BAND SPECTRUM REGRESSION FOR COINTEGRATED TIME SERIES WITH LONG MEMORY INNOVATIONS*

by

D Marinucci
London School of Economics and Political Science

Contents:

Discussion Paper
No.EM/98/353
July 1998

The Suntory Centre
Suntory and Toyota International Centres
for Economics and Related Disciplines
London School of Economics and Political Science
Houghton Street
London WC2A 2AE
Tel.: 0171-405 7686

---

* I am grateful to P M Robinson for many useful comments on an earlier version. Usual disclaimers apply.

Abstract

Band spectrum regression is considered for cointegrated time series with long memory innovations. The estimates we advocate are shown to be consistent when cointegrating relationships among stationary variables are investigated, while OLS are inconsistent due to correlation between the regressor and the cointegrating residuals; in the presence of unit roots, these estimates share the same asymptotic distribution as OLS. As a corollary of the main result, we provide a functional central limit theorem for quadratic forms in nonstationary fractionally integrated processes.

Keywords: Long-range dependence; band spectrum regression; cointegration.

JEL No.: C32

1. INTRODUCTION

Let $\{a_t\}$ be a covariance stationary sequence of random variables such that $Ea_1 = 0$, $Ea_t^2 = \gamma(0) < \infty$, $Ea_0 a_\tau = \gamma(\tau)$; introduce the spectral density function satisfying

$$\gamma(\tau) = \int_{-\pi}^{\pi} f(\lambda) \exp(i\lambda\tau) d\lambda \ , \ \tau = 0, \pm 1, \pm 2, \ldots \ .$$

We term the sequence $\{a_t\}$ short range dependent or $I(0)$ if

$$0 < f(0) < \infty \ ; \qquad (1)$$

otherwise we term $\{a_t\}$ long range dependent, or fractionally integrated of order d_a, $-\frac{1}{2} < d_a < \frac{1}{2}$ if

$$f(\lambda) \sim G\lambda^{-2d_a} \text{ as } \lambda \to 0^+, \ 0 < G < \infty \ , \qquad (2)$$

where " $\sim$ " signifies that the ratio of the left- and right-hand side tends to one. More precisely, the definition of long memory is adopted for the case $d_a > 0$, while for negative d_a the process is termed antipersistent. Models with long range dependent errors have been considered in various fields of applications, e.g. geophysics, hydrology and economics; statistical inference under long range dependence has been largely investigated in recent years, for instance by Deo (1997), Giraitis and Taqqu (1998a), Hurvich, Deo and Brodsky (1998), and Robinson and Hidalgo (1997).

For many applications, it is also of interest to focus on autoregressive processes which combine long memory innovations and unit roots, as in Chan and Terrin (1995), Hurvich and Ray (1995) and Velasco (1997). Consider the two-dimensional observations (y_t, x_t), $t = 1, 2, \ldots.$, where

$$\begin{cases} y_t = \beta x_t + e_t \\ x_t = \phi x_{t-1} + u_t \end{cases} , \ u_t \sim I(d_u) \ , \ e_t \sim I(d_e) \ , \ |\phi| \leq 1 \ , \qquad (3)$$

with $0 \leq d_e, d_u < \frac{1}{2}$. We cover dependence between x_t and e_t allowing correlation between u_t and e_t, while we allow for nonstationarity in x_t, y_t by including the possibility that ϕ equals unity, i.e. x_t is a partial sum of long memory innovations (for $|\phi| < 1$ and x_t independent from e_t efficient estimates of β are provided by Robinson and Hidalgo (1997)). When the innovations u_t, e_t are short range dependent and $\phi = 1$, y_t and x_t are integrated of order 1 (written $I(1)$) and the bivariate vector sequence (y_t, x_t) is cointegrated of order $(1, 0)$ (written $CI(1, 0)$) in the sense of Engle and Granger (1987), the cointegrating vector being $(1, -\beta)$. The asymptotic theory for $I(0)$ and $I(1)$ processes has been thoroughly analyzed in the econometric and probabilistic literature, functional central limit theorems for normalized functionals of such processes have been established (e.g. Phillips (1988), Hansen (1992)), providing the basis for the asymptotic statistical theory of the many estimates proposed for cointegrating parameters in the $CI(1, 0)$ case, see for instance Watson (1994) for a review. Much less is known, on the other hand, on statistical inference for cointegrated variables when the short range dependence condition (1) is relaxed.

When the sequences e_t, u_t are not short range dependent, we define (y_t, x_t) a cointegrated vector with long memory (or fractionally integrated) innovations (cf. Jeganathan (1996)), if either *a)* $\phi = 1$ or *b)* $d_u > d_e$, or both. Indeed for $|\phi| < 1$ it can be verified easily that x_t and u_t share the same order of fractional integration, $d_x = d_u$, and hence under both *a)* and *b)* e_t has less "memory" than (y_t, x_t), so that (3) characterizes a long-run equilibrium relationship which can be viewed as a generalization of the $CI(1,0)$ case. An alternative approach for the generalization of cointegration analysis to fractional processes is investigated by Robinson and Marinucci (1998), where a different definition of fractional integration is introduced for the nonstationary case, such that the sequences of first differences $\Delta x_t = x_t - x_{t-1}$, $\Delta y_t = y_t - y_{t-1}$ are not second order stationary; cointegration under fractional circumstances is considered also by Dolado and Marmol (1996) and others.

The purpose of this paper is to analyze the behaviour of a frequency-domain semiparametric estimate of the cointegrating parameter β. More precisely, for zero-mean sequences of scalars $\{a_t\}$, $\{b_t\}$ introduce the discrete Fourier transforms

$$w_a(\lambda) = \frac{1}{\sqrt{2\pi n}} \sum_{t=1}^{n} a_t \exp(-i\lambda t) \ , \ w_b(\lambda) = \frac{1}{\sqrt{2\pi n}} \sum_{t=1}^{n} b_t \exp(-i\lambda t) \ ,$$

and the (cross-) periodogram $I_{ab}(\lambda) = w_a(\lambda)\overline{w_b(\lambda)}$, the bar denoting complex conjugation. For $a_t = y_t$, $b_t = x_t$ and $-\pi < \omega < \pi$ we consider the statistic

$$\widetilde{\beta}_M(\omega) = Re\left\{\arg\min_{\beta} \int_{-\pi}^{\pi} K_M(\lambda)|w_y(\omega-\lambda) - \beta w_x(\omega-\lambda)|^2 d\lambda\right\} \ ,$$

where for $M = 1, 2, ...,$ $K_M(\lambda)$ represents a frequency-domain kernel such that $K_M(-\lambda) = K_M(\lambda)$, $-\pi \leq \lambda \leq \pi$, with $M = 1, 2, ...$ a bandwidth parameter such that

$$M < n \ , \ \frac{1}{M} + \frac{M}{n} \to 0 \text{ as } n \to \infty \ . \tag{4}$$

The statistic $\widetilde{\beta}_M(\omega)$ relates to works by Hannan (1963) and others and can be given the closed form expression

$$\widetilde{\beta}_M(\omega) = \frac{\int_{-\pi}^{\pi} K_M(\lambda) I_{xy}(\omega-\lambda) d\lambda}{\int_{-\pi}^{\pi} K_M(\lambda) I_{xx}(\omega-\lambda) d\lambda} \ . \tag{5}$$

