JETEM

Journal of Economic Theory and Econometrics

The Journal of the Korean Econometric Society

( 계 량 경 제 학 보 )

September 2008

Volume 19. Issue 3. pp.1-23

# Log Periodogram Estimation with Nonstationary Process*

**Chang Sik Kim**†

**Abstract** This paper studies a new semiparametric estimation procedure of the memory parameter (*d*) in models of fractional integration. The procedure is called modified log periodogram regression and it is especially useful for nonstationary time series where $d \geq \frac{1}{2}$.

The modified estimator is shown to be consistent and asymptotically normally distributedwith variance $\frac{\pi^2}{24}$ under mild regularity conditions. Simulations reveal that the estimatorperforms well for all d $\geq \frac{1}{2}$.

**Keywords** Discrete Fourier transform, fractional Brownian motion, fractional integration, log periodogram regression, long memory, nonstationarity, semiparametric estimation

**JEL Classification** C22

* We would like to thank two anonymous referees for many helpful discussions and suggestions that greatly improved this paper. This work was supported by the Korea Research Foundation Grant funded by the Korean Government (MOEHRD), Basic Research Promotion Fund 2006-B00042.

† School of Economics, Sungkyunkwan University, 53 Myeongnyun-dong 3-ga Jongno-gu, Seoul, Korea 110-745 Phone : 02-760-0611, Fax : 02-744-5717 Email: skimcs@skku.edu

*Received November 7, 2007, Revised July 1, 2008, Accepted September 18, 2008*

# 1. Introduction

Statistical inference about the memory parameter $d$ of a fractional process has been an ongoing topic of research on which there is now an extensive literature. Most previous studies have concentrated on stationary processes with long memory parameter in the range $-\frac{1}{2} < d < \frac{1}{2}$, and there are now two main semiparametric methods for estimating the fractional parameter $d$ within this range - log periodogram (or LP) regression and local Whittle estimation. The popular method in practical work is log periodogram (LP) regression (Geweke and Porter-Hudak, 1983, and Robinson, 1995a). While many procedures for estimating $d$ involve numerical optimization methods (including local Whittle estimation- see Künsch, 1987, and Robinson, 1995b, Phillips and Shimotsu, 2004), LP regression is appealing because it involves familiar least squares regression and it treats short memory components in a nonparametric way.

Not surprisingly, therefore, there has been substantial previous interest in LP regression, in both theoretical and empirical work. Following early efforts to develop an asymptotic theory (Geweke and Porter-Hudak, 1983; Kashyap and Eom, 1988; Hassler, 1993), Robinson (1995b) provided a rigorous development under Gaussian assumptions and using Künsch's (1986) suggestion of trimming out frequencies immediately adjacent to the origin. Hurvich, Deo, and Brodsky (1998) later extended this result without omitting any low frequencies ordinates. Andrews and Guggenberger (2002) suggested a biased reduced log periodogram regression by eliminating the first and higher order of biases from original LP regression estimator.

Although most of the theoretical literature has focused on stationary long-memory processes, on the empirical side there have been several attempts to analyze apparently nonstationary macroeconomic time series in terms of fractional processes (see the review by Baillie, 1996, for a detailed discussion) and there is now a growing interest in the analysis of nonstationary series by fractional integration methods. For example, Gil-Alaña and Robinson (1997) tested the Nelson-Plosser's data (1982) and concluded that most of these series are nonstationary, many of them of fractional order. While interest has grown, a systematic analysis of nonstationary fractional processes for $d \geq \frac{1}{2}$ has been elusive and technical methods used in the analysis of the stationary case seem to be of limited usefulness for values of $d$ around unity. Some representation techniques developed in Phillips (1999) offer a new path forward in the analysis of the nonstationary case and this paper draws extensively on that approach in developing a theory of inference for LP regression for nonstationary processes.

In this paper, utilizing the form of the data generating process in the frequency domain discovered in Phillips (1999) for the nonstationary case, we study a modified form of LP regression that was suggested there and which is based on a simple approximation to the data generating mechanism. The new approach to estimating the fractional parameter when $d > \frac{1}{2}$ is called modified LP regression, because the periodogram ordinates are modified to take into account the correct form of the data generating process for the discrete Fourier transforms (dft's). The modification is simple and easy to make and involves no unknown parameters. After this modification is made, we are left with a new LP regression that is much easier to analyze and which has good asymptotic properties over a wide range of values of $d$.

The main advantage of modified LP regression is that it arises in a natural way from the exact representation of the dft of a fractional process, whereas the conventional formulation of LP regression is inspired only by a moment condition relating the spectra of the data and the short memory components near the origin. Moreover, as noted before, the modified estimation procedure is attractive in that it is basically nothing more than a least squares regression with a transformed dependent variable, and hence it is very easy to implement in practice. The finite sample performance of the modified estimator given in the present paper confirms that the procedure is reliable, and supports the nonparametric treatment of innovations in a data generating process.

The paper is organized as follows. The following section briefly reviews the representation of theory given in Phillips (1999) for the dft of a fractional process and sharpens some of those results for our use here. Section 3 defines the modified estimator and presents our main results. A finite sample performance of the modified estimator is given in section 4, and concluding remarks are made in section 5. Proofs and graphs are given in section 6.

## 2. The DFT of Fractionally Integrated Processes

This section briefly reviews some representation and approximation theory for the dft of nonstationary fractional processes based on Phillips (1999) and Kim and Phillips (2006). This theory will then be used in the construction of the modified LP regression estimator, following lines suggested in Phillips (1999).

A process $X_t$ of fractional order $d$ is usually defined in terms of a general model in the operator form

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

where $u_t$ is stationary with zero mean and continuous bounded spectral density $f_u(\lambda) > 0$ and $L$ is the lag operator for which $Lu_t = u_{t-1}$. If it is assumed that $u_j = 0$ for all $j \leq 0$, then (1) can be written in (finite order) expanded form in either autoregressive or moving average formats. More explicit conditions on $u_t$ $(t > 0)$ are given in the following condition.

