Consistency of the KPSS unit root test against fractionally integrated alternative

Hyung S. Lee\textsuperscript{a}, Christine Amsler\textsuperscript{b,}\* \\
\textsuperscript{a}Regional Economy Division, Economic Policy Bureau, Ministry of Finance and Economy, 1, Joongang-dong, Kwacheon, Kyunggi-do, 427-760, Korea \\
\textsuperscript{b}Economics Department, Marshall Hall, Michigan State University, East Lansing, MI 48824 USA

Received 21 October 1996; accepted 14 January 1997

Abstract

We derive the asymptotic distribution of the Kwiatkowski et al. (1992) statistic under nonstationary long memory $(1/2 < d < 1)$. Its order in probability is the same under nonstationary long memory as under a unit root. It cannot, therefore, distinguish consistently between the two cases. © 1997 Elsevier Science S.A.

Keywords: Nonstationary long memory; Unit root; Stationarity test statistic

JEL classification: C12

1. Introduction

The enormous literature on testing for a unit root seeks to determine whether an economic time series is stationary or has a unit root. Dickey, Fuller (1979) tests are most commonly used to test the hypothesis of a unit root against the alternative of stationarity. In this paper we consider the test of Kwiatkowski et al. (1992)—hereafter KPSS—which is designed to test the hypothesis of stationarity against the alternative of a unit root. The KPSS statistic with different critical values can also be used as the basis for a unit root test, as suggested by Shin, Schmidt (1992).

More recently, it has been recognized that the null hypothesis for tests of “stationarity” (and the alternative for unit root tests) is actually short memory (or weak dependence) since the series or its difference is required to satisfy a functional central limit theorem for convergence to Brownian motion, and this result requires a limit on the memory or dependence of the series. Long memory series are also empirically relevant, and form a middle ground between short memory and unit root series; see, e.g., the survey by Baillie (1996). A popular and useful model is the fractionally integrated, or \( I(d) \), model of Granger (1980), Granger, Joyeux (1980) and Hosking (1981). In this framework short memory corresponds to \( d = 0 \) and a unit root to \( d = 1 \), but fractional values are

\*Corresponding author: Tel.: 001-517-355-3774; fax: 001-517-432-1068; e-mail amsler@pilot.msu.edu.

0165-1765/97/$17.00 © 1997 Elsevier Science S.A. All rights reserved.

PII S0165-1765(97)00066-9
meaningful. In particular, the series is stationary but has long memory for $0 < d < 1/2$, while it is nonstationary and long memory (but still mean-reverting) for $1/2 < d < 1$.

The asymptotic distribution of the KPSS statistic under short memory and under a unit root were given by KPSS and by Shin, Schmidt (1992). Lee, Schmidt (1996) derived the asymptotic distribution of the KPSS statistic for the stationary long memory case ($0 < d < 1/2$). In this paper we derive the asymptotic distribution for the nonstationary long memory case ($1/2 < d < 1$). We show that the order in probability of the test statistic is the same under nonstationary long memory as under a unit root, so that the KPSS statistic cannot distinguish consistently between these two cases. In other words, of the four cases discussed above, the KPSS statistic can distinguish consistently between short memory, stationary long memory, and either nonstationary long memory or unit root. We also present some simulations that show the relevance of these asymptotic results in finite samples.

2. Asymptotic results

A. Notation

We consider the data generating process:

$$y_t = \phi + \xi t + \epsilon_t, \ t = 1, 2, \ldots, T,$$

(1)

where $\{y_t\}$ is the observed series and $\{\epsilon_t\}$ is the deviation from trend. Let $e_t$ be the residuals from a regression of $y_t$ on intercept and trend ($t$), and let $S_t$ be their partial sum: $S_t = \Sigma_{j=1}^t e_j$. Let $s^2(\ell)$ be the Newey–West estimator of the long-run variance of the $\epsilon_t$:

$$s^2(\ell) = T^{-1} \sum_{t=1}^T e_t^2 + 2T^{-1} \sum_{s=1}^\ell [1 - s/(\ell + 1)] \sum_{t=s+1}^T e_t e_{t-s}.$$  

(2)

For the case $\ell = 0$, the second term on the right hand side of (2) is simply set to zero. Then the KPSS statistic $\hat{\eta}(\ell)$ is defined as

$$\hat{\eta}(\ell) = T^{-2} \sum_{t=1}^T S_t^2 / s^2(\ell).$$  

(3)

The KPSS statistic $\hat{\eta}(\ell)$ is defined similarly except that we set $\xi = 0$ in (1), which implies use of the residuals $\epsilon_t = y_t - \bar{y}$ in defining $S_t$ and $s^2(\ell)$.

