Econometric Theory, **22**, 2006, 765–814. Printed in the United States of America.
DOI: 10.1017/S0266466606060361

UNBALANCED COINTEGRATION

JAVIER HUALDE
Universidad de Navarra

Recently, increasing interest in the issue of fractional cointegration has emerged from theoretical and empirical viewpoints. Here, as opposed to the traditional prescription of unit root observables with weak dependent cointegrating errors, the orders of integration of these series are allowed to take real values, but, as in the traditional framework, equality of the orders of at least two observable series is necessary for cointegration. This assumption, in view of the real-valued nature of these orders, could pose some difficulties, and in the present paper we explore some ideas related to this issue in a simple bivariate framework. First, in a situation of "near-cointegration," where the only difference with respect to the "usual" fractional cointegration is that the orders of the two observable series differ in an asymptotically negligible way, we analyze properties of standard estimates of the cointegrating parameter. Second, we discuss the estimation of the cointegrating parameter in a situation where the orders of integration of the two observables are truly different but their corresponding balanced versions (with same order of integration) are cointegrated in the usual sense. A Monte Carlo study of finite-sample performance and simulated series is included.

1. INTRODUCTION

Cointegration has traditionally focused on the case of unit root (or integrated of order one) observable processes with weak dependent (or integrated of order zero) cointegrating errors. Formally, we consider that a zero-mean scalar covariance stationary process ζ_t, $t \in Z$, $Z = \{t : t = 0, \pm 1, \dots\}$, with spectral density $f_\zeta(\lambda)$, is integrated of order zero, denoted $\zeta_t \sim I(0)$, if

$$0 < f_\zeta(0) < \infty,$$

whereas a zero-mean scalar process is $I(d)$ if it could be represented as an $I(0)$ process after differencing it d times. More precisely, let ζ_t, $t \in Z$, be an $I(0)$ process and then define the Type II fractionally integrated of order d process ξ_t, denoted $\xi_t \sim I(d)$, as

I thank Adrian Pagan, James Davidson, and seminar participants at the 2004 Econometric Society European Meeting and the 2004 Simposio de Análisis Económico for helpful comments. I also thank two referees and a co-editor whose comments led to improvements of the paper. This research was supported by the Spanish Ministerio de Educación y Ciencia through a contract Juan de la Cierva and ref. SEJ2005-07657/ECON, and also by the Universidad de Navarra, ref. 16037001. Address correspondence to Javier Hualde, Department of Quantitative Methods, School of Economics and Business Administration, Universidad de Navarra, Edificio Bibliotecas (Entrada Este), 31080 Pamplona (Navarra), Spain; e-mail: jhualde@unav.es.

 0266-4666/06 $12.00

$$\xi_t = \Delta^{-d}\zeta_t^{\#}, \qquad t \in Z, \tag{1}$$

where $\Delta = 1 - L$, L is the lag operator,

$$(1-L)^{-\alpha} = \sum_{j=0}^{\infty} a_j(\alpha)L^j, \qquad a_j(\alpha) = \frac{\Gamma(j+\alpha)}{\Gamma(\alpha)\Gamma(j+1)}, \tag{2}$$

where $\Gamma(\cdot)$ represents the gamma function, taking $\Gamma(\alpha) = \infty$ for $\alpha = 0, -1, -2, \ldots$, and $\Gamma(0)/\Gamma(0) = 1$. The # superscript attached to a scalar or vector sequence h_t has the meaning

$$h_t^{\#} = h_t 1(t > 0), \tag{3}$$

where $1(\cdot)$ is the indicator function. Note that because of the truncation on the right-hand side of (1), $\xi_t = 0$, $t \leq 0$, whereas integer values of d provide the typical definition of integrated process with a particular initial condition. For example, if $d = 1$, as $a_j(1) = 1$, $j \geq 0$,

$$\xi_t = \sum_{j=1}^{t} \zeta_j 1(t > 0).$$

For an alternative definition of fractionally integrated process (the Type I class) see Marinucci and Robinson (1999). Note that for $d < \frac{1}{2}$ the truncation on the right-hand side of (1) is not strictly necessary and implies that ξ_t is only asymptotically stationary, in the precise sense defined by Robinson and Marinucci (2001). If $d \geq \frac{1}{2}$, this truncation implies that the variance of ξ_t is finite (albeit evolving at rate $O(t^{2d-1})$), so that ξ_t is well defined in mean-squared sense.

The traditional situation mentioned previously was denoted by Engle and Granger (1987) as $CI(1,1)$, first and second arguments referring to the integration orders of the observables and the reduction of these orders by certain linear combinations, respectively. Recently, increasing interest about a wider framework where the orders of integration of both observables and cointegrating errors could be real, but perhaps not integer, numbers has emerged. This, in view of the work of Granger and Joyeux (1980) and Hosking (1981) on fractionally integrated processes, represents a natural generalization of the $CI(1,1)$ framework, which was already anticipated by Engle and Granger (1987) because their $CI(d,b)$ definition did not necessarily require d, b to be integers. This general setting considers $CI(1,1)$ as a particular situation, noting that $I(1)$ and $I(0)$ are very specific cases of nonstationarity and stationarity, respectively, whereas it also allows for consideration of "stationary cointegration," where the observables are (perhaps only asymptotically) covariance stationary long-memory processes, with cointegrating errors also being long memory with strictly less memory than the observables or weak dependent $I(0)$ processes. Theoretical works on estimation of the relation of cointegration in fractional frameworks include Robinson (1994), Jeganathan (1999), Kim and Phillips (2000), Robinson and Marinucci (2001, 2003), Hualde and Robinson (2006), Robinson and Hualde (2003), and Hualde and Robinson (2004a).

This new setting introduces additional challenges with respect to the traditional $CI(1,1)$ situation. In particular, assuming that the integration orders are real numbers, the precise knowledge of their values seems difficult to justify, even after pretesting, which contradicts the usual practice in the standard $CI(1,1)$ framework, where the assumed knowledge of these orders is used to derive estimates with optimal asymptotic properties (see, e.g., Phillips, 1991). In the fractional setting, some of the works mentioned earlier deal explicitly with the important issue of unknown integration orders, although, as in the traditional prescription, standard estimates that do not rely on this knowledge could enjoy good asymptotic properties. Among those, the most common ones are the ordinary least squares (OLS) and the narrow band least squares (NBLS), whose asymptotic properties in possibly fractional circumstances were studied by Robinson and Marinucci (2001, 2003).

Furthermore, the real-valued condition of the integration orders also poses additional problems, which can be presented by the next simple example. Suppose we observe the processes y_t, x_t, $t = 1,\ldots,n$, which are $I(\delta_1)$, $I(\delta_2)$, being generated by the $I(0)$ processes ζ_{1t}, ζ_{2t}, respectively. Then, provided $\delta_1 = \delta_2$, a necessary and sufficient condition for the two series y_t, x_t to be cointegrated is that the long-run variance-covariance matrix of the bivariate process $(\zeta_{1t},\zeta_{2t})'$, defined as 2π times its spectral density evaluated at frequency 0, is singular. Needless to say, the existence of cointegration or no cointegration between y_t and x_t has an enormous impact on the behavior of different statistics and estimates. For example, if $\delta_1 = \delta_2 = 1$ and there is no cointegration, Phillips (1986) showed that the OLS estimate $\hat{\nu}_O = \sum_{t=1}^{n} x_t y_t / \sum_{t=1}^{n} x_t^2$ is not consistent for the fundamental coefficient (see Park, Ouliaris, and Choi, 1988)

$$\nu = \frac{\omega_{12}}{\omega_{22}}, \tag{4}$$

where ω_{ij} is the (i,j)th element of the long-run variance-covariance matrix of $(\zeta_{1t},\zeta_{2t})'$. Alternatively, if there is cointegration and $(\zeta_{1t},\zeta_{2t})'$ has a structure such that the linear combination $y_t - \nu x_t$ is an $I(0)$ process, then $\hat{\nu}_O$ is an n-consistent estimate of ν, with nonstandard limiting distribution (Phillips and Durlauf, 1986).

The present paper does not focus on departures from the standard notion of cointegration affecting the singularity of this long-run variance-covariance matrix (see, e.g., Jansson and Haldrup, 2002), but on consequences of relaxing the condition of equality of the integration orders of the observables, which is necessary for the existence of cointegration. For example, let $\delta_2 = 1$ and δ_1 be a fixed value such that $0 < \delta_1 \neq 1$ in the previous example. Clearly, any linear combination of y_t and x_t will be $I(1)$ or $I(\delta_1)$ depending on whether $\delta_1 < 1$ or $\delta_1 > 1$, respectively. Thus, irrespective of the possible singularity of the long-run variance-covariance matrix of ζ_{1t}, ζ_{2t}, following Robinson and Marinucci (2001),

$$\hat{\nu}_O = O_p(n^{\delta_1 - 1}),$$

so that the OLS estimate either converges to zero or diverges as n tends to infinity, in the case $\delta_1 < 1$ or $\delta_1 > 1$, respectively. We find two very relevant issues related to this discussion. First, even in the situation where the singularity condition on the input error process holds and the difference between the orders δ_1, δ_2 is very small, the asymptotic theory predicts that the OLS estimate does not converge at all to the fundamental coefficient ν. This fact is somehow counterintuitive, as one could suspect that if δ_1 and δ_2 are very close to each other and the long-run variance-covariance matrix is singular, we are in fact in a "near cointegration" situation, so that estimates like the OLS should have a closer finite-sample behavior to the proper case of cointegration than to the one predicted by the theory with distinct orders of integration. The Monte Carlo experiment reported in Section 5 supports this guess, which we find very appealing when dealing with fractional orders. Here, as mentioned before, there could be situations where it is certainly not realistic to assume that the orders of integration present in the model are known. If this is the case, any testing procedure for cointegration should include a pretest for equality of the orders of integration. This was considered theoretically by Robinson and Yajima (2002) and Hualde (2002) and empirically by Marinucci and Robinson (2001). But even if we conclude that the orders of integration of two processes are statistically equal, their real-valued essence could, in certain circumstances, make us suspect that the orders are perhaps not strictly equal but only very close to each other. Strictly speaking, this would not be a situation of cointegration, but in practice, properties of the estimates might not be very much affected by minor differences in orders of integration. In this case, the typical cotrending that cointegration implies would be only approximate but could be sufficient to infer sensible statistical results.

