Econometric Theory, **23**, 2007, 355–363. Printed in the United States of America.
DOI: 10.1017/S0266466607070156

NOTES AND PROBLEMS

MODIFIED KPSS TESTS FOR NEAR INTEGRATION

DAVID HARRIS
University of Melbourne

STEPHEN LEYBOURNE
University of Nottingham

BRENDAN MCCABE
University of Liverpool

This note suggests a simple modification to the Kwiatkowski, Phillips, Schmidt, and Shin (1992, *Journal of Econometrics*, 54, 159–178) test (KPSS test) so that it is applicable to testing the null hypothesis of near integration against a unit root alternative. The modified KPSS test is shown not to suffer from the asymptotic size distortion problems of the original KPSS test that are described by Müller (2005, *Journal of Econometrics* 128, 195–213). The test also has good asymptotic and finite-sample properties relative to the point optimal tests of Müller (2005) and Elliott and Müller (2006, *Journal of Econometrics* 135, 285–310).

1. MOTIVATION AND RESULTS

The problem of deciding whether a strongly autocorrelated time series is best described as an integrated process or a stationary process with largest root near to one (i.e., a near integrated process) is a long-standing one in time series econometrics. Many observed macroeconomic time series appear to be well described as either integrated or near integrated; see Tables 6 and 7 of Stock and Watson (1999) for some well-known examples. From a hypothesis testing perspective, the null hypothesis may be chosen to be either integration or near integration. Müller (2005) investigated the properties of the Kwiatkowski, Phillips, Schmidt, and Shin (1992) test (KPSS test) applied to testing the null hypothesis that a time series is near integrated against the alternative hypothesis that it is integrated. His results show that the KPSS test as usually implemented, using a bandwidth parameter in the long-run variance estimator that increases at a slower rate than the sample size, has an asymptotic size of unity under this null hypothesis. This finding is taken to explain the increased finite-sample size

Address correspondence to David Harris, Department of Economics, University of Melbourne, Victoria 3010, Australia; e-mail: harris@unimelb.edu.au.

© 2007 Cambridge University Press 0266-4666/07 $12.00

https://doi.org/10.1017/S0266466607070156 Published online by Cambridge University Press

of the KPSS test under strongly autocorrelated stationary data generating processes (DGPs); see, for example, Caner and Kilian (2001, Table 1). Alternately, Müller showed that choosing the bandwidth parameter to increase at the same rate as the sample size results in a test with size of less than unity but with a nonstandard asymptotic distribution, possibly dependent on nuisance parameters. He compared these asymptotic properties of the KPSS test with those of a point optimal test designed specifically for testing for near integration against a unit root, and he found the point optimal test to be superior.

In this note we suggest a simple modification to the KPSS test to make it applicable to testing for near integration against a unit root. Briefly, if the model for a time series w_t is

$$w_t = (1 - c/T)w_{t-1} + \nu_t, \tag{1}$$

where $c \geq 0$ and ν_t is stationary, then the near integrated null hypothesis is $H_0: c \geq \bar{c} > 0$ for some fixed $\bar{c}$, to be tested against the unit root alternative $H_1: c = 0$. Instead of applying the KPSS test to w_t, we suggest that it be applied instead to the filtered series $w_t - (1 - \bar{c}/T)w_{t-1}$. This has the effect of removing the near unit root under the null hypothesis, resulting in a KPSS test with controlled asymptotic size. Under the alternative hypothesis, the filtering does not completely remove the unit root, and so the test retains nontrivial power.

The asymptotic properties of the transformed KPSS test are derived and compared with those of the point optimal test of Müller (2005) and another point optimal test derived by Elliott and Müller (2006). These point optimal tests are based on differing treatments of the initial value, w_1, in (1). The modified test is found to have superior properties for a range of assumptions about the initial value and also to have superior finite-sample size properties in the presence of stationary autocorrelation in ν_t. We leave for future research the use of the modified test for confidence interval construction for the largest autoregressive root, as discussed by Elliott and Stock (2001) using point optimal tests.

