COWLES FOUNDATION FOR RESEARCH IN ECONOMICS
AT YALE UNIVERSITY

Box 2125, Yale Station
New Haven, Connecticut 06520

COWLES FOUNDATION DISCUSSION PAPER NO. 1244

UNIT ROOT LOG PERIODOGRAM REGRESSION

Peter C. B. Phillips

December 1999

Unit Root Log Periodogram Regression*

Peter C. B. Phillips
Cowles Foundation for Research in Economics
Yale University

1 October 1999

Abstract

Log periodogram (LP) regression is shown to be consistent and to have a mixed normal limit distribution when the memory parameter $d = 1$. Gaussian errors are not required. Tests of $d = 1$ based on LP regression are consistent against $d < 1$ alternatives but inconsistent against $d > 1$ alternatives. A test based on a modified LP regression that is consistent in both directions is provided.

JEL Classification: C22

Key words and phrases: Discrete Fourier transform, fractional Brownian motion, fractional integration, log periodogram regression, long memory parameter, nonstationarity, semiparametric estimation and testing, unit root.

1 Introduction

Fractional processes are growing in popularity with empirical researchers in economics and finance. In part, this is explained by their capacity to capture long range characteristics of economic data that elude other models, a feature that is particularly important in modeling the volatility of financial asset returns. In part also, models of fractional integration are attractive to empirical researchers because they provide liberation from the classical dichotomy of I(0) and I(1) time series and applications now cover a range of different time series from asset returns and exchange rates to interest rates and inflation (see Baillie, 1995, for an overview). A natural goal in many of these studies is to assess the extent to which the series under study may depart from a simple unit root model. To make such an assessment, an inferential theory that includes the unit root case is desirable, especially one that retains generality with regard to the short memory component of the series. One of the goals of the present paper is to provides such a theory.

*Thanks go to the NSF for research support under Grant No. SBR 97-30295. Computations were performed in GAUSS and the paper was typed by the author in SW2.5.

Our focus of interest is therefore the estimation of the memory parameter 'd' of a fractionally integrated process X_t in a model of the form

$$(1-L)^d X_t = u_t, \tag{1}$$

where u_t is stationary with zero mean and continuous spectral density $f_u(\lambda) > 0$. In such models, a commonly used estimator in empirical work is the log periodogram (LP) estimator (Geweke and Porter-Hudak, 1983) $\widehat{d}$ obtained from the following least squares regression

$$\log\left(I_x\left(\lambda_s\right)\right) = \widehat{c} - \widehat{d}\log\left|1-e^{i\lambda_s}\right|^2 + \text{residual} \tag{2}$$

taken over fundamental frequencies $\left\{\lambda_s = \frac{2\pi s}{n} : s = 1, ..., m\right\}$ for some $m < n$. Here

$$\widehat{d} = \frac{1}{2}\frac{\sum_{s=1}^{m} x_s \log I_x\left(\lambda_s\right)}{\sum_{s=1}^{m} x_s^2}, \tag{3}$$

where $I_x(\lambda_s) = w_x(\lambda_s)w_x(\lambda_s)^*$ is the periodogram and $w_x(\lambda_s)$ is the discrete Fourier transform (dft), $w_x(\lambda_s) = \frac{1}{\sqrt{2\pi n}}\sum_{t=1}^{n} X_t e^{it\lambda_s}$ of the time series X_t. The regression (2) is motivated by the form of the log spectrum of X_t and has appeal because of its nonparametric treatment of u_t and the convenience of linear least squares. Under Gaussian assumptions and in the stationary case, where $d \in (-\frac{1}{2}, \frac{1}{2})$, Robinson (1995) developed consistency and asymptotic normality results for a version of $\widehat{d}$ which trims out low frequencies periodogram ordinates (i.e. takes $s \geq l$, for some $m > l > 1$), as suggested by Künsch (1986). Hurvich and Beltrao (1993) have developed data-driven criteria for the selection of m; and Hurvich, Deo, and Brodsky (1998) extend Robinson's (1995) results to include low frequencies ordinates and find an optimal choice of the number of periodogram ordinates in the regression. This work provides a foundation of asymptotic theory validating (2) for use with Gaussian data in the stationary case. However, there is presently no theory for the unit root case where $d = 1$ in (1).

The present paper studies LP regression in this unit root case. This case is of interest for several reasons. First, the unit root model is an important special case of (1) which has received substantial attention in itself, but which is presently not covered by the existing theory of semiparametric estimation of d. The case is particularly important in economic applications, where, as indicated earlier, one is often concerned about assessing the extent of the departure from a simple unit root model. The intensive interest in testing the unit root hypothesis in the econometric literature has focussed on tests against stationary alternatives within the context of autoregressive models, rather than tests against nonstationary fractional alternatives. Second, it is well known that the corresponding semiparametric unit root limit theory is nonstandard (Phillips, 1987) and it is of interest to discover whether there are any unusual features to the limit theory in the semiparametric estimation of d when $d = 1$. Third, it has recently been shown by Kim and Phillips (1999a) that $\widehat{d}$ is inconsistent when $d > 1$. So $d = 1$ turns out to be the boundary case for consistent estimation by log periodogram regression. Moreover, Kim and Phillips show that $\widehat{d} \to_p 1$ when $d > 1$, and this asymptotic bias in estimation is certain to bias inference about departures from unit root model if it is not corrected.

The work reported in this paper complements some other recent research on LP regression and testing in the nonstationary case. Hurvich and Ray (1995) looked at the behavior of periodogram ordinates of a fractionally integrated process with memory parameter $d \in [0.5, 1.5)$ and found evidence of bias in log periodogram regression when $d > 1$. Velasco (1999a) showed consistency of an LP estimator that trims out low frequency ordinates, under Gaussian assumptions and for $\frac{1}{2} < d < 1$. Kim and Phillips (1999a) showed log periodogram regression is consistent for $\frac{1}{2} < d < 1$ without requiring Gaussianity or trimming, as well as demonstrating inconsistency for $d > 1$. Some simulation results covering nonstationary cases were reported in Hurvich and Ray (1995) and Velasco (1999b), both revealing evidence of estimation bias when $d > 1$. Finally, Robinson (1994) and Tanaka (1999) show how to use Lagrange multiplier and Wald theory for testing values of d in parametric models that include both stationary and non stationary cases, but exclude weak nonparametric dependence; and Robinson's (1994) tests were applied by Gil-Alaña and Robinson (1997) to the extended version of the Nelson and Plosser (1982) data set.

The approach in the present paper draws on an exact representation of the discrete Fourier transform (dft) in the unit root case. This representation was developed by the author in other work (Phillips, 1999, and Corbae, Ouliaris and Phillips, 1999) and the aspects of the general theory given in Phillips (1999) that we need here are reviewed briefly in Section 2. Section 3 gives the limit theory for $\widehat{d}$ in the unit root case. Gaussianity is not assumed and the approach to establishing a central limit theorem for linear combinations of nonlinear functions of periodogram ordinates is based on a new embedding argument that relies on a strong approximation for partial sums of linear processes, which is given in the Appendix (Lemma D) together with other technical results that are needed. This approach to a CLT for linear combinations of discrete Fourier transforms is useful outside the present context of log periodogram regression and is of some independent interest. Some additional theory and implications for testing are discussed in Section 4. Concluding remarks are made in Section 5 and proofs are given in the Appendix in Section 6.

2 Preliminaries

This section gives explicit assumptions for our development and briefly reviews some representations of the dft of a fractionally integrated time series given in Phillips (1999). While these representations are valid in both stationary and nonstationary cases, they will be used here only in the unit root case, where they take an especially simple form.

The fractionally integrated process X_t is defined as in (1), with $u_j = 0$ for all $j \leq 0$. Explicit conditions on u_t $(t > 0)$ are given in the following.

Assumption A *For all $t > 0$, u_t has Wold representation*

$$u_t = C(L)\,\varepsilon_t = \sum_{j=0}^{\infty} c_j \varepsilon_{t-j}, \quad \sum_{j=0}^{\infty} j\,|c_j| < \infty, \quad C(1) \neq 0, \tag{4}$$

with $\varepsilon_t = iid\left(0, \sigma^2\right)$ with $E(|\varepsilon_t|^q) < \infty$, for some $q > 4$.