This idea is very close in spirit to some well-established evidence in the literature. This mainly refers to the nearly nonstationary first-order autoregressive (AR(1)) model studied by, among others, Ahtola and Tiao (1984), Chan and Wei (1987), Phillips (1987, 1988), Cox and Llatas (1991), and Elliott (1998). These studies were motivated by the well-known fact that for the AR(1) process

$$w_t = \phi w_{t-1} + \zeta_t,$$

where ζ_t, $t = 0, \pm 1, \ldots$, is an independent and identically distributed sequence, when $|\phi| < 1$,

$$\hat{\phi} = \frac{\sum_{t=2}^{n} w_t w_{t-1}}{\sum_{t=2}^{n} w_{t-1}^2}$$

https://doi.org/10.1017/S0266466606060361 Published online by Cambridge University Press

is $\sqrt{n}$-consistent and asymptotically normal, but this limiting distribution provides a poor approximation to the actual finite-sample distribution of $\hat{\phi}$ for moderate n when ϕ is close to (but below) 1. Evans and Savin (1981) presented numerical evidence that the nonstandard limiting distribution of $\hat{\phi}$ when $\phi = 1$, which is n-consistent, provides a better approximation when ϕ is close to but below 1. This issue becomes very relevant if we are uncertain about whether a process has a root of unity or in the vicinity of unity. These authors studied the limiting distribution of $\hat{\phi}$ in the case

$$\phi = 1 - \frac{\alpha}{n},$$

which for a certain positive fixed real number α is smaller but approximating 1 as n tends to infinity, w_t in this case being a nearly nonstationary AR(1) process. Phillips (1988) denoted near cointegration the situation where a linear combination of $I(1)$ processes was nearly nonstationary. Jansson and Haldrup (2002) provided an alternative definition of near cointegration that complements to a certain extent the work of Phillips (1988). Taking our previous example, assuming $\delta_1 = \delta_2 = 1$, they analyzed the case where the long-run variance-covariance matrix of $(\zeta_{1t}, \zeta_{2t})'$ tends suitably to a singular matrix, and they examined asymptotic properties of different estimates of ν in this case. Thus, one of our aims seeks to complement these previous analyses in a particular sense that we find very relevant in the case of dealing with fractional orders, that is, the study of a near fractional cointegration situation, where the only departure from strict cointegration is the existence of very small differences in the orders of integration of the observables that tend to disappear as the sample size tends to infinity. We will refer to this situation as weakly unbalanced cointegration.

The relevance of the second issue we analyze in the paper can be also better motivated in the fractional setting. As we will see in Section 3, in a weakly unbalanced cointegrating situation, the main message of our work is that one should not worry about suspected small differences of the integration orders in the data, because even if this happens, standard estimates could retain the properties of the strictly balanced situation. However, there could be cases where the integration orders of the observables are substantially different, so that it is unrealistic to model the relation between the series as one covered by weakly unbalanced cointegration. Traditionally, different orders of integration implied that the relation between the series could not be captured by a cointegrating structure, because the necessary condition of equality of the orders of integration was missing. On the contrary, we find that there could be situations where the integration orders of the observables are substantially different but there still exists a strong intrinsic linkage between the two processes, in particular that the long-run variance-covariance matrix of the input error process generating the integrated processes is singular. Referring to the previous example, suppose $\delta_1 \neq \delta_2$, so that y_t and x_t would not be cointegrated but y_t and $x_t(\delta_2 - \delta_1)$ (which share the same integration order δ_1) are cointegrated in the usual sense, where for a scalar or vector process h_t and real number c,

$$h_t(c) = \Delta^c h_t^{\#}, \tag{5}$$

noting (2) and (3); that is, the linear combination $y_t - \nu x_t(\delta_2 - \delta_1)$ is integrated of an order strictly smaller than δ_1. Thus, it is readily seen that if $\delta_1 = \delta_2$ the relevant parameter to explain the long-run relationship between y_t and x_t is ν. If, on the contrary, the orders are different, but there is still cointegration between a raw and a filtered series, we find that two parameters are relevant to explain the long-run connection between y_t and x_t. These parameters are ν and $\delta_2 - \delta_1$. We will refer to this situation as strongly unbalanced cointegration. We believe that this concept could help practitioners in their task of unmasking anomalies appearing in some estimated models that contradict predictions from economic theory. An excellent motivating puzzle for our work is the forward premium anomaly. In short, this refers to surprising negative estimates from the regression of the change in the logarithm of the spot exchange rate on the forward premium, whereas the theory predicts a theoretical value of one for that slope (see Backus, Foressi, and Telmer, 1996; Bekaert, 1996; Bekaert, Hodrick, and Marshall, 1997). Baillie and Bollerslev (2000) describe this issue as mainly a statistical phenomenon, characterized by the fact that the integration orders of dependent and explanatory variables perhaps could be not the same. In particular, they suggest that the spot exchange rate could be $I(1)$, whereas there seems to be evidence of long-memory behavior of the forward premium. Maynard and Phillips (2001) gave formal theoretical justification to this phenomenon. With these ideas in mind, we hope to offer a sensible statistical solution to what might well be a statistical problem.

The paper is organized as follows. In the next section we present the model, assumptions, and particular estimates whose asymptotic properties will be analyzed in the different situations pinpointed before. Section 3 collects the main results, which are fully characterized in Appendix A and rigorously justified in Appendix B. Section 4 presents alternative estimates for the strongly unbalanced cointegration situation. Finally, Section 5 reports a Monte Carlo study of finite-sample behavior of the different estimates presented in the paper and some artificially generated figures that give further motivation to the strongly unbalanced cointegration situation.

2. MODEL, ASSUMPTIONS, AND PROPOSED ESTIMATES

Throughout the paper, we consider a bivariate triangular array $\{(y_{t,n}, x_{t,n})' : 1 \leq t \leq n\}_{n=1}^{\infty}$ generated by

$$y_{t,n} = \Delta^{-\delta} v_{1t}^{\#}, \tag{6}$$

$$x_{t,n} = \Delta^{-(\delta+\theta_n)} v_{2t}^{\#}, \tag{7}$$

https://doi.org/10.1017/S0266466606060361 Published online by Cambridge University Press

where θ_n is a sequence of real numbers and $v_t = (v_{1t}, v_{2t})'$, $t \in Z$, is a zero-mean bivariate process at least asymptotically stationary with bounded, possibly time-dependent, spectral density $f^{(t)}(\lambda)$. Model (6), (7) is very general and allows us to consider simultaneously different situations depending on θ_n and the structure of v_t, which will be the key element to assess whether or not there is cointegration between $y_{t,n}$ and $x_{t,n}$. We set subsequently specific conditions on θ_n and v_t that will determine different relations between $y_{t,n}$ and $x_{t,n}$, that is, whether they are cointegrated in the wide sense (balanced, weakly, strongly unbalanced cointegrated) or not cointegrated at all. We will use the simplifying notation $y_t = y_{t,n}$, because noting (6), $y_{t,n} = y_{t,n'}$ for $n \neq n'$, and similarly $x_t = x_{t,n}$ in the case $\theta_n = \theta$ for all n. As cointegration has been mainly considered among $I(1)$ or $I(2)$ processes, we will concentrate on the case where in (6), (7)

$$\delta > \tfrac{1}{2}, \tag{8}$$

so that for $\theta_n = 0$ both observables are purely nonstationary and the common cases $\delta = 1$ or 2 are covered by our theory. Note that we treat all purely nonstationary situations except $\delta = \frac{1}{2}$, a borderline case that would require a different approach.

As will be seen in the next section, for the noncointegrated cases we will concentrate on the situation where

$$\theta_n = \theta, \quad \text{for all } n. \tag{9}$$

Here, the particular case $\theta = 0$ was analyzed theoretically by Phillips (1986) for the unit root case $\delta = 1$ and by Marmol (1998) for general real δ in the nonstationary region. This situation was denoted spurious cointegration, and here the typical dimensionality reduction in the stochastic trends explaining jointly the evolution of the observables, which characterizes cointegration, is not present. The noncointegrated cases will be also characterized by the following set of regularity conditions on the process v_t. Throughout, we denote by I_p the $p \times p$ identity matrix.

ASSUMPTION NC (No cointegration). The process v_t, $t \in Z$, has representation

$$v_t = A(L)\varepsilon_t,$$

where

$$A(s) = I_2 + \sum_{j=1}^{\infty} A_j s^j$$

and

(i) $A(e^{i\lambda})$ is differentiable in λ with derivative in $Lip(\varrho)$, $\varrho > \frac{1}{2}$; in addition, with $\|\cdot\|$ denoting the euclidean norm,
(ii) the ε_t are independent and identically distributed vectors with mean zero, positive definite covariance matrix Σ, and $E\|\varepsilon_t\|^q < \infty$, $q \geq 4$,

https://doi.org/10.1017/S0266466606060361 Published online by Cambridge University Press

$q > \max\{2/(2\delta - 1), [2/(2(\delta+\theta) - 1)]1(\delta+\theta > \frac{1}{2})\}$; finally, defining the long-run variance-covariance matrix of v_t as

$$\Omega = A(1)\Sigma A'(1),$$

with (i,j)th element ω_{ij} and the squared correlation coefficient computed from Ω

$$\rho^2 = \frac{\omega_{12}^2}{\omega_{11}\omega_{22}},$$

(iii) $\omega_{11} > 0$, $\omega_{22} > 0$, $\rho^2 < 1$.

