Economics Letters 128 (2015) 1–5

Contents lists available at ScienceDirect

Economics Letters

journal homepage: www.elsevier.com/locate/ecolet

# Asymptotic theory for linear diffusions under alternative sampling schemes[☆]

Qiankun Zhou[a], Jun Yu[b,*]

[a] University of Southern California, United States

[b] Singapore Management University, Singapore

## HIGHLIGHTS

- Discuss three possibilities to make inference about the persistence parameter in diffusion models.
- Develop the in-fill asymptotic distribution for the persistence parameter.
- Show that the in-fill asymptotic distribution gives a more accurate approximation using simulated data.
- Use the in-fill asymptotic distribution for unit root testing using real data.

## ARTICLE INFO

*Article history:*
Received 15 August 2014
Received in revised form
28 November 2014
Accepted 9 December 2014
Available online 9 January 2015

*JEL classification:*
C12
C22
G12

*Keywords:*
Vasicek model
In-fill asymptotics
Long-span asymptotics
Double asymptotics
Unit root test

## ABSTRACT

The asymptotic distributions of the maximum likelihood estimator of the persistence parameter are developed in a linear diffusion model under three sampling schemes, long-span, in-fill and double. Simulations suggest that the in-fill asymptotic distribution gives a more accurate approximation to the finite sample distribution than the other two distributions. An empirical application highlights the difference in unit root testing based on the alternative asymptotic distributions.



## 1. Introduction

Consider the Ornstein–Uhlenbeck (OU) diffusion process which was first used by Vasicek (1977) to describe the movement of interest rates over time,

$$dX(t) = \kappa(\mu - X(t))dt + \sigma dW(t), \quad X(0) = X_0, \tag{1}$$

where $W(t)$ is a standard Brownian motion, $\mu$ is the intercept, $\kappa$ captures the persistence of $X(t)$, and $\sigma$ is the instantaneous volatility. The same model was also used to model the dynamic behavior of log volatility; see for example, Hull and White (1987). In practice, $X(t)$ is observed at discrete points in time, say $t = 0, \delta, 2\delta, \dots, n\delta (:= T)$, where $n$ is the sample size, $\delta$ the sampling interval, and $T$ the time span of the data. In this paper, we develop and compare the asymptotic theory for the maximum likelihood (ML) estimator of $\kappa$ under the following three sampling schemes:

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

$$T \to \infty, \quad \delta \to 0 \text{ and hence } n \to \infty \tag{A2}$$

$$\delta \to 0, \quad T \text{ is fixed and hence } n \to \infty. \tag{A3}$$

Obviously the continuous time model facilitates the discussions and derivations of three alternative sampling schemes.

Scheme (A1) assumes that the sampling interval is fixed and the sample size increases as the time span increases. This widely

☆ Yu gratefully acknowledge financial support from the Ministry of Education AcRF Tier 2 fund under Grant No. MOE2011-T2-2-096.

* Correspondence to: School of Economics and Lee Kong Chian School of Business, Singapore Management University, 90 Stamford Road, Singapore 178903, Singapore. Tel.: +65 68280858.

*E-mail address:* yujun@smu.edu.sg (J. Yu).

http://dx.doi.org/10.1016/j.econlet.2014.12.015
0165-1765/

adopted sampling scheme in time series analysis leads to the *long-span* asymptotics. In practical applications in economics, $T$ measures the number of years. Typical values for $T$ are not very large (between 1 and 50). In some cases, even if $T$ is large, a smaller value for $T$ may be used to avoid possible structural breaks in Model (1). Scheme (A3) allows the sample size to go to infinity by decreasing the sampling interval but keeping the time span fixed. This sampling scheme leads to the *in-fill* asymptotics. In practice, 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. Scheme (A2) combines both the long-span scheme and the in-fill scheme and leads to the *double* asymptotics.

