10-1-2006

Log Periodogram Regression: The Nonstationary Case

Chang Sik Kim

Peter C.B. Phillips

LOG PERIODOGRAM REGRESSION: THE NONSTATIONARY CASE

By

Chang Sik Kim and Peter C.B. Phillips

October 2006

COWLES FOUNDATION DISCUSSION PAPER NO. 1587

COWLES FOUNDATION FOR RESEARCH IN ECONOMICS
YALE UNIVERSITY
Box 208281
New Haven, Connecticut 06520-8281

http://cowles.econ.yale.edu/

Log Periodogram Regression: The Nonstationary Case*

Chang Sik Kim
Department of Economics
Ewha Women's University

and

Peter C. B. Phillips
Cowles Foundation for Research in Economics
Yale University
&
University of York and University of Auckland

September 2006

Abstract

Estimation of the memory parameter (d) is considered for models of nonstationary fractionally integrated time series with $d > \frac{1}{2}$. It is shown that the log periodogram regression estimator of d is inconsistent when $1 < d < 2$ and is consistent when $\frac{1}{2} < d \leq 1$. For $d > 1$, the estimator is shown to converge in probability to unity.

JEL Classification: C22

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

1. Introduction

Statistical inference in models of fractionally integrated time series has been an active field of recent research. Much of the literature has focused on estimating the memory parameter 'd' of a fractionally integrated process X_t satisfying a general model of the form

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

*The original version of this paper was circulated in 1999 and the present version has benefitted from comments of referees and an associate editor. Phillips acknowledges partial research support from the NSF under Grant Nos. SBR 97-30295 and SES 04-142254.

where u_t is stationary with zero mean and continuous spectral density $f_u(\lambda) > 0$, and L is a lag operator for which $Lu_t = u_{t-1}$. A variety of estimation methods of d have been suggested and asymptotic theories for them have recently been developed in the case of stationary long memory time series like (1) with $|d| < \frac{1}{2}$. A commonly used estimator in applied work is the log periodogram estimator, suggested in Geweke and Porter-Hudak (1983), which is appealing because of its nonparametric treatment of u_t and the convenience of linear least squares regression. Under Gaussian assumptions, Robinson (1995a) developed consistency and asymptotic normality results for a log periodogram estimator which trims out some low frequencies periodogram ordinates in the regression, following a suggestion of Künsch (1987). Hurvich, Deo, and Brodsky (1998) derive an expression for the mean squared error of this estimator without omitting low frequencies ordinates, again under Gaussianity, and obtain asymptotic normality results and an optimal choice of the number of periodogram ordinates to include in the regression.

Most of the theory of statistical inference for log periodogram regression has been developed for the stationary long memory case with fractional parameter $-\frac{1}{2} < d < \frac{1}{2}$, however, in practice, log periodogram regression has frequently been applied in apparently nonstationary cases (e.g. Bloomfield, 1991, Agiakloglou et al., 1993); and the importance of nonstationarity in practical work is borne out in many empirical studies including those of Cheung and Lai (1993), Phillips (1998/2005), and Maynard and Phillips (2002). Practically speaking, of course, there is seldom any prior information about the range of d before estimation, so that analysis of log periodogram regression for $d > \frac{1}{2}$ is important from both theoretical and practical points of view. Hurvich and Ray (1995) study the asymptotic behavior of periodogram ordinates of a fractionally integrated process with fractional parameter $d \in [0.5, 1.5)$ and argue that log periodogram regression can be badly biased for nonstationary processes with $d > 1$. These authors also illustrate that the estimator is not invariant to first differencing, a phenomenon that was earlier reported in Agiakloglou et al. (1993). Extending the work of Robinson (1995a), Velasco (1999a) has shown consistency of a log periodogram regression estimator that trims out low frequency ordinates, when $\frac{1}{2} < d < 1$ and under Gaussianity. Velasco (2000) showed the asymptotic normality of non-Gaussian pooled log periodogram regression in the stationary case. Local and exact local Whittle estimation for nonstationary fractional integrated processes has most recently been developed by Phillips and Shimotsu (2004) and Shimotsu and Phillips (2005), the latter showing consistent estimation for all values of d. A related approach based on a piecewise extension of local Whittle estimation to the nonstationary case has been given in Abadir, Distaso and Giraitis (2005).

Some intriguing simulation results were reported in Hurvich and Ray (1995). According to table III in their paper, the log periodogram estimates are very close to unity regardless of the true fractional value of d in a range of values over the interval $(1.0, 1.4)$. A later simulation in Velasco (1999b) revealed an estimated probability density for the log periodogram estimate when $d = 1.8$ that is sharply peaked around unity and has a long tail to the right. These simulations indicate that, in cases where $d > 1$, log periodogram regression generally produces estimates of d that are very close to unity, irrespective of the true value of d when $d \geq 1$. The present paper provides an explanation for this pattern of simulation results. Specifically, we show that log periodogram regression is inconsistent when $d > 1$ and that the probability limit of the estimate is unity for all values of $d \in [1, 2)$. The reason for the inconsistency is that the formulation of the log periodogram regression 'model', which is inspired by the local behavior of the spectrum near the origin, omits terms that become dominant in the nonstationary case when $d > 1$. These results correspond with the inconsistency of local Whittle estimation for $d > 1$ found in Phillips and Shimotsu (2004).

We make use of a representation of the discrete Fourier transform (dft) of a fractionally integrated time series under assumptions on the short memory component u_t that are quite weak. This representation and some related results were originally obtained in unpublished work of Phillips (1999) and are briefly reviewed here. Utilizing the representation of the dft, under some restrictive assumptions on the number of periodogram ordinates in the regression, but with no distributional restrictions, we provide here an inconsistency result for log periodogram regression when $d > 1$, showing that the estimator converges in probability to unity, and we give a consistency result that applies when $\frac{1}{2} < d \leq 1$.

The paper is organized as follows. The following section gives some useful alternate representations of the dft of a fractionally integrated time series. Section 3 contains our main results, and some concluding remarks are made in Section 4. Proofs are given in Section 5.

2. Representation of the DFT of a Fractionally Integrated process

This section briefly reviews some representations of the dft of a fractionally integrated time series obtained in Phillips (1999). These are valid in both the stationary, long memory case and the nonstationary case, and they turn out to be particularly helpful in analyzing regressions in the nonstationary case.

The fractionally integrated process X_t in this paper is defined as in (1), with $u_t = 0$ for all $t \leq 0$. More explicit conditions on u_t $(t > 0)$ are given in the following assumption that applies throughout this paper.

2.1 Assumption (Error Condition) *For all $t > 0$, u_t has Wold representation*

$$u_t = C(L)\varepsilon_t = \sum_{j=0}^{\infty} c_j \varepsilon_{t-j}, \quad \sum_{j=0}^{\infty} j|c_j| < \infty, \quad C(1) \neq 0, \tag{2}$$

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

The linear process error condition in (2) 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. Nonetheless, (2) is restrictive and implies the spectrum of u_t is continuously differentiable for all frequencies, which may be regarded as a stronger than necessary in the context of a narrow band semiparametric regression.

To proceed, we expand the fractional process (1) as

$$X_t = (1-L)^{-d} u_t = \sum_{k=0}^{t} \frac{(d)_k}{k!} u_{t-k}, \quad \text{with } u_t = 0 \text{ for all } t \leq 0 \tag{3}$$

where

$$(d)_k = \frac{\Gamma(d+k)}{\Gamma(d)}$$

is Pochammer's symbol for the forward factorial function. Next, define the operator $D_n(L;d) = \sum_{k=0}^{n} \frac{(-d)_k}{k!} L^k$, and expand $D_n(L;d)$ about $L = e^{i\lambda}$ as in Phillips (1999) as

$$D_n(L;d) = D_n\left(e^{i\lambda};d\right) + \widetilde{D}_{n\lambda}\left(e^{-i\lambda}L;d\right)\left(e^{-i\lambda}L - 1\right)$$

where $\widetilde{D}_{n\lambda}\left(e^{-i\lambda}L;d\right) = \sum_{p=0}^{n-1} \widetilde{d}_{\lambda p} e^{-ip\lambda} L^p$ and $\widetilde{d}_{\lambda p} = \sum_{k=p+1}^{n} \frac{(-d)_k}{k!} e^{ik\lambda}$. Writing the dft of X_t as $w_x(\lambda) = \frac{1}{\sqrt{2\pi n}} \sum_{t=1}^{n} X_t e^{it\lambda}$, the result below gives an exact representation of $w_x(\lambda)$ in terms of the dft, $w_u(\lambda)$, of the error process u_t.

2.2 Lemma (Phillips, 1999)

$$w_u(\lambda) = w_x(\lambda) D_n\left(e^{i\lambda};d\right) + \frac{1}{\sqrt{2\pi n}}\left(\widetilde{X}_{\lambda 0}(d) - e^{in\lambda}\widetilde{X}_{\lambda n}(d)\right) \tag{4}$$

where

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

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, expression in (4) becomes

$$\begin{aligned} w_u(\lambda) &= w_x(\lambda) D_n\left(e^{i\lambda}; d\right) - \frac{e^{in\lambda}}{\sqrt{2\pi n}} \widetilde{D}_{n\lambda}\left(e^{-i\lambda}L; d\right) X_n \\ &= w_x(\lambda) D_n\left(e^{i\lambda}; d\right) - \frac{1}{\sqrt{2\pi n}} e^{in\lambda} \widetilde{X}_{\lambda n}(d). \end{aligned} \tag{5}$$

Equation (5) shows that the exact relation between $w_x(\lambda)$ and $w_u(\lambda)$ involves a correction term that depends on $\widetilde{X}_{\lambda n}(d)$. This term therefore needs to be considered in studying the asymptotic behavior of $w_x(\lambda)$, and any function of it, like the log periodogram regression estimator. The asymptotic behavior of $\widetilde{X}_{\lambda n}(d)$ at the fundamental frequencies λ_s is given in lemma 3.1 of Phillips (1999) and is shown to be sensitive to the value of s in $\lambda_s = \frac{2\pi s}{n}$. Here, our main focus of interest is to develop the behavior of $\widetilde{X}_{\lambda n}(d)$ when $\lambda = \lambda_s = \frac{2\pi s}{n} \to 0$, the situation we allow for in log periodogram regression. The following lemma is based on theorem 3.2 of Phillips (1999), extends that theorem to the case where $d > 1$, and shows the asymptotic form of $\widetilde{X}_{\lambda n}(d)$ when X_t is a nonstationary fractional process.