**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 finite fourth moment $\mu_4$.*

The linear process error condition in (2) is general enough to cover a fairly wide class of weakly stationary processes. The summability condition ensures $|C(1)| < \infty$ and enables us to use the algebraic decomposition techniques of Phillips (1999) to develop a convenient representation of the dft of a fractionally integrated process. In fact, the linear process assumptions imply that the spectrum of $u_t$ is continuously differentiable at all frequencies, whereas much of the earlier work has made only local assumptions on spectrum. Expanding the inverse form of (1) we have

$$X_t = (1-L)^{-d} u_t = \sum_{k=0}^{t} \frac{(d)_k}{k!} u_{t-k},$$

where $(d)_k = \frac{\Gamma(d+k)}{\Gamma(d)}$ is Pochhammer's symbol for the forward factorial function. Then, defining the operator $D_n(L; d) = \sum_{k=0}^{n} \frac{(-d)_k}{k!} L^k$, and writing the dft of $a_t$ as $w_a(\lambda) = \frac{1}{\sqrt{2\pi n}} \sum_{t=1}^{n} a_t e^{it\lambda}$, the result below gives the exact representation of $w_x(\lambda)$ in terms of $w_u(\lambda)$, the dft of the error process $u_t$. The periodogram that will be used in the estimation of fractional parameter $d$ is just the norm of dft, therefore dft is a key element of inference in spectrum. We first look at the dft of fractionally integrated processes following the results in Phillips (1999) for later use.

**2.1 Lemma (Phillips, 1999)**

$$w_u(\lambda) = w_x(\lambda) D_n\left(e^{i\lambda}; d\right) + \frac{1}{\sqrt{2\pi n}}\left(\widetilde{X}_{\lambda 0}(d) - e^{in\lambda}\widetilde{X}_{\lambda n}(d)\right) \tag{3}$$

*where*

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

*and*

$$\widetilde{D}_{n\lambda}\left(e^{-i\lambda}L; d\right) = \sum_{p=0}^{n-1} \widetilde{d}_{\lambda p} e^{-ip\lambda} L^p \text{ and } \widetilde{d}_{\lambda p} = \sum_{k=p+1}^{n} \frac{(-d)_k}{k!} e^{ik\lambda}.$$

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, the expression in (3) becomes

$$\begin{aligned} w_u(\lambda) &= w_x(\lambda) D_n\left(e^{i\lambda}; d\right) - \frac{e^{in\lambda}}{\sqrt{2\pi n}} \widetilde{D}_{n\lambda}\left(e^{-i\lambda}L; d\right) X_n \\ &= w_x(\lambda) D_n\left(e^{i\lambda}; d\right) - \frac{1}{\sqrt{2\pi n}} e^{in\lambda} \widetilde{X}_{\lambda n}(d). \end{aligned} \tag{4}$$

Equation (4) 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 should be taken into account in analyzing the asymptotic behavior of $w_x(\lambda)$ and the periodogram $I_x(\lambda)$. LP regression (Geweke, Porter Hudak, 1983, and Kashyap and Eom, 1988) involves a linear in logarithms regression between the periodogram $I_x(\lambda) = w_x(\lambda) w_x(\lambda)^*$ and frequency $\lambda$ without considering the term $\widetilde{X}_{\lambda n}(d)$. However, as shown in Phillips (1999) and in the following lemma, the order of magnitude of the term $\widetilde{X}_{\lambda n}(d)$ gets larger as the value of $d$ increases so that neglecting it has greater consequences.

Thus, to avoid this difficulty and remain faithful to the data generating mechanism it is important to work directly from (4) or a satisfactory approximation to it. Our first objective is to characterize the behavior of $\widetilde{X}_{\lambda n}(d)$ for those frequencies $\lambda = \lambda_s = \frac{2\pi s}{n} \to 0$ adjacent to the origin that are used in LP regression. The following lemma gives the limit behavior of $\widetilde{X}_{\lambda n}(d)$ when $X_t$ is a nonstationary fractional process.

**2.2 Lemma** *For* $\lambda = \lambda_s = \frac{2\pi s}{n} \to 0$ *as* $n \to \infty$,

(a) *If* $X_t$ *follows* (1) *with* $\frac{1}{2} < d < 1$ *and* $\frac{n^\alpha}{s} \to 0$ *for some* $\alpha \in (\frac{1}{2}, 1)$, *or* $1 < d < 2$, *and* $u_t$ *is defined in* (2), *then*

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

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

$$\frac{1}{n}\frac{\widetilde{X}_{\lambda_s n}(d)}{\sqrt{n}} = -\frac{e^{i\lambda_s}}{n}\frac{X_n}{\sqrt{n}}. \tag{5}$$

This lemma is based on theorem 3.2 of Phillips (1999) for $\frac{1}{2} < d < 1$, and lemma 2.3 of Kim and Phillips (2006) for $1 < d < 2$. The formula (5) is particularly simple at $d = 1$, and it is worth noting that the restriction on the range of $s$ in the case of $\frac{1}{2} < d < 1$ is relaxed for $d > 1$. Observe that the leading term in the asymptotic approximation of $\frac{1}{n^d}\frac{\widetilde{X}_{\lambda_s n}(d)}{\sqrt{n}}$ involves the last sample observation $X_n$, which is the same as when $d \in (\frac{1}{2}, 1)$.

## 3. Modified Log-Periodogram Regression

We first give an idea for the construction of LP regression. The spectral density of the fractional process $X_t$ can be written as

$$f_x(\lambda) = |1 - e^{i\lambda}|^{-2d} f_u(\lambda)$$

from the definition of fractional process $X_t$. Taking log gives us

$$\ln f_x(\lambda) = -2d \ln |1 - e^{i\lambda}| + \ln f_u(\lambda)$$

which leads to the following linear least squares regression equation,

$$\ln I_x(\lambda) = f_u(0) - 2d \ln |1 - e^{i\lambda}| + \ln (f_u(\lambda)/f_u(0)) + \ln (I_x(\lambda)/f_x(\lambda))$$

Here $\lambda_s$ is defined as frequencies $\lambda_s = \frac{2\pi s}{n}, \ \ s = 1, ..., m$, and $m$ is the number of frequencies included in the regression. As we can see, regression error in this regression equation is $\ln (I_x(\lambda)/f_x(\lambda))$, which generates difficulties in the analysis of LP regression. Clearly, nonlinearity of logarithmic functions causes difficulties and the regression error $(I_x(\lambda)/f_x(\lambda))$ are not asymptotically uncorrelated even in the fixed number of frequencies, which shows different properties from weakly stationary case.

Therefore, we take a different approach in the analysis of LP regression in the nonstationary case. We first give some exact representation of periodogram $I_x(\lambda)$ at frequency $\lambda$ as in lemma 2.1, and then we specify the component of this representation as in lemma 2.2. We then analyze these components of the periodogram of fractional process $X_t$ rather than working with $I_x(\lambda)$ directly.