To test the null of short memory, KPSS require $\ell \to \infty$ but $\ell/T \to 0$ as $T \to \infty$. For the KPSS unit root test, Shin, Schmidt (1992) suggest $\ell = 0$. We will consider both of these cases.

B. KPSS under short memory

Let $Z_t = \Sigma_{j=1}^t \epsilon_j$ represent the partial sum of the $\epsilon_t$. Following Lee, Schmidt (1996), we can say that $\epsilon_t$ is a short memory process if it satisfies:

$$\sigma^2 = \lim_{T \to \infty} T^{-1} E(Z_t^2)$$

exists and is non-zero

$$\forall \ r \in [0, 1], \ T^{-1/2} Z_{[rT]} \to \sigma W(r)$$

Here $[rT]$ denotes the integer part of $rT$, $\to$ denotes weak convergence, and $W(r)$ is the standard
Wiener process (Brownian motion). KPSS assume the mixing and moment conditions of (Phillips, Perron, 1988, p.336), which imply (A1) and (A2).

Under the hypothesis that \( \epsilon_i \) is a short memory process, KPSS show that

\[
\hat{\eta}_i(\ell) \to \int_0^1 V_\epsilon(r)^2 \, dr, \quad (4)
\]

provided \( \ell \to \infty \) and \( \ell/T \to 0 \) as \( T \to \infty \). Here \( V_\epsilon(r) \) is a second-level Brownian bridge, as defined by KPSS, Eq. 16. For \( \ell = 0 \), \( \hat{\eta}_i(0) \) converges to \((\sigma_i^2/\sigma_e^2)\) times the limit in (4). Thus, in either case \( \hat{\eta}_i \), is \( O_p(1) \) under short memory. Similar statements hold for \( \hat{\eta}_\mu \), with \( V_\epsilon(r) \) replaced by the standard Brownian bridge, \( V_r(W(r) - rW(1)) \).

C. KPSS under unit root

Suppose that \( \epsilon_i \) has a unit root, and more specifically that \( \Delta \epsilon_i \) satisfies the short-memory conditions (A1) and (A2). Then, for the case that \( \ell \to \infty \) and \( \ell/T \to 0 \) as \( T \to \infty \), KPSS show that

\[
(\ell/T)\hat{\eta}_i(l) \to \int_0^1 \left( \int_0^1 W^*(s) \, ds \right)^2 \frac{d\alpha}{\int_0^1 W^*(s)^2 \, ds}, \quad (5)
\]

where \( W^*(s) \) is a demeaned and detrended Weiner process, as defined in (Park, Phillips, 1988, p. 474). A similar statement holds for \( \hat{\eta}_\mu(\ell) \), with \( W^*(s) \) replaced by the demeaned Brownian motion, \( W(s) = W(s) - \int_0^1 W(r) \, dr \). Thus \( \hat{\eta}_i(\ell) \) and \( \hat{\eta}_\mu(\ell) \) are \( O_p(T/\ell) \) for the case that \( \ell \to \infty \), \( \ell/T \to 0 \), as \( T \to \infty \). Shin, Schmidt (1992) show that \( T^{-1}\hat{\eta}_i(0) \) has the same distribution as given in (5), with a similarly modified result for \( \hat{\eta}_\mu(0) \). Thus \( \hat{\eta}_i(0) \) and \( \hat{\eta}_\mu(0) \) are \( O_p(T) \).