$\widetilde{\beta}_M(\omega)$ can be interpreted as resulting from a form of continuously averaged least squares regression of $w_y(.)$ on $w_x(.)$ around frequency ω, a technique known as "band spectrum regression" (Hannan and Robinson (1973), Engle (1974)); a discretely averaged version of $\widetilde{\beta}_M(\omega)$ is considered by Robinson (1994a), Robinson and Marinucci (1998). For short range dependent processes a well-known estimate of the spectral density matrix at frequency ω is given by

$$\widetilde{f}(\omega) = \int_{-\pi}^{\pi} K_M(\lambda) I(\omega-\lambda) d\lambda \ , \tag{6}$$

whence we can rewrite $\widetilde{\beta}_M(\omega)$ (formally) as $\widetilde{\beta}_M(\omega) = \widetilde{f}_{xy}(\omega)/\widetilde{f}_{xx}(\omega)$. Although it is also possible to focus on cointegration at seasonal frequencies, in the sequel we shall concentrate on the case $\omega = 0$ and write for brevity $\widetilde{\beta}_M(0) = \widetilde{\beta}_M$. Write $f_{ab}(\lambda) = (2\pi)^{-1}\sum_{\tau=-\infty}^{\infty}\gamma_{ab}(\tau)e^{-i\lambda\tau}$, $\gamma_{ab}(\tau) = Ea_0b_\tau$, for the cross-spectral density of the covariance stationary, zero-mean sequences a_t, b_t. We have

$$\widetilde{f}_{ab}(0) = \int_{-\pi}^{\pi} K_M(\lambda)I_{ab}(\lambda)d\lambda \ , \ a, b = x, y \ , \tag{7}$$

and it can be verified (Brockwell and Davis (1991), p.358-360) that (7) is equivalent to

$$\widetilde{f}_{ab}(0) = \frac{1}{2\pi}\sum_{\tau=-n+1}^{n-1} k_M(\tau)c_{ab}(\tau) \ , \tag{8}$$

where $k_M(.)$ is the lag window defined by $k_M(\tau) = \int_{-\pi}^{\pi} K_M(\lambda)e^{i\tau\lambda}d\lambda$, and

$$c_{ab}(\tau) = \begin{cases} \sum_{t=1}^{n-\tau} a_t b_{t+\tau} \ , & \tau \geq 0 \\ \sum_{t=|\tau|+1}^{n} a_t b_{t-\tau} \ , & \tau < 0 \end{cases} .$$

Therefore we can rewrite for (5)

$$\widetilde{\beta}_M = \frac{\sum_{\tau=-n+1}^{n-1} k_M(\tau)c_{xy}(\tau)}{\sum_{\tau=-n+1}^{n-1} k_M(\tau)c_{xx}(\tau)} \ , \ 1 \leq M \leq n-1 \ ,$$

and adopt for $\widetilde{\beta}_M$ the natural definition of Weighted Covariance Estimate (WCE). For short range dependent (u_t, e_t) and $\phi = 1$, $\widetilde{\beta}_M$ was previously considered by Phillips (1991), where the use of other spectral regression procedures, more efficient than $\widetilde{\beta}_M$ when the $CI(1,0)$ assumption is correct, is advocated; the following sections analyzes the behaviour of $\widetilde{\beta}_M$ in the fractional circumstances considered in this paper.

Consider now the OLS estimates $\widehat{\beta} = c_{xx}(0)^{-1}c_{xy}(0)$. For $|\phi| < 1$ and assuming ergodicity conditions hold on x_t and e_t, we have the convergence

$$\widehat{\beta} - \beta = (Ex_1^2)^{-1}Ex_1e_1 \text{ as } n \to \infty \ ,$$

and hence OLS (and indeed other procedures for cointegration analysis) are inconsistent in the presence of non-zero correlation between x_t and e_t. It is likely to be extremely difficult in practice to distinguish, on the basis of a finite sample of observations, between a unit root process and a stationary autoregression with long memory innovations and roots close to the unit circle; it is therefore remarkable that in Section 2 we are able to prove consistency for $\widetilde{\beta}_M$ under stationary circumstances (cf. Robinson (1994a)). In Section 3, we go on to characterize the limit distribution of $\widehat{\beta}$ and $\widetilde{\beta}_M$ when ϕ equals unity and u_t, e_t are stationary long memory processes; companion to this derivation is a functional central limit theorem for a class of quadratic forms in nonstationary variables, a result which may have some independent interest and can be extended to more general quadratic forms; most proofs are collected in the Appendix.

Throughout this paper, we restrict our attention to the bivariate case for simplicity; multivariate generalizations require in the nonstationary case extensions of functional central limit theorems from Gorodetskii (1977) and Chan and Terrin (1995), and these extensions are currently under investigation. In the sequel, C denotes a generic, positive constant, which need not be the same all the time it is used.

2. THE STATIONARY CASE

When ϕ is in absolute value smaller than unity, we find it notationally convenient to specify a model for the covariance stationary sequence (x_t, e_t) rather than for (u_t, e_t), and to write d_x for d_u.

Assumption A (3) holds, with $|\phi| < 1$ and

$$(x_t, e_t)' = \Psi(L)\varepsilon_t \ , \ \Psi(L) = \sum_{k=0}^{\infty} \Psi_k L^k \ ,$$

where L is the lag operator, $\Psi_0 = I_p$, and for $k = 1, 2, ...$ Ψ_k has (i, j)-th element

$$\psi_{1jk} \sim c_{1j} k^{d_x - 1} \ , \ \psi_{2jk} \sim c_{2j} k^{d_e - 1} \ , \text{ as } k \to \infty \ , \ 0 < c_{ij} < \infty \ , \tag{9}$$

for $i, j = 1, 2$, $0 \leq d_e < d_x < \frac{1}{2}$, and where $\varepsilon_t = (\varepsilon_{1t}, \varepsilon_{2t})'$ represents a zero-mean, independent and identically distributed (*i.i.d.*) sequence that satisfies $E\|\varepsilon_t\|^4 < \infty$, $\|.\|$ denoting Euclidean norm.

Assumption A characterizes the bivariate sequence (x_t, e_t) as a linear stationary long memory process with innovations satisfying a mild integrability condition (cf. Davydov (1970)), and it is for instance verified if x_t and e_t are generated by stationary autoregressive fractionally integrated moving averages processes (Granger and Joyeux (1980)) driven by innovations with finite fourth moments; such processes satisfy (2). As a consequence of (9), as $\tau \to \infty$,

$$Ex_t x_{t+\tau} \sim g_x \tau^{2d_x - 1}, \ 0 < g_x < \infty, \ Ex_t e_{t+\tau} \sim g_{xe} \tau^{d_e + d_x - 1} \ , \tag{10}$$

where $g_{xe} \equiv 0$ if $E\varepsilon_{1t}\varepsilon_{2t} = 0$. In the sequel, we find it convenient to set $k_M(.) \stackrel{def}{=} k(\tau/M)$, and to introduce

Assumption B The kernel $k(.)$ is a real-valued, Lebesgue-measurable function that for $v \in R$ satisfies

$$\int_{-1}^{1} k(v)dv = 1 \ , \ 0 \leq k(v) \leq C \ , \ k(v) = 0 \text{ for } |v| > 1 \ . \tag{11}$$

Assumption B is common for spectral estimates, and it is satisfied by (normalized versions of) truncated lag windows such as the Bartlett, modified Bartlett, Parzen, and many others; see Brillinger (1981) for a review.