From (4), we have

$$w_x(\lambda_s) = D_n\left(e^{i\lambda_s};d\right)^{-1}\left[w_u(\lambda_s) + \frac{1}{\sqrt{2\pi n}}\widetilde{X}_{\lambda_s n}(d)\right]. \tag{6}$$

Equation (6) holds for all frequencies without restriction and can be rewritten as

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

which effectively replaces the dft $w_x(\lambda_s)$ by the modified quantity

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

giving the exact relation

$$\upsilon_x(\lambda_s;d) = D_n\left(e^{i\lambda_s};d\right)^{-1} w_u(\lambda_s). \tag{7}$$

Using the approximation for $\widetilde{X}_{\lambda n}(d)$ suggested in lemma 2.2, we have

$$\frac{\widetilde{X}_{\lambda_s n}(d)}{\sqrt{2\pi n}} = -\frac{e^{i\lambda_s}}{(1-e^{i\lambda_s})^{1-d}}\frac{X_n}{\sqrt{2\pi n}} + o_p\left(\frac{s^{d-1}}{n^d}\right), \tag{8}$$

under $\frac{s}{n} \to 0$ as $n \to \infty$. Moreover, from lemma 3.1 of Phillips (1999), we have the following simplified form for the sinusoidal polynomial factor

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

for $\lambda = \lambda_s = \frac{2\pi s}{n} \to 0$ and $s \to \infty$ as $n \to \infty$, giving the approximation

$$D_n\left(e^{i\lambda_s};d\right)^{-1} = \left(1-e^{i\lambda_s}\right)^{-d} + O\left(\frac{1}{n^d s}\right). \tag{9}$$

From (8) and (9), the modified quantity $\upsilon_x(\lambda_s;d)$ in (3) can be approximated by the following observable quantity at frequency $\lambda_s$

$$\begin{aligned}
\upsilon_x(\lambda_s;d) &= w_x(\lambda_s) - D_n\left(e^{i\lambda_s};d\right)^{-1}\frac{1}{\sqrt{2\pi n}}\widetilde{X}_{\lambda n}(d) \\
&= w_x(\lambda_s) - \left[\left(1-e^{i\lambda_s}\right)^d + O\left(\frac{1}{n^d s}\right)\right]^{-1}\left[-\frac{e^{i\lambda_s}}{(1-e^{i\lambda_s})^{1-d}}\frac{X_n}{\sqrt{2\pi n}} + o_p\left(\frac{s^{d-1}}{n^d}\right)\right] \\
&= w_x(\lambda_s) - \left(1-e^{i\lambda_s}\right)^{-d}\left(-\frac{e^{i\lambda_s}}{(1-e^{i\lambda_s})^{1-d}}\frac{X_n}{\sqrt{2\pi n}}\right) + o_p\left(\frac{1}{(1-e^{i\lambda_s})^d}\frac{s^{d-1}}{n^d}\right) \\
&= w_x(\lambda_s) + \frac{e^{i\lambda_s}}{(1-e^{i\lambda_s})}\frac{X_n}{\sqrt{2\pi n}} + o_p\left(\frac{1}{s}\right) \qquad (10) \\
&=: \upsilon_x(\lambda_s) + o_p\left(\frac{1}{s}\right).
\end{aligned}$$

The quantity

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

is called the modified dft. It differs from the dft $w_x(\lambda_s)$ by a term involving the complex exponential function $\frac{e^{i\lambda_s}}{(1-e^{i\lambda_s})}$ with coefficient $\frac{X_n}{\sqrt{2\pi n}}$.

Since the latter function is proportional to the dft of a linear time trend, the modified dft $\upsilon_x(\lambda_s)$ has the interesting property of being invariant to a linear trend, which is of some consequence in applications. In particular, suppose the data generating process for $X_t$ were of the form

$$X_t = \beta t + X_t^0,$$

where $X_t^0$ follows (1), i.e. $(1-L)^d X_t^0 = u_t$ with the same stationary input sequence $u_t$ as in (1). Then,

$$\begin{aligned} w_x(\lambda_s) &= \beta w_t(\lambda_s) + w_{x^0}(\lambda_s) \\ &= \beta\frac{e^{i\lambda_s}}{e^{i\lambda_s}-1}\sqrt{\frac{n}{2\pi}} + w_{x^0}(\lambda_s), \end{aligned}$$

as can be verified by direct computation (see also Lemma A of Corbae, Ouliaris and Phillips, 2002). It follows that

$$\begin{aligned} \upsilon_x(\lambda_s) &= \beta\frac{e^{i\lambda_s}}{e^{i\lambda_s}-1}\sqrt{\frac{n}{2\pi}} + w_{x^0}(\lambda_s) + \frac{e^{i\lambda_s}}{(1-e^{i\lambda_s})}\frac{X_n}{\sqrt{2\pi n}} \\ &= w_{x^0}(\lambda_s) + \frac{e^{i\lambda_s}}{(1-e^{i\lambda_s})}\frac{X_n-\beta n}{\sqrt{2\pi n}} \\ &= w_{x^0}(\lambda_s) + \frac{e^{i\lambda_s}}{(1-e^{i\lambda_s})}\frac{X_n^0}{\sqrt{2\pi n}} \\ &= \upsilon_{x^0}(\lambda_s), \end{aligned}$$

which is invariant to $\beta$. Thus, regressions in the frequency domain that involve the modified dft $\upsilon_x(\lambda_s)$ turn out to be invariant to the coefficient in a linear trend in the data.

Taking the squared modulus in (7), we have the exact periodogram relation

$$I_{\upsilon_x}(\lambda_s; d) = \left|D_n\left(e^{i\lambda_s}; d\right)\right|^{-2} I_u(\lambda_s), \tag{11}$$

with $I_{\upsilon_x}(\lambda_s; d) = \upsilon_x(\lambda_s; d)\,\upsilon_x(\lambda_s; d)^*$. Setting $I_{\upsilon_x}(\lambda_s) = \upsilon_x(\lambda_s)\,\upsilon_x(\lambda_s)^*$, we have

$$\begin{aligned} I_{\upsilon_x}(\lambda_s; d) &= \left|\upsilon_x(\lambda_s) + o_p\left(\frac{1}{s}\right)\right|^2 \\ &= I_{\upsilon_x}(\lambda_s)\left|1 + \upsilon_x(\lambda_s)^{-1} o_p\left(\frac{1}{s}\right)\right|^2 \\ &= I_{\upsilon_x}(\lambda_s)\left|1 + o_p\left(\frac{1}{n^d}\right)\right|^2. \end{aligned} \tag{12}$$

Taking logarithms of (11) leads to

$$\log I_{\upsilon_x}(\lambda_s; d) = c_1 - 2d \log \left|1 - e^{i\lambda_s}\right| + a(\lambda_s) \tag{13}$$

where

$$a(\lambda_s) = \log (I_u(\lambda_s) / f_u(0)), \text{ and } c_1 = \log f_u(0).$$