2.3 Lemma *For $\lambda = \lambda_s = \frac{2\pi s}{n} \to 0$ as $n \to \infty$, and for all $s \leq m$, $m = o(n)$,*

(a) *When $\frac{1}{2} < d < 1$ and $\frac{n^\alpha}{s} \to 0$ for some $\alpha \in (\frac{1}{2}, 1)$, or $1 < d < 2$, then*

$$\frac{1}{n^d} \frac{\widetilde{X}_{\lambda_s n}(d)}{\sqrt{n}} = -\frac{1}{n^d} \frac{e^{i\lambda_s}}{(1 - e^{i\lambda_s})^{1-d}} \frac{X_n}{\sqrt{n}} + o_p\left(\frac{1}{n^d} \frac{e^{i\lambda_s}}{(1 - e^{i\lambda_s})^{1-d}} \frac{X_n}{\sqrt{n}}\right) \tag{6}$$

(b) *If X_t follows (1) with $d = 1$,* then

$$\frac{1}{n} \frac{\widetilde{X}_{\lambda_s n}(d)}{\sqrt{n}} = -\frac{e^{i\lambda_s}}{n} \frac{X_n}{\sqrt{n}}. \tag{7}$$

This result shows that the same formulae as those given in Phillips (1999) for the case $d \in (\frac{1}{2}, 1)$ also apply when $d > 1$. The formula (7) is particularly simple for $d = 1$, and the restriction on the range of s in case of $\frac{1}{2} < d < 1$ is relaxed for $d > 1$. As might be expected, the leading term in the asymptotic approximation of $\frac{1}{n^d} \frac{\widetilde{X}_{\lambda_s n}(d)}{\sqrt{n}}$ is the same as when $d \in (\frac{1}{2}, 1)$. When $d = \frac{3}{2}$, the formula (6) is still valid because we assume $u_t = 0,\ t \leq 0$. As shown in Liu (1998), if we allow for prehistorical influence in the fractional process X_t, then the order of X_n when $d = \frac{3}{2}$ is $n^{d-\frac{1}{2}}\sqrt{\ln n}$, i.e., $n^{\frac{1}{2}-d}(\ln n)^{-\frac{1}{2}} X_n = O_p(1)$, whereas $n^{\frac{1}{2}-d} X_n = O_p(1)$ for $d \in (\frac{1}{2}, \frac{3}{2})$. In that case, a minor change in part (a) of the lemma 2.3 is needed to incorporate the effect of the slowly varying factor $\sqrt{\ln n}$. However, if $u_t = 0,\ t \leq 0$, then we have the MA representation as in (3) above, and the order of magnitude of X_t is $\frac{1}{n}$, as is easily determined (c.f., Gourieroux, Maurel, and Monfort, 1987).

Lemma 2.4 below gives an asymptotic expression for the dft $w_x(\lambda)$ in term of $w_u(\lambda)$ and X_n. Our main interest is in the case where $\lambda = \lambda_s = \frac{2\pi s}{n} \to 0$. From (5)

$$w_x(\lambda_s) = D_n\left(e^{i\lambda}; d\right)^{-1}\left[w_u(\lambda_s) + \frac{1}{\sqrt{2\pi n}}\widetilde{X}_{\lambda n}(d)\right], \tag{8}$$

and using lemma 2.3 and an expansion for the sinusoidal polynomial $D_n\left(e^{i\lambda}; d\right)$ given in lemma 3.1 of Phillips (1999), we have the following asymptotic representation for $w_x(\lambda_s)$.

2.4 Lemma *For $\lambda = \lambda_s = \frac{2\pi s}{n} \to 0$ as $n \to \infty$, and for all $s \leq m$, $m = o(n)$, and when $\frac{1}{2} < d < 1$ and $\frac{n^\alpha}{s} \to 0$ for some $\alpha \in (\frac{1}{2}, 1)$ or $1 < d < 2$, then*

$$\frac{1}{n^d}w_x(\lambda_s) = \left(1 - e^{i\lambda_s}\right)^{-d}\frac{1}{n^d}w_u(\lambda_s) - \frac{1}{\sqrt{2\pi}}\frac{1}{n}\frac{e^{i\lambda_s}}{(1 - e^{i\lambda_s})}\frac{X_n}{n^{d-\frac{1}{2}}} + o_p\left(\frac{1}{\sqrt{2\pi}}\frac{1}{n}\frac{e^{i\lambda_s}}{(1 - e^{i\lambda_s})}\frac{X_n}{n^{d-\frac{1}{2}}}\right) \tag{9}$$

The representation (9) facilitates the development of an asymptotic theory for fractionally integrated time series. It shows that the dft of a fractionally integrated process is composed of two separate components. The first of these involves the dft of the innovation u_t, the second involves the value of the final sample observation X_n. The limit behavior of these two components is already known. Hence, we can develop dft asymptotics for fractionally integrated processes by analyzing these two terms, rather than by attempting to work directly with the dft of X_t itself. A second advantage of the new representation is that it follows by algebraic simplification and does not depend upon distributional specifications like Gaussianity. All that is needed to obtain (9) is the general linear process formulation (2). A third advantage is that the representation in lemma 2.2 holds for all frequencies $\lambda_s = \frac{2\pi s}{n}$, $s = 0, 1, ..., n-1$, making it helpful in the asymptotic analysis of a wide variety of quantities that arise in the study of fractional processes. The asymptotic representations in lemma 2.3 and 2.4 hold for the frequencies near the origin, which is enough for most semiparametric analyses of fractionally integrated processes. We can, in fact, go further and develop asymptotic forms for the dft of X_t when s is fixed, and when $s \to \infty$ and $\lambda_s \to \phi \neq 0$, as well as when $\lambda_s \to 0$ as $n \to \infty$. These forms are given in Phillips (1999) for $d \in (\frac{1}{2}, 1)$. However, only those where $\lambda_s \to 0$ as $n \to \infty$, as in lemma 2.4, are needed in the present paper.

3. Log-Periodogram Regression: the nonstationary case

(a) Inconsistency over $1 < d < 2$

Start by writing the normalized dft of X_t according to the lemma 2.4 as

$$\frac{1}{n^d}w_x(\lambda_s) = \left(1-e^{i\lambda_s}\right)^{-d}\frac{1}{n^d}w_u(\lambda_s) - \frac{1}{\sqrt{2\pi}}\frac{1}{n}\frac{e^{i\lambda_s}}{(1-e^{i\lambda_s})}\frac{X_n}{n^{d-\frac{1}{2}}} + o_p\left(\frac{1}{\sqrt{2\pi}}\frac{1}{n}\frac{e^{i\lambda_s}}{(1-e^{i\lambda_s})}\frac{X_n}{n^{d-\frac{1}{2}}}\right) \tag{10}$$

Then, the normalized periodogram of X_t can be written as

$$\frac{I_x(\lambda_s)}{n^{2d}} = \left|\left(1-e^{i\lambda_s}\right)^{-d}\frac{1}{n^d}w_u(\lambda_s) - \frac{1}{\sqrt{2\pi}}\frac{1}{n}\frac{e^{i\lambda_s}}{(1-e^{i\lambda_s})}\frac{X_n}{n^{d-\frac{1}{2}}} + o_p\left(\frac{1}{\sqrt{2\pi}}\frac{1}{n}\frac{e^{i\lambda_s}}{(1-e^{i\lambda_s})}\frac{X_n}{n^{d-\frac{1}{2}}}\right)\right|^2 \tag{11}$$

Here, the dominant term is $\frac{1}{\sqrt{2\pi}}\frac{1}{n}\frac{e^{i\lambda_s}}{(1-e^{i\lambda_s})}\frac{X_n}{n^{d-\frac{1}{2}}}$ as we will show in the appendix, and all the other terms are of lesser order. Therefore, (11) can be put in the form

$$\frac{I_x(\lambda_s)}{n^{2d}} = \left(\frac{1}{2\pi}\right)\left(\frac{\left|e^{i\lambda_s}\right|^2}{\left|1-e^{i\lambda_s}\right|^2}\right)\left(\frac{1}{n}\right)^2\left(\frac{X_n}{n^{d-\frac{1}{2}}}\right)^2\left|1+\zeta_{ns}\right|^2, \tag{12}$$

where

$$\begin{aligned}
\zeta_{ns} &= \left((2\pi)\frac{\left(1-e^{i\lambda_s}\right)^{1-d}}{e^{i\lambda_s}}n^{1-d}\frac{w_u(\lambda_s)}{\left(\frac{X_n}{n^{d-\frac{1}{2}}}\right)}\right) + o_p(1) \\
&= 2\pi\xi_{ns} + o_p(1),
\end{aligned}$$

and $o_p(1)$ is asymptotically negligible uniformly in s.

$$\xi_{ns} = \frac{\left(1-e^{i\lambda_s}\right)^{1-d}}{e^{i\lambda_s}}n^{1-d}\frac{w_u(\lambda_s)}{\left(\frac{X_n}{n^{d-\frac{1}{2}}}\right)}.$$

Rewriting (12) as

$$I_x(\lambda_s) = \left(\frac{1}{2\pi}\right)\left(\frac{1}{\left|1-e^{i\lambda_s}\right|^2}\right)\left(\frac{n^d}{n}\right)^2\left(\frac{X_n}{n^{d-\frac{1}{2}}}\right)^2\left|1+\zeta_{ns}\right|^2,$$

we obtain the following representation of the logarithm of the periodogram:

$$\ln(I_x(\lambda_s)) = \ln\left(\frac{1}{2\pi}\right) + 2\ln\left(\frac{n^d}{n}\right) - 2\ln\left(\left|1-e^{i\lambda_s}\right|\right) + 2\ln\left(\left|\frac{X_n}{n^{d-\frac{1}{2}}}\right|\right) + 2\ln\left|1+\zeta_{ns}\right|. \tag{13}$$

The log periodogram regression estimator of the memory parameter d is based on linear least squares regression of $\log I_x(\lambda_s)$ on $\log\lambda_s$ over frequencies $\{\lambda_s, s=1,...,m\}$. It has the explicit form

$$2\widehat{d} = -\left[\sum_{s=1}^{m} x_s^2\right]^{-1}\left[\sum_{s=1}^{m} x_s \log I_x\left(\lambda_s\right)\right], \tag{14}$$

where $x_s = \log\left(\left|1-e^{i\lambda_s}\right|\right) - \overline{\log\left(\left|1-e^{i\lambda_s}\right|\right)}$, and $\overline{\log\left(\left|1-e^{i\lambda_s}\right|\right)} = \frac{1}{m}\sum_{s=1}^{m}\log\left(\left|1-e^{i\lambda_s}\right|\right)$. From (13) we deduce that

$$2(\widehat{d}-1) = -2\left[\sum_{s=1}^{m} x_s^2\right]^{-1}\left[\sum_{s=1}^{m} x_s \log\left|1+\zeta_{ns}\right|\right], \tag{15}$$

where $v_n = \log\left(\left|\frac{X_n}{n^{d-\frac{1}{2}}}\right|\right)$, and noting that $\sum_{s=1}^{m} x_s = 0$.

The following result gives the inconsistency of $\widehat{d}$ when $1 < d < 2$.

3.1 Theorem *For $1 < d < 2$, if $\frac{m}{n} + \frac{m\log m}{n^{\frac{1}{2}-\frac{1}{p}}} \to 0$, then $\widehat{d} \xrightarrow{p} 1$.*

The conditions on the frequency band $\{\lambda_s,\ 1 < s < m\}$ where $\frac{m\log m}{n^{\frac{1}{2}-\frac{1}{p}}} \to 0$ restrict the range of the effective sample size m (the number of periodogram ordinates) in the regression (14). In particular, the restriction on m is far more restrictive than commonly used rules like $m = O_p(n^{\frac{2}{3}})$, $O_p(n^{\frac{1}{2}})$, or the optimal choice, $m = O_p(n^{\frac{4}{5}})$, suggested by Hurvich, Deo, and Brodsky (1998). This restriction is mainly due to our strong approximation approach without assuming Gaussianity, and to the treatment of reciprocal and logarithmic functions of random variables shown in appendix. Trimming low frequencies (Künsch, 1987; Robinson, 1995a) is not necessary, as shown in Hurvich, Deo, and Brodsky (1998) for stationary fractional processes (see also Velasco, 1999a).

Theorem 3.1 shows that the log periodogram regression estimator is inconsistent and has unity as its limit in probability over the interval $1 < d < 2$. The estimator can therefore be expected to be systematically biased when the true value of d is greater than unity, and severely biased when d is well above unity. This behavior is apparent in the simulation results of Hurvich and Ray (1995) and Velasco (1999b).

We conclude that the estimator $\widehat{d}$ has unity as its probability limit over the whole interval $1 < d < 2$. On the other hand, the log periodogram estimator is consistent over the nonstationary domain $\frac{1}{2} \le d < 1$, as the following section shows.