Lemma 1 Under Assumptions A and B, as $M \to \infty$ for $M = o(n^2)$ we have

$$\begin{aligned}\left\{\sum_{\tau=-M}^{M} k(\frac{\tau}{M})\gamma_{xx}(\tau)\right\}^{-1} \sum_{\tau=-M}^{M} k(\frac{\tau}{M})c_{xx}(\tau) &= 1 + o_p(1) , \\ \left\{\sum_{\tau=-M}^{M} k(\frac{\tau}{M})\gamma_{xe}(\tau)\right\}^{-1} \sum_{\tau=-M}^{M} k(\frac{\tau}{M})c_{xe}(\tau) &= 1 + o_p(1) .\end{aligned}$$

For $d_x > 0$, the spectral density of x_t has a singularity at frequency zero and cannot be estimated there. For $\lambda_j = 2\pi j/n,\ j = 1, 2, ...m$, denote by $\widehat{F}_{ab}(\lambda_m)$ the real part of the discretely averaged periodogram, i.e. $\widehat{F}_{ab}(\lambda_m) = (2\pi/n)\sum_{j=1}^{m} I_{ab}(\lambda_j)$. Assuming

$$m < n\ ,\ \frac{1}{m} + \frac{m}{n} \to 0 \text{ as } n \to \infty\ ,$$

and regularity conditions on (x_t, e_t) (such that Assumption A is covered), it was shown by Robinson (1994a) and Lobato (1997) that for $a, b = x, e$,

$$p\lim \frac{\widehat{F}_{ab}(\lambda_m)}{F_{ab}(\lambda_m)} = 1\ ,\ \text{as } n \to \infty\ ,\ F_{ab}(\lambda_m) = \int_0^{\lambda_m} f_{ab}(\lambda)d\lambda\ . \tag{12}$$

Lemma 1 is similar to (12), relating however to the case when the continuously averaged periodogram (6) is considered.

Under Assumption A and in view of (10), (11), as $M \to \infty$ we have, by the dominated convergence theorem

$$M^{-2d_x} \sum_{\tau=-M}^{M} k(\frac{\tau}{M})\gamma_{xx}(\tau) = \sum_{\tau=-M}^{M} k(\frac{\tau}{M})\frac{\gamma_{xx}(\tau)}{M^{2d_x-1}}\frac{1}{M} \sim B_{xx}\ , \tag{13}$$

$$M^{-d_x-d_e} \sum_{\tau=-M}^{M} k_M(\tau)\gamma_{xe}(\tau) = \sum_{\tau=-M}^{M} k(\frac{\tau}{M})\frac{\gamma_{xe}(\tau)}{M^{d_x+d_e-1}}\frac{1}{M} \sim B_{xe}\ , \tag{14}$$

where

$$B_{xx} = g_x \int_{-1}^{1} k(v)v^{2d_x-1}dv\ ,\ B_{xe} = g_{xe}\int_{-1}^{1} k(v)v^{d_x+d_e-1}dv\ .$$

Hence B_{xe} can be equal to zero if $E\varepsilon_{1t}\varepsilon_{2t}$ is, in which case the left-hand side of (14) is $o_p(1)$.

As an application of Lemma 1 we consider the statistic

$$\ln|\sum_{\tau=-M}^{M} k(\frac{\tau}{M})c_{xx}(\tau)| - 2d_x \ln M = \ln B_{xx} + o_p(1) \text{ as } n \to \infty\ ,$$

whence a consistent estimate of the parameter d_x can be obtained under Assumptions A, B and (4) by

$$\widehat{d}_x = \frac{\ln |\sum_{\tau=-M}^{M} k_M(\tau) c_{xx}(\tau)|}{2 \ln M} .$$

This estimate is likely to be severely biased in finite samples, though, and rather than investigating in more detail its properties we concentrate on (3), for which we introduce the following result.

Theorem 1 Under (3), Assumptions A, B and $M^2 = o(n)$, as $M \to \infty$

$$M^{d_x - d_e}(\widetilde{\beta}_M - \beta) = \frac{B_{xe}}{B_{xx}} + o_p(1) .$$

Theorem 1 suggests that the presence of correlation between x_t and e_t does not prevent consistency of $\widetilde{\beta}_M$, the rate of convergence being determined by the "strength" of the cointegrating relationship $d_x - d_e$. We delay to future research the investigation of issues such as the determination of optimal bandwidth parameter M (cf. Robinson (1994b)), the choice of an optimal kernel $k(.)$, the estimation of the noncentrality parameter B_{xe}/B_{xx}, the implementation of bias reduction techniques, and the derivation of the asymptotic distribution for the adjusted estimate; we focus instead on the unit root case, which is dealt with in the next section.

3. THE UNIT ROOT CASE

The unit root case is characterized by the identification $\phi = 1$ in (3), so to obtain (after the initialization $x_0 = 0$)

$$x_t = \sum_{s=1}^{t} u_s \ , \ t = 1, 2, \ldots . \tag{15}$$

We consider first the $CI(1,0)$ case. For convenience, we write $\omega_{ab} = 2\pi f_{ab}(0)$, $a, b = u, e$, with $|\omega_{ab}| < \infty$ by (1). Let $\Rightarrow$ denote weak convergence in the sense of Billingsley (1968) and $B(r, \omega)$ denote scaled Brownian motion, i.e. the Gaussian zero-mean process with independent increments and $EB^2(r, \omega) = \omega r$, $\omega > 0$. The following result is proved by Phillips (1991).

Lemma 2 *(Phillips (1991))* Let (3) hold for $\phi = 1$, and assume that as $n \to \infty$

$$\frac{1}{n^2} \sum_{t=1}^{n} (\sum_{s=1}^{t} u_s)^2 \ \Rightarrow \ \int_0^1 B^2(r; \omega_{uu}) dr \ , \tag{16}$$

$$\frac{1}{n} \sum_{t=1}^{n} (\sum_{s=1}^{t} u_s) e_t \ \Rightarrow \ \int_0^1 B(r; \omega_{uu}) dB(r; \omega_{ee}) + \sum_{\tau=0}^{\infty} \gamma_{ue}(\tau) . \tag{17}$$

Assume also that Assumption B holds. Then under (4), as $n \to \infty$

$$n(\widetilde{\beta}_M - \beta) \Rightarrow \left\{ \int_0^1 B^2(r;\omega_{uu})dr \right\}^{-1} \left\{ \int_0^1 B(r;\omega_{uu})dB(r;\omega_{ee}) + \sum_{\tau=0}^{\infty} \gamma_{ue}(\tau) \right\} . \tag{18}$$

Under (1), conditions for (16)/(17) to hold are given for instance by Phillips (1988) and Hansen (1992). On the other hand, when (2) holds, u_t and e_t are not short range dependent and the asymptotics for $\widetilde{\beta}_M$ depends on functional central limit theorems for normalized partial sums of long memory innovations. Such results have now been given under a variety of different conditions, for instance by Davydov (1970), Gorodetskii (1977), and more recently by Chan and Terrin (1995). For our purposes, we introduce the following

Assumption C (15) holds, where for $-\frac{1}{2} < d_u < \frac{1}{2}$

$$\begin{aligned} u_t &= \sum_{j=0}^{\infty} \psi_k \varepsilon_{t-k} \ , \ \sum_{k=0}^{\infty} |\psi_k|^2 < \infty \ , \ \psi_k \sim ck^{d_u - 2} \text{ as } k \to \infty \ , \ 0 < c < \infty \ , \\ \varepsilon_t &\equiv i.i.d.(0, \sigma_\varepsilon^2) \ , \ \sigma_\varepsilon^2 < \infty \ , \ E|\varepsilon_t|^\delta < \infty \ , \ \delta > \frac{1}{2d_u + 1} \ . \end{aligned}$$

We have allowed here for the possibility that u_t is antipersistent, i.e. $d_u < 0$ (the condition on δ is clearly redundant if $d_u > -1/4$). Some of the results of this section need somewhat stronger assumptions than C, and therefore we introduce also

Assumption D (3) holds with $\phi = 1$, and for covariance stationary sequences a_t, $a = u, e$, we have

$$\begin{aligned} a_t &= \int_{-\pi}^{\pi} \exp(it\lambda) f_{aa}(\lambda)^{1/2} dM_a(\lambda) \ , \\ f_{aa}(\lambda) &\sim |\lambda|^{-2d_a} G_a \ , \ \text{as } \lambda \to 0^+, \ 0 < G_a < \infty \ , \end{aligned}$$

where $M_u(.)$, $M_e(.)$ are complex-valued, Gaussian random measures which satisfy

$$\begin{aligned} dM_a(\lambda) &= \overline{dM_a(-\lambda)} \\ EdM_a(\lambda) &= 0 \\ EdM_a(\lambda)\overline{dM_b(\lambda)} &= \begin{cases} 0, & \lambda \neq \mu \\ d\lambda, & \lambda = \mu \end{cases} \ , \ a, b = u, e \ . \end{aligned}$$

Because by Wold representation theorem any Gaussian covariance stationary sequence can be viewed as a linear process with $i.i.d.$ innovations, Assumption D entails stricter conditions on u_t than Assumption C. In the sequel, for notational convenience we shall occasionally use the identification $d_x = d_u + 1$; although x_t is not covariance stationary, condition (2) can be granted a broader interpretation in this case, see Hurvich and Ray (1995).