D. KPSS under stationary long memory

Lee, Schmidt consider the case that \( \epsilon_i \) follows an I(\(d\)) process: \((1-L)^d \epsilon_i = u_i \), where \( u_i \) is normal white noise: This process is stationary for \(-1/2 < d < 1/2\). Here we are concerned only with \( d \) in the range \( 0 < d < 1/2 \), but the results cited hold for \(-1/2 < d < 1/2\). If \( Z_i \) is the partial sum of the \( \epsilon_i \), it satisfies the invariance principle

\[
T^{-d} Z_{[T]} \to k(d)W_d(r), \quad (6)
\]

where \( k(d) \) is a scalar whose value is irrelevant for our purposes, and \( W_d(r) \) is a fractional Brownian motion as defined by, e.g., (Beran, 1994, p. 56).

Lee, Schmidt (1996) establish the following results:

\[
T^{-2d} \hat{\eta}_i(0) \to (k(d)/\sigma_e^2) \int_0^1 V_\epsilon(r)^2 \, dr \quad (\ell = 0), \quad (7a)
\]

\[
(\ell/T)^{-2d} \hat{\eta}_i(\ell) \to \int_0^1 V_\epsilon(r)^2 \, dr \quad (\ell \to \infty \text{ but } \ell/T \to 0 \text{ as } T \to \infty), \quad (7b)
\]

where \( V_\epsilon(r) \) is defined by Lee, Schmidt, Lemma 2, p. 291. Similar statements hold for \( \hat{\eta}_\mu \), with \( V_\epsilon(r) \)
where replaced by $B_{g}(r)=W_{g}(r)−rW_{g}(1)$. Thus $\hat{\eta}_{g}(0)$ and $\hat{\eta}_{g}(0)$ are $O_{p}(T^{2d})$; and $\hat{\eta}_{g}(\ell)$ and $\hat{\eta}_{g}(\ell)$ are $O_{p}(\ell T^{2d})$ when $\ell \to \infty$, $\ell T \to 0$ as $T \to \infty$.

Each of the KPSS statistics has a different order in probability under short memory, stationary long memory, and unit root, respectively. Thus the KPSS test can distinguish consistently between these three possibilities.

E. KPSS Under Nonstationary Long Memory

We now turn to the main theoretical contribution of this paper, which is the derivation of the asymptotic distribution of the KPSS statistics when $\epsilon_{i}$ is a nonstationary long memory process. Thus we wish to consider the case that $\epsilon_{i}$ is $I(d)$ with $1/2<d<1$.

Define $d^{*}=d-1$, so that $\Delta \epsilon_{i}$ is $I(d^{*})$ with $|d^{*}|<1/2$; that is, $\Delta \epsilon_{i}$ is a stationary long memory process. Then, assuming that $u_{i}=(1-L)^{d} \Delta \epsilon_{i}=(1-L)^{d} \epsilon_{i}$ is a normal white noise process, we have the invariance principle

$$T^{-(d^{*}+1/2)} \epsilon_{(rT)}=T^{-(d^{*}-1/2)} \epsilon_{(rT)} \rightarrow k(d^{*})W_{d^{*}}(r).$$

This is essentially the same result as (6) above, treating $\epsilon_{i}$ as the cumulation of the $\Delta \epsilon_{i}$, which are $I(d^{*})$.

Using this result, we can prove the following theorem.

**Theorem 1.** Suppose that Eq. (1) holds, with $\xi=0$, and that $(1-L)^{d} \epsilon_{i}=u_{i}$, $1/2<d<1$, where $u_{i}$ is normal white noise. Then

$$T^{-1} \hat{\eta}_{g}(0) \rightarrow \left\{ \int_{0}^{1} W_{d}(a) \, da \right\}^{2} \beta K \left\{ \int_{0}^{1} W_{d}(a)^{2} \, da \right\}$$

(9a)

$$\ell T^{-1} \hat{\eta}_{g}(\ell) \rightarrow \left\{ \int_{0}^{1} W_{d}(a) \, da \right\}^{2} \beta K \left\{ \int_{0}^{1} W_{d}(a)^{2} \, da \right\}$$

(9b)

where $W_{d}(a)=W_{d}(a)-\int_{0}^{1} W_{d}(b) \, db$. Furthermore, the same results hold for $\hat{\eta}_{g}(0)$ and $\hat{\eta}_{g}(\ell)$, with $W_{d}(a)$ replaced by

$$W_{d^{*}}(a)=W_{d}(a)+(6a-4) \int_{0}^{1} W_{d}(b) \, db+(-12a+6) \int_{0}^{1} bW_{d}(b) \, db.$$

(10)

The results for $\hat{\eta}_{g}$ do not require the assumption that $\xi=0$.