The q'th moment condition on the errors ε_t is useful in the embedding argument used in the proof of our main theorem and in lemmas D and E of Section 5. The linear process formulation and summability condition in (4) covers a wide class of short memory processes and, as in Phillips (1999), enables us to use a decomposition technique to develop a convenient representation of the dft of a fractionally integrated process. In particular, from theorem 2.2 of Phillips (1999) we have

$$w_u(\lambda) = w_x(\lambda)\, D_n\left(e^{i\lambda}; d\right) + \frac{1}{\sqrt{2\pi n}}\left(\widetilde{X}_{\lambda 0}(d) - e^{in\lambda}\widetilde{X}_{\lambda n}(d)\right), \tag{5}$$

where

$$\widetilde{X}_{\lambda n}(d) = \widetilde{D}_{n\lambda}\left(e^{-i\lambda}L; d\right) X_n = \sum_{p=0}^{n-1} \widetilde{d}_{\lambda p} e^{-ip\lambda} X_{n-p}. \tag{6}$$

When $u_t = 0$ for $t \leq 0$, as is assumed above, $X_t = 0$ for $t \leq 0$ and, hence, $\widetilde{X}_{\lambda 0}(d) = 0$. In this case, (5) becomes

$$\begin{aligned} w_u(\lambda) &= w_x(\lambda)\, D_n\left(e^{i\lambda}; d\right) - \frac{e^{in\lambda}}{\sqrt{2\pi n}} \widetilde{D}_{n\lambda}\left(e^{-i\lambda}L; d\right) X_n \\ &= w_x(\lambda)\, D_n\left(e^{i\lambda}; d\right) - \frac{1}{\sqrt{2\pi n}} e^{in\lambda}\widetilde{X}_{\lambda n}(d). \end{aligned} \tag{7}$$

Equation (7) shows that the exact relation between $w_x(\lambda)$ and $w_u(\lambda)$ involves a correction term that depends on $\widetilde{X}_{\lambda n}(d)$. When $d = 1$, we have the simplification $\widetilde{X}_{\lambda n}(1) = -e^{i\lambda}X_n$, and (5) becomes

$$w_u(\lambda) = \left(1 - e^{i\lambda}\right) w_x(\lambda) + \frac{e^{i\lambda}}{\sqrt{2\pi n}}\left(e^{in\lambda}X_n - X_0\right), \tag{8}$$

or simply

$$w_u(\lambda) = \left(1 - e^{i\lambda}\right) w_x(\lambda) + \frac{e^{i\lambda}}{\sqrt{2\pi n}} e^{in\lambda} X_n,$$

when $X_0 = 0$, a result given earlier in Corbae, Ouliaris and Phillips (1999). Rewriting this last equation as

$$w_x(\lambda) = \frac{w_u(\lambda)}{1 - e^{i\lambda}} - \frac{e^{i(n+1)\lambda}}{1 - e^{i\lambda}} \frac{X_n}{\sqrt{2\pi n}} \tag{9}$$

$$= \frac{1}{1 - e^{i\lambda}}\left[w_u(\lambda) - e^{i(n+1)\lambda}\frac{X_n}{\sqrt{2\pi n}}\right], \tag{10}$$

it is immediately apparent that both components will influence the asymptotic behavior of the data dft $w_x(\lambda)$ when $d = 1$.

The representation (5) shows that the dft of a fractionally integrated process comprises two distinct components. The first of these is the dft of the innovations u_t, scaled by the transfer function of the differencing filter, $D_n\left(e^{i\lambda}; d\right)$. The second involves a weighted sinusoidal sum, $\widetilde{X}_{\lambda n}(d)$, of the observations X_t. When $d = 1$, both components simplify. The transfer function is then simply $D_n\left(e^{i\lambda}; 1\right) = 1 - e^{i\lambda}$ and the sinusoidal sum (6) becomes $\widetilde{X}_{\lambda n}(1) = -e^{i\lambda}X_n$, which depends only on the final sample observation.

Since the limit behavior of these two components is rather easily obtained in the unit root case, we can expect to develop dft asymptotics for integrated processes by analyzing the two terms in (9), rather than by attempting to work directly with the dft of X_t itself. Some results along these lines are presented in Phillips (1999) and Corbae, Ouliaris and Phillips (1999). Moreover, the representations follow by algebraic simplification and so these, and results obtained from them, turn out not to depend upon distributional specifications like Gaussianity.

At the fundamental frequencies $\lambda_s = \frac{2\pi s}{n}$, $s = 1, ..., n$, (10) becomes

$$w_x(\lambda_s) = \frac{1}{1 - e^{i\lambda_s}}\left[w_u(\lambda_s) - e^{i\lambda_s}\frac{X_n}{\sqrt{2\pi n}}\right]. \tag{11}$$

The component $(2\pi n)^{-1/2}X_n$ in (11) is independent of frequency and influences the asymptotic behavior of $w_x(\lambda_s)$ for all values of s. This property means that $w_x(\lambda_s)$ is spatially correlated across all the fundamental frequencies. In effect, there is leakage across all the fundamental frequencies from the zero frequency ($s = 0$), i.e. from $w_u(\lambda_0) = (2\pi n)^{-1/2}X_n$.

A further complicating factor in log periodogram regression is that one needs to work with a logarithmic function of the periodogram ordinates. In effect, the model underlying the empirical regression (2) involves the logarithm of the squared modulus of (11), i.e.,

$$\log|w_x(\lambda_s)|^2 = -2\log\left|1 - e^{i\lambda_s}\right| + \log\left|w_u(\lambda_s) - e^{i\lambda_s}\frac{X_n}{\sqrt{2\pi n}}\right|^2. \tag{12}$$

The second term in (12) involves the dft of the stationary errors, $w_u(\lambda_s)$, coupled with the leakage factor $w_u(\lambda_0) = (2\pi n)^{-1/2}X_n$ from $s = 0$. The asymptotic analysis of log periodogram regression requires that we treat both the nonlinearity in (12) and the spatial correlation arising from this leakage.

3 Log-Periodogram Regression in the Unit Root Case

The approach we adopt in resolving the technical difficulties just mentioned is to expand the probability space in which the processes are defined in such a way that the data can be represented almost surely and up to a negligible error in terms of a Brownian motion that is defined on the same space. The argument invokes an almost sure invariance principle and embedding.

More specifically, define the partial sum process $S_k = \sum_{j=1}^k u_j$ for $k \geq 1$, and $S_0 = 0$, for $k = 0$. Since u_j has finite moments of order $q > 2p > 4$, we can expand the probability space as necessary to set up a partial sum process that is distributionally equivalent to S_k and a Brownian motion $B(\cdot)$ with variance $2\pi f_u(0)$ on the same space for which

$$\sup_{0\leq k\leq n} |S_k - B(k)| = o_{a.s}(n^{\frac{1}{p}}), \tag{13}$$

giving a uniform approximation to S_k over $0 \leq k \leq n$ in terms of the Brownian motion B. In (13), p satisfies $q > 2p > 4$, so that $p > 2$. When q and p are large, the error order of magnitude in this approximation becomes small and it is bounded, or $O_{a.s}(1)$, when S_k is Gaussian. When the components u_j in S_k are independent, almost sure invariance principles or strong approximations of the type (13) have been proved by many authors using a variety of techniques, a popular recent approach being the Hungarian construction, e.g. Shorack and Wellner (1986) and Csörgõ, M. and L. Horváth (1993). A strong approximation result justifying the representation (13) in the case where u_t is a linear processes satisfying Assumption A is given in lemma D in the Appendix. Setting $S_{nk} = n^{-1/2}\Sigma_{j=1}^k u_j$, we can write this approximation in the form

$$\sup_{0\leq k\leq n} \left| S_{nk} - B(\frac{k}{n}) \right| = o_{a.s}\left(\frac{1}{n^{\frac{1}{2}-\frac{1}{p}}}\right). \tag{14}$$

As shown in (36) and (37) in the Appendix, the approximation (14) leads to the simultaneous representation of the dft

$$\begin{aligned} w_u(\lambda_s) &= \frac{1}{\sqrt{2\pi}} \int_0^1 e^{2\pi i s r} dB(r) + O_{a.s}\left(\frac{m}{n}\right) + o_{a.s}\left(\frac{1}{n^{\frac{1}{2}-\frac{1}{p}}}\right) \\ &= \xi_s + O_{a.s}\left(\frac{m}{n}\right) + o_{a.s}\left(\frac{1}{n^{\frac{1}{2}-\frac{1}{p}}}\right), \end{aligned} \tag{15}$$

with the error magnitude holding uniformly over $s = 1, ..., m$, and the representation

$$\frac{X_n}{\sqrt{2\pi n}} = \frac{1}{\sqrt{2\pi}} B(1) + o_{a.s}\left(\frac{1}{n^{\frac{1}{2}-\frac{1}{p}}}\right) = \eta + +o_{a.s}\left(\frac{1}{n^{\frac{1}{2}-\frac{1}{p}}}\right). \tag{16}$$

The variables $(\xi_s)_{s=1}^m$ and η that appear in (15 and (16) are Gaussian and independent.

The simple form of (15) and (16) enables us to develop a limit theory for frequency averages of the second term in (12). Log periodogram asymptotics follow directly. The outcome is the following result, which gives the asymptotic distribution of $\widehat{d}$ when $d = 1$.

3.1 Theorem *Let X_t follow* (1) *with $d = 1$ and u_t satisfy* (4). *If*

$$\frac{m^{\frac{3}{2}}}{n} + \frac{\sqrt{m}}{n^{\frac{1}{2}-\frac{1}{p}}} \to 0, \tag{17}$$

then

$$\sqrt{m}\left(\widehat{d}-d\right)\to_d MN\left(0,\frac{1}{4}\sigma^2\left(W\right)\right)\equiv\int_0^\infty N\left(0,\frac{1}{4}\sigma^2\left(W\right)\right)\mathrm{pdf}\left(W\right)dW \qquad (18)$$

Here, W is χ_1^2 with $\mathrm{pdf}\left(W\right)=[2^{\frac{1}{2}}\Gamma(\frac{1}{2})]^{-1}e^{-W/2}W^{-\frac{1}{2}}$ *and*

$$\sigma^2\left(W\right)=e^{-W}\sum_{j=0}^{\infty}\frac{W^j}{j!}\left\{\psi\left(1+j\right)^2+\psi'\left(1+j\right)\right\}-\left[e^{-W}\sum_{j=0}^{\infty}\frac{W^j}{j!}\psi\left(1+j\right)\right]^2,$$

where $\psi(z)=\Gamma'(z)/\Gamma(z)$ is the psi function, the logarithmic derivative of the gamma function Γ, and $\psi'(z)$ is the trigamma function, its first derivative.