Thus, absence of cointegration will be characterized by v_t being a bivariate covariance stationary process with spectral density matrix $f^{(t)}(\lambda) = f(\lambda)$ such that

$$f(\lambda) = \frac{1}{2\pi} A(e^{i\lambda})\Sigma A'(e^{-i\lambda}),$$

with (i,j)th element $f_{ij}(\lambda)$ and

$$rank(f(0)) = rank(\Omega) = 2,$$

as is implied by (iii), which, as mentioned in the Introduction, rules out the possibility of cointegration between y_t and $x_{t,n}$ even if $\theta_n = 0$. Notice that (i) implies $\sum_{j=1}^{\infty} j\|A_j\| < \infty$, because the derivative of $A(e^{i\lambda})$ has Fourier coefficients jA_j, whence Zygmund (1977, p. 240) can be applied. Further, this also implies $\sum_{j=1}^{\infty} j\|A_j\|^2 < \infty$, which, along with the condition in (ii), enables us to apply the functional limit theorem of Marinucci and Robinson (2000) (developing earlier work of Akonom and Gourieroux, 1987; Silveira, 1991), as is required to characterize the limit distribution of our estimates. Also, the moment assumption on ε_t is satisfied, for any $\delta, \delta + \theta > \frac{1}{2}$, by Gaussianity. Finally, note that (i) and (iii) imply that both individual processes v_{1t}, v_{2t}, are $I(0)$.

Next, we will focus on the situation where there exists cointegration in the wide sense. The key here is to characterize properly the structure of v_t in (6), (7), as we do in the next assumption.

ASSUMPTION C (Cointegration). There exist real numbers $\nu \neq 0$ and γ such that

$$0 \leq \gamma < \delta \qquad \textbf{(10)}$$

and a certain process u_{1t}, $t \in Z$, where

https://doi.org/10.1017/S0266466606060361 Published online by Cambridge University Press

(i)

$$v_{1t} = \nu v_{2t} + u_{1t}(\delta - \gamma); \quad \textbf{(11)}$$

(ii) the bivariate process $u_t = (u_{1t}, v_{2t})'$, $t \in Z$ can be represented as

$$u_t = B(L)\epsilon_t, \quad \textbf{(12)}$$

where the conditions on $A(L)$ and ε_t set in Assumptions NC(i) and (ii) apply to $B(L)$, ϵ_t, respectively, with $E(\epsilon_t \epsilon_t') = \Psi$ and $E\|\epsilon_t\|^q < \infty$, $q \geq 4$, $q > \max\{2/(2\delta - 1), [2/(2\gamma - 1)]1(\gamma > \frac{1}{2})\}$; finally, denoting ϕ_{ij} the (i,j)th element of the long-run variance-covariance matrix of u_t, namely,

$$\Phi = B(1)\Psi B'(1),$$

(iii) $\phi_{11} > 0$, $\phi_{22} > 0$.

This assumption has very important implications. First, (ii) and (iii) imply that both u_{1t} and v_{2t} are $I(0)$ with differentiable spectral density matrix. Further, by (i), v_{1t} is decomposed as the sum of an $I(0)$ process and an overdifferenced asymptotically stationary process $u_{1t}(\delta - \gamma)$, noting (5) and (10). This is the most relevant and distinctive condition of Assumption C. An alternative representation could have been to consider

$$\tilde{v}_{1t} = \nu v_{2t} + \Delta^{\delta-\gamma} u_{1t} \quad \textbf{(13)}$$

instead of v_{1t}, where the second term on the right-hand side of (13) is now covariance stationary, so that $\tilde{v}_{1t}$ also shares this condition. However, we find it more justifiable to use the particular v_t characterized by Assumption C, because under this condition (6) and (7) imply that

$$y_t = \nu x_{t,n}(\theta_n) + u_{1t}(-\gamma),$$

$$x_{t,n} = v_{2t}(-(\delta + \theta_n)),$$

which if $\theta_n = 0$ for all n is the bivariate cointegrated system involving Type II fractionally integrated processes considered by Hualde and Robinson (2006) and Robinson and Hualde (2003). Note also that even if y_t and $x_{t,n}$ are not cointegrated in the usual sense, standard cointegration occurs between y_t and $x_{t,n}(\theta_n)$. Thus, the time dependence of v_t seems natural when dealing with Type II processes. In fact, denoting

$$h(\lambda) = \frac{1}{2\pi} B(e^{i\lambda})\Psi B'(e^{-i\lambda}), \quad \textbf{(14)}$$

with (i,j)th element $h_{ij}(\lambda)$, it is straightforward to show that under Assumption C the time-dependent spectral density of v_t, $f^{(t)}(\lambda)$ is

$$f^{(t)}(\lambda) = \begin{pmatrix} a^t(-\beta;\lambda) & \nu \\ 0 & 1 \end{pmatrix} h(\lambda) \begin{pmatrix} a^t(-\beta;-\lambda) & 0 \\ \nu & 1 \end{pmatrix},$$

where

$$a^t(c;\lambda) = \sum_{j=0}^{t-1} a_j(c) e^{ij\lambda}$$

and

$$\beta = \delta - \gamma,$$

which is the cointegrating gap. Clearly, $f^{(t)}(\lambda)$ can be decomposed as

$$f^{(t)}(\lambda) = g(\lambda) + R(\lambda) + R^{(t)}(\lambda), \tag{15}$$

where

$$g(\lambda) = \begin{pmatrix} \nu^2 & \nu \\ \nu & 1 \end{pmatrix} h_{22}(\lambda),$$

$$R(\lambda) = \begin{pmatrix} |a(-\beta;\lambda)|^2 h_{11}(\lambda) + 2\nu \operatorname{Re}\{h_{12}(\lambda) a(-\beta;\lambda)\} & h_{12}(\lambda) a(-\beta;\lambda) \\ h_{21}(\lambda) a(-\beta;-\lambda) & 0 \end{pmatrix},$$

$$R^{(t)}(\lambda) = \begin{pmatrix} b^t(-\beta;\lambda) h_{11}(\lambda) - 2\nu \operatorname{Re}\{h_{12}(\lambda) \bar{a}^t(-\beta;\lambda)\} & -h_{12}(\lambda) \bar{a}^t(-\beta;\lambda) \\ -h_{21}(\lambda) \bar{a}^t(-\beta;-\lambda) & 0 \end{pmatrix},$$

$$a(c;\lambda) = \sum_{j=0}^{\infty} a_j(c) e^{ij\lambda}, \qquad \bar{a}^t(c;\lambda) = \sum_{j=t}^{\infty} a_j(c) e^{ij\lambda},$$

$$b^t(c;\lambda) = |a^t(c;\lambda)|^2 - |a(c;\lambda)|^2.$$

Defining

$$f(\lambda) = g(\lambda) + R(\lambda), \tag{16}$$

$f(\lambda)$ is the spectral density function of $\tilde{v}_t = (\tilde{v}_{1t}, v_{2t})'$, which could be referred to as the covariance stationary process version of v_t. Our cointegrating model implies

$$rank(f(0)) < 2, \tag{17}$$

because the rank of $g(\lambda)$ is reduced for all λ, whereas noting that $a(c;0) = 0$, for any $c < 0$, $R(0) = 0$. Also

$$R(\lambda) = O(\lambda^{\delta-\gamma}) \quad \text{as } \lambda \to 0, \tag{18}$$

https://doi.org/10.1017/S0266466606060361 Published online by Cambridge University Press

where the $O(\cdot)$ notation in (18) does not have the usual meaning of upper bound but exact rate. As mentioned before, the type of cointegration between y_t and $x_{t,n}$ depends crucially on the structure of the process v_t and more specifically on the behavior of the component $R(\lambda)$ in the vicinity of frequency 0. The speed at which this element vanishes as λ tends to 0 is the key to characterize the strength of the cointegrating relationship, and, in view of (18), this is completely determined by the cointegrating gap β. Further, by Stirling's approximation it can be seen that the rate at which $R^{(t)}(\lambda)$ vanishes depends also on this cointegrating gap, specifically

$$R^{(t)}(\lambda) = O(t^{-\beta}) \quad \text{as } t \to \infty, \tag{19}$$

uniformly in λ. We believe that (15)–(19), which are implied by model (6), (7) under Assumption C, are the key intrinsic features of any cointegrating model involving Type II fractionally integrated processes. Finally, note that the real number ν has the meaning anticipated in (4), where ω_{ij} now denotes the (i,j)th element of the long-run variance-covariance matrix of the covariance stationary version of v_t (note that under Assumption NC is v_t itself).

For balanced cointegration we assume that $\theta_n = 0$ in (7), so that the stochastic trends driving y_t and x_t are strictly of the same order, but, as opposed to the case of absence of cointegration, there is a common trend of order δ driving the behavior of both processes. Weakly unbalanced cointegration will denote situations where

$$\theta_n \to 0 \quad \text{as } n \to \infty,$$

and the previously referred cotrending, which is fully characterized by the order of integration of the observables and the structure of v_t, is only approximate. Strictly speaking, the triangular array $x_{t,n}$ is $I(\delta + \theta_n)$, so that even if the behavior of v_t implies $CI(\delta, \beta)$ cointegration in the case $\theta_n = 0$, for fixed n, any linear combination of y_t and $x_{t,n}$ is $I(\delta)$ or $I(\delta + \theta_n)$ depending on whether $\theta_n < 0$ or $\theta_n > 0$, respectively. Note that in practice this situation could be indistinguishable from that of balanced cointegration, so that from the viewpoint of modeling purposes the applicability of the idea of weakly unbalanced cointegration is limited. However, this concept was not introduced with this objective but mainly to stress the fact that the effect of minor suspected differences in the integration orders of the observables could be negligible asymptotically. In fact, one of the points of the paper is to show that even if in finite samples y_t and $x_{t,n}$ are not cointegrated in the strict sense, the presence of a small perturbation (converging to 0 as n tends to infinity) in one of the integration orders in a bivariate system (which could be thought of as natural in a framework where the orders of integration are real numbers) may not affect first-order asymptotic properties of different estimates and statistics, whose finite-sample performance could display behavior closer to a $CI(\delta, \beta)$ situation than to the one predicted by the theory if the orders of integration of the observables

Table 1. Cointegration between y_t and $x_{t,n}$

$f(0)$	$\theta_n = 0$	$\theta_n \to 0$	$\theta_n = \theta \neq 0$
Singular	yes	weakly unbalanced	strongly unbalanced
Full rank	no	no	no

are different. As will be seen later, our asymptotic results and Monte Carlo experiment support this idea and also could help to explain the unexpected values (over what economic theory predicts) of some estimated parameters that appear in some empirical studies.