Lemma 3 Let $u_t = \sum_{j=0}^{\infty} \alpha_j \xi_{t-j}$, for $t = 1, 2, ...$, where

$$\sum_{j=0}^{\infty} |\alpha_j|^2 < \infty \ , \ \sum_{j=0}^{\infty} \alpha_j \alpha_{j+\tau} \sim c\tau^{2d_u - 1}, \ 0 < c < \infty, \ \text{as } \tau \to \infty \ , \tag{19}$$

$$E\xi_t = 0 \ , \ E\xi_t^2 < C \ , \ E\xi_t \xi_s = 0 \ , \ t \neq s \ . \tag{20}$$

Under (4), (15) and for $k(.)$ such that (11) holds, we have

$$c_{xx}(0) - \frac{1}{M} \sum_{\tau=-M}^{M} k(\frac{\tau}{M}) c_{xx}(\tau) = o_p(n^{2d_x - 1}) \ .$$

Because ξ_t need not be independent, or identically distributed, or satisfy any moment condition of order greater than two, (19)/(20) are weaker than Assumption C.

For the following result we need to narrow the focus and impose Assumption D.

Lemma 4 Under Assumption D, $d_u + d_e > 0$, (4) and (11)

$$c_{xe}(0) - \frac{1}{M} \sum_{\tau=-M}^{M} k(\frac{\tau}{M}) c_{xe}(\tau) = o_p(n^{d_x + d_e - 1}) \ .$$

Let us introduce fractional Brownian motion, which we present as the Gaussian a.s. continuous process with harmonizable representation (Samorodnitsky and Taqqu (1994))

$$B(r; d_a) = \int_R \frac{\exp(i\lambda r) - 1}{i\lambda} f_{aa}(\lambda)^{1/2} dM_a(\lambda) \ .$$

Here we consider also the compound processes, for $d_a + d_b > 0$, $a, b = u, e$

$$P(d_a) = \int_0^1 B^2(r; d_a) dr \ ,$$

$$Q(d_a, d_b) = \int_0^1 B(r; d_a) dB(r; d_b) + C(d_a, d_b) \ , \tag{21}$$

where

$$C(d_a, d_b) = \int_R \left\{ \int_0^s \frac{1 - \exp(-it\mu)}{i\mu} dt \right\} f_{ab}(\mu) d\mu \ .$$

The stochastic integral on the right-hand side of (21) is defined only in a formal sense to be equal to

$$\int_{R^2}'' \left\{ \int_0^s \exp(it\lambda) \frac{\exp(it\mu) - 1}{i\mu} \right\} f_{uu}(\mu)^{1/2} f_{ee}(\lambda)^{1/2} dM_u(\mu) dM_e(\lambda) \ , \tag{22}$$

where $\int''_{R^2}$ signifies that the integral excludes the diagonals $\mu = \pm\lambda$. (22) is a multiple Wiener-Ito stochastic integral in the sense of Major (1981), but it cannot be defined as an Ito integral with respect to $B(r; d_e)$ because fractional Brownian motion is not a semimartingale.

Lemma 5 *(Gorodetskii (1977)), Chan and Terrin (1995))* As $n \to \infty$, under Assumptions C and D

$$\left(n^{2d_x} G_u\right)^{-1} \sum_{t=1}^{n} x_t^2 \Rightarrow P(d_u) \ , \tag{23}$$

Also, under Assumption D, $d_x + d_e > 1$

$$\left(n^{d_x+d_e} G_u^{1/2} G_e^{1/2}\right)^{-1} \sum_{t=1}^{n} x_t e_t \Rightarrow Q(d_u, d_e) \ . \tag{24}$$

Proof (23) follows under Assumption C from Gorodetskii (1977) and the continuous mapping theorem; (24) is given in Chan and Terrin (1995).

It follows from Lemma 2 that in the $CI(1,0)$ case $\widetilde{\beta}_M$ shares the same asymptotic distribution as OLS; when the innovation are long memory a stronger result holds, namely the difference between the two estimates is asymptotically $o_p(n^{d_e-d_x})$. More precisely,

Theorem 2 Under Assumption D, (3), (4) and (11), as $n \to \infty$

$$n^{d_x-d_e}(\widehat{\beta} - \beta) = \frac{G_e^{1/2}}{G_u^{1/2}} P(d_u)^{-1} Q(d_u, d_e) \ , \tag{25}$$

and

$$|\widetilde{\beta}_M - \widehat{\beta}| = o_p(n^{d_e-d_x}) \ , \tag{26}$$

$$n^{d_x-d_e}(\widetilde{\beta}_M - \beta) = \frac{G_e^{1/2}}{G_u^{1/2}} P(d_u)^{-1} Q(d_u, d_e) \ . \tag{27}$$

Proof (25) follows from Lemma 5 and the continuous mapping theorem. For (26) we can rewrite (cf. Robinson and Marinucci (1998))

$$\begin{aligned}
\widetilde{\beta}_M - \widehat{\beta} &= \left\{\frac{1}{M}\sum_{\tau=-M}^{M} k_M(\tau) c_{xx}(\tau)\right\}^{-1} \left\{\frac{1}{M}\sum_{\tau=-M}^{M} k_M(\tau) c_{xe}(\tau) - c_{xe}(0)\right\} + \\
&\quad \left\{\frac{1}{M}\sum_{\tau=-M}^{M} k_M(\tau) c_{xx}(\tau)\right\}^{-1} \left\{c_{xx}(0) - \frac{1}{M}\sum_{\tau=-M}^{M} k_M(\tau) c_{xx}(\tau)\right\} (c_{xx}(0))^{-1} c_{xe}(0) \\
&= O_p(n^{1-2d_x}) o_p(n^{d_x+d_e-1}) + O_p(n^{1-2d_x}) o_p(n^{2d_x-1}) O_p(n^{d_e-d_x}) \\
&= o_p(n^{d_e-d_x})
\end{aligned}$$

in view of Lemmas 3 and 4, so that the proof of (26) is completed. (27) follows immediately.

The constant $C(d_u, d_e)$ at the numerator in (25) is due to the non-zero correlation between u_t and e_t. The left-hand side of (25) generalizes in an intuitive way the rate of convergence and the asymptotic distribution of the $CI(1,0)$ case, which is provided by (18).