Next, we approximate $I_{\upsilon_x}(\lambda_s; d)$ by $I_{\upsilon_x}(\lambda_s) = \upsilon_x(\lambda_s)\upsilon_x(\lambda_s)^*$ using the approximation $\upsilon_x(\lambda_s)$ for $\upsilon_x(\lambda_s; d)$. That is, using the relation in (12), we propose to estimate the memory parameter $d$ in (13) by a linear regression in which $I_{\upsilon_x}(\lambda_s; d)$ is replaced by $I_{\upsilon_x}(\lambda_s)$, following the suggestion in Phillips (1999, section 4). The resulting estimator is called modified LP regression. In particular, the modified LP regression estimator of $d$ is based on linear least squares regression of $\log I_{\upsilon_x}(\lambda_s)$ on $\log\left|1 - e^{i\lambda_s}\right|$ over frequencies $\{\lambda_s, s = 1, ..., m\}$, and has the explicit form

$$2\widehat{d} = -\left[\sum_{s=1}^{m} x_s^2\right]^{-1}\left[\sum_{s=1}^{m} x_s \log I_{\upsilon_x}\right], \tag{14}$$

where $x_s = \log\left|1 - e^{i\lambda_s}\right| - \overline{\log(|1 - e^{i\lambda}|)}$, and $\overline{\log(|1 - e^{i\lambda}|)} = \frac{1}{m}\sum_{s=1}^{m} \log\left(\left|1 - e^{i\lambda_s}\right|\right)$. From (14) we deduce that

$$2(\widehat{d} - d) = -\left[\sum_{s=1}^{m} x_s^2\right]^{-1}\left[\sum_{s=1}^{m} x_s b(\lambda_s)\right]. \tag{15}$$

where

$$b(\lambda_s) = a(\lambda_s) + \log \frac{I_{\upsilon_x}(\lambda_s; d)}{I_{\upsilon_x}(\lambda_s)}.$$

The modified LP regression has a very similar form as the traditional LP regression estimator. The $x_s$ is just a deterministic regressor that is the same the LP regression. The only difference will be the periodogram $I_{\upsilon_x}$ that differs from the dft $w_x(\lambda_s)$ by a term involving the deterministic exponential function $\frac{e^{i\lambda_s}}{(1-e^{i\lambda_s})}$ with coefficient $\frac{X_n}{\sqrt{2\pi n}}$ that includes the last observation. We can notice that the information of the last observation $X_n$ is crucial in the modified dft and hence the estimation of fractional parameter $d$ in the nonstationary case.

Kim and Phillips (2006) showed that the conventional LP regression estimator when $1 < d < 2$ is inconsistent and has a probability limit equal to unity. The following theorem shows that the modified LP regression estimator is consistent under mild conditions and without any distributional specification like Gaussianity.

**3.1 Theorem (Consistency)** *Let $X_t$ follow* (1) *with* $1 < d < 2$, $u_t$ *satisfy* (2) *and* $\varepsilon_t$ *fulfill Cramér's condition, i.e.*

$$\exists\, \delta > 0,\ b > 0,\ \forall\, |t| > b\ \ |\mathbf{E} \exp(it\varepsilon_t)| \leq 1 - \delta, \tag{16}$$

*and*

$$\int |\mathbf{E} \exp(it\varepsilon_t)|^p \, dt < \infty \text{ for some integer } p \geq 1. \tag{17}$$

*Then, if* $\frac{m \log m}{n^{\frac{1}{2}-\delta}} \to 0$, *we have* $\widehat{d} \xrightarrow{p} d$.

Note that the modified LP estimator involves only linear least squares regression and therefore has the same appeal of convenient computation and nonparametric treatment of the short memory component as the original LP estimator. From the computational standpoint, all that is needed is the additional frequency domain data correction given by (10). Also, no distributional assumptions are needed to validate the asymptotic theory. The main reason for the latter is that the dft representations effectively replace the dft of the fractional process $X_t$ by the dft of the short memory component $u_t$ and the final sample observation $X_n$. These two components are much easier to handle because their asymptotic behavior is already known.

Another simplification is attained by way of the BN decomposition used in Phillips and Solo (1992). In particular, it is shown in the Appendix that the following relation holds

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

under the condition $\frac{m^{\frac{3}{2}} \log m}{n^{\frac{1}{2}-\delta}} \to 0$ for any $\delta > 0$. The advantage of this additional reduction is that we can make use of the exact uncorrelatedness of $I_\varepsilon(\lambda_s)$ and $I_\varepsilon(\lambda_p)$ at different frequencies $s \neq p$, and we can therefore utilize the more favorable properties of the quantities $I_\varepsilon(\lambda_s)$ in place of $I_u(\lambda_s)$.

Some additional technical difficulties in LP regression arise from the presence of non-linear transforms of the periodogram, so that the favorable properties of $I_\varepsilon(\lambda_s)$ can not be used directly. Instead, we follow Velasco (2000)'s approach and utilize the asymptotic expansions in Janas and Von Sachs (1995). Using this expansion, Velasco (2000) recently derived the consistency of the pooled LP estimator by trimming out low frequencies in the stationary case ($0 < d < \frac{1}{2}$) without relying upon Gaussianity.

Next, we turn to the limit distribution of the modified LP estimator. Robinson (1995b) established a central limit theorem(CLT) for linear combinations of the ordinates $I_\varepsilon(\lambda_s)$'s using a martingale CLT. As noted above, however, the non-linear periodogram transforms that appear in LP regression make it helpful to use stronger assumptions in obtaining a limit distribution theory. One convenient facilitating assumption is Gaussianity of $\varepsilon_t$, as is assumed in much of the previous literature on LP regression, which implies the normality of the dft $w_\varepsilon(\lambda_s)$. Under normality, we are able to derive the limit theory for the modified LP estimator in a straightforward way, as seen in the following theorem.

A very different approach that avoids the assumption of Gaussianity was developed recently by Phillips (2007) and relies on a strong approximation and embedding argument for the dft's. This approach requires only the existence of higher moments of $\varepsilon_t$, and a slightly stronger restriction on the number of periodogram ordinates $m$ to derive a limit distribution theory.

**3.2 Theorem (Asymptotic Normality)** *Let $X_t$ follow* (1) *with $1 < d < 2$ and $u_t$ satisfy* (2). *Assume if $\frac{m^5}{n^4} + \frac{n^\alpha}{m} + \frac{m^{\frac{3}{2}}}{n^d} \to 0$ for small $\alpha > 0$ as $n \to \infty$, and $\varepsilon_t$ is Gaussian Then*

$$\sqrt{m}\left(\widehat{d} - d\right) \xrightarrow{d} N\left(0, \frac{\pi^2}{24}\right). \tag{18}$$

The conditions of Theorem 3.2 involve Gaussianity on $\varepsilon_t$ and stronger restrictions on $m$ compared to those in Theorem 3.1. In fact, replacing the Gaussianity assumption by a weaker moment condition on $\varepsilon_t$ with a more restrictive choice of $m$ can be done by following the approach of strong approximations developed in Phillips (2007). The condition on $m$ will be highly restrictive than those established in the previous literature under Gaussian errors, for example, Hurvich et al (1998) because of technical difficulties without Gaussianity.

