

ON PHILLIPS–PERRON-TYPE TESTS FOR SEASONAL UNIT ROOTS

JÖRG BREITUNG
Humboldt University

PHILIP HANS FRANSES
Erasmus University

In this paper we consider a semiparametric version of the test for seasonal unit roots suggested by Hylleberg, Engle, Granger, and Yoo (1990, *Journal of Econometrics* 44, 215–238). The asymptotic theory is based on the analysis of a simple regression problem, and the results apply to tests at any given frequency in the range $(0, \pi]$. Monte Carlo simulations suggest that the test may have more power than the parametric test of Hylleberg et al. (1990). On the other hand, the semiparametric version suffers from severe size distortions in some situations.

1. INTRODUCTION

Seasonally observed economic time series often display persistent changes in seasonal fluctuations. A class of models that is able to describe such patterns is the autoregressive time series process with seasonal unit roots. In a frequency domain such processes exhibit spectral poles at seasonal frequencies. Given the current focus on seasonal cointegration (e.g., Lee, 1992; Engle, Granger, Hylleberg, and Lee, 1993) and model based seasonal adjustment procedures (see Sections 12–14 of Hylleberg, 1992; Breitung, 1994), it is important that one can rely on adequate test statistics for such roots in univariate processes. Hylleberg, Engle, Granger, and Yoo (1990) (henceforth HEGY) and Beaulieu and Miron (1993) employed an autoregressive framework to account for serial correlations of the errors.

Because the empirical performance of the tests critically depends on an appropriate lag augmentation of the auxiliary model (cf. Ghysels, Lee, and Noh, 1994; Hylleberg, 1995) it is desirable to consider a semiparametric approach in the spirit of Phillips (1987), Phillips and Perron (1988), and Schmidt and Phillips (1992). An obvious advantage of such an approach is that we do not need to assume that the data are generated by an autoregressive model with known order

The first author thanks the European Union for the financial support of an HCM fellowship at the University of Amsterdam. The second author thanks the Royal Netherlands Academy of Arts and Sciences for its financial support. Moreover, we are indebted to a co-editor and two anonymous referees for their efforts in eliminating the errors and improving the readability of the earlier versions. Of course all remaining errors are our own. Address correspondence to: Jörg Breitung, Institute of Statistics and Econometrics, Humboldt University Berlin, Spandauer Strasse 1, D-10178 Berlin, Germany.

as in, e.g., HEGY. Furthermore, this approach can be shown to be robust against particular forms of structural breaks (Amsler and Lee, 1995).

When testing for integration at a frequency $0 < \omega_\kappa < \pi$, the implied hypothesis is two-dimensional. This means that the test regression involves two regressors for the same frequency. We show that the Wald test for the joint hypothesis is asymptotically equivalent to a simple function of the t -statistics from two bivariate regressions. Thus, a simple nonparametric correction of the joint test is obtained from correcting the corresponding t -statistics. Whereas most of the literature is concerned with tests at quarterly frequencies, our results apply to tests at any given frequency.

There are further advantages of Phillips–Perron-type tests. The limiting distribution of the test statistic does not depend on the frequency in the range $(0, \pi]$. Thus, we need not apply different sets of critical values as in HEGY or Franses (1990). Moreover, tests at different seasonal frequencies are independent and, therefore, it is not necessary to include all nonstationary terms associated with the range of seasonal frequencies. Our simulations suggest that this may lead to a substantial gain in power. However, similarly to the Phillips–Perron type of tests for (nonseasonal) unit roots, our tests may perform poorly if there is a near MA unit root corresponding to the root under test.

This paper is organized as follows. In Section 2, we outline the testing problem and provide details of the relevant null and alternative hypotheses. We show that the test statistics can be based on simple regression models involving only one or two variables. In Section 3, a score-type test statistic is derived. The asymptotic theory for the bivariate regression problem is considered in Section 4, and in Section 5 a Wald-type test for the joint hypothesis is suggested. The application for the case of multiple unit roots is discussed in Section 6. Section 7 presents some results concerning the small sample properties of the new test, and Section 8 concludes. All proofs can be found in the Appendix.

2. THE TEST PROBLEM

Let $\{Z_t\}$ ($t = 0, \pm 1, \pm 2, \dots$) be a seasonal time series with S seasonal periods. It is assumed that the “seasonal differences” admit the (Wold) representation

$$y_t = (1 - L^S)Z_t = \mu + u_t, \quad (1)$$

where μ is an unknown constant and u_t is a random variable with $E(u_t) = 0$. The model implies that the mean of Z_t is of the form

$$E(Z_t) = bt + \sum_{j=1}^S m_{jt},$$

where m_{jt} is a dummy variable with $m_{jt} = a_j$ for $(t - 1) \bmod S + 1 = j$ and $m_{jt} = 0$ otherwise. Similar deterministic are considered, e.g., in HEGY and Beaulieu and Miron (1993). For the random sequence $\{u_t\}$ we make the following assumption.

Assumption 1. The term u_t is generated by a linear process with $u_t = \epsilon_t + c_1\epsilon_{t-1} + c_2\epsilon_{t-2} + \dots = c(L)\epsilon_t$, $\sum_{j=1}^{\infty} j^2 c_j^2 < \infty$, and $\{\epsilon_t\}$ is a white noise sequence with $E(\epsilon_t) = 0$, $E(\epsilon_t^2) = \sigma^2$, and $\sup_t E|\epsilon_t|^{2+\delta} < \infty$ for some $\delta > 0$.

A similar set of assumptions is used, e.g., by Ahn (1993). The additional assumption of square-summability is required for the existence of the Beveridge–Nelson decomposition that is used below (Phillips and Solo, 1992).

With respect to representation (1) it may be the case that the series is “over-differenced” in the sense that the application of the seasonal difference operator $(1 - L^S)$ imposes unit roots on the lag polynomial $c(L)$. The roots of $|1 - z^S| = 0$ are of the form $z_\kappa = e^{i\omega_\kappa}$ with $\omega_\kappa = 2\pi\kappa/S$ ($\kappa = 0, \dots, S - 1$). For even S , which we will assume in the following discussion,¹ the pairs $(\omega_\kappa, \omega_{S-\kappa})$ are associated with complex conjugate roots for $\kappa = 1, \dots, (S/2 - 1)$, whereas the roots corresponding to $\omega_0 = 0$ and $\omega_{S/2} = \pi$ are real. Accordingly, we factorize the seasonal difference filter as

$$\begin{aligned} (1 - L^S) &= (1 - L)(1 + L) \prod_{\kappa=1}^{S/2-1} (1 - e^{i\omega_\kappa L})(1 - e^{-i\omega_\kappa L}) \\ &= (1 - L)(1 + L) \prod_{\kappa=1}^{S/2-1} [1 - 2 \cos(\omega_\kappa)L + L^2] \\ &= \prod_{\kappa=0}^{S/2} \nabla_\kappa(L), \end{aligned}$$

where

$$\nabla_\kappa(L) = \begin{cases} 1 - L & \text{for } \kappa = 0 \\ 1 - 2 \cos(\omega_\kappa)L + L^2 & \text{for } \kappa = 1, 2, \dots, S/2 - 1 \\ 1 + L & \text{for } \kappa = S/2. \end{cases}$$

Under the null hypothesis we assume that Z_t is integrated at frequency ω_κ , i.e., the factor $\nabla_\kappa(L)$ is necessary to obtain a bounded spectrum at ω_κ . More precisely, we will employ the following definition of integration at frequency ω_κ :

DEFINITION 1. A series $\{\xi_t^\kappa\}$ is (first order) integrated at frequency $0 \leq \omega_\kappa < \pi$ if $x_t = \nabla_\kappa(L)\xi_t^\kappa$ possesses a spectral density $f_x(\omega)$ with $0 < f_x(\omega_\kappa) < \infty$.

Similar definitions were suggested by HEGY, Phillips and Solo (1992), and Gregoir and Laroque (1993). To obtain a test of the hypothesis

H_0 : Z_t is integrated at frequency ω_κ

we write

$$\begin{aligned} y_t &= (1 - L^S)Z_t \\ &= \nabla_\kappa(L)\nabla_\kappa(L)^{-1}(1 - L^S)Z_t \\ &= \nabla_\kappa(L)\xi_t^\kappa, \end{aligned} \tag{2}$$

where $\xi_t^\kappa = \nabla_\kappa(L)^{-1}(1 - L^S)Z_t$. Whenever Z_t is integrated at frequency ω_κ , ξ_t^κ is integrated at ω_κ as well. However, Z_t may be integrated at some other seasonal frequencies, whereas ξ_t^κ has a bounded spectrum for $\omega \neq \omega_\kappa$ by Assumption 1. Under the alternative hypothesis it is assumed that Z_t has a bounded spectrum at ω_κ so that the spectral density of $x_t = \nabla_\kappa(L)Z_t$ satisfies $f_x(\omega_\kappa) = 0$.