(b) Consistency over $\frac{1}{2} < d < 1$

From lemma 2.4, the representation of $\frac{w_x(\lambda_s)}{n^d}$ is the same as (9) for the $\frac{1}{2} < d < 1$ case and, therefore, the formula for the normalized periodogram is also the same as (11). Thus,

$$\begin{aligned} \frac{I_x(\lambda_s)}{n^{2d}} &= \left|\left(1-e^{i\lambda_s}\right)^{-d}\frac{1}{n^d}w_u(\lambda_s) - \frac{1}{\sqrt{2\pi}}\frac{1}{n}\frac{e^{i\lambda_s}}{(1-e^{i\lambda_s})}\frac{X_n}{n^{d-\frac{1}{2}}} + o_p\left(\frac{1}{\sqrt{2\pi}}\frac{1}{n}\frac{e^{i\lambda_s}}{(1-e^{i\lambda_s})}\frac{X_n}{n^{d-\frac{1}{2}}}\right)\right|^2 \quad (16) \\ &= \left(\frac{1}{2\pi}\right)\left(\frac{\left|e^{i\lambda_s}\right|^2}{\left|1-e^{i\lambda_s}\right|^2}\right)\left(\frac{1}{n}\right)^2 \left|1+\zeta_{ns}\right|^2, \end{aligned}$$

where the same formula for ζ_{ns} applies. The dominant term in (16) is $\left|\left(1-e^{i\lambda_s}\right)^{-d}\frac{1}{n^d}\right|^{2d}\left|w_u(\lambda_s)\right|^2$ as we will show in the appendix, and it is the other terms that are now of lesser order. Arranging (16) gives

$$\begin{aligned} \frac{I_x(\lambda_s)}{n^{2d}} &= \frac{1}{n^{2d}}\left|1-e^{i\lambda_s}\right|^{-2d}\left(\frac{X_n}{n^{d-\frac{1}{2}}}\right)^2 \left|\frac{w_u(\lambda_s)}{\left(\frac{X_n}{n^{d-\frac{1}{2}}}\right)} + \frac{1}{\sqrt{2\pi}}\frac{n^d}{n}\frac{e^{i\lambda_s}}{\left(1-e^{i\lambda_s}\right)^{1-d}}(1+o_p(1))\right|^2 \\ &= \frac{1}{n^{2d}}\left|1-e^{i\lambda_s}\right|^{-2d}\left(\frac{X_n}{n^{d-\frac{1}{2}}}\right)^2 \left|\xi_{ns} + \frac{1}{\sqrt{2\pi}}\frac{n^d}{n}\frac{e^{i\lambda_s}}{\left(1-e^{i\lambda_s}\right)^{1-d}}(1+o_p(1))\right|^2, \quad (17) \end{aligned}$$

where

$$\xi_{ns} = \frac{w_u(\lambda_s)}{\left(\frac{X_n}{n^{d-\frac{1}{2}}}\right)} \text{ and } \upsilon_{ns} = \frac{1}{\sqrt{2\pi}}\frac{n^d}{n}\frac{e^{i\lambda_s}}{\left(1-e^{i\lambda_s}\right)^{1-d}}(1+o_p(1)).$$

Now, (17) can be written as

$$I_x(\lambda_s) = \left|1-e^{i\lambda_s}\right|^{-2d}\left(\frac{X_n}{n^{d-\frac{1}{2}}}\right)^2 \left|\xi_{ns}+\upsilon_{ns}\right|^2,$$

and the log periodogram regression equation can be formulated as

$$\ln(I_x(\lambda_s)) = -2d\ln\lambda_s + 2\upsilon_n + 2\ln\left|\xi_{ns}+\upsilon_{ns}\right|, \quad (18)$$

where $\upsilon_n = \ln\left[\left|\frac{X_n}{n^{d-\frac{1}{2}}}\right|\right]$. The formulation in (18) holds strictly over frequencies λ_s with $s \geq l$ and $\frac{n^\alpha}{l} \to 0$ for some $\alpha \in (\frac{1}{2}, 1)$. However, it turns out that we can relax this trimming restriction (viz. $s > l$) in our asymptotic development, so that the log periodogram estimator we consider has the usual definition as the linear least squares regression of $\ln I_x(\lambda_s)$ on λ_s over the full set of frequencies $\{\lambda_s,\ s = 1, ..., m\}$. As we show in the appendix, the representation of the logarithm of the periodogram is a little different from (16) over the frequencies $\{\lambda_s,\ s = 1, ..., l\}$. The following result gives the consistency of $\widehat{d}$ over $\frac{1}{2} < d < 1$.

3.2 Theorem *If X_t follows (1) with $\frac{1}{2} < d < 1$, if u_t satisfies (2) and ε_t fulfills a Cramér type condition, i.e.*

$$\exists\,\delta > 0,\ p > 0,\ \textit{such that}\ \forall\,|t| > p\ \ |\mathbf{E}\exp(it\varepsilon_t)| \leq 1 - \delta, \tag{19}$$

and

$$\int |\mathbf{E}\exp(it\varepsilon_t)|^p\,dt < \infty \textit{for some integer } p \geq 1, \tag{20}$$

and if $\frac{m}{n} + \frac{l(\ln n)^2}{m} + \frac{m\log m}{n^{\frac{1}{2}-\delta}} \to 0 \to 0$ *for any* $0 < \delta < \frac{1}{2}$, *then* $\widehat{d} \xrightarrow{p} d$, *where* $\widehat{d}$ *is defined in* (14).

The two additional conditions (19) and (20) on ε_t are needed for the proof of the consistency of the estimator. Neither is very restrictive. Condition (19) is a form of Cramér condition (see, e.g. Bhattacharya and Rao, 1976), and holds for distributions with a non zero absolutely continuous part. Condition (20) ensures that the density of $\sum_{t=1}^{n}\varepsilon_t$ exists whenever $n \geq p$. Some further discussion is given in the appendix.

A related consistency result for log periodogram regression over $\frac{1}{2} < d < 1$ has been established by Velasco (1999a) under different conditions. In that work, the regression estimator trims out low frequency ordinates, the restrictions on the number of ordinates in the regression are a little stronger than those of Robinson (1995a), and Gaussianity is required, as in earlier analysis of log periodogram regression. The results in the present paper rely, in the main, on the representation of lemma 2.4 and are free from specific distributional assumptions whereas u_t is a linear process defined in (2). However, our linear process assumption is somewhat restrictive in the sense that the condition (2) implies the spectrum of u_t is continuously differentiable for all frequencies, whereas much of earlier work has made only local smoothness of the spectrum near origin.

4. Concluding Remarks

This paper has addressed consistency issues for log periodogram regression with nonstationary, fractionally integrated time series, showing that there is a major difference between the two nonstationary cases where $\frac{1}{2} < d < 1$ and $d > 1$. When $d > 1$, the log periodogram regression estimator is inconsistent, converges to unity in probability for all $d \in (1, 2)$, and, as previous simulation experiments have shown, appears to be seriously biased in finite samples. On the other hand, when $\frac{1}{2} < d < 1$, the log periodogram regression estimator is shown to be consistent with no Gaussianity assumption, but with strong restrictions on the number ordinates included in the regression. The case $d = 1$ has been studied by Phillips (2006), and in this case the estimator is consistent and has a mixed normal limit distribution.

In all of these cases, the time series are nonstationary. But, there is an important difference between nonstationary series with $\frac{1}{2} < d < 1$ and those with $d \geq 1$. In particular, when $d < 1$, the constituent innovations in the time series are not persistent, in the sense that the impact of a unit innovation at time t on the process eventually vanishes; whereas for nonstationary processes with $d \geq 1$, the effects of the innovations do not eventually vanish.

From the practical standpoint of empirical research the inconsistency for $d > 1$ has the most important consequences. In practice, we rarely have any prior information about the magnitude of the memory parameter and it is therefore desirable to have procedures of estimation and inference that have satisfactory properties over a range of plausible parameter values. For some series, like prices and monetary aggregates, the range of plausible parameter values seem most likely to include the region $d > 1$. Inference about the memory parameter for the levels of such series using log periodogram regression using conventional log periodogram regression is clearly inappropriate.

The formulae given in Lemma 2.4 reveal the modifications to the log periodogram estimator that are needed to avoid the inconsistency over $d > 1$. In particular, the second term on the right side of the dft representation (9) suggests that we may replace the dft $w_x()$ in log periodogram calculations by the observable quantity

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

which directly eliminates the term that is responsible for the bias, leading to a procedure that is essentially equivalent to first differencing the series, applying log periodogram regression and adding back unity. However, as discussed in Phillips (1999, 2006) consistent semiparametric estimation for all values of d by log periodogram regression without trimming or tapering can be accomplished using an exact log periodogram regression (ELP) procedure, which takes into account the exact form of the dft given in (8). An asymptotic theory for this estimator has not been developed, but the approach is entirely analogous to the exact local Whittle (ELW) estimator studied in Shimotsu and Phillips (2005). Since both estimators involve nonlinear optimization, the ELP procedure loses the advantage of linear regression that makes log periodogram regression appealing in practice, and ELW is in any event likely to be the preferred choice in terms of asymptotic efficiency.

5. Appendix

5.1 Proof of Lemma 2.3 When $\frac{1}{2} < d < 1$, and $d = 1$ the results for parts (a) and (b) are given in Phillips (1999). So we need only show the $d > 1$ case for the proof of part (a).

Start with part (a) when $d > 1$. Following the proof of theorem 3.2 of Phillips (1999), we write $\widetilde{X}_{\lambda_s n}(d)$ as the sum of two components with $p \leq \ell$ and $p > \ell$. The choice of ℓ will be discussed later. We have

$$\widetilde{X}_{\lambda_s n}(d) = \sum_{p=0}^{n-1} \widetilde{d}_{\lambda_s p} e^{-ip\lambda_s} X_{n-p} = \sum_{p=0}^{n-1} \left(\sum_{k=p+1}^{n} \frac{(-d)_k}{k!} e^{ik\lambda_s} \right) e^{-ip\lambda_s} X_{n-p}$$

$$= \sum_{p=0}^{\ell} \left(\sum_{k=p+1}^{n} \frac{(-d)_k}{k!} e^{ik\lambda_s} \right) e^{-ip\lambda_s} X_{n-p} + \sum_{p=\ell+1}^{n-1} \left(\sum_{k=p+1}^{n} \frac{(-d)_k}{k!} e^{ik\lambda_s} \right) e^{-ip\lambda_s} X_{n-p}. \quad (21)$$

Then, as in the proof of theorem 3.2 in Phillips (1999), we get

$$\frac{1}{n^{1-d}} \sum_{p=0}^{\ell} \left(\sum_{k=p+1}^{n} \frac{(-d)_k}{k!} e^{ik\lambda_s} \right) e^{-ip\lambda_s}$$

$$= e^{i\lambda_s} \frac{1}{n^{1-d}} \sum_{p=0}^{\infty} \frac{(-d)_{p+1}}{(p+1)!} \, {}_2F_1 \left(1+p-d, 1; p+2; e^{i\lambda_s} \right) + O\left(\frac{\ell}{ns} \right)$$

$$+ e^{i\lambda_s} \frac{1}{n^{1-d}} \sum_{p=\ell+1}^{\infty} \frac{(-d)_{p+1}}{(p+1)!} \, {}_2F_1 \left(1+p-d, 1; p+2; e^{i\lambda_s} \right)$$

where ${}_2F_1$ denotes the hypergeometric function. Now

$$\sum_{p=0}^{\infty} \frac{(-d)_{p+1}}{(p+1)!} \, {}_2F_1 \left(1+p-d, 1; p+2; e^{i\lambda_s} \right)$$

$$= \sum_{p=0}^{\infty} \frac{(-d)_{p+1}}{(p+1)!} \sum_{k=0}^{\infty} \frac{(1+p-d)_k \, (1)_k}{(k)! \, (p+2)_k} e^{i\lambda_s k}$$

$$= \sum_{p=0}^{\infty} \sum_{k=0}^{\infty} \frac{\Gamma(p+1-d)}{\Gamma(-d)\,\Gamma(p+2)} \frac{\Gamma(k+p+1-d)}{\Gamma(p+1-d)} \frac{\Gamma(p+2)}{\Gamma(k+p+2)} e^{i\lambda_s k}$$

$$= \sum_{p=0}^{\infty} \sum_{k=0}^{\infty} \frac{\Gamma(k+p+1-d)}{\Gamma(-d)\,\Gamma(k+p+2)} e^{i\lambda_s k}.$$