Strongly unbalanced cointegration will be characterized by (9), where $\theta \neq 0$ in (7). As mentioned in the Introduction, here y_t, x_t are not cointegrated, but y_t, $x_t(\theta)$ are, so that estimates taking into account the imbalance between the orders of integration of the observables could enjoy good asymptotic properties.

In short, we present the relevant cases in Table 1.

The study of the different situations will mainly focus on analyzing the asymptotic properties of a particular class of estimates of the fundamental parameter ν. To specify this class, we need some preliminary definitions. First, for a scalar or vector process ζ_t, let

$$w_\zeta(\lambda) = \frac{1}{\sqrt{2\pi n}} \sum_{t=1}^{n} \zeta_t e^{it\lambda}$$

be the discrete Fourier transform and given another sequence (possibly the same one) ξ_t, let

$$I_{\zeta\xi}(\lambda) = w_\zeta(\lambda) w'_\xi(-\lambda), \qquad I_\zeta(\lambda) = I_{\zeta\zeta}(\lambda)$$

be the cross-periodogram and periodogram, respectively. Thus, we will consider members of the class of estimates of ν given by

$$\bar{\nu}_a(m) = \frac{\operatorname{Re}\left\{ \sum_{j=0}^{m} c_j I_{x(a)y}(\lambda_j) \right\}}{\sum_{j=0}^{m} c_j I_{x(a)}(\lambda_j)}, \tag{20}$$

where $\lambda_j = 2\pi j/n$, $j = 0, \ldots, m$, are the Fourier frequencies, m is an integer such that $1 \leq m \leq n/2$, $c_j = 1, j = 0, n/2$, $c_j = 2$ otherwise, and a is a possibly random scalar. Note that because of the orthogonal properties of the complex exponential

$$\sum_{t=0}^{n-1} e^{it\lambda_j} = n, \qquad j = 0, \bmod n; \qquad = 0, \text{ otherwise},$$

https://doi.org/10.1017/S0266466606060361 Published online by Cambridge University Press

and by the symmetry of the real part of the (cross-) periodogram about $\lambda = 0, \pi$, for the particular choice $m = [n/2]$, where $[\cdot]$ denotes integer part,

$$\bar{\nu}_0([n/2]) = \frac{\sum_{t=1}^{n} x_{t,n} y_t}{\sum_{t=1}^{n} x_{t,n}^2},$$

which is the OLS estimate. In this case, the numerator of (20) is real, so our notation in (20) may be redundant. When $m < [n/2]$, we will only consider the case where

$$m^{-1} + m/n \to 0 \quad \text{as } n \to \infty, \tag{21}$$

so that under (21), (20) is the NBLS estimate. This is motivated by the fact that cointegration is a long-run phenomenon and hence neglecting components of the observable series associated with high (short-run) frequencies could improve the estimation. For cases where $\theta_n = \theta$ for all n, we will also consider estimates $\bar{\nu}_\theta(m)$ (note that for $\theta = 0$ this estimate is, depending on m, the OLS or NBLS) and $\bar{\nu}_{\hat{\theta}}(m)$, where $\hat{\theta}$ is a consistent estimate of θ such that the following regularity condition holds. Throughout K denotes a generic positive constant.

ASSUMPTION PE (Preliminary estimate). Provided $\theta_n = \theta$ for all n, there exists an estimate $\hat{\theta}$ of θ such that

$$|\hat{\theta}| \le K, \tag{22}$$

and for a real sequence g_n, such that

$$g_n \log^{-1} n \to \infty \quad \text{as } n \to \infty, \tag{23}$$

and random variable Λ,

$$g_n(\hat{\theta} - \theta) \to_d \Lambda. \tag{24}$$

The parameter θ reflects the difference between the integration orders of processes x_t and y_t, so that an estimate of θ will be naturally based on the difference of the estimates of their respective orders. As both processes y_t and x_t are observable, (24) with (23) is very mild, holding even if $\hat{\theta}$ is based on semiparametric estimates of the orders of x_t, y_t, such as the log-periodogram or Gaussian semiparametric, analyzed by Robinson (1995a, 1995b), respectively, and extended by Velasco (1999a, 1999b) to cover arbitrarily large but finite orders of integration. The only difficulty here is that these results do not apply directly to Type II processes, although in this case the bounds measuring the distance between Type I and Type II processes established by Robinson (2005) suffice. For parametric estimates of θ, $g_n = n^\alpha$ with $\alpha = \frac{1}{2}$ is achievable. The term Λ could be Gaussian or non-Gaussian, depending on the method of estimation of

θ, and also on various things, such as, for example, the degree of tapering applied in the estimation (see Velasco, 1999a, 1999b), a user-chosen number (see Lobato and Robinson, 1996), or even δ, θ (see Velasco and Robinson, 2000). Condition (22) is innocuous if $\hat{\theta}$ is based on estimates that optimize over bounded sets, this being standard for implicitly defined estimates.

The estimates $\bar{\nu}_\theta(m)$, $\bar{\nu}_{\hat{\theta}}(m)$ take explicitly into account the possible imbalance between the orders of integration of the series, but, in general, $\bar{\nu}_\theta(m)$ will be infeasible, because knowledge of the particular integration orders of the observables in empirical situations is difficult to justify. On the contrary, $\bar{\nu}_{\hat{\theta}}(m)$ does not require this knowledge, and in view of Hualde and Robinson (2006) and Robinson and Hualde (2003), one could suspect that provided $\hat{\theta}$ converges fast enough to θ, $\bar{\nu}_{\hat{\theta}}(m)$ could enjoy the same rate of convergence and limiting distribution of $\bar{\nu}_\theta(m)$. The fact that generally θ is unknown could be a strong point in favor of the feasible estimate, even if we suspect that $\theta = 0$, although, as shown in the next section, the insertion of $\hat{\theta}$ could importantly affect first-order asymptotic properties of the estimates.

Finally, note that we could have equally considered the class of estimates

$$\tilde{\nu}_a(m) = \frac{\operatorname{Re}\left\{ \sum_{j=0}^{m} c_j I_{xy(a)}(\lambda_j) \right\}}{\sum_{j=0}^{m} c_j I_x(\lambda_j)},$$

with $a = -\theta, -\hat{\theta}$, which in certain cases could be superior to $\bar{\nu}_a(m)$. More precisely, under Assumption C, y_t and $x_t(\theta)$ are $CI(\delta,\gamma)$ cointegrated, whereas, noting that under this condition

$$y_t(-\theta) = \nu x_t + u_{1t}(-\gamma - \theta),$$

$y_t(-\theta)$ and x_t are $CI(\delta+\theta, \gamma+\theta)$ cointegrated. This implies that when $\theta > 0$, $\tilde{\nu}_a(m)$ could enjoy a faster convergence rate than $\bar{\nu}_a(m)$, the opposite happening when $\theta < 0$ (see, e.g., Robinson and Marinucci, 2001). For simplicity, we only consider $\bar{\nu}_a(m)$, although the treatment of $\tilde{\nu}_a(m)$ could be addressed in a similar way.

3. MAIN ASYMPTOTIC RESULTS

First, we concentrate on the situation of balanced orders, with

$$\theta_n = 0, \quad \text{for all } n. \tag{25}$$

We collect results corresponding to this case in two theorems that cover the noncointegrating and cointegrating situations, respectively. Most of these results are well known in the literature and are presented just for completeness.

Denote by $W(r)$ the 2×1 vector Brownian motion with covariance matrix I_2 and define the (Type II; see Marinucci and Robinson, 1999) fractional Brownian motion

$$W(r;\delta) = \int_0^r \frac{(r-s)^{\delta-1}}{\Gamma(\delta)}\, dW(s)$$

and the 2×1 vectors

$$\zeta = (1,0)', \qquad \xi = (0,1)'.$$

By $\Rightarrow$ we will mean convergence in the Skorohod space $D[0,1]$ endowed with the J_1 topology and related to the different analyzed estimates denote

$$\bar{\nu}_1 = \bar{\nu}_0([n/2]), \qquad \bar{\nu}_2 = \bar{\nu}_0(m), \qquad \bar{\nu}_3 = \bar{\nu}_\theta([n/2]),$$

$$\bar{\nu}_4 = \bar{\nu}_\theta(m), \qquad \bar{\nu}_5 = \bar{\nu}_{\hat{\theta}}([n/2]), \qquad \bar{\nu}_6 = \bar{\nu}_{\hat{\theta}}(m),$$

under (21), noting that when $\theta = 0$, $\bar{\nu}_1 = \bar{\nu}_3$, $\bar{\nu}_2 = \bar{\nu}_4$. The expressions $\bar{\nu}_i$, $i = 1,2$ and $i = 3,\dots,6$ will be referred to as undifferenced-regressor (U) and differenced-regressor (D) estimates, respectively, noting the previous caution about the equality of the different types of estimates when $\theta = 0$. The main results will be given in a sequence of theorems linked to different cointegrating situations and estimates. For most of the theorems the full characterization of the limiting distributions and rates of convergence is given in Appendix A.