Following Ahtola and Tiao (1987), a Dickey–Fuller type of test for integration at frequency ω_κ can be obtained by means of a simple least-squares regression. Assuming the deterministic part $D_t^\kappa = E(\xi_t^\kappa)$ to be known, the regression equation for $0 < \omega_\kappa < \pi$ takes the form

$$(y_t - \mu) = \lambda_1(\xi_{t-1}^\kappa - D_{t-1}^\kappa) + \lambda_2(\xi_{t-2}^\kappa - D_{t-2}^\kappa) + e_t, \quad (3)$$

where e_t is a stationary error term obeying Assumption 1. Using (2) the regression (3) can be rewritten as

$$(\xi_t^\kappa - D_t^\kappa) = (\lambda_1 + 2 \cos \omega_\kappa)(\xi_{t-1}^\kappa - D_{t-1}^\kappa) + (\lambda_2 - 1)(\xi_{t-2}^\kappa - D_{t-2}^\kappa) + e_t, \quad (4)$$

and, thus, the parameters λ_1 and λ_2 measure the deviation from the polynomial $\nabla_\kappa(L)$. For $\omega_\kappa = \pi$ the regression is

$$(y_t - \mu) = \lambda_1(\xi_{t-1}^\kappa - D_{t-1}^\kappa) + e_t.$$

The mean function is defined as

$$D_t^\kappa = \begin{cases} \gamma_0 + \gamma_1 \cos t\omega_\kappa + \gamma_2 \sin t\omega_\kappa & \text{for } 0 < \omega_\kappa < \pi \\ \gamma_0 + \gamma_1(-1)^t & \text{for } \omega_\kappa = \pi. \end{cases}$$

This function has the property that $\nabla_\kappa(L)D_t^\kappa$ is a constant.

It follows from Definition 1 that H_0 implies the parameter restrictions

$$H'_0: \lambda_1 = \lambda_2 = 0$$

in (3). An attractive and simple test for this kind of null hypothesis is obtained by applying the score principle.

3. A SCORE-TYPE TEST STATISTIC

Following Schmidt and Phillips (1992) we first consider the score test for the special situation that u_t is a Gaussian white noise process. We confine ourselves to the cases $0 < \omega_\kappa < \pi$. The case $\omega_\kappa = \pi$ can be treated straightforwardly.

First, we need to specify appropriate initial conditions for ξ_t^κ . For convenience we will assume that the process starts at $t = -1$ so that $\xi_t^\kappa - D_t^\kappa = 0$ for $t \leq -2$ and by using (4) we get

$$\begin{aligned} \xi_{-1}^\kappa &= D_{-1}^\kappa + e_{-1} \\ &= \gamma_0 + \gamma_1 \cos \omega_\kappa - \gamma_2 \sin \omega_\kappa + e_{-1}, \\ \xi_0^\kappa &= D_0^\kappa + e_0 + (\lambda_1 + 2 \cos \omega_\kappa)e_{-1} \\ &= \gamma_0 + \gamma_1 + e_0 + (\lambda_1 + 2 \cos \omega_\kappa)e_{-1}, \end{aligned}$$

where we assume $e_{-1} \sim N(0, \sigma_1^2)$, $e_0 \sim N(0, \sigma_0^2)$, and $E(e_{-1}e_0) = 0$.

Apart from a constant the concentrated log-likelihood function for model (3) is

$$\begin{aligned} \mathcal{L}_c(\lambda_1, \lambda_2, \gamma_0, \gamma_1, \gamma_2) &= -\frac{1}{2} \log(\xi_{-1}^\kappa - D_{-1}^\kappa)^2 \\ &\quad - \frac{1}{2} \log[\xi_0^\kappa - D_0^\kappa - (\lambda_1 + 2 \cos \omega_\kappa)(\xi_{-1}^\kappa - D_{-1}^\kappa)]^2 \\ &\quad - \frac{T}{2} \log \left\{ T^{-1} \sum_{t=1}^T [y_t - \mu - \lambda_1(\xi_{t-1}^\kappa - D_{t-1}^\kappa) \right. \\ &\quad \quad \left. - \lambda_2(\xi_{t-2}^\kappa - D_{t-2}^\kappa)]^2 \right\}. \end{aligned}$$

Under the null hypothesis we have $\mu = (2 - 2 \cos \omega_\kappa)\gamma_0$ and the maximum likelihood (ML) estimates of the restricted likelihood function $\mathcal{L}_c(\lambda_1 = 0, \lambda_2 = 0, \gamma_0, \gamma_1, \gamma_2)$ are obtained as

$$\begin{aligned} \hat{\gamma}_0 &= \frac{1}{2 - 2 \cos \omega_\kappa} \bar{y} \\ \hat{\gamma}_1 &= \xi_0^\kappa - \hat{\gamma}_0 \\ \hat{\gamma}_2 &= -\frac{\xi_{-1}^\kappa - \hat{\gamma}_0 - \hat{\gamma}_1 \cos \omega_\kappa}{\sin \omega_\kappa}, \end{aligned}$$

where $\bar{y} = T^{-1} \sum_{t=1}^T y_t$. Let $\mathcal{L}^0 \equiv \mathcal{L}(\lambda_1, \lambda_2, \hat{\gamma}_0, \hat{\gamma}_1, \hat{\gamma}_2)$, where $\mathcal{L}(\cdot)$ denotes the log-likelihood function. It can be shown that

$$\begin{aligned} \left. \frac{\partial \mathcal{L}^0}{\partial \lambda_1} \right|_{H_0} &\propto \sum_{t=1}^T \{ \xi_{t-1}^\kappa - \hat{\gamma}_0 - \hat{\gamma}_1 \cos(t-1)\omega_\kappa - \hat{\gamma}_2 \sin(t-1)\omega_\kappa \} \\ &\quad \times \{ y_t - (2 - 2 \cos \omega_\kappa)\hat{\gamma}_0 \} \\ \left. \frac{\partial \mathcal{L}^0}{\partial \lambda_2} \right|_{H_0} &\propto \sum_{t=1}^T \{ \xi_{t-2}^\kappa - \hat{\gamma}_0 - \hat{\gamma}_1 \cos(t-2)\omega_\kappa - \hat{\gamma}_2 \sin(t-2)\omega_\kappa \} \\ &\quad \times \{ y_t - (2 - 2 \cos \omega_\kappa)\hat{\gamma}_0 \}. \end{aligned}$$

Analogously to Schmidt and Phillips (1992), these expressions can be rewritten as

$$\left. \frac{\partial \mathcal{L}_c(\cdot)}{\partial \lambda_1} \right|_{H_0} \propto \sum_{t=2}^T \tilde{Y}_{t-1}^\kappa \tilde{y}_t, \tag{5}$$

$$\left. \frac{\partial \mathcal{L}_c(\cdot)}{\partial \lambda_2} \right|_{H_0} \propto \sum_{t=3}^T \tilde{Y}_{t-2}^\kappa \tilde{y}_t, \tag{6}$$

where

$$\begin{aligned} \tilde{y}_t &= y_t - \bar{y}, \\ \tilde{Y}_t^\kappa &= 2 \cos(\omega_\kappa)\tilde{Y}_{t-1}^\kappa - \tilde{Y}_{t-2}^\kappa + \tilde{y}_t, \quad \text{for } t = 1, \dots, T, \end{aligned} \tag{7}$$

and $\tilde{Y}_t^\kappa = 0$ for $t \leq 0$.

The expressions in (5) and (6) indicate that the scores are similar to the scores for testing the hypotheses $\alpha_1 = 0$ and $\alpha_2 = 0$ in the auxiliary regression equation

$$\tilde{y}_t = \alpha_1 \tilde{Y}_{t-1}^\kappa + \alpha_2 \tilde{Y}_{t-2}^\kappa + \tilde{u}_t. \tag{8}$$

Hence, the score principle suggests testing the null hypothesis

$$H_0'': \alpha_1 = \alpha_2 = 0$$

by applying ordinary least squares (OLS) to (8). The difference between a test of this hypothesis and a test of H_0' is that in (8) the unknown parameters of the mean function are replaced by estimates. Under H_0'' we have $\tilde{u}_t = y_t - \bar{y}$, and by virtue of Assumption 1, the errors behave like a demeaned stationary series.

The case of testing against integration at frequency zero is considered by Schmidt and Phillips (1992) and Ahn (1993). In the following section we consider the asymptotic theory for the case $0 < \omega_\kappa \leq \pi$.