Condition (17) on the frequency band $\{\lambda_s,\ 1<s<m\}$ restricts the range of the effective sample size m (the number of periodogram ordinates) in the log periodogram regression (2). The condition holds when $m=O(n^{\frac{2}{3}-\delta})$ for any $\delta>0$, when $p>6$.

The variance of the limit distribution (18) is 1/4 times

$$\begin{aligned}
&\frac{1}{2^{\frac{1}{2}}\Gamma(\frac{1}{2})}\int_0^\infty\left(e^{-W}\sum_{j=0}^{\infty}\frac{W^j}{j!}\left\{\psi\left(1+j\right)^2+\psi'\left(1+j\right)\right\}-\left[e^{-W}\sum_{j=0}^{\infty}\frac{W^j}{j!}\psi\left(1+j\right)\right]^2\right)e^{-W/2}W^{-\frac{1}{2}}dW\\
=&\ \frac{1}{\sqrt{2\pi}}\sum_{j=0}^{\infty}\frac{\left\{\psi\left(1+j\right)^2+\psi'\left(1+j\right)\right\}}{j!}\int_0^\infty e^{-\frac{3}{2}W}W^{j-\frac{1}{2}}dW\\
&-\frac{1}{\sqrt{2\pi}}\sum_{j=0}^{\infty}\sum_{k=0}^{\infty}\frac{\psi\left(1+j\right)\psi\left(1+k\right)}{j!k!}\int_0^\infty e^{-\frac{5}{2}W}W^{j+k-\frac{1}{2}}dW\\
=&\ \frac{1}{\sqrt{2\pi}}\sum_{j=0}^{\infty}\frac{\left\{\psi\left(1+j\right)^2+\psi'\left(1+j\right)\right\}}{j!}\frac{\Gamma\left(j+\frac{1}{2}\right)}{\left(\frac{3}{2}\right)^{j+\frac{1}{2}}}\\
&-\frac{1}{\sqrt{2\pi}}\sum_{j=0}^{\infty}\sum_{k=0}^{\infty}\frac{\psi\left(1+j\right)\psi\left(1+k\right)}{j!k!}\frac{\Gamma\left(j+k+\frac{1}{2}\right)}{\left(\frac{5}{2}\right)^{j+k+\frac{1}{2}}}\\
=&\ \frac{\pi^2}{6}+\frac{1}{\sqrt{3}}\sum_{j=0}^{\infty}\frac{\left[\psi\left(1+j\right)^2-\sum_{k=1}^{j}\frac{1}{k^2}\right]\left(\frac{1}{2}\right)_j\left(\frac{2}{3}\right)^j}{j!}\\
&-\frac{1}{\sqrt{5}}\sum_{j=0}^{\infty}\sum_{k=0}^{\infty}\frac{\psi\left(1+j\right)\psi\left(1+k\right)}{j!k!}\left(\frac{1}{2}\right)_{j+k}\left(\frac{2}{5}\right)^{j+k}
\end{aligned} \qquad (19)$$

since

$$\frac{1}{\sqrt{2\pi}}\sum_{j=0}^{\infty}\frac{\psi\left(1+j\right)^2}{j!}\frac{\Gamma\left(j+\frac{1}{2}\right)}{\left(\frac{3}{2}\right)^{j+\frac{1}{2}}}=\frac{1}{\sqrt{3}}\sum_{j=0}^{\infty}\frac{\psi\left(1+j\right)^2\left(\frac{1}{2}\right)_j\left(\frac{2}{3}\right)^j}{j!},$$

$$\psi'\left(1+j\right)=\frac{\pi^2}{6}-\sum_{k=1}^{j}\frac{1}{k^2}, \qquad (20)$$

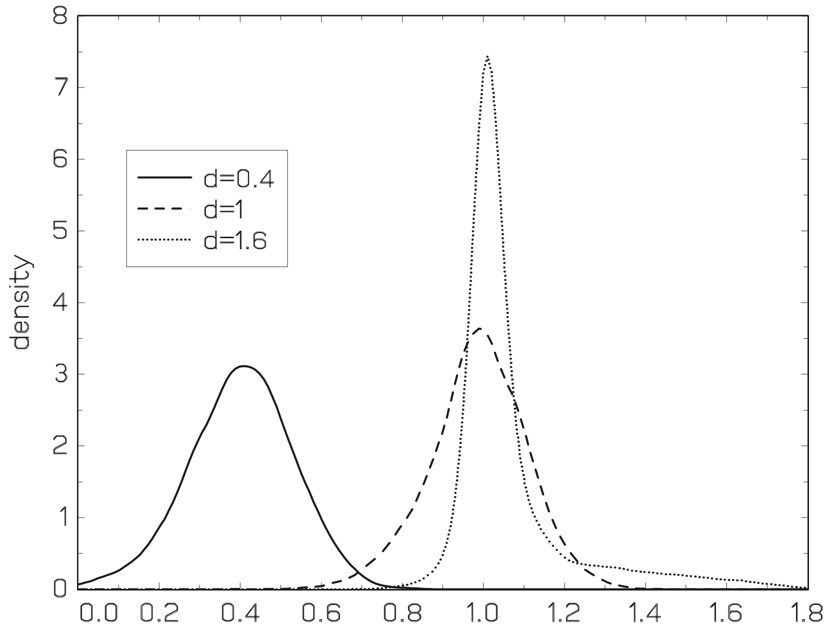

Figure 1: Densities of log periodogram estimator: $n = 200$, $m = [n^{0.65}]$.

and

$$\psi(1+j) = \begin{cases} -C + \sum_{k=1}^{j} \frac{1}{k} & j \geq 1 \\ -C & j = 0 \end{cases} \qquad C = 0.577215 \text{ (Euler's constant)}. \tag{21}$$

where (20) and (21) are standard - e.g. Gradshteyn and Ryzhik (1965, 8.365 & 8.366). The summation in the third term of (19) can be written as

$$\begin{aligned} &\sum_{j=0}^{\infty}\sum_{k=0}^{\infty} \frac{\psi(1+j)\,\psi(1+k)}{j!k!} \left(\frac{1}{2}\right)_{j+k} \left(\frac{2}{5}\right)^{j+k} \\ = \;& \sum_{j=0}^{\infty} \frac{\psi(1+j)\left(\frac{1}{2}\right)_j \left(\frac{2}{5}\right)^j}{j!} \sum_{k=0}^{\infty} \frac{\psi(1+k)}{k!} \left(\frac{1}{2}+j\right)_k \left(\frac{2}{5}\right)^k . \end{aligned}$$

Hence, the variance of the limit distribution of $\sqrt{m}\left(\widehat{d} - d\right)$ is

$$\begin{aligned} \sigma_d^2 \;=\; & \frac{\pi^2}{24} + \frac{1}{4\sqrt{3}} \sum_{j=0}^{\infty} \frac{\left[\psi(1+j)^2 - \sum_{k=1}^{j} \frac{1}{k^2}\right] \left(\frac{1}{2}\right)_j \left(\frac{2}{3}\right)^j}{j!} \\ & - \frac{1}{4\sqrt{5}} \sum_{j=0}^{\infty} \frac{\psi(1+j)\left(\frac{1}{2}\right)_j \left(\frac{2}{5}\right)^j}{j!} \sum_{k=0}^{\infty} \frac{\psi(1+k)}{k!} \left(\frac{1}{2}+j\right)_k \left(\frac{2}{5}\right)^k . \end{aligned}$$

Numerical evaluation gives $\sigma_d^2 = 0.3948$, which is slightly smaller than $\pi^2/24 = 0.4112$, the limiting variance of $\sqrt{m}\left(\widehat{d} - d\right)$ in the stationary case. Thus, the limit distribution of the log periodogram estimator in the unit root case, although mixed normal, has less dispersion than in the stationary case. This is confirmed in finite sample simulations. Figure 1 shows the sampling distributions of the log periodogram estimator (based on 10,000 replications) for several values of d when $n = 200$ and $m = \left[n^{0.65}\right]$. The reduction in the variance when $d = 1$ in comparison with $d = 0.4$ is apparent, as is the downward bias towards unity when $d > 1$. The latter is systematically explored in Kim and Phillips (1999a).