**Proof.** See Appendix A.

The Theorem implies that $\hat{\eta}_{d}(0)$ and $\hat{\eta}_{d}(0)$ are $O_{p}(T)$, while $\hat{\eta}_{d}(\ell)$ and $\hat{\eta}_{d}(\ell)$ are $O_{p}(T/\ell)$ for the case that $\ell \to \infty$ and $\ell T \to 0$. Therefore the KPSS unit root test is not consistent against nonstationary long memory alternatives, $I(d)$ for $1/2<d<1$, because the KPSS statistics have the same orders in probability under both the null and alternative hypotheses. In fact, for $1/2<d \leq 1$, the orders in probability of the KPSS statistics are independent of the value of $d$, even though the form of their asymptotic distributions is affected by the value of $d$. This is in contrast to the case of a stationary
long memory process, where both the order in probability and the form of the asymptotic distribution depend on $d$.

The KPSS short memory test is known to be consistent against stationary short memory alternatives and unit root alternatives. Our result shows that it is also consistent against nonstationary long memory alternatives. This is not surprising, but the inability to distinguish consistently between unit root and nonstationary long memory is perhaps surprising.

3. Simulation results

In this section we provide simulation evidence on the power of the KPSS stationarity (short memory) and unit root tests. The lag truncation parameters are chosen as $\ell_0 = 0$, $\ell_4 = \text{integer}[4(T/100)^{1/4}]$, and $\ell_{12} = \text{integer}[12(T/100)^{1/4}]$ as in Schwert (1989), KPSS (1992) and Lee, Schmidt 1996. We consider sample sizes 50, 150, 250, 500 and 1000, and the number of iterations is 10 000. All of our tests are based on the 5% significance level. The method of data generation is as described in Lee, Schmidt (1996).

Table 1 gives the powers of the 5% upper tail KPSS short memory tests against the alternatives $d = 0.0, 0.1, 0.2, \ldots, 0.9, 1.0$, and also $d = 0.45$ and 0.499. These are an elaboration of the values considered by Lee, Schmidt (1996). For fixed $d$, power increases with $T$, reflecting the consistency of the tests. For fixed $T$, power increases with $d$, a result that is not surprising, even though it is not transparent from the relevant asymptotics for $1/2 < d \leq 1$.

Table 2 gives the power of the lower tail KPSS unit root test against $I(d)$ alternatives with $1/2 \leq d \leq 1$. The most important result is that, with $d$ fixed, power does not approach one as $T$ increases. For example, for $d = 0.7$, power grows from 0.169 with $T = 50$ to only 0.256 with $T = 1000$. This is a reflection of the inconsistency of the KPSS unit root test against nonstationary long memory processes; power would not be expected to approach one even for arbitrarily large values of $T$. We can also see that, for fixed $T$, power increases as $d$ decreases. This is not surprising, but also not transparent from the relevant asymptotics for $1/2 < d \leq 1$.