4. ASYMPTOTIC THEORY FOR THE BIVARIATE REGRESSION

In this section we analyze the asymptotic properties of the least-squares estimator of α_l in the regression

$$\tilde{y}_t = \alpha_l \tilde{Y}_{t-l}^\kappa + \tilde{v}_t, \tag{9}$$

where $l = 1$ or $l = 2$. The null hypothesis H_0'' implies $\alpha_l = 0$, $l = 1, 2$, and $\tilde{v}_t = \tilde{y}_t$ so that under the null hypothesis the least-squares estimator is

$$\hat{\alpha}_l = \frac{\sum_{t=l+1}^T \tilde{Y}_{t-l}^\kappa \tilde{y}_t}{\sum_{t=l+1}^T (\tilde{Y}_{t-l}^\kappa)^2}.$$

As will become apparent in Section 5, the asymptotic theory for such bivariate regressions can be used to derive the null distribution of a test based on the multiple regression (8).

First we consider the asymptotic behavior of the sequence \tilde{Y}_t^κ . For $0 < \omega_\kappa < \pi$, the generating scheme for \tilde{Y}_t^κ is given by

$$\tilde{Y}_t^\kappa = \frac{1}{\sin \omega_\kappa} \sum_{j=1}^t \tilde{y}_j \sin(t - j + 1) \omega_\kappa \tag{10}$$

$$= \sum_{j=0}^{t-1} \psi_j^\kappa \tilde{y}_{t-j} \quad \text{with } \psi_j^\kappa = (\sin \omega_\kappa)^{-1} \sin(j + 1) \omega_\kappa \tag{11}$$

$$= \frac{1}{\sin \omega_\kappa} \{ \tilde{A}_t^\kappa \sin(t + 1) \omega_\kappa - \tilde{B}_t^\kappa \cos(t + 1) \omega_\kappa \}, \tag{12}$$

where $\tilde{A}_t^\kappa = \sum_{j=1}^t \tilde{y}_j \cos j\omega_\kappa$ and $\tilde{B}_t^\kappa = \sum_{j=1}^t \tilde{y}_j \sin j\omega_\kappa$ (cf. Chan and Wei, 1988, p. 385). For $\omega_\kappa = \pi$ we have

$$\begin{aligned} \tilde{Y}_t^{S/2} &= \sum_{j=1}^t \tilde{y}_j \cos(t-j)\pi \\ &= \sum_{j=0}^{t-1} \psi_j^{S/2} \tilde{y}_{t-j}, \quad \psi_j^{S/2} = \cos j\pi \\ &= -\tilde{A}_t^{S/2} \cos(t+1)\pi. \end{aligned}$$

In contrast to the case of testing for a unit root at frequency zero it can be shown that the inclusion of a constant term does not affect the limiting behavior by considering

$$\begin{aligned} \tilde{A}_t^\kappa &= \sum_{j=1}^t (y_j - \mu) \cos j\omega_\kappa - \left(\sum_{j=1}^t \cos j\omega_\kappa \right) (\bar{y} - \mu) \\ &= A_t^\kappa + O_p(T^{-1/2}) \quad \text{for } 0 < \omega_\kappa \leq \pi, \end{aligned} \tag{13}$$

where $A_t^\kappa = \sum_{j=1}^t (y_j - \mu) \cos j\omega_\kappa$. Similarly we get $\tilde{B}_t^\kappa = B_t^\kappa + O_p(T^{-1/2})$, where $B_t^\kappa = \sum_{j=1}^t (y_j - \mu) \sin j\omega_\kappa$. The following lemma follows easily from Chan and Wei (1988, Sect. 3.3).

LEMMA 1. *Let \tilde{y}_t be a demeaned white noise sequence obeying Assumption 1 and \tilde{Y}_t is constructed as in (10). Then, as $T \rightarrow \infty$,*

(i)

$$T^{-1} \sum_{t=2}^T \tilde{Y}_{t-1}^\kappa \tilde{y}_t \Rightarrow \frac{\sigma^2}{2 \sin \omega_\kappa} \left[\int_0^1 W_1(r) dW_2(r) - \int_0^1 W_2(r) dW_1(r) \right],$$

(ii)

$$T^{-1} \sum_{t=3}^T \tilde{Y}_{t-2}^\kappa \tilde{y}_t \Rightarrow \frac{\sigma^2}{2 \sin \omega_\kappa} \left[\int_0^1 W_1(r) dV_2^\kappa(r) - \int_0^1 W_2(r) dV_1^\kappa(r) \right],$$

(iii)

$$T^{-2} \sum_{t=1}^T (\tilde{Y}_t^\kappa)^2 \Rightarrow \frac{\sigma^2}{4 \sin^2 \omega_\kappa} \left[\int_0^1 W_1(r)^2 dr + \int_0^1 W_2(r)^2 dr \right],$$

(iv)

$$T^{-2} \sum_{t=2}^T \tilde{Y}_{t-1}^\kappa \tilde{Y}_t^\kappa \Rightarrow \frac{\sigma^2 \cos \omega_\kappa}{4 \sin^2 \omega_\kappa} \left[\int_0^1 W_1(r)^2 dr + \int_0^1 W_2(r)^2 dr \right],$$

where

$$dV_1^\kappa(r) = \cos \omega_\kappa dW_1(r) + \sin \omega_\kappa dW_2(r), \tag{14}$$

$$dV_2^\kappa(r) = \cos \omega_\kappa dW_2(r) - \sin \omega_\kappa dW_1(r), \tag{15}$$

and $W_1(r), W_2(r)$ are independent standard Brownian motions.

Using these results it is straightforward to derive the limiting distributions of the least-squares estimates of α_1 and α_2 in (9) for the special case that $y_t = (1 - L^S)Z_t$ is a white noise sequence.

To derive the limiting distribution for the more general case of Assumption 1 we state the following lemma, which can be seen as a Beveridge–Nelson type of decomposition for seasonally integrated processes.

LEMMA 2. *Let $u_t = c(L)\epsilon_t$ be a linear process obeying Assumption 1 with spectral density $0 < f_u(\omega) < \infty$ for $0 \leq \omega \leq \pi$. There exists a decomposition $u_t = x_t + v_t$ such that x_t is white noise with $\text{Var}(x_t) = 2\pi f_u(\omega_\kappa)$ and*

$$\begin{aligned} \zeta_t^\kappa &= (\sin \omega_\kappa)^{-1} \sum_{j=1}^t v_j \sin(t - j + 1) \omega_\kappa \quad \text{for } 0 < \omega_\kappa < \pi \\ &= \sum_{j=1}^t v_j \cos(t - j) \pi \quad \text{for } \omega_\kappa = \pi \end{aligned}$$

has a bounded spectral density for $0 \leq \omega \leq \pi$.

It is important to note that x_t may be correlated with v_{t+k} for $k \in \mathbb{Z}$. If the spectral density of u_t has a global minimum at ω_κ , then an orthogonal decomposition is possible. In general such a decomposition is not unique. For our purpose it is sufficient to know that at least one valid decomposition exists. It is not necessary to construct the series x_t and v_t empirically.

Using Lemmas 1 and 2 the limiting distribution of the least-squares estimator of α_1 in (9) can be obtained as follows.

THEOREM 1. *Assume that y_t can be represented as in (1), where u_t obeys Assumption 1. Under the null hypothesis H_0'' we have for $0 < \omega_\kappa < \pi$ and $T \rightarrow \infty$*

$$\begin{aligned} T\tilde{\alpha}_1 &\Rightarrow \frac{2 \sin \omega_\kappa \left[\int W_1^*(r) dW_2^*(r) - \int W_2^*(r) dW_1^*(r) \right] + (\varrho_1^\kappa / \theta^\kappa)}{\int W_1^*(r)^2 dr + \int W_2^*(r)^2 dr}, \\ t_{\tilde{\alpha}_1} &\Rightarrow \frac{\sqrt{2\pi f_u(\omega_\kappa)} \left[\int W_1^*(r) dW_2^*(r) - \int W_2^*(r) dW_1^*(r) \right] + (\varrho_1^\kappa / \sqrt{\theta^\kappa})}{\sigma \sqrt{\int W_1^*(r)^2 dr + \int W_2^*(r)^2 dr}}, \end{aligned}$$

where all integrals run from 0 to 1, $W_1^*(r)$, $W_2^*(r)$ are two independent Brownian motions,

$$\theta^\kappa = \frac{\pi f_u(\omega_\kappa)}{2 \sin^2(\omega_\kappa)},$$

$$\varrho_1^\kappa = \sum_{j=1}^\infty \psi_{j-1}^\kappa E(u_1 u_{1+j}),$$

and ψ_j^κ is defined in (11).