2. THE MODEL AND MODIFIED KPSS TEST

Consider the following DGP (cf. Müller and Elliott, 2003) for an observed series y_t

$$\begin{aligned} y_t &= \mu + w_t, \qquad t = 1, \ldots, T, \\ w_t &= \rho_{c,T} w_{t-1} + \nu_t, \qquad t = 2, \ldots, T, \\ w_1 &= \xi, \end{aligned} \tag{2}$$

where ν_t is a stationary process. Here $\rho_{c,T} = 1 - cT^{-1}$, $c \geq 0$. We wish to test the null hypothesis of local to unit root stationarity against the unit root alternative. Following Müller (2005) we state these as $H_0: c \geq \bar{c} > 0$ against $H_1: c = 0$ where $\bar{c}$ specifies the minimal amount of mean reversion under the stationary null hypothesis.

https://doi.org/10.1017/S0266466607070156 Published online by Cambridge University Press

As regards the initial value, which is important under H_0, following Müller and Elliott (2003), we assume that it is fixed and set $\xi = \alpha\omega/(1-\rho_{c,T}^2)^{1/2}$, for $c > 0$. Here α controls the magnitude of the initial value relative to the standard deviation of a stationary AR(1) process with parameter $\rho_{c,T} < 1$. We assume that ν_t is a zero mean stationary process with finite autocovariances $E(\nu_t \nu_{t-j}) = \gamma_j$ and long-run variance $\omega^2 = \sum_{j=-\infty}^{\infty} \gamma_j$ that is finite and nonzero.

We consider the following generalized least squares (GLS) type transformation of the y_t based on our hypothesized value $\bar{c}$ under the stationary null:

$$y_t - \rho_{\bar{c},T} y_{t-1} = \mu(1-\rho_{\bar{c},T}) + (w_t - \rho_{\bar{c},T} w_{t-1})$$

for $t = 2, \ldots, T$. To use the $y_t - \rho_{\bar{c},T} y_{t-1}$ as the basis for a KPSS test we need to make them invariant to $\mu(1-\rho_{\bar{c},T})$. So, we utilize the ordinary least squares (OLS) estimator of $\mu(1-\rho_{\bar{c},T})$ given by

$$m_{\bar{c}} = \frac{\sum_{t=2}^{T} (y_t - \rho_{\bar{c},T} y_{t-1})}{(T-1)}$$

and calculate the OLS residuals

$$r_{\bar{c},t} = (y_t - \rho_{\bar{c},T} y_{t-1}) - m_{\bar{c}}.$$

Our modified KPSS test is then the standard KPSS statistic constructed from the $r_{\bar{c},t}$. That is,

$$S^{\mu}(\bar{c}) = \frac{T^{-2} \sum_{t=2}^{T} \left(\sum_{i=2}^{t} r_{\bar{c},i} \right)^2}{\hat{\omega}_{\bar{c}}^2}.$$

Here $\hat{\omega}_{\bar{c}}^2$ is any standard long-run variance estimator of the form

$$\hat{\omega}_{\bar{c}}^2 = \hat{\gamma}_{\bar{c},0} + 2 \sum_{j=1}^{T-1} \lambda(j/l) \hat{\gamma}_{\bar{c},j}, \qquad \hat{\gamma}_{\bar{c},j} = T^{-1} \sum_{t=j+1}^{T} r_{\bar{c},t} r_{\bar{c},t-j}, \tag{3}$$

where $\lambda(.)$ is a kernel function. The following theorem provides the asymptotic distribution of $S^{\mu}(\bar{c})$ under the null and alternative hypotheses.

THEOREM 1. *Under the preceding assumption for* ν_t,

$$S^{\mu}(\bar{c}) \Rightarrow \int_0^1 H_{\alpha,c,\bar{c}}(r)^2 \, dr,$$

https://doi.org/10.1017/S0266466607070156 Published online by Cambridge University Press

where

$$H_{\alpha,c,\bar{c}}(r) = K_{\alpha,c}(r) + \bar{c}\int_0^r K_{\alpha,c}(s)\,ds - r\left\{K_{\alpha,c}(1) + \bar{c}\int_0^1 K_{\alpha,c}(s)\,ds\right\},$$

$$K_{\alpha,c}(r) = \begin{cases} \alpha(e^{-rc}-1)(2c)^{-1/2} + \displaystyle\int_0^r e^{-(r-s)c}\,dW(s), & c > 0 \\ W(r), & c = 0, \end{cases}$$

and $W(r)$ *is a standard Wiener process. Also, when* $\bar{c} = c > 0$, $H_{\alpha,c,\bar{c}}(r) = W(r) - rW(1)$.