For empirically reasonable values of $\delta$, $T$ and parameters, what we find in the present paper is that among the three asymptotic distributions the in-fill distribution provides the best approximation to the finite sample distribution. Moreover, like the finite sample distribution but unlike the other two asymptotic distributions, the in-fill distribution is continuous in $\kappa$ when the persistence of the process passes the unity (i.e. when $\kappa = 0$) and thus the confidence regions based on the in-fill distribution are connected. The property of continuity provides an answer to the Bayesian criticism to the unit root asymptotics; see for example, Sims (1988) and Sims and Uhlig (1991). Consequently, we advocate doing unit test based on the in-fill asymptotics.

## 2. The model and the asymptotics

The exact discrete time model corresponding to (1) is a first order autoregressive (AR(1)) model with intercept:

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

$$\phi = e^{-\kappa\delta},\ X_{0\delta} = X_0,\ \epsilon_t \overset{\text{i.i.d}}{\sim} N(0,1). \tag{2}$$

When there is no confusion, we simply write $X_{t\delta}$ as $X_t$. The ML estimator of $\phi$ is: $\hat{\phi} = \sum(X_{t-1}-\overline{X}_-)(X_t-\overline{X})/\sum(X_{t-1}-\overline{X})^2$, where $\sum := \sum_{t=1}^{n}$, $\overline{X}_- = \frac{1}{n}\sum X_{t-1}$, and $\overline{X} = \frac{1}{n}\sum X_t$. Correspondingly, the ML estimator of $\kappa$ is $\hat{\kappa} = -\ln\hat{\phi}/\delta$.

### 2.1. The long-span asymptotics

When $\kappa > 0$ (i.e. $\phi < 1$), this is the stationary case and the long-span asymptotic distribution of $\hat{\kappa}$ (i.e. under Scheme (A1)) is known as (see for example, Tang and Chen, 2009):

$$\sqrt{T}(\hat{\kappa} - \kappa) \xrightarrow{d} N\left(0, \frac{e^{2\kappa\delta} - 1}{\delta}\right), \quad \text{as } T \to \infty. \tag{3}$$

When $\kappa$ is close to zero, $\phi$ is close to 1 and hence the root of the AR(1) model (2) is near unity. It is known in the literature that when the root is near unity, the finite sample distribution of $\hat{\phi}$ is not close to the normal distribution; see for example Ahtola and Tiao (1984). The same feature applies to $\kappa$ when $\kappa$ is close to 0 and $\mu$ is known in Model (1); see for example, Yu (2012, 2014). We expect this property of discrepancy holds true when $\mu$ is unknown.

When $\kappa = 0$ (i.e. $\phi = 1$), this is the unit root case and the long-span asymptotic distribution of $\hat{\kappa}$ is known as (see for example, Phillips, 1987a,b):

$$T(\hat{\kappa} - 0) \xrightarrow{d} -\frac{\int WdW - W(1)\int W}{\int W^2 - \left[\int W\right]^2}, \quad \text{as } T \to \infty, \tag{4}$$

where $W(r)$ is a standard Brownian motion with $r \in [0, 1]$ and $\int = \int_0^1$.

Clearly, under Scheme (A1) both the asymptotic distribution of $\hat{\kappa}$ and its rate of convergence depend on the true value of $\kappa$, i.e., whether $\kappa = 0$ or $\kappa > 0$. This feature of discontinuity is the same as the well-known discontinuity in the long-span asymptotic theory of $\widehat{\phi}$ and suggests that the confidence intervals obtained from the asymptotic distributions ((3) and (4)) are disjoint pieces (Sims, 1988). On the other hand, the confidence intervals obtained from the finite sample distribution of $\hat{\kappa}$ is always connected because the finite sample distribution is continuous in $\kappa$. This observation has generated some criticisms on making inference of persistence based on the nonstationary asymptotic theory (Sims and Uhlig, 1991). See also the critique of the criticisms (Phillips, 1991).

