An I(d) model with trend and cycles[*]

Karim M. Abadir

Imperial College Business School, Imperial College London, London SW7 2AZ, UK
(k.m.abadir@imperial.ac.uk)

Walter Distaso

Imperial College Business School, Imperial College London, London SW7 2AZ, UK
(w.distaso@imperial.ac.uk)

Liudas Giraitis

Department of Economics, Queen Mary, University of London, London E14 NS, UK
(l.giraitis@qmul.ac.uk)

March 14, 2011

Abstract

This paper deals with models allowing for trending processes and cyclical component with error processes that are possibly nonstationary, nonlinear, and non-Gaussian. Asymptotic confidence intervals for the trend, cyclical component, and memory parameters are obtained. The confidence intervals are applicable for a wide class of processes, exhibit good coverage accuracy, and are easy to implement.

JEL Classification: C22.

Keywords: fractional integration, trend, cycle, nonlinear process, Whittle objective function.

---

[*]Corresponding author. *E-mail address:* `k.m.abadir@imperial.ac.uk`.

1 Introduction

To start, consider the basic model

$$X_t = \beta_1 + \beta_2 t + u_t, \quad t = 1, 2, ..., n, \tag{1.1}$$

where $\mathrm{E}(u_t) = 0$. When $\{u_t\}$ is a long-memory stationary process, with memory parameter $d \in (-1/2, 1/2)$, the problem of estimating such models has been studied extensively in the literature. Yajima (1988, 1991) derived conditions for consistency and asymptotic normality of Least Squares (LS) estimators of the parameters of a regression model with nonstochastic regressors, when the errors $\{u_t\}$ have long memory. Dahlhaus (1995) suggested an efficient weighted least squares estimator for β_1 and β_2 and investigated its asymptotic properties in the case of a polynomial regression with stationary errors. Nonlinear regression models with long memory errors have been investigated by Ivanov and Leonenko (2004, 2008). The estimation of a trend when $\{u_t\}$ has $d \in [0, 3/2)$ was discussed by Deo and Hurvich (1998), but they did not estimate d and they required $\{u_t\}$ to have a linear structure with restrictive asymptotic weights.

There is a large literature on the estimation of d in the case of long memory. Fewer papers have so far considered an extended range for d to include regions of nonstationarity. Assuming that $\{u_t\}$ is observed, to estimate d, Velasco (1999, 2003) used data differencing and data tapering, and he noted that this inflates the estimator's variance. Robinson (2005) suggested an adaptive semiparametric estimation method for the case of a polynomial regression with fractionally-integrated errors, employing in his Monte-Carlo study a tapered estimate of d. An alternative approach was developed by Shimotsu and Phillips (2005) who introduced an exact local Whittle estimation method based of fractional diffencing of $\{u_t\}$, which is valid when a nonstationary process $\{u_t\}$ is generated by a linear process. Abadir, Distaso, and Giraitis (2007) extended the classical Whittle estimator to the Fully-Extended Local Whittle (FELW) estimator that is valid for a wider range of d values, allowing for nonstationary $\{u_t\}$ but not for deterministic components.

The present papers focuses on the estimation of the linear regression model (1.1) and its extended version

$$X_t = \beta_1 + \beta_2 t + \beta_3 \sin(\omega t) + \beta_4 \cos(\omega t) + u_t, \quad t = 1, 2, ..., n, \tag{1.2}$$

which allows for stationary and nonstationary errors $\{u_t\}$ and a cyclical component with $\omega \in (0, \pi)$. We assume that ω is known in (1.2), and so we treat separately the boundary case $\omega = 0$ as model (1.1), effectively covering $\omega \in [0, \pi)$ in the paper but not the unrealistic case of $\omega = \pi$ that leads to $X_t = \beta_1 + \beta_2 t + \beta_4 (-1)^t + u_t$. We do not propose a method of estimation for ω; see Nandi and Kundu (2003) and references therein for the estimation of ω in the context of a short-memory linear process and no linear trend. Estimating ω is beyond the scope of this paper, though (as we will show) our procedure allows for time-varying ω and/or multiple cyclical components with different frequencies $\omega_\bullet$. For expository purposes, we refrain from writing these features into the model given in this introduction.

In this paper, we estimate $\boldsymbol{\beta} := (\beta_1, \beta_2, \beta_3, \beta_4)'$ by LS and generalize (for the presence of trend and cycle) the Fully-Extended Local Whittle (FELW) estimator of d given in Abadir, Distaso, and Giraitis (2007). We also provide a simpler alternative form for the

FELW estimator. We show that our estimators are consistent and we obtain their rates of convergence and limiting distributions, as well as confidence intervals based on them. The asymptotic properties of our LS estimators of β_1, β_2 turn out to be unaffected by (and robust to) the unknown cyclical component.

The papers listed earlier require the assumptions of linearity or Gaussianity of the error process. However, our estimation procedure allows for a wide range of permissible values of the memory parameter d and for possibly nonstationary, nonlinear, and nonnormal processes $\{u_t\}$. By virtue of $\{u_t\}$ being modeled semiparametrically, the procedure also allows for seasonality and other effects to be present in $\{u_t\}$ at nonzero spectral frequencies.

In Section 2, we investigate the LS estimators of $\beta_1, \beta_2, \beta_3, \beta_4$, while Section 3 is concerned with the parameters of the process $\{u_t\}$. Section 4 contains the results of simulation experiments on the performance of the estimators suggested earlier. It is technically straightforward to extend our results to higher-order polynomials and to values of d outside the interval $-1/2 < d < 3/2$ to which we restrict our attention in this paper. We do not report such extensions in order to simplify the exposition and because most economic series will not require more than a linear trend or d outside $-1/2 < d < 3/2$. The proofs of the main results are given in Section 5.

We use $\overset{p}{\to}$ and $\overset{d}{\to}$ to denote convergence in probability and in distribution, respectively. We write i for the imaginary unit, 1_A for the indicator of a set A, $\lfloor \nu \rfloor$ for the integer part of ν, C for a generic constant but $c_\bullet$ for specific constants. The lag operator is denoted by L, such that $Lu_t = u_{t-1}$, and the backward difference operator by $\nabla := 1 - L$. We define $a \wedge b := \min\{a, b\}$ and $a \vee b := \max\{a, b\}$.

DEFINITION 1.1. *Let $d = k + d_\xi$, where $k = 0, 1, 2, ...$ and $d_\xi \in (-1/2, 1/2)$. We say that $\{u_t\}$ is an I(d) process (denoted by $u_t \sim \text{I}(d)$) if*

$$\nabla^k u_t = \xi_t, \quad t = 1, 2, ...,$$

where the generating process $\{\xi_t\}$ is a second order stationary sequence with spectral density

$$f_\xi(\lambda) = b_0 |\lambda|^{-2d_\xi} + o(|\lambda|^{-2d_\xi}), \ as \ \lambda \to 0 \tag{1.3}$$

where $b_0 > 0$.

Notice that there are two parameters of interest in this definition, b_0 and d_ξ.

2 Estimation of $\boldsymbol{\beta}$

We will use Ordinary LS (OLS) estimation of $\boldsymbol{\beta}$, because of its ease of application, its consistency, and its asymptotic normality. Feasible Generalized LS (GLS) applied to (1.1) would require us to specify the autocovariance structure explicitly, which is not usually known, so OLS is more in line with the semiparametric approach of our paper. Even so, assuming the autocovariance structure is known and is correctly specified, it has been shown that the loss of efficiency will not be substantial in this context. For example, Table 1 of Yajima (1988) implies that the maximal loss of asymptotic efficiency by OLS compared to the BLUE is 11% when estimating β_1 and β_2, and 2% when estimating the mean of

the differenced data (hence β_2 of the original data). These will correspond to our cases $d \in (-1/2, 1/2)$ and $d \in (1/2, 3/2)$, respectively, as will be seen later. These efficiency bounds apply to GLS as well, since it is a linear estimator, thus limiting the efficiency loss of OLS relative to GLS.

Below it will be shown that the rates of convergence of the OLS estimators depend on the order of integration d of u_t, and their limits depend on the the long run variance s_ξ^2 of $\{\xi_t\}$ which needs to be estimated. Property (1.3) of the spectral density f_ξ implies that

$$
\begin{aligned}
s_\xi^2 &= \lim_{n\to\infty} \mathrm{E}\left(n^{-1/2-d_\xi}\sum_{t=1}^{n}\xi_t\right)^2 = \lim_{n\to\infty} n^{-1-2d_\xi}\int_{-\pi}^{\pi}\left(\frac{\sin(n\lambda/2)}{\sin(\lambda/2)}\right)^2 f_\xi(\lambda)\mathrm{d}\lambda \quad (2.1)\\
&= p(d_\xi)b_0,
\end{aligned}
$$

where b_0 is defined in (1.3) and

$$
p(d) := \int_{-\infty}^{\infty}\left(\frac{\sin(\lambda/2)}{\lambda/2}\right)^2 |\lambda|^{-2d}\mathrm{d}\lambda = \begin{cases} 2\frac{\Gamma(1-2d)\sin(\pi d)}{d(1+2d)}, & \text{if } d \neq 0,\\ 2\pi, & \text{if } d = 0. \end{cases}
$$

To derive the asymptotic distribution of estimators of (β_1, β_2), we introduce the following condition on the generating process $\{\xi_t\}$ of Definition 1.1.

ASSUMPTION FDD. The finite-dimensional distributions of the process

$$
Y_n(r) := n^{-1/2-d_\xi}\sum_{t=1}^{\lfloor nr\rfloor+1}\xi_t, \quad 0 \le r \le 1 \tag{2.2}
$$

converge to those of the Gaussian process $Y_\infty(r)$, that is,

$$
Y_n(r) \xrightarrow{d} Y_\infty(r), \text{ as } n \to \infty.
$$

Assumption FDD together with asymptotic (1.3) of spectral density f_ξ imply that

$$
Y_\infty(r) = s_\xi J_{1/2+d_\xi}(r), \; 0 \le r \le 1,
$$

where $J_{1/2+d_\xi}(r)$ is a fractional Brownian motion. By definition, $J_{1/2+d_\xi}(r)$ is a Gaussian process with zero mean and covariance function

$$
R_d(r,s) := \mathrm{E}\left(J_{1/2+d_\xi}(r)J_{1/2+d_\xi}(s)\right) = \frac{1}{2}(r^{1+2d_\xi} + s^{1+2d_\xi} - |r-s|^{1+2d_\xi}), \quad 0 \le r,s \le 1. \tag{2.3}
$$

2.1 Model (1.1)

In order to estimate the slope parameter β_2 and the location parameter β_1 of model (1.1), we use the OLS estimators

$$
\widehat{\beta}_2 = \frac{\sum_{t=1}^{n}(X_t - \bar{X})(t - \bar{t})}{\sum_{t=1}^{n}(t - \bar{t})^2} \tag{2.4}
$$

4

and

$$\widehat{\beta}_1 = \bar{X} - \widehat{\beta}_2 \bar{t}, \tag{2.5}$$

where $\bar{X} = n^{-1} \sum_{t=1}^n X_t$ and $\bar{t} = n^{-1} \sum_{t=1}^n t = (n+1)/2$ are the sample means of the variables. We shall show that the estimator $(\widehat{\beta}_1, \widehat{\beta}_2)$ is valid for both models, (1.1) and (1.2), and the cyclical component of model (1.2) does not affect the asymptotic properties of this estimator. The regressors of model (1.1) are asymptotically orthogonal to the additional cyclical ones in model (1.2), when we normalize them all to make the matrix of sum-of-squares asymptotically nonsingular. In standard regression analysis, where $\{u_t\}$ has short memory, the asymptotic distribution of the OLS estimators of (β_1, β_2) is unaffected by the choice of models. Here, we show inter alia that this still holds for these distribiutions for the values of d that we consider.

Define the following functions of d:

$$\sigma_{2,d}^2 := \begin{cases} (12)^2 \left(\frac{1}{2d+3} - \frac{1}{4} \right), & \text{if } d \in [-1/2, 1/2], \\ (12)^2 \frac{2d-1}{8d(2d+1)(2d+3)}, & \text{if } d \in (1/2, 3/2], \end{cases} \tag{2.6}$$

and

$$\sigma_{1,d}^2 := 1 + 36 \left(\frac{1}{2d+3} - \frac{1}{4} \right), \quad \sigma_{12,d} := \frac{18(2d-1)}{3+2d}. \tag{2.7}$$

Set

$$\tau_{\widehat{\beta}_2} := n^{3/2-d}(\widehat{\beta}_2 - \beta_2), \quad \tau_{\widehat{\beta}_1} := n^{1/2-d}(\widehat{\beta}_1 - \beta_1).$$

THEOREM 2.1. *Assume that X_t follows model (1.1), and $\{u_t\}$ is an I(d) process with $d \in (-1/2, 3/2)$ and $d \neq 1/2$.*

(i) Then, as $n \to \infty$,

$$\mathrm{E}(\tau_{\widehat{\beta}_2}^2) \to s_\xi^2 \sigma_{2,d}^2, \text{ if } d \in (-1/2, 3/2) \tag{2.8}$$

$$\mathrm{E}(\tau_{\widehat{\beta}_1}^2) \to s_\xi^2 \sigma_{1,d}^2, \text{ if } d \in (-1/2, 1/2). \tag{2.9}$$

(ii) If the generating process $\{\xi_t\}$ satisfies Assumption FDD, then, as $n \to \infty$,

$$\tau_{\widehat{\beta}_2} \xrightarrow{d} \mathrm{N}(0, s_\xi^2 \sigma_{2,d}^2), \text{ if } d \in (-1/2, 3/2), \tag{2.10}$$

$$(\tau_{\widehat{\beta}_1}, \tau_{\widehat{\beta}_2}) \xrightarrow{d} \mathrm{N}(\mathbf{0}, \boldsymbol{\Sigma}_2), \text{ if } d \in (-1/2, 1/2), \tag{2.11}$$

where $\mathrm{N}(\mathbf{0}, \boldsymbol{\Sigma}_2)$ is a zero mean Gaussian vector with the covariance matrix

$$\boldsymbol{\Sigma}_2 := s_\xi^2 \begin{pmatrix} \sigma_{1,d}^2 & \sigma_{12,d} \\ \sigma_{12,d} & \sigma_{2,d}^2 \end{pmatrix}.$$

(iii) Results (i) and (ii) remain valid if X_t follows the model (1.2) with unknown ω.

(iv) Results (i) and (ii) hold if $X_t = \beta_1 + \beta_2 t + z_t + u_t$ with any deterministic trend z_t such that $\sum_{t=1}^n z_t = o(n^{1/2+d})$, as $n \to \infty$.