THEOREM BNCUD (Balanced orders, no cointegration, U and D estimates). *Let (6)–(8), (21), (25), and Assumptions NC and PE hold. Then, as $n \to \infty$,*

$$\bar{\nu}_1 \Rightarrow \frac{\xi' A(1)\Sigma^{1/2} \int_0^1 W(r;\delta)W'(r;\delta)\, dr \Sigma^{1/2} A'(1)\zeta}{\xi' A(1)\Sigma^{1/2} \int_0^1 W(r;\delta)W'(r;\delta)\, dr \Sigma^{1/2} A'(1)\xi}, \tag{26}$$

$$\bar{\nu}_1 - \bar{\nu}_2 = O_p(m^{1-2\min\{\delta,1\}}), \tag{27}$$

$$\bar{\nu}_5 - \bar{\nu}_1 = O_p(g_n^{-1} \log n), \tag{28}$$

$$\bar{\nu}_6 - \bar{\nu}_2 = O_p(g_n^{-1} \log n). \tag{29}$$

Theorem BNCUD refers to the situation of spurious cointegration. We omit the proof of (26) and (27) because the first result is a straightforward application of Theorem 1 of Marinucci and Robinson (2000) and the continuous mapping theorem, whereas (27) is implied by Lemma 1 of Marmol and Velasco

(2004). The proof of (28) and (29) is given in Appendix B. Note that as $A(1)\Sigma A'(1) = LL'$, where

$$L = \begin{pmatrix} \omega_{11}^{1/2}(1-\rho^2)^{1/2} & \omega_{12}\omega_{22}^{-1/2} \\ 0 & \omega_{22}^{1/2} \end{pmatrix},$$

from the Cholesky decomposition. We could equally represent the limiting distribution of $\bar{\nu}_1$ by (see Marmol and Velasco, 2004)

$$\bar{\nu}_1 \Rightarrow \nu + \frac{\omega_{11}^{1/2}(1-\rho^2)^{1/2}\int_0^1 \xi' W(r;\delta)W'(r;\delta)\zeta dr}{\omega_{22}^{1/2}\int_0^1 \xi' W(r;\delta)W'(r;\delta)\xi dr}. \tag{30}$$

Note also that (27)–(29) under Assumption PE imply that all estimates enjoy the same limiting distribution, so these estimates are not consistent for the fundamental parameter ν. As is well known, this inconsistency turns into consistency with reduced rank of Ω, as is readily seen from (30) when $\rho^2 = 1$. However, this information is not sufficient to derive the limiting distribution of the estimates, this being precisely the reason why the characterization in Assumption C(i) comes into play. Defining

$$\Upsilon(\delta) = \xi' B(1)\Psi^{1/2}\int_0^1 W(r;\delta)W'(r;\delta)\, dr \Psi^{1/2}B'(1)\xi,$$

the next theorem, given just for completeness, collects results given in Robinson and Marinucci (2001) for OLS and NBLS estimates.

THEOREM BCU (Balanced orders, cointegration, U estimates). *Let (6)–(8), (21), (25), and Assumption C hold. Then, as $n \to \infty$,*

$$p_i(n)(\bar{\nu}_i - \nu) \Rightarrow \Xi_i(\gamma,\delta)/\Upsilon(\delta), \qquad i = 1,2.$$

See Appendix A for full characterization of $p_i(n)$, $\Xi_i(\gamma,\delta)$, $i = 1,2$. As in Robinson and Marinucci (2001), we were unable to characterize the precise limiting distribution of our estimates in the case $\delta + \gamma > 1$, $0 < \gamma \leq \frac{1}{2}$. It is also important to note that Assumption C is sufficient for the set of conditions needed for the different results given in Robinson and Marinucci (2001) to hold. In particular, Assumption C(ii) is sufficient for the cumulant spectral density related conditions and also implies square integrability of the individual spectra of u_{1t}, v_{2t}, and fourth-order stationarity of u_t. We do not consider in this theorem the behavior of $\bar{\nu}_5$, $\bar{\nu}_6$, because this is covered by the more general Theorem UCD, given subsequently, which assumes $\theta_n = \theta$, where θ is not necessarily 0.

Now, we turn to the weakly unbalanced cointegration case, where $\theta_n \to 0$ as $n \to \infty$. Although it is not needed for our proofs, it is convenient to visualize θ_n

as a monotonic sequence, and to have a neater interpretation of the asymptotically negligible departures from the equality of the integration orders, we specify further the θ_n sequence, setting

$$\theta_n = \frac{\eta}{h_n}, \tag{31}$$

where η is a nonzero real finite number and h_n is a positive sequence such that

$$h_n \log^{-1} n \to \infty \text{ as } \quad n \to \infty. \tag{32}$$

Note that η reflects the direction of the asymptotically negligible perturbation of the order of integration of $x_{t,n}$, as depending on whether $\eta > 0$ or $\eta < 0$, θ_n tends to 0 from above or below, respectively. We only consider the infeasible estimates $\bar{\nu}_i$, $i = 1,2$, hence avoiding the difficulty associated with the fact that now $\hat{\theta}$ would estimate not a fixed parameter θ but a sequence of parameters, whereas it does not seem realistic to assume knowledge of θ_n. In any case, the main point of the theorem that follows is that in a situation of standard fractional cointegration, a small perturbation of the order of integration of one of the observables could be asymptotically negligible.

THEOREM ABCU (Asymptotically balanced orders, cointegration, U estimates). *Let (6)–(8), (21), (31), (32), and Assumption C hold. Then, as $n \to \infty$,*

$$q_i(n)(\bar{\nu}_i - \nu) \Rightarrow F_i(\gamma, \delta, \nu, \eta), \qquad i = 1,2.$$

See Appendix A for full characterization of $q_i(n)$, $F_i(\gamma, \delta, \nu, \eta)$, $i = 1,2$, and Appendix B for the formal derivations of those results. The interpretation of this theorem is straightforward. The small perturbation in the order of integration of $x_{t,n}$ produces some additional terms in the expansion of $\bar{\nu}_i - \nu$. Only one of these terms is competitive with the difference $\bar{\nu}_i - \nu$ when $\theta_n = 0$. Thus, the speed at which θ_n tends to 0 determines the relative importance of these two terms. Heuristically, on the one hand, if θ_n converges relatively fast to 0, the presence of the small perturbation θ_n in the order of $x_{t,n}$ does not affect the limiting distribution of the estimates with respect to the situation where $\theta_n = 0$. On the other hand, if θ_n converges slowly, the term previously referred to arising by the presence of the perturbation dominates, and the estimates have a degenerate limiting distribution. Apart from these situations, there is an exact rate for θ_n at which those two terms are balanced, so that contributions from both appear in the limiting distributions of the estimates. In this case, the effect on the situation where $\theta_n = 0$ is simply to shift the asymptotic distribution by $-\nu\eta$, whereas the rates of convergence remain unchanged. Finally, note that (32) is a consistency condition, and as long as it holds, the OLS and NBLS will be consistent estimates with possibly very slow rates of convergence.

Finally, we discuss the situation where (9) holds, so in the case $\theta \neq 0$, the integration orders of the observables are truly different. The first theorem devoted to this situation collects results, mainly based on Robinson and Marinucci (2001), for OLS and NBLS estimates. As we might guess, because of the imbalance between the orders, these estimates are not consistent for ν. We derive all results under Assumption NC, although we did not actually use condition $\rho^2 < 1$, and the results given subsequently are also valid in the case $\rho^2 = 1$, which would denote a situation of cointegration. It can be seen easily that if $\rho^2 = 1$, some terms in the asymptotic distributions given later could disappear, although this is not always the case. For the cases where it is possible to fully characterize the limiting distributions, we also give in Appendix A an alternative representation of them highlighting this fact and also stressing the dependence of these distributions on the fundamental parameter ν.

THEOREM UU (Unbalanced orders, U estimates). *Let (6)–(9) and (21) with $\theta \neq 0$ and Assumption NC (without the requirement $\rho^2 < 1$) hold. Then, as $n \to \infty$,*

$$r_i(n)\bar{\nu}_i \Rightarrow \frac{\Theta_i(\theta,\delta)}{\Pi_i(\delta+\theta)}, \qquad i = 1,2.$$

See Appendix A for full characterization of $r_i(n)$, $\Theta_i(\theta,\delta)$, $\Pi_i(\delta+\theta)$, $i = 1,2$. These results are mainly taken from Robinson and Marinucci (2001), but we indicate and justify in Appendix B the exact steps that do not follow directly from this reference. As expected, U estimates are not consistent for ν because of the imbalance between the orders of integration, so, in most cases, $\bar{\nu}_i$, $i = 1,2$, converge to zero or infinity depending on the different region of the (θ,δ) space on which we focus.

Finally, we propose a sensible solution to the problem of unbalanced series based on the previously defined D estimates, for which the following theorem holds.

THEOREM UCD (Unbalanced orders, cointegration, D estimates). *Let (6)–(9), (21), and Assumptions C and PE hold. Then, as $n \to \infty$,*

$$s_i(n)(\bar{\nu}_i - \nu) \Rightarrow G_i(\gamma,\delta,\nu,\Lambda), \qquad i = 3,\dots,6.$$

See Appendix A for full characterization of $s_i(n)$, $G_i(\gamma,\delta,\nu,\Lambda)$, $i = 3,\dots,6$. We omit the proof of this theorem because it is almost identical to that of Theorem ABCU. Trivially, results for $\bar{\nu}_i$, $i = 3,4$, are identical to those of Theorem BCU, because taking θ differences on the regressor simply balances the series to turn the situation into one of traditional cointegration, $\bar{\nu}_i$, $i = 3,4$, behaving accordingly. Results for $\bar{\nu}_i$, $i = 5,6$, are similar to those of Theorem ABCU, and they also have a straightforward interpretation. The feasible filtering of the regressor ($x_t(\hat{\theta})$ instead of $x_t(\theta)$) makes some additional terms appear in the expansion of $\bar{\nu}_i - \nu$, $i = 5,6$. One of these terms is in all cases $O_p(g_n^{-1} \log n)$, so that unless this term converges faster to 0 than the usual term in $\bar{\nu}_i - \nu$, $i = 3,4$, it dominates, obtaining for the feasible estimates in all cases the limiting distribution

$$g_n \log^{-1} n(\bar{\nu}_i - \nu) \to_d \nu\Lambda, \qquad i = 5,6, \tag{33}$$

which is simply the asymptotic distribution of the estimate of θ, with a slightly different convergence rate (because of the $\log^{-1} n$ factor), and premultiplied by the unknown parameter ν. Note that if Λ is distributed as a zero-mean Gaussian random variable, the value of ν affects the limiting variance of the feasible estimates. In any case, it is worth stressing that as long as (23) holds, the feasible estimates are always consistent. Note finally that if θ is estimated semiparametrically, we expect g_n to be relatively slow, this being transmitted to the rate of convergence of the feasible estimates of ν. In the next section we will propose alternative estimates of ν that do not suffer from the serious drawback of having slow rates of convergence. On the contrary, they achieve optimal rates.