### 2.2. The double asymptotics

When $\kappa > 0$, taking $\delta$ towards 0 in (3), we obtain the following double asymptotic distribution of $\kappa$:

$$\sqrt{T}(\hat{\kappa} - \kappa) \xrightarrow{d} N(0, 2\kappa), \quad \text{as } T \to \infty \text{ and } \delta \to 0. \tag{5}$$

If $\kappa = 0$, the double asymptotic distribution is

$$T(\hat{\kappa} - 0) \xrightarrow{d} -\frac{\int WdW - W(1)\int W}{\int W^2 - \left[\int W\right]^2},$$
$$\text{as } T \to \infty \text{ and } \delta \to 0. \tag{6}$$

As in Scheme (A1), under Scheme (A2), the asymptotic distribution of $\hat{\kappa}$ and its rate of convergence depend on the true value of $\kappa$. So the discontinuity of the asymptotic distribution of $\hat{\kappa}$ in $\kappa$ remains true under the double asymptotic scheme.

### 2.3. The in-fill asymptotics

Perron (1991) and Yu (2014) have obtained some interesting results for the OU process with a known $\mu$. In particular, Perron derived the in-fill asymptotic distribution of $\kappa$ which depends on the initial condition and holds true for all values of $\kappa$. Yu compared the performance of this in-fill asymptotic distribution with the long-span and double asymptotic distributions and found that the in-fill distribution can better approximate the finite sample distribution when $\kappa$ is close to 0. However, when $\mu$ is unknown the in-fill asymptotic distribution has not been derived in the literature. Theorem 2.1 presents the result and the proof of the theorem may be found in the working paper version of Zhou and Yu (2010).

**Theorem 2.1.** *For Model (1), under Scheme (A3), regardless of the value of $\kappa$, the in-fill asymptotic distribution of $\hat{\kappa}$ is*

$$T(\hat{\kappa} - \kappa) \xrightarrow{d} -\frac{A(\gamma_0, c)}{B(\gamma_0, c)}, \quad \text{as } \delta \to 0 \tag{7}$$

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

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

**Remark 2.1.** The in-fill asymptotic theory in (7) is analogous to that in Theorem 1 of Perron (1991). The asymptotic distribution is continuous in $\kappa$ for all values of $\kappa$, including $\kappa = 0$. It depends on the initial condition, the intercept parameter $\mu$, and the time span $T$. Since both the in-fill distribution and the finite sample distribution depend on the initial condition, we expect the in-fill distribution performs better than the other two asymptotic distributions.

**Remark 2.2.** If $\kappa \to 0$ and $X_0 = 0$, the numerator in (7) becomes

$$\frac{b}{c}\int c_1 dW(r) + \int J_c(r)dW(r) - \int dW(r)\left(c_2 b + \int J_c(r)dr\right)$$
$$\to b\int\left(r-\frac{1}{2}\right)dW(r) + \int W(r)dW(r) - \int dW(r)\int W(r)dr,$$

and the denominator becomes

$$c_3 b^2 + \frac{2b}{c}\int c_1 J_c(r)dr - \left(c_2 b + \int J_c(r)dr\right)^2 + \int J_c^2(r)dr$$
$$\to \frac{b^2}{12} + 2b\int\left(r-\frac{1}{2}\right)W(r)dr + \int W^2(r)dr - \left(\int W(r)dr\right)^2.$$

By the *Delta* method, the in-fill asymptotic distribution of $\hat{\phi}$ is

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

This distribution is the same as that obtained in Haldrup and Hylleberg (1995).

**Remark 2.3.** As it was shown earlier, there is a discontinuity in the long-span asymptotics of both $\widehat{\phi}$ (at $\phi = 1$) and $\widehat{\kappa}$ (at $\kappa = 0$), creating difficulties in unit root testing. However, the unit root tests can be also performed in continuous time based on the in-fill asymptotics. Since the in-fill asymptotic distribution is continuous in $\kappa$, it provides a unified framework to make statistical inference about $\kappa$. In particular, when $\kappa$ is close to 0 the limiting distribution in (7) is skewed and behaves similarly to the unit root limiting distribution. Consequently, our answer to the Bayesian criticisms is that the disconnecting confidence intervals are caused by the poor approximation of the long-span asymptotic distribution (3) and the double asymptotic distribution (5) to the finite sample distribution and the confidence interval, if constructed from the in-fill asymptotic distribution, is connected. Extensive simulations will be carried out in the next section to support this claim.