The results of Theorem 2.1 are derived under unrestrictive Assumption FDD that requires only a finite second moment of $\{\xi_t\}$. They do not assume that the generating process $\{\xi_t\}$ has higher moments, or a linear structure with regularly decaying weights $a_t \sim C t^{-\gamma}$ as in Deo and Hurvich (1998). Notice that the mean of X_t cannot be estimated consistently if $d \geq 1/2$, a general feature of integrated models, which is why the results on $\widehat{\beta}_1$ are limited to the case $d < 1/2$. Constructing confidence intervals for β_2 below will use the estimate $\sigma^2_{2,\widehat{d}}$ of $\sigma^2_{2,d}$, where $\widehat{d}$ is a consistent estimate of d. Although we exclude the value $d = 1/2$ by assumption, $\widehat{d}$ may take value $d = 1/2$ with a negligible probability. For this reason, definition (2.6) of $\sigma^2_{2,d}$ includes the boundary value $d = 1/2$.

REMARK 2.1. Results (i)-(ii) of Theorem 2.1 remain valid for a process $X_t = \beta_1 + \beta_2 t + z_t + u_t$ where

$$z_t = \sum_{j=1}^{p} (\beta_{3,j} \sin(\omega_j t) + \beta_{4,j} \cos(\omega_j t))$$

with unknown parameters $\beta_{3,j}, \beta_{4,j}, \omega_j$, assuming that $0 < \omega_j < \pi$, $j = 1, ..., p$. They also remain valid when the ω_j's vary deterministically over time, so long as they satisfy the condition in (iv). Such changes include a finite number of breaks as a special case.

2.2 Model (1.2)

To estimate the parameter vector $\boldsymbol{\beta} = (\beta_1, \beta_2, \beta_3, \beta_4)'$ of model (1.2), we shall use the OLS estimator

$$\widetilde{\boldsymbol{\beta}} := (\widetilde{\beta}_1, \widetilde{\beta}_2, \widetilde{\beta}_3, \widetilde{\beta}_4)' = \boldsymbol{S}^{-1} \boldsymbol{t}, \tag{2.12}$$

where $\boldsymbol{S} = (s_{ij})_{i,j=1,...,4}$ and $\boldsymbol{t} = (t_{n,j})_{j=1,...,4}$, with elements

$$s_{ij} := \sum_{t=1}^{n} z_{ti} z_{tj}, \ \ t_{n,j} := \sum_{t=1}^{n} u_t z_{tj} \tag{2.13}$$

such that

$$z_{t1} := 1, \ z_{t2} := t, \ z_{t3} := \sin(\omega t), \ z_{t4} := \cos(\omega t).$$

Theorem 2.1 and Theorem 2.2 below show that, in spite of the cyclical component, the earlier LS estimator $(\widehat{\beta}_1, \widehat{\beta}_2)$ given in (2.4)-(2.5) can be applied to model (1.2) without loss of efficiency and that, under the unrestrictive Assumption FDD, it has the same asymptotic distribution as the LS estimator $(\widetilde{\beta}_1, \widetilde{\beta}_2)$ of the larger model.

Deriving the asymptotic normality of the LS estimators of the parameters β_3, β_4 requires an additional assumption, that the generating process $\{\xi_t\}$ is a linear process

$$\xi_t = \sum_{j=0}^{\infty} a_j \varepsilon_{t-j}, \tag{2.14}$$

where $\{a_j\}$ are real nonrandom weights, $\sum_{j=0}^{\infty} a_j^2 < \infty$, and $\{\varepsilon_j\}$ are i.i.d. variates with zero mean and unit variance. This assumption is because $\widetilde{\beta}_3$ and $\widetilde{\beta}_4$ involve trigonometric sums that cannot be transformed into integrals of the process $\{Y_n(r)\}$, whence the need to impose assumptions directly on $\{\xi_t\}$ rather than on the convergence of its partial sum

process $\{Y_n(r)\}$. In the case here, asymptotic normality is obtained by assuming that $\{\xi_t\}$ is a linear process.

Define

$$\tau_{\widetilde{\beta}_2} := n^{3/2-d}(\widetilde{\beta}_2 - \beta_2), \quad \tau_{\widetilde{\beta}_1} := n^{1/2-d}(\widetilde{\beta}_1 - \beta_1),$$

$$\tau_{\widetilde{\beta}_k} := \begin{cases} \frac{n^{1/2}}{\sqrt{A_k}}(\widetilde{\beta}_k - \beta_k), & \text{if } d < 1, \\ \frac{n^{3/2-d}}{\sqrt{A_k}}(\widetilde{\beta}_k - \beta_k), & \text{if } d \geq 1, \end{cases}$$

for $k = 3, 4$, where A_k is such that $\tau_{\widetilde{\beta}_k}$ has a limiting variance of 1. It will be derived in Theorem 2.2(i) as

$$A_k = \begin{cases} 4\pi f_\xi(\omega), & \text{if } -\frac{1}{2} < d < \frac{1}{2}, \\ \frac{\pi f_\xi(\omega)}{\sin^2(\omega/2)}, & \text{if } \frac{1}{2} < d < 1, \\ \frac{n^{2(1-d)}\pi f_\xi(\omega)}{\sin^2(\omega/2)} + s_\xi^2(\theta_{1k}^2\psi_{11} + \theta_{2k}^2\psi_{22} + 2\theta_{1k}\theta_{2k}\psi_{12} \\ \quad + 4\gamma_k(\frac{\theta_{1k}}{2} + v\theta_{2k}) + 4\gamma_k^2), & \text{if } 1 \leq d < \frac{3}{2}, \end{cases} \tag{2.15}$$

with

$$\psi_{11} := \frac{1}{1+2d}, \quad \psi_{12} := \frac{6d-1}{8d(1+2d)}, \quad \psi_{22} := \frac{1}{2(1+2d)}\left(1 - \frac{1}{(3+2d)d}\right) \tag{2.16}$$

$$\theta_{1k} := -8s_{1k} + 12n^{-1}s_{2k}, \quad \theta_{2k} := 12s_{1k} - 24n^{-1}s_{2k},$$

$$v := \frac{1}{4} + \frac{1}{1+2d} - \frac{1}{4d}, \quad \gamma_3 := -\frac{\cos(\omega(n+1/2))}{2\sin(\omega/2)}, \quad \gamma_4 := \frac{\sin(\omega(n+1/2))}{2\sin(\omega/2)}.$$

Observe that, for $1 \leq d < 3/2$, A_k is bounded but oscillates with n. Although for $d > 1$ in (2.15) the term $n^{2(1-d)}\pi f_\xi(\omega)/\sin^2(\omega/2)$ is asymptotically negligible, we found that preserving it improves the coverage probabilities in finite samples.

THEOREM 2.2. *Assume that X_t follows model (1.2). Suppose that $\{u_t\}$ is an I(d) process with $d \in (-1/2, 3/2)$ and $d \neq 1/2$, such that the spectral density f_ξ is continuous at ω. In the case of $d \in (1/2, 3/2)$, assume that $E(u_0^2) < \infty$.*

(i) Then, $E(\tau_{\widetilde{\beta}_1}^2)$ and $E(\tau_{\widetilde{\beta}_2}^2)$ satisfy (2.9) and (2.8) with $\widehat{\beta}_1, \widehat{\beta}_2$ replaced by $\widetilde{\beta}_1, \widetilde{\beta}_2$. For $k = 3, 4$, as $n \to \infty$,

$$E(\tau_{\widetilde{\beta}_k}^2) \sim 1 \text{ for } d \in (-1/2, 3/2).$$

(ii) If the generating process $\{\xi_t\}$ satisfies Assumption FDD, then $\tau_{\widetilde{\beta}_2}$ and $(\tau_{\widetilde{\beta}_1}, \tau_{\widetilde{\beta}_2})$ satisfy (2.10) and (2.11), respectively.

(iii) If the generating process $\{\xi_t\}$ is a linear process (2.14), then for $d \in (-1/2, 1/2)$,

$$(\tau_{\widetilde{\beta}_1}, \tau_{\widetilde{\beta}_2}, \tau_{\widetilde{\beta}_3}, \tau_{\widetilde{\beta}_4}) \xrightarrow{d} N(\mathbf{0}, \boldsymbol{\Sigma}_4), \tag{2.17}$$

where $N(\mathbf{0}, \boldsymbol{\Sigma}_4)$ is a zero mean Gaussian vector with covariance matrix

$$\boldsymbol{\Sigma}_4 := \begin{pmatrix} s_\xi^2\sigma_{1,d}^2 & s_\xi^2\sigma_{12,d} & 0 & 0 \\ s_\xi^2\sigma_{12,d} & s_\xi^2\sigma_{2,d}^2 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}.$$

For $d \in (1/2, 3/2)$,

$$\tau_{\widetilde{\beta}_2} \xrightarrow{d} \mathrm{N}\left(0, s_\xi^2 \sigma_{2,d}^2\right), \tag{2.18}$$

$$\tau_{\widetilde{\beta}_k} \xrightarrow{d} \mathrm{N}(0,1), \quad k = 3, 4. \tag{2.19}$$

In particular, for $1/2 < d < 1$,

$$(\tau_{\widetilde{\beta}_2}, \tau_{\widetilde{\beta}_3}, \tau_{\widetilde{\beta}_4}) \xrightarrow{d} \mathrm{N}\left(\mathbf{0}, \mathrm{diag}(s_\xi^2 \sigma_{2,d}^2, 1, 1)\right). \tag{2.20}$$

To make this theorem operational, we need to estimate the unknown parameters d and s_ξ. Some actual estimators will be proposed in the next section, but in the meantime we have the following result which is valid for any estimators satisfying a weak condition for consistency.

COROLLARY 2.1. *Suppose that the assumptions of Theorem 2.2(iii) are satisfied, and that we have estimators $\widehat{d}$, $\widehat{s}_\xi^2$, and $\widehat{f}_\xi(\omega)$ with the property*

$$\widehat{d} = d + o_p(1/\log n), \quad \widehat{s}_\xi^2 = s_\xi^2 + o_p(1), \quad \widehat{f}_\xi(\omega) = f_\xi(\omega) + o_p(1).$$

Then, as $n \to \infty$,

$$\frac{n^{3/2-\widehat{d}}}{\widehat{s}_\xi \, \sigma_{2,\widehat{d}}} (\widetilde{\beta}_2 - \beta_2) \xrightarrow{d} \mathrm{N}(0,1), \quad \text{if} \quad d \in (-1/2, 3/2), \tag{2.21}$$

and, for $k = 3, 4$,

$$\sqrt{\frac{n}{\widehat{A}_k^+}} (\widetilde{\beta}_k - \beta_k) \xrightarrow{d} \mathrm{N}(0,1), \quad \text{if} \quad d \in (-1/2, 1)$$

$$\frac{n^{3/2-d}}{\sqrt{\widehat{A}_k^+}} (\widetilde{\beta}_k - \beta_k) \xrightarrow{d} \mathrm{N}(0,1), \quad \text{if} \quad d \in [1, 3/2).$$

with $\widehat{A}_k$ the corresponding estimate of A_k and $\widehat{A}_k^+ := \max\{0, \widehat{A}_k\}$.

Using the above results, we can write an asymptotic confidence interval (CI) of size $1 - \gamma$ for β_1 and β_2 as

$$\left[\widetilde{\beta}_1 - \frac{\sigma_{1,\widehat{d}} \widehat{s}_\xi}{n^{1/2-\widehat{d}}} c_\gamma, \ \widetilde{\beta}_1 + \frac{\sigma_{1,\widehat{d}} \widehat{s}_\xi}{n^{1/2-\widehat{d}}} c_\gamma\right], \quad \left[\widetilde{\beta}_2 - \frac{\sigma_{2,\widehat{d}} \widehat{s}_\xi}{n^{3/2-\widehat{d}}} c_\gamma, \ \widetilde{\beta}_2 + \frac{\sigma_{2,\widehat{d}} \widehat{s}_\xi}{n^{3/2-\widehat{d}}} c_\gamma\right]$$

valid for $|d| < 1/2$ and $-1/2 < d < 3/2$, respectively, where c_γ denotes here the quantile of the standard normal distribution satisfying $\Pr(|\mathrm{N}(0,1)| > c_\gamma) = \gamma$. Similarly for β_3 and β_4, the corresponding intervals in the cases $-1/2 < d < 1$ and $1 \le d < 3/2$ are

$$\left[\widetilde{\beta}_k - \sqrt{\frac{\widehat{A}_k^+}{n}} c_\gamma, \ \widetilde{\beta}_k + \sqrt{\frac{\widehat{A}_k^+}{n}} c_\gamma\right], \quad \left[\widetilde{\beta}_k - \frac{\sqrt{\widehat{A}_k^+}}{n^{3/2-d}} c_\gamma, \ \widetilde{\beta}_k + \frac{\sqrt{\widehat{A}_k^+}}{n^{3/2-d}} c_\gamma\right], \ k = 3, 4;$$

see Section 4 concerning the positivity of estimates of A_k. If we have $\widehat{A}_k = 0$, then the Corollary's statistics based on $\hat{A}_k^+$ are infinite but this happens with probability 0 asymptotically. Also, we interpret the CIs to become the whole real line in this case.

The formulae imply that the length of the CIs for β_1 and β_2 increase when the parameter d approaches the bounds $1/2$ and $3/2$. The length of CIs for β_3, β_4 does not depend on d when $d < 1$ and increases when d approaches $3/2$. Theorem 2.2 shows that as long as $\{u_t\}$ is an I(d) process, the LS estimators $\widetilde{\beta}_1, \widetilde{\beta}_2, \widetilde{\beta}_3, \widetilde{\beta}_4$ are consistent and the rate of convergence is known. As in the previous subsection which dealt with the simpler model (1.1), consistent estimation of the intercept β_1 is not possible when $d > 1/2$. However, and perhaps surprisingly, we can still estimate consistently β_3 and β_4 of the bounded cyclical component when $d > 1/2$.

The asymptotic normality of $\widetilde{\beta}_1$ and $\widetilde{\beta}_2$ holds under the weak Assumption FDD, whereas asymptotic normality of $\widetilde{\beta}_3, \widetilde{\beta}_4$ follows under the assumption that the generating process $\{\xi_t\}$ is linear. Only a finite second moment of u_t, ξ_t is required.

REMARK 2.2. Assumption FDD with a Gaussian limit is satisfied for a wide class of generating processes $\{\xi_t\}$. In the case of a linear process, convergence of the finite-dimensional distributions of $Y_n(r)$ to those of a Gaussian process $s_\xi J_{1/2+d_\xi}$ has been known for some time; e.g. see Ibragimov and Linnik (1971, Theorem 18.6.5) and Davydov (1970). In the case of EGARCH processes, it was shown by Surgailis and Viano (2002). Taqqu (1979) and Giraitis and Surgailis (1985) have shown that Assumption FDD is satisfied for wide class on nonlinear transformations $\xi_t = h(\zeta_t)$ of stationary Gaussian I(d) sequences $\{\zeta_t\}$.

Although FDD is satisfied for a large class of nonlinear process, it does not cover nonlinear processes whose partial sum process has a non-Gaussian limit Y_∞. The latter case is specific and of limited interest in applications since, although it allows consistent OLS estimation of (β_1, β_2), one does not have a simple data-based procedure for the computation of its critical values.