Note that

$$\sum_{p=0}^{\infty} \sum_{k=0}^{\infty} \left| \frac{\Gamma(k+p+1-d)}{\Gamma(-d)\,\Gamma(k+p+2)} e^{i\lambda_s k} \right| < \infty,$$

for $d > 1$, which is a sufficient condition for exchanging summation and convergence of double summation w.r.t. p, k. Following the manipulation in Phillips (1999), we have

$$\sum_{p=0}^{\infty} \frac{(-d)_{p+1}}{(p+1)!} \, {}_2F_1 \left(1+p-d, 1; p+2; e^{i\lambda_s} \right) = -\frac{e^{i\lambda_s}}{\left(1-e^{i\lambda_s}\right)^{1-d}},$$

which is finite for all values of λ_s for $d > 1$. Moreover

$$\sum_{p=\ell+1}^{\infty} \frac{(-d)_{p+1}}{(p+1)!} \, {}_2F_1 \left(1+p-d, 1; p+2; e^{i\lambda_s} \right) = o(1)$$

since it is a tail sum of a convergent series. We therefore have

$$\frac{1}{n^{1-d}}\sum_{p=0}^{\ell}\left(\sum_{k=p+1}^{n}\frac{(-d)_k}{k!}e^{ik\lambda_s}\right)e^{-ip\lambda_s}=-\frac{1}{n^{1-d}}\frac{e^{i\lambda_s}}{\left(1-e^{i\lambda_s}\right)^{1-d}}+O\left(\frac{\ell}{ns}\right)+o\left(n^{1-d}\right) \tag{22}$$

Hence,

$$\begin{aligned}\frac{1}{n^d}\left[\frac{1}{n^{1-d}}\sum_{p=0}^{\ell}\left(\sum_{k=p+1}^{n}\frac{(-d)_k}{k!}e^{ik\lambda_s}\right)e^{-ip\lambda_s}\right] &= -\frac{1}{n^d}\frac{1}{n^{1-d}}\frac{e^{i\lambda_s}}{\left(1-e^{i\lambda_s}\right)^{1-d}}+O\left(\frac{1}{n^d}\frac{\ell}{ns}\right)+o\left(n^{1-2d}\right) \\ &= O\left(\frac{s^{d-1}}{n^d}\right)+O\left(\frac{1}{n^d}\frac{\ell}{ns}\right)+o\left(n^{1-2d}\right),\end{aligned} \tag{23}$$

which holds in the $d>1$ case. The first term in (23) is $O\left(s^{d-1}/n^d\right)$, which clearly dominates the second term when $\ell=n^{1-\beta}$ with $\beta>\frac{1}{2}$ (see below) and also dominates the third term. Therefore, the limit behavior of the first term in (21) can be written as

$$\begin{aligned}&\frac{1}{n^d}\left[\sum_{p=0}^{\ell}\left(\sum_{k=p+1}^{n}\frac{(-d)_k}{k!}e^{ik\lambda_s}\right)e^{-ip\lambda_s}\frac{X_{n-p}}{\sqrt{n}}\right] \\ =\ &\frac{1}{n^d}\left[\frac{1}{n^{1-d}}\sum_{p=0}^{\ell}\left(\sum_{k=p+1}^{n}\frac{(-d)_k}{k!}e^{ik\lambda_s}\right)e^{-ip\lambda_s}\frac{X_{n-p}}{n^{d-\frac{1}{2}}}\right] \\ =\ &\frac{1}{n^d}\left[\frac{1}{n^{1-d}}\sum_{p=0}^{\ell}\left(\sum_{k=p+1}^{n}\frac{(-d)_k}{k!}e^{ik\lambda_s}\right)e^{-ip\lambda_s}\left[\frac{X_n}{n^{d-\frac{1}{2}}}+o_p(1)\right]\right] \\ =\ &\frac{1}{n^d}\left[\frac{1}{n^{1-d}}\sum_{p=0}^{\ell}\left(\sum_{k=p+1}^{n}\frac{(-d)_k}{k!}e^{ik\lambda_s}\right)e^{-ip\lambda_s}\frac{X_n}{n^{d-\frac{1}{2}}}\right]+o_p\left(\frac{s^{d-1}}{n^d}\right) \\ =\ &-\frac{1}{n}\frac{e^{i\lambda_s}}{\left(1-e^{i\lambda_s}\right)^{1-d}}\frac{X_n}{n^{d-\frac{1}{2}}}+o_p\left(\frac{s^{d-1}}{n^d}\right),\end{aligned}$$

since $\frac{X_{n-p}}{n^{d-\frac{1}{2}}}=\frac{X_n}{n^{d-\frac{1}{2}}}+o_p(1)$ uniformly over $p<\ell$ with $\ell=n^{1-\beta}$, $\beta>\frac{1}{2}$, a property that can be shown to hold in the same way as in the proof of theorem 3.2 in Phillips (1999). It remains to show that the second term in (21) is of lesser order than the first term. Observe that for $d>1$,

$$\sum_{p=\ell+1}^{n-1}\left(\sum_{k=p+1}^{n}\frac{(-d)_k}{k!}e^{ik\lambda_s}\right)e^{-ip\lambda_s} = \sum_{p=0}^{\infty}\left(\sum_{k=p+1}^{n}\frac{(-d)_k}{k!}e^{ik\lambda_s}\right)e^{-ip\lambda_s}-\sum_{p=0}^{\ell}\left(\sum_{k=p+1}^{n}\frac{(-d)_k}{k!}e^{ik\lambda_s}\right)e^{-ip\lambda_s} \tag{24}$$

$$-\sum_{p=n}^{\infty}\left(\sum_{k=p+1}^{n}\frac{(-d)_k}{k!}e^{ik\lambda_s}\right)e^{-ip\lambda_s}. \tag{25}$$

Since the difference between two terms in (24) is negligible according to (22), and so is the tail sum in (25), it follows that the order of the second term in (21) is $o_p\left(\frac{s^{d-1}}{n^d}\right)$, and hence it may be neglected.

To conclude part (a), we may therefore extend the result in theorem 3.2(b) in Phillips (1999) to the $d > 1$ case as follows. For $\lambda = \lambda_s = \frac{2\pi s}{n} \to 0$ as $n \to \infty$, we have

$$\begin{aligned}\frac{1}{n^d}\frac{\widetilde{X}_{\lambda_s n}(d)}{\sqrt{n}} &= -\frac{1}{n^d}\frac{e^{i\lambda_s}}{\left(1-e^{i\lambda_s}\right)^{1-d}}\frac{X_{n-p}}{\sqrt{n}} + o_p\left(\frac{1}{n^d}\frac{e^{i\lambda_s}}{\left(1-e^{i\lambda_s}\right)^{1-d}}\frac{X_{n-p}}{\sqrt{n}}\right) \\ &= -\frac{1}{n}\frac{e^{i\lambda_s}}{\left(1-e^{i\lambda_s}\right)^{1-d}}\frac{X_n}{n^{d-\frac{1}{2}}} + o_p\left(\frac{s^{d-1}}{n^d}\right),\end{aligned}$$

as required. ■

5.2 Proof of Lemma 2.4 From lemma 3.1 in Phillips (1999), we have the following relation for the sinusoidal polynomial

$$D_n\left(e^{i\lambda_s}; d\right) = \left(1-e^{i\lambda_s}\right)^d + \frac{1}{\Gamma(-d)\, n^d}\frac{1}{2\pi i s}\left[1 + O\left(\frac{1}{s}\right)\right],$$

for $\lambda = \lambda_s = \frac{2\pi s}{n} \to 0$ and $s \to \infty$ as $n \to \infty$. With this behavior of $D_n\left(e^{i\lambda}; d\right)$ and lemma 2.3(a), it is easily deduced that

$$\begin{aligned}\frac{1}{n^d}w_x(\lambda_s) &= D_n\left(e^{i\lambda_s}; d\right)^{-1}\left[\frac{1}{n^d}w_u(\lambda_s) + \frac{1}{n^d}\frac{1}{\sqrt{2\pi n}}\widetilde{X}_{\lambda_s n}(d)\right] \\ &= \left[\left(1-e^{i\lambda_s}\right)^d + O\left(\frac{1}{n^d}\frac{1}{s}\right)\right]^{-1}\left[\frac{1}{n^d}w_u(\lambda_s) + \frac{1}{\sqrt{2\pi}}\frac{1}{n^d}\frac{\widetilde{X}_{\lambda_s n}(d)}{\sqrt{n}}\right] \\ &= \left(1-e^{i\lambda_s}\right)^{-d}\frac{1}{n^d}w_u(\lambda_s) - \frac{1}{\sqrt{2\pi}}\frac{1}{n}\frac{e^{i\lambda_s}}{\left(1-e^{i\lambda_s}\right)}\frac{X_n}{n^{d-\frac{1}{2}}} + o_p\left(\frac{1}{s}\right) \\ &= \left[\frac{1}{(-2\pi i s)^d}w_u(\lambda_s) - \frac{1}{\sqrt{2\pi}}\frac{1}{(-2\pi i s)}\frac{X_n}{n^{d-\frac{1}{2}}}\right]\left[1 + o\left(\frac{s}{n}\right)\right] + o_p\left(\frac{1}{s}\right), \quad (26)\end{aligned}$$

as stated in lemma 2.4 under the condition $\frac{s}{n} \to 0$. ■

5.3 Proof of Theorem 3.1 As is well known, $\frac{1}{m}\sum_{s=1}^m x_s^2 \to 1$ since

$$\sum_{s=1}^m x_s^2 \sim m - \left(\frac{1}{2} + \frac{1}{4}m^{-1}\right)(\ln m)^2 + \left(1 - 2D_m - m^{-1}D_m\right)\ln m + C_m + 2D_m - m^{-1}D_m^2,$$

where C_m and D_m are constants (e.g., equation (6) in Geweke and Porter-Hudak, 1983). Noting that

$$\frac{1}{(\ln m)\, m^{1-\alpha}}\sum_{j=1}^m \frac{\ln j}{j^\alpha} \sim \frac{1}{1-\alpha} + O\left(\frac{1}{\ln m}\right), \quad \alpha < 1,$$

we have

$$\frac{1}{(\ln m)\, m^{2-d}} \sum_{s=1}^{m} \frac{\ln s}{s^{d-1}} \approx \frac{1}{2-d} + o(1),$$

and then

$$\frac{(\ln m)\, m^{2-d}}{m} \frac{1}{(\ln m)\, m^{2-d}} \sum_{s=1}^{m} \ln s \frac{1}{s^{d-1}} \to 0, \tag{27}$$

for all $d > 1$. We also have

$$\frac{1}{m^{2-d}} \sum_{s=1}^{m} \frac{1}{s^{d-1}} = \frac{1}{2-d} + o(1),$$

so that

$$\frac{(\ln m)\, m^{2-d}}{m} \frac{1}{m^{2-d}} \sum_{s=1}^{m} \frac{1}{s^{d-1}} \to 0.$$