## 3. Monte Carlo simulations

We design several Monte Carlo experiments to compare the accuracy of the alternative asymptotic distributions of $\widehat{\kappa}$ for approximating the finite sample distribution. To obtain the density of the limiting distributions, we use the method proposed by Chan (1988). According to Chan, the in-fill asymptotic distributions expressed in (7) may be approximated by the Riemann sums and $dW(r)$ by $\epsilon_i/\sqrt{n}$, where $\{\epsilon_i\}$ is a sequence of standard normal random variables, and $n$ the sample size. Consequently, the limiting distribution $\int J_c(r)dW(r)/\int J_c^2(r)dr$ may be approximated by $n\left(\sum_{i=1}^{n}\sum_{k=1}^{i} e^{c(i-k)/n}\epsilon_k\epsilon_{i+1}\right)/\left(\sum_{i=1}^{n}\left(\sum_{k=1}^{i} e^{c(i-k)/n}\epsilon_k\right)^2\right)$. In this paper we choose $n = 10{,}000$ to approximate the Wiener process.

The true value of $\kappa$ is set at 0.01, 0.1 and 1, respectively. The first two values are empirically realistic for interest rates while the last value is empirically realistic for volatility. The true value of $\mu$ is set to 0.1, $\sigma$ to 0.1 and $X_0 \sim N(\mu, \sigma^2/2\kappa)$. The value of the sampling interval $\delta$ 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 number of replications is set at 10,000.

The Monte Carlo simulation results, including the 1%, 5%, 10%, 50%, 90%, 95%, and 99% percentiles of the four distributions (i.e., the finite sample distribution, the asymptotic distributions developed under Schemes (A1), (A2) and (A3)) based on $T(\hat{\kappa}-\kappa)$, are reported in Tables 1–3, for $\kappa = 0.01, 0.1, 1$, respectively.

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 estimation and the power of 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 $\kappa$ and to the initial condition. The smaller the value of $\kappa$, the more sensitive the percentiles to the initial condition. This feature is related to the role that the initial condition plays in unit root tests; see, for example, Phillips (1987a) and Müller and Elliott (2003). Second, normality always provides inaccurate approximations of the finite sample distribution, suggesting that (A1) and (A2) should not be used in practice to make statistical inference of $\kappa$. The percentiles of the limiting distribution under Schemes (A1) and (A2) are very different from those of 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 double 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 sampling interval, 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 double asymptotic distributions. Finally, the in-fill asymptotic distribution is less accurate when $\kappa$ and $\delta$ become larger and hence a root is further away from unity. However, the in-fill asymptotic distribution continues to outperform the long-span and the double asymptotic distributions.

To examine the impact of initial condition on the asymptotic approximations, we repeat the above experiment only for $\kappa = 0.1$ but fix $X_0 = 0.1$. This initial condition is different from the mean of the marginal distribution. Table 4 shows the 1%, 5%, 10%, 50%, 90%, 95%, and 99% percentiles of the four distributions. Interestingly, the difference between the finite sample distribution and the two asymptotic normal approximations become bigger than the case when $X_0 \sim N(\mu, \sigma^2/2\kappa)$ (i.e. Table 2). However, the in-fill asymptotic distribution continues to provide much more adequate approximations to the finite sample distribution than the other two asymptotic distributions. This finding is not surprising as the in-fill asymptotic distribution depends on the initial condition while the two asymptotic normal approximations do not.