The second part of the theorem shows that $S^{\mu}(c)$ has the standard intercept case KPSS limiting null distribution. Note that $S^{\mu}(c)$ is invariant to the initial value even in finite samples.[1]

The modified KPSS test is related to the prewhitened long-run variance estimator suggested by Sul, Phillips, and Choi (2005) for the KPSS test. Instead of the usual fixed upper bound of 0.97 for the AR(1) prewhitening coefficient, they suggest using $1 - T^{-1/2}$. The difference is that our AR(1) filtering uses $1 - \bar{c}T^{-1}$ and this is employed in *both* numerator and denominator of the test. Sul et al. (2005) note that the rate T^{-1} at which $1 - \bar{c}T^{-1}$ approaches unity makes it inappropriate for prewhitening when used in the denominator alone, but clearly it is valid in our context.

3. COMPARISONS WITH OPTIMAL STATIONARITY TESTS

We compare the asymptotic performance of $S^{\mu}(\bar{c})$ with that of the stationarity test of Müller (2005), denoted $Q^{\mu}(g)$ in the notation of that paper (with g corresponding to $\bar{c}$). This test is asymptotically optimal in a Gaussian setting when $g = c$, in the situation where the initial value is generated with $\alpha = 1$. The test was originally suggested as a point optimal test of the unit root null in Elliott (1999) but can equally be considered as a test of the null of stationarity, simply by considering the opposite tail of the distribution to the unit root case.

When testing the unit root null, Müller and Elliott (2003) consider the effect that the magnitude of the initial value in a stationary series has on the power of unit root tests. In the current context of stationarity testing, this translates into an issue of size control. Elliott and Müller (2005), in the unit root testing context, derive an asymptotically optimal test (based on a weighted average power criterion) whose power varies little across different initial values. Again, it can equally be considered as a stationarity test, where size should be fairly robust across differing initial values. We denote this test $Q^{\mu}(g,k)$. As regards choosing parameters, we follow the previous authors and set $\bar{c} = g = 10$ ($k = 3.8$).

https://doi.org/10.1017/S0266466607070156 Published online by Cambridge University Press

Limit distributions of the three statistics are simulated by approximating the Wiener process functionals involved using i.i.d.$N(0,1)$ variables, approximating the integrals by normalized sums of 5,000 steps. All experiments are based on 10,000 replications.

As in Müller (2005), we compare the tests by determining critical values for each such that rejection rates coincide at some prespecified value for $c = 0$ (power) and then examining their size for $c > 0$. In this way, one test dominates another if the former's rejection profile for $c > 0$ (size) lies consistently below that of the latter.[2]

Figure 1 shows the rejection profiles of the tests across c for a rejection rate of 0.50 when $c = 0$. When $\alpha = 1$, $Q^{\mu}(10)$ is the optimal test, and so it dominates the other two tests everywhere. Between $Q^{\mu}(10, 3.8)$ and $S^{\mu}(10)$, it is $Q^{\mu}(10, 3.8)$ that dominates, though not substantially. In fact, there is only ever a very modest amount of difference between the sizes of all three tests, with all having size close to zero when $c = 10$. The size of each test is also monotonically decreasing in c. For $\alpha > 1$, the rejection profiles are no longer monotonic in c. There is always a region of c for which the size of each test is greater than power. However, this effect differs significantly across tests. As α increases the size of $Q^{\mu}(10)$ rapidly approaches 1.00 for most c, demonstrating its sensitivity to departures of the initial value from its optimality point. The other two test sizes appear much less sensitive to increasing α. Between these two, whereas there is virtually nothing to choose for $\alpha = 2$, for all $\alpha \geq 3$ it is $S^{\mu}(10)$ that clearly dominates $Q^{\mu}(10, 3.8)$ in terms of exhibiting the least size distortion.

To compare finite-sample size, we consider the DGP (2) with $\mu = 0$ and ν_t generated by the MA(1) model

$$\nu_t = \epsilon_t - \theta\epsilon_{t-1}$$

with $\epsilon_t \sim$ i.i.d.$N(0,1)$. For $Q^{\mu}(10)$ and $Q^{\mu}(10, 3.8)$, we follow Müller (2005) and estimate ω^2 using residuals from an AR(1) regression of demeaned y_t in place of the $r_{\bar{c},t}$ in (3). For all tests we use the quadratic spectral (QS) kernel for $\lambda(.)$ and employ the automatic bandwidth selection of Newey and West (1994).