We have

$$2(\widehat{d}-1) \ = -2 \left[\sum_{s=1}^{m} x_s^2\right]^{-1} \left[\sum_{s=1}^{m} x_s v_n\right] - 2 \left[\sum_{s=1}^{m} x_s^2\right]^{-1} \left[\sum_{s=l+1}^{m} x_s \ln|1+\zeta_{ns}|\right] \tag{28}$$

from the representation of the logarithm of the periodogram in (13) and (15). Now, observe that the first term of (28) is identically zero, and therefore we need to show that the second term in (28) converges to zero in order to establish the inconsistency of the log periodogram estimator. First, we show that

$$\frac{1}{m} \sum_{s=1}^{m} x_s \ln|1+\zeta_{ns}| = o_p(1). \tag{29}$$

Throughout the following proof we use a domination argument to establish (29). That is, we will show

$$\left|\frac{1}{m} \sum_{s=1}^{m} x_s \ln|1+\zeta_{ns}|\right| = o_p(1).$$

Note that the following inequality holds for all x

$$|\ln|1+x|| \leq |x| + \frac{|x|}{|1+x|}.$$

Therefore, we have

$$\begin{aligned} \left|\frac{1}{m} \sum_{s=1}^{m} x_s \ln|1+\zeta_{ns}|\right| \ &\leq \ \frac{1}{m} \sum_{s=1}^{m} |x_s|\, |\ln|1+\zeta_{ns}|| \\ &\leq \ \frac{1}{m} \sum_{s=1}^{m} |x_s|\, |\zeta_{ns}| \left[1 + \frac{1}{|1+\zeta_{ns}|}\right] \\ &= \ \frac{1}{m} \sum_{s=1}^{m} |x_s|\, |\zeta_{ns}| + \frac{1}{m} \sum_{s=1}^{m} |x_s| \frac{1}{|1+\zeta_{ns}|}. \end{aligned} \tag{30}$$

Now, we have to show that both terms in (30) converge to zero in probability. To prove the convergence in the first term in (30), it suffices to show that

$$\frac{1}{m}\sum_{s=1}^{m}|x_s|\,|\xi_{ns}| \xrightarrow{p} 0, \tag{31}$$

where ξ_{ns} is already defined. The proof of (31) follows and will be called *Step (i)* for future reference.

Step (i).

Let

$$\xi_{ns} = \frac{\left(1-e^{i\lambda_s}\right)^{1-d}}{e^{i\lambda_s}} n^{1-d} \frac{w_u(\lambda_s)}{\left(\frac{X_n}{n^{d-\frac{1}{2}}}\right)} =: \frac{\left(1-e^{i\lambda_s}\right)^{1-d}}{e^{i\lambda_s}} n^{1-d} \upsilon_{ns},$$

then we need to show that

$$\frac{1}{m}\sum_{s=1}^{m}|x_s|\,|\xi_{ns}| = \frac{1}{m}\sum_{s=1}^{m}|x_s|\left|\frac{1}{s^{d-1}}\upsilon_{ns}\right| \xrightarrow{p} 0. \tag{32}$$

Note that

$$\begin{aligned}
& \frac{1}{m}\sum_{s=1}^{m}|x_s|\, n^{1-d}\left|1-e^{i\lambda_s}\right|^{1-d}|\upsilon_{ns}| \\
= \;& \frac{1}{m}\sum_{s=1}^{m}|x_s|\left(n\left|1-e^{i\lambda_s}\right|\right)^{1-d}|w_u(\lambda_s)|\left(\frac{1}{\left|\frac{X_n}{n^{d-\frac{1}{2}}}\right|}\right) \\
\leq \;& \left(\frac{1}{\left|\frac{X_n}{n^{d-\frac{1}{2}}}\right|}\right)\left(\frac{1}{m}\sum_{s=1}^{m}\left(|x_s|\left(n\left|1-e^{i\lambda_s}\right|\right)^{1-d}\right)^2\right)^{\frac{1}{2}}\left(\frac{1}{m}\sum_{s=1}^{m}|w_u(\lambda_s)|^2\right)^{\frac{1}{2}}.
\end{aligned}$$

From (27), we deduce that

$$\frac{1}{m}\sum_{s=1}^{m}\left||x_s|\left(n\left|1-e^{i\lambda_s}\right|\right)^{1-d}\right| \to 0,$$

since

$$\frac{1}{m}\sum_{s=1}^{m}\left|\frac{x_s}{s^{d-1}}\right| = \frac{1}{m}\sum_{s=1}^{m}\frac{\left|\ln s-\overline{\ln s}\right|}{s^{d-1}} = O\left(\frac{\ln(m)m^{2-d}}{m}\right), \text{ where } \overline{\ln s} = \frac{1}{m}\sum_{s=1}^{m}\ln s,$$

and hence it follows that

$$\frac{1}{m}\sum_{s=1}^{m}\left|\frac{x_s}{s^{d-1}}\right| \to 0 \text{ and hence } \frac{1}{m}\sum_{s=1}^{m}\left|\frac{x_s}{s^{d-1}}\right|^2 \to 0. \tag{33}$$

It is known (e.g., Marinucci and Robinson, 2000) that

$$\frac{X_n}{n^{d-\frac{1}{2}}} \xrightarrow{d} \frac{\omega}{\Gamma(d)} \int_0^1 (1-s)^{d-1} \, dW(s),$$

where ω^2 is the long run variance of u_t and W is a standard Brownian motion. Hence, $\left| X_n / n^{d-\frac{1}{2}} \right|^{-1} = O_p(1)$. We now need to show that

$$\frac{1}{m} \sum_{s=1}^{m} |w_u(\lambda_s)|^2 = O_p(1),$$

a result which can be obtained by a strong approximation approach. Phillips (2006) showed that asymptotically infinite collections of dft's of u_t at the fundamental frequencies in the vicinity of the origin can be treated as asymptotically independent normal variates. A suitable strong approximation that applies when u_t is a linear processes satisfying the condition (2) is given in Phillips (2006). Setting $S_{nk} = n^{-1/2} \Sigma_{j=1}^{k} u_j$, we can write this approximation in the form

$$\sup_{0 \le k \le n} \left| S_{nk} - B(\frac{k}{n}) \right| = o_p\left(\frac{1}{n^{\frac{1}{2}-\frac{1}{p}}} \right). \tag{34}$$

The following argument is based on the proof of Lemma C and the proof of theorem 3.3 in Phillips (2006). Using the embedding (34), we can write for $s = 1, ..., m$

$$w_u(\lambda_s) = \frac{1}{\sqrt{2\pi}} \int_0^1 e^{2\pi i s r} dB(r) + o_p\left(\frac{m}{n^{\frac{1}{2}-\frac{1}{p}}} \right), \tag{35}$$

where the error magnitude holds uniformly in $s \le m$. Let

$$\zeta_s =: \frac{1}{\sqrt{2\pi}} \int_0^1 e^{2\pi i s r} dB(r),$$

then the collection of random variates $\{\zeta_s\}_{s=1}^{m}$ are independent complex Gaussian $N_c(0, f_u(0))$,where $f_u(0) = (2\pi)^{-1} |C(1)|^2 \sigma^2$ is the spectral density of u_t at origin. Writing

$$w_u(\lambda_s) = \zeta_s + o_p\left(\frac{m}{n^{\frac{1}{2}-\frac{1}{p}}} \right),$$

uniformly over $s \le m$, then we have

$$\begin{aligned} \frac{1}{m} \sum_{s=1}^{m} |w_u(\lambda_s)|^2 &= \frac{1}{m} \sum_{s=1}^{m} |\zeta_s + A_s|^2 = \frac{1}{m} \sum_{s=1}^{m} |\zeta_s|^2 \left| 1 + \frac{A_s}{\zeta_s} \right|^2 \\ &\le \sup_s \left| 1 + \frac{A_s}{\zeta_s} \right|^2 \frac{1}{m} \sum_{s=1}^{m} |\zeta_s|^2, \end{aligned}$$

where $A_s = o_p\left(\frac{m}{n^{\frac{1}{2}-\frac{1}{p}}}\right)$ uniformly in $s \leq m$. By the application of lemma C in Phillips (2006), it follows that

$$\sup_s \left|\frac{A_s}{\zeta_s}\right| \leq \sup_s |A_s| \sup_s \left|\frac{1}{\zeta_s}\right| \leq \sup_s |A_s| \sum_{s\leq m} \frac{1}{|\zeta_s|} = O_p\left(\frac{m\log m}{n^{\frac{1}{2}-\frac{1}{p}}}\right),$$

and note that

$$\frac{1}{m}\sum_{s=1}^{m} |\zeta_s|^2 = O_p(1),$$

since $\{\zeta_s\}_{s=1}^m$ are independent complex Gaussian $N_c(0, f_u(0))$. Therefore, we have

$$\frac{1}{m}\sum_{s=1}^{m} |w_u(\lambda_s)|^2 = O_p\left(\frac{m\log m}{n^{\frac{1}{2}-\frac{1}{p}}}\right) O_p(1) = o_p(1)$$

if $\frac{m\log m}{n^{\frac{1}{2}-\frac{1}{p}}} \to 0$, which leads

$$\frac{1}{m}\sum_{s=1}^{m} |x_s| |\xi_{ns}| = o(1)O_p(1) O_p(1) = o_p(1), \tag{36}$$

giving (32) and completing Step (i).

Step (ii).

Next, we need to show that the second term of (30) converges to zero in probability. It suffices to show that

$$\frac{1}{m}\sum_{s=1}^{m} |x_s| \frac{1}{|1+\zeta_{ns}|} = o_p(1). \tag{37}$$

The proof proceeds in a similar way to Step (i). Note that

$$\begin{aligned}
\frac{1}{m}\sum_{s=1}^{m} |x_s| \frac{1}{|1+\zeta_{ns}|} &= \frac{1}{m}\sum_{s=1}^{m} \left|\frac{x_s}{s^{d-1}}\right| \frac{1}{\left|\frac{1}{s^{d-1}} + \frac{1}{s^{d-1}}\zeta_{ns}\right|} \\
&\leq \sup_s \frac{1}{\left|\frac{1}{s^{d-1}} + \frac{1}{s^{d-1}}\zeta_{ns}\right|} \frac{1}{m}\sum_{s=1}^{m} \left|\frac{x_s}{s^{d-1}}\right| \\
&= \sup_s \frac{1}{\left|\frac{1}{s^{d-1}} + \left(\frac{X_n}{n^{d-\frac{1}{2}}}\right)^{-1} w_u(\lambda_s)\right|} \frac{1}{m}\sum_{s=1}^{m} \left|\frac{x_s}{s^{d-1}}\right|
\end{aligned}$$

Now we have

$$\inf_s \left(\frac{X_n}{n^{d-\frac{1}{2}}}\right)^{-1} w_u(\lambda_s) = \left(\frac{X_n}{n^{d-\frac{1}{2}}}\right)^{-1} \inf_s w_u(\lambda_s) = O_p(1) \tag{38}$$

uniformly in $s \leq m$ by the strong approximation given in (35) as long as $\frac{m}{n^{\frac{1}{2}-\frac{1}{p}}} \to 0$, and $\frac{1}{s^{d-1}}$ is bounded above zero for all $s \leq m$. By (33) and (38), we have

$$\sup_s \frac{1}{\left| \frac{1}{s^{d-1}} + \left(\frac{X_n}{n^{d-\frac{1}{2}}} \right)^{-1} w_u(\lambda_s) \right|} \frac{1}{m} \sum_{s=1}^{m} \left| \frac{x_s}{s^{d-1}} \right| = O_p(1)\, o(1)$$

and hence

$$\frac{1}{m} \sum_{s=1}^{m} |x_s| \frac{1}{|1+\zeta_{ns}|} = o_p(1) \tag{39}$$

as required, completing Step (ii).