Remark A. Similarly, for $\omega_\kappa = \pi$ it can be shown that

$$T\tilde{\alpha}_1 \Rightarrow \frac{-\int W_1^*(r) dW_1^*(r) + [\varrho_1^\kappa/2\pi f_u(\pi)]}{\int W_1^*(r)^2 dr}$$

$$t_{\tilde{\alpha}_1} \Rightarrow \frac{-\sqrt{2\pi f_u(\pi)} \int W_1^*(r) dW_1^*(r) + (\varrho_1^\kappa/\sqrt{2\pi f_u(\pi)})}{\sigma \sqrt{\int W_1^*(r)^2 dr}}$$

Remark B. It should be noted that the use of Lemma 2 implies that the Brownian motions involved are different from those of Lemma 1. As pointed out by a co-editor, it is possible to construct the Beveridge–Nelson decomposition using the results of Phillips and Solo (1992, Sect. 4). However, the resulting nonstationary sequence has a phase shift so that our initial conditions are violated. Although it is possible to adapt the limit theory for alternative initial conditions, we prefer to use a decomposition that satisfies the required initial conditions. The drawback of our decomposition is, however, that it implies different Brownian motions and, thus, cannot be related directly to the limiting expressions of Lemma 1. Nevertheless, for practical purposes like simulating critical values this does not cause any problems.

Remark C. The limiting distributions for $T\tilde{\alpha}_2$ and $t_{\tilde{\alpha}_2}$ are obtained from replacing $dW_1^*(r)$ and $dW_2^*(r)$ by

$$dV_1^{*\kappa}(r) = \cos \omega_\kappa dW_1^*(r) + \sin \omega_\kappa dW_2^*(r),$$

$$dV_2^{*\kappa}(r) = \cos \omega_\kappa dW_2^*(r) - \sin \omega_\kappa dW_1^*(r).$$

Furthermore, ϱ_1^κ is replaced by $\varrho_2^\kappa = \sum_{j=2}^\infty \psi_{j-2}^\kappa E(u_1 u_{1+j})$.

Remark D. We can construct test statistics asymptotically, not depending on nuisance parameters, as

$$Q(\tilde{\alpha}_l) = \frac{\left(T^{-1} \sum_{t=l+1}^T \tilde{Y}_{t-l}^\kappa \tilde{y}_t \right) - \varrho_l^\kappa}{T^{-2} \sum_{t=l+1}^T (\tilde{Y}_{t-l}^\kappa)^2}, \tag{16}$$

$$Q(t_{\tilde{\alpha}_l}) = \frac{\left(\sum_{t=l+1}^T \tilde{Y}_{t-l}^\kappa \tilde{y}_t \right) - T\varrho_l^\kappa}{\sqrt{2\pi f_u(\omega_\kappa) \sum_{t=l+1}^T (\tilde{Y}_{t-l}^\kappa)^2}} \tag{17}$$

for $l = 1, 2$.

Remark E. In practice the unknown quantities $f_u(\omega_\kappa)$ and ϱ_l^κ have to be replaced by estimates. Following Phillips (1987) the spectral density can be estimated by applying an appropriate window, e.g.,

$$2\pi - \hat{f}_u(\omega_\kappa) = c_0 + 2 \sum_{j=1}^m \cos(j\omega_\kappa) \frac{m-j+1}{m+1} c_j, \tag{18}$$

where $c_j = T^{-1} \sum_{t=j+l+1}^T (\tilde{y}_t - \tilde{\alpha}_t \tilde{Y}_{t-l}^\kappa)(\tilde{y}_{t-j} - \tilde{\alpha}_t \tilde{Y}_{t-l-j}^\kappa)$ and m is the truncation lag. Similarly, the parameter ϱ_l^κ can be estimated using

$$\hat{\varrho}_l^\kappa = \sum_{j=l}^m \frac{m-j+1}{m+1} \psi_{j-l}^\kappa c_j. \tag{19}$$

5. THE JOINT TEST

In this section we consider the joint test of the hypothesis $\alpha_1 = \alpha_2 = 0$ in the auxiliary regression model (8). For the case of uncorrelated errors, the following theorem gives the limiting distributions of the least-squares estimators for α_1 and α_2 and the Wald statistic for the null hypothesis $\alpha_1 = \alpha_2 = 0$.

THEOREM 2. *Let y_t be generated by (1), where u_t satisfies Assumption 1 with $c(L) = 1$. Then, the least-squares estimators of (8) are asymptotically distributed as*

(i)

$$T\hat{\alpha}_1 \Rightarrow -2 \frac{\int W_1(r) d\bar{V}_1^\kappa(r) + \int W_2(r) d\bar{V}_2^\kappa(r)}{\int W_1(r)^2 dr + \int W_2(r)^2 dr},$$

(ii)

$$t_{\hat{\alpha}_1} \Rightarrow - \frac{\int W_1(r) d\bar{V}_2^\kappa(r) + \int W_2(r) d\bar{V}_1^\kappa(r)}{\sqrt{\int W_1(r)^2 dr + \int W_2(r)^2 dr}},$$

(iii)

$$T\hat{\alpha}_2 \Rightarrow -2 \frac{\int W_1(r) dW_1(r) + \int W_2(r) dW_2(r)}{\int W_1(r)^2 dr + \int W_2(r)^2 dr},$$

(iv)

$$t_{\hat{\alpha}_2} \Rightarrow - \frac{\int W_1(r) dW_1(r) + \int W_2(r) dW_2(r)}{\sqrt{\int W_1(r)^2 dr + \int W_2(r)^2 dr}},$$

where

$$d\bar{V}_1^\kappa(r) = -\cos \omega_\kappa dW_1(r) - \sin \omega_\kappa dW_2(r),$$

$$d\bar{V}_2^\kappa(r) = \sin \omega_\kappa dW_1(r) - \cos \omega_\kappa dW_2(r).$$

For the Wald statistic of a joint test of $H_0: \alpha_1 = \alpha_2 = 0$ we have

(v)

$$\Lambda(\hat{\alpha}_1, \hat{\alpha}_2) \Rightarrow \left(\frac{\int W_1(r) dW_1(r) + \int W_2(r) dW_2(r)}{\sqrt{\int W_1(r)^2 dr + \int W_2(r)^2 dr}} \right)^2 + \left(\frac{\int W_1(r) dW_2(r) - \int W_2(r) dW_1(r)}{\sqrt{\int W_1(r)^2 dr + \int W_2(r)^2 dr}} \right)^2.$$

Part (iii) of this theorem was already established by Chan and Wei (1988, Corollary 3.3.8). In the Appendix we give a slightly different proof for completeness. As a result, the limiting distributions of $\hat{\alpha}_2$ and $\Lambda(\hat{\alpha}_1, \hat{\alpha}_2)$ do not depend on the frequency ω_κ , whereas the limiting distribution of $\hat{\alpha}_1$ depends on ω_κ . Similar results, using a different representation of the limiting distributions, were obtained by Ahtola and Tiao (1987). Simulated critical values for the Wald statistic are given in Table 1.

The following lemma considers the asymptotic relationship between the Wald statistic and the t -statistics of the bivariate regression (9). This relationship is valid for the case of correlated errors so that the result can be used to construct a modified statistic based on the analysis of Section 4.

LEMMA 3. *Let y_t be generated as in (1), where the error u_t obeys Assumption 1. For the Wald statistic of $H_0'': \alpha_1 = \alpha_2 = 0$ we have*

$$\Lambda(\hat{\alpha}_1, \hat{\alpha}_2) = t_{\hat{\alpha}_1}^2 - 2 \cos(\omega_\kappa) t_{\hat{\alpha}_1} t_{\hat{\alpha}_2} + t_{\hat{\alpha}_2}^2 + o_p(1).$$

TABLE 1. Critical values for the Wald statistic

Significance level	0.10	0.05	0.01
$T = 50$	4.89	6.44	10.2
$T = 100$	4.86	6.32	9.83
$T = 150$	4.85	6.28	9.70
$T = 200$	4.85	6.26	9.64
$T = 400$	4.84	6.23	9.55
$T = \infty$	4.83	6.20	9.46

Note: The critical values are obtained by fitting a regression line to 30 empirical critical values for $T \in [30, 500]$ using 10,000 Monte Carlo replications each. The regressors are T^{-1} and a constant. Other powers of T turned out to be negligible. For details, see Davidson and McKinnon (1993).

For the special case $\omega_\kappa = \pi/2$, the Wald statistic has the same asymptotic distribution as $(t_{\tilde{\alpha}_1}^2 + t_{\tilde{\alpha}_2}^2)$, which was already noticed by Engle et al. (1993).

Combining this result with the results of Section 4, a modified test statistic can be constructed as

$$\Lambda^* = Q(t_{\tilde{\alpha}_1})^2 - 2 \cos(\omega_\kappa) Q(t_{\tilde{\alpha}_1}) Q(t_{\tilde{\alpha}_2}) + Q(t_{\tilde{\alpha}_2})^2 \tag{20}$$

such that the limiting distribution is the same as in the case of uncorrelated errors.