For $T = 200$ and $c = 10$, Table 1 shows the empirical size of the tests using nominal 0.05 level asymptotic null critical values for various values of θ and

Table 1. Empirical sizes at the nominal 0.05 level

θ	0.0			0.6			−0.6		
α	1	3	5	1	3	5	1	3	5
$S^{\mu}(10)$	0.046	0.046	0.046	0.021	0.021	0.021	0.051	0.051	0.051
$Q^{\mu}(10)$	0.060	0.592	0.999	0.000	0.000	0.016	0.096	0.711	0.999
$Q^{\mu}(10, 3.8)$	0.060	0.091	0.130	0.000	0.000	0.000	0.088	0.116	0.149

https://doi.org/10.1017/S0266466607070156 Published online by Cambridge University Press

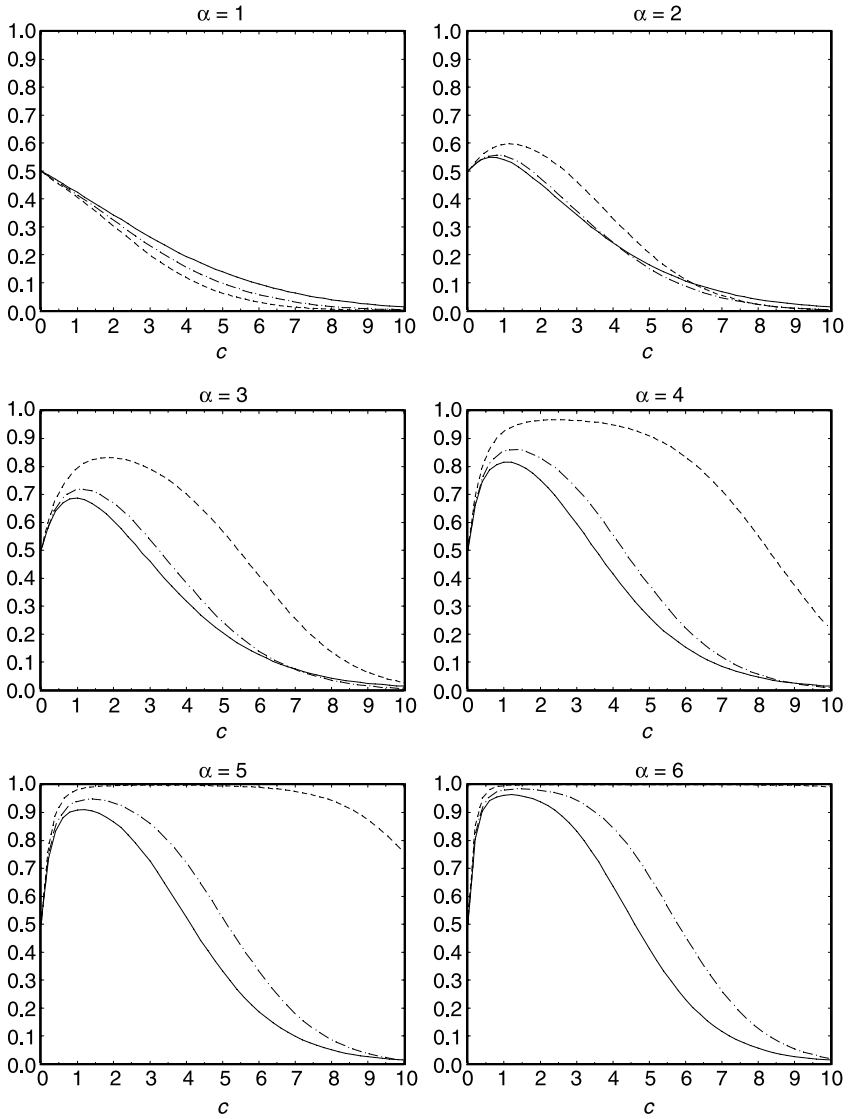


FIGURE 1. Asymptotic size and power: ———, $S^{\mu}(10)$; - - - - -, $Q^{\mu}(10)$; —·—·—, $Q^{\mu}(10, 3.8)$.