3   Estimation of d and s_ξ^2

In this section we discuss estimators for the unknown parameter d of the I(d) process $\{u_t\}$ and the long-run variance s_ξ^2 of $\{\xi_t\}$. We start with a model with no deterministic components, then generalize it to models (1.1) and (1.2).

3.1   Model with no deterministic components

Assume in this subsection that we observe the process $\{u_t\}$, which follows $u_t \sim \mathrm{I}(d_0)$ with

$$d_0 \in (a, b) \subset (-1/2,\ \infty), \quad d_0 \neq 1/2, 3/2, 5/2, \cdots . \tag{3.1}$$

Then, d_0 can be estimated using the FELW estimator developed in Abadir *et al.* (2007) which extends to nonstationarity the classical local Whittle estimator (see Robinson (1995b)). In this subsection, we write the estimator in a form that is equivalent to the original FELW estimator, but in a way that is simpler to use. The estimators are identical due to the algebraic relation of the periodograms presented in Lemma 4.4 of Abadir *et al.* (2007).

Denote by

$$I_{n,u}(\lambda_j) := |w_u(\lambda_j)|^2, \qquad w_u(\lambda_j) := (2\pi n)^{-1/2} \sum_{t=1}^{n} e^{it\lambda_j} u_t$$

the periodogram and discrete Fourier transform of $\{u_t\}$, where $\lambda_j = 2\pi j/n$, $j = 1, \ldots, n$ denote the Fourier frequencies. For $k = 0, 1, 2, \ldots$, define

$$I_{n,u}(\lambda_j, d) := |1 - e^{i\lambda_j}|^{-2k} I_{n, \nabla^k u}(\lambda_j)$$

for all $d \in (k - 1/2, \ k + 1/2]$, $k = 0, 1, 2, \ldots$. The points $d_0 = k + 1/2$ are excluded from (3.1) because they lead to the spectral density of $\nabla^k u_t$ not being integrable.

The FELW estimator $\widehat{d} \equiv \widehat{d}_u$ of d_0 is defined as

$$\widehat{d} := \operatorname{argmin}_{d \in [a,b]} \ U_n(d), \tag{3.2}$$

where

$$U_n(d) := \log\left(\frac{1}{m} \sum_{j=1}^{m} j^{2d}|1 - e^{i\lambda_j}|^{-2k} \ I_{n,\nabla^k u}(\lambda_j)\right) - \frac{2d}{m} \sum_{j=1}^{m} \log j, \tag{3.3}$$

for $d \in (k - 1/2, \ k + 1/2]$, $k = 0, 1, 2, \ldots$, and the bandwidth parameter m is such that

$$m \to \infty, \ m = o(n)$$

as $n \to \infty$. We can think of k as determined by any given d: for any d chosen to evaluate the objective function $U_n(d)$, we have a corresponding k such that $d \in (k - 1/2, \ k + 1/2]$. For example,

$$U_n(d) := \begin{cases} \log\left(\frac{1}{m} \sum_{j=1}^{m} j^{2d} \ I_{n,u}(\lambda_j)\right) - \frac{2d}{m} \sum_{j=1}^{m} \log j, & \text{if } d \in (-1/2, 1/2], \\ \log\left(\frac{1}{m} \sum_{j=1}^{m} j^{2d}|1 - e^{i\lambda_j}|^{-2} \ I_{n,\nabla u}(\lambda_j)\right) - \frac{2d}{m} \sum_{j=1}^{m} \log j, & \text{if } d \in (1/2, 3/2]. \end{cases}$$

REMARK 3.1. In applying the estimation method, the sample data must be enumerated to allow to compute the difference $\nabla^k u_1$. For example, if $[a, b] = [-1/2, 3/2]$ then the data should be enumerated as $u_0, \ldots, u_n$, whereas if $[a, b] = [-1/2, 5/2]$ then it should be $u_{-1}, u_0, \ldots, u_n$.

In the case of $d_0 \in (-1/2, 3/2)$, we shall estimate the scale parameter b_0 of (1.3) by

$$\widehat{\beta}_{m,u}(\widehat{d}) := \frac{1}{m} \sum_{j=1}^{m} \lambda_j^{2\widehat{d}} I_{n,u}(\lambda_j, \widehat{d})$$

and the long-run variance s_ξ^2 by

$$\widehat{s}_\xi^2 \equiv \widehat{s}_{m,u}^2(\widehat{d}) = p(\widehat{d} - k)\widehat{\beta}_{m,u}(\widehat{d}), \quad \text{for } \widehat{d} \in (k - 1/2, \ k + 1/2], \ k = 0, 1. \tag{3.4}$$

In Dalla, Giraitis, and Hidalgo (2006) for $|d| < 1/2$ and Abadir *et al.* (2007) for general d, it was shown that under weak (ergodicity type) assumptions on the generating process $\{\xi_t\}$, the FELW estimator is consistent at a rate faster than $1/\log n$,

$$\widehat{d} - d_0 = o_p(1/\log n). \tag{3.5}$$

10

This is sufficient to write $n^d = n^{\widehat{d}}(1+o_p(1))$. Consistency of $\widehat{d}$ did not require the assumption that $\{\xi_t\}$ be a Gaussian or a linear process. Under the same conditions, in Theorem 2.1 of Abadir *et al.* (2007), consistency of the estimator of b_0 was shown,

$$\widehat{\beta}_{m,u}(\widehat{d}) \xrightarrow{p} b_0,$$

which together with (3.5) implies consistency of the estimator of the long-run variance, $\widehat{s}^2_{m,u}(\widehat{d}) \xrightarrow{p} s^2_\xi$. Asymptotic normality of the FELW estimator in Abadir *et al.* (2007) was derived for the case of a linear generating process $\{\xi_t\}$.

3.2 The two models with deterministic components

Assume now that $\{u_t\}$ is not observed. In model (1.1), to estimate d we use the detrended observations

$$\widehat{u}_t = X_t - \widehat{\beta}_2 t = u_t + \beta_1 + (\beta_2 - \widehat{\beta}_2)t, \quad t = 0, 1, ..., n, \tag{3.6}$$

where $\widehat{\beta}_2$ is the OLS estimator (2.4). In regression (1.2), we use instead

$$\begin{aligned}
\widehat{u}_t &= X_t - \widetilde{\beta}_2 t - \widetilde{\beta}_3 \sin(\omega t) - \widetilde{\beta}_4 \cos(\omega t) \\
&= u_t + \beta_1 + (\beta_2 - \widetilde{\beta}_2)t + (\beta_3 - \widetilde{\beta}_3)\sin(\omega t) + (\beta_4 - \widetilde{\beta}_4)\cos(\omega t), \ t = 0, 1, ..., n, \tag{3.7}
\end{aligned}$$

where $\widetilde{\beta}_2, \widetilde{\beta}_3, \widetilde{\beta}_4$ is the OLS estimator (2.12). The FELW estimator $\widehat{d}_{\widehat{u}}$ is computed using the series $\{\widehat{u}_t\}$. Henceforth, we write $\widehat{d}$ for $\widehat{d}_{\widehat{u}}$ only.

Recall that for $d_0 > -1/2$, the FELW estimator is exactly (not just asymptotically) invariant with respect to the mean because of the property $\sum_{t=1}^n e^{i\lambda_j t} g_t = 0$ when g_t is constant; and for $d_0 > 1/2$, it is so with respect to a linear trend because of the inherent differencing in the estimation procedure; see Abadir *et al.* (2007). In regression models (1.1) and (1.2), we assume that $d \in (-1/2, 3/2)$ and hence the minimization in (3.2) that yields $\widehat{d}$ is carried out over $[-1/2, 3/2]$.

We show below that the estimator of d based on $\{\widehat{u}_t\}$ has the same asymptotic properties as in case when $\{u_t\}$ is observed. We shall need the following assumption.

ASSUMPTION L. $\{\xi_t\}$ is a linear process (2.14) such that: the i.i.d. noise $\{\varepsilon_t\}$ in (2.14) has finite fourth moment, the spectral density

$$f_\xi(\lambda) = |\lambda|^{-2d_\xi}(b_0 + \lambda^2 b_1 + o(\lambda^2)), \text{ as } \lambda \to 0, \tag{3.8}$$

for some $d_\xi \in (-1/2, 1/2)$, $b_0 > 0$, and the transfer function $\alpha(\lambda) := \sum_{j=0}^\infty e^{ij\lambda} a_j$, is such that

$$\frac{d\alpha(\lambda)}{d\lambda} = O(|\alpha(\lambda)|/\lambda), \quad \text{as} \quad \lambda \to 0^+.$$

The i.i.d. assumption imposed on $\{\varepsilon_t\}$ in (2.14) is only needed for expository purposes and can be relaxed to martingale difference.

THEOREM 3.1. *Assume that $u_t \sim I(d_0)$, where $d_0 \in (-1/2, 3/2)$, $d_0 \neq 1/2$. Suppose that the generating process $\{\xi_t\}$ is a linear process (2.14) satisfying Assumption L.*

(i) Then,

$$\sqrt{m}(\widehat{d} - d_0) \xrightarrow{d} \mathrm{N}(0, 1/4), \quad if\ m = o(n^{4/5}), \tag{3.9}$$

and, if $n^{4/5}/m = o(1)$, then

$$(n/m)^2(\widehat{d} - d_0) \xrightarrow{p} \begin{cases} \frac{b_1}{9b_0(2\pi)^2}, & if\ d_0 \in (-1/2, 1/2), \\ \frac{b_1}{9b_0(2\pi)^2} + \frac{1}{108(2\pi)^2}, & if\ d_0 \in (1/2, 3/2). \end{cases} \tag{3.10}$$

If m is proportional to $n^{4/5}$, we get a CLT but the limiting mean is not zero.

(ii) The results of (i) remain valid if $\widehat{d} = \widehat{d}_{\widehat{u}}$ is computed using:

a) $\widehat{u}_t$ of (3.6) obtained from model (1.1),

b) $\widehat{u}_t$ of (3.7) obtained from model (1.2).

The estimation of β_1 and β_2 in Corollary 2.1 requires only $o_p(1/\log n)$ consistency of an estimator $\widehat{d}$ of d_0, and does not assume $\{\xi_t\}$ to be a linear process. However, if one is interested in obtaining the asymptotic normality of $\widehat{d}$, then this is derived in Theorem 3.1 under the assumption that $\{\xi_t\}$ is a linear process.

4   Simulation results

In this section, the finite sample performance of the asymptotic results given earlier are assessed through simulations. We let the errors $\{\xi_t\}$ be generated from a Gaussian fractional ARIMA$(1, d_\xi, 0)$ noise with unit standard deviation, where the autoregressive parameter ρ is equal to $-0.5, 0, 0.5$. Throughout the simulation exercise, the number of replications is 10,000, the sample size $n = 500$. The memory parameter d is estimated using the Local Whittle bandwidth parameter $m = \lfloor n^{.65} \rfloor = 56$. We have also tried two other bandwidths which are not reported for space considerations: $m = \lfloor n^{.7} \rfloor = 77$ whose results are almost indistinguishable from $\lfloor n^{.65} \rfloor$, and $m = \lfloor n^{.8} \rfloor = 144$ whose results are dominated by $\lfloor n^{.65} \rfloor$. For more details on our choice of bandwidth, see Abadir et al. (2007), Dalla et al. (2006), and Shimotsu and Phillips (2005). Under some additional restrictions, the optimal m can be chosen using the data-driven methods of Henry and Robinson (1996) or Andrews and Sun (2004).

We start with the case when the data generating process follows model (1.1). Using asymptotic CIs, we report Coverage Probabilities (CPs) in Table 1. The CPs are satisfactory, with a few exceptions: low values of ρ and d, and high values of ρ.

We then move to the estimation of the regression parameters $\beta_1, \beta_2, \beta_3, \beta_4$. Corollary 2.1 established consistency and a central limit theorem for the LS estimators $\widetilde{\beta}_2, \widetilde{\beta}_3, \widetilde{\beta}_4$ when $d \in (-1/2, 3/2)$ and $\widetilde{\beta}_1$ when $d \in (-1/2, 1/2)$. Using asymptotic CIs, we report CPs in Table 2. CPs for β_1 and β_2 are very similar to those reported in Table 1. For β_3 and β_4, we observe a deterioration of CPs at the opposite extremes of the table: low ρ and d, and high ρ and d. Notice that in the Monte Carlo exercise, apart from ω, all the remaining parameters needed to construct CIs are estimated. A consistent estimator of the spectral density at ω is given as follows. Define

$$w_u(\lambda_j; d) := \begin{cases} (2\pi n)^{-1/2} \sum_{t=1}^{n} \mathrm{e}^{\mathrm{i}t\lambda_j} u_t, & -1/2 < d < 1/2; \\ (2\pi n)^{-1/2} \sum_{t=1}^{n} \mathrm{e}^{\mathrm{i}t\lambda_j} (u_t - u_{t-1}), & 1/2 < d < 3/2. \end{cases}$$

12

Then,

$$\widehat{f}_\xi(\omega) = \frac{2}{n} \sum_{j=1}^{\lfloor n/2 \rfloor - 1} b_n(\lambda_j) \left| w_{\widehat{u}}(\lambda_j - \omega; \widehat{d}) \right|^2, \tag{4.1}$$

where

$$b_n(\lambda_j) = q^{-1} \left(\frac{\sin(q\lambda_j/2)}{\sin(\lambda_j/2)} \right)^2, \quad q = \left\lfloor n^{1/3} \right\rfloor.$$

For $-1/2 < d < 1$, the sign $A_k > 0$ is obvious by (2.15). Since A_k may be close to zero for $1 \le d < 3/2$, we conducted a numerical check with $n = 1, \cdots, 1000$ and a dense grid of $d \in [1, 1.5)$ and $\omega \in (0, \pi)$: the minimum of the function

$$\theta_{1k}^2 \psi_{11} + \theta_{2k}^2 \psi_{22} + 2\theta_{1k}\theta_{2k}\psi_{12} + 4\gamma_k\left(\frac{\theta_{1k}}{2} + v\theta_{2k}\right) + 4\gamma_k^2$$

(appearing in A_k) is positive. It is also possible to use $\widehat{A}_k^+ := \max\{0, \widehat{A}_k\}$, although simulations show that this is not necessary.

It is worth mentioning that the CPs improve and become very close to the nominal ones when infeasible CIs (ones based on the "true" memory parameter and spectral density) are calculated. A full set of tables is not reported for space reasons and is available upon request.