4. Conclusions

In this paper we have asked whether the KPSS statistic can be used to distinguish between the following possibilities: (i) short memory; (ii) stationary long memory; (iii) nonstationary long memory; and (iv) unit root. We find that it can consistently distinguish (i) from (ii) from the combination of (iii) and (iv), but it cannot distinguish consistently between (iii) and (iv). An interesting question is whether another single statistic can distinguish consistently between these four possibilities. The central feature of our result is that the order in probability of the KPSS statistic does not depend on $d$ for $d$ in some range, and this possibility is not peculiar to the KPSS statistic. For example, Sowell (1990) shows that the orders in probability of the Dickey–Fuller (1979) statistics depend on $d$ for $d < 1$ but not for $1 \leq d < 3/2$. A broader investigation of this phenomenon would certainly be worthwhile.
<table>
<thead>
<tr>
<th>(d)</th>
<th>0.0</th>
<th>0.1</th>
<th>0.2</th>
<th>0.3</th>
<th>0.4</th>
<th>0.45</th>
<th>0.49</th>
<th>0.499</th>
<th>0.5</th>
<th>0.51</th>
<th>0.56</th>
<th>0.7</th>
</tr>
</thead>
<tbody>
<tr>
<td>(T)</td>
<td>(\eta_{m}) Test</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>50</td>
<td>0.044</td>
<td>0.037</td>
<td>0.124</td>
<td>0.072</td>
<td>0.012</td>
<td>0.240</td>
<td>0.387</td>
<td>0.032</td>
<td>0.072</td>
<td>0.613</td>
<td>0.278</td>
<td>0.072</td>
</tr>
<tr>
<td>100</td>
<td>0.053</td>
<td>0.047</td>
<td>0.165</td>
<td>0.100</td>
<td>0.033</td>
<td>0.356</td>
<td>0.542</td>
<td>0.093</td>
<td>0.198</td>
<td>0.773</td>
<td>0.384</td>
<td>0.198</td>
</tr>
<tr>
<td>150</td>
<td>0.050</td>
<td>0.044</td>
<td>0.195</td>
<td>0.119</td>
<td>0.039</td>
<td>0.402</td>
<td>0.631</td>
<td>0.121</td>
<td>0.255</td>
<td>0.801</td>
<td>0.468</td>
<td>0.255</td>
</tr>
<tr>
<td>250</td>
<td>0.046</td>
<td>0.051</td>
<td>0.216</td>
<td>0.131</td>
<td>0.041</td>
<td>0.542</td>
<td>0.730</td>
<td>0.158</td>
<td>0.329</td>
<td>0.884</td>
<td>0.549</td>
<td>0.329</td>
</tr>
<tr>
<td>500</td>
<td>0.052</td>
<td>0.051</td>
<td>0.262</td>
<td>0.169</td>
<td>0.049</td>
<td>0.631</td>
<td>0.845</td>
<td>0.211</td>
<td>0.455</td>
<td>0.958</td>
<td>0.714</td>
<td>0.455</td>
</tr>
<tr>
<td>1000</td>
<td>0.053</td>
<td>0.051</td>
<td>0.323</td>
<td>0.195</td>
<td>0.048</td>
<td>0.685</td>
<td>0.928</td>
<td>0.282</td>
<td>0.576</td>
<td>0.693</td>
<td>0.797</td>
<td>0.576</td>
</tr>
</tbody>
</table>

| \(T\) | \(\eta_{c}\) Test | | | | | | | | | | | |
| 50  | 0.051 | 0.041 | 0.140 | 0.071 | 0.044 | 0.267 | 0.418 | 0.060 | 0.072 | 0.639 | 0.259 | 0.072 |
| 100 | 0.052 | 0.045 | 0.215 | 0.122 | 0.034 | 0.380 | 0.608 | 0.086 | 0.154 | 0.870 | 0.490 | 0.154 |
| 150 | 0.052 | 0.048 | 0.265 | 0.145 | 0.044 | 0.461 | 0.714 | 0.113 | 0.219 | 0.941 | 0.600 | 0.219 |
| 250 | 0.053 | 0.053 | 0.331 | 0.196 | 0.048 | 0.577 | 0.829 | 0.149 | 0.476 | 0.991 | 0.794 | 0.476 |
| 500 | 0.053 | 0.053 | 0.311 | 0.226 | 0.044 | 0.722 | 0.932 | 0.218 | 0.626 | 0.999 | 0.877 | 0.626 |
| 1000| 0.052 | 0.052 | 0.311 | 0.212 | 0.048 | 0.839 | 0.983 | 0.295 | 0.913 | 1.000 | 0.970 | 0.913 |