6. MULTIPLE UNIT ROOTS

So far we have considered tests for the hypothesis that the time series is integrated at a single frequency. In practice, however, it is likely that seasonal time series are integrated at several frequencies. In this section we therefore consider the case of multiple unit roots.

Assume that a series Z_t has two unit roots such that $u_t = \nabla_1(L)\nabla_2(L)$ is stationary with nonzero spectral densities at ω_1 and ω_2 . In other words, the series Z_t is integrated at frequencies ω_1 and ω_2 . It is assumed that Z_t is tested for a unit root at ω_1 with the technique suggested previously. Thus, we construct $\tilde{y}_t = \nabla_1(L)Z_t - \text{mean}[\nabla_1(L)Z_t]$ and \tilde{Y}_t^κ as in (10). Then two bivariate regressions according to (9) are performed to test the hypotheses $\alpha_1 = 0$ and $\alpha_2 = 0$. Because in this case \tilde{y}_t and \tilde{Y}_t^κ are integrated at frequency ω_2 , $T\tilde{\alpha}_1$ and $T\tilde{\alpha}_2$ will diverge as $T \rightarrow \infty$. As a consequence the Wald test suggested in Section 5 will reject with probability one as T approaches infinity, whether Z_t is integrated at frequency ω_1 or not. To perform the test for a single unit root it is therefore important to remove all unit roots at other frequencies.

In many applications, however, the number and frequencies of unit roots are unclear. It may therefore be desirable to perform a joint test including all possible unit roots of the series. A straightforward generalization of the tests suggested in Section 3 is

$$\tilde{y}_t = \alpha_{0,1}\tilde{Y}_{t-1}^0 + \alpha_{S/2,1}\tilde{Y}_{t-1}^{S/2} + \sum_{\kappa=1}^{(S/2)-1} \alpha_{\kappa,1}\tilde{Y}_{t-1}^\kappa + \alpha_{\kappa,2}\tilde{Y}_{t-2}^\kappa + \tilde{v}_t, \tag{21}$$

where

$$\tilde{Y}_t^0 = \sum_{j=1}^t \tilde{y}_j,$$

$$\tilde{Y}_t^{S/2} = \sum_{j=1}^t (-1)^{t-j} \tilde{y}_j.$$

The hypothesis that Z_t is integrated at all frequencies ω_κ , $\kappa = 0, \dots, S/2$ implies $\alpha_{0,1} = \dots = \alpha_{S/2,1} = 0$ and $\alpha_{1,2} = \dots = \alpha_{S/2-1,2} = 0$.

To discuss the properties of a test based on (21), the following lemma is useful.

LEMMA 4. Let \tilde{y}_t be a demeaned stationary process with nonzero spectral density at the seasonal frequencies ω_κ and ω_l , $k \neq l$. Then, for the sequences

$$\begin{aligned} \tilde{Y}_t^k &= (\sin \omega_k)^{-1} \sum_{j=1}^t \tilde{y}_j \sin(t - j + 1)\omega_k && \text{for } 0 < \omega_k < \pi \\ &= \sum_{j=1}^t (-1)^{t-j} \tilde{y}_j && \text{for } \omega_k = \pi \end{aligned}$$

$$\tilde{Y}_t^l = (\sin \omega_l)^{-1} \sum_{j=1}^t \tilde{y}_j \sin(t - j + 1)\omega_l \quad \text{for } 0 < \omega_l < \pi$$

we have

$$\sum_{t=1}^T \tilde{Y}_t^k \tilde{Y}_t^l = O_p(T).$$

From this lemma it follows that if Z_t is integrated at all seasonal frequencies, the regressors in (21) are “asymptotically orthogonal.” Hence, the multivariate estimates of the parameters have the same asymptotic distribution as the bivariate regressions considered in Section 4. In other words, in the bivariate regressions (9) with a particular frequency ω_κ , the coefficients α_l , with $l \neq k$ can be seen as nuisance parameters.²

7. SMALL SAMPLE PROPERTIES

In this section we report some results on the small sample properties of the new test. In order to compare the performance of the Phillips–Perron-type tests with the parametric HEGY tests, we consider three different data-generating processes:

- DGP1: $Z_t = \mu - \rho Z_{t-1} + u_t$,
- DGP2: $Z_t = \mu - \rho Z_{t-2} + u_t$,
- DGP3: $Z_t = \mu + \rho Z_{t-4} + u_t$.

If $\rho = 1$, DGP1 and DGP2 have single seasonal unit roots at frequency $\omega_\kappa = \pi$ and $\omega_\kappa = \pi/2$, respectively, and DGP3 is integrated at frequencies 0, π , and $\pi/2$. To study the size of the tests we therefore set $\rho = 1$, whereas we let $|\rho| < 1$ to study the power. As in Ghysels et al. (1994), the errors are generated by the MA process

$$u_t = \epsilon_t - \theta_1 \epsilon_{t-1} - \theta_4 \epsilon_{t-4},$$

where the values for θ_1 and θ_4 are the same as in Ghysels et al. (1994). The errors ϵ_t are uncorrelated $N(0,1)$ random variables.

To apply the test procedure of HEGY, the MA process for the errors is approximated by an AR process of sufficient order. For the subsequent simulations we consider AR(4) and AR(8) approximations. The corresponding tests are labeled as HEGY(4) and HEGY(8). According to the standard model formulation we include seasonal dummies in all cases and add a linear time trend for DGP3.

For the nonparametric estimates of $f_u(\omega_\kappa)$ and q_i^κ we use the Newey–West weighting scheme with $m = 4$ and $m = 8$. For testing at frequency π we use the statistic $Q(t_{\hat{\alpha}_1})$ defined in (17). It follows from Remark A in Section 4 that $-Q(t_{\hat{\alpha}_1})$ has the same asymptotic distribution as the Dickey–Fuller t -statistic for a test at frequency zero without an intercept. Accordingly, we apply the critical values given in Fuller (1976). For a test at $\omega_\kappa = \pi/2$ we use the modified Wald statistic $\Lambda^*(m)$ given in (20), where the truncation lag is given in parentheses. For all simulations we set $T = 100$, the nominal significance level is 0.05, and 5,000 Monte Carlo replications are used.

Table 2 presents the empirical sizes of the tests at frequencies π and $\pi/2$. Letting $\theta_1 = \theta_4 = 0$ we observe a moderate size bias for both test procedures. Because the critical values for the tests do not account for the lag order, such differences between actual and nominal significance levels may occur (see, e.g., Cheung, 1995).

Setting θ_1 and θ_4 different to zero, we are able to assess the ability of the tests to correct for correlated errors. If $\theta_1 = -1$ the MA polynomial has a unit root corresponding to frequency π , whereas for $\theta_4 = 1$ the MA polynomial has unit roots according to the frequencies 0, $\pi/2$, and π . In these cases the corresponding unit roots of the AR polynomial cancel, and the null hypothesis is violated. Hence, we expect a poor performance of the respective tests whenever θ_1 is close to -1 or θ_4 is close to 1. This is what we observe in Table 2. If there is a “near cancellation” of the relevant root, the Phillips–Perron-type tests perform very poorly. A similar observation is made in the case of testing for a unit root at frequency zero (Schwert, 1989; Pantula, 1991). On the other hand, the parametric test (HEGY) seems to be able to deal with such situations by choosing a sufficient augmentation lag. A theoretic explanation for this finding is given by Perron and Ng (1996).

If $\theta_1 = 0.9$ or $\theta_4 = -0.9$ the roots of the MA polynomial are quite different from the respective roots of the AR polynomial. Therefore, it does not come as a surprise that in these cases the tests perform much better. Although the Phillips–Perron-type tests still have a tendency to overreject the null hypothesis, the bias is moderate and only slightly larger than for the case of uncorrelated errors.

Table 3 presents the power of the tests against alternatives with $\rho < 1$. As a result of the size distortions reported in Table 2, we confine ourselves to the case $\theta_1 = \theta_4 = 0$, for which the size bias is small in all cases. From the simulation results it emerges that the Phillips–Perron-type tests are much more powerful than the HEGY test procedure. This is in particular true for DGP1 and DGP2. There are different reasons that may explain this finding. First, if there is a unit root at a single frequency, it is expected that a test considering only one frequency—such as the Phillips–Perron type of test—is superior. Second, it is known that the power of parametric tests deteriorates rapidly with an increasing augmentation lag, whereas the power of the Phillips–Perron variant is not as sensitive with respect to the truncation lag. Third, estimating the deterministic part under the null hypothesis leads to a more parsimonious specification of the regression. Instead of including S seasonal dummies (and a time trend), only a constant is required in our test.