The limit theory for the modified estimator for $\frac{1}{2} < d < 1$ can also be established in a similar way for $1 < d < 2$, and it therefore can be skipped here to save space. Minor trimming is necessary to develop the asymptotic theory for $\frac{1}{2} < d < 1$ because of a technical difficulty, it does not affect the use of modified log periodogram regression estimator in practice since minor trimming does not change the estimate. From a practical point of view, the modified estimator has more important advantages than simply the relaxation of restrictions in LP regression. The inconsistency of LP regression when $d > 1$ makes it desirable to have a method that has satisfactory properties over a broader range of values for the memory parameter, especially one that includes values of $d > 1$ such as $1 < d < 2$ as well as $\frac{1}{2} < d < 1$. One rarely has any prior information about the fractional parameter $d$ in empirical work and it is therefore useful to have a procedure that works reliably over a broad band of realistic values of $d$. Test results from Robinson and Gil-Alaña (1996) show that the fractional parameters of most series in the Nelson-Plosser data set, for instance, appear to lie in the range between $\frac{1}{2}$ and 2. Empirical work in Phillips (1998) for interest rates, real interest rates and inflation further confirms the importance of this range of values for macroeconomic time series.

Velasco (1999b) showed consistency of the LP regression estimator for $\frac{1}{2} < d < 1$ under Gaussianity and by trimming out low frequencies ordinates up to some value $\ell$, which goes to infinity more slowly than $m$. Theorem 3.3 shows that the modified estimator has asymptotic normality can be obtained under additional restrictions by trimming out low frequencies. This is necessary for avoiding the technical difficulties for asymptotic normality, but in practice trimming is not necessary since it does not affect the performances of the modified LP estimator as confirmed by simulation. In particular, we define a trimmed version of the modified LP regression estimator as

$$2\widehat{d}_{\ell} = -\left[\sum_{s=\ell}^{m} x_s^2\right]^{-1}\left[\sum_{s=\ell}^{m} x_s \log I_{v_x}(\lambda_s)\right], \tag{19}$$

where frequencies $s = 1, ..., \ell - 1$ have been trimmed out of the regression. We present the following asymptotic results of the modified LP estimation for $\frac{1}{2} < d < 1$ without proof to save space.

**3.3 Theorem** *Let $X_t$ follow* (1) *with $\frac{1}{2} < d < 1$ and $u_t$ satisfy* (2). *Assume either condition (i) or (ii) below:*

*(i) if $\frac{m^5}{n^4} + \frac{n^\alpha}{m} + \frac{m^{d+\frac{1}{2}}}{n^d} \to 0$ for small $\alpha > 0$ as $n \to \infty$, and $\varepsilon_t$ is Gaussian; or*

*(ii) if $\frac{m^{\frac{3}{2}}}{n} + \frac{m^{\frac{3}{2}} \log m}{n^{\frac{1}{2}-\frac{1}{p}}} + \frac{m^{d+\frac{1}{2}}}{n^d} \to 0$ for $p > 2$, and $\mathbf{E}|\varepsilon_t|^q < \infty$ for some $q > 2p > 4$;*

*and if $\frac{\sqrt{n}}{\ell} + \frac{\ell}{m} \to 0$. Then*

$$\sqrt{m}\left(\widehat{d}_{\ell} - d\right) \xrightarrow{d} N\left(0, \frac{\pi^2}{24}\right). \tag{20}$$

Thus, the limit theory for the modified estimator seems favorable for $\frac{1}{2} < d < 1$ as well as for $1 < d < 2$ and can be established without assuming Gaussianity. The use of the modified estimator can also be expected to be satisfactory even in the stationary case. Notice that the order of the correction term $\frac{e^{i\lambda_s}}{(1-e^{i\lambda_s})}\frac{X_n}{\sqrt{2\pi n}}$ is $\frac{\sqrt{n}}{s}$, since $X_n$ itself is $O_p(1)$ in the stationary case. Therefore, the correction term is negligible for sufficiently large $s$, and hence the modified estimator may be expected to work well even for $d \leq \frac{1}{2}$. We do not provide the proof for this argument, but modified estimator appears to work as well as the original LP estimator in the $0 < d \leq \frac{1}{2}$ case based on simulations, when $m$ is sufficiently large.

As recently shown in Phillips (2007), the LP regression estimator for the unit root process is consistent and has a mixed normal asymptotic distribution. However, the modified LP estimator in this case is asymptotically normal with the same variance in (20), not mixed normal. When $d = \frac{1}{2}$, Liu (1998) obtained the asymptotic behavior of fractionally integrated processes with $d = \frac{1}{2}$.

## 4. Finite Sample Performance of the Modified Log Periodogram Regression

The following subsections corroborate the theoretical findings of the present paper, and show the finite sample performance of the modified LP estimator. We tried four values of $d$ ($d = 0.8$, 1.25, 1.45, 1.8), varying the sample size $n$ ($n = 256$, 512) and the number of ordinates $m$ (setting $m = n^{0.4}, \dots, m = n^{0.7}$). White noise was used for the data generating process, and 1,000 replications were run in the simulation. Table 1 summarizes the results for the modified LP regression, showing the mean and standard deviation of the estimates.