α.[3] Here $S^{\mu}(10)$ controls size quite well across θ (note that its size here is invariant to α). In comparison, both $Q^{\mu}(10)$ and $Q^{\mu}(10, 3.8)$ are oversized for $\theta = 0.0, -0.6$ and undersized for $\theta = 0.6$. The oversizing problems also increase with the magnitude of α—and at a drastic rate in the case of $Q^{\mu}(10)$, in keeping with what our asymptotic results would predict.

https://doi.org/10.1017/S0266466607070156 Published online by Cambridge University Press

Overall, our findings indicate that despite its optimality, $Q^{\mu}(10)$ is too susceptible to severe oversizing problems when $\alpha > 1$ to be recommended for practical use. Although $Q^{\mu}(10, 3.8)$ suffers rather less in this same situation, it is still the case that $S^{\mu}(10)$ generally displays the more robust size control asymptotically for a given level of power and also has better finite-sample size properties. We therefore conclude that, in circumstances when there is some uncertainty about the generation of the initial value, the modified KPSS test is likely to provide the most dependable inference of those considered here.

4. PROOF OF THEOREM

As the $r_{\bar{c},t}$ are invariant to μ, we can set $\mu = 0$ without loss of generality, so that $y_t = w_t$. It also proves convenient to define $z_t = w_t - w_1$ and rewrite $r_{\bar{c},t}$ in the form

$$r_{\bar{c},t} = (z_t - \rho_{\bar{c},T} z_{t-1}) - m_{\bar{c},z},$$

$$m_{\bar{c},z} = \frac{\sum_{t=2}^{T} (z_t - \rho_{\bar{c},T} z_{t-1})}{(T-1)}.$$

Then

$$T^{-1/2} \sum_{i=2}^{t} r_{\bar{c},i} = T^{-1/2} \sum_{i=2}^{t} (z_i - \rho_{\bar{c},T} z_{i-1}) - tT^{-1} \cdot T^{1/2} m_{\bar{c},z}, \tag{4}$$

and we may write the first right-hand side term of (4) as

$$\begin{aligned} T^{-1/2} \sum_{i=2}^{t} (z_i - \rho_{\bar{c},T} z_{i-1}) &= T^{-1/2} \sum_{i=2}^{t} (z_i - \{1 - \bar{c}T^{-1}\} z_{i-1}) \\ &= T^{-1/2} \sum_{i=2}^{t} \Delta z_i + \bar{c}T^{-3/2} \sum_{i=2}^{t} z_{i-1} \\ &= T^{-1/2}(z_t - z_2) + \bar{c}T^{-3/2} \sum_{i=3}^{t} z_{i-1} \end{aligned}$$

on noting that $z_1 = 0$. It is shown in Elliott (1999) that $T^{-1/2} z_{[rT]} \Rightarrow \omega K_{\alpha,c}(r)$ where $K_{\alpha,c}(r)$ is defined as in Theorem 1 and $W(r)$ is the Wiener process defined by $\sum_{i=2}^{[Tr]} \nu_i \Rightarrow \omega W(r)$. So, setting $t = [rT]$ we find that

$$T^{-1/2} \sum_{i=2}^{[rT]} (z_i - \rho_{\bar{c},T} z_{i-1}) \Rightarrow \omega \left\{ K_{\alpha,c}(r) + \bar{c} \int_0^r K_{\alpha,c}(s)\, ds \right\}.$$

https://doi.org/10.1017/S0266466607070156 Published online by Cambridge University Press

For the second right-hand side term of (4), note that

$$[rT]T^{-1}\cdot T^{1/2}m_{\bar{c},z} = [rT]T^{-1}.\frac{T}{(T-1)}\,T^{-1/2}\sum_{t=2}^{T}(z_t - \rho_{\bar{c},T}\,z_{t-1})$$

$$\Rightarrow \omega r\left\{K_{\alpha,c}(1) + \bar{c}\int_0^1 K_{\alpha,c}(s)\,ds\right\}.$$

Hence,

$$T^{-1/2}\sum_{i=2}^{[rT]} r_{\bar{c},i} \Rightarrow \omega\left[K_{\alpha,c}(r) + \bar{c}\int_0^r K_{\alpha,c}(s)\,ds - r\left\{K_{\alpha,c}(1) + \bar{c}\int_0^1 K_{\alpha,c}(s)\,ds\right\}\right]$$

$$= \omega H_{\alpha,c,\bar{c}}(r).$$

Thus, via the continuous mapping theorem,

$$T^{-1}\sum_{t=2}^{T}\left(T^{-1/2}\sum_{i=2}^{t} r_{\bar{c},i}\right)^2 \Rightarrow \omega^2\int_0^1 H_{\alpha,c,\bar{c}}(r)^2\,dr.$$

