Testing for a unit root by frequency domain regression*

In Choi
Ohio State University, Columbus, OH 43210-1172, USA

Peter C.B. Phillips
Yale University, New Haven, CT 06520, USA

Frequency domain tests for the presence of a unit root are developed. Their limit distributions are derived under the assumption of weakly stationary errors and are free of nuisance parameters. Results on test consistency are also reported. Monte Carlo simulations are performed to study the size and power of the proposed tests in finite samples. The computations indicate that the frequency domain tests have stable size and good power in finite samples for a variety of error-generating mechanisms. We conclude that the frequency domain tests have some good performance characteristics in relation to time domain procedures, although they are also susceptible to size distortion when there is negative serial correlation in the errors.

1. Introduction

Testing for the presence of a unit root in autoregressive time series models has been a popular topic in both the recent econometric and statistical literature. The testing procedures outlined in Fuller (1976) and Dickey and Fuller (1979) under iid errors have become standard and have been employed in various empirical applications. Recent articles by Fuller (1984) and Dickey, Bell, and Miller (1986) review the literature in the field up to around 1985. Since then there has been a large and growing literature on time series with a unit root, and many new tests have been developed.

Said and Dickey (1984), extending the t-ratio method of Dickey and Fuller (1979, 1981), proposed a test for the presence of a unit root in models with ARMA errors of an unknown order based on long autoregression. The same authors (1985) also applied the maximum likelihood method to develop tests for a unit root in ARIMA models of a known order and reported some simulations results. Phillips (1987) and Phillips and Perron (1988) took a semiparametric approach and developed unit root test statistics that were applicable to models with quite general weakly dependent errors. The relative performance of the Phillips–Perron and Said–Dickey (1984) test statistics are studied in Schwert (1987) and Phillips and Perron (1988). Asymptotic results favor the $Z(\hat{\alpha})$ procedure of Phillips and Perron, but simulation results indicate that this method suffers size distortions in finite samples when there is negative serial correlation in the errors. Unfortunately, the Said–Dickey test also suffers size distortions and has lower power than the Phillips–Perron test when the errors follow moving average processes. Therefore there is a need for new procedures which overcome these deficiencies of existing tests.

The present paper adopts a frequency domain approach to develop tests for a unit root. The frequency domain (or spectral regression) approach has been used in the past to efficiently estimate the parameters in regression models with fixed or strictly exogenous regressors. Hannan's (1963) efficient estimator is the cornerstone of subsequent work. The rationale for the approach is that only minimal assumptions like stationarity and weak dependence are required for the theory to apply. Engle and Gardner (1976) took advantage of this feature to estimate a coefficient parameter in a standard regression model under various dynamic specifications for the errors. It was found that the frequency domain estimator performs quite well at moderate sample sizes. The spectral regression method was also applied to regression models with dynamic regression in Espasa and Sargan (1977) and Engle (1980). Readers are referred to Granger and Engle (1985) for a review of related applications. Recently, Phillips (1991) has shown that spectral regression methods may be successfully used in models with nonstationary regressors. In that paper the Hannan efficient estimator is used to obtain consistent and asymptotically efficient estimators of long-run equilibrium parameters in error correction models. We shall demonstrate that a similar approach works well in the present context of unit root tests.

The merits of frequency domain methods are numerous. First, the non-parametric treatment of the errors means that it is not necessary to be explicit about the dynamic specification of the errors. Second, the limit distributions of the frequency domain test statistics have no nuisance parameters since the problem of temporal dependence becomes one of heteroskedasticity in the frequency domain and this is eliminated by the GLS transform. Third, we may test the hypothesis of a root on the unit circle at zero frequency. Both full-band and limited-band spectral estimators can be employed for this purpose. Fourth, simulation results, which we will report in section 6, show that the full-band

frequency domain tests display good power and stable size for various errors specifications. But like the $Z(\hat\alpha)$ and Said–Dickey tests, the new tests also suffer size distortion when there is negative serial correlation in the errors.

This paper is organized as follows. Section 2 introduces our models and assumptions. Section 3 defines the tests in the frequency domain. Section 4 develops the limiting distribution theory. Section 5 considers consistency of the tests. Section 6 reports simulation results concerning the power and size of the new test statistics in finite samples. Section 7 concludes the paper. Appendix 1 explains how to calculate the Hannan efficient estimator, and appendix 2 contains proofs.

A few words on our notation: All the limits are taken as $T \to \infty$ unless otherwise specified. W denotes a standard Brownian motion; W, W_1, $\int_0^1 r\,dW$, $\int_0^1 rW$, $\int_0^1 W^2$, $\int_0^1 W$, and $\int_0^1 W d W$ are understood to be $W(r)$, $W(1)$, $\int_0^1 r\,dW(r)$, $\int_0^1 rW(r)$, $\int_0^1 W^2(r)dr$, $\int_0^1 W(r)dr$, and $\int_0^1 W(r)dW(r)$, respectively. Brownian motion with variance ζ^2 is written as $BM(\zeta^2)$. $\sum_{t=1}^{T}$ is denoted simply as $\sum$, unless otherwise specified. The symbol $\Rightarrow$ signifies weak convergence, and the symbols $\equiv$ stands for equality in distribution.

2. Preliminaries

The univariate time series models that concern us are

$$y_t = \alpha y_{t-1} + u_t, \qquad t = 1, 2, \ldots, T, \tag{1}$$

$$y_t = \mu + \alpha y_{t-1} + u_t, \tag{2}$$

$$y_t = \mu + \beta t + \alpha y_{t-1} + u_t. \tag{3}$$

The error process $\{u_t\}$ is stationary with continuous spectral density $f_{uu}(\lambda) > 0$ over $-\pi < \lambda \leq \pi$. We shall take an interest in testing for the presence of a unit root in models (1), (2), and (3) against the alternative of stationarity, so that the null and alternative hypotheses are $H_0: \alpha = 1$ and $H_1: |\alpha| < 1$.

We shall assume that the partial sum process $S_t = \sum_{j=1}^{T} u_j$ satisfies the invariance principle

$$T^{-1/2} S_{[Tr]} \Rightarrow B(r) \equiv BM(\zeta^2), \qquad 0 \leq r \leq 1, \tag{4}$$

where $\zeta^2 = 2\pi f_{uu}(0) > 0$ is the 'long-run' variance of u_t. We decompose ζ^2 as

$$\zeta^2 = \sigma^2 + 2\gamma,$$

where

$$\sigma^2 = E(u_0^2), \qquad \gamma = \sum_{k=1}^{\infty} E(u_0 u_k), \tag{5}$$

and we define $\delta = \sigma^2 + \gamma$. The series that defines γ in (5) is assumed to converge absolutely so that the spectrum $f_{uu}(\lambda)$ is uniformly continuous over $[-\pi, \pi]$. Under condition (4), we have the following weak convergence of the sample covariance between S_t and u_t, viz.

$$T^{-1}\sum S_t u_t \Rightarrow \int_0^1 B \, dB + \delta \, . \tag{6}$$

The conditions under which (4) holds are quite weak. They involve rather mild moment and weak dependence requirements which are satisfied by a wide class of time series, including stationary ARMA models whose innovations have finite · variance. These conditions are discussed in detail in earlier work [see Phillips (1987), and Phillips (1988) for a review].