Theorem 2 might be extended to allow for $d_u < 0$, provided $d_x + d_e > 1$, i.e. $d_u + d_e > 0$. However we refrain from the analysis of this case here, both for the sake of brevity and to maintain symmetry with the stationary case where $d_u < 0$ was ruled out. The possibility of an "antipersistent" behaviour in the innovation sequence u_t seems moreover less relevant for applications. Here as in Section 2 we leave several issues for future research; in particular from the point of view of practitioners it seems important to analyze the case of deterministic components in (y_t, x_t), including a non-zero mean. Also, it seems possible to make asymptotic statistical inference on $\widetilde{\beta}_M$ viable by nonparametric estimation of G_a and $f_{ab}(\lambda)$, $a, b = u, e$, possibly by a two-step procedure, and then by tabulation of the left-hand side of (27). Although under (4) the asymptotic distribution of $\widetilde{\beta}_M$ does not depend on the bandwidth parameter M, some guidance must be provided for applied research, cf. Robinson and Marinucci (1998).

As a final point, recall that from Lemmas 3-5 we learn, under (4),

$$\frac{n^{1-2d_x}}{M} \sum_{\tau=-M}^{M} k(\frac{\tau}{M}) c_{xx}(\tau) \Rightarrow G_u P(d_u) + o_p(1) \tag{28}$$

and

$$\frac{n^{1-d_x-d_e}}{M} \sum_{\tau=-M}^{M} k(\frac{\tau}{M}) c_{xe}(\tau) \Rightarrow G_u^{1/2} G_e^{1/2} Q(d_u, d_e) + o_p(1) \ . \tag{29}$$

In view of (8), (28)/(29) provide the asymptotic distribution of the weighted covariance estimate of the (cross-) spectral density at zero frequency for the variables x_t and e_t when the former is nonstationary. This result can have some independent applications, for instance for estimates of the differencing parameter d_x under the same circumstances as in Hurvich and Ray (1995), and Velasco (1997). Moreover, the same argument as in Lemmas 3 and 4 can be exploited in the analysis of the behaviour of more general quadratic forms in nonstationary variables (for the stationary case, cf. for instance Giraitis and Taqqu (1998b) and the references mentioned therein). Consider the quadratic form $\sum_{t=1}^{n} \sum_{s=1}^{n} b_{M,n}(t-s) x_t x_s$, where $b_{M,n}(\tau) = k^*(\tau/M)$ (say), for $k^*(.)$ such that Assumption B holds and x_t satisfying Assumption D; hence

$$\begin{aligned} \frac{1}{n} \sum_{t=1}^{n} \sum_{s=1}^{n} b_{M,n}(t-s) x_t x_s &= \sum_{\tau=-n+1}^{n-1} b_{M,n}(\tau) \frac{1}{n} \sum_{t=1}^{n-|\tau|} x_t x_{t+|\tau|} \\ &= \sum_{\tau=-M}^{M} k^*(\frac{\tau}{M}) c_{xx}(\tau) \ . \end{aligned}$$

Then under (4) we have

$$\frac{n^{-2d_x-1}}{M}\sum_{t=1}^{n}\sum_{s=1}^{n}b_{M,n}(t-s)x_t x_s \Rightarrow G_u\int_0^1 B^2(r;d_u)dr\ ,\ \text{as } n\to\infty\ .$$

As a consequence of nonstationarity the weight function $b(.)$ is not invariant with respect to n and M.

4. APPENDIX

Proof of Lemma 1 Recall that

$$\begin{aligned} E\sum_{p=-M}^{M}k(\frac{p}{M})c_{xx}(p) &= \sum_{p=-M}^{M}k(\frac{p}{M})(1-\frac{p}{n})\gamma_{xx}(p)\ , \\ E\sum_{p=-M}^{M}k(\frac{p}{M})c_{xe}(p) &= \sum_{p=-M}^{M}k(\frac{p}{M})(1-\frac{p}{n})\gamma_{xe}(p)\ , \end{aligned}$$

where, under (4), (13) and (14)

$$\lim_{M\to\infty}\frac{\sum_{p=-M}^{M}(1-\frac{p}{n})k(\frac{p}{M})\gamma_{xx}(p)}{\sum_{p=-M}^{M}k(\frac{p}{M})\gamma_{xx}(p)} = \lim_{M\to\infty}\frac{\sum_{p=-M}^{M}(1-\frac{p}{n})k(\frac{p}{M})\gamma_{xe}(p)}{\sum_{p=-M}^{M}k(\frac{p}{M})\gamma_{xe}(p)} = 1\ , \tag{30}$$

by the dominated convergence theorem. Hence it is enough to prove that under Assumption A we have

$$\sum_{p=-M}^{M}k(\frac{p}{M})\left\{c_{xx}(p)-(1-\frac{p}{n})\gamma_{xx}(p)\right\} = o_p(M^{2d_x})\ , \tag{31}$$

$$\sum_{p=-M}^{M}k(\frac{p}{M})\left\{c_{xe}(p)-(1-\frac{p}{n})\gamma_{xe}(p)\right\} = o_p(M^{d_x+d_e})\ . \tag{32}$$

For (31)/(32), it is sufficient to show that

$$\sum_{p=-M+1}^{M-1}\sum_{q=-M+1}^{M-1}|Cov(c_{xa}(p),c_{xa}(q))| = o(M^{2d_x+2d_a})\ ,\ a=x,e\ .$$

From Hannan (1970), p.209, we have that $Cov(c_{xa}(p),c_{xa}(q))$ is equal to

$$\frac{1}{n}\sum_{r=-n+1}^{n-1}(1-\frac{|r|}{n})(\gamma_{xx}(r)\gamma_{aa}(r+q-p)+\gamma_{xa}(r+q)\gamma_{ax}(r-p) \tag{33}$$

$$+\frac{1}{n^2}\sum_{r=-n+1}^{n-1}\sum_{s=1-r}^{n-r}cum_{xaxa}(s,s+p,s+r,s+r+q) \tag{34}$$

where $cum_{xaxa}(q,r,s,t)$ represents the fourth-order cumulant of x_q, a_r, x_s, a_t. Now for (33), when $d_x + d_a < \frac{1}{2}$

$$\begin{aligned} |\sum_{r=-n+1}^{n-1} \gamma_{xx}(r)\gamma_{aa}(r+p-q) + \gamma_{xa}(r+p)\gamma_{ax}(r-q)| &\leq C \sum_{r=-\infty}^{\infty} \left\{\gamma_{xx}^2(r) + \gamma_{xa}^2(r)\right\} \\ &< \infty \end{aligned}$$

because $\gamma(r)\gamma(s) \leq \frac{1}{2}\left\{\gamma^2(r) + \gamma^2(s)\right\}$, $r,s = 0, \pm 1, \pm 2, ...$ and hence

$$\begin{aligned} &\sum_{p=-M}^{M} \sum_{q=-M}^{M} \frac{1}{n} |\sum_{r=-n+1}^{n-1} (1 - \frac{|r|}{n}) \left\{\gamma_{xx}(r)\gamma_{aa}(r+p-q) + \gamma_{xa}(r+p)\gamma_{ax}(r-q)\right\}| \\ = \; & O(\frac{M^2}{n}) = o(1) \; . \end{aligned}$$

For $d_x + d_a \geq \frac{1}{2}$, and in view of (10),

$$\begin{aligned} &\sum_{p=-M}^{M} \sum_{q=-M}^{M} \frac{1}{n} |\sum_{r=-n+1}^{n-1} (1 - \frac{|r|}{n}) \gamma_{xx}(r)\gamma_{aa}(r+p-q)| \\ \leq \; & \frac{M}{n} \sum_{\tau=M-1}^{M-1} \sum_{r=-n+1}^{n-1} |\gamma_{xx}(r)\gamma_{aa}(r+\tau)| \leq C\frac{M}{n} \sum_{\tau=M-1}^{M-1} \sum_{r=-n+1}^{n-1} |r|^{2d_x-1}|r+\tau|^{2d_a-1} \\ = \; & O(\frac{M}{n} M^{2d_x+2d_a}) = o(M^{2d_x+2d_a}) \; . \end{aligned}$$