Combining (36) and (39) from these two steps we get

$$\frac{1}{m} \sum_{s=1}^{m} |x_s| |\zeta_{ns}| + \frac{1}{m} \sum_{s=1}^{m} |x_s| \frac{1}{|1+\zeta_{ns}|} \xrightarrow{p} 0,$$

and hence

$$\left| \frac{1}{m} \sum_{s=1}^{m} x_s \ln |1+\zeta_{ns}| \right| \xrightarrow{p} 0,$$

by the inequality in (30), which further implies that

$$\frac{1}{m} \sum_{s=1}^{m} x_s \ln |1+\zeta_{ns}| \xrightarrow{p} 0. \tag{40}$$

From (40), the stated inconsistency result follows, viz.,

$$2(\widehat{d}-1) \xrightarrow{p} 0, \tag{41}$$

when $1 < d < 2$. ■

Next, we give a lemma which enables us to calculate the moments of the logarithmic function of the periodogram, which is needed for the proof of Theorem 3.3. The statistical properties of non-linear functions of the periodogram of stationary processes have been explored earlier in the literature, notably by Chen and Hannan (1980), Von Sachs (1994), and Janas and Von Sachs (1995), and their results for the moments of such non-linear functions are not dependent upon Gaussianity assumptions. We will use the following lemma, which is a slightly modified version of Lemma A.1 in Janas and Von Sachs (1995).

5.4 Lemma *Assume that i.i.d. sequence ε_t satisfies the Cramér condition* (19) *and condition* (20) *and has unit variance and finite fourth moments. Then*

(i) $\mathbf{E} \ln(I_\varepsilon(\lambda_j)) = \mathbf{E}[\ln Z] + O(n^{-1}) = \gamma + O(n^{-1})$, *uniformly in* λ_j,

(ii) $\mathbf{Var} \ln \left(I_{\varepsilon}\left(\lambda_j\right)\right)=\mathbf{Var}[\ln Z]+O\left(n^{-1}\right)=\frac{\pi^2}{6}+O\left(n^{-1}\right)$, *uniformly in* λ_j,

(iii) $\mathbf{Cov}\left[\ln \left(I_{\varepsilon}\left(\lambda_i\right)\right), \ln \left(I_{\varepsilon}\left(\lambda_j\right)\right)\right]=O\left(n^{-1}\right)$, *uniformly in* $\lambda_i \neq \pm \lambda_j$,

where Z denotes a standard exponentially distributed random variable (i.e. with parameter 1) and γ is the Euler's Gamma. The frequency index j can be any number such that $1<j<\frac{n}{2}$, i.e., the lemma holds irrespective of j.

The Cramér condition is needed for the approximation of the joint density of discrete Fourier transforms and for non-linear functions of the dft, but is not enough for the logarithmic function because of the singular behavior of $\ln x$ at $x=0$, as discussed in the proof of Corollary 3.4 of Janas and Von Sachs (1995). The additional assumption (20) takes care of this difficulty by ensuring that the distribution of $\sum I_{\varepsilon}\left(\lambda_j\right)$ is absolutely continuous for sufficiently large n.

To extend the results of lemma 5.4 to linear processes, we use the spectral form of the BN decomposition used in Phillips and Solo (1992), as we now demonstrate. In particular, we may write

$$\frac{1}{m} \sum_{s=1}^m I_u\left(\lambda_s\right)=\frac{1}{m} \sum_{s=1}^m I_{\varepsilon}\left(\lambda_s\right)+o_p(1),$$

which is deduced as follows. As in Phillips and Solo (1992), decompose the operator $C(L)$ as

$$C(L)=C\left(e^{i \lambda_s}\right)+\widetilde{C}\left(e^{-i \lambda_s} L\right)\left(e^{-i \lambda_s} L-1\right), \quad \widetilde{C}(L)=\sum_{j=0}^{\infty}\left(\sum_{k=j+1}^{\infty} c_k e^{i \lambda_s k}\right) L^j,$$

where $\sum_{j=0}^{\infty}\left|\sum_{k=j+1}^{\infty} c_k\right|<\infty$ in view of the summability condition in (2). The dft of u_t can then be written as

$$w_u\left(\lambda_s\right)=C\left(e^{i \lambda_s}\right) w_{\varepsilon}\left(\lambda_s\right)+\frac{1}{\sqrt{2 \pi n}}\left(\varepsilon_{\lambda_s 0}-\varepsilon_{\lambda_s n}\right), \tag{42}$$

with

$$\begin{aligned}
\varepsilon_{\lambda_s n} & =\widetilde{C}\left(e^{-i \lambda} L\right) \varepsilon_n=\sum_{j=0}^{\infty}\left(\sum_{k=j+1}^{\infty} c_k e^{i \lambda_s k}\right) e^{-i \lambda_s j} \varepsilon_{n-j} \\
& =\sum_{j=0}^{\infty}\left(\sum_{k=j+1}^{\infty} c_k\right) e^{i \lambda_s(k-j)} \varepsilon_{n-j}=\sum_{j=0}^{\infty} \widetilde{c}_{j \lambda_s} e^{-i \lambda_s j} \varepsilon_{n-j},
\end{aligned}$$

where $\widetilde{c}_{j \lambda_s}=\sum_{k=j+1}^{\infty} c_k e^{i \lambda_s k}$. The following lemma shows that the second component in this decomposition is negligible uniformly over s and is needed in our log periodogram regression application.

5.5 Lemma *If Assumption 2.1 holds,* $\max_s\left|\frac{1}{n^{\delta}} \varepsilon_{\lambda_s n}\right| \xrightarrow{p} 0$ *for all* $\delta>0$.

Proof We have

$$\begin{aligned}\max_s \left|\varepsilon_{\varepsilon_{\lambda_s n}}\right| &= \max_s \left|\sum_{j=0}^{\infty} \widetilde{c}_{j\lambda_s} e^{-i\lambda_s j} \varepsilon_{n-j}\right| \\ &\leq \max_s \left[\sum_{j=0}^{\infty} \left|\widetilde{c}_{j\lambda_s} \varepsilon_{n-j}\right|\right] \leq \left[\sum_{j=0}^{\infty} \left|\overline{c}_j \varepsilon_{n-j}\right|\right],\end{aligned}$$

where $\overline{c}_j = \sum_{k=j+1}^{\infty} |c_k|$. So,

$$\mathbf{E} \max_s \left|\varepsilon_{\lambda_s n}\right| \leq \mathbf{E}\left[\sum_{j=0}^{\infty} \left|\overline{c}_j \varepsilon_{n-j}\right|\right] = \mathbf{E}\left[\sum_{j=0}^{\infty} \left|\overline{c}_j \varepsilon_{-j}\right|\right].$$

It follows that, for any $\eta, \delta > 0$

$$\begin{aligned}P\left(\frac{1}{n^\delta} \max_s \left|\varepsilon_{\lambda_s n}\right| > \eta\right) &< \frac{\mathbf{E} \max_s \left|\varepsilon_{\lambda_s n}\right|}{\eta n^\delta} \leq \frac{\mathbf{E}\left[\sum_{j=0}^{\infty} \left|\overline{c}_j \varepsilon_{-j}\right|\right]}{\eta n^\delta} \\ &\leq \frac{\sum_{j=0}^{\infty} \left|\overline{c}_j\right| \mathbf{E}\left|\varepsilon_0\right|}{\eta n^\delta} = \frac{\left(\sum_{k=0}^{\infty} k\left|c_k\right|\right) \mathbf{E}\left|\varepsilon_0\right|}{\eta n^\delta} \to 0,\end{aligned}$$

in view of (2), so that

$$\max_s \left|\frac{1}{n^\delta} \varepsilon_{\lambda_s n}\right|, \max_s \left|\frac{1}{n^\delta} \varepsilon_{\lambda_s 0}\right| \xrightarrow{p} 0.$$

as required. ■

The next lemma applies the Phillips and Solo (1992) device to the log periodogram $\ln I_u(\lambda_s)$.

5.6 Lemma *If Assumption 2.1 and the assumptions in theorem 3.3 hold and* $\frac{m \log m}{n^{\frac{1}{2}-\delta}} \to 0$ *for any* $\delta > 0$*, then*

$$\frac{1}{m} \sum_{s=1}^{m} x_s \ln I_u(\lambda_s) = \frac{1}{m} \sum_{s=1}^{m} x_s \ln I_\varepsilon(\lambda_s) + o_p(1).$$

Proof Using (42), we have

$$\begin{aligned}\frac{1}{m} \sum_{s=1}^{m} x_s \ln I_u(\lambda_s) &= \frac{1}{m} \sum_{s=1}^{m} x_s \ln \left|C\left(e^{i\lambda_s}\right) w_\varepsilon(\lambda_s) + \frac{1}{\sqrt{2\pi n}}\left(\varepsilon_{\lambda_s 0} - \varepsilon_{\lambda_s n}\right)\right|^2 \\ &= \frac{1}{m} \sum_{s=1}^{m} x_s \ln \left|C\left(e^{i\lambda_s}\right)\right|^2 I_\varepsilon(\lambda_s) + \frac{1}{m} \sum_{s=1}^{m} x_s \ln \left|1 + \Psi_s\right|^2,\end{aligned}$$

where

$$\Psi_s = \frac{\frac{1}{\sqrt{2\pi n}}\left(\varepsilon_{\lambda_s 0} - \varepsilon_{\lambda_s n}\right)}{C\left(e^{i\lambda_s}\right) w_\varepsilon(\lambda_s)}$$

We need to show that

$$\frac{1}{m}\sum_{s=1}^{m} x_s \ln|1+\Psi_s| \xrightarrow{p} 0.$$

Note that

$$\left|\frac{1}{m}\sum_{s=1}^{m} x_s \ln|1+\Psi_s|\right| \le \frac{1}{m}\sum_{s=1}^{m}|x_s|\left|\ln|1+\Psi_s|\right| \le \frac{1}{m}\sum_{s=1}^{m}|x_s|\sup_s\left|\ln|1+\Psi_s|\right|.$$

On the other hand, from the inequality $|\ln(1+Y)| \le 2|Y|$ for $|Y| \le \frac{1}{2}$, we deduce that

$$\mathbf{P}\left[|\ln(1+Y)| > \epsilon\right] \le \mathbf{P}\left[|Y| > \epsilon/2\right] \text{ for } \epsilon \le 1,$$

which holds for nonnegative $1+Y$, from Robinson (1995a). Then, if

$$\sup_s |\Psi_s| \xrightarrow{p} 0, \tag{43}$$

it follows that

$$\sup_s \left|\ln|1+\Psi_s|\right| \xrightarrow{p} 0. \tag{44}$$

Observe that

$$\begin{aligned}\sup_s|\Psi_s| &= \sup_s\left|\frac{\frac{1}{\sqrt{n}}\left(\varepsilon_{\lambda_s 0}-\varepsilon_{\lambda_s n}\right)}{C\left(e^{i\lambda_s}\right)w_\varepsilon(\lambda_s)}\right| = \sup_s\left|\frac{\frac{1}{n^\delta}\left(\varepsilon_{\lambda_s 0}-\varepsilon_{\lambda_s n}\right)}{n^{\frac{1}{2}-\delta}C\left(e^{i\lambda_s}\right)w_\varepsilon(\lambda_s)}\right| \\ &\le \frac{1}{n^{\frac{1}{2}-\delta}}\sup_s\left|\frac{1}{n^\delta}\left(\varepsilon_{\lambda_s 0}-\varepsilon_{\lambda_s n}\right)\right|\sup_s\left|\frac{1}{C\left(e^{i\lambda_s}\right)w_\varepsilon(\lambda_s)}\right|\end{aligned}$$

for $0<\delta<\frac{1}{2}$. From lemma 5.5, we have

$$\sup_s\left|\frac{1}{n^\delta}\left(\varepsilon_{\lambda_s 0}-\varepsilon_{\lambda_s n}\right)\right| \xrightarrow{p} 0.$$