3. Frequency domain estimators and unit root tests

The spectral estimates have the same general form:

$$\hat{f}_{ab}(\lambda) = \frac{1}{2\pi} \sum_{n=-M}^{M} k\left(\frac{n}{M}\right) C_{ab}(n) e^{-in\lambda} \, ,$$

where

$$C_{ab}(n) = T^{-1} \sum_{t=1}^{T-n} a_t b_{t+n} \, ,$$

and where the convergence factor or lag window $k(x)$ is a bounded even function with $k(0) = 1$, vanishing outside the domain $[-1, 1]$. Using this general form, the frequency domain estimates of α for the regression models (1), (2), and (3) are defined as in Phillips (1991) and Hannan (1963) as follows, respectively:

$$\hat{\alpha} = Y_{y_L y}/X_{y_L y_L} \, , \tag{7}$$

$$\tilde{\alpha} = (X_{11} Y_{y_L y} - X_{1 y_L} Y_{1y})/(X_{11} X_{y_L y_L} - X_{1y_L}^2) \, , \tag{8}$$

$$\alpha^* = \{(X_{1t}X_{ty_L} - X_{1y_L}X_{tt})Y_{1y} - (X_{11}X_{ty_L} - X_{1y_L}X_{1t})Y_{ty}$$

$$+ (X_{11}X_{tt} - X_{1t}^2)Y_{yy_L}\}/(X_{11}X_{tt}X_{y_L y_L} + X_{1t}X_{ty_L}X_{1y_L}$$

$$+ X_{1y_L}X_{1t}X_{ty_L} - X_{1y_L}X_{tt}X_{1y_L} - X_{ty_L}^2 X_{11} - X_{y_L y_L}X_{1t}^2), \tag{9}$$

where

$$X_{11} = \frac{1}{2M} \sum_{j=-M+1}^{M} \hat{f}_{11}(\omega_j) \hat{f}_{uu}(\omega_j)^{-1},$$

$$X_{1t} = \frac{1}{2M} \sum_{j=-M+1}^{M} \hat{f}_{1t}(\omega_j) \hat{f}_{uu}(\omega_j)^{-1},$$

$$X_{1y_L} = \frac{1}{2M} \sum_{j=-M+1}^{M} \hat{f}_{1y_L}(\omega_j) \hat{f}_{uu}(\omega_j)^{-1},$$

$$X_{tt} = \frac{1}{2M} \sum_{j=-M+1}^{M} \hat{f}_{tt}(\omega_j) \hat{f}_{uu}(\omega_j)^{-1},$$

$$X_{ty_L} = \frac{1}{2M} \sum_{j=-M+1}^{M} \hat{f}_{ty_L}(\omega_j) \hat{f}_{uu}(\omega_j)^{-1},$$

$$X_{y_L y_L} = \frac{1}{2M} \sum_{j=-M+1}^{M} \hat{f}_{y_L y_L}(\omega_j) \hat{f}_{uu}(\omega_j)^{-1},$$

$$Y_{1y} = \frac{1}{2M} \sum_{j=-M+1}^{M} \hat{f}_{1y}(\omega_j) \hat{f}_{uu}(\omega_j)^{-1},$$

$$Y_{ty} = \frac{1}{2M} \sum_{j=-M+1}^{M} \hat{f}_{ty}(\omega_j) \hat{f}_{uu}(\omega_j)^{-1},$$

$$Y_{y_L y} = \frac{1}{2M} \sum_{j=-M+1}^{M} \hat{f}_{y_L y}(\omega_j) \hat{f}_{uu}(\omega_j)^{-1}.$$

Note that the subscript '1' denotes a unit vector and 't' a vector of linear time trend.

Because our main concern lies in estimating the autoregressive coefficient and testing for a unit root, we may define the estimates of the autoregressive coefficient using detrended models. That is, we transform models (2) and (3) as

$$\tilde{y}_t = \alpha \tilde{y}_{t-1} + \tilde{u}_t,$$
(2)'

$$y_t^* = \alpha y_{t-1}^* + u_t^*,$$
(3)'

where

$$y_t = \tilde{\gamma}_0 + \tilde{y}_t,$$

$$y_t = \hat{\gamma}_0 + \hat{\gamma}_1 t + y_t^*.$$

Using these models, we may define the frequency domain estimates of α as follows:

$$\tilde{\alpha}_{DT} = Y_{\tilde{y}_L \tilde{y}} / X_{\tilde{y}_L \tilde{y}_L} \,, \tag{10}$$

$$\alpha^*_{DT} = Y_{y_L^* y^*} / X_{y_L^* y_L^*} \,, \tag{11}$$

where the spectral density estimates are defined in the same way as before using the transformed variables $\tilde{y}_t$ and y_t^*. One merit of these estimators is that the number of estimated spectral densities is substantially reduced relative to the estimators $\tilde{\alpha}$ and α^*.

Under the null hypothesis the spectral density of y_t is

$$f_{yy}(\lambda) = |1 - e^{-i\lambda}|^{-2} f_{uu}(\lambda) \,,$$

which has a pole at the origin $(\lambda = 0)$ characterized by the local behavior $f_{yy}(\lambda) \sim \zeta^2 / 2\pi\lambda^2$ as $\lambda \to 0$. The singularity of $f_{yy}(\lambda)$ at $\lambda = 0$ is of course the manifestation in the frequency domain of the nonstationarity in y_t under the null. Interestingly, although $f_{yy}(\lambda)$ is undefined at $\lambda = 0$, we may still construct conventional spectral estimates at the origin. Upon restandardization, the estimate $\hat{f}_{yy}(0)$ is meaningful and converges weakly, but not in probability, to well-defined random elements. Correspondingly, we define

$$f_{yu}(\lambda) = [1 - e^{i\lambda}]^{-1} f_{uu}(\lambda) \,,$$

$$f_{y_L y_L}(\lambda) = |e^{i\lambda}|^2 f_{yy}(\lambda) = f_{yy}(\lambda) \,,$$

$$f_{y_L u}(\lambda) = e^{i\lambda} f_{yu}(\lambda) \,.$$

None of these are defined at $\lambda = 0$, but the conventional spectral estimates likewise converge weakly to well-defined random elements upon suitable standardization.

In (7), (8), (9), (10), and (11), we use the fundamental frequencies

$$\omega_j = \pi j / M, \qquad j = -M + 1, \ldots, M \,,$$

for M integer. The spectral estimates that appear in these formulae may then be regarded as applying within a band of width π / M centered on ω_j. Thus, to obtain $\hat{f}_{uu}(\omega_j)$ we may use the smoothed periodogram estimate

$$\hat{f}_{uu}(\omega_j) = \frac{M}{T} \sum_{\mathcal{A}_j} k(\omega_j) w_u(\lambda_s) w_u(\lambda_s)^* \,,$$

where $k(\omega_j)$ is a lag window, $w_u(\lambda) = (2\pi T)^{-1/2}\sum u_t e^{it\lambda}$, and the summation is over $\lambda_s \in \mathscr{B}_j = (\omega_j - \pi/2M, \omega_j + \pi/2M)$. Then $\hat{f}_{uu}(\omega_j)$ is, in effect, a weighted average of $m = [T/M]$ neighboring periodogram ordinates around the frequency ω_j, as interpreted in Hannan (1970, p. 274). As usual, we shall require the bandwidth parameter $M \to \infty$, but in such a way that $M = o(T^{1/2})$ as $T \to \infty$ [as in Hannan (1970, p. 489)].

The estimators we have defined are conventional spectral regression estimates and follow from formulae given in Hannan (1963). In popular parlance, they are the Hannan efficient estimates. What does differ from convention is the autoregressive context in which the estimates are being used and the asymptotic theory that applies to them. The autoregressive context is of importance since y_{t-1} and u_t are in general coherent series, due to the temporal dependence in u_t. This is a major departure from the regression model context in which spectral regression estimators were first developed.