4. ALTERNATIVE FEASIBLE ESTIMATES UNDER COINTEGRATION

In this section, we consider the model (6), (7) with (9) and work under Assumption C. Noting that the results presented in Theorem UCD may be unsatisfactory in some cases because of the distortive effect of the insertion of $\hat{\theta}$ instead of the true imbalance parameter θ, in the spirit of Hualde and Robinson (2006) and Robinson and Hualde (2003), we propose similar estimates to theirs, considering the possible imbalance of the integration orders of the observables. As will become clear subsequently, these estimates enjoy the same optimal asymptotic properties of the ones previously referred to, namely, optimal rates of convergence and standard limiting distribution, implying, among other consequences, that straightforward inference on the value of ν is readily available.

First, we discuss the case where $\beta > \frac{1}{2}$ (termed strong fractional cointegration by Hualde and Robinson, 2004a). For the sake of a clear exposition we consider initially the situation where $B(L) = I_2$ in (12), so that u_t is a white noise. Denoting by ψ_{ij} the (i,j)th element of Ψ and by ϵ_{it}, $i = 1,2$, the ith element of ϵ_t, assuming γ, δ, θ are known real numbers, the pseudo maximum likelihood estimate of ν is identical to the OLS estimate of ν in the equation

$$y_t(\gamma) = \nu x_t(\gamma + \theta) + \tau x_t(\delta + \theta) + \epsilon_{1.2t}, \tag{34}$$

where

$$\tau = \psi_{12}/\psi_{22}, \qquad \epsilon_{1.2t} = \epsilon_{1t} - \tau\epsilon_{2t}$$

(see, e.g., Phillips, 1991, in which this result is derived for the case $\gamma = \theta = 0$, $\delta = 1$). Robinson and Hualde (2003) showed that under mild regularity conditions this estimate is n^{β}-consistent with mixed-normal limiting distribution. However, in practice it is unrealistic to assume knowledge of γ, δ, and/or θ, and therefore our proposed estimate is in general infeasible. Fortunately, the results of Robinson and Hualde (2003) imply that as long as the estimates of the orders, say, $\hat{\gamma}$, $\hat{\delta}$, $\hat{\theta}$, are n^{κ}-consistent with

$$\kappa > \max(0, 1-\beta), \tag{35}$$

estimating ν from the OLS regression of $y_t(\hat{\gamma})$ on $x_t(\hat{\gamma}+\hat{\theta})$, $x_t(\hat{\delta}+\hat{\theta})$ is asymptotically equivalent (to first-order properties) to estimating ν from (34). As in Robinson and Hualde (2003) the main difficulty here is to find estimates of the orders for which (35) holds, because almost $\sqrt{n}$-consistency could be required. Our present framework adds more difficulties because of the unknown nature of θ, and whereas δ and $\delta+\theta$ can be estimated easily from y_t and x_t, respectively, to estimate γ from residuals, preliminary estimates of the parameters explaining the long-run linkage between the series, namely, ν, θ, are needed, as opposed to the Robinson and Hualde (2003) situation, where only preliminary estimation of ν was required (although $\theta = 0$ was assumed to be known). Thus, although showing rigorously the properties of the estimates of the orders goes beyond the scope of the present paper, we propose several sensible estimation procedures that are relatively simple to implement. The simplest one is to base a semiparametric procedure on the residuals $y_t - \bar{\nu}_{\hat{\theta}}(m)x_t(\hat{\theta})$, where $\hat{\theta}$ is the difference between semiparametric estimates of the integration orders of x_t and y_t. Provided $\beta \geq 1$, this is a valid strategy, but for $\beta < 1$ this procedure does not ensure obtaining estimates of γ with the required rates of convergence. However, we could easily improve upon this method by exploiting parametric assumptions on u_t. Keeping the discussion within the white noise framework, $\delta+\theta$ is $\sqrt{n}$-consistently estimable from x_t by various methods including

$$\widehat{\delta+\theta} = \underset{d \in \mathcal{D}}{\arg\min} \sum_{t=1}^{n} x_t^2(d), \tag{36}$$

where $\mathcal{D}$ is a compact set such that $\delta+\theta \in \mathcal{D}$ (see Hualde and Robinson, 2004b). Similarly, we could estimate γ and θ simultaneously by

$$\hat{\gamma}, \hat{\theta} = \underset{c \in \mathcal{C}, a \in \mathcal{A}}{\arg\min} \sum_{t=1}^{n} (y_t(c) - \bar{\nu}_a(m)x_t(c+a))^2, \tag{37}$$

for an appropriate choice of m, where $\mathcal{C}$, $\mathcal{A}$ are compact sets such that $\gamma \in \mathcal{C}$, $\theta \in \mathcal{A}$. The exact properties of $\hat{\gamma}$ and $\hat{\theta}$ could be difficult to justify, but in view of Hualde and Robinson (2004b) and (33), our guess is that the rate $\sqrt{n}/\log n$ is achievable for both estimates, and hence (35) is satisfied for any $\beta > \frac{1}{2}$.

In the general $I(0)$ case, we can adapt the estimation procedure proposed in Robinson and Hualde (2003) to our present framework. In particular, if u_t is $I(0)$ with spectral density $h(\lambda)$ (see (14)) depending on a vector of short-memory parameters φ, such that

$$h(\lambda;\varphi) = h(\lambda), \qquad B(L;\varphi) = B(L), \qquad \Psi(\varphi) = \Psi,$$

we could define

$$\tilde{\nu}(c,d,e,k) = \frac{\tilde{a}(c,d,e,k)}{\tilde{b}(c+e,k)}, \qquad \hat{\nu}(c,d,e,k) = \frac{a(c,d,e,k)}{b(c+e,k)},$$

where setting $z_t(c,d) = (y_t(c), x_t(d))'$,

$$\tilde{a}(c,d,e,k) = \sum_{t=1}^{n} \{B(L;k)\zeta x_t(c+e)\}'\Psi(k)^{-1}\{B(L;k)z_t(c,d+e)\},$$

$$\tilde{b}(c,k) = \sum_{t=1}^{n} \{B(L;k)\zeta x_t(c)\}'\Psi(k)^{-1}\{B(L;k)\zeta x_t(c)\},$$

and

$$p(\lambda;k) = \zeta' h(\lambda;k)^{-1}, \qquad q(\lambda;k) = \zeta' h(\lambda;k)^{-1}\zeta,$$

$$a(c,d,e,k) = \sum_{j=1}^{n} p(\lambda_j;k) w_{x(c+e)}(-\lambda_j) w_{z(c,d+e)}(\lambda_j),$$

$$b(c,k) = \sum_{j=1}^{n} q(\lambda_j;k) |w_{x(c)}(\lambda_j)|^2.$$

Thus, for an estimate of the short-memory parameters $\hat{\varphi}$ and estimates of the orders $\hat{\gamma}$, $\hat{\delta}$, $\hat{\theta}$, we propose

$$\nu^*(\gamma,\delta,\theta,\varphi),\ \nu^*(\gamma,\delta,\theta,\hat{\varphi}),\ \nu^*(\hat{\gamma},\delta,\theta,\hat{\varphi}),$$
$$\nu^*(\gamma,\hat{\delta},\theta,\hat{\varphi}),\ \nu^*(\hat{\gamma},\hat{\delta},\theta,\hat{\varphi}),\ \nu^*(\hat{\gamma},\hat{\delta},\hat{\theta},\hat{\varphi}), \tag{38}$$

where ν^* denotes $\tilde{\nu}$ or $\hat{\nu}$. Each estimate in (38) reflects situations of different knowledge about the structure of the model. Note that the first four estimates with $\theta = 0$ are identical to those presented in Robinson and Hualde (2003). Again, the problem is obtaining $\hat{\gamma}$, $\hat{\delta}$, $\hat{\theta}$, $\hat{\varphi}$ with the required properties, but procedures incorporating the ideas developed in the white noise situation and those in Velasco and Robinson (2000), hence employing as loss functions possibly tapered parametric Whittle likelihoods, should provide estimates with rate of convergence $\sqrt{n}/\log n$. In any case, in view of Hualde and Robinson (2004a), even if only n^{κ}-consistent estimates of the nuisance parameters with arbitrarily small but positive κ are available, narrow band versions of the frequency domain estimates in (38) could also enjoy optimal asymptotic properties for an adequate choice of the bandwidth, semiparametric extensions also being possible.

The situation where $\beta < \frac{1}{2}$ was referred to by Hualde and Robinson (2006) as weak fractional cointegration. In view of the results of this paper, if u_t is a white noise and γ, δ, and θ are known, the OLS in (34) would produce $\sqrt{n}$-consistent and asymptotically normal estimates of ν, this result being easily extendable to the situation where u_t is an autoregressive process of finite order p, so that

$$B(L) = \sum_{j=0}^{p} B_j L^j \tag{39}$$

in (12). In the case where the orders are unknown, additional difficulties arise because of the presence of the unknown parameter θ, but as in the case of strong cointegration, and following the lines of Hualde and Robinson (2006), it seems possible to deal with this issue, although $\sqrt{n}$-consistent estimates of the orders are needed to estimate ν $\sqrt{n}$-consistently.

If u_t is a white noise, we could improve upon the method proposed in (37) by first estimating $\delta + \theta$ by (36) and then γ and $\gamma + \theta$ by

$$\hat{\gamma}, \widehat{\gamma + \theta} = \arg\min_{c \in \mathcal{C}, b \in \mathcal{B}} \sum_{t=1}^{n} (y_t(c) - \bar{\nu}(c, b, \widehat{\delta + \theta}) x_t(b) - \bar{\tau}(c, b, \widehat{\delta + \theta}) x_t(\widehat{\delta + \theta}))^2, \tag{40}$$

where $\bar{\nu}(c,b,d)$, $\bar{\tau}(c,b,d)$ are the estimated slopes corresponding to regressors $x_t(b)$, $x_t(d)$, respectively, in the OLS regression of $y_t(c)$ on $x_t(b)$ and $x_t(d)$. In this case the $\sqrt{n}$-consistent estimate of ν is $\bar{\nu}(\hat{\gamma}, \widehat{\gamma + \theta}, \widehat{\delta + \theta})$. As in Hualde and Robinson (2006), this procedure can be easily extended to the case where in (39) B_j is upper-triangular for all $j = 1, \ldots, p$, noting that this is also a valid strategy when $\beta > \frac{1}{2}$, in which case our proposed estimate of ν is n^{β}-consistent with mixed-normal limiting distribution and the estimates of the orders are also $\sqrt{n}$-consistent.