That $\hat{\omega}^2_{\bar{c}} \xrightarrow{p} \omega^2$ follows simply by noting that for $t = 2,\ldots,T$ we can write

$$y_t - \rho_{\bar{c},T}\,y_{t-1} = \nu_t + (c-\bar{c})T^{-1}w_{t-1},$$

and so

$$r_{\bar{c},t} = \nu_t + (c-\bar{c})T^{-1}w_{t-1} - \frac{\sum_{t=2}^{T}\{\nu_t + (c-\bar{c})T^{-1}w_{t-1}\}}{(T-1)}$$

$$= \nu_t + o_p(1).$$

To show the second part of the theorem note that when $\bar{c} = c > 0$ we have, for $t = 2,\ldots,T$,

$$r_{\bar{c},t} = (w_t - \rho_{c,T}\,w_{t-1}) - m_{c,w}$$

$$= v_t - m_{c,w},$$

$$m_{c,w} = \frac{\sum_{t=2}^{T}(w_t - \rho_{c,T}\,w_{t-1})}{(T-1)}$$

$$= \frac{\sum_{t=2}^{T} v_t}{(T-1)},$$

https://doi.org/10.1017/S0266466607070156 Published online by Cambridge University Press

and so $r_{\bar{c},t}$ invariant to α and using standard results

$$T^{-1/2}\sum_{i=2}^{[rT]} r_{\bar{c},i} \Rightarrow \omega[W(r) - rW(1)]. \quad \blacksquare$$

NOTES

1. If (2) also includes a linear trend term, the modified statistic is formed from the OLS detrended $y_t - \rho_{\bar{c},T} y_{t-1}$, $t = 2,\ldots,T$. Redefining $r_{\bar{c},t}$ accordingly, the analogue to Theorem 1 follows by replacing $H_{\alpha,c,\bar{c}}(r)$ with $H_{\alpha,c,\bar{c}}(r) - 6r(1-r)\int_0^1 H_{\alpha,c,\bar{c}}(s)\,ds$. Also, $S^{\mu}(c)$ has the usual trend case KPSS distribution.

2. Note that when $c = 0$ all the tests are invariant to α, so that the same critical values apply for all α.

3. For $S^{\mu}(10)$ this is the standard KPSS value of 0.460. Unlike $S^{\mu}(10)$, neither $Q^{\mu}(10)$ nor $Q^{\mu}(10, 3.8)$ is invariant to α when $c = 10$. For these two tests we use asymptotic critical values calculated assuming $\alpha = 1$. This is appropriate for $Q^{\mu}(10)$ as $\alpha = 1$ is the value at which it is optimal. Also, it is not unreasonable for $Q^{\mu}(10, 3.8)$ because this test is designed to have critical values that are fairly insensitive to α.

REFERENCES

Caner, M. & L. Kilian (2001) Size distortions of tests of the null hypothesis of stationarity: Evidence and implications for the PPP debate. *Journal of International Money and Finance* 20, 639–657.

Elliott, G. (1999) Efficient tests for a unit root when the initial observation is drawn from its unconditional distribution. *International Economic Review* 40, 767–783.

Elliott, G. & U. Müller (2006) Minimizing the impact of the initial condition on testing for unit roots. *Journal of Econometrics* 135, 285–310.

Elliott, G. & J.H. Stock (2001) Confidence intervals for autoregressive coefficients near one. *Journal of Econometrics* 103, 155–181.

Kwiatkowski, D., P. Phillips, P. Schmidt, & Y. Shin (1992) Testing the null hypothesis of stationarity against the alternative of a unit root. *Journal of Econometrics* 54, 159–178.

Müller, U. (2005) Size and power of tests for stationarity in highly autocorrelated time series. *Journal of Econometrics* 128, 195–213.

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

Newey, W. & K. West (1994) Automatic lag selection in covariance matrix estimation. *Review of Economic Studies* 61, 631–653.

Stock, J.H. & M.W. Watson (1999) Business cycle fluctuations in U.S. macroeconomic time series. In J.B. Taylor & M. Woodford (eds.), *Handbook of Macroeconomics,* vol. 1, pp. 3–64. Elsevier.

Sul, D., P.C.B. Phillips, & C. Choi (2005) Prewhitening bias in HAC estimation. *Oxford Bulletin of Economics and Statistics* 67, 517–546.

https://doi.org/10.1017/S0266466607070156 Published online by Cambridge University Press