We may also formulate band spectrum estimators that employ only the zero frequency estimates. For instance, the band spectrum estimator for model (1) is simply

$$\hat{\alpha}_0 = \hat{f}_{y_L y}(0)/\hat{f}_{y_L y_L}(0) .$$

Test statistics based on $\hat{\alpha}_0$ may also be constructed. They are $T(\hat{\alpha}_0 - 1)$ and the spectral analogue of the t-ratio, $t(\hat{\alpha}_0) = T(\hat{\alpha}_0 - 1)/\{T\hat{f}_{y_L y_L}(0)^{-1}\hat{f}_{uu}(0)\}^{1/2}$. As we will show in section 5, these tests, in fact, do not possess good properties.

The test statistics we shall use are based on the full-band spectral estimates. They are $T(\hat{\alpha} - 1)$, $T(\tilde{\alpha} - 1)$, $T(\alpha^* - 1)$, $T(\tilde{\alpha}_{DT} - 1)$, and $T(\alpha^*_{DT} - 1)$. We also consider the spectral analogues of the regression t-statistics associated with $\hat{\alpha}$, $\tilde{\alpha}$, α^*, $\tilde{\alpha}_{DT}$, and α^*_{DT}. These are given by

$$t(\hat{\alpha}) = T(\hat{\alpha} - 1)/(TX_{y_L y_L}^{-1})^{1/2} ,$$

$$t(\tilde{\alpha}) = T(\tilde{\alpha} - 1)/\{T(X_{11}X_{y_L y_L} - X_{1y_L}^2)^{-1}X_{11}\}^{1/2} ,$$

$$t(\alpha^*) = T(\hat{\alpha}^* - 1)/\{T(X_{11}X_{tt}X_{y_L y_L} + X_{1t}X_{ty_L}X_{1y_L}$$

$$+ X_{1y_L}X_{1t}X_{ty_L} - X_{1y_L}X_{tt}X_{1y_L} - X_{ty_L}^2 X_{11}$$

$$- X_{y_L y_L}X_{1t}^2)^{-1}(X_{11}X_{tt} - X_{1t}^2)\}^{1/2},$$

$$t(\tilde{\alpha}_{DT}) = T(\tilde{\alpha}_{DT} - 1)/(TX_{\tilde{y}_L \tilde{y}_L}^{-1})^{1/2} ,$$

$$t(\alpha^*_{DT}) = T(\alpha^*_{DT} - 1)/(TX_{y_L^* y_L^*}^{-1})^{1/2} .$$

The variance estimates implicit in these t-ratios are based on the usual formulae for the estimated asymptotic variances of the spectral estimates. For instance, the variance estimate of $\hat{\alpha}$ is

$$\frac{1}{T}\left[\frac{1}{2M}\sum_{j=-M+1}^{M}\hat{f}_{y_L y_L}(\omega_j)/\hat{f}_{uu}(\omega_j)\right]^{-1},$$

as in Hannan (1970, p. 442).

4. Asymptotic theory

Our attention will concentrate on the test statistics $T(\hat{\alpha}-1)$, $T(\tilde{\alpha}-1)$, $T(\alpha^*-1)$, $T(\tilde{\alpha}_{DT}-1)$, $T(\alpha_{DT}^*-1)$, $t(\hat{\alpha})$, $t(\tilde{\alpha})$, $t(\alpha^*)$, $t(\tilde{\alpha}_{DT})$, and $t(\alpha_{DT}^*)$.

Theorem 1.   Suppose that the assumptions made in section 2 hold.

(a) *For the regression model (1), under $\alpha = 1$,*

(i) $\displaystyle T(\hat{\alpha}-1) \Rightarrow \int_0^1 W\,\mathrm{d}W \Big/ \int_0^1 W^2\,,$

(ii) $\displaystyle t(\hat{\alpha}) \Rightarrow \int_0^1 W\,\mathrm{d}W \Big/ \left[\int_0^1 W^2\right]^{1/2}.$

(b) *For the regression model (2), under $(\mu, \alpha) = (0, 1)$,*

(i) $\displaystyle T(\tilde{\alpha}-1),\, T(\tilde{\alpha}_{DT}-1) \Rightarrow \int_0^1 \tilde{W}\,\mathrm{d}W \Big/ \int_0^1 \tilde{W}^2\,,$

(ii) $\displaystyle t(\tilde{\alpha}),\, t(\tilde{\alpha}_{Dt}) \Rightarrow \int_0^1 \tilde{W}\,\mathrm{d}W \Big/ \left[\int_0^1 \tilde{W}^2\right]^{1/2}.$

(c) *For the regression model (3), under $(\beta, \alpha) = (0, 1)$,*

(i) $\displaystyle T(\alpha^*-1),\, T(\alpha_{DT}^*-1) \Rightarrow \int_0^1 W^*\,\mathrm{d}W \Big/ \int_0^1 W^{*2}\,.$

(ii) $\displaystyle t(\alpha^*),\, t(\alpha_{Dt}^*) \Rightarrow \int_0^1 W^*\,\mathrm{d}W \Big/ \left[\int_0^1 W^{*2}\right]^{1/2},$

where

$$\tilde{W} = W - \int_0^1 W\,,$$

$$W^* = W - 4\left(\int_0^1 W - \tfrac{3}{2}\int_0^1 rW\right) + 6r\left(\int_0^1 W - 2\int_0^1 rW\right).$$

Remarks

(i) The limit distributions given by (a)–(c) in the above theorem are all free of nuisance parameters. So no serial correlation corrections such as those employed in the tests of Phillips (1987) and Phillips and Perron (1988) are needed. The serial dependence in the error process u_t is, of course, automatically taken care of by the Fourier transformation of the data. What is of additional interest is that no correction is needed for the fact that y_{t-1} and u_t are coherent and even contemporaneously correlated when there is serial dependence in u_t. This is, as usual, explained by the fact that y_{t-1} is an integrated process and the signal that it imparts is correspondingly an order of magnitude larger (in $T^{1/2}$) than the covariance of y_{t-1} and u_t.

(ii) For the estimates analyzed above, we need to estimate the error spectrum $\hat{f}_{uu}(\lambda)$. Moreover, we use estimates of this spectrum at the $2M$ frequencies $\{\omega_j : j = -M + 1, \dots, M\}$. This is to be distinguished from the time domain procedures in earlier work [see Phillips (1987) and Phillips and Perron (1988)] where spectral estimates are required only at the origin to achieve the appropriate semiparametric corrections. The regression leading to the estimates above is, of course, a weighted regression across frequencies and the heterogeneity in the spectrum over the frequencies $\{w_j\}$ is used to obtain efficient estimates in conventional weighted regression for stationary time series.

5. Test consistency

Under H_1, y_t is stationary and it is of interest to examine the behavior of the power functions of the tests as $T \to \infty$. It is simplest to work with the band spectral tests $S(\hat{\alpha}_0)$ and $t(\hat{\alpha}_0)$. We start by observing that under stationarity

$$\hat{f}_{y_L y_L}(0) \xrightarrow{p} f_{yy}(0) = f_{uu}(0)/(1 - \alpha)^2$$

and

$$\hat{f}_{uy_L}(0) \xrightarrow{p} f_{uy_L}(0) = f_{uy}(0) = f_{uu}(0)/(1 - \alpha) .$$

Then

$$\hat{\alpha}_0 = \alpha + \hat{f}_{uy_L}(0)/\hat{f}_{y_L y_L}(0) \xrightarrow{p} \alpha + (1 - \alpha) = 1 .$$

Thus, $\hat{\alpha}_0$ tends to unity even under the alternative hypothesis. This suggests that the $S(\hat{\alpha}_0)$ and $t(\hat{\alpha}_0)$ tests are unlikely to have good power.