5. MONTE CARLO EVIDENCE

With the aim of analyzing the finite-sample equivalent of the asymptotic phenomena under wide-sense cointegration described in the paper, we carried out a small Monte Carlo experiment. We generated u_t in (12) as a white noise process of dimensions $n = 64$, 128, 256, with $\psi_{11} = \psi_{22} = 1$, varying the correlation ψ_{12} (taking values 0, 0.5, -0.5). Then, we considered (6), (7), and (11) with $\nu = 1$ and employed six combinations of (γ, δ) given by

$(\gamma, \delta) = (0, 0.6), (0, 1.2), (0, 2), (0.4, 0.8), (0.4, 1.2), (0.7, 1),$

noting that the fourth and sixth cases correspond to weak cointegration, whereas the rest represent different situations of strong cointegration. In the first part of the study, we analyze the case of weakly unbalanced cointegration, where θ_n in (7) takes four different values $\theta_n = \theta_n^{(i)}$, $i = 1,\ldots,4$, where

$$\theta_n^{(1)} = \log^{-1} n, \qquad \theta_n^{(2)} = -n^{-1/2}, \qquad \theta_n^{(3)} = n^{-1}, \qquad \theta_n^{(4)} = 0,$$

for which we study in terms of Monte Carlo bias (defined as the estimate minus ν) and standard deviation over 1,000 replications the behavior of $\bar{\nu}_1$, $\bar{\nu}_2$, noting that by Theorems ABCU and BCU, these estimates are consistent for all choices of θ_n except $\theta_n^{(1)}$. Results are given in Tables 2–4 (for the bias in Tables 2 and 4; for standard deviations in Tables 3 and 5). Clearly, in almost all cases bias decreases as θ_n decreases in absolute value, the sign of θ_n being inversely related to the sign of the bias because of the relative dominance of the denominator (numerator) of the estimates when θ_n is positive (negative). Although the inconsistency of $\bar{\nu}_1$, $\bar{\nu}_2$ when $\theta_n = \theta_n^{(1)}$ is evident by looking at the evolution of the bias, the decrease of the bias as n increases is very slow for $\theta_n = \theta_n^{(2)}$. Biases for $\theta_n = \theta_n^{(3)}$ are larger than when the orders of integration of the observables are strictly balanced, that is, when $\theta_n = 0$, but not far from them and reacting in a similar way when n increases. As expected, for $\theta_n = \theta_n^{(i)}$, $i = 3,4$, bias corresponding to both estimates tends to decrease as β increases, this not being the case when $\theta_n = \theta_n^{(2)}$ (most evident for $\rho = -0.5$). In fact, the very small bias for $(\gamma, \delta) = (0, 0.6)$, $(0.4, 0.8)$, $(0.7, 1)$ for $\rho = -0.5$ is the most surprising result of this part of the Monte Carlo experiment, some cancellations probably taking place, although in this case, at least when $\beta < \frac{1}{2}$, no clear evidence of decreases in bias as n increases is observed. For $\theta_n = \theta_n^{(i)}$, $i = 3,4$, bias tends to increase as $|\rho|$ increases, this also being true for $\theta_n = \theta_n^{(2)}$ when $\rho = 0.5$ (but not when $\rho = -0.5$). Finally, whereas for $\rho \neq 0$ and $\theta_n = \theta_n^{(i)}$, $i = 3,4$, $\bar{\nu}_2$ clearly beats $\bar{\nu}_1$ (most noticeably for $(\gamma, \delta) = (0, 0.6)$), this is not the case for $\theta_n = \theta_n^{(2)}$, especially when $\rho = -0.5$, when $\bar{\nu}_1$ dominates $\bar{\nu}_2$ (especially for small β).

Standard deviations are reported in Tables 3 and 5. In almost all cases they decrease as n increases, including the situation $\theta_n = \theta_n^{(1)}$, which suggests that the inconsistency of the estimates due to the (slowly converging to zero) imbalance between the orders is mainly due to a bias problem. In fact, in many cases, the smallest variances correspond to this inconsistent case. In general, standard deviations tend to decrease as β increases, the effect of changes in ρ not being very clear. Values of standard deviations for the two estimates are very similar, although some superiority of $\bar{\nu}_1$ when $\beta < \frac{1}{2}$ is noted, whereas for $(\gamma, \delta) = (0, 0.6)$, in general $\bar{\nu}_2$ beats $\bar{\nu}_1$.

The second part of the Monte Carlo experiment focuses on the strongly unbalanced cointegration situation. Here, the only difference with respect to the previous analysis is that we consider

$$\theta_n = \theta = 0.3, -0.3, \qquad \text{for all } n.$$

TABLE 2. Monte Carlo bias of $\bar{\nu}_1$

ρ	γ	n / δ	64 $\theta_n^{(1)}$	64 $\theta_n^{(2)}$	64 $\theta_n^{(3)}$	64 $\theta_n^{(4)}$	128 $\theta_n^{(1)}$	128 $\theta_n^{(2)}$	128 $\theta_n^{(3)}$	128 $\theta_n^{(4)}$	256 $\theta_n^{(1)}$	256 $\theta_n^{(2)}$	256 $\theta_n^{(3)}$	256 $\theta_n^{(4)}$
0	0	0.6	−0.454	0.195	−0.035	−0.005	−0.467	0.193	−0.020	−0.002	−0.479	0.173	−0.011	0.000
	0	1.2	−0.532	0.431	−0.047	−0.001	−0.544	0.374	−0.029	0.000	−0.556	0.311	−0.017	0.000
	0	2	−0.527	0.468	−0.047	0.000	−0.539	0.390	−0.029	0.000	−0.551	0.319	−0.017	0.000
	0.4	0.8	−0.509	0.293	−0.053	−0.015	−0.520	0.279	−0.033	−0.009	−0.531	0.250	−0.016	−0.002
	0.4	1.2	−0.534	0.424	−0.052	−0.007	−0.545	0.370	−0.031	−0.003	−0.557	0.310	−0.018	−0.001
	0.7	1	−0.540	0.350	−0.073	−0.031	−0.548	0.318	−0.049	−0.023	−0.553	0.288	−0.021	−0.005
0.5	0	0.6	−0.379	0.483	0.154	0.194	−0.411	0.423	0.136	0.160	−0.438	0.356	0.119	0.133
	0	1.2	−0.529	0.449	−0.039	0.007	−0.543	0.380	−0.026	0.002	−0.557	0.314	−0.016	0.001
	0	2	−0.526	0.466	−0.048	−0.001	−0.537	0.388	−0.029	0.000	−0.551	0.318	−0.017	0.000
	0.4	0.8	−0.421	0.592	0.144	0.192	−0.454	0.509	0.126	0.155	−0.486	0.416	0.103	0.120
	0.4	1.2	−0.518	0.497	−0.010	0.038	−0.536	0.409	−0.007	0.022	−0.553	0.329	−0.006	0.012
	0.7	1	−0.429	0.702	0.160	0.214	−0.459	0.602	0.149	0.182	−0.492	0.485	0.124	0.143
−0.5	0	0.6	−0.526	−0.075	−0.211	−0.191	−0.526	−0.026	−0.171	−0.157	−0.525	−0.006	−0.141	−0.133
	0	1.2	−0.534	0.424	−0.053	−0.007	−0.547	0.375	−0.031	−0.002	−0.557	0.312	−0.018	−0.001
	0	2	−0.524	0.470	−0.047	0.001	−0.539	0.393	−0.029	0.000	−0.552	0.320	−0.017	0.000
	0.4	0.8	−0.582	0.034	−0.222	−0.193	−0.579	0.080	−0.171	−0.151	−0.576	0.090	−0.132	−0.120
	0.4	1.2	−0.547	0.375	−0.082	−0.038	−0.554	0.347	−0.049	−0.021	−0.561	0.296	−0.028	−0.011
	0.7	1	−0.621	0.073	−0.253	−0.220	−0.614	0.105	−0.198	−0.176	−0.608	0.104	−0.156	−0.142

Note: $\bar{\nu}_1 = \bar{\nu}_0([n/2])$, $\theta_n^{(1)} = \log^{-1} n$, $\theta_n^{(2)} = -n^{-1/2}$, $\theta_n^{(3)} = n^{-1}$, $\theta_n^{(4)} = 0$.

TABLE 3. Monte Carlo standard deviation of $\bar{\nu}_1$

ρ	γ	n / δ	64 $\theta_n^{(1)}$	64 $\theta_n^{(2)}$	64 $\theta_n^{(3)}$	64 $\theta_n^{(4)}$	128 $\theta_n^{(1)}$	128 $\theta_n^{(2)}$	128 $\theta_n^{(3)}$	128 $\theta_n^{(4)}$	256 $\theta_n^{(1)}$	256 $\theta_n^{(2)}$	256 $\theta_n^{(3)}$	256 $\theta_n^{(4)}$
0	0	0.6	0.106	0.151	0.084	0.086	0.091	0.106	0.051	0.052	0.078	0.072	0.033	0.033
	0	1.2	0.056	0.113	0.025	0.025	0.044	0.071	0.011	0.010	0.040	0.047	0.005	0.005
	0	2	0.044	0.076	0.007	0.003	0.037	0.048	0.003	0.001	0.032	0.033	0.002	0.000
	0.4	0.8	0.130	0.242	0.167	0.171	0.101	0.181	0.126	0.128	0.077	0.128	0.092	0.093
	0.4	1.2	0.070	0.151	0.076	0.078	0.050	0.092	0.045	0.046	0.041	0.057	0.024	0.025
	0.7	1	0.202	0.428	0.312	0.322	0.165	0.358	0.273	0.278	0.121	0.269	0.211	0.214
0.5	0	0.6	0.143	0.101	0.106	0.100	0.117	0.062	0.079	0.075	0.099	0.038	0.059	0.057
	0	1.2	0.056	0.098	0.025	0.022	0.045	0.067	0.010	0.009	0.037	0.044	0.004	0.004
	0	2	0.045	0.073	0.006	0.003	0.038	0.050	0.003	0.001	0.030	0.031	0.001	0.000
	0.4	0.8	0.142	0.196	0.160	0.160	0.112	0.145	0.127	0.127	0.088	0.101	0.093	0.092
	0.4	1.2	0.072	0.118	0.067	0.067	0.051	0.077	0.040	0.040	0.040	0.047	0.022	0.022
	0.7	1	0.185	0.372	0.275	0.283	0.150	0.311	0.243	0.247	0.115	0.235	0.190	0.192
−0.5	0	0.6	0.071	0.189	0.088	0.097	0.059	0.152	0.069	0.074	0.054	0.119	0.055	0.058
	0	1.2	0.053	0.117	0.020	0.022	0.043	0.072	0.009	0.010	0.039	0.047	0.004	0.004
	0	2	0.045	0.072	0.007	0.003	0.036	0.047	0.003	0.001	0.030	0.031	0.001	0.000
	0.4	0.8	0.099	0.244	0.147	0.154	0.078	0.193	0.118	0.122	0.062	0.145	0.090	0.092
	0.4	1.2	0.057	0.154	0.062	0.066	0.044	0.096	0.039	0.040	0.039	0.060	0.022	0.022
	0.7	1	0.155	0.374	0.260	0.270	0.132	0.316	0.231	0.237	0.105	0.244	0.186	0.188

Note: $\bar{\nu}_1 = \bar{\nu}_0([n/2])$; $\theta_n^{(1)} = \log^{-1} n$, $\theta_n^{(2)} = -n^{-1/2}$, $\theta_n^{(3)} = n^{-1}$, $\theta_n^{(4)} = 0$.

