ correspond to the long-span, the in-fill and the doubt asymptotics, respectively.
3. Percentiles for the exact distribution and the infill asymptotic distribution are based on 10000 times of replications.

Table 7. Summary statistics and unit root tests for monthly Federal fund rates

Number of Observations	432
Mean	0.0698
Standard Deviation	0.0319
Autocorrelation ρ_1	0.977
Autocorrelation ρ_2	0.939
Autocorrelation ρ_3	0.901
Autocorrelation ρ_4	0.868
Autocorrelation ρ_5	0.841
Autocorrelation ρ_6	0.817
Autocorrelation ρ_7	0.797
$Z(t)$ test	-2.53
10% critical value	-2.57
P value	0.1081

Table 8. Estimate of κ, and 90% and 95% confidence intervals

	(A1)		(A2)		(A3)
	$\kappa > 0$	$\kappa = 0$	$\kappa > 0$	$\kappa = 0$	
90% CI	(0.0609, 0.4616)	(-0.1277, 0.2576)	(0.0631, 0.4594)	(-0.1277, 0.2576)	(-0.1579 0.3551)
95% CI	(0.0225, 0.4999)	(-0.2054, 0.2729)	(0.0251, 0.4973)	(-0.2054, 0.2729)	(-0.2430, 0.3795)