As $T \to \infty$, the power properties depend on the behavior of the spectral estimates $\hat{f}_{uy_L}(0)$ and $\hat{f}_{y_L y_L}(0)$. Define the matrix of spectral estimates

$$g_T = \begin{bmatrix} \hat{f}_{uu}(0) & \hat{f}_{uy_L}(0) \\ \hat{f}_{y_L u}(0) & \hat{f}_{y_L y_L}(0) \end{bmatrix},$$

and, under H_1, set

$$g = \begin{bmatrix} f_{uu}(0) & f_{uy_L}(0) \\ f_{y_L u}(0) & f_{y_L y_L}(0) \end{bmatrix}$$

and

$$v = 2T/M \int_{-\infty}^{\infty} k(x)^2 \, dx,$$

where M is the bandwidth parameter and $k(\cdot)$ is the lag window employed in the spectral estimates in g_T. Then, from the asymptotic theory of spectral estimates for stationary time series [e.g., Hannan (1970, p. 289)], we have the following limit theory:

$$v^{1/2}(g_T - g) \Rightarrow N(0, V),$$

where

$$V = \left(\int_{-\infty}^{\infty} k(x) \, dx \right) g \otimes g.$$

Now we obtain

$$v^{1/2}(\hat{\alpha}_0 - 1) \Rightarrow (1/f_{y_L y_L}(0)) A_{01} - (f_{uy_L}(0)/(f_{y_L y_L}(0)^2) A_{11}$$

$$= \frac{(1 - \alpha)^2}{f_{uu}(0)} \{ A_{01} - (1 - \alpha) A_{11} \},$$

(12)

where

$$A = \begin{bmatrix} A_{00} & A_{01} \\ A_{10} & A_{11} \end{bmatrix} \equiv N(0, V).$$

We deduce that

$$T^{1/2}(\hat{\alpha}_0 - 1) = O_p(M^{1/2}),$$

under H_1. It follows that a two-sided test of H_0 using the statistic $T(\hat{\alpha}_0 - 1)$ is consistent as $T \to \infty$. But, in view of the symmetry of the limit distribution (12) about the origin, the power of a one-sided test of H_0 based on $T(\hat{\alpha}_0 - 1)$ tends to 0.50. So this test of H_0 is inconsistent.

In a similar way we find that

$$t(\hat{\alpha}_0) = O_p(M^{1/2}) \,,$$

under H_1. The power properties of the t-ratio test $t(\hat{\alpha}_0)$ would therefore seem to be worse than those of the coefficient-based test $T(\hat{\alpha}_0 - 1)$. For example, when $M = O(T^{1/5})$, which is a bandwidth choice that minimizes a mean squared error criterion [see Bartlett (1966, p. 368)], we have $T(\hat{\alpha}_0 - 1) = O_p(T^{3/5})$ and $t(\hat{\alpha}_0) = O_p(T^{1/10})$ under H_1.

Given these asymptotic results, neither $T(\hat{\alpha}_0 - 1)$ nor $t(\hat{\alpha}_0)$ can be expected to yield good power for the usual one-sided tests of a unit root against stationary alternatives. Moreover, $t(\hat{\alpha}_0)$ can be expected to have even lower power than $T(\hat{\alpha}_0 - 1)$.

The behavior of the full-band spectral tests is more complicated. Here we investigate the behavior of the statistics $T(\hat{\alpha} - 1)$ and $t(\hat{\alpha})$. Analysis of the other estimators follows analogously. The results depend on the manner of estimation of the error spectrum $f_{uu}(\omega)$ as employed in $\hat{\alpha}$. If we use differences $u_t = \Delta y_t$ in constructing $\hat{f}_{uu}(\omega)$, then it is easy to see that, under the stationary alternative H_1,

$$\hat{f}_{\Delta y \Delta y}(0) \xrightarrow{\text{p}} 0 \,. \tag{13}$$

This means that the $\omega_j = 0$ term (i.e., $j = 0$) dominates both the numerator and denominator of $\hat{\alpha}$. Multiplying through the $\hat{f}_{\Delta y \Delta y}(\omega)$ in both numerator and denominator then shows that, under H_1,

$$\hat{\alpha} \sim \hat{f}_{y_L y}(0)/\hat{f}_{y_L y_L}(0) = \hat{\alpha}_0 \,,$$

as $T \to \infty$. Thus, when the differences of y_t are used to estimate f_{uu}, we find that $\hat{\alpha}$ is asymptotically equivalent to $\hat{\alpha}_0$. Then the tests $T(\hat{\alpha} - 1)$ and $t(\hat{\alpha})$ behave like $T(\hat{\alpha}_0 - 1)$ and $t(\hat{\alpha}_0)$, respectively, and are therefore inconsistent.

However, when the error spectrum f_{uu} is estimated using regression residuals $\hat{u}_t = y_t - \hat{\alpha}_{OLS} y_{t-1}$, the results are quite different because (13) no longer applies. In general, because of serial dependence in the error process u_t, we find that the least squares coefficient $\hat{\alpha}_{OLS}$ is not consistent for α, and hence $\hat{f}_{uu}(\omega)$ is not consistent for $f_{uu}(\omega)$. In fact, under the alternative for stationary and ergodic y_t, we have

$$\hat{\alpha}_{OLS} \xrightarrow{\text{p}} E(y_t y_{t-1})/E(y_t^2) = \bar{\alpha}(|\bar{\alpha}| < 1) \,,$$

and then

$$\hat{f}_{uu}(\omega) \xrightarrow{\text{p}} |1 - \bar{\alpha}e^{i\omega}|^2 f_{yy}(\omega) = \bar{f}_{uu}(\omega) .$$

Intuitively, the spectral regression in this case amounts to estimating $\bar{\alpha}$ in the regression equation

$$y_t = \bar{\alpha}y_{t-1} + u_t .$$

Since $\hat{f}_{uu}(\omega)$ is a consistent estimator of the spectral density $\bar{f}_{uu}(\omega)$ for $u_t = y_t - \bar{\alpha}y_{t-1}$, it is easy to find, by using the same methods as in the proof of Lemma A.1, that $\hat{\alpha} \xrightarrow{\text{p}} \bar{\alpha}$ ($|\alpha| < 1$) under the alternative. Hence, we find that the tests based on $T(\hat{\alpha} - 1)$ and $t(\hat{\alpha})$ are consistent.

6. Experimental evidence

In this section, we report some simulation results investigating the power and size of frequency domain tests. The data were generated by models (1), (2), and (3) with the ARMA errors and with the initial value $e_0 = 0$. The sequence $\{e_t\}$ was generated as iid $N(0, 1)$ by the IMSL subroutine RNNOA. The sample size is $T = 100$, and 1,000 iterations were made to calculate the empirical power of the test statistics in one-sided tests. The nominal size was set to be 5%. Critical values at $T = 100$ were taken from Fuller (1976, pp. 371, 373). It is assumed that $\mu = 0$ for model (2) and $(\mu, \beta) = (0, 0)$ for model (3) both under the null and alternative.