Table 3 provides a check for Remark 2.1. Comparing the mean squared errors (MSEs) of both misspecified and correctly specified LS estimators of β_1 and β_2, it is immediate to see that there is no loss of precision caused by misspecification here. Therefore, if the purpose is only the consistent estimation of location and trend parameters, then one can ignore modelling cyclical components and just estimate a simple mean plus trend model. Of course, this is not true in the case of testing hypotheses on β_1 and β_2, since in the latter case one has to estimate s_ξ^2, which requires a correct specification of the model. The specification

$$
\begin{aligned}
X_t &= \beta_1 + \beta_2 t + z_t + u_t, \\
z_t &= \sum_{j=1}^{3} (\beta_{3,j} \sin(\omega_j t) + \beta_{4,j} \cos(\omega_j t)),
\end{aligned}
$$

with $\beta_{3,1} = \beta_{4,1} = 1$, $\beta_{3,2} = \beta_{4,2} = 1/2$, $\beta_{3,3} = \beta_{4,3} = 3/2$, $\omega_1 = \pi/2$, $\omega_2 = \pi/4$, $\omega_3 = 3\pi/2$, was also tried and gave very similar results.

Table 4 reports the results of estimation of the fractional integration parameter d when the series are detrended and the extended local Whittle estimator $\widehat{d}_{\widehat{u}}$ of the previous section is used. We check the performance of the estimator for a range of values of $d \in (-1/2, 3/2)$. The table confirms some previous findings of the literature on Whittle estimation of the memory parameter. In particular, we observe a slight negative bias for $\rho = -.5$ and $\rho = 0$ and a higher, positive bias when $\rho = .5$. The MSEs of $\widehat{d}_{\widehat{u}}$ are generally slightly higher than those of $\widehat{d}_X$ in Abadir et $al.$ (2007), reflecting the contribution of the estimation of deterministic components.

These are general features that hold for the whole range considered for d, except at the lower boundary of our interval. For $d = -.4$ and $\rho = -0.5$, $\rho = 0$ the estimator of d is less biased and, in the case where $\rho = 0$, it is (slightly) positively biased.

5 Appendix. Proofs

Preliminary facts. Assume that $u_t \sim \mathrm{I}(d)$, where $-1/2 < d < 3/2$ and $d \neq 1/2$. Property (1.3) together with definition of $Y_n(r)$ and (2.1) implies that, for any $0 < r \leq 1$, as $n \to \infty$,

$$\mathrm{E}\left(Y_n^2(r)\right) \to s_\xi^2 r^{1+2d_\xi} = \mathrm{E}\left(Y_\infty^2(r)\right), \quad \max_{0 \leq r \leq 1} \mathrm{E}\left(Y_n^2(r)\right) \leq C. \tag{5.1}$$

Also, for any $0 < r \leq s \leq 1$, one has the equality

$$
\begin{aligned}
\mathrm{E}\left(Y_n^2(s)\right) &= \mathrm{E}(([Y_n(s) - Y_n(r)] + Y_n(r))^2) \\
&= \mathrm{E}((Y_n(s) - Y_n(r))^2) + 2\mathrm{E}(Y_n(s)Y_n(r)) - \mathrm{E}(Y_n^2(r)),
\end{aligned}
$$

where $\mathrm{E}((Y_n(s) - Y_n(r))^2) \sim \mathrm{E}((Y_n(s-r))^2)$, which together with (5.1) yields the equality

$$\mathrm{E}\left(Y_n(r)Y_n(s)\right) \to s_\xi^2 R_d(r,s) = \mathrm{E}\left(Y_\infty(r)Y_\infty(s)\right), \tag{5.2}$$

where $R_d(r,s) = \mathrm{E}\left(J_{1/2+d_\xi}(r)J_{1/2+d_\xi}(s)\right)$ is given by (2.3).

For $0 < \omega < \pi$, we compute the elements s_{ij} of the symmetric matrix $\boldsymbol{S}$ of (2.13) as

$$s_{11} = \sum_{t=1}^n 1 = n; \quad s_{12} = \sum_{t=1}^n t = n^2/2 + n/2;$$

$$s_{13} = \sum_{t=1}^n \sin(\omega t) = \frac{\cos(\omega/2) - \cos(\omega(n+1/2))}{2\sin(\omega/2)} = O(1); \tag{5.3}$$

$$s_{14} = \sum_{t=1}^n \cos(\omega t) = \frac{\sin(\omega(n+1/2))}{2\sin(\omega/2)} - \frac{1}{2} = O(1); \tag{5.4}$$

$$s_{22} = \sum_{t=1}^n t^2 = n^3/3 + n^2/2 + n/6;$$

$$s_{23} = \sum_{t=1}^n t\sin(\omega t) = n\frac{\sin(\omega n) - \sin(\omega(n+1))}{4\sin^2(\omega/2)} + \frac{\sin(\omega n)}{4\sin^2(\omega/2)} = O(n);$$

$$s_{24} = \sum_{t=1}^n t\cos(\omega t) = n\frac{\cos(\omega n) - \cos(\omega(n+1))}{4\sin^2(\omega/2)} - \frac{\sin^2(\omega n/2)}{2\sin^2(\omega/2)} = O(n);$$

$$s_{33} = \sum_{t=1}^n \sin^2(\omega t) = \frac{1}{2}\sum_{t=1}^n (1 - \cos(2\omega t)) = n/2 + O(1);$$

$$s_{34} = \sum_{t=1}^n \sin(\omega t)\cos(\omega t) = (1/2)\sum_{t=1}^n \sin(2\omega t) = O(1);$$

$$s_{44} = \sum_{t=1}^n \cos^2(\omega t) = \sum_{t=1}^n (1 - \sin^2(\omega t)) = n/2 + O(1); \tag{5.5}$$

by means of $\sin(x) = (\mathrm{e}^{\mathrm{i}x} - \mathrm{e}^{-\mathrm{i}x})/(2\mathrm{i})$ and $\cos(x) = (\mathrm{e}^{\mathrm{i}x} + \mathrm{e}^{-\mathrm{i}x})/2$. Denote $t_{n,p;u} = \sum_{t=1}^n z_{tp} u_t$, $p = 1, ..., 4$ and set

$$Z_{n,k} \equiv Z_{n,k;u} := m_k^{-1} t_{n,k;u}, \quad k = 1, ..., 4,$$

14

where

$$\begin{pmatrix} m_1 \\ m_2 \\ m_3 \\ m_4 \end{pmatrix} = \begin{pmatrix} n^{1/2+d} \\ n^{3/2+d} \\ n^{(1/2)\vee(d-1/2)} \\ n^{(1/2)\vee(d-1/2)} \end{pmatrix}.$$

Set $Z_n^* = |Z_{n,1}| + ... + |Z_{n,4}|$.

By the well-known property of LS estimators,

$$\widetilde{\boldsymbol{\beta}} \equiv (\widetilde{\beta}_1, \widetilde{\beta}_2, \widetilde{\beta}_3, \widetilde{\beta}_4)' = \boldsymbol{\beta} + \boldsymbol{S}^{-1} \begin{pmatrix} t_{n,1;u} \\ t_{n,2;u} \\ t_{n,3;u} \\ t_{n,4;u} \end{pmatrix}.$$

Inverting the symmetric matrix $\boldsymbol{S}$, defined by (2.13), and bearing in mind the formulae for s_{ij} above which are valid for $0 < \omega < \pi$, we obtain a symmetric matrix

$$\boldsymbol{S}^{-1} = n^{-1} \begin{pmatrix} 4 & -6n^{-1} & (-8s_{13}+12s_{23}n^{-1})n^{-1} & (-8s_{14}+12s_{24}n^{-1})n^{-1} \\ ... & 12n^{-2} & (12s_{13}-24s_{23}n^{-1})n^{-2} & (12s_{14}-24s_{24}n^{-1})n^{-2} \\ ... & ... & 2 & 0 \\ ... & ... & ... & 2 \end{pmatrix}$$

$$+ n^{-1} \begin{pmatrix} O(n^{-1}) & O(n^{-2}) & O(n^{-2}) & O(n^{-2}) \\ ... & O(n^{-4}) & O(n^{-3}) & O(n^{-3}) \\ ... & ... & O(n^{-2}) & O(n^{-1}) \\ ... & ... & ... & O(n^{-2}) \end{pmatrix}.$$

Then,

$$\boldsymbol{S}^{-1} \begin{pmatrix} t_{n,1;u} \\ t_{n,2;u} \\ t_{n,3;u} \\ t_{n,4;u} \end{pmatrix} = \boldsymbol{S}^{-1} \begin{pmatrix} n^{1/2+d}Z_{n,1;u} \\ n^{3/2+d}Z_{n,2;u} \\ n^{1/2\vee(d-1/2)}Z_{n,3;u} \\ n^{1/2\vee(d-1/2)}Z_{n,4;u} \end{pmatrix}$$

$$= \begin{pmatrix} n^{-1/2+d}[t_{\beta_1} + O(n^{-1\wedge(1+d)})Z_n^*] \\ n^{-3/2+d}[t_{\beta_2} + O(n^{-1\wedge(1+d)})Z_n^*] \\ n^{-3/2+d}[t_{\beta_3}] + n^{-1/2\wedge(3/2-d)}[2Z_{n,3;u} + O(n^{-1})Z_n^*] \\ n^{-3/2+d}[t_{\beta_4}] + n^{-1/2\wedge(3/2-d)}[2Z_{n,4;u} + O(n^{-1})Z_n^*] \end{pmatrix},$$

where

$$\begin{aligned} t_{\beta_1} &= 4Z_{n,1;u} - 6Z_{n,2;u}, \\ t_{\beta_2} &= -6Z_{n,1;u} + 12Z_{n,2;u}, \\ t_{\beta_3} &= s_{13}(-8Z_{n,1;u} + 12Z_{n,2;u}) + (s_{23}/n)(12Z_{n,1;u} - 24Z_{n,2;u}), \\ t_{\beta_4} &= s_{14}(-8Z_{n,1;u} + 12Z_{n,2;u}) + (s_{24}/n)(12Z_{n,1;u} - 24Z_{n,2;u}). \end{aligned}$$

Hence,

$$\begin{pmatrix} \tau_{\widetilde{\beta}_1} \\ \tau_{\widetilde{\beta}_2} \end{pmatrix} = \begin{pmatrix} t_{\beta_1} + O(n^{-1\wedge(1+d)})R_{n,1} \\ t_{\beta_2} + O(n^{-1\wedge(1+d)})R_{n,2} \end{pmatrix}; \tag{5.6}$$

15

$$\tau_{\widetilde{\beta}_j} = 2Z_{n,3;u} + n^{-1+d}t_{\beta_j} + n^{-1}R_{n,j}, \quad \text{if } -1/2 < d < 1 \tag{5.7}$$
$$= 2Z_{n,3;u} + t_{\beta_j} + n^{-1}R_{n,j}, \quad \text{if } 1 \le d < 3/2,$$

for $j = 3, 4$, where $\mathrm{E}|R_{n,j}|^2 < \infty$ for $j = 1, ..., 4$ because $\mathrm{E}|Z_{n,j;u}|^2 < \infty$ in view of Lemma 5.1 below.

In the case of the LS estimator $(\widehat{\beta}_1, \widehat{\beta}_2)$, we have

$$\begin{pmatrix} \widehat{\beta}_1 \\ \widehat{\beta}_2 \end{pmatrix} = \begin{pmatrix} \beta_1 \\ \beta_2 \end{pmatrix} + \begin{pmatrix} s_{11} & s_{12} \\ s_{21} & s_{22} \end{pmatrix}^{-1} \begin{pmatrix} t_{n,1;u*} \\ t_{n,2;u*} \end{pmatrix}, \tag{5.8}$$

where $u_t^* = u_j + \beta_3 \sin(t\omega) + \beta_4 \cos(t\omega)$. Then,

$$\begin{pmatrix} t_{n,1;u*} \\ t_{n,2;u*} \end{pmatrix} = \begin{pmatrix} t_{n,1;u} \\ t_{n,2;u} \end{pmatrix} + \begin{pmatrix} s_{13} + s_{14} \\ s_{23} + s_{24} \end{pmatrix} = \begin{pmatrix} t_{n,1;u} \\ t_{n,2;u} \end{pmatrix} + \begin{pmatrix} O(1) \\ O(n) \end{pmatrix}, \tag{5.9}$$

where $s_{13}, s_{14}, s_{23}, s_{24}$ are elements of $\boldsymbol{S}$ satisfying $|s_{13}| + |s_{14}| = O(1), |s_{23}| + |s_{24}| = O(n)$, as it was shown above in the derivation of the elements of $\boldsymbol{S}$. Since

$$\begin{pmatrix} s_{11} & s_{12} \\ s_{21} & s_{22} \end{pmatrix}^{-1} = n^{-1} \begin{pmatrix} 4 + O(n^{-2}) & -6n^{-1} + O(n^{-3}) \\ -6n^{-1} + O(n^{-2}) & 12n^{-2} + O(n^{-4}) \end{pmatrix},$$

then

$$\begin{pmatrix} s_{11} & s_{12} \\ s_{21} & s_{22} \end{pmatrix}^{-1} \begin{pmatrix} t_{n,1} \\ t_{n,2} \end{pmatrix} = \begin{pmatrix} s_{11} & s_{12} \\ s_{21} & s_{22} \end{pmatrix}^{-1} \begin{pmatrix} n^{1/2+d}Z_{n,1;u} \\ n^{3/2+d}Z_{n,2;u} \end{pmatrix}$$
$$= \begin{pmatrix} n^{-1/2+d}[4Z_{n,1;u} - 6Z_{n,2;u} + O(Z_n'/n)] \\ n^{-3/2+d}[-6Z_{n,1;u} + 12Z_{n,2;u} + O(Z_n'/n)] \end{pmatrix},$$

where $Z_n' = |Z_{n,1}| + |Z_{n,2}|$. Hence,

$$\begin{pmatrix} \widehat{\beta}_1 \\ \widehat{\beta}_2 \end{pmatrix} = \begin{pmatrix} \beta_1 \\ \beta_2 \end{pmatrix} + \begin{pmatrix} n^{-1/2+d}[t_{\beta_1} + O(Z_n'/n)] \\ n^{-3/2+d}[t_{\beta_2} + O(Z_n'/n)] \end{pmatrix}, \tag{5.10}$$

$$\begin{pmatrix} \tau_{\widehat{\beta}_1} \\ \tau_{\widehat{\beta}_2} \end{pmatrix} = \begin{pmatrix} t_{\beta_1} + n^{-1}r_{n,1} \\ t_{\beta_2} + n^{-1}r_{n,2} \end{pmatrix}, \tag{5.11}$$

where $\mathrm{E}|r_{n,j}| < \infty$ for $j = 1, 2$ since $\mathrm{E}|Z_{n,j;u}|^2 < \infty$ in view of Lemma 5.1 below. Hence, the LS estimators $(\widehat{\beta}_1, \widehat{\beta}_2)$ and $(\widetilde{\beta}_1, \widetilde{\beta}_2)$ are equivalent:

$$\begin{pmatrix} \tau_{\widetilde{\beta}_1} \\ \tau_{\widetilde{\beta}_2} \end{pmatrix} = \begin{pmatrix} \tau_{\widehat{\beta}_1} + O_p(n^{-1\vee(1+d)}) \\ \tau_{\widehat{\beta}_2} + O_p(n^{-1\vee(1+d)}) \end{pmatrix}. \tag{5.12}$$

Define the 4×4 matrix $\boldsymbol{\Psi}_d \equiv \boldsymbol{\Psi} = (\psi_{kj})_{k,j=1,...,4}$ as

$$\boldsymbol{\Psi} := \begin{pmatrix} s_\xi^2 & s_\xi^2/2 & 0 & 0 \\ s_\xi^2/2 & s_\xi^2(3+2d)^{-1} & 0 & 0 \\ 0 & 0 & \pi f_\xi(\omega) & 0 \\ 0 & 0 & 0 & \pi f_\xi(\omega) \end{pmatrix} \tag{5.13}$$

16

if $-1/2 < d < 1/2$, and

$$\boldsymbol{\Psi} := \begin{pmatrix} \frac{s_\xi^2}{1+2d} & \frac{s_\xi^2(6d-1)}{8d(1+2d)} & 0 & 0 \\ \frac{s_\xi^2(6d-1)}{8d(1+2d)} & \frac{s_\xi^2}{2(1+2d)}\left(1 - \frac{1}{(3+2d)d}\right) & 0 & 0 \\ 0 & 0 & \pi f_\xi(\omega)/4\sin^2(\omega/2) & 0 \\ 0 & 0 & 0 & \pi f_\xi(\omega)/4\sin^2(\omega/2) \end{pmatrix}$$
(5.14)

if $1/2 < d < 1$. Lemma 5.1 will show that, under general assumptions,

$$\left(\mathrm{cov}(Z_{n,k;u}, Z_{n,j;u}) \right)_{k,j=1,\dots,4} \to \boldsymbol{\Psi}.$$
(5.15)

Proof of Theorem 2.1. Since the estimator $\widehat{\beta}_2$, (2.4) is invariant with respect to a shift of X_t, we can set $u_0 = 0$ for $1/2 < d < 3/2$.