TABLE 2. Empirical sizes

DGP1, $\omega_\kappa = \pi$					
θ_1	θ_4	HEGY(4)	HEGY(8)	$-Q(t_{\bar{\alpha}_1}), m = 4$	$-Q(t_{\bar{\alpha}_1}), m = 8$
-0.9	0	0.282	0.099	0.956	0.970
0	-0.9	0.104	0.025	0.012	0.022
0	0	0.035	0.034	0.053	0.054
0.9	0	0.040	0.036	0.029	0.025
0	0.9	0.678	0.325	0.793	0.739
DGP2, $\omega_\kappa = \pi/2$					
θ_1	θ_4	HEGY(4)	HEGY(8)	$\Lambda^*(4)$	$\Lambda^*(8)$
-0.9	0	0.048	0.043	0.081	0.091
0	-0.9	0.122	0.047	0.142	0.103
0	0	0.050	0.046	0.065	0.078
0.9	0	0.049	0.042	0.084	0.090
0	0.9	0.915	0.553	0.893	0.843
DGP3, $\omega_\kappa = \pi$					
θ_1	θ_4	HEGY(4)	HEGY(8)	$-Q(t_{\bar{\alpha}_1}), m = 4$	$-Q(t_{\bar{\alpha}_1}), m = 8$
-0.9	0	0.606	0.161	0.961	0.974
0	-0.9	0.091	0.022	0.013	0.023
0	0	0.032	0.030	0.052	0.052
0.9	0	0.032	0.028	0.030	0.026
0	0.9	0.702	0.307	0.806	0.754
DGP3, $\omega_\kappa = \pi/2$					
θ_1	θ_4	HEGY(4)	HEGY(8)	$\Lambda^*(4)$	$\Lambda^*(8)$
-0.9	0	0.051	0.040	0.087	0.093
0	-0.9	0.113	0.037	0.143	0.101
0	0	0.044	0.042	0.064	0.077
0.9	0	0.046	0.038	0.081	0.089
0	0.9	0.935	0.562	0.892	0.843

Note: The entries present the rejection frequencies of 5,000 Monte Carlo replications of the respective model with $\rho = 1$. HEGY(k) refers to the t (for $\omega_\kappa = \pi$) and F (for $\omega_\kappa = \pi/2$) statistics as defined in HEGY, where the regression includes k lagged seasonal differences, a constant, seasonal dummies, and, for DGP3, a linear time trend. The statistic $\Lambda^*(m)$ is defined in (20), where m indicates the truncation lag of the estimates $\hat{\varrho}_1^*$ and $\hat{f}_\kappa(\omega_\kappa)$ defined in (18) and (19).

8. CONCLUDING REMARKS

Following Schmidt and Phillips (1992), we consider score-type tests for integration at seasonal frequencies. Because the resulting test statistics are asymp-

TABLE 3. Empirical power

DGP1, $\omega_\kappa = \pi$				
ρ	HEGY(4)	HEGY(8)	$-Q(t_{\bar{\alpha}_1}), m = 4$	$-Q(t_{\bar{\alpha}_1}), m = 8$
0.95	0.073	0.057	0.300	0.306
0.90	0.135	0.095	0.648	0.652
0.80	0.301	0.173	0.896	0.888
0.60	0.543	0.276	0.949	0.948
DGP2, $\omega_\kappa = \pi/2$				
ρ	HEGY(4)	HEGY(8)	$\Lambda^*(4)$	$\Lambda^*(8)$
0.95	0.081	0.069	0.192	0.211
0.90	0.125	0.094	0.404	0.427
0.80	0.291	0.183	0.745	0.739
0.60	0.667	0.375	0.910	0.893
DGP3, $\omega_\kappa = \pi$				
ρ	HEGY(4)	HEGY(8)	$-Q(t_{\bar{\alpha}_1}), m = 4$	$-Q(t_{\bar{\alpha}_1}), m = 8$
0.95	0.044	0.037	0.083	0.085
0.90	0.049	0.044	0.133	0.134
0.80	0.077	0.062	0.252	0.251
0.60	0.202	0.127	0.520	0.504
DGP3, $\omega_\kappa = \pi/2$				
ρ	HEGY(4)	HEGY(8)	$\Lambda^*(4)$	$\Lambda^*(8)$
0.95	0.061	0.056	0.115	0.133
0.90	0.079	0.069	0.188	0.209
0.80	0.142	0.101	0.357	0.376
0.60	0.396	0.233	0.641	0.624

Note: This table presents the rejection frequencies of 5,000 Monte Carlo replications of the respective model with various values of ρ and $\theta_1 = \theta_4 = 0$. HEGY(k) refers to the t (for $\omega_\kappa = \pi$) and F (for $\omega_\kappa = \pi/2$) statistics as defined in HEGY, where the regression includes k lagged seasonal differences, a constant, seasonal dummies, and, for DGP3, a linear time trend. The statistic $\Lambda^*(m)$ is defined in (20), where m indicates the truncation lag of the estimates $\hat{\rho}^*$ and $\hat{f}_i(\omega_\kappa)$ defined in (18) and (19).

totically equivalent to a function of t -statistics from bivariate regressions, the asymptotic theory is based on the analysis of a simple regression problem. In order to account for serially correlated errors, a nonparametric correction similar to the one suggested by Phillips and Perron (1988) is used.

In a limited Monte Carlo investigation we found that our tests share the well-known merits and deficiencies of unit root tests involving nonparametric correction for short run dynamics. Although they are able to outperform tests using a

parametric correction based on an autoregressive approximation with respect to power, the semiparametric tests may suffer from a tremendous size bias in some (empirically relevant) cases. Hence such tests cannot be recommended for general use. It might therefore be attractive to consider modifications as considered by Perron and Ng (1996). Whenever the parametric tests require a high augmentation lag, which is often the case when using monthly data, the semiparametric variants are an attractive alternative because the power of the test is more robust to the choice of the truncation lag.

NOTES

1. As pointed out by a referee, it is possible to extend the analysis to odd values of S to cover, e.g., day-of-the-week effects. However, for the ease of exposition we follow the previous literature and focus on even S .

2. We are indebted to an anonymous referee, who suggested this interpretation to us.

REFERENCES

- Ahn, S.K. (1993) Some tests for unit roots in autoregressive-integrated-moving average models with deterministic trends. *Biometrika* 80, 855–868.
- Ahtola, J. & G.C. Tiao (1987) Distributions of least squares estimators of autoregressive parameters for a process with complex roots on the unit circle. *Journal of Time Series Analysis* 8, 1–14.
- Amsler, C. & J. Lee (1995) An LM test for a unit root in the presence of a structural break. *Econometric Theory* 11, 359–368.
- Beaulieu, J.J. & J.A. Miron (1993) Seasonal unit roots in aggregate U.S. data. *Journal of Econometrics* 55, 305–328.
- Breitung, J. (1994) A model based seasonal adjustment method using the Beveridge-Nelson decomposition. *Allgemeines Statistisches Archiv* 78, 365–385.
- Chan, N.H. & C.Z. Wei (1988) Limiting distributions of least squares estimates of unstable autoregressive processes, *Annals of Statistics* 16, 367–401.
- Cheung, K.S. (1995) Lag order and critical values of the augmented Dickey-Fuller test. *Journal of Business and Economic Statistics* 13, 277–280.
- Davidson, R. & J.G. MacKinnon (1993) Estimation and inference in econometrics. New York: Oxford University Press.
- Engle, R.F., C.W.J. Granger, S. Hylleberg, & H.S. Lee (1993) Seasonal cointegration: The Japanese consumption function. *Journal of Econometrics* 55, 275–299.
- Franses, P.H. (1990) Testing for Seasonal Unit Roots in Monthly Data. Tinbergen Institute, Report 9032/A, Erasmus University, Rotterdam.
- Fuller, W.A. (1976) *Introduction to Statistical Time Series*. New York: Wiley.
- Ghysels, E., H.S. Lee, & J. Noh (1994) Testing for unit roots in seasonal time series. *Journal of Econometrics* 62, 415–442.
- Gregoir, S. & G. Laroque (1993) Multivariate time series: A polynomial error correction representation theorem. *Econometric Theory* 9, 329–342.
- Hylleberg, S. (1992) *Modelling Seasonality*. Oxford: Oxford University Press.
- Hylleberg, S. (1995) Tests for seasonal unit roots: General to specific or specific to general. *Journal of Econometrics* 69, 5–25.
- Hylleberg, S., R.F. Engle, C.W.J. Granger, & B.S. Yoo (1990) Seasonal integration and co-integration. *Journal of Econometrics* 44, 215–238.
- Lee, H.S. (1992) Maximum likelihood inference on cointegration and seasonal cointegration. *Journal of Econometrics* 54, 1–47.