Computation involved the following steps:

(a) Run an OLS regression in the time domain and compute residuals.
(b) Using the OLS residuals from step (a), calculate $\hat{f}_{uu}(\lambda)$.
(c) Detrend the series by regression on time polynomials, as necessary.
(d) Compute $\hat{f}_{y_L y}(\lambda)$ and $\hat{f}_{y_L y_L}(\lambda)$ using the original series or the detrended series form step (c).
(e) Calculate the test statistics using the spectral density estimates from steps (b) and (d).

Note that step (c) is skipped for model (1). The IMSL subroutines SSWD and CSSWD were used to compute the spectral density estimates. The use of these subroutines and the calculation of Hannan's efficient estimates are explained in detail in appendix 1.

The frequency domain test statistics we consider are $SC_1 = T(\hat{\alpha} - 1)$, $ST_1 = t(\hat{\alpha})$, $SC_2 = T(\tilde{\alpha}_{DT} - 1)$, $ST_2 = t(\tilde{\alpha}_{DT})$, $SC_3 = T(\alpha^*_{DT} - 1)$, and $ST_3 = t(\alpha^*_{DT})$. The tests based on $\tilde{\alpha}$ and α^* are asymptotically equivalent to those based on $\tilde{\alpha}_{DT}$ and α^*_{DT}. However, the tests using detrended series are more convenient in terms of computation, since we do not need to estimate as many spectral densities in computing these tests. We used the Tukey–Hanning window to estimate the spectral densities. We chose $M = \sqrt{T}/2$. As is discussed in Hannan (1970), we need $M/T \to 0$ and $M = O(T^{1/2})$. The above rule satisfies these conditions. Hence, for sample size $T = 100$, we set $M = 5$. Obviously, it would also be of interest to consider the performance of various data-based choices of M. But this is left for subsequent work.

In addition to the frequency domain tests, we report the empirical size and power of the Dickey–Fuller tests in table 5, which is based on 2,000 iterations. The finite sample performance of the frequency domain tests will be compared with that of the Dickey–Fuller tests with iid errors by using the computation results reported in table 5. More comprehensive simulation results are reported in Dickey, Bell, and Miller (1986).

In table 1(1), the size and power of the SC_1 and ST_1 tests under first-order moving average errors are reported. The nominal size is well maintained except when the moving average coefficients are negative, and the power is essentially the same as that of the Dickey–Fuller tests with iid errors (see table 5 for this). The experiments are extended to ARMA errors in parts (2), (3), and (4) of table 1. We observe that the nominal size is again well maintained except when the moving average coefficients are negative, or when the value of the AR coefficient of the error process is -0.6. The power is essentially the same as that of the Dickey–Fuller tests with iid errors as in part (1). We also report simulation results for the Said–Dickey (1984) t-test in table 2 to compare this with our frequency domain tests. The simulation results are based on 2,000 iterations. We find that the frequency domain tests are more powerful than the Said–Dickey test in general. However, the Said–Dickey test displays more stable size when the MA coefficient takes negative values. In table 3, we report the size and power of the SC_2 and ST_2 test. Again, we find that the power and size are virtually equivalent to those of the Dickey–Fuller tests with iid errors, and that there are size distortions when the moving average coefficients are negative. We may use $T(\tilde{\alpha} - 1)$ and $t(\tilde{\alpha})$ instead SC_2 and ST_2, but simulations (not reported here) showed these tests to be more susceptible to size distortions. In table 4, we report the performance of the SC_3 and ST_3 test. When the errors follow a first-order moving average process, we find that the power and size are almost equivalent to those of the Dickey–Fuller tests with iid errors, though again there are size distortions with zero and negative moving average coefficients. Also, we observe that the SC_3 and ST_3 become less powerful as the moving average coefficient takes larger positive values. When the errors follow an autoregressive process, we observe

Table 1

Monte Carlo power of SC_1 and ST_1 at $T = 100$ and $M = 5$.

Model: $y_t = \alpha y_{t-1} + u_t$

θ		α 0.85	1.00
		(1) $u_t = e_t + \theta e_{t-1}$	
0.5	SC_1	0.910	0.032
	ST_1	0.899	0.034
0.2	SC_1	0.956	0.050
	ST_1	0.947	0.048
0.0	SC_1	0.978	0.067
	ST_1	0.975	0.070
− 0.5	SC_1	1.000	0.300
	ST_1	1.000	0.295
− 0.8	SC_1	1.000	0.871
	ST_1	1.000	0.870
		(2) $u_t = 0.2u_{t-1} + e_t + \theta e_{t-1}$	
0.5	SC_1	0.819	0.021
	ST_1	0.811	0.025
0.2	SC_1	0.879	0.026
	ST_1	0.876	0.028
− 0.5	SC_1	1.000	0.185
	ST_1	1.000	0.191
− 0.8	SC_1	1.000	0.759
	ST_1	1.000	0.753
		(3) $u_t = -0.2u_{t-1} + e_t + \theta e_{t-1}$	
0.5	SC_1	0.956	0.050
	ST_1	0.948	0.049
0.2	SC_1	0.978	0.067
	ST_1	0.975	0.070
− 0.5	SC_1	1.000	0.426
	ST_1	1.000	0.424
− 0.8	SC_1	1.000	0.942
	ST_1	1.000	0.936

Table 1 (continued)

		α	
θ		0.85	1.00

(4) $u_t = -0.6u_{t-1} + e_t + \theta e_{t-1}$

θ		0.85	1.00
0.5	SC_1	0.981	0.074
	ST_1	0.980	0.073
0.2	SC_1	0.997	0.120
	ST_1	0.994	0.116
−0.5	SC_1	1.000	0.715
	ST_1	1.000	0.700
−0.8	SC_1	1.000	1.000
	ST_1	1.000	0.999

Table 2

Monte Carlo power of the Said–Dickey t-test at $T = 100$.

Model: $y_t = \alpha y_{t-1} + u_t$

	Size				Power			
θ	$l=1$	$l=2$	$l=5$	$l=7$	$l=1$	$l=2$	$l=5$	$l=7$
	(1) $u_t = 0.2u_{t-1} + e_t + \theta e_{t-1}$							
0.5	0.236	0.023	0.039	0.039	1.000	0.901	0.616	0.358
0.2	0.125	0.033	0.038	0.040	0.996	0.862	0.441	0.264
−0.5	0.205	0.068	0.040	0.040	0.992	0.542	0.163	0.107
−0.8	0.825	0.431	0.163	0.097	0.994	0.986	0.471	0.236
	(2) $u_t = -0.2u_{t-1} + e_t + \theta e_{t-1}$							
0.5	0.100	0.032	0.037	0.039	0.989	0.662	0.365	0.214
0.2	0.045	0.038	0.036	0.039	0.870	0.570	0.259	0.172
−0.5	0.673	0.112	0.045	0.038	0.994	0.592	0.123	0.078
−0.8	0.992	0.576	0.209	0.122	0.994	0.994	0.553	0.259
	(3) $u_t = -0.6u_{t-1} + e_t + \theta e_{t-1}$							
0.5	0.047	0.041	0.038	0.040	0.833	0.562	0.233	0.156
0.2	0.209	0.043	0.039	0.042	0.979	0.403	0.175	0.124
−0.5	0.982	0.145	0.046	0.035	0.994	0.659	0.111	0.066
−0.8	0.994	0.654	0.235	0.128	0.999	0.994	0.599	0.270

Table 3

Monte Carlo power of SC_2 and ST_2 at $T = 100$ and $M = 5$.