(i) Since the estimators $\widehat{\beta}_1$ and $\widehat{\beta}_2$ can be written as linear combinations of variables $Z_{n,1}$ and $Z_{n,2}$, (5.10) together with Lemma 5.1 implies the limits (2.8)-(2.9).

(ii)-(iii). In view of expansion (5.10) and (5.15),

$$\begin{pmatrix} \tau_{\widehat{\beta}_1} \\ \tau_{\widehat{\beta}_2} \end{pmatrix} = \begin{pmatrix} 4Z_{n,1} - 6Z_{n,2} \\ -6Z_{n,1} + 12Z_{n,2} \end{pmatrix} + \begin{pmatrix} o_p(1) \\ o_p(1) \end{pmatrix}.$$

Denote by (V_1, V_2) a zero mean Gaussian vector with the covariances

$$\mathrm{cov}(V_k, V_j) = \psi_{kj}, \quad k,j = 1,2.$$

We shall show that

$$(Z_{n,1;u}, Z_{n,2;u}) \overset{d}{\to} (V_1, V_2).$$
(5.16)

By the Cramer-Wold device, (5.16) holds if and only if for any real numbers a, b,

$$aZ_{n,1;u} + bZ_{n,2;u} \overset{d}{\to} aV_1 + bV_2.$$
(5.17)

Convergence (5.16) implies that

$$(\tau_{\widehat{\beta}_1}, \tau_{\widehat{\beta}_2}) \overset{d}{\to} (4V_1 - 6V_2, -6V_1 + 12V_2).$$

Since $\mathrm{var}(4V_1 - 6V_2) = \sigma_{1,d}^2$, $\mathrm{var}(-6V_1 + 12V_2) = \sigma_{2,d}^2$, and $\mathrm{cov}(4V_1 - 6V_2, -6V_1 + 12V_2) = \sigma_{12,d}$, where $\sigma_{1,d}^2$, $\sigma_{12,d}$, and $\sigma_{2,d}^2$ are given by (2.7)-(2.6), this completes proof of (2.10)-(2.11).

It remains to prove (5.17).

Case of $|d| < 1/2$. Then, $u_j = \xi_j$. Using the notation $Y_n(r)$, (2.2), and summation by parts, we can write,

$$t_{n,1;u} = \sum_{k=1}^{n} \xi_k = n^{1/2+d} Y_n(1) - \xi_{n+1},$$

$$t_{n,2;u} = \sum_{t=1}^{n} t\xi_t = -\sum_{t=1}^{n-1}\sum_{k=1}^{t} \xi_k + n\sum_{k=1}^{n} \xi_k$$

17

$$= n^{3/2+d}[-\int_0^1 Y_n(r)\mathrm{d}r + Y_n(1)] + [\sum_{k=1}^n \xi_k - n\xi_{n+1}]. \qquad (5.18)$$

Whence

$$Z_{n,1;u} = Y_n(1) + o_p(1), \quad Z_{n,2;u} = -\int_0^1 Y_n(r)\mathrm{d}r + Y_n(1) + o_p(1).$$

Case of $1/2 < d < 3/2$. Then, $d = 1 + d_\xi$, and the definition of the I(d) model implies $u_t = \sum_{j=1}^t \xi_j + u_0$, where by assumption $\mathrm{E}\left(u_0^2\right) < \infty$. Using notation (2.2), write

$$t_{n,1;u} = \sum_{t=1}^n u_t = \sum_{t=1}^n S_{t,1;\xi} + nu_0 = n^{3/2+d_\xi}\int_0^1 Y_n(r)\mathrm{d}r + nu_0, \qquad (5.19)$$

$$t_{n,2;u} = \sum_{t=1}^n tu_t = \sum_{t=1}^n tS_{t,1;\xi} + u_0\sum_{t=1}^n t \qquad (5.20)$$

$$= n^{3/2+d}\int_0^1 \frac{(\lfloor rn\rfloor + 1)}{n}Y_n(r)\mathrm{d}r + u_0 O(n^2)$$

$$= n^{3/2+d}[\int_0^1 rY_n(r)\mathrm{d}r + O_p(n^{-1})] + u_0 O(n^2),$$

in view of (5.1), observing that $u_0 O(n^2) = o_p(n^{3/2+d})$, since $\mathrm{E}\left(u_0^2\right) < \infty$ and $d > 1/2$. In this case,

$$Z_{n,1} = \int_0^1 Y_n(r)\mathrm{d}r + o_p(1), \quad Z_{n,2} = \int_0^1 rY_n(r)\mathrm{d}r + o_p(1).$$

Note that $\{Y_n(r), r \in [0,1]\}$, $n \geq 0$ is a sequence of real valued measurable processes with the paths from the space $L_2[0,1]$, i.e. $\int_0^1 Y_n^2(r)\mathrm{d}r < \infty$, and for any reals a, b, and c,

$$F_n := a\int_0^1 Y_n(r)\mathrm{d}r + b\int_0^1 rY_n(r)\mathrm{d}r + cY_n(1)$$

is continuous real valued bounded functional from $L_2[0,1]$ to $\mathbb{R}$. Relations (5.1) and (5.2) and Assumption FDD imply that the process $Y_n(r)$ satisfies the conditions of the weak convergence criterion in the space $L_2[0,1]$ by Cremers and Kadelka (1986), which implies the convergence to the Gaussian limit

$$F_n \xrightarrow{d} a\int_0^1 Y_\infty(r)\mathrm{d}r + b\int_0^1 rY_\infty(r)\mathrm{d}r + cY_\infty(1)$$

which proves (5.17).

(iv) Set $Z_{n,1} := \sum_{t=1}^n z_t$, $Z_{n,2} := \sum_{t=1}^n tz_t$. In (iv) it is assumed that $Z_{n,1} = o(n^{1/2+d})$. Summation by parts yields

$$Z_{n,2} = -\sum_{k=1}^{n-1} Z_{k,1} + nZ_{n,1} + \sum_{k=1}^{n-1} o(k^{1/2+d}) + no(n^{1/2+d}) = o(n^{3/2+d}).$$

Denote $\widehat{\beta}_1^0, \widehat{\beta}_2^0$ the OLS estimators corresponding to case $z_t = 0$. Then (5.8)-(5.9) imply that $\widehat{\beta}_i = \widehat{\beta}_i^0 + R_i$, $i = 1, 2$ where

$$
\begin{aligned}
|R_1| &\leq C(n^{-1}|Z_{n,1}| + n^{-2}|Z_{n,2}|) = o(n^{-1/2+d}), \\
|R_2| &\leq C(n^{-2}|Z_{n,1}| + n^{-3}|Z_{n,2}|) = o(n^{-3/2+d}).
\end{aligned}
$$

Since results (i) and (ii) of Theorem 2.1 are valid for $\widehat{\beta}_i^0$, $i = 1, 2$, they remain true for $\widehat{\beta}_i$, $i = 1, 2$, because the terms R_1 and R_2 are negligible.

LEMMA 5.1. *Assume that $u_t \sim \mathrm{I}(d)$, with $-1/2 < d < 3/2$, $d \neq 1/2$, and let $n \to \infty$.*
  If $-1/2 < d < 1$, then

$$
\mathrm{cov}(Z_{n,k;u}, Z_{n,j;u}) \to s_\xi^2 \psi_{kj}, \quad k, j = 1, ..., 4, \tag{5.21}
$$

where ψ_{kj} are entries of the matrix $\mathbf{\Psi}$, defined by (5.13)-(5.14).
  If $1 \leq d < 3/2$, then

$$
\begin{aligned}
\mathrm{cov}(Z_{n,k;u}, Z_{n,j;u}) &\to s_\xi^2 \psi_{kj}, \quad k, j = 1, 2, & (5.22) \\
\mathrm{var}(Z_{n,k;u}) &= \frac{n^{2(1-d)}\pi f_\xi(\omega)}{4\sin^2(\omega/2)}(1 + o(1)) + s_\xi^2 \gamma_k^2, \quad k = 3, 4, & (5.23) \\
\mathrm{cov}(Z_{n,1;u}, Z_{n,k;u}) &= s_\xi^2 \gamma_k/2 + o(1), \quad k = 3, 4, & (5.24) \\
\mathrm{cov}(Z_{n,2;u}, Z_{n,k;u}) &= s_\xi^2 v\gamma_k + o(1), \quad k = 3, 4, & (5.25)
\end{aligned}
$$

where ψ_{kj} are entries of matrix (5.14), and γ_k, v are defined in (2.16).

  Proof of Lemma 5.1. Denote for $k, j = 1, ..., 4$,

$$
R_{n,k,j} := \mathrm{E}\left(Z_{n,k;u} Z_{n,j;u}\right) = m_k^{-1} m_j^{-1} \mathrm{E}\left(t_{n,k;u} t_{n,j;u}\right) = m_k^{-1} m_j^{-1} \mathrm{E}\left(\sum_{t=1}^n z_{tk} u_t \sum_{s=1}^n z_{sj} u_s\right).
$$

  Case of $|d| < 1/2$. Then, $u_j = \xi_j$, $d = d_\xi$, and $t_{n,k;u} \equiv t_{n,k;\xi} = \sum_{t=1}^n z_{tk}\xi_t$. Properties (5.1)-(5.2) of $Y_n(r)$ and the dominated convergence theorem imply that, for $p, l = 0, 1$, as $n \to \infty$,

$$
\begin{aligned}
\mathrm{E}\left(\int_0^1 r^p Y_n(r)\mathrm{d}r\right)\left(\int_0^1 Y_n(v)v^l \mathrm{d}v\right) &\to s_\xi^2 \int_0^1 \int_0^1 r^p R_d(r, v)v^l \mathrm{d}r\mathrm{d}v; & (5.26) \\
\mathrm{E}\left(Y_n(1)\int_0^1 r^p Y_n(r)\mathrm{d}r\right) &\to s_\xi^2 \int_0^1 r^p R_d(r, 1)\mathrm{d}r; \\
\mathrm{E}\left(Y_n^2(1)\right) &\to s_\xi^2 R_d(1, 1) = s_\xi^2.
\end{aligned}
$$

Writing $t_{n,1;\xi}$, $t_{n,2;\xi}$ as in (5.18), (5.26) implies

$$
\begin{aligned}
R_{n,1,1} &= n^{-1-2d}\mathrm{E}(t_{n,1;\xi})^2 \to s_\xi^2; \\
R_{n,2,2} &= \mathrm{E}(n^{-3/2-d}t_{n2})^2 = \mathrm{E}\left(Y_n(1) - \int_0^1 Y_n(r)\mathrm{d}r\right)^2 + o(1)
\end{aligned}
$$

19

$$\rightarrow \quad s_\xi^2\Big(R_d(1,1) - 2\int_0^1 R_d(1,r)\mathrm{d}r + \int_0^1\int_0^1 R_d(r,v)\mathrm{d}r\mathrm{d}v\Big)$$

$$= \quad s_\xi^2(1 - 1 + (3 + 2d)^{-1}) = s_\xi^2(3 + 2d)^{-1};$$

$$R_{n,1,2} \quad = \quad \mathrm{E}\Big(n^{-1/2-d}t_{n,1}n^{-3/2-d}t_{n,2}\Big)$$

$$= \quad \mathrm{E}\Big(Y_n(1) - n^{-1/2-d}\xi_{n+1}\Big)\Big(-\int_0^1 Y_n(r)\mathrm{d}r + Y_n(1)\Big) + o(1)$$

$$\rightarrow \quad s_\xi^2\Big(-\int_0^1 R_d(1,r)\mathrm{d}r + R_d(1,1)\Big) = s_\xi^2(-1/2 + 1) = s_\xi^2/2.$$

To obtain the remaining $R_{n,k,j}$ with $k,j = 3,4$, set $D_{n,k} := \sum_{t=1}^n \mathrm{e}^{\mathrm{i}tx}z_{tk}$, $k = 1,...,4$ and write

$$R_{n,k,j} = (m_k m_j)^{-1}\sum_{t,s=1}^n z_{tk}z_{sj}\mathrm{E}\,(\xi_t\xi_s) = (m_k m_j)^{-1}\int_{-\pi}^\pi D_{n,k}(x)D_{n,j}(-x)f_\xi(x)\mathrm{d}x. \quad (5.27)$$

Denote by

$$D_n(x) := \sum_{t=1}^n \mathrm{e}^{\mathrm{i}tx} = \mathrm{e}^{\mathrm{i}x(n+1)/2}\frac{\sin(nx/2)}{\sin(x/2)}$$

the Dirichlet kernel. Then, since $\sin(x) = (\mathrm{e}^{\mathrm{i}x} - \mathrm{e}^{-\mathrm{i}x})/\,(2\mathrm{i})$ and $\cos(x) = (\mathrm{e}^{\mathrm{i}x} + \mathrm{e}^{-\mathrm{i}x})/2$,

$$D_{n,1}(x) \quad = \quad D_n(x) \tag{5.28}$$

$$D_{n,3}(x) \quad = \quad \sum_{t=1}^n \mathrm{e}^{\mathrm{i}tx}\sin(t\omega) = (D_n(x+\omega)\mathrm{e}^{\mathrm{i}\omega} - D_n(x-\omega)\mathrm{e}^{-\mathrm{i}\omega})/(2\mathrm{i}), \tag{5.29}$$

$$D_{n,4}(x) \quad = \quad \sum_{t=1}^n \mathrm{e}^{\mathrm{i}tx}\cos(t\omega) = (D_n(x+\omega)\mathrm{e}^{\mathrm{i}\omega} + D_n(x-\omega)\mathrm{e}^{-\mathrm{i}\omega})/2,$$

which implies

$$|D_{n,k}(x)|^2 = (|D_n(x+\omega)|^2 + |D_n(x-\omega)|^2)/4 + O(|D_n(x+\omega)|\,|D_n(x-\omega)|),$$
$$|\mathrm{Re}[D_{n,3}(x)D_{n,4}(-x)]| \le C|D_n(x-\omega)D_n(x+\omega)|.$$