| \(\eta_{c}\) Test | 50  | 100  | 150  | 250  | 500  | 1000 |
| 0.044 | 0.053 | 0.050 | 0.046 | 0.052 | 0.053 |
| 0.036 | 0.047 | 0.046 | 0.044 | 0.051 | 0.051 |
| 0.012 | 0.033 | 0.039 | 0.041 | 0.049 | 0.048 |
| 0.124 | 0.165 | 0.195 | 0.216 | 0.262 | 0.323 |
| 0.072 | 0.100 | 0.119 | 0.131 | 0.169 | 0.195 |
| 0.020 | 0.055 | 0.075 | 0.085 | 0.117 | 0.141 |
| 0.240 | 0.356 | 0.402 | 0.485 | 0.592 | 0.693 |
| 0.123 | 0.182 | 0.221 | 0.265 | 0.348 | 0.418 |
| 0.032 | 0.093 | 0.121 | 0.158 | 0.211 | 0.282 |
| 0.200 | 0.356 | 0.402 | 0.485 | 0.592 | 0.693 |

| \(\eta_{m}\) Test | 50  | 100  | 150  | 250  | 500  | 1000 |
| 0.053 | 0.050 | 0.046 | 0.052 | 0.053 | 0.053 |
| 0.047 | 0.046 | 0.044 | 0.051 | 0.051 | 0.051 |
| 0.033 | 0.039 | 0.041 | 0.049 | 0.048 | 0.048 |
| 0.165 | 0.195 | 0.216 | 0.262 | 0.323 | 0.323 |
| 0.100 | 0.119 | 0.131 | 0.169 | 0.195 | 0.195 |
| 0.055 | 0.075 | 0.085 | 0.117 | 0.141 | 0.141 |
| 0.356 | 0.402 | 0.485 | 0.592 | 0.693 | 0.693 |
| 0.182 | 0.221 | 0.265 | 0.348 | 0.418 | 0.418 |
| 0.093 | 0.121 | 0.158 | 0.211 | 0.282 | 0.282 |
| 0.356 | 0.402 | 0.485 | 0.592 | 0.693 | 0.693 |
Acknowledgments

We thank Peter Schmidt for helpful comments and suggestions.

Appendix A

In this Appendix we give a sketch of the proof of Theorem 1 for the \( \hat{\eta}_\mu(0) \) and \( \hat{\eta}_\mu(\ell) \) statistics. More detail, and the similar proof for \( \hat{\eta}_\ell(0) \) and \( \hat{\eta}_\ell(\ell) \), can be found in Lee (1995).

Table 2

| Power of KPSS Lower Tail Unit Root Test Against \( l(d) \), \( d \in [0.5, 1.0] \) |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| \( T \) | \( \hat{\eta}_\mu(0) \) Test | \( \hat{\eta}_\ell(0) \) Test |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 50 | 100 | 150 | 250 | 500 | 1000 | 50 | 100 | 150 | 250 | 500 | 1000 |
| \( d = 0.5 \) | 0.358 | 0.426 | 0.450 | 0.485 | 0.535 | 0.576 | 0.360 | 0.499 | 0.528 | 0.590 | 0.647 | 0.714 |
| \( d = 0.51 \) | 0.338 | 0.397 | 0.433 | 0.465 | 0.515 | 0.549 | 0.350 | 0.456 | 0.510 | 0.566 | 0.634 | 0.687 |
| \( d = 0.6 \) | 0.243 | 0.295 | 0.298 | 0.322 | 0.347 | 0.364 | 0.250 | 0.328 | 0.351 | 0.379 | 0.430 | 0.456 |
| \( d = 0.7 \) | 0.169 | 0.190 | 0.193 | 0.207 | 0.213 | 0.223 | 0.169 | 0.209 | 0.212 | 0.226 | 0.240 | 0.256 |
| \( d = 0.8 \) | 0.108 | 0.119 | 0.126 | 0.125 | 0.129 | 0.127 | 0.107 | 0.122 | 0.130 | 0.130 | 0.144 | 0.150 |
| \( d = 0.9 \) | 0.072 | 0.074 | 0.082 | 0.075 | 0.078 | 0.078 | 0.068 | 0.073 | 0.076 | 0.075 | 0.078 | 0.077 |
| \( d = 0.95 \) | 0.063 | 0.059 | 0.061 | 0.059 | 0.058 | 0.059 | 0.049 | 0.056 | 0.058 | 0.060 | 0.055 | 0.057 |
| \( d = 0.99 \) | 0.048 | 0.050 | 0.049 | 0.050 | 0.049 | 0.045 | 0.043 | 0.044 | 0.044 | 0.046 | 0.048 | 0.049 |
| \( d = 1.0 \) | 0.048 | 0.049 | 0.045 | 0.044 | 0.045 | 0.048 | 0.042 | 0.040 | 0.042 | 0.044 | 0.041 | 0.044 |
As in the main text, $Z_t$ is the partial sum of the $\varepsilon_t$ and $S_t$ is the partial sum of $\varepsilon_t$; for the $\hat{\eta}_t$ tests, $e_t = \varepsilon_t - \bar{\varepsilon}$.