Model: $y_t = \mu + \alpha y_{t-1} + u_t,\ u_t = e_t = e_t + \theta e_{t-1}$

θ		α	
		0.85	1.00
0.5	SC_2	0.515	0.020
	ST_2	0.344	0.037
0.2	SC_2	0.621	0.029˙
	ST_2	0.468	0.044
0.0	SC_2	0.775	0.046
	ST_2	0.627	0.052
− 0.5	SC_2	1.000	0.404
	ST_2	1.000	0.369
− 0.8	SC_2	1.000	0.978
	ST_2	1.000	0.969

that the nominal size is not well maintained, and that the power is lower than it is for moving average errors.

In summary, we find that the frequency domain tests show quite reasonable finite-sample performance in comparison with previous parametric tests under various error specifications. However, it needs to be borne in mind that the performance of the frequency domain tests is sensitive to how the Hannan efficient estimator is calculated. Different spectral windows (and different choices of bandwidth) will give different experimental results. Last, the SC_3 and ST_3 tests do not show good finite-sample performance when the errors follow an autoregressive process, which seems to accord with many other unit root tests, as the simulation results in DeJong, Nankervis, Savin, and Whiteman (1989) indicate.

7. Conclusion

We have proposed tests for a unit root that use frequency domain methods. The proposed test statistics do not involve nuisance parameters in their limiting distributions, and quite general temporal dependence is permitted for the errors. Simulation results show that the frequency domain tests have good power and reasonable size characteristics in finite samples, although, like previous parametric tests, there are size distortions under negative serially

Table 4

Monte Carlo power of SC_3 and ST_3 at $T = 100$ and $M = 5$.

Model: $y_t = \mu + \beta t + \alpha y_{t-1} + u_t, \mu = \beta = 0$

θ		α	
		0.85	1.00
		(1) $u_t = e_t + \theta e_{t-1}$	
0.5	SC_3	0.208	0.019
	ST_3	0.174	0.035
0.2	SC_3	0.333	0.038
	ST_3	0.281	0.051
0.0	SC_3	0.516	0.068
	ST_3	0.466	0.080
− 0.5	SC_3	0.997	0.611
	ST_3	0.997	0.589
− 0.8	SC_3	1.000	1.000
	ST_3	1.000	1.000
		(2) $u_t = \phi u_{t-1} + e_t$	
0.5	SC_3	0.046	0.004
	ST_3	0.033	0.015
0.2	SC_3	0.272	0.034
	ST_3	0.224	0.045
− 0.5	SC_3	0.971	0.328
	ST_3	0.959	0.318
− 0.8	SC_3	1.000	0.783
	ST_3	1.000	0.754

correlated errors. The tests are also conveniently robust to a wide class of dynamic specifications of the error terms, and hence can be used even when it is difficult to identify a parametric structure of the error process. We conclude, therefore, that the frequency domain tests have many convenient and appealing properties in relation to existing test procedures like the Said–Dickey t-test and the Phillips $Z(\alpha)$ and $Z(t)$ tests. They also have some interesting multivariate generalizations, which will be reported in later work.

Appendix 1: Calculating the Hannan efficient estimator

We used the IMSL subroutines SSWD and CSSWD to calculate spectral densities. Following the notation of the IMSL Stat/Library, we report the values

Table 5

Monte Carlo power of the Dickey–Fuller tests at $T = 100$.

	α	
	0.85	1.00
$(1)\ y_t = \alpha y_{t-1} + e_t$		
$T(\hat{\alpha} - 1)$	0.969	0.052
$t(\hat{\alpha})$	0.967	0.049
$(2)\ y_t = +\mu + \alpha y_{t-1} + e_t,\ \mu = 0$		
$T(\hat{\alpha} - 1)$	0.792	0.048
$t(\hat{\alpha})$	0.631	0.051
$(3)\ y_t = \mu + \beta t + \alpha y_{t-1} + e_t,\ \mu = \beta = 0$		
$T(\hat{\alpha} - 1)$	0.469	0.044
$t(\hat{\alpha})$	0.402	0.049

of the major input parameters for the subroutines, which we used in order to obtain the simulation results of section 6.

$$NOBS = 100,\quad XCNTR = YCNTR = 0,\quad NPAD = NOB - 1,$$

$$IFSCAL = 0\ ,$$

$$NF = 6,\quad F = [0, \pi/5, 2\pi/5, 3\pi/5, 4\pi/5, \pi],\quad TINT = 1,\ ISWVER = 4\ ,$$

$$NM = 1,\quad M = 5.$$

Programs interrupt due to overflow when we calculate the zero frequency using the IMSL subroutines. This problem can be avoided by using the IBM subroutine ERRSET(207, 251, − 1, 1), which suppresses interrupts due to overflow.

After calculating the spectral densities at the fundamental frequencies, we calculated the test statistics using eq. (7), (10), and (11) and the corresponding definitions for the t-ratios. We exploited the symmetry of spectral densities to calculate the test statistics, and hence calculating spectral densities at negative frequencies is not required.

Appendix 2: Proofs

Lemma A.1. Suppose that $\{u_t\}$ satisfies the conditions in Theorem 1. Then we have

(a) $\quad \dfrac{1}{2M}\displaystyle\sum_{j=-M+1}^{M}\hat{f}_{11}(\omega_j)f_{uu}^{-1}(\omega_j)\ \rightarrow\ 1/\zeta^2$,

(b) $\quad T^{-2}\dfrac{1}{2M}\displaystyle\sum_{j=-M+1}^{M}\hat{f}_{tt}(\omega_j)f_{uu}^{-1}(\omega_j)\ \rightarrow\ 1/(3\zeta^2)$,

(c) $\quad T^{-1}\dfrac{1}{2M}\displaystyle\sum_{j=-M+1}^{M}\hat{f}_{t1}(\omega_j)f_{uu}^{-1}(\omega_j)\ \rightarrow\ 1/(2\zeta^2)$,

(d) $\quad T^{-1/2}\dfrac{1}{2M}\displaystyle\sum_{j=-M+1}^{M}\hat{f}_{1y_L}(\omega_j)f_{uu}^{-1}(\omega_j)\ \Rightarrow\ \displaystyle\int_0^1 W/\zeta$,

(e) $\quad T^{-3/2}\dfrac{1}{2M}\displaystyle\sum_{j=-M+1}^{M}\hat{f}_{ty_L}(\omega_j)f_{uu}^{-1}(\omega_j)\ \Rightarrow\ \displaystyle\int_0^1 rW/\zeta$,

(f) $\quad T^{-1}\dfrac{1}{2M}\displaystyle\sum_{j=-M+1}^{M}\hat{f}_{y_Ly_L}(\omega_j)f_{uu}^{-1}(\omega_j)\ \Rightarrow\ \displaystyle\int_0^1 W^2$,

(g) $\quad T^{1/2}\dfrac{1}{2M}\displaystyle\sum_{j=-M+1}^{M}\hat{f}_{1u}(\omega_j)f_{uu}^{-1}(\omega_j)\ \Rightarrow\ W(1)/\zeta$,

(h) $\quad T^{-1/2}\dfrac{1}{2M}\displaystyle\sum_{j=-M+1}^{M}\hat{f}_{t,u}(\omega_j)f_{uu}^{-1}(\omega_j)\ \Rightarrow\ \displaystyle\int_0^1 r\,dW/\zeta$,

(i) $\quad \dfrac{1}{2M}\displaystyle\sum_{j=-M+1}^{M}\hat{f}_{y_Lu}(\omega_j)f_{uu}^{-1}(\omega_j)\ \Rightarrow\ \displaystyle\int_0^1 W\,dW$.

Proof. The proof is closely related to the proof of Theorem 3.1 of Phillips (1988a), and so we shall only give the essential details here. In addition, (d)–(h) use the weak convergence results in Phillips and Perron (1988).