4 Testing a Unit Root against Fractional Alternatives

Theorem 3.1 shows that the conventional LP regression theory, viz. the limit distribution $N\left(0, \frac{\pi^2}{24}\right)$ for $\sqrt{m}\left(\widehat{d} - d\right)$ that is known to apply in the stationary case (Robinson, 1995, Hurvich et al, 1998), is invalid at $d = 1$. Nevertheless, the limit theory in theorem 3.1 could easily be used for testing $d = 1$ simply by adjusting critical values to conform with the correct mixed normal limit distribution given in (18). Such critical values could readily be computed by numerical methods. The resulting test would involve the computation of $\sqrt{m}\left(\widehat{d} - 1\right)$ and calibration against the critical values from the limit distribution (18). Since it is known (Kim and Phillips, 1999a) that LP regression is consistent for $-\frac{1}{2} < d < 1$, such a test would be consistent against such alternatives. However, it is also known (Kim and Phillips, 1999a) that LP regression is inconsistent when $d > 1$, and indeed $\widehat{d} \to_p 1$ for all $d > 1$. Thus, the test would be inconsistent against alternatives $d > 1$.

The theory in Sections 2 and 3 gives us a simple way of constructing a valid test. We need an estimator of d that is consistent for both $d \leq 1$ and $d > 1$ to accomplish this. Consistent estimation of d around unity is possible by using a new estimator called modified log periodogram regression suggested in Phillips (1999). Here, the estimator can be obtained quite simply from (9), which at the fundatmental frequencies $\{\lambda_s : s = 1, ..., m\}$ is

$$w_x(\lambda_s) = \frac{w_u(\lambda_s)}{1 - e^{i\lambda_s}} - \frac{e^{i\lambda_s}}{1 - e^{i\lambda_s}} \frac{X_n}{\sqrt{2\pi n}}. \tag{22}$$

We construct the modified dft

$$v_x(\lambda_s) = w_x(\lambda_s) + \frac{e^{i\lambda_s}}{1 - e^{i\lambda_s}} \frac{X_n}{\sqrt{2\pi n}}, \tag{23}$$

and associated periodogram ordinates $I_v(\lambda_s) = v_x(\lambda_s)\, v_x(\lambda_s)^*$. Observe that $v_x(\lambda_s)$ and $I_v(\lambda_s)$ are both observable functions of the data. With this simple data transformation we have the following exact model

$$\log I_v(\lambda_s) = c - 2\log\left|1 - e^{i\lambda_s}\right| + \log \frac{I_u(\lambda_s)}{f_u(0)},$$

with $c = \log f_u(0)$. Modified LP regression is simply a least squares regression of $\log I_v(\lambda_s)$ on $\log\left|1-e^{i\lambda_s}\right|$. Set $a_s = \log\left|1-e^{i\lambda_s}\right|$ and $x_s = a_s - \overline{a}$, with $\overline{a} = m^{-1}\sum_{s=1}^{m} a_s$. Then, the modified LP estimator is simply

$$\widetilde{d} = \frac{1}{2}\frac{\sum_{s=1}^{m} x_s \log I_v(\lambda_s)}{\sum_{s=1}^{m} x_s^2}. \tag{24}$$

The limit distribution of this estimator when $d = 1$ is given in the following result.

4.1 Theorem *Let X_t follow* (1) *with $d = 1$ and u_t satisfy* (4)*. If*

$$\frac{m^{\frac{3}{2}}}{n} + \frac{\sqrt{m}}{n^{\frac{1}{2}-\frac{1}{p}}} \to 0,$$

then

$$\sqrt{m}\left(\widetilde{d} - d\right) \to_d N\left(0, \frac{\pi^2}{24}\right). \tag{25}$$

Thus, $\widetilde{d}$ has the same limit distribution at $d = 1$ as the well known limit distribution of the LP estimator in the stationary case (Robinson, 1995, and Hurvich et al., 1998). In fact, Kim and Phillips (1999b) show that the limit theory (25) for $\widetilde{d}$ applies for values of d over the range $0 < d < 2$. In particular, $\widetilde{d}$ is consistent for values of d around unity and not equal to unity. Thus, a semiparametric test for a unit root against fractional alternatives can be based simply on the statistic

$$Z_d = \frac{\sqrt{m}\left(\widetilde{d} - 1\right)}{\pi/\sqrt{24}},$$

with critical values obtained from the standard normal distribution. This test is consistent against both $d < 1$ and $d > 1$ fractional alternatives.

5 Concluding Remarks

Log periodogram regression is shown to be consistent and to have a mixed normal limit distribution when $d = 1$. This case turns out to define the boundary for consistent estimation of d by log periodogram regression. For larger values of d, it is known that $\widehat{d} \to_p 1$ (Kim and Phillips, 1999a). Thus, LP regression encounters nonstandard limit theory at the unit root boundary, just like the serial correlation coefficient, but of a different form.

In view of the inconsistency of LP regression when the memory parameter $d > 1$, the LP estimator $\widehat{d}$ is unsuitable for testing a unit root against two-sided fractional alternatives. Fortunately, there is a simple alternative procedure based on modified LP regression which produces consistent estimates of d for $d > 1$ and for $d \leq 1$. Tests for a unit root based on the modified LP estimator are easy to construct, involve only the standard normal distribution and apply under nonparametric weak dependence for the short memory component. They therefore seem suitable for general implementation in empirical research. As in the case of conventional unit root tests, when there are deterministic trends in the model, it is desirable to extract the trends under the null as in Schmidt and Phillips (1992) before implementing the Z_d test given here. Additionally, issues relating to finding an 'optimal' choice of the number of ordinates m in the construction of the estimator $\widetilde{d}$, as well as developing data-determined methods of choosing m are still to be considered.

6 Appendix

The first two lemmas report moment expressions which can be obtained straightforwardly from known results. They are given here for convenience. Lemma C gives the limit distribtion of a sample average of nonlinear functions of correlated Gaussian variates, which is useful in proving the main theorem. Lemma D gives a strong approximation result for partial sums of linear processes satisfying Assumption A with $q > 4$. A weaker approximation result that holds when $q > 2$ is given in Lemma E.

6.1 Lemma A *If* $\mathrm{Re}\,(a)\,, \mathrm{Re}\,(\nu) > 0$, *then*

$$\int_0^\infty (\log z)\, e^{-az} z^{\nu-1} dz = \frac{\Gamma(\nu)}{a^\nu}\left[\psi(\nu) - \log(a)\right],$$

and

$$\int_0^\infty (\log z)^2\, e^{-az} z^{\nu-1} dz = \frac{\Gamma(\nu)}{a^\nu}\left\{\left[\psi(\nu) - \log(a)\right]^2 + \psi'(\nu)\right\},$$

where $\psi(\nu)$ *is the psi function, the logarithmic derivative of the gamma function* $\Gamma(\nu)$, *and* $\psi'(\nu)$, *the first derivative of* $\psi(\nu)$, *is the trigamma function.*

6.2 Proof Start by differentiating with respect to ν the equation

$$\frac{\Gamma(\nu)}{a^\nu} = \int_0^\infty e^{-az} z^{\nu-1} dz,$$

giving the first result

$$\frac{\Gamma'(\nu)}{a^\nu} - \frac{\Gamma(\nu)\log a}{a^\nu} = \frac{\Gamma(\nu)}{a^\nu}\left[\psi(\nu) - \log(a)\right] = \int_0^\infty e^{-az} z^{\nu-1} \log(z)\, dz,$$

where $\psi(\nu)$ is the logarithmic derivative of the gamma function. Subsequent differentiation yields the second result.

$$\frac{\Gamma(\nu)}{a^\nu}\left\{\left[\psi(\nu) - \log(a)\right]^2 + \psi'(\nu)\right\} = \int_0^\infty e^{-az} z^{\nu-1} (\log z)^2\, dz.$$

■

6.3 Lemma B

$$E\left(\log\left(\chi_\nu^2(\delta)\right)\right) = e^{-\delta/2}\sum_{j=0}^{\infty}\frac{\left(\frac{\delta}{2}\right)^j}{j!}\psi\left(\frac{\nu}{2}+j\right)+\log 2,$$

$$\begin{aligned}\mathrm{Var}\left(\log\left(\chi_\nu^2(\delta)\right)\right) &= e^{-\delta/2}\sum_{j=0}^{\infty}\frac{\left(\frac{\delta}{2}\right)^j}{j!}\left\{\psi\left(\frac{\nu}{2}+j\right)^2+\psi'\left(\frac{\nu}{2}+j\right)\right\}\\ &\quad -\left[e^{-\delta/2}\sum_{j=0}^{\infty}\frac{\left(\frac{\delta}{2}\right)^j}{j!}\psi\left(\frac{\nu}{2}+j\right)\right]^2.\end{aligned}$$