**Table 1. Mean and Standard Deviation for the Modified Log Periodogram Estimator**

(i) ARMA(0,0)

| $d$ | $n$ | $m = n\hat{}0.4$ | $m = n\hat{}0.5$ | $m = n\hat{}0.6$ | $m = n\hat{}0.7$ |
|---|---|---|---|---|---|
| 0.8 | 256 | 0.8102 (0.3207) | 0.8113 (0.2138) | 0.7917 (0.1439) | 0.7910 (0.1033) |
| | 512 | 0.8108 (0.2612) | 0.7964 (0.1701) | 0.7995 (0.1135) | 0.7919 (0.0835) |
| 1.25 | 256 | 1.2630 (0.3205) | 1.2655 (0.2093) | 1.2541 (0.1497) | 1.2402 (0.1099) |
| | 512 | 1.2794 (0.2566) | 1.2541 (0.1671) | 1.2590 (0.1160) | 1.2445 (0.0819) |
| 1.45 | 256 | 1.4873 (0.3038) | 1.4826 (0.2116) | 1.4847 (0.1513) | 1.4548 (0.1108) |
| | 512 | 1.4737(0.2619) | 1.4901 (0.1738) | 1.4750 (0.1187) | 1.4541 (0.0840) |
| 1.8 | 256 | 1.8133 (0.3115) | 1.8236 (0.2050) | 1.8141 (0.1557) | 1.8004 (0.1098) |
| | 512 | 1.8296 (0.2513) | 1.8263 (0.1751) | 1.8235 (0.1127) | 1.8064 (0.0854) |

(ii). AR(1)

| $d$ | $n$ | $m = n\hat{}0.4$ | $m = n\hat{}0.5$ | $m = n\hat{}0.6$ | $m = n\hat{}0.7$ |
|---|---|---|---|---|---|
| 0.8 | 256 | 0.8080 (0.2934) | 0.8155 (0.2136) | 0.8289 (0.1509) | 0.8758 (0.1062) |
| | 512 | 0.7968 (0.2542) | 0.8050 (0.1685) | 0.8227 (0.1117) | 0.8503 (0.0785) |
| 1.25 | 256 | 1.2620 (0.3154) | 1.2756 (0.2107) | 1.2915 (0.1538) | 1.3273 (0.1074) |
| | 512 | 1.2728 (0.2565) | 1.2778 (0.1652) | 1.2812 (0.1178) | 1.3073 (0.0812) |
| 1.45 | 256 | 1.4961 (0.3164) | 1.4970 (0.2110) | 1.5045 (0.1499) | 1.5336 (0.1048) |
| | 512 | 1.4726 (0.2553) | 1.4907 (0.1769) | 1.4991 (0.1211) | 1.5119 (0.0841) |
| 1.8 | 256 | 1.8444 (0.3085) | 1.8330 (0.2072) | 1.8492 (0.1472) | 1.8768 (0.1007) |
| | 512 | 1.8443 (0.2634) | 1.8353 (0.1711) | 1.8452 (0.1212) | 1.8592 (0.0840) |

(iii). MA(1)

| $d$ | $n$ | $m = n\hat{}0.4$ | $m = n\hat{}0.5$ | $m = n\hat{}0.6$ | $m = n\hat{}0.7$ |
|---|---|---|---|---|---|
| 0.8 | 256 | 0.7942 (0.3186) | 0.7965 (0.1982) | 0.7616 (0.1524) | 0.7008 (0.1033) |
| | 512 | 0.8036 (0.2568) | 0.7999 (0.1713) | 0.7776 (0.1186) | 0.7348 (0.0782) |
| 1.25 | 256 | 1.2670 (0.3184) | 1.2356 (0.2228) | 1.2218 (0.1512) | 1.1544 (0.1052) |
| | 512 | 1.2681 (0.2579) | 1.2504 (0.1699) | 1.2349 (0.1104) | 1.1862 (0.0820) |
| 1.45 | 256 | 1.4997 (0.3078) | 1.4682 (0.2123) | 1.4262 (0.1531) | 1.3669 (0.1095) |
| | 512 | 1.4811 (0.2648) | 1.4779 (0.1762) | 1.4548 (0.1207) | 1.3982 (0.0830) |
| 1.8 | 256 | 1.8253 (0.3004) | 1.8180 (0.2100) | 1.7901 (0.1524) | 1.7188 (0.1128) |
| | 512 | 1.8290 (0.2598) | 1.8208 (0.1741) | 1.8016 (0.1198) | 1.7534 (0.0887) |

Asymptotic standard deviation for each case;

$n = 256 : 0.2115\ (m = n\hat{}0.4)\,, 0.1603\ (m = n\hat{}0.5)\,, 0.1214\ (m = n\hat{}0.6)\,, 0.0920\ (m = n\hat{}0.7)$

$n = 512 : 0.1841\ (m = n\hat{}0.4)\,, 0.1348\ (m = n\hat{}0.5)\,, 0.0986\ (m = n\hat{}0.6)\,, 0.0722\ (m = n\hat{}0.7)$

Velasco(1999b) show rather obvious inconsistency behavior for both the LP estimator and the Gaussian Semiparametric method (e.g. Robinson, 1995a) for the $d = 1.8$ case. Velasco shows that these two semiparametric estimates appear to concentrate to a value close to 1, although their distributions also appear to have a long right tail.

Table 1 indicates a striking difference between the modified estimator and the original LP estimator shown in the previous literature. The average value of the modified estimator in most cases is very close to its true value, as the theory in the previous section predicts. As $m$ gets larger, the convergence rate is faster, as expected from theorem 3.2-3. When the data generating process is white noise, the modified estimator performs extremely well and the asymptotic theory seems relevant even for the smaller sample size. Finite sample performance in the $AR(1)$ and $MA(1)$ cases is not as good as in the $ARMA(0,0)$ case, showing that biases in the $AR(1)$ case tend to be positive, whereas negative biases arise in the $MA(1)$ case, corresponding to previous findings in Agiakloglou, Newbold, and Wohar (1993) for LP regression. Overall, these results show that the modified LP estimator performs fairly well in finite samples. When $d = 0.8$, the modified estimator works better than the original LP estimator. In fact, the modified estimator has less variance and is less biased in most cases for different $m$ and $n$.

## 5. Concluding Remarks

Utilizing the exact representation of the dft of a nonstationary fractionally integrated process, this paper analyses the properties of a new estimation method called modified LP regression.

The central idea of the method is simple. From the representation of the dft, we find that the inconsistency of LP regression (when $d > 1$) is due to an omitted term. We need only include this term in the LP regression to obtain a new consistent estimator with the same limit distribution as that of the original LP regression estimator in the stationary case. With this modification, the new estimator has good asymptotic behavior over a wide band of values of the memory parameter and this favorable performance carries over to finite samples as well.

The limit theory of the modified LP estimator is obtained and compared to those in previous studies of LP regression. Like LP regression, modified LP regression is appealing because it does not require numerical optimization procedures and retains its validity for a wide class of short memory process inputs. The simulations in section 4 show that the finite sample performance of the modified LP estimator seems to be satisfactory and reliable for a range of different short memory inputs.

A remaining issue in our approach, like that of LP regression, is the selection of the number of ordinates, $m$, included in the regression. For the Gaussian stationary long memory time series case, Hurvich, Deo, and Brodsky (1998) calculated the optimal rate, $m = Cn^{\frac{4}{5}}$, by minimizing the asymptotic mean squared error of the LP regression estimator. Recently Hurvich and Deo (1999) suggested an estimation method for the constant $C$ in this optimal rate. However, as it presently stands, this rule has not been justified for nonstationary processes or for the modified LP regression. Since the choice of $m$ is important in the use of our methods in empirical research, further work on this issue now seems to be important.