For any $\epsilon > 0$,

$$|D_n(x)| \le C, \quad \epsilon \le |x| \le \pi; \quad |D_n(x)| \le C\frac{n}{1 + n|x|}, \quad |x| \le 2\pi - \epsilon. \tag{5.30}$$

If f_ξ is continuous at ω, then a standard argument, using (1.3) and (5.30), implies that, as $n \to \infty$,

$$n^{-1-2d_\xi}\int_{-\pi}^\pi |D_n(x)|^2 f_\xi(x)\mathrm{d}x \to s_\xi^2,$$

$$n^{-1}\int_{-\pi}^\pi |D_n(x\pm\omega)|^2 f_\xi(x)\mathrm{d}x \to 2\pi f_\xi(\omega),$$

$$\int_{-\pi}^\pi |D_n(x+\omega)|\,|D_n(x-\omega)|f_\xi(x)\mathrm{d}x \le C\log n,$$

$$\int_{-\pi}^{\pi} |D_n(x)|\,|D_n(x \pm \omega)| f_\xi(x) \mathrm{d}x \le C(\log n + n^{2d_\xi \vee 0}).$$

Therefore,

$$
\begin{aligned}
R_{n,k,k} &= n^{-1} \int_{-\pi}^{\pi} |D_{n,k}(x)|^2 f_\xi(x) \mathrm{d}x = n^{-1} \int_{-\pi}^{\pi} \Big[(|D_n(x+\omega)|^2 + |D_n(x-\omega)|^2)/4 \\
&+ O(|D_n(x+\omega)|\,|D_n(x-\omega)|) \Big] f_\xi(x) \mathrm{d}x \to \pi f_\xi(\omega), \quad k = 3,4, \qquad (5.31)
\end{aligned}
$$

$$
\begin{aligned}
|R_{n,3,4}| &\le C n^{-1} \int_{-\pi}^{\pi} |D_n(x+\omega)|\,|D_n(x-\omega)| f_\xi(x) \mathrm{d}x \to 0, \\
|R_{n,1,k}| &\le C n^{-1-d} \int_{-\pi}^{\pi} |D_n(x)|\,(|D_n(x+\omega)| + |D_n(x-\omega)|) f_\xi(x) \mathrm{d}x \qquad (5.32) \\
&\to 0, \quad k = 3,4.
\end{aligned}
$$

It remains to obtain $R_{n,2,k}, k = 3,4$. By (5.18), write $t_{n,2;u}$ as

$$t_{n,2;\xi} = -\sum_{j=1}^{n-1} S_{j,1;\xi} + n t_{n,1;\xi}.$$

By (5.30),
$$\max_{1 \le j \le n} |D_j(x)| \le C \max_{1 \le j \le n} j/(1 + j|x|) \le n/(1 + n|x|).$$

Hence, by (5.27), (5.28), and (1.3),

$$
\begin{aligned}
\max_{1 \le j \le n} |\mathrm{E}\,(S_{j,1;\xi} t_{n,k;\xi})| &\le \max_{1 \le j \le n} \int_{-\pi}^{\pi} |D_j(x)|\,(|D_n(x+\omega)| + |D_n(x-\omega)|) f_\xi(x) \mathrm{d}x \\
&\le C \int_{-\pi}^{\pi} \frac{n}{1+n|x|}\,(|D_n(x+\omega)| + |D_n(x-\omega)|) f_\xi(x) \mathrm{d}x \le C(\log n + n^{2d_\xi \vee 0}). \quad (5.33)
\end{aligned}
$$

Therefore, for $k = 3,4$,

$$|R_{n,2,k}| \le n^{-2-d_\xi} \Big(\sum_{j=1}^{n-1} |\mathrm{E}\,(S_{j,1;\xi} t_{n,k;\xi})| + n|\mathrm{E}\,(t_{n,1;\xi} t_{n,k;\xi})| \Big)$$

$$\le C n^{-2-d_\xi}(n \log n + n^{2d_\xi \vee 0} n) \to 0,$$

which completes proof of (5.21) with $\mathbf{\Psi}$ as in (5.13), corresponding to $|d| < 1/2$.

   Case of $1/2 < d < 3/2$. Recall that $d = 1 + d_\xi$ and $\mathrm{E}\,(u_0^2) = O(1)$. Consider first $R_{n,k,j}, \ k,j = 1,2$. Then by (5.19), (5.20), and (5.26), for $k,j = 1,2$, with $m_1 = n^{1/2+d}$ and $m_2 = n^{3/2+d}$,

$$R_{n,k,j} = (m_k m_j)^{-1} \mathrm{E}\Big[\Big(\sum_{t=1}^{n} t^{k-1} u_t\Big)\Big(\sum_{s=1}^{n} s^{j-1} u_s\Big) \Big] \to \psi_{kj} = s_\xi^2 \int_0^1 \int_0^1 r^{k-1} R_{d_\xi}(r,v) v^{j-1} \mathrm{d}r \mathrm{d}v,$$

which yields

$$\psi_{11} = \frac{s_\xi^2}{1+2d}; \quad \psi_{12} = s_\xi^2 \frac{6d-1}{8d(1+2d)}; \quad \psi_{22} = s_\xi^2 \frac{1}{2(1+2d)}\left(1 - \frac{1}{d(3+2d)}\right).$$

21

Similarly, we obtain the convergence

$$m_k^{-1} n^{-1/2-d_\xi} \mathrm{E}\left(t_{n,k;u} t_{n,1;\xi}\right) \quad \sim \quad \mathrm{E}\left[\left(\sum_{t=1}^{n} t^{k-1} u_t\right)(u_n - u_0)\right] \tag{5.34}$$

$$\rightarrow \quad s_\xi^2 \int_0^1 r^{k-1} R_{d_\xi}(r,1)\mathrm{d}r = s_\xi^2/2, \text{ if } k = 3;$$

$$= s_\xi^2\left(\frac{1}{4} + \frac{1}{1+2d} - \frac{1}{4d}\right) = s_\xi^2 v, \text{ if } k = 4,$$

which will be used below.

Next we obtain $R_{n,k,j}$, $k,j = 3,4$. The definition of the I(d) model implies $u_t = \sum_{j=1}^{t} \xi_j + u_0$. Then,

$$t_{n,k;u} = \sum_{t=1}^{n} z_{tk} u_t = \sum_{t=1}^{n} z_{tk} \sum_{j=1}^{t} \xi_j + u_0 \sum_{t=1}^{n} z_{tk} = \sum_{j=1}^{n}\left(\sum_{t=j}^{n} z_{tk}\right)\xi_j + u_0 \sum_{t=1}^{n} z_{tk}. \tag{5.35}$$

For $k = 3,4$, $z_{t3} = \sin(\omega t)$ and $z_{t4} = \cos(\omega t)$. Using (5.3)-(5.4), it follows that $|\sum_{t=1}^{n} z_{tk}| \leq C$ and

$$\sum_{t=j}^{n} \sin(\omega t) = \frac{\cos(\omega(j-1/2)) - \cos(\omega(n+1/2))}{2\sin(\omega/2)};$$

$$\sum_{t=j}^{n} \cos(\omega t) = \frac{\sin(\omega(n+1/2)) - \sin(\omega(j-1/2))}{2\sin(\omega/2)}.$$

Since $\cos(\omega(j-1/2)) = \cos(\omega j)\cos(\omega/2) + \sin(\omega j)\sin(\omega/2)$ and $\sin(\omega(j-1/2)) = \sin(\omega j)\cos(\omega/2) - \cos(\omega j)\sin(\omega/2)$, then

$$t_{n,3;u} = \frac{t_{n,4;\xi}\cos(\omega/2) + t_{n,3;\xi}\sin(\omega/2) - \cos(\omega(n+1/2))t_{n,1;\xi}}{2\sin(\omega/2)} + |u_0|O(1), \tag{5.36}$$

$$t_{n,4;u} = \frac{\sin(\omega(n+1/2))t_{n,1;\xi} - t_{n,3;\xi}\cos(\omega/2) + t_{n,4;\xi}\sin(\omega/2)}{2\sin(\omega/2)} + |u_0|O(1). \tag{5.37}$$

Using (5.31)-(5.32), we obtain

$$\mathrm{E}\left(t_{n,3;u}^2\right) = \frac{\mathrm{E}\left(t_{n,4;\xi}^2\right)\cos^2(\omega/2) + \mathrm{E}\left(t_{n,3;\xi}^2\right)\sin^2(\omega/2) + \cos^2(\omega(n+1/2))\mathrm{E}\left(t_{n,1;\xi}^2\right) + o(n)}{4\sin^2(\omega/2)}$$

$$= \frac{n\pi f_\xi(\omega) + \cos^2(\omega(n+1/2))s_\xi^2 n^{1+2d_\xi}}{4\sin^2(\omega/2)} + o(n)$$

$$= n\frac{\pi f_\xi(\omega)}{4\sin^2(\omega/2)} + n^{2d-1}s_\xi^2\gamma_3^2 + o(n).$$

A similar argument shows that

$$\mathrm{E}\left(t_{n,4;u}^2\right) = n\frac{\pi f_\xi(\omega)}{4\sin^2(\omega/2)} + n^{2d-1}s_\xi^2\gamma_4^2 + o(n),$$

22

which proves that for $k = 3, 4$,

$$
\begin{aligned}
R_{n,k,k} \quad &\rightarrow \quad \frac{\pi f_\xi(\omega)}{4\sin^2(\omega/2)}, \quad \text{if} \quad (1/2 < d < 1) \\
&= \quad \frac{n^{2(1-d)}\pi f_\xi(\omega)}{4\sin^2(\omega/2)}(1 + o(1)) + s_\xi^2\gamma_k^2, \quad \text{if} \quad (1 \le d < 3/2),
\end{aligned}
$$

which proves (5.22) and (5.23).

Next, for $1/2 < d < 1$,

$$
\begin{aligned}
|R_{n,3,4}| &= n^{-1}|\mathrm{E}\,(t_{n,3;u}t_{n,4;u})\,| \\
&\le \quad Cn^{-1}(|\mathrm{E}\,(t_{n,4;\xi}^2) - \mathrm{E}\,(t_{n,3;\xi}^2)\,| + |\mathrm{E}\,(t_{n,3;\xi}t_{n,4;\xi})\,| + |\mathrm{E}\,(t_{n,1;\xi}t_{n,3;\xi})\,| + |\mathrm{E}\,(t_{n,1;\xi}t_{n,4;\xi})\,|) \\
&\le \quad Cn^{-1}(o(n) + O(\log n) + o(n)) = o(1).
\end{aligned}
$$

Finally, we show that $R_{n,k,j} = o(1)$ for $k = 1, 2$, $j = 3, 4$, when $1/2 < d < 1$. By (5.36)-(5.37),

$$
|\mathrm{E}\,(t_{n,k;u}t_{n,j;u})\,| \le C(|\mathrm{E}\,(t_{n,k;u}t_{n,1,\xi})\,| + |\mathrm{E}\,(t_{n,k;u}t_{n,3;\xi})\,| + |\mathrm{E}\,(t_{n,k;u}t_{n,4;\xi})\,| + \mathrm{E}\,(|u_0 t_{n,k;u}|)).
$$

By (5.34),

$$
|\mathrm{E}\,(t_{n,k;u}t_{n,1,\xi})\,| \le Cm_k n^{d-1/2}.
$$

Observe that for $j = 1, ..., 4$, $|\mathrm{E}\,(u_0^2 t_{n,l;u})\,| \le (\mathrm{E}\,(u_0^2)\,\mathrm{E}\,(t_{n,l;u}^2))^{1/2} \le Cm_l$, since $\mathrm{E}\,(u_0^2) < \infty$ and $\mathrm{E}\,(t_{n,l;u}^2) \le Cm_l^2$, as it was shown above.

Using (5.19)-(5.20) and (5.33), we obtain for $j = 3, 4$,

$$
\begin{aligned}
|\mathrm{E}\,(t_{n,k;u}t_{n,j;\xi})\,| &\le C\sum_{t=1}^{n} t^{k-1}|\mathrm{E}\,(S_{t,k;\xi}t_{n,j;\xi})\,| + O(n^k)\mathrm{E}\,(|u_0 S_{t,j;\xi}|) \\
&\le \quad C\left(n^{k-1}\sum_{t=1}^{n}(\log n + n^{2d_\xi \vee 0}) + n^{k+1/2}\right) = O(n^k(n^{2d_\xi \vee 0} + n^{1/2})),
\end{aligned}
$$

since $m_3 = m_4 = n^{1/2}$. Since for $j = 3, 4$, $m_1 m_j = n^{1+d}$, $m_2 m_j = n^{2+d}$, then

$$
|R_{n,k,j}| \le C(m_k m_j)^{-1}(m_k n^{d-1/2} + n^{k+2d_\xi \vee 0} + n^{k+1/2} + m_k) = o(1),
$$

which completes proof of (5.21).

To prove (5.24) and (5.25) in the case of $d \ge 1$, recall that $m_3 = m_4 = n^{d-1/2}$, $m_1 = n^{1/2+d}$, and $m_2 = n^{3/2+d}$. Then, by (5.36)-(5.37),

$$
\begin{aligned}
Z_{n,3;u} &= \quad \frac{1}{2n^{d-1}}\left(Z_{n,4;\xi}\frac{\cos(\omega/2)}{\sin(\omega/2)} + Z_{n,3;\xi}\right) + \gamma_3 Z_{n,1;\xi} + |u_0|n^{-d+1/2}; \\
Z_{n,4;u} &= \quad \frac{1}{2n^{d-1}}\left(Z_{n,4;\xi} - Z_{n,3;\xi}\frac{\cos(\omega/2)}{\sin(\omega/2)}\right) + \gamma_4 Z_{n,1;\xi} + |u_0|n^{-d+1/2}.
\end{aligned}
$$

In (5.34) it was shown that

$$
\mathrm{E}\,(Z_{n,1;u}Z_{n,1;\xi}) \rightarrow s_\xi^2/2, \quad \mathrm{E}\,(Z_{n,2;u}Z_{n,1;\xi}) \rightarrow s_\xi^2 v.
$$

23

We show that
$$\mathrm{E}\left(Z_{n,j;u}Z_{n,k;\xi}\right) \to 0, \quad j = 1, 2, \ k = 3, 4, \tag{5.38}$$
which together with estimates above implies (5.24) and (5.25).