## 4. An empirical application

In this section, we apply the alternative asymptotic theory based on real monthly time series data on short term interest rates. 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

**Table 1**
Percentiles of $T\left(\hat{\kappa}-\kappa\right)$ when $\kappa=0.01$, $\mu=0.1$, $\sigma=0.1$, and $X_0\sim N(\mu,\sigma^2/2\kappa)$.

| Percentile | 1% | 5% | 10% | 50% | 90% | 95% | 99% |
|---|---|---|---|---|---|---|---|
| exact$^M$ | −1.1271 | 0.1504 | 0.8658 | 4.4434 | 11.7888 | 14.7151 | 21.6592 |
| (A3)$^M$ | −1.8638 | −0.3987 | 0.3551 | 3.9760 | 12.1459 | 15.8847 | 25.0705 |
| (A1)$^M$ | −1.0402 | −0.7360 | −0.5749 | 0 | 0.5749 | 0.7360 | 1.0402 |
| (A2)$^M$ | −1.0398 | −0.7347 | −0.5722 | 0 | 0.5722 | 0.7347 | 1.0398 |
| exact$^W$ | −1.1123 | 0.0904 | 0.8315 | 4.3354 | 11.2763 | 14.4341 | 20.8579 |
| (A3)$^W$ | −1.2444 | 0.0323 | 0.7148 | 4.1809 | 11.1811 | 14.2994 | 21.9489 |
| (A1)$^W$ | −1.0399 | −0.7357 | −0.5747 | 0 | 0.5747 | 0.7357 | 1.0399 |
| (A2)$^W$ | −1.0398 | −0.7347 | −0.5722 | 0 | 0.5722 | 0.7347 | 1.0398 |
| exact$^D$ | −1.0901 | 0.0594 | 0.7844 | 4.3352 | 11.1289 | 13.7981 | 20.1534 |
| (A3)$^D$ | −1.1520 | 0.0877 | 0.8015 | 4.3983 | 11.6374 | 14.6220 | 21.0115 |
| (A1)$^D$ | −1.0398 | −0.7357 | −0.5747 | 0 | 0.5747 | 0.7357 | 1.0398 |
| (A2)$^D$ | −1.0398 | −0.7347 | −0.5722 | 0 | 0.5722 | 0.7347 | 1.0398 |

Note: The superscripts, $M$, $W$ and $D$, denote statistics calculated from the monthly, weekly and daily data, respectively. (A3), (A1) and (A2) correspond to the long-span, the in-fill and the double asymptotics, respectively.

**Table 2**
Percentiles of $T\left(\hat{\kappa}-\kappa\right)$ when $\kappa=0.1$, $\mu=0.1$, $\sigma=0.1$, and $X_0\sim N(\mu,\sigma^2/2\kappa)$.

| Percentile | 1% | 5% | 10% | 50% | 90% | 95% | 99% |
|---|---|---|---|---|---|---|---|
| exact$^M$ | −1.5283 | −0.2825 | 0.5070 | 4.2512 | 11.7567 | 14.7229 | 22.2292 |
| (A3)$^M$ | −1.5624 | −0.2878 | 0.4333 | 4.0707 | 11.2104 | 14.0524 | 20.9886 |
| (A1)$^M$ | −3.3018 | −2.3361 | −1.8249 | 0 | 1.8249 | 2.3361 | 3.3018 |
| (A2)$^M$ | −3.2880 | −2.3264 | −1.8173 | 0 | 1.8173 | 2.3264 | 3.2880 |
| exact$^W$ | −1.6288 | −0.2974 | 0.4416 | 4.0834 | 11.3671 | 14.3156 | 21.2052 |
| (A3)$^W$ | −1.5715 | −0.3574 | 0.4079 | 4.0211 | 11.1196 | 14.0089 | 20.2376 |
| (A1)$^W$ | −3.2912 | −2.3286 | −1.8190 | 0 | 1.8190 | 2.3286 | 3.2912 |
| (A2)$^W$ | −3.2880 | −2.3264 | −1.8173 | 0 | 1.8173 | 2.3264 | 3.2880 |
| exact$^D$ | −1.5512 | −0.3467 | 0.3867 | 4.1199 | 11.0154 | 13.8110 | 21.0254 |
| (A3)$^D$ | −1.4917 | −0.2597 | 0.4643 | 4.1457 | 11.5266 | 14.4463 | 20.9275 |
| (A1)$^D$ | −3.2887 | −2.3268 | −1.8176 | 0 | 1.8176 | 2.3268 | 3.2887 |
| (A2)$^D$ | −3.2880 | −2.3264 | −1.8173 | 0 | 1.8173 | 2.3264 | 3.2880 |

See notes of Table 1.