## 6. Appendix

The following three lemmas are needed in the proofs of theorems 3.1 and 3.2. We first give the spectral form of the BN decomposition from Phillips and Solo (1992), which enables us to work with i.i.d. processes instead of general linear processes. As in Phillips and Solo (1992), we 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

$$w_u(\lambda_s) = C\left(e^{i\lambda_s}\right) w_\varepsilon(\lambda_s) + \frac{1}{\sqrt{2\pi n}}\left(\varepsilon_{\lambda_s 0} - \varepsilon_{\lambda_s n}\right), \tag{21}$$

with

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

and where $\widetilde{c}_{j\lambda_s} = \sum_{k=j+1}^{\infty} c_k e^{i\lambda_s k}$. The following lemma, which is needed in the modified LP regression application, shows that the second component in (21) is negligible uniformly over $s$.

**6.1 Lemma** *If the error conditions in* (2) *hold,* $\max_s \left|\frac{1}{n^\delta}\varepsilon_{\lambda_s n}\right| \xrightarrow{p} 0$ *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,

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

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

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

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 following lemma can be established in the same way as in Kim and Phillips (2006), and here skipped to save space.

**6.2 Lemma** *Under the error conditions in* (2) *and* $\frac{m^{\frac{3}{2}}\log m}{n^{\frac{1}{2}-\delta}} \to 0$ *for any* $\delta > 0$, *we have*

$$\frac{1}{\sqrt{m}}\sum_{s=1}^{m} x_s \ln I_u(\lambda_s) = \frac{1}{\sqrt{m}}\sum_{s=1}^{m} x_s \ln I_\varepsilon(\lambda_s) + o_p(1). \tag{22}$$

Next, we give a lemma which enables us to calculate the moments of the log function of periodogram, which is needed for the proof of Theorem 3.1. We will use the following lemma, which is a slightly modified version of Lemma A.1 in Janas and Von Sachs (1995). The proof of this lemma is given in Janas and Von Sachs (1995).

**6.3 Lemma** *Assume that the iid sequence* $\varepsilon_t$ *satisfies Cramér's condition* (16) *and condition* (17) *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* $\lambda_j$,

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

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

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

**6.4. Proof of theorem 3.1** From (15), we have

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

where

$$b(\lambda_s) = a(\lambda_s) + \log\frac{I_{v_x}(\lambda_s; d)}{I_{v_x}(\lambda_s)} = \log\left(\frac{I_u(\lambda_s)}{f_u(0)}\right) + \log\frac{I_{v_x}(\lambda_s; d)}{I_{v_x}(\lambda_s)}.$$

We need to show that

$$\left[\sum_{s=1}^{m} x_s^2\right]^{-1}\left[\sum_{s=1}^{m} x_s \log\left(\frac{I_u(\lambda_s)}{f_u(0)}\right)\right] \xrightarrow{p} 0, \tag{23}$$

and

$$\left[\sum_{s=1}^{m} x_s^2\right]^{-1}\left[\sum_{s=1}^{m} x_s \log\frac{I_{v_x}(\lambda_s; d)}{I_{v_x}(\lambda_s)}\right] \xrightarrow{p} 0 \tag{24}$$

for the consistency of the modified LP regression estimator. (24) can be easily deduced from the result (30) in the following proof of theorem 3.2. In view of the BN decomposition in (22), we only have to show

$$\frac{1}{m}\sum_{s=1}^{m} x_s \ln I_\varepsilon(\lambda_s) \xrightarrow{p} 0, \tag{25}$$

for (23). To show (25), we evaluate the first two moments of the log function of periodogram $I_\varepsilon(\lambda_s)$. By lemma 6.3, we have

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

The variance term can be written as

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

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

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

by the fact that $\frac{1}{m}\sum_{s=1}^{m}|x_s| = O(1)$, which is given in Robinson (1995b), and the (iii) in the lemma 6.3. Therefore, we have

$$\mathbf{Var}\left[\frac{1}{m}\sum_{s=1}^{m} x_s \ln I_\varepsilon(\lambda_s)\right] = o_p(1),$$

which implies (25). Thus, $\widehat{d} - d = o_p(1)$. ■

**6.5. Proof of theorem 3.2** We first prove the asymptotic normality of the modified estimator under condition (i). As before, from (15) we have

$$\widehat{d} - d = -\frac{1}{2}\left[\sum_{s=1}^{m} x_s^2\right]^{-1}\left[\sum_{s=1}^{m} x_s\left\{\log\left(\frac{I_u(\lambda_s)}{f_u(0)}\right) + \log\frac{I_{v_x}(\lambda_s; d)}{I_{v_x}(\lambda_s)}\right\}\right].$$

We need to show a CLT for

$$\left[\frac{1}{m}\sum_{s=1}^{m} x_s^2\right]^{-1}\left[\frac{1}{\sqrt{m}}\sum_{s=1}^{m} x_s \log\left(I_u(\lambda_s)/f_u(0)\right)\right], \tag{26}$$

and prove that the remainder

$$\left[\frac{1}{m}\sum_{s=1}^{m} x_s^2\right]^{-1}\left[\frac{1}{\sqrt{m}}\sum_{s=1}^{m} x_s \log\frac{I_{v_x}(\lambda_s; d)}{I_{v_x}(\lambda_s)}\right] \tag{27}$$

can be neglected. From (12), we have

$$\begin{aligned}\left|\frac{1}{\sqrt{m}}\sum_{s=1}^{m}x_s\log\frac{I_{\upsilon_x}(\lambda_s;d)}{I_{\upsilon_x}(\lambda_s)}\right| &= \frac{1}{\sqrt{m}}\sum_{s=1}^{m}|x_s|\left|\log\frac{I_{\upsilon_x}(\lambda_s;d)}{I_{\upsilon_x}(\lambda_s)}\right| \\ &= 2\frac{1}{\sqrt{m}}\sum_{s=1}^{m}|x_s|\left|\log\left|1+\upsilon_x(\lambda_s)^{-1}o_p\left(\frac{1}{s}\right)\right|\right| \\ &\le 2\frac{1}{m}\sum_{s=1}^{m}|x_s|\sqrt{m}\sup_s\log\left|1+\upsilon_x(\lambda_s)^{-1}o_p\left(\frac{1}{s}\right)\right|. \quad (28)\end{aligned}$$