Pantula, S.G. (1991) Asymptotic distributions of unit-root tests when the process is nearly stationary. *Journal of Business and Economic Statistics* 9, 63–71.

Perron, P. & S. Ng (1996) Useful modifications to some unit root tests with dependent errors and their local asymptotic properties. *Review of Economic Studies* 63, 435–463.

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

Phillips, P.C.B. & P. Perron (1988) Testing for a unit root in time series regression. *Biometrika* 75, 335–346.

Phillips, P.C.B. & V. Solo (1992) Asymptotics for Linear Processes. *Annals of Statistics* 20, 971–1001.

Schmidt, P. & P.C.B. Phillips (1992) LM test for a unit root in the presence of deterministic trends. *Oxford Bulletin of Economics and Statistics* 54, 257–287.

Schwert, G.W. (1989) Tests for unit roots: A Monte Carlo investigation. *Journal of Business and Economic Statistics* 7, 147–159.

APPENDIX

Proof of Lemma 2. Consider the spectral representation of u_t :

$$\begin{aligned} u_t &= \int_{-\pi}^{\pi} e^{it\omega} dG_u(\omega) \\ &= \int_{-\pi}^{\pi} \lambda(\omega) e^{it\omega} dG_u(\omega) + \int_{-\pi}^{\pi} [1 - \lambda(\omega)] e^{it\omega} dG_u(\omega) \\ &= \int_{-\pi}^{\pi} e^{it\omega} dG_x(\omega) + \int_{-\pi}^{\pi} e^{it\omega} dG_v(\omega), \end{aligned}$$

where $dG_x(\omega) = \lambda(\omega) dG_u(\omega)$, $dG_v(\omega) = [1 - \lambda(\omega)] dG_u(\omega)$, and

$$\lambda(\omega) = \sqrt{\frac{f_u(\omega_\kappa)}{f_u(\omega)}}.$$

Hence, we have $E[|dG_x(\omega)|^2] = f_u(\omega_\kappa) d\omega$ for $0 \leq \omega \leq \pi$. Accordingly, components with the desired properties can be found by letting $x_t = \int_{-\pi}^{\pi} e^{it\omega} dG_x(\omega)$ and $v_t = \int_{-\pi}^{\pi} e^{it\omega} dG_v(\omega)$. The variance of x_t is obtained as

$$E(x_t^2) = \int_{-\pi}^{\pi} f_u(\omega_\kappa) d\omega = 2\pi f_u(\omega_\kappa),$$

which completes the proof. ■

Proof of Theorem 1. Let $\tilde{y}_t = \tilde{x}_t + \tilde{v}_t$, where \tilde{x}_t and \tilde{v}_t are demeaned analogs of x_t and v_t , having the properties as stated in Lemma 2, and $\tilde{X}_t^\kappa, \tilde{\zeta}_t^\kappa$ are defined as in (10), where \tilde{y}_t is replaced by \tilde{x}_t and \tilde{v}_t , respectively. Then, $\tilde{Y}_t^\kappa = \tilde{X}_t^\kappa + \tilde{\zeta}_t^\kappa$. From Lemma 2 it follows that $\tilde{\zeta}_t^\kappa$ is $O_p(1)$ and from Lemma 1(iii) we deduce

$$\begin{aligned} \frac{1}{T^2} \sum_{t=2}^T (\tilde{Y}_{t-1}^\kappa)^2 &= \frac{1}{T^2} \sum_{t=2}^T (\tilde{X}_{t-1}^\kappa)^2 + o_p(1) \\ &\Rightarrow \frac{2\pi f_u(\omega_\kappa)}{4 \sin^2 \omega_\kappa} \left[\int W_1^*(r)^2 dr + \int W_2^*(r)^2 dr \right], \end{aligned}$$

where $W_1^*(r)$ and $W_2^*(r)$ represent two independent Brownian motions.

Next, we give a useful lemma.

LEMMA 5. *Let $a(L)$ be a lag polynomial of order $m < \infty$ and x_t and y_t are two mean zero sequences such that*

$$\lim_{T \rightarrow \infty} \sup_{i,j \in \{1, \dots, T\}} \{T^{-2\delta} E(x_i y_j)^2\} = 0$$

and

$$T^{-\delta} \sum_{t=m+1}^T [a(L)x_t]y_t$$

converges in probability to a constant for some $\delta > 0$. Then the sequence

$$T^{-\delta} \sum_{t=1}^{T-m} x_t [a(L^{-1})y_t]$$

converges in probability to the same limit.

Proof of Lemma 5. The lemma follows immediately from

$$\begin{aligned} T^{-\delta} \left(\sum_{t=m+1}^T a_0 x_t y_t + a_1 x_{t-1} y_t + \dots + a_m x_{t-m} y_t \right) \\ = T^{-\delta} \left(\sum_{t=1}^{T-m} a_0 x_t y_t + a_1 x_t y_{t+1} + \dots + a_m x_t y_{t+m} \right) \\ + \sum_{i=0}^{m-1} \sum_{j=0}^{m-i-1} a_i x_{T-j-i} y_{T-j} - \sum_{i=0}^{m-1} \sum_{j=1}^{m-i} a_i x_j y_{j+i} \\ = T^{-\delta} \sum_{t=1}^{T-m} x_t [a(L^{-1})y_t] + o_p(1). \quad \blacksquare \end{aligned}$$

Using Lemma 5 and $L^2 \nabla_\kappa(L^{-1}) = \nabla_\kappa(L)$ for frequencies $0 < \omega_\kappa < \pi$ and $L \nabla_\kappa(L^{-1}) = \nabla_\kappa(L)$ for $\omega_\kappa = \pi$ we have for $0 < \omega_\kappa < \pi$

$$\begin{aligned} \sum_{t=2}^T \tilde{Y}_{t-1}^\kappa \tilde{y}_t &= \sum_{t=2}^T (\tilde{X}_{t-1}^\kappa + \tilde{\zeta}_{t-1}^\kappa)(\tilde{x}_t + \tilde{v}_t) \\ &= \sum_{t=2}^T (\tilde{X}_{t-1}^\kappa \tilde{x}_t + \tilde{\zeta}_{t-1}^\kappa \tilde{x}_t + \tilde{\zeta}_{t-1}^\kappa \tilde{v}_t) + \sum_{t=1}^{T-1} \tilde{x}_{t+1} \tilde{\zeta}_t^\kappa + O_p(1). \end{aligned}$$

From Lemma 1(i) it follows that

$$T^{-1} \sum_{t=2}^T \tilde{X}_{t-1}^\kappa \tilde{x}_t \Rightarrow \frac{\sigma^2}{2 \sin \omega_\kappa} \left[\int W_1^*(r) dW_2^*(r) - \int W_2^*(r) dW_1^*(r) \right].$$

Because $E(\tilde{X}_{t-1}^\kappa \tilde{x}_t) = 0$ and $T^{-1} \sum_{t=2}^T (\tilde{Y}_{t-1}^\kappa \tilde{y}_t - \tilde{X}_{t-1}^\kappa \tilde{x}_t)$ converges in probability to $E[T^{-1} \sum_{t=2}^T (\tilde{Y}_{t-1}^\kappa \tilde{y}_t - \tilde{X}_{t-1}^\kappa \tilde{x}_t)] = E[T^{-1} \sum_{t=2}^T \tilde{Y}_{t-1}^\kappa \tilde{y}_t]$ we get

$$T^{-1} \sum_{t=2}^T \tilde{Y}_{t-1}^\kappa \tilde{y}_t \Rightarrow \frac{\sigma^2}{2 \sin \omega_\kappa} \left[\int W_1^*(r) dW_2^*(r) - \int W_2^*(r) dW_1^*(r) \right] \\ + E \left[T^{-1} \sum_{t=2}^T \tilde{Y}_{t-1}^\kappa \tilde{y}_t \right].$$

Finally,

$$E \left[T^{-1} \sum_{t=2}^T \tilde{Y}_{t-1}^\kappa \tilde{y}_t \right] = \lim_{T \rightarrow \infty} E(\tilde{Y}_{T-1}^\kappa \tilde{y}_T) \\ = \lim_{T \rightarrow \infty} E \left[\left(\sum_{j=1}^{\infty} \psi_{j-1}^\kappa \tilde{y}_{T-j} \right) \tilde{y}_T \right] \\ = \sum_{j=1}^{\infty} \psi_{j-1}^\kappa E(u_T u_{T-j}) \\ = \varrho_1^\kappa.$$

With these results the limiting distributions for $\tilde{\alpha}_1$ and $t_{\tilde{\alpha}_1}$ follow easily. \blacksquare

Proof of Theorem 2.