**Table 3**
Percentiles of $T\left(\hat{\kappa}-\kappa\right)$ when $\kappa=1$, $\mu=0.1$, $\sigma=0.1$, and $X_0\sim N(\mu,\sigma^2/2\kappa)$.

| Percentile | 1% | 5% | 10% | 50% | 90% | 95% | 99% |
|---|---|---|---|---|---|---|---|
| exact$^M$ | −5.4859 | −3.6285 | −2.4241 | 3.6119 | 13.7670 | 17.6815 | 27.0131 |
| (A3)$^M$ | −5.2775 | −3.5098 | −2.4004 | 3.1535 | 11.9858 | 14.8782 | 21.3779 |
| (A1)$^M$ | −10.8464 | −7.6741 | −5.9947 | 0 | 5.9947 | 7.6741 | 10.8464 |
| (A2)$^M$ | −10.3977 | −7.3567 | −5.7467 | 0 | 5.7467 | 7.3567 | 10.3977 |
| exact$^W$ | −5.3858 | −3.6281 | −2.4647 | 3.3337 | 12.6606 | 16.4740 | 24.0561 |
| (A3)$^W$ | −5.3106 | −3.5741 | −2.3879 | 3.2850 | 12.3413 | 15.7487 | 22.5331 |
| (A1)$^W$ | −10.4985 | −7.4280 | −5.8024 | 0 | 5.8024 | 7.4280 | 10.4985 |
| (A2)$^W$ | −10.3977 | −7.3567 | −5.7467 | 0 | 5.7467 | 7.3567 | 10.3977 |
| exact$^D$ | −5.2109 | −3.4059 | −2.3081 | 3.3241 | 12.3800 | 15.7653 | 23.3141 |
| (A3)$^D$ | −5.3526 | −3.4908 | −2.2837 | 3.4331 | 12.7952 | 16.1574 | 23.8631 |
| (A1)$^D$ | −10.4184 | −7.3713 | −5.7581 | 0 | 5.7581 | 7.3713 | 10.4184 |
| (A2)$^D$ | −10.3977 | −7.3567 | −5.7467 | 0 | 5.7467 | 7.3567 | 10.3977 |

See notes of Table 1.

**Table 4**
Percentiles of $T\left(\hat{\kappa}-\kappa\right)$ when $\kappa=0.1$, $\mu=0.1$, $\sigma=0.1$, and $X_0=0.1$.

| Percentile | 1% | 5% | 10% | 50% | 90% | 95% | 99% |
|---|---|---|---|---|---|---|---|
| exact$^M$ | −1.4311 | −0.0849 | 0.7183 | 4.6260 | 12.3435 | 15.3660 | 23.4456 |
| (A3)$^M$ | −1.4552 | −0.1171 | 0.6667 | 4.3946 | 11.5336 | 14.3703 | 20.4988 |
| (A1)$^M$ | −3.3018 | −2.3361 | −1.8249 | 0 | 1.8249 | 2.3361 | 3.3018 |
| (A2)$^M$ | −3.2880 | −2.3264 | −1.8173 | 0 | 1.8173 | 2.3264 | 3.2880 |
| exact$^W$ | −1.4102 | −0.1402 | 0.6721 | 4.4192 | 11.8036 | 14.5898 | 21.1528 |
| (A3)$^W$ | −1.3698 | −0.1320 | 0.6340 | 4.3541 | 11.5998 | 14.2569 | 20.9383 |
| (A1)$^W$ | −3.2912 | −2.3286 | −1.8190 | 0 | 1.8190 | 2.3286 | 3.2912 |
| (A2)$^W$ | −3.2880 | −2.3264 | −1.8173 | 0 | 1.8173 | 2.3264 | 3.2880 |
| exact$^D$ | −1.3563 | −0.0609 | 0.6526 | 4.4665 | 11.4818 | 14.3163 | 21.0762 |
| (A3)$^D$ | −1.3741 | −0.0626 | 0.6802 | 4.5302 | 11.9893 | 15.1922 | 21.9786 |
| (A1)$^D$ | −3.2887 | −2.3268 | −1.8176 | 0 | 1.8176 | 2.3268 | 3.2887 |
| (A2)$^D$ | −3.2880 | −2.3264 | −1.8173 | 0 | 1.8173 | 2.3264 | 3.2880 |

See notes of Table 1.