6.4 Proof The density of $\chi_\nu^2(\delta)$ is

$$pdf(z) = \frac{e^{-\delta/2-z/2}}{\Gamma\left(\frac{\nu}{2}\right)2^{\frac{\nu}{2}}}\sum_{j=0}^{\infty}\left(\frac{\delta}{4}\right)^j\frac{z^{\frac{\nu}{2}-1+j}}{j!\left(\frac{\nu}{2}\right)_j} = \frac{e^{-\delta/2-z/2}z^{\frac{\nu}{2}-1}}{\Gamma\left(\frac{\nu}{2}\right)2^{\frac{\nu}{2}}}\,{}_0F_1\left(\frac{\nu}{2},\frac{\delta}{4}z\right),$$

where ${}_0F_1(a,z) = \sum_{j=0}^{\infty}\frac{z^j}{j!(a)_j}$. So

$$\begin{aligned}E\left(\log\left(\chi_\nu^2(\delta)\right)\right) &= \frac{e^{-\delta/2}}{\Gamma\left(\frac{\nu}{2}\right)2^{\frac{\nu}{2}}}\sum_{j=0}^{\infty}\frac{\left(\frac{\delta}{4}\right)^j}{j!\left(\frac{\nu}{2}\right)_j}\int_0^{\infty}e^{-z/2}z^{\frac{\nu}{2}-1+j}\log(z)\,dz\\ &= \frac{e^{-\delta/2}}{\Gamma\left(\frac{\nu}{2}\right)2^{\frac{\nu}{2}}}\sum_{j=0}^{\infty}\frac{\left(\frac{\delta}{4}\right)^j}{j!\left(\frac{\nu}{2}\right)_j}\frac{\Gamma\left(\frac{\nu}{2}+j\right)}{\left(\frac{1}{2}\right)^{\frac{\nu}{2}+j}}\left[\psi\left(\frac{\nu}{2}+j\right)+\log 2\right]\\ &= e^{-\delta/2}\sum_{j=0}^{\infty}\frac{\left(\frac{\delta}{2}\right)^j}{j!}\left[\psi\left(\frac{\nu}{2}+j\right)+\log 2\right]\\ &= e^{-\delta/2}\sum_{j=0}^{\infty}\frac{\left(\frac{\delta}{2}\right)^j}{j!}\psi\left(\frac{\nu}{2}+j\right)+\log 2.\end{aligned}$$

Next

$$\begin{aligned}E\left(\log\left(\chi_\nu^2(\delta)\right)\right)^2 &= \frac{e^{-\delta/2}}{\Gamma\left(\frac{\nu}{2}\right)2^{\frac{\nu}{2}}}\sum_{j=0}^{\infty}\frac{\left(\frac{\delta}{4}\right)^j}{j!\left(\frac{\nu}{2}\right)_j}\int_0^{\infty}e^{-z/2}z^{\frac{\nu}{2}-1+j}(\log z)^2\,dz\\ &= \frac{e^{-\delta/2}}{\Gamma\left(\frac{\nu}{2}\right)2^{\frac{\nu}{2}}}\sum_{j=0}^{\infty}\frac{\left(\frac{\delta}{4}\right)^j}{j!\left(\frac{\nu}{2}\right)_j}\frac{\Gamma\left(\frac{\nu}{2}+j\right)}{\left(\frac{1}{2}\right)^{\frac{\nu}{2}+j}}\left\{\left[\psi\left(\frac{\nu}{2}+j\right)+\log(2)\right]^2+\psi'\left(\frac{\nu}{2}+j\right)\right\}\\ &= e^{-\delta/2}\sum_{j=0}^{\infty}\frac{\left(\frac{\delta}{2}\right)^j}{j!}\left\{\left[\psi\left(\frac{\nu}{2}+j\right)+\log(2)\right]^2+\psi'\left(\frac{\nu}{2}+j\right)\right\},\end{aligned}$$

and then

$$\begin{aligned}
\operatorname{Var}\left(\log\left(\chi^2_\nu(\delta)\right)\right) &= E\left(\log\left(\chi^2_\nu(\delta)\right)\right)^2 - \left(E\left(\log\left(\chi^2_\nu(\delta)\right)\right)\right)^2 \\
&= e^{-\delta/2}\sum_{j=0}^{\infty}\frac{\left(\frac{\delta}{2}\right)^j}{j!}\left\{\left[\psi\left(\frac{\nu}{2}+j\right)+\log(2)\right]^2+\psi'\left(\frac{\nu}{2}+j\right)\right\} \\
&\quad -\left[e^{-\delta/2}\sum_{j=0}^{\infty}\frac{\left(\frac{\delta}{2}\right)^j}{j!}\psi\left(\frac{\nu}{2}+j\right)+\log(2)\right]^2 \\
&= e^{-\delta/2}\sum_{j=0}^{\infty}\frac{\left(\frac{\delta}{2}\right)^j}{j!}\left\{\psi\left(\frac{\nu}{2}+j\right)^2+\psi'\left(\frac{\nu}{2}+j\right)\right\} \\
&\quad -\left[e^{-\delta/2}\sum_{j=0}^{\infty}\frac{\left(\frac{\delta}{2}\right)^j}{j!}\psi\left(\frac{\nu}{2}+j\right)\right]^2,
\end{aligned}$$

giving the two stated results. ■

6.5 Lemma C *Let* $\xi_s = \frac{1}{\sqrt{2\pi}}\int_0^1 e^{2\pi isr}dB(r)$ *and* $\eta = \frac{1}{\sqrt{2\pi}}\int_0^1 dB(r)$, *where* B *is Brownian motion with variance* $2\pi f_u(0)$. *Then, if* $x_s = a_s - \overline{a}$ *where* $a_s = \log\left|1-e^{i\lambda_s}\right|$ *and* $\overline{a} = m^{-1}\sum_{s=1}^m a_s$,

$$\frac{1}{\sqrt{m}}\sum_{j=1}^{m} x_s \log|\xi_s-\eta|^2 \to_d MN\left(0,\sigma^2(W)\right) \equiv \int_0^\infty N\left(0,\sigma^2(W)\right)\text{pdf}(W)\,dW,$$

where

$$\sigma^2(W) = e^{-W}\sum_{j=0}^{\infty}\frac{W^j}{j!}\left\{\psi(1+j)^2+\psi'(1+j)\right\} - \left[e^{-W}\sum_{j=0}^{\infty}\frac{W^j}{j!}\psi(1+j)\right]^2,$$

and W *is chi-squared with one degree of freedom and* $\text{pdf}(W) = [2^{\frac{1}{2}}\Gamma(\frac{1}{2})]^{-1}e^{-W/2}W^{-\frac{1}{2}}$.

6.6 Proof Set

$$\xi_s = \frac{1}{\sqrt{2\pi}}\int_0^1 e^{2\pi isr}dB(r) = \sqrt{f_u(0)}\int_0^1 e^{2\pi isr}dW(r) = \sqrt{f_u(0)}Z_s, \text{ say,}$$

and

$$\eta = \sqrt{f_u(0)}Y$$

where W is standard Brownian motion, and $\{Z_s\}_1^m \equiv iid\ N_c(0,1)$ and is independent of Y, which is $N(0,1)$. It is convenient to write $Z_s = \zeta_{1s}+\zeta_{2s}i$. The components ζ_{1s}, ζ_{2s} in this decomposition are independent and each is $N\left(0,\frac{1}{2}\right)$. Then

$$\log|\xi_s-\eta|^2 = \log(f_u(0)) + \log\left[(\zeta_{1s}-Y)^2+\zeta_{2s}^2\right]$$

$$
\begin{aligned}
&= \log\left(\frac{1}{2}f_u(0)\right) + \log\left[2\left\{(\zeta_{1s}-Y)^2+\zeta_{2s}^2\right\}\right] \\
&= \log\left(\frac{1}{2}f_u(0)\right) + \log\left[G_{sY}\right], \text{ say.} \qquad (26)
\end{aligned}
$$

Conditional on Y, $\zeta_{1s}-Y$ is $N\left(-Y,\frac{1}{2}\right)$, and so, conditional on Y,

$$
G_{sY} = \frac{(\zeta_{1s}-Y)^2+\zeta_{2s}^2}{1/2} = 2\left\{(\zeta_{1s}-Y)^2+\zeta_{2s}^2\right\} \equiv \chi_2^2(\delta).
$$

Thus, conditional on Y, the family $\{G_{sY}\}_1^m$ are independent and identically distributed noncentral chi-squared variates with two degrees of freedom and noncentrality parameter δ where

$$
\delta = \left(\frac{-Y}{1/\sqrt{2}}\right)^2 = 2Y^2.
$$

It follows from Lemma B and (21) that

$$
\begin{aligned}
E(\log G_{sY}|Y) &= E\left(\log\left(\chi_2^2(\delta)\right)|Y\right) = e^{-\delta/2}\sum_{j=0}^{\infty}\frac{\left(\frac{\delta}{2}\right)^j}{j!}\psi(1+j)+\log 2 \\
&= e^{-\delta/2}\sum_{j=1}^{\infty}\frac{\left(\frac{\delta}{2}\right)^j}{j!}\left(\sum_{k=1}^{j}\frac{1}{k}-C\right) - e^{-\delta/2}C+\log 2 \\
&= e^{-\delta/2}\sum_{j=1}^{\infty}\frac{\left(\frac{\delta}{2}\right)^j}{j!}\sum_{k=1}^{j}\frac{1}{k} - Ce^{-\delta/2} - C + \log 2 \\
&= \mu_Y, \text{ say.}
\end{aligned}
$$