(a) Using the Fourier series $f_{uu}^{-1}(\lambda) = (1/2\pi)\sum_{g=-\infty}^{\infty}d_g e^{ig\lambda}$, we have

$$\dfrac{1}{2M}\sum_{j=-M+1}^{M}\hat{f}_{11}(\omega_j)f_{uu}^{-1}(\omega_j)$$

$$=\dfrac{1}{2\pi}\sum_{g=-\infty}^{\infty}d_g\dfrac{1}{2M}\sum_{j=-M+1}^{M}e^{ig\pi j/M}\hat{f}_{11}(\pi j/M)$$

$$=\left[\dfrac{1}{2\pi}\right]^2\sum_{g=-\infty}^{\infty}d_g C_{11}(g)k(g/M)\ ,$$

where $g + 2lM = g$, $-M + 1 \leq g \leq M$ for some integer l, and $C_{11}(g) = T^{-1}\sum_{t=1}^{T-g} 1$. We have

$$k(g/M) \to 1 , \tag{A.1}$$

for all fixed g as $T \to \infty$, and hence, $M \to \infty$. From the Fourier series of $f_{uu}^{-1}(\lambda)$, we deduce that

$$\left[\frac{1}{2\pi}\right]^2 \sum_{-\infty}^{\infty} d_g = (2\pi f_{uu}(0))^{-1} = \zeta^{-2} . \tag{A.2}$$

Since $C_{11}(g) \to 1$ for any fixed g, (A.1) and (A.2) yield the required result.

(b) Writing

$$T^{-2}\frac{1}{2M} \sum_{j=-M+1}^{M} \hat{f}_{tt}(\omega_j) f_{uu}^{-1}(\omega_j)$$

$$= T^{-2}\frac{1}{2\pi} \sum_{g=-\infty}^{\infty} d_g \frac{1}{2M} \sum_{j=-M+1}^{M} e^{ig\pi j/M} \hat{f}_{tt}(\pi j/M)$$

$$= T^{-2}\left[\frac{1}{2\pi}\right]^2 \sum_{g=-\infty}^{\infty} d_g C_{tt}(g)k(g/M),$$

and noting that $T^{-2}C_{tt}(g) = T^{-3}\sum_{t=1}^{T-g} t(t + g) \to \frac{1}{3}$, the result follows from (A.1) and (A.2).

(c) We have

$$T^{-1}\frac{1}{2M} \sum_{j=-M+1}^{M} \hat{f}_{t1}(\omega_j) f_{uu}^{-1}(\omega_j)$$

$$= T^{-1}\frac{1}{2\pi} \sum_{g=-\infty}^{\infty} d_g \frac{1}{2M} \sum_{j=-M+1}^{M} e^{ig\pi j/M} \hat{f}_{t1}(\pi j/M)$$

$$= T^{-1}\left[\frac{1}{2\pi}\right]^2 \sum_{g=-\infty}^{\infty} d_g C_{t1}(g)k(g/M) .$$

The result follows, because $T^{-1}C_{t1}(g) \to \frac{1}{2}$.

(d) We have

$$T^{-1/2} \frac{1}{2M} \sum_{j=-M+1}^{M} \hat{f}_{1y_L}(\omega_j) f_{uu}^{-1}(\omega_j)$$

$$= T^{-1/2} \frac{1}{2\pi} \sum_{g=-\infty}^{\infty} d_g \frac{1}{2M} \sum_{j=M+1}^{M} e^{ig\pi j/M} \hat{f}_{1y_L}(\pi j/M)$$

$$= T^{-1/2} \left[\frac{1}{2\pi} \right]^2 \sum_{g=-\infty}^{\infty} d_g C_{1y_L}(g) k(g/M),$$

and $T^{-1/2} C_{1y_L}(g) = T^{-3/2} \sum y_{t-1+g} \Rightarrow \zeta \int_0^1 W$. Hence the result follows.

(e) Simply note that $T^{-3/2} C_{ty_L}(g) = T^{-5/2} \sum t y_{t-1+g} \Rightarrow \zeta \int_0^1 r W$.

(f) This follows from $T^{-1} C_{y_L y_L}(g) \Rightarrow \zeta^2 \int_0^1 W^2$.

(g) Note that $T^{1/2} C_{1u}(g) \Rightarrow \zeta W(1)$.

(h) The result follows from $T^{-1/2} C_{tu}(g) \Rightarrow \zeta \int_0^1 r dW$.

(i) Write

$$\frac{1}{2M} \sum_{j=-M+1}^{M} \hat{f}_{y_L u}(\omega_j) f_{uu}^{-1}(\omega_j) = \frac{1}{2\pi} \sum_{g=-\infty}^{\infty} \left[\frac{1}{2M} \sum_{j=-M+1}^{M} \hat{f}_{y_L u}(\omega_j) e^{ig\omega_j} \right] d_g$$

$$= \left[\frac{1}{2\pi} \right]^2 \frac{1}{T} \sum_{g=-\infty}^{\infty} C_{y_L u}(g) k(g/M) d_g .$$

Now we have $C_{y_L u}(n) = T^{-1} \sum_{t=1}^{T-n} y_{t-1} u_{t+n} \Rightarrow \zeta^2 \int_0^1 W dW + \Delta(n+1)$, where $\Delta(n+1) = \sum_{j=0}^{\infty} E u_0 u_{j+n+1}$. Defining $u_j = \sum_{g=-\infty}^{\infty} u_{g+1+j} d_g$, we have

$$\left[\frac{1}{2\pi} \right]^2 \sum_{g=-\infty}^{\infty} \Delta(g+1) d_g = \left[\frac{1}{2\pi} \right]^2 \sum_{g=-\infty}^{\infty} \left[\sum_{j=0}^{\infty} E u_0 u_{g+1+j} \right] d_g$$

$$= \left[\frac{1}{2\pi} \right]^2 \sum_{g=-\infty}^{\infty} E u_0 u_j . \tag{A.3}$$

But, using the inverse transform, we have the representation $Eu_0 u_j = \int_{-\pi}^{\pi} e^{ij\lambda} f_{uu}(\lambda) d\lambda$ for all j, and

$$f_{uu}(\lambda) = \sum_{g=-\infty}^{\infty} f_{uu}(\lambda) e^{i(g+1)\lambda} d_g = f_{uu}(\lambda) e^{i\lambda} 2\pi f_{uu}(\lambda)^{-1} = 2\pi e^{i\lambda} .$$

Thus,

$$E(u_0 u_j) = 2\pi \int_{-\pi}^{\pi} e^{i(j+1)\lambda} d\lambda$$

$$= \begin{cases} (2\pi)^2 , & j = -1 , \\ 0, & \text{otherwise.} \end{cases} \tag{A.4}$$

Using (A.3) and (A.4), we deduce that

$$\left[\frac{1}{2\pi} \right]^2 \sum_{g=-\infty}^{\infty} \Delta(g+1) d_g = 0 . \tag{A.5}$$

Also,

$$\left[\frac{1}{2\pi} \right]^2 \sum_{g=-\infty}^{\infty} \zeta^2 \left(\int_0^1 W dW \right) d_g = \int_0^1 W dW . \tag{A.6}$$

Now the result follows using (A.1), (A.5), and (A.6).