Also, by Cauchy-Schwarz and elementary inequalities

$$\begin{aligned} &\sum_{p=-M}^{M} \sum_{q=-M}^{M} \frac{1}{n} |\sum_{r=-n+1}^{n-1} (1 - \frac{|r|}{n}) \gamma_{xa}(r+p)\gamma_{ax}(r-q)| \\ \leq \; & C\frac{M^2}{n} \sum_{r=-n+1}^{n-1} \gamma_{xa}^2(r) = O(M^2 n^{2d_x+2d_a-2}) = o(M^{2d_x+2d_a}) \; . \end{aligned}$$

For (34), by Hannan (1970), p.211 and Assumption A we have that

$$cum_{xaxa}(p,q,r,s) \leq C \sum_{d=0}^{\infty} g(p+d)g(d+q-p)g(d+r-p)g(d+s-p)$$

with $g(u) = (|u|+1)^{d_x-1}$. Hence (34) is bounded by

$$\begin{aligned} &\frac{C}{n^2} \sum_{r=0}^{n-1} \sum_{s=1-r}^{n-r} \sum_{d=0}^{\infty} g(d)g(d+p)g(d+r)g(d+r+q) \\ \leq \; & \frac{C}{n} \sum_{r=-n+1}^{n-1} \sum_{d=-\infty}^{\infty} g(d)g(d+p)g(d+r)g(d+r+q) \\ \leq \; & \frac{C}{n} \sum_{r=-n+1}^{n-1} \sum_{d=-\infty}^{\infty} 4g^4(d) = O(1) \; , \end{aligned}$$

where the last inequality follows from $ABCD \leq A^4 + B^4 + C^4 + D^4$, which holds for real-valued A, B, C, D.

Proof of Theorem 1 We can rewrite

$$\begin{aligned} \widetilde{\beta}_M - \beta &= \left\{ \sum_{p=-M}^{M} k(\frac{p}{M}) c_{xx}(p) \right\}^{-1} \sum_{p=-M}^{M} k(\frac{p}{M}) c_{xe}(p) \\ &= A^{-1} b \end{aligned}$$

for

$$\begin{aligned} A &= \sum_{p=-M}^{M} (1 - \frac{p}{n}) k(\frac{p}{M}) \gamma_{xx}(p) + \sum_{p=-M}^{M} k(\frac{p}{M}) \left\{ c_{xx}(p) - (1 - \frac{p}{n}) \gamma_{xx}(p) \right\} \\ b &= \sum_{p=-M}^{M} (1 - \frac{p}{n}) k(\frac{p}{M}) \gamma_{xe}(p) + \sum_{p=-M}^{M} k(\frac{p}{M}) \left\{ c_{xe}(p) - (1 - \frac{p}{n}) \gamma_{xe}(p) \right\} \end{aligned}$$

Now by Lemma 1 and Assumption A

$$\begin{aligned} \sum_{p=-M}^{M} (1 - \frac{p}{n}) k(\frac{p}{M}) \gamma_{xx}(p) A^{-1} &= 1 + o_p(1) , \\ \left\{ \sum_{p=-M}^{M} (1 - \frac{p}{n}) k(\frac{p}{M}) \gamma_{xe}(p) \right\}^{-1} b &= 1 + o_p(1) , \end{aligned}$$

and hence the result follows by Slutsky's theorem and (30).

Proof of Lemma 3 Because by Assumption B $M^{-1} \sum_{\tau=-M}^{M} k(\tau/M) \sim 1$ as $M \to \infty$, it is sufficient to prove that

$$\frac{1}{M} \sum_{\tau=1}^{M} k(\frac{\tau}{M}) \{ c_{xx}(0) - c_{xx}(\tau) \} = o_p(n^{2d_x - 1}) . \quad (35)$$

The right hand side of (35) is equal to $2/M$ times

$$\begin{aligned} & \frac{1}{n} \sum_{\tau=1}^{M} k(\frac{\tau}{M}) \left\{ \sum_{t=1}^{n} x_t^2 - \sum_{t=\tau+1}^{n} x_t x_{t-\tau} \right\} \\ = \; & \frac{1}{n} \sum_{\tau=1}^{M} k(\frac{\tau}{M}) \sum_{t=1}^{\tau} x_t^2 \quad (36) \\ & + \frac{1}{n} \sum_{\tau=1}^{M} k(\frac{\tau}{M}) \sum_{t=\tau+1}^{n} x_t (x_t - x_{t-\tau}) . \quad (37) \end{aligned}$$

For (36), we have easily

$$\sum_{t=1}^{M} \left\{ \sum_{\tau=t}^{M} k(\frac{\tau}{M}) \right\} x_t^2 \leq CM \sum_{t=1}^{M} x_t^2 ,$$

where

$$\sum_{\tau=1}^{M} x_t^2 = \sum_{t=1}^{M} \left\{ \sum_{s=1}^{t} \sum_{j=0}^{\infty} \alpha_j \xi_{s-j} \right\}^2 = \sum_{t=1}^{M} \sum_{s=1}^{t} \sum_{k=1}^{t} \sum_{j=0}^{\infty} \sum_{i=0}^{\infty} \alpha_j \alpha_i \xi_{s-j} \xi_{k-i} \, .$$

To bound the expected value of the above (non-negative) random variable, we use (19)/(20) to obtain

$$\begin{aligned} \sum_{t=1}^{M} \sum_{s=1}^{t} \sum_{k=1}^{t} \sum_{j=0}^{\infty} \alpha_j \alpha_{k-s+j} &\leq C \sum_{t=1}^{M} \sum_{s=1}^{t} \sum_{k=1}^{t} (|k-s|+1)^{2d_u-1} \\ &= C \sum_{t=1}^{M} t \sum_{v=0}^{t-1} (1 - \frac{|v|}{t}) v^{2d_u-1} \leq CM^{2d_x} \, , \end{aligned}$$

whence it follows that $\sum_{t=1}^{M} x_t^2 = o_p(n^{2d_x})$. By Cauchy-Schwarz inequality, (37) is bounded by

$$\frac{1}{n} \left\{ \sum_{t=1}^{n} x_t^2 \right\}^{1/2} \sum_{\tau=1}^{M} k(\frac{\tau}{M}) \left\{ \sum_{t=\tau+1}^{n} (x_t - x_{t-\tau})^2 \right\}^{1/2} \, .$$

The last element has stochastic order of magnitude

$$\begin{aligned} & O_p(\frac{1}{n} \left\{ \sum_{t=1}^{n} x_t^2 \right\}^{1/2}) O_p(E \sum_{\tau=1}^{M} k(\frac{\tau}{M}) \left\{ \sum_{t=\tau+1}^{n} (x_t - x_{t-\tau})^2 \right\}^{1/2}) \\ = \; & O_p(n^{d_x-1} E \sum_{\tau=1}^{M} \left\{ \sum_{t=\tau+1}^{n} (x_t - x_{t-\tau})^2 \right\}^{1/2}) \, , \end{aligned} \tag{38}$$

in view of Assumption B and because $\{n^{-1} \sum_{t=1}^{n} x_t^2\}^{1/2} = O_p(n^{d_x-1})$ follows from previous calculations. From Jensen's inequality

$$C \sum_{\tau=1}^{M} E \left\{ \sum_{t=\tau+1}^{n} (x_t - x_{t-\tau})^2 \right\}^{1/2} \leq C \sum_{\tau=1}^{M} \left\{ \sum_{t=\tau+1}^{n} E(x_t - x_{t-\tau})^2 \right\}^{1/2} ,$$

where

$$\begin{aligned} \sum_{t=\tau+1}^{n} E(x_t - x_{t-\tau})^2 &= \sum_{t=\tau+1}^{n} E \left\{ \sum_{s=t-\tau+1}^{t} u_s \right\}^2 \\ &= \sum_{t=\tau+1}^{n} E \sum_{s=t-\tau+1}^{t} \sum_{k=t-\tau+1}^{t} \sum_{j=0}^{\infty} \sum_{i=0}^{\infty} \alpha_j \alpha_i \xi_{s-j} \xi_{k-i} \\ &< C \sum_{t=\tau+1}^{n} \sum_{s=0}^{\tau-1} \sum_{k=0}^{\tau-1} \sum_{j=0}^{\infty} \alpha_j \alpha_{k-s+j} \\ &\leq C\tau \sum_{t=\tau+1}^{n} \sum_{v=0}^{\tau-1} v^{2d_u-1} \leq Cn\tau^{2d_x-1} \, . \end{aligned}$$