(i) From the results of Chan and Wei (1988, p. 386) and Lemma 1(iv) it follows that

$$T^{-2} \begin{bmatrix} \sum (\tilde{Y}_{t-1}^\kappa)^2 & \sum \tilde{Y}_{t-1}^\kappa \tilde{Y}_{t-2}^\kappa \\ \sum \tilde{Y}_{t-1}^\kappa \tilde{Y}_{t-2}^\kappa & \sum (\tilde{Y}_{t-2}^\kappa)^2 \end{bmatrix} \\ = T^{-2} \sum_{t=1}^T (\tilde{Y}_t^\kappa)^2 \begin{bmatrix} 1 & \cos \omega_\kappa \\ \cos \omega_\kappa & 1 \end{bmatrix} + o_p(1) \\ \Rightarrow \frac{\sigma^2}{4 \sin^2 \omega_\kappa} \left[\int W_1(r)^2 dr + \int W_2(r)^2 dr \right] \begin{bmatrix} 1 & \cos \omega_\kappa \\ \cos \omega_\kappa & 1 \end{bmatrix} \quad (\text{A.1})$$

and, thus,

$$\begin{bmatrix} \hat{\alpha}_1 \\ \hat{\alpha}_2 \end{bmatrix} = \left\{ \sin^2 \omega_\kappa \sum (\tilde{Y}_{t-1}^\kappa)^2 \right\}^{-1} \begin{bmatrix} \sum \tilde{Y}_{t-1}^\kappa \tilde{y}_t - \cos \omega_\kappa \sum \tilde{Y}_{t-2}^\kappa \tilde{y}_t \\ \sum \tilde{Y}_{t-2}^\kappa \tilde{y}_t - \cos \omega_\kappa \sum \tilde{Y}_{t-1}^\kappa \tilde{y}_t \end{bmatrix} + o_p(T^{-1}). \quad (\text{A.2})$$

Using Lemma 1 we get

$$T^{-1} \left[\sum \tilde{Y}_{t-2}^\kappa \tilde{y}_t - \cos \omega_\kappa \sum \tilde{Y}_{t-1}^\kappa \tilde{y}_t \right] \\ \Rightarrow -(\sigma^2/2) \left[\int W_1(r) dW_1(r) + \int W_2(r) dW_2(r) \right].$$

From (A.2) and Lemma 1 (iii) we obtain

$$T\hat{\alpha}_2 \Rightarrow -2 \frac{\int W_1(r) dW_1(r) + \int W_2(r) dW_2(r)}{\int W_1(r)^2 dr + \int W_2(r)^2 dr}.$$

- (ii) The limiting distribution for $t_{\hat{\alpha}_1}$ follows immediately from (i) and Lemma 1.
- (iii) Using (A.2) and Lemma 1 yields

$$\begin{aligned} T^{-1} & \left[\sum \tilde{Y}_{t-1}^\kappa \tilde{y}_t - \cos \omega_\kappa \sum \tilde{Y}_{t-2}^\kappa \tilde{y}_t \right] \\ & \Rightarrow \frac{\sigma^2}{2 \sin \omega_\kappa} \left[\int W_1(dW_2 - \cos \omega_\kappa dV_2) + \int W_2(-dW_1 + \cos \omega_\kappa dV_1) \right] \\ & = \frac{\sigma^2}{2} \left[\int W_1(\sin \omega_\kappa dW_2 + \cos \omega_\kappa dW_1) \right. \\ & \quad \left. + \int W_2(-\sin \omega_\kappa dW_1 + \cos \omega_\kappa dW_2) \right] \\ & = -(\sigma^2/2) \left[\int W_1 d\bar{V}_1^\kappa + \int W_2 d\bar{V}_2^\kappa \right], \end{aligned}$$

where, for notational convenience, we omit the argument (r) in the integrals.

- (iv) The limiting distribution for $t_{\tilde{\alpha}_1}$ follows immediately from (iii) and Lemma 1.
- (v) It is easy to verify that

$$\begin{bmatrix} 1 & \cos \omega_\kappa \\ \cos \omega_\kappa & 1 \end{bmatrix} = RR'$$

with

$$R = \begin{bmatrix} 1 & 0 \\ \cos \omega_\kappa & \sin \omega_\kappa \end{bmatrix}$$

so that using (A.1) we have

$$\Lambda(\hat{\alpha}_1, \hat{\alpha}_2) = \sigma^{-2} [\sin^2 \omega_\kappa \hat{\alpha}_2^2 + (\hat{\alpha}_1 + \hat{\alpha}_2 \cos \omega_\kappa)^2] \sum_{i=2}^T (\tilde{Y}_{t-1}^\kappa)^2 + o_p(1). \quad (\mathbf{A.3})$$

Using (A.2) we get

$$\hat{\alpha}_1 + \cos \omega_\kappa \hat{\alpha}_2 = \tilde{\alpha}_1 + o_p(T^{-1}).$$

From Theorem 1 we obtain the distribution of $\tilde{\alpha}_1$ as a special case with $W_1^*(r) = W_1(r)$, $W_2^*(r) = W_2(r)$, $2\pi f_u(\omega_\kappa) = \sigma^2$, and $\varrho_1^\kappa = 0$. Hence,

$$\tilde{\alpha}_1 \Rightarrow \frac{2 \sin \omega_\kappa \left[\int W_1(r) dW_2(r) - \int W_2(r) dW_1(r) \right]}{\int W_1(r)^2 dr + \int W_2(r)^2 dr}.$$

Finally, using Theorem 2(iii) the limiting distribution of $\Lambda(\hat{\alpha}_1, \hat{\alpha}_2)$ can easily be derived. ■

Proof of Lemma 3. From (A.2) we get

$$\hat{\alpha}_1 \sqrt{\sum_{t=2}^T (\tilde{Y}_{t-1}^\kappa)^2} = \frac{\sigma}{\sin \omega_\kappa} [t_{\hat{\alpha}_1} - \cos(\omega_\kappa)t_{\hat{\alpha}_2}] + o_p(1),$$

$$\hat{\alpha}_2 \sqrt{\sum_{t=2}^T (\tilde{Y}_{t-1}^\kappa)^2} = \frac{\sigma}{\sin \omega_\kappa} [t_{\hat{\alpha}_2} - \cos(\omega_\kappa)t_{\hat{\alpha}_1}] + o_p(1).$$

Inserting these expressions in (A.3) gives, after some simple manipulations,

$$\Lambda(\hat{\alpha}_1, \hat{\alpha}_2) = t_{\hat{\alpha}_1}^2 - 2 \cos(\omega_\kappa)t_{\hat{\alpha}_1}t_{\hat{\alpha}_2} + t_{\hat{\alpha}_2}^2 + o_p(1).$$

Proof of Lemma 4. It is convenient to introduce the “vector of seasons notation,” where we stack the seasonal periods of a year into a vector. To do this we explicitly indicate the year by $n = 1, \dots, N$ and the seasonal period by $s = 1, \dots, S$, so that $y_i = \mu + u_i$ is equivalent to $y_{n,s}$ and \tilde{Y}_i^k is equivalent to $\tilde{Y}_{n,s}^k$. For convenience, let the sample period start at $n = 1, s = 1$ and end at $n = N, s = S$. Define the $S \times 1$ vector $u(n) = [u_{n,1}, \dots, u_{n,S}]'$ and the multivariate partial sum process

$$U(n) = \begin{bmatrix} \sum_{\tau=1}^n u_{\tau,1} \\ \sum_{\tau=1}^n u_{\tau,2} \\ \vdots \\ \sum_{\tau=1}^n u_{\tau,S} \end{bmatrix} = \sum_{\tau=1}^n u(\tau)$$

and $U(0) = 0$. Then, by using (11) we have under the null hypothesis

$$\tilde{Y}^k(n) = \begin{bmatrix} \tilde{Y}_{n,1}^k \\ \tilde{Y}_{n,2}^k \\ \vdots \\ \tilde{Y}_{n,S}^k \end{bmatrix} = \Psi^k U(n-1) + \Phi^k u(n) + O_p(T^{-1/2}), \tag{A.4}$$

where the (i, j) elements of the matrices Ψ^k and Φ^k are given by

$$\Psi_{ij}^k = \psi_{i-j-1}^k$$

and

$$\Phi_{ij}^k = \begin{cases} \psi_{i-j-1}^k & \text{for } i - j - 1 > 0 \\ 0 & \text{otherwise.} \end{cases}$$

Because $\Psi^{k'}\Psi^l = 0$ for $k \neq l$ we get

$$\begin{aligned} \frac{1}{N} \sum_{n=1}^N \tilde{Y}^k(n)' \tilde{Y}^l(n) &= \frac{1}{N} \sum_{n=1}^N u(n)' \Phi^{k'} \Phi^l u(n) + \frac{2}{N} \sum_{n=1}^N U(n-1)' \Psi^{k'} \Phi^l u(n) + o_p(1) \\ &= O_p(1). \end{aligned}$$

■