By (5.19), (5.20), and (5.33),

$$
\begin{aligned}
\left|\mathrm{E}\left(Z_{n,j;u}Z_{n,k;\xi}\right)\right| &\leq m_j^{-1}m_k^{-1}\Big(\sum_{t=1}^{n} t^{j-1}\left|\mathrm{E}\left(S_{t,j;u}t_{n,k;\xi}\right)\right| + n^j\left|\mathrm{E}\left(u_0 t_{n,k;\xi}\right)\right|\Big) \\
&\leq Cn^{-2d}\sum_{t=1}^{n}(\log n + n^{2d_\xi \vee 0}) + m_j^{-1}n^j(\mathrm{E}\left(u_0^2\right)\mathrm{E}\left(Z_{n,k;\xi}^2\right))^{1/2} \to 0,
\end{aligned}
$$

which proves (5.38) and completes the proof of Lemma 5.1.

Proof of Theorem 2.2.

1. Proof of (i)-(ii). By (5.6),

$$
\begin{aligned}
\tau_{\widetilde{\beta}_1} &= 4Z_{n,1;u} - 6Z_{n,2;u} + O(n^{-1\wedge(1+d)})R_{n,1} = 4Z_{n,1;u} - 6Z_{n,2;u} + o_p(1), \\
\tau_{\widetilde{\beta}_2} &= -6Z_{n,1;u} + 12Z_{n,2;u} + O(n^{-1\wedge(1+d)})R_{n,2} = -6Z_{n,1;u} + 12Z_{n,2;u} + o_p(1),
\end{aligned}
$$

where $\mathrm{E}|R_{n,j}^2| \leq C$, $j = 1, 2$. Therefore the same argument used in the proof of Theorem 2.1 implies statements (i) and (ii) of Theorem 2.2 about $\widetilde{\beta}_1$ and $\widetilde{\beta}_2$.

By (5.7), for $k = 3, 4$ and $-1/2 < d < 3/2$,

$$\sqrt{A_k}\tau_{\widetilde{\beta}_k} = n^{0\wedge(-1+d)}(\theta_{1k}Z_{n,1;u} + \theta_{2k}Z_{n,2;u}) + 2Z_{n,k;u} + n^{-1}R_{n,j}, \tag{5.39}$$

where θ_{1k}, θ_{2k} are defined in (2.16), and $\mathrm{E}|R_{n,k}^2| \leq C$. Hence, for $k = 3, 4$, (i) follows from (5.39), applying Lemma 5.1.

2. Proof of (iii). Assume that $-1/2 < d < 1$. Since $\tau_{\widetilde{\beta}_1}$, $\tau_{\widetilde{\beta}_2}$, $\tau_{\widetilde{\beta}_3}$, and $\tau_{\widetilde{\beta}_4}$ are linear combinations of the variables $Z_{n,j;u}, j = 1, ..., 4$, convergence (2.17) and (2.20) follow from (5.40) of Lemma 5.2 below by the same argument as that used in the proof of Theorem 2.1.

If $1 \leq d < 3/2$, then (2.18) and (2.19) follow from (5.41) and Lemma 5.1.

LEMMA 5.2. *Assume that the conditions of Theorem 2.2 (iii) are satisfied.*

(i) If $-1/2 < d < 1$, then for any real numbers a_j, $j = 1, ..., 4$,

$$t_n := \sum_{k=1}^{4} a_k Z_{n,k;u} \xrightarrow{d} \sum_{k=1}^{4} a_k V_k =: Z, \quad n \to \infty, \tag{5.40}$$

where $(V_1, ..., V_4) \sim \mathrm{N}(0, \boldsymbol{\Psi})$ with $\boldsymbol{\Psi}$ given by (5.13).

(ii) If $1 \leq d < 3/2$, then

$$(\mathrm{var}(t_n))^{-1/2}t_n \xrightarrow{d} \mathrm{N}(0, 1). \tag{5.41}$$

Proof. (i) For $-1/2 \leq d < 1$, by Lemma 5.1, the matrix $\boldsymbol{\Psi}$ is the limit covariance matrix of the vector $\boldsymbol{t}$. Therefore $\sigma_n^2 := \mathrm{var}(t_n) \to \sigma^2 := \mathrm{var}(Z)$ as $n \to \infty$. It remains to show

$$\sigma_n^{-1}t_n \xrightarrow{d} \mathrm{N}(0, 1). \tag{5.42}$$

24

We shall show that t_n can be written as

$$t_n = \sum_{j=-\infty}^{n} c_{n,j}\varepsilon_j + o_p(1) \qquad (5.43)$$

with some weights $c_{n,j}$ such that

$$c_n^* := \max_{-\infty < j \le n} |c_{n,j}| = o(1). \qquad (5.44)$$

First we show that (5.43) and (5.44) imply (5.42). Assume first that $\mathrm{E}|\varepsilon_t|^k < \infty$ for all $k \ge 1$. To prove (5.42), it suffices to show for any $k \ge 3$, the k-th order cumulant of $\sigma_n^{-1}t_n$ satisfies

$$\kappa_k \to 0, \quad n \to \infty.$$

Since $\{\varepsilon_t\}$ is a sequence of i.i.d. variables, then

$$\kappa_k = \kappa_k(\varepsilon_0) \sum_{j=-\infty}^{n} c_{n,j}^k,$$

where $\kappa_k(\varepsilon_0)$ is the k-th order cumulant of the variable ε_0. Since $\sigma_n^2 \to \sigma^2 > 0$, then by (5.44), for $k \ge 3$,

$$|\kappa_k| \le |\kappa_k(\varepsilon_0)|(c_n^*)^{k-2}\sigma_n^{-k} \sum_{j=-\infty}^{n} c_{n,j}^2 \le |\kappa_k(\varepsilon_0)|(c_n^*/\sigma_n)^{-k+2} \to 0,$$

to prove (5.42).

Assume now that not all moments of ε_j are finite. Fix $K > 0$, set $\varepsilon_j^- = 1_{|\varepsilon_j| \le K}\varepsilon_j$, $\varepsilon_j^+ = 1_{|\varepsilon_j| > K}\varepsilon_j$, and write

$$t_n = S_n^- + S_n^+, \quad S_n^- := \sum_{j=-\infty}^{n} c_{n,j}\varepsilon_j^-, \quad S_n^- := \sum_{j=-\infty}^{n} c_{n,j}\varepsilon_j^+.$$

As we showed above, for any fixed $K > 0$, $S_n^- \xrightarrow{d} Z_K$, where $Z_K \sim \mathrm{N}(0, \sigma^2\mathrm{var}(\varepsilon_j^-))$, and $\mathrm{var}(\varepsilon_j^-) \to \mathrm{var}(\varepsilon_j) = 1$ as $K \to \infty$. On the other hand, as $n \to \infty$,

$$\mathrm{var}(S_n^+) = \mathrm{var}(\varepsilon_1^+) \sum_{j=-\infty}^{n} c_{n,j}^2 \sim \mathrm{var}(\varepsilon_1^+)\sigma^2 \to 0, \quad K \to \infty$$

which in view of a standard argument implies (5.42).

To prove (5.43) and (5.44) it suffices to show that $Z_{n,k;u}$, $k = 1, ..., 4$ can be written as

$$Z_{n,k;u} = \sum_{j=-\infty}^{n} c_{n,j}^{(k)}\varepsilon_j + o_p(1), \quad k = 1, ..., 4$$

with weights $c_{n,j}^{(k)}$ satisfying

$$\max_{-\infty < j \le n} |c_{n,j}^{(k)}| = o(1). \qquad (5.45)$$

25

1. Assume that $-1/2 < d < 1/2$. Then $u_t = \xi_t$, where $\{\xi_t\}$ is a linear process (2.14), and

$$Z_{n,k;u} = Z_{n,k;\xi} = m_k^{-1} \sum_{t=1}^{n} z_{tk}\xi_t = \sum_{j=-\infty}^{n} c_{n,j}^{(k)}\varepsilon_j$$

where

$$c_{n,j}^{(k)} = m_k^{-1} \sum_{t=j\vee 1}^{n} z_{tk}a_{t-j}.$$

First, we show that

$$\max_{1\le k\le n} |\sum_{t=1}^{k} a_t| \le C(\log n + n^{d_\xi\vee 0}).$$

Since a_t is square summable, it can be written in the form $a_t = \int_{-\pi}^{\pi} e^{itx}\widehat{a}(x)dx$, where $\widehat{a}(x)$ is a square integrable function. In addition, the spectral density of ξ_t can be written as $f_\xi(x) = 2\pi|\widehat{a}(x)|^2$. Hence, for $1 \le k \le n$, by (5.30) and (1.3),

$$
\begin{aligned}
|\sum_{t=1}^{k} a_t| &= |\int_{-\pi}^{\pi} \sum_{t=1}^{k} e^{itx}\widehat{a}(x)dx| \le \int_{-\pi}^{\pi} |D_k(x)||\widehat{a}(x)|dx \le C\int_{-\pi}^{\pi} \frac{k}{1+k|x|}\sqrt{f_\xi(x)}dx \\
&\le C(\log k + k^{d_\xi\vee 0}) \le C(\log n + n^{d_\xi\vee 0}).
\end{aligned}
\tag{5.46}
$$

Therefore,

$$|c_{n,j}^{(1)}| = n^{-1/2-d_\xi}|\sum_{t=j\vee 1}^{n} a_{t-j}| \le Cn^{-1/2-d_\xi}(\log n + n^{d_\xi\vee 0}) = o(1).$$

Using summation by parts and (5.46), we obtain that

$$|\sum_{t=j\vee 1}^{n} ta_{t-j}| \le \sum_{k=1}^{n-1}|\sum_{t=j\vee 1}^{k} a_{t-j}| + n|\sum_{t=j\vee 1}^{n} a_{t-j}| \le Cn(\log n + n^{d_\xi\vee 0}),$$

which implies

$$|c_{n,j}^{(2)}| = n^{-3/2-d_\xi}|\sum_{t=j\vee 1}^{n} ta_{t-j}| \le Cn^{-1/2-d_\xi}(\log n + n^{d_\xi\vee 0}) = o(1).$$

Finally, for $k = 3, 4$, bearing in mind that $z_{t3} = \sin(\omega t)$, $z_{t4} = \cos(\omega t)$,

$$
\begin{aligned}
|c_{n,j}^{(k)}| &= n^{-1/2}|\sum_{t=j\vee 1}^{n} z_{tk}a_{t-j}| \le n^{-1/2}\sum_{t=j\vee 1}^{n} |a_{t-j}| \\
&\le n^{-1/2}\sum_{t=j\vee 1:|t-j|\le\log n}^{n} |a_{t-j}| + n^{-1/2}\sum_{t=j\vee 1:|t-j|>\log n}^{n} |a_{t-j}| \\
&\le Cn^{-1/2}\log n + (\sum_{s>\log n}^{\infty} |a_s|^2)^{1/2} \to 0, \quad n \to \infty,
\end{aligned}
$$

26

to complete proof of (5.45).

2. Assume that $1/2 < d < 1$. By (5.35), we can write

$$Z_{n,k;u} = m_k^{-1} \sum_{t=1}^{n} z_{tk}^* \xi_t + o_p(1) = \sum_{j=-\infty}^{n} c_{n,j}^{(k)} \varepsilon_j + o_p(1),$$

where

$$c_{n,j}^{(k)} = m_k^{-1} \sum_{t=j\vee 1}^{n} z_{tk}^* a_{t-j}, \quad z_{tk}^* = \sum_{s=t}^{n} z_{sk}.$$

It remains to show that $c_{n,j}^{(k)}$ satisfies (5.45). For $k = 1, 2$, we have $z_{j1} = 1$ and $z_{jk} = j$, and summation by parts gives

$$|c_{n,k}(k)| \le m_k^{-1}\Big(\sum_{s=1}^{n-1} |z_{sk}|| \sum_{t=j\vee 1}^{s} a_{t-j}| + |z_{nk}|| \sum_{t=j\vee 1}^{n} a_{t-j}|\Big) \le C m_k^{-1} n^k (\log n + n^{d_\xi \vee 0}) = o(1).$$

For $k = 3, 4$, (5.3)-(5.4) imply that $|z_{tk}^*| \le C$, and (5.45) follows using the same argument as in the case $-1/2 < d < 1/2$.

(ii) The proof of (5.41) follows using the same argument as (5.40) in the case of $1/2 < d < 1$. Differently from case (i), for case (ii) covering $1 \le d < 3/2$, $\mathrm{var}(t_n)$ remains bounded but it may oscillate with n in view of Lemma 5.1, which requires normalization by $\sqrt{\mathrm{var}(t_n)}$.

Proof of Theorem 3.1. (i) Asymptotic results (3.9)-(3.10) follow from Theorem 2.2 and Corollary 2.1 of Abadir *et al.* (2007).

(ii) a) Write

$$\widehat{u}_t = g_{n,t} + \beta_1 + u_t, \tag{5.47}$$

where $h_t = g_{n,t}$ if $-1/2 < d_0 < 1/2$, $h_t = g_{n,t} - g_{n,t-1}$ if $1/2 < d_0 < 3/2$. In Theorem 2.5 and Corollary 2.1 of Abadir *et al.* (2007), it was shown that (3.9)-(3.10) are valid if $u_t \sim \mathrm{I}(d_0)$, $d_0 \in (-1/2, 3/2)$ and $d_0 \ne 1/2$, u_t satisfies Assumption L, and $g_{n,t}$ is such that

$$\sum_{t=1}^{n} |h_{n,t} - h_{n,t+1}| = O_p(n^{-1/2+d_\xi}), \quad (1/2 < d_0 < 3/2). \tag{5.48}$$

To prove (3.9)-(3.10) in case (ii) a), it suffices to note that $\widehat{u}_t$ can be written as (5.47) with $g_{n,t} = (\beta_2 - \widehat{\beta}_2)t$ and, by Theorem 2.1, $\widehat{\beta}_2 - \beta_2 = O_p(n^{-3/2+d_0})$, which implies the validity of (5.48).

(ii) b) Then, (5.47) holds with

$$g_{n,t} = \sum_{j=1}^{3} g_{n,t;j}, \quad g_{n,t;1} = (\beta_2 - \widehat{\beta}_2)t, \quad g_{n,t;2} = (\beta_3 - \widehat{\beta}_3)\sin(\omega t), \quad g_{n,t;3} = (\beta_4 - \widehat{\beta}_4)\cos(\omega t). \tag{5.49}$$

Write $w_h(\lambda_j) := (2\pi n)^{-1/2} \sum_{t=1}^{n} e^{it\lambda_j} h_t$.