Further,

$$
\begin{aligned}
\operatorname{Var}(\log G_{sY}|Y) &= \operatorname{Var}\left(\left(\log\left(\chi_2^2(\delta)\right)|Y\right)\right) \\
&= e^{-\delta/2}\sum_{j=0}^{\infty}\frac{\left(\frac{\delta}{2}\right)^j}{j!}\left\{\psi(1+j)^2+\psi'(1+j)\right\} \\
&\quad - \left[e^{-\delta/2}\sum_{j=0}^{\infty}\frac{\left(\frac{\delta}{2}\right)^j}{j!}\psi(1+j)\right]^2 \\
&= \sigma_Y^2, \text{ say.}
\end{aligned}
$$

Thus, conditional on Y, $\log G_{sY}$ is *iid* $\left(\mu_Y,\sigma_Y^2\right)$. It follows that, conditional on Y,

$$
\frac{1}{\sqrt{m}}\sum_{j=1}^{m}x_s\log G_{sY} = \frac{1}{\sqrt{m}}\sum_{j=1}^{m}x_s\left(\log G_{sY}-\mu_Y\right)
$$

satisfies the Lindeberg-Feller central limit theorem (c.f. Robinson, 1995, p. 1070) and we have

$$\frac{1}{\sqrt{m}}\sum_{j=1}^{m}x_s\log G_{sY}\bigg|_Y \xrightarrow{d} N\left(0,\sigma_Y^2\left(\lim_{m\to\infty}\frac{1}{m}\sum_{j=1}^{m}x_s^2\right)\right) = N\left(0,\sigma_Y^2\left(\lim_{m\to\infty}\frac{1}{m}\sum_{j=1}^{m}x_s^2\right)\right) = N\left(0,\sigma_Y^2\right),$$

since

$$\lim_{m\to\infty}\frac{1}{m}\sum_{j=1}^{m}x_s^2=\int_0^1(\log x)^2\,dx-\left(\int_0^1(\log x)\,dx\right)^2=1.$$

Unconditionally, we therefore have

$$\frac{1}{\sqrt{m}}\sum_{j=1}^{m}x_s\log G_{sY}\xrightarrow{d} MN\left(0,\sigma_Y^2\right) = \int_{-\infty}^{\infty}N\left(0,\sigma_Y^2\right)\operatorname{pdf}(Y)\,dY = \int_0^{\infty}N\left(0,\sigma^2(W)\right)\operatorname{pdf}(W)\,dW, \tag{27}$$

where

$$\sigma^2(W) = e^{-W}\sum_{j=0}^{\infty}\frac{W^j}{j!}\left\{\psi(1+j)^2+\psi'(1+j)\right\} - \left[e^{-W}\sum_{j=0}^{\infty}\frac{W^j}{j!}\psi(1+j)\right]^2,$$

and

$$W=Y^2=\chi_1^2.$$

It follows from (38), (26) and (27) that

$$\frac{1}{\sqrt{m}}\sum_{j=1}^{m}x_s\log|\xi_s-\eta|^2=\frac{1}{\sqrt{m}}\sum_{j=1}^{m}x_s\log G_{sY}\xrightarrow{d}\int_0^{\infty}N\left(0,\sigma^2(W)\right)\operatorname{pdf}(W)\,dW,$$

as stated. ■

6.7 Lemma D *Let $S_k=\sum_{j=1}^{k}u_j$ for $k\geq 1$, and $S_0=0$, for $k=0$, where u_j satisfies Assumption A but with $E|\varepsilon_t|^q<\infty$ for some $q>2p>4$. Then, the probability space on which the u_j and S_k are defined can be expanded in such a way that there is a process distributionally equivalent to S_k and a Brownian motion $B(\cdot)$ with variance $2\pi f_u(0)$ on the new space for which*

$$\sup_{0\leq k\leq n}\left|\frac{S_k}{\sqrt{n}}-B(\frac{k}{n})\right|=o_{a.s}(\frac{1}{n^{\frac{1}{2}-\frac{1}{p}}}). \tag{28}$$

6.8 Proof In view of Assumption A, we may use the BN decomposition (see Phillips and Solo, 1992) to write

$$C\left(L\right)=C\left(1\right)+\widetilde{C}\left(L\right)\left(L-1\right)$$

where $\widetilde{C}\left(L\right)=\sum_{j=0}^{\infty}\widetilde{c}_jL^j$ with $\widetilde{c}_j=\sum_{s=j+1}^{\infty}c_s$ and $\sum_{j=0}^{\infty}\left|\widetilde{c}_j\right|<\infty$. Then,

$$u_t=C\left(1\right)\varepsilon_t+\widetilde{\varepsilon}_{t-1}-\widetilde{\varepsilon}_t,$$

with $\widetilde{\varepsilon}_t=\widetilde{C}\left(L\right)\varepsilon_t$, and

$$S_t=C\left(1\right)\sum_{j=1}^{t}\varepsilon_j+\widetilde{\varepsilon}_0-\widetilde{\varepsilon}_t=S_{\eta t}+\widetilde{\varepsilon}_0-\widetilde{\varepsilon}_t,$$

where $S_{\eta t}=\sum_{j=1}^{t}\eta_j$ and $\eta_j=C\left(1\right)\varepsilon_j$. Next, since η_j is *iid* with mean zero and finite moments of order $q>2p>4$, we may use a strong approximation to the partial sum process $S_{\eta t}$ of η_j. In particular, by a result of Komlós, Major and Tusnády (see Csörgő and Horvath, 1993, p.18) we can expand the probability space as necessary to set up a partial sum process that is distributionally equivalent to $S_{\eta k}$ and a Brownian motion $B(\cdot)$ with variance $2\pi f_u(0)$ on the same space for which

$$\sup_{0\leq k\leq n}\left|S_{\eta k}-B(k)\right|=o_{a.s}(n^{\frac{1}{q}}),\tag{29}$$

giving a uniform approximation to $S_{\eta k}$ over $0\leq k\leq n$ in terms of the Brownian motion B. Next, since

$$\left|S_k-B(k)\right|\leq\left|S_{\eta k}-B(k)\right|+\left|\widetilde{\varepsilon}_0-\widetilde{\varepsilon}_k\right|$$

we have

$$\sup_{0\leq k\leq n}\left|\frac{S_k}{\sqrt{n}}-B(\frac{k}{n})\right|\leq\sup_{0\leq k\leq n}\left|\frac{S_{\eta k}}{\sqrt{n}}-B(\frac{k}{n})\right|+2\sup_{0\leq k\leq n}\frac{\left|\widetilde{\varepsilon}_k\right|}{\sqrt{n}}.\tag{30}$$

Now

$$\sup_{0\leq k\leq n}\frac{\left|\widetilde{\varepsilon}_k\right|}{\sqrt{n}}=o_{a.s.}\left(\frac{1}{n^{\frac{1}{2}-\frac{1}{p}}}\right)\tag{31}$$

holds if

$$\max_{0\leq k\leq n}\frac{\left|\widetilde{\varepsilon}_k\right|}{n^{\frac{1}{p}}}=o_{a.s.}\left(1\right).\tag{32}$$

But

$$\begin{aligned}
P\left[\max_{0\leq k\leq n}\frac{\left|\widetilde{\varepsilon}_k\right|}{n^{\frac{1}{p}}}>\delta\right] &= P\left[\max_{0\leq k\leq n}\frac{\left|\widetilde{\varepsilon}_k\right|^q}{n^{\frac{q}{p}}}>\delta^q\right]\\
&= P\left[\frac{1}{n^{\frac{q}{p}}}\sum_{k=1}^{n}\left|\widetilde{\varepsilon}_k\right|^q\mathbf{1}\left[\left|\widetilde{\varepsilon}_k\right|^q>n^{\frac{q}{p}}\delta^q\right]>\delta^q\right]\\
&< \frac{E\left(\sum_{k=1}^{n}\left|\widetilde{\varepsilon}_k\right|^q\mathbf{1}\left[\left|\widetilde{\varepsilon}_k\right|^q>n^{\frac{q}{p}}\delta^q\right]\right)}{n^{\frac{q}{p}}\delta^q}\\
&= \frac{E\left(\left|\widetilde{\varepsilon}_k\right|^q\mathbf{1}\left[\left|\widetilde{\varepsilon}_k\right|^q>n^{\frac{q}{p}}\delta^q\right]\right)}{n^{\frac{q}{p}-1}\delta^q},
\end{aligned}$$

by stationarity of $\widetilde{\varepsilon}_k$. It follows that

$$\sum_{n=1}^{\infty} P\left[\max_{0\le k\le n} \frac{|\widetilde{\varepsilon}_k|}{n^{\frac{1}{p}}} > \delta\right] < \sum_{n=1}^{\infty} \frac{E\left(|\widetilde{\varepsilon}_k|^q\right)}{n^{\frac{q}{p}-1}} < \infty$$

since $q > 2p$. Result (32) then follows by the Borel Cantelli lemma. We deduce from (29), (30) and (31).that

$$\sup_{0\le k\le n}\left|\frac{S_k}{\sqrt{n}} - B(\frac{k}{n})\right| = o_{a.s.}\left(\frac{1}{n^{\frac{1}{2}-\frac{1}{p}}}\right),$$

giving the stated result. ■

6.9 Remark It is apparent from the above proof that if all we need is

$$\sup_{0\le k\le n}\left|\frac{S_k}{\sqrt{n}} - B(\frac{k}{n})\right| = o_p\left(\frac{1}{n^{\frac{1}{2}-\frac{1}{p}}}\right), \tag{33}$$

then in place of (32) the following is sufficient

$$\max_{0\le k\le n} \frac{|\widetilde{\varepsilon}_k|}{n^{\frac{1}{p}}} = o_p(1). \tag{34}$$