**Table 5**
Estimate of $\kappa$, 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) |

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).

Assuming $X_0$ is the same as the first observation, the ML estimates of the three parameters $\kappa$, $\mu$ and $\sigma$ 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 $\kappa$ under the three schemes, which are reported in Table 5. Under Schemes (A1) and (A2), the limiting distribution when $\kappa > 0$ is different from the limiting distribution when $\kappa = 0$. So two sets of confidence intervals are reported in the two cases.

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 (1996) 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) with $\kappa > 0$, 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) with $\kappa = 0$ 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 form of the in-fill asymptotic distribution is the same, whether $\kappa = 0$ or not. Hence, only one confidence interval is needed regardless of $\kappa$. 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 $\kappa$. Consequently, we believe the empirical result based on Scheme (A3) is more reliable.

## References


Ahtola, J., Tiao, G.C., 1984. Parameter inference for a nearly nonstationary first order autoregressive model. Biometrika 71, 263–272.

Aït-Sahalia, Y., 1996. Testing continuous-time models of spot interest rate derivative securities. Rev. Financ. Stud. 9, 385–426.

Aït-Sahalia, Y., 1999. Transition densities for interest rate and other non-linear diffusions. J. Finance 54, 1361–1395.

Chan, N.H., 1988. The parameter inference for nearly nonstationary time series. J. Amer. Statist. Assoc. 83, 857–862.

Haldrup, N., Hylleberg, S., 1995. A note on the distribution of the least squares estimator of a random walk with drift: some analytical evidence. Econom. Lett. 48, 221–228.

Hull, J., White, A., 1987. The pricing of options on assets with stochastic volatilities. J. Finance 42, 281–300.

Müller, U.K., Elliott, G., 2003. Tests for unit roots and the initial condition. Econometrica 71, 1269–1286.

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

Phillips, P.C.B., 1987a. Time series regression with a unit root. Econometrica 2, 277–301.

Phillips, P.C.B., 1987b. Toward a unified asymptotic theory for autoregression. Biometrika 74, 533–547.

Phillips, P.C.B., 1991. To criticize the critics: an objective Bayesian analysis of stochastic trends. J. Appl. Econometrics 6, 333–364.

Phillips, P.C.B., Yu, J., 2005. Jackknifing bond option prices. Rev. Financ. Stud. 18, 707–742.

Sims, C.A., 1988. Bayesian skepticism on unit root econometrics. J. Econom. Dynam. Control 12, 463–474.

Sims, C.A., Uhlig, H., 1991. Understanding unit rooters: a helicopter tour. Econometrica 59, 1591–1599.

Stock, J., Watson, M., 1988. Testing for common trends. J. Amer. Statist. Assoc. 83, 1097–1107.

Tang, C.Y., Chen, S.X., 2009. Parameter estimation and bias correction for diffusion processes. J. Econometrics 149, 65–81.

Vasicek, O., 1977. An equilibrium characterization of the term structure. J. Financ. Econ. 5, 177–186.

Yu, J., 2012. Bias in the estimation of the mean reversion parameter in continuous time models. J. Econometrics 169, 114–122.

Yu, J., 2014. Econometric analysis of continuous time models: a survey of Peter Phillips' work and some new results. Econometric Theory 30, 737–774.

Zhou, Q., Yu, J., 2010. Asymptotic distributions of the least squares estimator for diffusion processes. Working Paper 02-2010, Centre for Financial Econometrics, Singapore Management University.