Proof of Theorem 1

For the estimators $\tilde{\alpha}_{DT}$, α_{DT}^* and the corresponding t-ratios, the stated results follow using the weak convergence results

$$T^{-2} \sum_{t=1}^{T-n+1} \tilde{y}_{t-1} \tilde{y}_{t-1+n} \Rightarrow \zeta^2 \int_0^1 \tilde{W}^2 ,$$

$$T^{-1} \sum_{t=1}^{T-n} \tilde{y}_{t-1} u_{t+n} \Rightarrow \zeta^2 \int_0^1 \tilde{W} dW + \Delta(n+1) ,$$

$$T^{-2} \sum_{t=1}^{T-n+1} y_{t-1}^* y_{t-1+n}^* \Rightarrow \zeta^2 \int_0^1 W^{*2} ,$$

$$T^{-1} \sum_{t=1}^{T-n} y_{t-1}^* u_{t+n} \Rightarrow \zeta^2 \int_0^1 W dW + \Delta(n+1) ,$$

which we obtain working from Phillips (1988b, p. 85), and changing parts (f) and (i) of Lemma A.1 in an analogous way. The rest of the results are proven in the following:

(a) (i) As in Phillips (1988a) we have $\max_{\omega_j} |\hat{f}_{uu}(\omega_j) - f_{uu}(\omega_j)| \underset{p}{\to} 0$. Hence, we may replace $\hat{f}_{uu}(\omega_j)$ with $f_{uu}(\omega_j)$ for asymptotic analysis in what follows. Write

$$T(\hat{\alpha} - 1) = Y_{y_L u}/(T^{-1}X_{1y_L}),$$

where

$$Y_{y_L u} = \frac{1}{2M} \sum_{j=-M+1}^{M} \hat{f}_{y_L u}(\omega_j)\hat{f}_{uu}(\omega_j)^{-1}.$$

The result follows from Lemma A.1.

(ii) This follows from part (a)(i) immediately.

(b) (i) Simply note that

$$T(\tilde{\alpha} - 1) = (X_{11}Y_{y_L u} - X_{1y_L}Y_{1u})/\{T^{-1}(X_{11}X_{y_L y_L} - X_{1y_L}^2)\},$$

and apply Lemma A.1. Here

$$Y_{1u} = \frac{1}{2M} \sum_{j=-M+1}^{M} \hat{f}_{1u}(\omega_j)\hat{f}_{uu}(\omega_j)^{-1}.$$

(ii) This follows from part (b)(i).

(c) (i) We have

$$T(\alpha^* - 1) = \{T^{-2}(X_{1t}X_{ty_L} - X_{1y_L}X_{tt})Y_{1u} - T^{-2}(X_{11}X_{ty_L} - X_{1y_L}X_{1t})Y_{tu}$$

$$+ T^{-2}(X_{11}X_{tt} - X_{1t}^2)Y_{y_L u}\}$$

$$/\{T^{-3}(X_{11}X_{tt}X_{y_L y_L} + X_{1t}X_{ty_L}X_{1y_L} + X_{1y_L}X_{1t}X_{ty_L}$$

$$- X_{1y_L}X_{tt}X_{1y_L} - X_{ty_L}^2X_{11} - X_{y_L y_L}X_{1t}^2)\},$$

where

$$Y_{tu} = \frac{1}{2M} \sum_{j=-M+1}^{M} \hat{f}_{tu}(\omega_j)\hat{f}_{uu}(\omega_j)^{-1}.$$

The result is an easy consequence of Lemma A.1.

(ii) This follows straightforwardly from part (c)(i).

References

Bartlett, M.S., 1966, Stochastic processes (Cambridge University Press, Cambridge).

Billingsley, P., 1968, Convergence of probability measures (Wiley, New York, NY).

Billingsley, P., 1986, Probability and measure (Wiley, New York, NY).

Chan, N.H. and C.Z. Wei, 1988, Limiting distributions of least squares estimates of unstable autoregressive processes, Annals of Statistics 16, 367–401.

DeJong, D.N., J.C. Nankervis, N.E. Savin, and C.H. Whiteman, 1989, Unit root tests or coin tosses for time series with autoregressive errors?, Working paper no. 89–14 (University of Iowa, Iowa City, IA).

Dickey, D.A. and W.A. Fuller, 1979, Distributions of estimators for autoregressive time series with a unit root, Journal of the American Statistical Association 74, 427–431.

Dickey, D.A. and W.A. Fuller, 1981, Likelihood ratio statistics for autoregressive time series with a unit root, Econometrica 49, 1057–1072.

Dickey, D.A., W.R. Bell, and R.B. Miller, 1986, Unit roots in time series models: Tests and implications, American Statistician 40, 12–26.

Engle, R.F., 1980, Exact maximum likelihood methods for dynamic regressions and band spectrum regressions, International Economic Review 21, 391–407.

Engle, R.F. and R. Gardner, 1980, Some finite sample properties of spectral estimators of a linear regression, Econometrica 44, 149–165.

Espasa, A. and J.D. Sargan, 1977, The spectral estimation of simultaneous equation systems with lagged endogenous variables, International Economic Review 18, 583–605.

Fuller, W.A., 1976, Introduction to statistical time series (Wiley, New York, NY).

Fuller, W.A., 1984, Nonstationary autoregressive time series, in: Handbook of statistics, Vol. 5 (North-Holland, Amsterdam) 1–23.

Granger, C.W.J. and R.F. Engle, 1983, Application of spectral analysis in econometrics, in: Handbook of statistics, Vol. 3 (North-Holland, Amsterdam) 93–109.

Grenander, U. and M. Rosenblatt, 1957, Statistical analysis of stationary time series (Wiley, New York, NY).

Hall, A., 1989, Testing for a unit root in the presence of moving average errors, Biometrika 76, 49–56.

Hall, P. and C.C. Heyde, 1980, Martingale limit theory and its application (Academic Press, New York, NY).

Hannan, E.J., 1963, Regression for time series, in: M. Rosenblatt, ed., Time series analysis (Wiley, New York, NY).

Hannan, E.J., 1970, Multiple time series (Wiley, New York, NY).

Hannan, E.J., 1973, The asymptotic theory of linear time series models, Journal of Applied Probability 10, 130–145.

IMSL, 1989, Stat/Library (IMSL, Houston, TX).

Park, J.Y. and P.C.B. Phillips, 1989, Statistical inference in regressions with integrated processes: Part 1, Econometric Theory 4, 468–497.

Phillips, P.C.B., 1987, Time series regression with a unit root, Econometrica 55, 277–301.

Phillips, P.C.B., 1988, Multiple regression with integrated time series, in: N.U. Prabhu, ed., Statistical inference from stochastic processes, Contemporary Mathematics 80, 79–106.

Phillips, P.C.B., 1991, Spectral regression for cointegrated time series, in: W. Barnett, ed., Nonparametric and semiparametric methods in economics and statistics (Cambridge University Press, Cambridge).

Phillips, P.C.B. and S. Ouliaris, 1990, Asymptotic properties of residual based tests for cointegration, Econometrica 58, 165–193.

Phillips, P.C.B. and P. Perron, 1988, Testing for a unit root in time series regression, Biometrika 74, 535–547.

Priestley, M.B., 1981, Spectral analysis and time series, Vol. 2 (Academic Press, New York, NY).

Said, E.S. and D.A. Dickey, 1984, Testing for unit roots in autoregressive–moving average models of unknown order, Biometrika 71, 599–607.

Said, E.S. and D.A. Dickey, 1985, Hypothesis testing in ARIMA $(p, 1, q)$ models, Journal of the American Statistical Association 80, 369–374.

Schwert, W.G., 1987, Tests for unit roots: A Monte Carlo investigation, Journal of Business and Economic Statistics 7, 147–160.