Apparently, (34) holds if

$$\begin{aligned}
P\left[\max_{0\le k\le n} \frac{|\widetilde{\varepsilon}_k|}{n^{\frac{1}{p}}} > \delta\right] &= P\left[\max_{0\le k\le n} \frac{|\widetilde{\varepsilon}_k|^p}{n} > \delta^p\right] \\
&= P\left[\frac{1}{n}\sum_{k=1}^{n} |\widetilde{\varepsilon}_k|^p \mathbf{1}\left[|\widetilde{\varepsilon}_k|^p > n\delta^p\right] > \delta^p\right] \\
&< \frac{E\left(\sum_{k=1}^n |\widetilde{\varepsilon}_k|^p \mathbf{1}\left[|\widetilde{\varepsilon}_k|^p > n\delta^p\right]\right)}{n^{\frac{q}{p}}\delta^q} \\
&= \frac{E\left(|\widetilde{\varepsilon}_k|^p \mathbf{1}\left[|\widetilde{\varepsilon}_k|^p > n\delta^p\right]\right)}{\delta^p} \to 0,
\end{aligned}$$

which will be so when $E\left(|\widetilde{\varepsilon}_k|^p\right) < \infty$. By Minkowski's inequality, we have

$$\begin{aligned}
E\left(|\widetilde{\varepsilon}_k|^p\right) &= E\left(\left|\sum_{j=0}^{\infty} \widetilde{c}_j \varepsilon_{k-j}\right|^p\right) < \left(\sum_{j=0}^{\infty} \left(|\widetilde{c}_j|^p E\left|\varepsilon_{k-j}\right|^p\right)^{\frac{1}{p}}\right)^p \\
&= \left(\sum_{j=0}^{\infty} |\widetilde{c}_j|\right)^p E\left|\varepsilon_k\right|^p < \left(\sum_{j=0}^{\infty} j\left|c_j\right|\right)^p E\left|\varepsilon_k\right|^p.
\end{aligned}$$

Thus, (33) holds if the moment condition in Assumption A is simply $E\left(|\varepsilon_k|^p\right) < \infty$ for any $p > 2$. Stated formally, the result is given as Lemma E below. Akonom (1993, theorem 3) gave a similar result, using a different method of proof.

6.10 Lemma E *Let $S_k = \sum_{j=1}^{k} u_j$ for $k \geq 1$, and $S_0 = 0$, for $k = 0$, where u_j satisfies Assumption A but with $E\left|\varepsilon_t\right|^p < \infty$* for some $p > 2$. *Then, the probability space on which the u_j and S_k are defined can be expanded in such a way that there is a process distributionally equivalent to S_k and a Brownian motion $B(\cdot)$ with variance $2\pi f_u(0)$ on the new space for which*

$$\sup_{0\leq k\leq n}\left|\frac{S_k}{\sqrt{n}} - B(\frac{k}{n})\right| = o_p(\frac{1}{n^{\frac{1}{2}-\frac{1}{p}}}).$$

6.11 Proof of Theorem 3.1 Set $a_s = \log\left|1 - e^{i\lambda_s}\right|$ and $x_s = a_s - \overline{a}$, where $\overline{a} = m^{-1}\sum_{s=1}^{m} a_s$. Then,

$$2\widehat{d} = \frac{\sum_{s=1}^{m} x_s \log\left|w_x\left(\lambda_s\right)\right|^2}{\sum_{s=1}^{m} x_s^2},$$

and, from (12),

$$\log\left|w_x\left(\lambda_s\right)\right|^2 = -2\log\left|1 - e^{i\lambda_s}\right| + \log\left|w_u\left(\lambda_s\right) - e^{i\lambda_s}\frac{X_n}{\sqrt{2\pi n}}\right|^2$$

so that

$$2\sqrt{m}\left(\widehat{d} - 1\right) = \frac{\frac{1}{\sqrt{m}}\sum_{s=1}^{m} x_s \log\left|w_u\left(\lambda_s\right) - \frac{1}{\sqrt{2\pi n}} e^{i\lambda_s} X_n\right|^2}{\frac{1}{m}\sum_{s=1}^{m} x_s^2}. \tag{35}$$

Next, we proceed to find a more convenient representation for $w_u\left(\lambda_s\right)$ and $X_n/\sqrt{n}$ in (35). Using partial summation, we write for $s = 1, ..., m$

$$\begin{aligned}
w_u\left(\lambda_s\right) &= \frac{1}{\sqrt{2\pi n}}\sum_{t=1}^{n} e^{\frac{2\pi i s t}{n}} u_t = \Delta\left(\frac{1}{\sqrt{2\pi n}}\sum_{t=1}^{n} e^{\frac{2\pi i s t}{n}} S_t\right) - \frac{1}{\sqrt{2\pi n}}\sum_{t=1}^{n} S_{t-1}\Delta\left(e^{\frac{2\pi i s t}{n}}\right) \\
&= \frac{1}{\sqrt{2\pi n}} S_n - \frac{1}{\sqrt{2\pi n}}\sum_{t=1}^{n} S_{t-1} e^{\frac{2\pi i s(t-1)}{n}}\left(e^{\frac{2\pi i s}{n}} - 1\right) \\
&= \frac{1}{\sqrt{2\pi}}\frac{S_n}{\sqrt{n}} - \frac{1}{\sqrt{2\pi}}\sum_{t=1}^{n}\frac{S_{t-1}}{\sqrt{n}} e^{\frac{2\pi i s(t-1)}{n}}\left(e^{\frac{2\pi i s}{n}} - 1\right).
\end{aligned}$$

In view of the embedding (14)

$$\frac{S_{t-1}}{\sqrt{n}} = B\left(\frac{t-1}{n}\right) + o_{a.s}\left(\frac{1}{n^{\frac{1}{2}-\frac{1}{p}}}\right),$$

where the error magnitude holds uniformly in $t = 1, .., n$. Then, we have

$$\begin{aligned}
w_u\left(\lambda_s\right) &= \frac{1}{\sqrt{2\pi}}\left[B\left(1\right) + o_{a.s}\left(\frac{1}{n^{\frac{1}{2}-\frac{1}{p}}}\right)\right] - \frac{1}{\sqrt{2\pi}}\sum_{t=1}^{n}\left[B\left(\frac{t-1}{n}\right) + o_{a.s}\left(\frac{1}{n^{\frac{1}{2}-\frac{1}{p}}}\right)\right] e^{\frac{2\pi i s(t-1)}{n}}\left(e^{\frac{2\pi i s}{n}} - 1\right) \\
&= \frac{1}{\sqrt{2\pi}} B\left(1\right) - \frac{1}{\sqrt{2\pi}}\sum_{t=1}^{n}\left[B\left(\frac{t-1}{n}\right) + o_{a.s}\left(\frac{1}{n^{\frac{1}{2}-\frac{1}{p}}}\right)\right] e^{\frac{2\pi i s(t-1)}{n}}\left(e^{\frac{2\pi i s}{n}} - 1\right) + o_{a.s}\left(\frac{1}{n^{\frac{1}{2}-\frac{1}{p}}}\right) \\
&= \frac{1}{\sqrt{2\pi}} B\left(1\right) - \frac{2\pi s i}{\sqrt{2\pi}}\int_0^1\left[B\left(r\right) + o_{a.s}\left(\frac{1}{n^{\frac{1}{2}-\frac{1}{p}}}\right)\right] e^{2\pi i s r} dr\left(1 + O\left(\frac{m}{n}\right)\right) + o_{a.s}\left(\frac{1}{n^{\frac{1}{2}-\frac{1}{p}}}\right)
\end{aligned}$$

$$= \frac{1}{\sqrt{2\pi}} B(1) - \frac{2\pi s i}{\sqrt{2\pi}} \int_0^1 B(r) e^{2\pi i s r} dr + O_{a.s}\left(\frac{m}{n}\right) + o_{a.s}\left(\frac{1}{n^{\frac{1}{2}-\frac{1}{p}}}\right)$$

$$= \frac{1}{\sqrt{2\pi}} \int_0^1 e^{2\pi i s r} dB(r) + O_{a.s}\left(\frac{m}{n}\right) + o_{a.s}\left(\frac{1}{n^{\frac{1}{2}-\frac{1}{p}}}\right), \tag{36}$$

where the error magnitude holds uniformly in $s \leq m$. When $s = 0$, we have in the same way

$$\frac{X_n}{\sqrt{2\pi n}} = \frac{1}{\sqrt{2\pi}} B(1) + o_{a.s}\left(\frac{1}{n^{\frac{1}{2}-\frac{1}{p}}}\right). \tag{37}$$

Now use the notation

$$\xi_s = \frac{1}{\sqrt{2\pi}} \int_0^1 e^{2\pi i s r} dB(r), \quad \eta = \frac{1}{\sqrt{2\pi}} B(1).$$