As before, by the application of the lemma C and the proof of theorem 3.2 in Phillips (1996), we have

$$\sup_s\left|\frac{1}{C\left(e^{i\lambda_s}\right)w_\varepsilon(\lambda_s)}\right| = O_p(m\log m),$$

and therefore, for all $s<m$, it follows that

$$\sup_s|\Psi_s| = o_p\left(\frac{m\log m}{n^{\frac{1}{2}-\delta}}\right).$$

Therefore, we have the desired result in (43), and (44) follows as long as $\frac{m\log m}{n^{\frac{1}{2}-\delta}} \to 0$. Since $\frac{1}{m}\sum_{s=1}^{m}|x_s| = O(1)$, as shown in Robinson (1995b), it follows that

$$\frac{1}{m}\sum_{s=1}^{m} x_s \ln\left|1+\frac{\frac{1}{\sqrt{2\pi n}}\left(\varepsilon_{\lambda_s 0}-\varepsilon_{\lambda_s n}\right)}{C\left(e^{i\lambda_s}\right)w_\varepsilon(\lambda_s)}\right| = o_p(1).$$

Therefore, we have

$$\begin{aligned}\frac{1}{m}\sum_{s=1}^{m} x_s \ln I_u(\lambda_s) &= \frac{1}{m}\sum_{s=1}^{m} x_s \ln \left|C\left(e^{i\lambda_s}\right)\right|^2 I_\varepsilon(\lambda_s) + o_p(1) \\ &= \frac{1}{m}\sum_{s=1}^{m} x_s \ln I_\varepsilon(\lambda_s) + o_p(1),\end{aligned}$$

since $|C(1)|^2 < \infty$ and

$$\frac{1}{\sqrt{m}}\sum_{s=1}^{m} x_s \ln \left|C\left(e^{i\lambda_s}\right)\right|^2 \to 0$$

as shown in lemma 1 of Hurvich, Deo, and Brodsky (1998) under $\frac{m^5}{n^4} \to 0$.

5.7 Proof of Theorem 3.2 As defined in (14), the log periodogram regression estimator employs the frequencies $\{\lambda_s,\ s = 1, ..., m\}$. From (13) in section 3 of the paper, we have the following representation of the periodogram over frequencies $\{\lambda_s,\ s = l+1, ..., m\}$ where $\frac{n^\alpha}{l} \to 0$ for some $\alpha \in (\frac{1}{2}, 1)$,

$$\ln(I_x(\lambda_s)) = \ln\left(\frac{1}{2\pi}\right) + 2\ln\left(\frac{n^d}{n}\right) - 2\ln(\lambda_s) + 2\ln\left(\left|\frac{X_n}{n^{d-\frac{1}{2}}}\right|\right) + 2\ln|1 + \zeta_{ns}|.$$

The representation over frequencies $\{\lambda_s,\ s = 1, ..., l\}$ should be slightly changed as the following argument shows for $\frac{1}{2} < d < 1$. We work from the representation of $w_x(\lambda_s)$ given in (8) and the representation of $\widetilde{X}_{\lambda_s n}(d)$ given in (21). Proceeding as in the proof of lemma 2.3 and using the proof of theorem 3.2 in Phillips (1999), the first term of (21) has a factor of the form

$$\begin{aligned}&\frac{1}{n^{1-d}}\sum_{p=0}^{\ell}\left(\sum_{k=p+1}^{n}\frac{(-d)_k}{k!}e^{ik\lambda_s}\right)e^{-ip\lambda_s} \\ =\ & -\frac{1}{n^{1-d}}\frac{e^{i\lambda_s}}{(1-e^{i\lambda_s})^{1-d}} + \frac{1}{n^{1-d}}\sum_{p=\ell+1}^{\infty}\frac{(-d)_{p+1}}{(p+1)!}\,{}_2F_1\left(1+p-d, 1; p+2; e^{i\lambda_s}\right) + O\left(\frac{\ell}{ns}\right).\end{aligned} \tag{45}$$

However, unlike the proof of lemma 2.3, we will not here assume that $\frac{n^\alpha}{s} \to 0$ (for some $\alpha \in (\frac{1}{2}, 1)$). Hence, the first term in (45) does not necessarily dominate the second term in (45). Let $C(\ell, d, s) = \sum_{p=\ell+1}^{\infty}\frac{(-d)_{p+1}}{(p+1)!}\,{}_2F_1\left(1+p-d, 1; p+2; e^{i\lambda_s}\right)$ for notational simplicity.

Then, we have

$$
\begin{aligned}
& \frac{1}{n^d}\left[\sum_{p=0}^{\ell}\left(\sum_{k=p+1}^{n}\frac{(-d)_k}{k!}e^{ik\lambda_s}\right)e^{-ip\lambda_s}\frac{X_{n-p}}{\sqrt{n}}\right] \\
= & \frac{1}{n^d}\left[\frac{1}{n^{1-d}}\sum_{p=0}^{\ell}\left(\sum_{k=p+1}^{n}\frac{(-d)_k}{k!}e^{ik\lambda_s}\right)e^{-ip\lambda_s}\left[\frac{X_n}{n^{d-\frac{1}{2}}}+o_p(1)\right]\right] \\
= & \frac{1}{n^d}\left[-\frac{1}{n^{1-d}}\frac{e^{i\lambda_s}}{\left(1-e^{i\lambda_s}\right)^{1-d}}+\frac{1}{n^{1-d}}C(\ell,d,s)\right]\frac{X_n}{n^{d-\frac{1}{2}}}+O_p\left(\frac{1}{n^d}\frac{\ell}{ns}\right), \qquad (46)
\end{aligned}
$$

since $\frac{X_{n-p}}{n^{d-\frac{1}{2}}}=\frac{X_n}{n^{d-\frac{1}{2}}}+o_p(1)$ uniformly over $p<\ell$ such that $\ell=n^{1-\beta},\ \beta>\frac{1}{2}$, as before. Moreover, the second term in (21) is of lesser order than the first term, as we now show. In particular, using lemma C(b) in Phillips (1999), we have

$$
\begin{aligned}
& \frac{1}{n^{1-d}}\sum_{p=\ell+1}^{n-1}\left(\sum_{k=p+1}^{n}\frac{(-d)_k}{k!}e^{ik\lambda_s}\right)e^{-ip\lambda_s}\frac{X_{n-p}}{n^{d-\frac{1}{2}}} \\
= & O_p\left(\frac{1}{n^{1-d}}\sum_{p=\ell+1}^{n-1}\frac{1}{p^d s}\frac{X_{n-p}}{n^{d-\frac{1}{2}}}\right)=O_p\left(\frac{1}{s}\frac{1}{n^{1-d}}\left(\sum_{p=\ell+1}^{\infty}\frac{1}{p^d}-\sum_{p=n}^{\infty}\frac{1}{p^d}\right)\frac{X_{n-p}}{n^{d-\frac{1}{2}}}\right) \\
= & O_p\left(\frac{1}{s}\frac{1}{n^{1-d}}\frac{1}{\ell^{d-1}}\right)=O_p\left(\frac{1}{s}\left(\frac{n}{\ell}\right)^{d-1}\right).
\end{aligned}
$$

Therefore, the order of the second term in (21) can be written as

$$
\frac{1}{n^d}\left[\frac{1}{n^{1-d}}\sum_{p=\ell+1}^{n-1}\left(\sum_{k=p+1}^{n}\frac{(-d)_k}{k!}e^{ik\lambda_s}\right)e^{-ip\lambda_s}\frac{X_{n-p}}{n^{d-\frac{1}{2}}}\right]=O_p\left(\frac{1}{s}\frac{1}{\ell^{d-1}}\frac{1}{n}\right), \qquad (47)
$$

which may be neglected because the order of the first term in (46) exceeds the order of (47).

Therefore, for $s=1,...,\ell$, we have

$$
\frac{1}{n^d}\frac{\widetilde{X}_{\lambda_s n}(d)}{\sqrt{n}}=-\frac{1}{n}\left[\frac{e^{i\lambda_s}}{\left(1-e^{i\lambda_s}\right)^{1-d}}-C(\ell,d,s)\right]\frac{X_n}{n^{d-\frac{1}{2}}}+O_p\left(\frac{1}{n^d}\frac{\ell}{ns}\right),
$$

which includes the additional term $C(\ell,d,s)$ compared to the representation given in the lemma 2.3. Now, the dft over frequencies $\{\lambda_s,\ s=1,...,\ell\}$ will be

$$
\begin{aligned}
\frac{1}{n^d}w_x(\lambda_s) = & D_n\left(e^{i\lambda_s};d\right)^{-1}\left[\frac{1}{n^d}w_u(\lambda_s)+\frac{1}{n^d}\frac{1}{\sqrt{2\pi n}}\widetilde{X}_{\lambda_s n}(d)\right] \\
= & \left(1-e^{i\lambda_s}\right)^{-d}\frac{1}{n^d}w_u(\lambda_s) \\
& -\frac{1}{n}\left(\frac{e^{i\lambda_s}}{\left(1-e^{i\lambda_s}\right)}-\left(1-e^{i\lambda_s}\right)^{-d}C(\ell,d,s)\right)\frac{X_n}{\sqrt{2\pi}n^{d-\frac{1}{2}}}+O_p\left(\frac{\ell}{ns^{d+1}}\right). \qquad (48)
\end{aligned}
$$

As shown in Phillips (1999), we have the following representation

$$\begin{aligned} C(\ell,d,s) &= \sum_{k=0}^{\infty}\left[\sum_{p=\ell+1}^{\infty}\frac{(-d)_{p+1}}{(p+1)!}\frac{(1-d+k)_p\,(2)_p}{(1-d)_p\,(k+2)_p}\right]\frac{(1-d)_k}{(2)_k}e^{i\lambda_s k} \\ &= \sum_{k=0}^{\infty}\left[(-d)\sum_{p=\ell+1}^{\infty}\frac{(1-d+k)_p}{(k+2)_p}\right]\frac{(1-d)_k}{(2)_k}e^{i\lambda_s k} \\ &= O\left(\sum_{k=0}^{\infty}\left[\frac{(-d)\,\Gamma(k+2)}{\Gamma(1-d+k)}\sum_{p=\ell+1}^{\infty}\frac{1}{p^{1+d}}\right]\frac{(1-d)_k}{(2)_k}e^{i\lambda_s k}\right) \\ &= O\left(\frac{1}{\ell^d}\frac{(-d)}{\Gamma(1-d)}\sum_{k=0}^{\infty}e^{i\lambda_s k}\right)=O\left(\frac{1}{\ell^d}\frac{1}{(1-e^{i\lambda_s})}\right). \end{aligned}$$