It follows that

$$\sum_{\tau=1}^{M}\left\{\sum_{t=\tau+1}^{n}E(x_t-x_{t-\tau})^2\right\}^{1/2}\le C\sum_{\tau=1}^{M}\left\{n\tau^{2d_x-1}\right\}^{1/2}=o(n^{1/2}M^{d_x+1/2})\ ;$$

hence in view of (38), (37) is $o_p(n^{2d_x})$, which completes the proof of Lemma 3.

Proof of Lemma 5 In the sequel, we repeatedly use the inequality

$$Ea_tb_s\le C(1+|t-s|)^{d_a+d_b-1},\ a,b=u,e\ , \tag{39}$$

which holds under Assumption D. We have

$$\begin{aligned}c_{xe}(0)-\frac{1}{M}\sum_{\tau=-M}^{M}k(\frac{\tau}{M})c_{xe}(\tau) &= \frac{1}{M}\sum_{\tau=1}^{M}k(\frac{\tau}{M})\left\{c_{xe}(0)-c_{xe}(\tau)\right\}\\ &\quad+\frac{1}{M}\sum_{\tau=-m}^{-1}k(\frac{\tau}{M})\left\{c_{xe}(0)-c_{xe}(\tau)\right\}\\ &\quad+c_{xe}(0)\left\{1-\frac{1}{M}\sum_{\tau=-M}^{M}k(\frac{\tau}{M})\right\}\\ &= (I)+(II)+(III)+(IV)+(V)\end{aligned}$$

with

$$\begin{aligned}I &= \frac{1}{Mn}\sum_{\tau=1}^{M}k_M(\tau)\sum_{t=\tau+1}^{n}x_t(e_t-e_{t-\tau})\ ,\ II=\frac{1}{Mn}\sum_{\tau=1}^{M}k_M(\tau)\sum_{t=1}^{\tau}x_te_t\\ III &= \frac{1}{Mn}\sum_{\tau=-M}^{-1}k_M(\tau)\sum_{t=\tau+1}^{n}(x_t-x_{t-\tau})e_t\ ,\ IV=\frac{1}{Mn}\sum_{\tau=-M}^{-1}k_M(\tau)\sum_{t=1}^{\tau}x_te_t\ ,\\ V &= c_{xe}(0)\left\{1-\frac{1}{M}\sum_{\tau=-M}^{M}k(\frac{\tau}{M})\right\}\ .\end{aligned}$$

Define

$$\widetilde{\sum}=\sum_{t=1}^{M}\sum_{s=1}^{M}\sum_{v=s}^{M}\sum_{\tau=t}^{M}\ .$$

In view of Isserlis formula (Brillinger (1981), page 21) which for zero-mean Gaussian variables gives

$$Ex_1x_2x_3x_4=Ex_1x_2Ex_3x_4+Ex_1x_3Ex_3x_2+Ex_1x_4Ex_3x_2\ ,$$

the expected value of the square of (II) is bounded by

$$E\left\{\sum_{t=1}^{M}\left\{\sum_{\tau=t}^{M}k_M(\tau)\right\}x_te_t\right\}^2\le\frac{C}{(Mn)^2}E\widetilde{\sum}x_te_tx_se_s=\frac{C}{(Mn)^2}\left\{\Gamma_1+\Gamma_2+\Gamma_3\right\}\ ,$$

for

$$\Gamma_1 = \widetilde{\sum} Ex_te_tEx_se_s \ , \ \Gamma_2 = \widetilde{\sum} Ex_tx_sEe_te_s \ , \ \Gamma_3 = \widetilde{\sum} Ex_se_tEx_te_s \ .$$

Now in view of (39), Γ_1 is bounded by

$$\widetilde{\sum} E\sum_{j=1}^{t} u_je_tE\sum_{i=1}^{s} u_ie_s \le C\widetilde{\sum} t^{d_u+d_e}s^{d_u+d_e} = O(M^{2d_u+2d_e+4}) = o(M^2n^{2d_x+2d_e}) \ ,$$

and Γ_2 is bounded by

$$\begin{aligned} & CM^2\sum_{t=1}^{M}\sum_{s=1}^{M}(|t-s|+1)^{2d_e-1}\sum_{i=1}^{t}\sum_{j=1}^{s} Eu_iu_j \\ \le \ & CM^2\sum_{t=1}^{M}\sum_{s=1}^{M}(|t-s|+1)^{2d_e-1}M\sum_{k=1}^{M}|k|^{2d_u-1} \\ = \ & O(M^{2d_u+1}\sum_{t=1}^{M}\sum_{s=1}^{M}(|t-s|+1)^{2d_e-1}) = o(M^2n^{2d_x+2d_e}) \ , \end{aligned}$$

As far as Γ_3 is concerned we have that

$$\begin{aligned} \Gamma_3 & \le CM^2\sum_{t=1}^{M}\sum_{s=1}^{M}\sum_{j=1}^{M}\sum_{i=1}^{M}(|s-j|+1)^{d_u+d_e-1}(|t-i|+1)^{d_u+d_e-1} \\ & = O(M^{2d_u+2d_e+4}) = o(M^2n^{2d_x+2d_e}) \ . \end{aligned}$$

Hence $\Gamma_i = o(M^2n^{2d_x+2d_e})$, $i = 1, 2, 3$, and (II) is $o_p(n^{d_x+d_e-1})$; same argument can be applied to (IV). For (I), we can rewrite

$$\sum_{\tau=1}^{M} k(\frac{\tau}{M})\left\{\sum_{t=\tau+1}^{n} x_te_t - \sum_{t=\tau+1}^{n}(x_t - x_{t-\tau})e_{t-\tau} - \sum_{t=\tau+1}^{n} x_{t-\tau}e_{t-\tau}\right\} = \Delta_1 - \Delta_2 - \Delta_3 \ ,$$

where

$$\begin{aligned} \Delta_1 & = \sum_{\tau=1}^{M} k(\frac{\tau}{M})\sum_{t=n-\tau+1}^{n} x_te_t \ , \ \Delta_2 = \sum_{\tau=1}^{M} k(\frac{\tau}{M})\sum_{t=1}^{\tau} x_te_t \ , \\ \Delta_3 & = \sum_{\tau=1}^{M} k(\frac{\tau}{M})\sum_{t=\tau+1}^{n}(x_t - x_{t-\tau})e_{t-\tau} \ . \end{aligned}$$

Now Δ_1 and Δ_2 are, apart from a change of index, proportional to (II) which we analyzed before. On the other hand, Δ_3 can be dealt exactly as (III), again with a change of index; therefore we analyze (III). In view of (15), we have

$$\sum_{t=\tau+1}^{n}(x_t - x_{t-\tau})e_t = \sum_{t=\tau+1}^{n}\sum_{i=t-\tau+1}^{t} u_ie_t \ .$$