From the proof of lemma 6.2, we only need to show that

$$\sqrt{m}\sup_s\left|\upsilon_x(\lambda_s)^{-1}o_p\left(\frac{1}{s}\right)\right|\xrightarrow{p}0, \quad (29)$$

for (27) to be negligible. Observe that

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

and

$$\sup_s\left|\sqrt{m}\left[\upsilon_x(\lambda_s)\right]^{-1}\right| = \sup_s\left|\frac{1}{\frac{1}{\sqrt{m}}w_x(\lambda_s)+\frac{1}{\sqrt{m}}\frac{e^{i\lambda_s}}{(1-e^{i\lambda_s})}\frac{X_n}{\sqrt{2\pi n}}}\right|.$$

However, note that the orders of magnitude of $\frac{1}{\sqrt{m}}w_x(\lambda_s)$ and $\frac{1}{\sqrt{m}}\frac{e^{i\lambda_s}}{(1-e^{i\lambda_s})}\frac{X_n}{\sqrt{2\pi n}}$ are $O_p\left(\frac{1}{\sqrt{m}}\frac{n^d}{s^d}\right)$ and $O_p\left(\frac{1}{\sqrt{m}}\frac{n^d}{s}\right)$ respectively, and hence $1/\left|\frac{1}{\sqrt{m}}w_x(\lambda_s)+\frac{1}{\sqrt{m}}\frac{e^{i\lambda_s}}{(1-e^{i\lambda_s})}\frac{X_n}{\sqrt{2\pi n}}\right|$ converges to zero regardless of $s$ if $\frac{n^d}{m^{1+\frac{1}{2}}}\to\infty$, which is already imposed. Therefore, it can be deduced that

$$\sup_s\left|\sqrt{m}\left[\upsilon_x(\lambda_s)\right]^{-1}\right|\xrightarrow{p}0,$$

and hence (29) holds. From (28), it follows that

$$\frac{1}{\sqrt{m}}\sum_{s=1}^{m}x_s\log\frac{I_{\upsilon_x}(\lambda_s;d)}{I_{\upsilon_x}(\lambda_s)}\xrightarrow{p}0.$$

Notice that

$$\frac{1}{m}\sum_{s=1}^{m}x_s\log\frac{I_{\upsilon_x}(\lambda_s;d)}{I_{\upsilon_x}(\lambda_s)}\xrightarrow{p}0 \quad (30)$$

without the condition $\frac{m^{\frac{3}{2}}}{n^d}\to 0$.

Now, we need to show CLT for (26). By the BN decomposition in lemma 6.2, we have

$$\frac{1}{\sqrt{m}}\sum_{s=1}^{m}x_s\ln I_u(\lambda_s) = \frac{1}{\sqrt{m}}\sum_{s=1}^{m}x_s\ln I_\varepsilon(\lambda_s)+o_p(1). \quad (31)$$

Hence, for the asymptotic normality of the modified estimator, we need a CLT for the first term of (31). Since $\varepsilon_t$ is a Gaussian process under condition (i) in theorem 3.2, its dft $w_\varepsilon(\lambda_s)$ is also Gaussian. Moreover, the dft's at different frequencies are exactly uncorrelated, i.e., we have

$$\mathbf{Cov}\left[w_\varepsilon(\lambda_s), w_\varepsilon(\lambda_t)\right] = 0$$

for $s \neq t$. Therefore, the collection $\{w_\varepsilon(\lambda_s)\}_{s=1}^m$ are independent, implying the independence of $\ln I_\varepsilon(\lambda_1), \ldots, \ln I_\varepsilon(\lambda_m)$. We can apply the Lindeberg-Feller CLT under the condition

$$\frac{\max_{1\leq s\leq m} x_s^2}{\sum_{s=1}^m x_s^2} \to 0,$$

which is clearly satisfied here. Therefore, it follows that

$$\frac{1}{\sqrt{m}}\sum_{s=1}^m x_s \ln I_\varepsilon(\lambda_s) \xrightarrow{d} N\left(0, \frac{\pi^2}{6}\right),$$

and

$$\sqrt{m}\left(\widehat{d} - d\right) \xrightarrow{d} N\left(0, \frac{\pi^2}{24}\right),$$

since $\frac{1}{m}\sum_{s=1}^m x_s^2 \to 1$. ■

## References

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.

Andrews, D.W.K. and P. Guggenberger, 2003, A Bias-Reduced Log-Periodogram Regression Estimator for the Long-Memory Parameter, *Econometrica* 71, 675-712.

Baillie, R.T., 1996, Long Memory Processes and Fractional Integration in Econometrics, *Journal of Econometrics* 73, 5-59.

Corbae, D., S. Ouliaris and P. C. B. Phillips, 2002, Band Spectral Regression with Trending Data, *Econometrica* 70, 1067-1109.

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.

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.

Hassler, U., 1993, Regression of Spectral Estimators with Fractionally Integrated Time Series, *Journal of Time Series Analysis* 14, 369-380.

Hassler, U. and J. Wolters, 1995, Long Memory in Inflation Rates: International Evidence, *Journal of Business and Economic Statistics* 13, 37-45.

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.

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.

Kashyap, R. L., and K.-B. Eom, 1988, Estimation in Long-Memory Time Series Model, *Journal of Time Series Analysis* 9, 35-41.

Kim, C. S. and P. C. B. Phillips, 2006, Log Periodogram Regression: The Nonstationary Case, *Cowles Foundation Discussion Paper*, Yale University.

Künsch, H. R., 1986, Discrimination between Monotonic Trends and Long-Range Dependence, *Journal of Applied Probability* 23, 1025-1030.

Liu, M., 1998, Asymptotics of Nonstationary Fractional Integrated Series, *Econometric Theory* 14, 641-662.

Nelson, C. R. and C. I. Plosser, 1982, Trends and Random Walks in Macroeconomic Time Series, *Journal of Monetary Economics* 10, 139-162.

Phillips, P. C. B., 1998, Econometric Analysis of Fisher's Equation, *Cowles Foundation Discussion Paper* #1180, Yale University.

Phillips, P. C. B., 1999, Discrete Fourier Transforms of Fractional Processes, Yale University.

Phillips, P. C. B., 2007, Unit Root Log Periodogram Regression, *Journal of Econometrics* 138, 104-124.

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

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

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

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

Shea, G. S., 1991, Uncertainty and Implied Variance Bounds in Long Memory Models of the Interest Rate Term Structure, *Empirical Economics* 16, 287-312.

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

Velasco, C., 1999a, Non-Stationary Log-Periodogram Regression, *Journal of Econometrics* 91, 325-371.

Velasco, C., 1999b, Gaussian Semiparametric Estimation of Non-Stationary Time Series, *Journal of Time Series Analysis* 20, 87-127.

Velasco, C., 2000, Non-Gaussian Log Periodogram Regression, *Econometric Theory* 16, 44-79.