That is,

$$C(\ell,d,s)\sim\frac{1}{\ell^d}\frac{1}{(1-e^{i\lambda_s})}c,$$

where c is a constant. Therefore, the dft in (48) can be rewritten as

$$\begin{aligned} \frac{1}{n^d}w_x(\lambda_s) &= \left(1-e^{i\lambda_s}\right)^{-d}\frac{1}{n^d}w_u(\lambda_s) \\ &\quad -\left[\frac{1}{n}\frac{e^{i\lambda_s}}{(1-e^{i\lambda_s})}-\frac{c}{n}\left(1-e^{i\lambda_s}\right)^{-d}\frac{1}{\ell^d}\frac{1}{(1-e^{i\lambda_s})}\right]\frac{X_n}{\sqrt{2\pi}n^{d-\frac{1}{2}}}+O_p\left(\frac{\ell}{ns^{d+1}}\right) \\ &= \left[\frac{1}{(-2\pi is)^d}w_u(\lambda_s)-\left(\frac{1}{(-2\pi is)}-\frac{c_1}{s^{1+d}}\frac{n^d}{\ell^d}\right)\frac{X_n}{\sqrt{2\pi}n^{d-\frac{1}{2}}}\right]\left[1+o\left(\frac{s}{n}\right)\right]+O_p\left(\frac{\ell}{ns^{d+1}}\right), \end{aligned}$$

where $c_1=c/(-2\pi i)^{1+d}$. Then, the periodogram is

$$\frac{I_x(\lambda_s)}{n^{2d}} = \left|\left[\frac{1}{(-2\pi is)^d}w_u(\lambda_s)-\left(\frac{1}{(-2\pi is)}-\frac{c_1}{s^{1+d}}\frac{n^d}{\ell^d}\right)\frac{X_n}{\sqrt{2\pi}n^{d-\frac{1}{2}}}\right]\left[1+o\left(\frac{s}{n}\right)\right]+O_p\left(\frac{\ell}{ns^{d+1}}\right)\right|^2. \tag{49}$$

From (49), the periodogram over frequencies $\{\lambda_s,\ s=1,...,\ell\}$ can be rearranged as

$$I_x(\lambda_s)=\left(\frac{n}{2\pi s}\right)^{2d}\left(\frac{n^d}{\ell^d}\right)^2\left(\frac{X_n}{n^{d-\frac{1}{2}}}\right)^2\left|\xi_{ns}\frac{\ell^d}{n^d}+O\left(\frac{1}{s}\right)+O\left(\frac{1}{s^{1-d}}\frac{\ell^d}{n^d}\right)\right|^2, \tag{50}$$

where

$$\xi_{ns}=\frac{w_u(\lambda_s)}{\left(\frac{X_n}{n^{d-\frac{1}{2}}}\right)},$$

and the relation (50) holds uniformly in $s \leq m$. Next, break the estimator $\widehat{d}$ in (14) down into the following two components

$$
\begin{aligned}
2\widehat{d} &= -\left[\sum_{s=1}^{m} x_s^2\right]^{-1}\left[\sum_{s=1}^{m} x_s \ln I_x(\lambda_s)\right] \\
&= -\left[\sum_{s=1}^{m} x_s^2\right]^{-1}\left[\sum_{s=1}^{l} x_s \ln I_x(\lambda_s) + \sum_{s=l+1}^{m} x_s \ln I_x(\lambda_s)\right].
\end{aligned}
$$

Using (17), (18) and (50), we have

$$
\begin{aligned}
2(\widehat{d} - d) &= -\left(\sum_{s=1}^{m} x_s^2\right)^{-1} \sum_{s=l+1}^{m} x_s \upsilon - \left(\sum_{s=1}^{m} x_s^2\right)^{-1} \sum_{s=l+1}^{m} x_s \ln|\xi_{ns}| \\
&\quad - \left(\sum_{s=1}^{m} x_s^2\right)^{-1} \sum_{s=1}^{l} x_s \upsilon_n - \left(\sum_{s=1}^{m} x_s^2\right)^{-1} \sum_{s=1}^{l} x_s \ln\left|\xi_{ns}\frac{\ell^d}{n^d}\right| + o_p(1) \\
&= -\left(\sum_{s=1}^{m} x_s^2\right)^{-1} \sum_{s=1}^{m} x_s \upsilon_n - \left(\sum_{s=1}^{m} x_s^2\right)^{-1} \sum_{s=1}^{m} x_s \ln|\xi_{ns}| \\
&\quad - \left(\sum_{s=1}^{m} x_s^2\right)^{-1} \sum_{s=1}^{l} x_s \ln\frac{\ell^d}{n^d} + o_p(1). \qquad (51)
\end{aligned}
$$

As before, the first term of (51) is zero, so we need only show that

$$
\frac{1}{m}\sum_{s=1}^{m} x_s \ln|\xi_{ns}| \xrightarrow{p} 0,
$$

and

$$
\frac{1}{m}\sum_{s=1}^{l} x_s \ln\frac{\ell^d}{n^d} \to 0.
$$

Observe that

$$
\begin{aligned}
\frac{1}{m}\sum_{s=1}^{m} x_s \ln|\xi_{ns}| &= \frac{1}{m}\sum_{s=1}^{m} x_s \ln\frac{|w_u(\lambda_s)|}{\left|\frac{X_n}{n^{d-\frac{1}{2}}}\right|} \\
&= \frac{1}{m}\sum_{s=1}^{m} x_s \ln|w_u(\lambda_s)| - \frac{1}{m}\sum_{s=1}^{m} x_s \ln\left|\frac{X_n}{n^{d-\frac{1}{2}}}\right|,
\end{aligned}
$$

where the second term is also zero. Using lemma 5.6, we have

$$
\frac{1}{m}\sum_{s=1}^{m} x_s \ln I_u(\lambda_s) = \frac{1}{m}\sum_{s=1}^{m} x_s \ln I_\varepsilon(\lambda_s) + o_p(1),
$$

and, hence, for

$$\frac{1}{m}\sum_{s=1}^{m} x_s \ln |w_u(\lambda_s)| = \frac{1}{2}\frac{1}{m}\sum_{s=1}^{m} x_s \ln |I_u(\lambda_s)| \xrightarrow{p} 0$$

to hold, we need only show that

$$\frac{1}{m}\sum_{s=1}^{m} x_s \ln I_\varepsilon(\lambda_s) \xrightarrow{p} 0.$$

To do so, we evaluate the first two moments. By lemma 5.4, we have

$$\mathbf{E}\left[\frac{1}{m}\sum_{s=1}^{m} x_s \ln I_\varepsilon(\lambda_s)\right] \to 0.$$

The variance term is

$$\mathbf{Var}\left[\frac{1}{m}\sum_{s=1}^{m} x_s \ln I_\varepsilon(\lambda_s)\right] = \frac{1}{m^2}\sum_{s=1}^{m} x_s^2 \mathbf{Var}\left[\ln I_\varepsilon(\lambda_s)\right] + 2\frac{1}{m^2}\sum_{s=1}^{m}\sum_{r=s+1}^{m} x_s x_r \mathbf{Cov}\left[\ln I_\varepsilon(\lambda_s), \ln I_\varepsilon(\lambda_r)\right],$$

the first term of which is clearly $o_p(1)$. Moreover,

$$\frac{1}{m^2}\sum_{s=1}^{m}\sum_{r=s+1}^{m} x_s x_r \mathbf{Cov}\left[\ln I_\varepsilon(\lambda_s), \ln I_\varepsilon(\lambda_r)\right] = o(1),$$

from result (iii) of lemma 5.4 and the fact that $\frac{1}{m}\sum_{s=1}^{m}|x_s| = O(1)$, which is given in Robinson (1995b). Therefore, we have

$$\frac{1}{m}\sum_{s=1}^{m} x_s \ln |w_u(\lambda_s)| = o_p(1). \tag{52}$$

It remains to show that the third term in (51) goes to zero, which clearly holds because

$$\sum_{s=1}^{l}|x_s| = \sum_{s=1}^{l}\left|\ln s - \overline{\ln s}\right| = O(l \ln l) + O(l \ln m),$$

and

$$\frac{1}{m}\sum_{s=1}^{l} x_s \ln \frac{\ell^d}{n^d} = O\left(\frac{l(\ln n)^2}{m}\right) = o(1), \tag{53}$$

under the assumption $\frac{l(\ln n)^2}{m} \to 0$. From (52) and (53), we have $\widehat{d} - d = o_p(1)$, giving the consistency of log periodogram regression for $\frac{1}{2} < d < 1$. ∎

References

[1] Abadir, K., W. Distaso, and L. Giraitis (2005). "Nonstationarity-extended local Whittle Estimation". mimeo, University of York.

[2] Agiakloglou, C., P. Newbold and M. Wohar (1993). "Bias in an Estimator of the Fractional Difference Parameter", *Journal of Time Series Analysis* 14, 235-246.

[3] Akonom, J. and C. Gourieroux. (1987). A Functional Central Limit Theorem for Fractional Processes, Discussion Paper #8801, CEPREMAP, Paris.

[4] Bhattacharya, R. N. and R. R. Rao (1976). *Normal Approximation and Asymptotic Expansions,* New York: Wiley.

[5] Bloomfield, P. (1991). "Time Series Methods", in *Statistical Theory and Modelling*, London, Chapman and Hall, 152-176.

[6] Chen, Z.-G. and E. J. Hannan. (1980). "The Distribution of Periodogram Ordinates", *Journal of Time Series Analysis* 1, 73-82.

[7] Cheung, Y-W. and K. S. Lai. (1993). "A Fractional Cointegration Analysis of Purchasing Power Parity", *Journal of Business and Economic Statistics* 11, 103-112.

[8] 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.

[9] Gourieroux, C., F., A. Maurel, and A. Monfort (1987). "Regression and Nonstationarity", No. 8708, Institut National de la Statistique et des Etudes Economiques, France.

[10] 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.

[11] 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.

[12] Janas, J. and R. Von Sachs (1995). "Consistency for Non-Linear Functions of the Periodogram of Tapered Data", *Journal of Time Series Analysis* 16, 585-606.

[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] Liu, M. (1998). "Asymptotics of Nonstationary Fractional Integrated Series", *Econometric Theory* 14, 641-662.

[15] Marinucci, D. and P. M. Robinson (2000). "Weak convergence of multivariate fractional processes". *Stochastic Processes and Their Applications*, 86, 103-120.

[16] Maynard, A. and P. C. B. Phillips. (1998). "Rethinking an Old Empirical Puzzle: Econometric Evidence on the Forward Discount Anomaly." *Journal of Applied Econometrics* 16, Issue 6, 671 - 708

[17] Phillips, P. C. B. (1998/2005). "Econometric Analysis of Fisher's Equation", *Cowles Foundation Discussion Paper* #1180, Yale University. Published in American Journal of Economics and Sociology, 69, 117

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

[19] Phillips, P. C. B. (2006). "Unit Root Log Periodogram Regression", Yale University. (download from http://cowles.econ.yale.edu), *Forthcoming in Journal of Econometrics.*.

[20] Phillips, P.C.B. and K. Shimotsu (2004), "Local Whittle Estimation in Nonstationary and Unit Root Cases," *Annals of Statistics* 32, 656-692.

[21] Phillips, P. C. B. and V. Solo (1992). "Asymptotics for Linear Processes," *The Annals of Statistics* 20, 971–1001.

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

[23] Robinson, P. M. (1995b). "Gaussian Semiparametric Estimation of Long Range Dependence", *The Annals of Statistics* 23, 1630-1661.

[24] Shimotsu, K. and P. C. B. Phillips (2005), "Exact Local Whittle Estimation of Fractional Integration", *Annals of Statistics*, 33, 1890-1933.

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

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

[27] Velasco, C. (2000). "Non-Gaussian Log Periodogram Regression", *Econometric Theory* 16, 44-79.

[28] Von Sachs, R. (1994). "Peak-Insensitive Nonparametric Spectrum Estimation", *Journal of Time Series Analysis* 15, 429-452.