The variates $\{\xi_s\}_{s=1}^m$ are independent complex Gaussian $N_c(0, f_u(0))$ and are independent of η, which is real Gaussian $N(0, f_u(0))$. Writing

$$w_u(\lambda_s) = \xi_s + O_{a.s}\left(\frac{m}{n}\right) + o_{a.s}\left(\frac{1}{n^{\frac{1}{2}-\frac{1}{p}}}\right),$$

and

$$\frac{X_n}{\sqrt{2\pi n}} = \eta + o_{a.s}\left(\frac{1}{n^{\frac{1}{2}-\frac{1}{p}}}\right),$$

and using the fact that $\sum_{s=1}^m |x_s| = O(m)$ (c.f. Robinson, 1995, p. 1067), we have

$$\frac{1}{\sqrt{m}} \sum_{s=1}^m x_s \log \left| w_u(\lambda_s) - \frac{1}{\sqrt{2\pi n}} e^{i\lambda_s} X_n \right|^2$$

$$= \frac{1}{\sqrt{m}} \sum_{s=1}^m x_s \log \left| \xi_s - \eta + O_{a.s}\left(\frac{m}{n}\right) + o_{a.s}\left(\frac{1}{n^{\frac{1}{2}-\frac{1}{p}}}\right) \right|^2$$

$$= \frac{1}{\sqrt{m}} \sum_{s=1}^m x_s \left[\log |\xi_s - \eta|^2 + \log \left| 1 + O_{a.s}\left(\frac{m}{n}\right) + o_{a.s}\left(\frac{1}{n^{\frac{1}{2}-\frac{1}{p}}}\right) \right|^2 \right]$$

$$= \frac{1}{\sqrt{m}} \sum_{s=1}^m x_s \log |\xi_s - \eta|^2 + O\left(\frac{1}{\sqrt{m}} \sum_{s=1}^m |x_s|\right) \left[O_{a.s}\left(\frac{m}{n}\right) + o_{a.s}\left(\frac{1}{n^{\frac{1}{2}-\frac{1}{p}}}\right) \right]$$

$$= \frac{1}{\sqrt{m}} \sum_{s=1}^m x_s \log |\xi_s - \eta|^2 + O(\sqrt{m}) \left[O_{a.s}\left(\frac{m}{n}\right) + o_{a.s}\left(\frac{1}{n^{\frac{1}{2}-\frac{1}{p}}}\right)^2 \right]$$

$$= \frac{1}{\sqrt{m}} \sum_{s=1}^m x_s \log |\xi_s - \eta|^2 + O_{a.s}\left(\frac{m^{\frac{3}{2}}}{n}\right) + o_{a.s}\left(\frac{\sqrt{m}}{n^{\frac{1}{2}-\frac{1}{p}}}\right)$$

$$= \frac{1}{\sqrt{m}} \sum_{s=1}^m x_s \log |\xi_s - \eta|^2 + o_{a.s.}(1) \tag{38}$$

when

$$\frac{m^{\frac{3}{2}}}{n} + \frac{\sqrt{m}}{n^{\frac{1}{2}-\frac{1}{p}}} \to 0. \tag{39}$$

Then, from (38) and Lemma C we deduce that

$$\frac{1}{\sqrt{m}}\sum_{s=1}^{m} x_s \log\left|w_u(\lambda_s) - \frac{1}{\sqrt{2\pi n}} e^{i\lambda_s} X_n\right|^2 \to_d \int_0^\infty N\left(0, \sigma^2(W)\right) \text{pdf}(W)\, dW. \tag{40}$$

Finally, using (40) and $m^{-1}\sum_{s=1}^{m} x_s^2 \to 1$ in (35), we obtain

$$\sqrt{m}\left(\widehat{d} - 1\right) \to_d MN\left(0, \frac{1}{4}\sigma^2(W)\right) \equiv \int_0^\infty N\left(0, \frac{1}{4}\sigma^2(W)\right) \text{pdf}(W)\, dW,$$

giving the required result. ■

6.12 Proof of Theorem 4.1 From (24) we have

$$\widetilde{d} = \frac{1}{2}\frac{\sum_{s=1}^{m} x_s \log I_v(\lambda_s)}{\sum_{s=1}^{m} x_s^2},$$

and so

$$2\sqrt{m}\left(\widetilde{d} - 1\right) = \frac{\frac{1}{\sqrt{m}}\sum_{s=1}^{m} x_s \log|w_u(\lambda_s)|^2}{\frac{1}{m}\sum_{s=1}^{m} x_s^2}. \tag{41}$$

Proceeding as in the proof of theorem 3.1, we have

$$\frac{1}{\sqrt{m}}\sum_{s=1}^{m} x_s \log|w_u(\lambda_s)|^2 = \frac{1}{\sqrt{m}}\sum_{s=1}^{m} x_s \log|\xi_s|^2 + o_{a.s.}(1), \tag{42}$$

under (39), and then, just as in the proof of Lemma C but with $Y = 0$, we get

$$\frac{1}{\sqrt{m}}\sum_{s=1}^{m} x_s \log|\xi_s|^2 \to_d N\left(0, \frac{\pi^2}{6}\right). \tag{43}$$

Note that this case follows directly from Lemma C when $W = 0$ *a.s.* and then the limit distribution is normal with variance

$$\sigma^2(0) = \psi'(1) = \frac{\pi^2}{6}.$$

from (20) or Gradshteyn and Ryzhik (1965, 8.366). Combining (42) and (43) with $m^{-1}\sum_{s=1}^{m} x_s^2 \to 1$ in (41), we obtain

$$2\sqrt{m}\left(\widetilde{d} - 1\right) \to_d N\left(0, \frac{\pi^2}{6}\right),$$

giving the stated result. ■

References

[1] Akonom, J. (1993). "Comportement asymptotique du temps d'occupation du processus des sommes partielles." *Annals of the Institute of Henri Poincaré*, 29, 57–81.

[2] Corbae, D., S. Ouliaris and P. C. B. Phillips (1999). "Band Spectral Regression with Trending Data", Yale University, mimeographed.

[3] Csörgő, M. And L. Horváth (1993). *Weighted Approximations in Probability and Statistics.* New York: Wiley.

[4] Geweke, J. and S. Porter-Hudak (1983). "The Estimation and Application of Long Memory Time Series Models", *Journal of Time Series Analysis* 4, 221-237.

[5] Gil-Alaña, L. A. and P. M. Robinson (1997). "Testing of Unit Root and Other Nonstationary Hypothesis in Macroeconomic Time Series", *Journal of Econometrics* 80, 241-268.

[6] Gradshteyn I. S. and I. M. Ryzhik (1965). *Tables of Integrals, Series and Products.* New York: Academic Press.

[7] Hurvich, C. M. and B. K. Ray (1995). "Estimation of the Memory Parameter for Nonstationary or Noninvertible Fractionally Integrated Processes", *Journal of Time Series Analysis* 16, 17-42.

[8] Hurvich, C. M. and K. I. Beltrao (1993). "Automatic Semiparametric Estimation of the Memory Parameter of a Long-Memory Time Series", *Journal of Time Series Analysis* 15, 285-302.

[9] Hurvich, C. M., R. Deo and J. Brodsky (1998). "The Mean Squared Error of Geweke and Porter Hudak's Estimator of the Memory Parameter of a Long Memory Time Series", *Journal of Time Series Analysis* 19, 19-46.

[10] Kim, C. S. and P. C. B. Phillips (1999a). "Log Periodogram Regression: the Nonstationary Case". Yale University, mimeographed.

[11] Kim, C. S. and P. C. B. Phillips (1999b). "Modified Log Periodogram Regression". Yale University, mimeographed.

[12] Künsch, H. R. (1986). "Discrimination between long range dependence and monotonic trends". *Journal of Applied Probability*, 23, 1025–1030.

[13] Künsch, H. R. (1987). "Statistical Aspects of Self-Similar Processes", *Proceedings of the First World Congress of the Bernoulli Society* Vol. 1, 67-74, VNU Science Press.

[14] Phillips, P. C. B. (1999). "Discrete Fourier Transforms of Fractional Processes", Yale University, mimeographed.

[15] Phillips, P. C. B. and V. Solo (1992). "Asymptotics for linear processes," *Annals of Statistics* 20, 971–1001.

[16] Robinson, P. M. (1994). "Efficient tests of nonstationary hypotheses," *Journal of the American Statistical Association*, Vol. 89, 1420-1437.

[17] Robinson, P. M. (1995). "Log Periodogram Regression of Time Series with Long Memory Dependence", *The Annals of Statistics* 23, 1048-1072.

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

[19] Shorack, G. R. And J. A. Wellner (1986). *Empirical Processes with Applications to Statistics.* New York: Wiley.

[20] Tanaka, K (1999), "The Nonstationary Fractional Unit Root", *Econometric Theory* 15, 549-582.

[21] Velasco, C. (1999a). "Non-Stationary Log-Periodogram Regression", *Journal of Econometrics* 91, 325-371.

[22] Velasco, C. (1999b). "Gaussian Semiparametric Estimation of Non-Stationary Time Series", *Journal of Time Series Analysis* 20, 87-127.