**Lemma 1.**

1. $T^{-(d^*+3/2)} Z^{(T)} \rightarrow \kappa(d^*) \int_0^1 W_d(a) da$;
2. $T^{-(d^*+1/2)} e^{(T)} \rightarrow \kappa(d^*) W_{d^*}(a)$, where $W_{d^*}(a) = W_d(a) - \int_0^1 W_d(b) db$;
3. $T^{-(d^*+3/2)} S^{(T)} \rightarrow \kappa(d^*) \int_0^1 W_{d^*}(a) da$;
4. $T^{-(2d^*+2)} \sum_{t=1}^T \varepsilon_t^2 = T^{-(d^*+1/2)} S^2(0) \rightarrow \kappa(d^*) \left\{ \int_0^1 W_d(a)^2 da \right\}$;
5. $T^{-(2d^*+4)} \sum_{t=1}^T S_t^2 \rightarrow \kappa(d^*) \left\{ \int_0^1 W_{d^*}(a) da \right\}^2 dr$

**Proof.** We begin with Eq. (8) of the text: $T^{-(d^*+1/2)} e^{(T)} \rightarrow \kappa(d^*) W_{d^*}(r)$; (i) follows directly. To establish (ii), note

$$T^{-(d^*+1/2)} e^{(T)} = T^{-(d^*+1/2)} e^{(T)} - T^{-(d^*+1/2)} e^{(T)}$$

(A1)

which implies (ii). Then (iii), (iv) and (v) follow from (ii). The proof of the first part of Theorem 1 is now straightforward. We have

$$T^{-1} \hat{\eta}_t(0) = T^{-(2d^*+4)} \sum_{t=1}^T S_t^2 / T^{-(2d^*+1)} S^2(0) \rightarrow \left\{ \int_0^1 W_d(a) da \right\}^2 dr / \left\{ \int_0^1 W_{d^*}(a) da \right\}^2 \text{ (} \ell = 0 \text{)}$$

(A2)

using (iv) and (v) of Lemma 1. This establishes (9A) of the text.

**Lemma 2.**

1. $T^{-(2+2d^*)} \sum_{t=1}^T \varepsilon_t \Delta \varepsilon_{s+t} \rightarrow 0$ for any nonnegative integer $s$.
2. $T^{-(2+2d^*)} \sum_{s=0}^{t-1} \sum_{t=s+1}^T \varepsilon_t \Delta \varepsilon_{t-s} \rightarrow 0$ for any nonnegative integer $s$.
3. $T^{-(2+2d^*)} \sum_{t=1}^T \varepsilon_t \Delta \varepsilon_{t-s} \rightarrow 0$ da.
4. When $\ell \rightarrow \infty$ and $\ell / T \rightarrow 0$ as $T \rightarrow \infty$, then $[s^2(\ell)]T^{-(1+2d^*)} \rightarrow k(d^*) \int_0^1 W_{d^*}(a)^2 da$.