Also

$$\left\{ \sum_{t=\tau+1}^{n} \sum_{i=t-\tau+1}^{t} u_i e_t \right\}^2 = \widehat{\sum} e_t u_s e_s u_j \ ,$$

with

$$\widehat{\sum} = \sum_{s=\tau+1}^{n} \sum_{t=\tau+1}^{n} \sum_{i=t-\tau+1}^{t} \sum_{j=s-\tau+1}^{s} \ .$$

Hence we obtain $E\widehat{\sum} e_t u_i e_s u_j = \Theta_1 + \Theta_2 + \Theta_3$, for

$$\Theta_1 = \widehat{\sum} E e_t e_s E u_i u_j, \ \Theta_2 = \widehat{\sum} E e_t u_i E e_s u_j, \ \Theta_3 = \widehat{\sum} E e_t u_j E u_i e_s \ .$$

Thus

$$\begin{aligned} \Theta_1 &= E \sum_{s=\tau+1}^{n} \sum_{t=\tau+1}^{n} e_t e_s E \sum_{i=t-\tau+1}^{t} \sum_{j=s-\tau+1}^{s} u_i u_j \\ &= O(n(n-\tau)^{2d_e}) O(\tau^{2d_u+1}) = o(n^{2d_x+2d_e}) \ , \end{aligned}$$

$$\begin{aligned} \Theta_2 &= \sum_{t=\tau+1}^{n} \sum_{i=t-\tau+1}^{t} E e_t u_i \sum_{s=\tau+1}^{n} \sum_{j=s-\tau+1}^{s} E e_s u_j \\ &\leq C \sum_{t=\tau+1}^{n} \sum_{i=t-\tau+1}^{t} (|t-i|+1)^{d_u+d_e-1} \sum_{s=\tau+1}^{n} \sum_{j=s-\tau+1}^{s} (|s-j|+1)^{d_u+d_e-1} \\ &= O(\left\{ (n-\tau)\tau^{d_u+d_e} \right\}^2) = o(n^{2d_x+2d_e}) \ , \end{aligned}$$

$$\begin{aligned} \Theta_3 &= \sum_{t=\tau+1}^{n} \sum_{j=s-\tau+1}^{s} E e_t u_j \sum_{s=\tau+1}^{n} \sum_{i=t-\tau+1}^{t} E u_i e_s \\ &= O(\left\{ (n-\tau)\tau^{d_u+d_e} \right\}^2) = o(n^{2d_x+2d_e}) \ . \end{aligned}$$

Because the expected value of the square of (III) is bounded by $Cn^{-2}\left\{\Theta_1 + \Theta_2 + \Theta_3\right\}$, we have easily $(III) = o_p(n^{d_x+d_e-1})$.

Finally, from Assumption B we have $(V) = o_p(c_{xe}(0)) = o_p(n^{d_x+d_e-1})$, the last equality following from (24).

REFERENCES

Billingsley, P. (1968) *Convergence of Probability Measures,* J.Wiley, New York

Brillinger, D.R. (1981) *Time Series: Data Analysis and Theory*, Holt, Rinehart & Winston, New York

Brockwell, P.J. and R.A.Davis (1991) *Time Series: Theory and Methods,* 2nd edition, New York, Springer and Verlag

Chan, N.H. and N.Terrin (1995) "Inference for Unstable Long-Memory Processes with Applications to Fractional Unit Root Autoregressions", *Ann. Stat.*, 23, 1662-1683

Davydov, Y.A. (1970) "The Invariance Principle for Stationary Processes", *Theory Probab. App.* 15, 487-498

Deo, R.S. (1997) "Asymptotic Theory for Certain Regression Models with Long Memory Errors", *J. Time Ser. Anal.,* 18, 385-393

Dolado, J. and F.Marmol (1996) "Efficient Estimation of Cointegrating Relationships Among Higher Order and Fractionally Integrated Processes", Banco de Espana, *Working Paper n.9617*

Engle, R. (1974) "Band Spectrum Regression", *Int.Ec.Rev.,* 15, 1-11

Engle, R. and C.Granger (1987) "Cointegration and Error Correction: Representation, Estimation and Testing", *Econometrica,* 55, 251-276

Giraitis, L. and M.S.Taqqu (1998a) "Whittle Estimator for Finite-Variance Non-Gaussian Time Series with Long Memory", *Ann. Stat.,* forthcoming

Giraitis, L. and M.S.Taqqu (1998b) "Convergence of Normalized Quadratic Forms", *J. Statist. Plann. Inf.,* forthcoming

Gorodetskii, V.V. (1977) "On Convergence to Semi-Stable Gaussian Processes", *Theory Prob. Appl.,* 22, 498-508

Granger, C.W.J. and R.Joyeux (1980) "An Introduction to Long Memory Time Series Models and Fractional Differencing", *J. Time Ser. Anal.* 1, 15-29

Hannan, E.J. (1963) "Regression for Time Series with Errors of Measurement", *Biometrika,* 50, 293-302

Hannan, E.J. (1970) *Multiple Time Series*, New York, John Wiley and Sons

Hannan, E.J. and P.M.Robinson (1973) "Lagged Regression with Unknown Lags", *J. Royal Stat. Soc.,* Series B, 35, 252-267

Hansen, B.E. (1992) "Convergence to Stochastic Integrals for Dependent Heterogenous Processes", *Econ. Theory,* 8, 489-501

Hurvich, C.M., R.Deo and J.Brodsky (1998) "The Mean Squared Error of Geweke and Porter-Hudak's Estimates of the Memory Parameter of a Long Memory Time Series", *J. Time Ser. Anal.,* 19, 19-46

Hurvich, C.M. and B.K.Ray (1995) "Estimation of the Memory Parameter for Nonstationary or Noninvertible Fractionally Integrated Processes", *J. Time Ser. Anal.,* 16, 17-41

Jeganathan, P. (1996) "On Asymptotic Inference in Cointegrated Time Series with Fractionally Integrated Errors", University of Michigan, Dep. of Statistics, *Technical Report #277*

Lobato, I. (1997) "Consistency of the Averaged Cross-Periodogram in Long Memory Series", *J. Time Ser. Anal.,* 18, 137-156

Major, P. (1981) "Limit Theorems for Non-Linear Functionals of Gaussian Sequences", *Zeit. Wahrsch. verw. Geb.,* 57, 129-158

Phillips, P.C.B. (1988) "Weak Convergence of Sample Covariance Matrix to Stochastic Integrals via Martingale Approximations", *Econ. Theory,* 4, 528-533

Phillips, P.C.B. (1991) "Spectral Regression for Cointegrated Time Series", in *Nonparametric and Semiparametric Methods in Economics and Statistics,* ed.by W.Barnett, J.Powell and G.Tauchen, New York, Cambridge University Press

Robinson, P.M. (1994a) "Semiparametric Analysis of Long-Memory Time Series", *Ann. Stat.,* 22, 515-539

Robinson, P.M. (1994b) "Rates of Convergence and Optimal Spectral Bandwidth for Long Range Dependence", *Prob. Theory Rel. Fields,* 99, 443-473

Robinson, P.M. and J.Hidalgo (1997) "Time Series Regression with Long Range Dependence", *Ann. Stat.,* 25, 77-104

Robinson, P.M. and D.Marinucci (1998) "Semiparametric Frequency Domain Analysis of Fractional Cointegration", *STICERD Discussion Paper #348*, LSE

Samorodnitsky, G. and M.S.Taqqu, (1994) *Stable Non-Gaussian Random Processes,* Chapman and Hall, New York

Velasco, C. (1997) "Gaussian Semiparametric Estimation of Non-Stationary Time Series", preprint

Watson, M.W. (1994) "Vector Autoregression and Cointegration", in *Handbook of Econometrics,* vol.4, R.F.Engle and D.L.McFadden eds., Elsevier Science