To prove the consistency $\widehat{d} \overset{p}{\to} d_0$, it suffices to check (see the proof of Theorem 2.5 (i) in Abadir *et al.* (2007), p. 1371 and Lemma 4.1) that for any $0 < \gamma < 1$

$$m^{-1} \sum_{j=1}^{m} (j/m)^{-\gamma} \lambda_j^{2d_\xi} |w_h(\lambda)|^2 = o_p(1). \tag{5.50}$$

27

To verify (5.50) for (5.49), it suffices to check that

$$\tau_{n,j} := m^{-1} \sum_{j=1}^{m} (j/m)^{-\gamma} \lambda_j^{2d_\xi} |w_{h_j}(\lambda)|^2 = o_p(1), \quad j = 1, 2, 3. \tag{5.51}$$

Asymptotics (3.9)-(3.10) follow (see (4.24)-(4.25) of the proof of Theorem 2.5 (iii) in Abadir *et al.* (2007), p. 1373), if the trend $g_{n,t}$ satisfies (5.50) and

$$m^{-1} \sum_{j=1}^{m} (j/m)^{-\gamma} |\log(j/m) + 1| (\lambda_j^{d_\xi} |w_h(\lambda)| + \lambda_j^{2d_\xi} |w_h(\lambda)|^2) = o_p(m^{-1/2}). \tag{5.52}$$

To prove (5.52), it suffices to show that

$$\tau_{n,j}^* := m^{-1} \sum_{j=1}^{m} |\log(j/m) + 1| (\lambda_j^{d_\xi} |w_{h_j}(\lambda)| + \lambda_j^{2d_\xi} |w_{h_j}(\lambda)|^2) = o_p(m^{-1/2}), \quad j = 1, 2, 3. \tag{5.53}$$

For $j = 1$, (5.51) and (5.53) were shown in Abadir *et al.* (2007), Lemma 4.1 and (4.25).

It remains to prove (5.51) and (5.53) for $j = 2$; for $j = 3$ the proofs following using the same argument. By Theorem 2.1, $\widehat{\beta}_3 - \beta_3 = O_p(n^{-1/2 \wedge (3/2 - d_0)})$, $j = 1, 2$. This implies that

$$|w_{h_1}(\lambda_j)| \leq C n^{-1/2} |\widehat{\beta}_3 - \beta_3| |D_{n,3}(\lambda_j)| = O_p(n^{-(1 \wedge (2 - d_0))})$$

and observing that (5.29) and (5.30) yields

$$|D_{n,3}(\lambda_j)| \leq |D_n(\lambda_j + \omega)| + |D_n(\lambda_j - \omega)| \leq Cn[(1 + n|\lambda_j + \omega|)^{-1} + (1 + n|\lambda_j - \omega|)^{-1}] \leq C$$

because $\lambda_j \leq \lambda_m \leq \omega/2$ as $n \to \infty$.

Hence, if $-1/2 < d \leq 1$, then

$$\begin{aligned}
\tau_{n,2} &= O_p(1) m^{-1} n^{-2} \sum_{j=1}^{m} (j/m)^{-\gamma} \lambda_j^{2d_\xi} \\
&= O_p(1) n^{-2 + 2|d_\xi|} m^{-1} \sum_{j=1}^{m} (j/m)^{-\gamma} = O_p(1) n^{-2 + 2|d_\xi|} = o_p(1), \\
\tau_{n,2}^* &= O_p(1) m^{-1} \sum_{j=1}^{m} (|\log(j/m)| + 1)(\lambda_j^{d_\xi} n^{-1} + \lambda_j^{2d_\xi} n^{-2}) = O_p(n^{|d_\xi| - 1}) = o_p(m^{-1/2}).
\end{aligned}$$

If $1 < d < 3/2$, then $d = 1 + d_\xi$ where $0 \leq d_\xi < 1/2$, and

$$\begin{aligned}
\tau_{n,2} &= O_p(1) m^{-1} n^{-2(2 - d_0)} \sum_{j=1}^{m} (j/m)^{-\gamma} \lambda_j^{2d_\xi} = O_p(1) n^{-2 + 4d_\xi} = o_p(1), \\
\tau_{n,2}^* &= O_p(1) m^{-1} \sum_{j=1}^{m} (|\log(j/m)| + 1)(\lambda_j^{d_\xi} n^{-2 + d_0} + \lambda_j^{2d_\xi} n^{-4 + 2d_0})
\end{aligned}$$

28

$$= O_p(1)m^{d_\xi}n^{-1}m^{-1}\sum_{j=1}^{m}(|\log(j/m)| + 1)(j/m)^{d_\xi} = O_p(1)m^{d_\xi}n^{-1} = o_p(m^{-1/2}),$$

which implies (5.53) for $j = 2$ and completes proof of Theorem 3.1.

References

Abadir, K.M., Distaso, W. and Giraitis, L., 2007, Nonstationarity-extended local Whittle estimation. Journal of Econometrics 141, 1353–1384.

Andrews, D.W.K. and Sun, Y. , 2004, Adaptive local polynomial Whittle estimation of long-range dependence. Econometrica 72, 569–614.

Cremers, H. and Kadelka, D., 1986, On weak convergence of integral functionals of stochastic processes with applications to processes taking paths in L_p^E. Stochastic Processes and their Applications 21, 305–317.

Dahlhaus, R., 1995, Efficient location and regression estimation for long range dependent regression models. Annals of Statistics 23, 1029–1047.

Dalla, V., Giraitis, L. and Hidalgo, J., 2006, Consistent estimation of the memory parameter for nonlinear time series. Journal of Time Series Analysis 27, 211–251.

Davydov, Y.A., 1970, The invariance principle for stationary processes. Theory of Probability and its Applications 15, 487–498.

Deo, R.S. and Hurvich, C.M., 1998, Linear trend with fractionally integrated errors. Journal of Time Series Analysis 19, 379–397.

Giraitis, L. and Surgailis, D., 1985, CLT and other limit theorems for functionals of Gaussian process. Zeitschrift für Wahrscheinlichkeitstheorie und verwandte Gebiete 70, 191–212.

Henry, M and Robinson, P.M., 1996, Bandwidth choice in Gaussian semiparametric estimation of long range dependence. In Athens Conference on Applied Probability and Time Series, vol. 2, edited by P.M. Robinson and M. Rosenblatt. Springer-Verlag, New York.

Ibragimov, I.A. and Linnik, Yu.V., 1977, Independent and Stationary Sequences of Random Variables. Wolters-Noordhoff, Groningen.

Ivanov, A.V. and Leonenko, N.N., 2004, Asymptotic theory of nonlinear regression with long-range dependence. Mathematical Methods of Statistics 13, 153–178.

Ivanov, A.V. and Leonenko, N.N., 2008, Semiparametric analysis of long-range dependence in nonlinear regression. Journal of Statistical Planning and Inference 138, 1733–1753.

Nandi, S. and Kundu, D., 2003, Estimating the fundamental frequency of a periodic function. Statistical Methods and Applications 12, 341–360.

Robinson, P.M., 1995a, Log-periodogram regression of time series with long range dependence. Annals of Statistics 23, 1048–1072.

Robinson, P.M., 1995b, Gaussian semiparametric estimation of long range dependence. Annals of Statistics 23, 1630–1661.

Robinson, P.M., 2005, Efficiency improvements in inference on stationary and nonstationary fractional time series. Annals of Statistics 33, 1800–1842.

Shimotsu, K. and Phillips, P.C.B., 2005, Exact local Whittle estimation of fractional integration. Annals of Statistics 33, 1890–1933.

Surgailis, D. and Viano, M.C., 2002, Long memory properties and covariance structure of the EGARCH model. ESAIM: Probability and Statistics 6, 311–329.

Taqqu, M.S., 1979, Convergence of integrated processes of arbitrary Hermite rank. Zeitschrift für Wahrscheinlichkeitstheorie und verwandte Gebiete 50, 53–83.

Velasco, C., 1999, Gaussian semiparametric estimation of nonstationary time series. Journal of Time Series Analysis 20, 87–127.

Velasco, C., 2003, Nonparametric frequency domain analysis of nonstationary multivariate time series. Journal of Statistical Planning and Inference 116, 209–247.

Yajima, Y., 1988, On estimation of a regression model with long-memory stationary errors. Annals of Statistics 16, 791–807.

Yajima, Y., 1991, Asymptotic properties of the LSE in a regression model with long-memory stationary errors. Annals of Statistics 19, 158–177.

Table 1: Coverage probabilities of 90% confidence intervals for the LS estimator.[a]

	$\rho = -.5$		$\rho = 0$		$\rho = .5$	
	β_1	β_2	β_1	β_2	β_1	β_2
$d = -.4$	0.97	0.96	0.95	0.96	0.88	0.89
$d = -.2$	0.91	0.91	0.86	0.87	0.88	0.87
$d = 0$	0.85	0.85	0.83	0.86	0.88	0.86
$d = .2$	0.85	0.86	0.84	0.85	0.87	0.87
$d = .4$	0.81	0.85	0.80	0.82	0.82	0.87
$d = .6$		0.85		0.83		0.89
$d = .8$		0.86		0.82		0.90
$d = 1$		0.86		0.83		0.90
$d = 1.2$		0.85		0.85		0.90
$d = 1.4$		0.81		0.80		0.79

[a] Notes: This Table reports coverage probabilities of the 90% confidence intervals constructed using the asymptotic theory developed in Section 2. In particular, the data generating process is given by

$$X_t = \beta_1 + \beta_2 t + u_t, \ \ u_t \sim \mathrm{I}(d).$$

β_1, β_2 are estimated using Ordinary Least Squares. The long memory parameter d and the long run variance s_ξ^2 are estimated using the FELW method and formula (3.4), respectively, applied to the residuals $\widehat{u}_t = X_t - \widehat{X}_t$.

Table 2: Coverage probabilities of 90% confidence intervals for the LS estimator.[a]

	$\rho = -.5$				$\rho = 0$				$\rho = .5$			
	β_1	β_2	β_3	β_4	β_1	β_2	β_3	β_4	β_1	β_2	β_3	β_4
$d = -.4$	0.95	0.95	0.84	0.84	0.95	0.95	0.86	0.88	0.88	0.88	0.92	0.91
$d = -.2$	0.89	0.91	0.84	0.85	0.86	0.86	0.87	0.88	0.86	0.86	0.93	0.92
$d = 0$	0.84	0.87	0.86	0.87	0.86	0.87	0.91	0.90	0.86	0.84	0.94	0.94
$d = .2$	0.84	0.84	0.88	0.87	0.82	0.84	0.91	0.92	0.89	0.86	0.93	0.93
$d = .4$	0.79	0.86	0.91	0.90	0.79	0.83	0.93	0.93	0.81	0.87	0.94	0.94
$d = .6$		0.86	0.85	0.86		0.86	0.88	0.87		0.86	0.91	0.91
$d = .8$		0.87	0.84	0.83		0.85	0.87	0.87		0.89	0.91	0.89
$d = 1$		0.87	0.85	0.88		0.85	0.89	0.88		0.90	0.92	0.92
$d = 1.2$		0.86	0.90	0.90		0.85	0.91	0.92		0.90	0.92	0.92
$d = 1.4$		0.82	0.93	0.93		0.82	0.92	0.92		0.80	0.82	0.82

[a] Notes: This Table reports coverage probabilities of the 90% confidence intervals constructed using the asymptotic theory developed in Section 2. In particular, the data generating process is given by

$$X_t = \beta_1 + \beta_2 t + \beta_3 \cos(\omega t) + \beta_4 \sin(\omega t) + u_t, \ \ \omega = \pi/2, \ \ u_t \sim \mathrm{I}(d).$$

$\beta_1, \beta_2, \beta_3, \beta_4$ are estimated using Ordinary Least Squares. The long memory parameter d and the long run variance s_ξ^2 are estimated using the FELW method and formula (3.4), respectively, applied to the residuals $\widehat{u}_t = X_t - \widetilde{X}_t$. The estimator of the spectral density $f_\xi(\omega)$ is given by formula (4.1).

Table 3: Root MSEs of LS estimators.[a]

	$\widetilde{\beta}_1$	$\widetilde{\beta}_2$	$\widehat{\beta}_1$	$\widehat{\beta}_2$
$d = -.4$	0.018138	6.70865e-5	0.018173	6.72490e-5
$d = -.2$	0.036291	0.000130	0.036311	0.000130
$d = 0$	0.089922	0.000309	0.089924	0.000309
$d = .2$	0.244466	0.000761	0.244473	0.000761
$d = .4$	0.901005	0.002061	0.901016	0.002061
$d = .6$		0.005568		0.005568
$d = .8$		0.016482		0.016482
$d = 1$		0.049259		0.049259
$d = 1.2$		0.169340		0.169340
$d = 1.4$		0.754884		0.754873

[a] Notes: This Table reports root mean squared errors of the LS estimators of deterministic components. In particular, the data generating process is given by

$$X_t = \beta_1 + \beta_2 t + \beta_3 \cos(\omega t) + \beta_4 \sin(\omega t) + u_t, \ \ \omega = \pi/2, \ \ u_t \sim \mathrm{I}(d).$$

$\widetilde{\beta}_1, \widetilde{\beta}_2$ are the LS estimators when the model is correctly specified, while $\widehat{\beta}_1, \widehat{\beta}_2$ denote the LS estimator for the misspecified model.

Table 4: Bias and Root MSEs of the estimator $\widehat{d}_{\widehat{u}}$ for different values of ρ.

	Bias	RMSE	Bias	RMSE	Bias	RMSE
	$\rho = -.5$		$\rho = 0$		$\rho = .5$	
$d = -0.4$	-0.004406	0.082613	0.005228	0.080757	0.101875	0.129679
$d = -0.2$	-0.026378	0.085891	-0.010943	0.080904	0.087739	0.119000
$d = 0$	-0.036340	0.087974	-0.019628	0.080196	0.082967	0.116947
$d = 0.2$	-0.041281	0.090194	-0.028645	0.089280	0.080431	0.113747
$d = 0.4$	-0.043690	0.090695	-0.028162	0.090983	0.084435	0.120408
$d = 0.6$	-0.030618	0.091283	-0.016453	0.089691	0.091507	0.125162
$d = 0.8$	-0.027204	0.089273	-0.012923	0.085367	0.096704	0.125939
$d = 1.0$	-0.027851	0.080885	-0.011396	0.077266	0.088296	0.120020
$d = 1.2$	-0.023571	0.082448	-0.009434	0.081665	0.090578	0.119118
$d = 1.4$	-0.022104	0.082663	-0.005929	0.077420	0.096853	0.123975

[a] Notes: This Table reports bias and Root MSE of the FELW estimator of d. In particular, the data generating process is given by

$$X_t = \beta_1 + \beta_2 t + \beta_3 \cos(\omega t) + \beta_4 \sin(\omega t) + u_t, \ \ \omega = \pi/2, \ \ u_t \sim \mathrm{I}(d).$$

The long memory parameter d is estimated using the FELW method applied to the residuals $\widehat{u}_t = X_t - \widetilde{X}_t$.