**Proof.** Let $\gamma_j$ be the $j$th autocovariance of $\Delta \varepsilon$. We will prove (i) for the case $s=0$; the case of $s > 0$ is similar but more complicated. Since $\sum_{t=1}^T \varepsilon_t \Delta \varepsilon_t = 1/2(\varepsilon_T^2 - \varepsilon_0^2 - \sum_{t=1}^T (\Delta \varepsilon_t^2))$,

$$T^{-(2+2d^*)} \sum_{t=1}^T \varepsilon_t \Delta \varepsilon_t = (2T)^{-1} [T^{-d^*/2}] [\varepsilon_T^2] - (2T)^{-1} [T^{-d^*/2}] [\varepsilon_0^2]$$

$$- (2^{-1} T^{-(1+2d^*)} \sum_{t=1}^T (\Delta \varepsilon_t^2) \rightarrow 0$$

(A3)

because $[T^{-d^*/2}] [\varepsilon_T^2] \rightarrow \omega_{d^*}^2 W_{d^*}(1)^2$, $[T^{-1} \sum_{t=1}^T (\Delta \varepsilon_t^2) \rightarrow \gamma_0$ and $\varepsilon_0$ is $O_p(1)$. To prove (ii),
\begin{equation}
T^{-(2+2d^*)} \sum_{j=0}^{T} \sum_{i=j+1}^{T} \epsilon_{i-j} \Delta \epsilon_{i-j} = T^{-(2+2d^*)} \sum_{j=0}^{T} [\epsilon_{i-j} \Delta \epsilon_{i-j} + \epsilon_{i-j} \Delta \epsilon_{i-1} + \cdots + \epsilon_{i-j} \Delta \epsilon_{i+1}] \end{equation}

(A4)

and each term $\frac{p}{T} \to 0$ by (i). To prove (iii), write

\begin{equation}
T^{-(2+2d^*)} \sum_{j=0}^{T} \epsilon_{i} \epsilon_{i-s} = T^{-(2+2d^*)} \sum_{j=0}^{T} \epsilon_{i-j}^2 + T^{-(2+2d^*)} \sum_{j=0}^{T} \sum_{i=j+1}^{T} \epsilon_{i-j} \Delta \epsilon_{i-j} .
\end{equation}

(A5)

The first term has the desired distribution by (iv) of Lemma 1, while the second term $\frac{p}{T} \to 0$ by (ii).

Finally, to prove (iv), define $\omega(s, \ell) = 1 - s/(\ell + 1)$, recall $e_{\ell} = e_{\ell} - \bar{e}$, and do some algebra to obtain

\begin{equation}
[s^2(\ell)][T^{-(1+2d^*)}] = \left[ T^{-1} \sum_{t=1}^{T} [(T^{-(d^*+1/2)})^2 - \left[T^{-1} \sum_{t=1}^{T} (T^{-(d^*+1/2)})^2 \right] \right] + 2 \sum_{s=1}^{\ell} \omega(s, \ell) \times \left[ \sum_{t=1}^{T} [(T^{-(2+2d^*)})(\epsilon_t \epsilon_{t-s} - e_{\ell} \bar{e} \epsilon_{t-s} + \bar{e}^2) \right] = (A) + 2 \sum_{s=1}^{\ell} \omega(s, \ell) \times (B_s).
\end{equation}

(A6)

Part (A) is the same as in the consideration of $s^2(0)$ and has the distribution given in (iv) of Lemma 1. Furthermore, $B_s$ has this same distribution for all $s$, using (iii) and the same argument that led to (iv) of Lemma 1. Since $1 + 2 \sum_{s=1}^{\ell} \omega(s, \ell) = \ell + 1$, for fixed $\ell$ we have

\begin{equation}
T^{-(1+2d^*)} s^2(\ell) \to (\ell + 1)k(d^*)^2 \int_0^1 W_d(a)^2 da .
\end{equation}

(A7)

Then (iv) follows by dividing by $\ell$, and letting $\ell \to \infty$ as $T \to \infty$. The proof of the second part of Theorem 1 is now straightforward. We have

\begin{equation}
(\ell/T) \hat{\eta}_p(\ell) = T^{-(2d^*+4)} \sum_{t=1}^{T} S_t^2 / \ell^{-1} T^{-(1+2d^*)} s^2(\ell)
\end{equation}

(A8)

and we obtain (9B) of the text using (v) of Lemma 1, and (iv) of Lemma 2.

References