Table 4. Monte Carlo bias of $\bar{\nu}_2$

ρ	γ	n / δ	64 $\theta_n^{(1)}$	64 $\theta_n^{(2)}$	64 $\theta_n^{(3)}$	64 $\theta_n^{(4)}$	128 $\theta_n^{(1)}$	128 $\theta_n^{(2)}$	128 $\theta_n^{(3)}$	128 $\theta_n^{(4)}$	256 $\theta_n^{(1)}$	256 $\theta_n^{(2)}$	256 $\theta_n^{(3)}$	256 $\theta_n^{(4)}$
0	0	0.6	−0.484	0.278	−0.040	−0.005	−0.490	0.248	−0.023	−0.002	−0.497	0.209	−0.013	0.000
	0	1.2	−0.533	0.438	−0.047	−0.001	−0.544	0.376	−0.029	0.000	−0.556	0.311	−0.017	0.000
	0	2	−0.527	0.468	−0.047	0.000	−0.539	0.390	−0.029	0.000	−0.551	0.319	−0.017	0.000
	0.4	0.8	−0.519	0.336	−0.056	−0.015	−0.527	0.302	−0.035	−0.009	−0.535	0.262	−0.017	−0.002
	0.4	1.2	−0.535	0.431	−0.052	−0.007	−0.545	0.372	−0.032	−0.003	−0.557	0.310	−0.018	−0.001
	0.7	1	−0.543	0.368	−0.075	−0.032	−0.550	0.325	−0.050	−0.023	−0.553	0.291	−0.021	−0.005
0.5	0	0.6	−0.442	0.467	0.074	0.116	−0.459	0.390	0.070	0.095	−0.474	0.317	0.062	0.077
	0	1.2	−0.531	0.446	−0.044	0.003	−0.544	0.379	−0.028	0.001	−0.557	0.313	−0.017	0.000
	0	2	−0.525	0.465	−0.047	−0.001	−0.537	0.387	−0.029	0.000	−0.550	0.318	−0.017	0.000
	0.4	0.8	−0.442	0.603	0.116	0.166	−0.467	0.508	0.108	0.137	−0.493	0.410	0.090	0.107
	0.4	1.2	−0.520	0.495	−0.015	0.033	−0.537	0.408	−0.009	0.020	−0.553	0.329	−0.006	0.011
	0.7	1	−0.436	0.712	0.151	0.206	−0.462	0.604	0.145	0.178	−0.493	0.485	0.122	0.141
−0.5	0	0.6	−0.525	0.110	−0.144	−0.115	−0.525	0.116	−0.111	−0.093	−0.523	0.103	−0.088	−0.077
	0	1.2	−0.534	0.439	−0.049	−0.003	−0.547	0.380	−0.029	−0.001	−0.557	0.313	−0.017	0.000
	0	2	−0.524	0.469	−0.047	0.001	−0.539	0.392	−0.029	0.000	−0.552	0.320	−0.017	0.000
	0.4	0.8	−0.583	0.110	−0.201	−0.169	−0.579	0.127	−0.155	−0.134	−0.576	0.120	−0.121	−0.107
	0.4	1.2	−0.546	0.389	−0.079	−0.035	−0.554	0.352	−0.048	−0.020	−0.561	0.298	−0.028	−0.011
	0.7	1	−0.621	0.098	−0.248	−0.214	−0.614	0.118	−0.194	−0.172	−0.608	0.111	−0.153	−0.139

Note: $\bar{\nu}_2 = \bar{\nu}_0(m)$, where $m = 10, 20, 40$, corresponding to $n = 64, 128, 256$, respectively; $\theta_n^{(1)} = \log^{-1} n$, $\theta_n^{(2)} = -n^{-1/2}$, $\theta_n^{(3)} = n^{-1}$, $\theta_n^{(4)} = 0$.

TABLE 5. Monte Carlo standard deviation of $\bar{\nu}_2$

ρ	γ	n δ	64 $\theta_n^{(1)}$	64 $\theta_n^{(2)}$	64 $\theta_n^{(3)}$	64 $\theta_n^{(4)}$	128 $\theta_n^{(1)}$	128 $\theta_n^{(2)}$	128 $\theta_n^{(3)}$	128 $\theta_n^{(4)}$	256 $\theta_n^{(1)}$	256 $\theta_n^{(2)}$	256 $\theta_n^{(3)}$	256 $\theta_n^{(4)}$
0	0	0.6	0.095	0.165	0.096	0.098	0.079	0.109	0.057	0.058	0.069	0.071	0.036	0.036
	0	1.2	0.055	0.108	0.025	0.025	0.044	0.069	0.011	0.010	0.040	0.046	0.005	0.004
	0	2	0.044	0.076	0.007	0.003	0.037	0.048	0.003	0.001	0.032	0.033	0.002	0.000
	0.4	0.8	0.129	0.264	0.181	0.187	0.099	0.190	0.134	0.137	0.075	0.132	0.096	0.097
	0.4	1.2	0.070	0.150	0.077	0.080	0.050	0.091	0.045	0.046	0.041	0.057	0.024	0.025
	0.7	1	0.204	0.451	0.323	0.335	0.166	0.369	0.279	0.285	0.121	0.273	0.213	0.216
0.5	0	0.6	0.109	0.127	0.095	0.093	0.090	0.077	0.063	0.061	0.078	0.046	0.043	0.042
	0	1.2	0.054	0.102	0.023	0.022	0.044	0.068	0.010	0.009	0.037	0.045	0.004	0.004
	0	2	0.045	0.073	0.006	0.003	0.038	0.050	0.003	0.001	0.030	0.031	0.001	0.000
	0.4	0.8	0.133	0.223	0.167	0.170	0.105	0.159	0.129	0.130	0.083	0.107	0.092	0.092
	0.4	1.2	0.070	0.122	0.067	0.068	0.051	0.078	0.040	0.040	0.039	0.047	0.022	0.022
	0.7	1	0.185	0.395	0.285	0.294	0.150	0.321	0.247	0.252	0.114	0.239	0.192	0.194
−0.5	0	0.6	0.076	0.176	0.082	0.088	0.062	0.130	0.056	0.060	0.056	0.095	0.039	0.041
	0	1.2	0.053	0.106	0.021	0.021	0.043	0.067	0.010	0.009	0.039	0.046	0.005	0.004
	0	2	0.045	0.072	0.007	0.003	0.036	0.047	0.003	0.001	0.030	0.031	0.001	0.000
	0.4	0.8	0.102	0.250	0.156	0.163	0.079	0.190	0.121	0.125	0.063	0.139	0.090	0.092
	0.4	1.2	0.057	0.147	0.062	0.066	0.044	0.093	0.039	0.040	0.039	0.058	0.022	0.022
	0.7	1	0.157	0.389	0.268	0.278	0.133	0.322	0.235	0.240	0.105	0.246	0.187	0.190

Note: $\bar{\nu}_2 = \bar{\nu}_0(m)$, where $m = 10, 20, 40$, corresponding to $n = 64, 128, 256$, respectively; $\theta_n^{(1)} = \log^{-1} n$, $\theta_n^{(2)} = -n^{-1/2}$, $\theta_n^{(3)} = n^{-1}$, $\theta_n^{(4)} = 0$.

We reported results corresponding to six different estimates, which are as follows: $\bar{\nu}_1$, $\bar{\nu}_2$, their feasible D estimate counterparts $\bar{\nu}_5$, $\bar{\nu}_6$, respectively, $\bar{\nu}_7$, which is the OLS estimate of ν in (34), and its feasible version, denoted by $\bar{\nu}_8$. Noting that OLS and NBLS do not rely on any parametric assumption about the structure of u_t, to compute $\bar{\nu}_5$ and $\bar{\nu}_6$, the estimate of θ was calculated as

$$\tilde{\theta} = \widetilde{\delta + \theta} - \tilde{\delta},$$

where $\widetilde{\delta + \theta}$ is the version of the log periodogram estimate of Geweke and Porter-Hudak (1983) proposed by Robinson (1995a) without pooling or trimming, applied to the series

$$x_t^* = x_t 1(\delta + \theta < 1) + x_t(1)1(1 \leq \delta + \theta < 2) + x_t(2)1(\delta + \theta \geq 2),$$

adding back to the estimate 1 or 2 in the case $x_t^* = x_t(1)$ or $= x_t(2)$, respectively. Similarly, $\tilde{\delta}$ was computed in the same way by applying the log periodogram to the series

$$y_t^* = y_t 1(\delta < 1) + y_t(1)1(1 \leq \delta < 2) + y_t(2)1(\delta \geq 2)$$

instead. In all cases the bandwidths of the estimates of the orders (and in fact also of $\bar{\nu}_2$ and $\bar{\nu}_6$) were $m = 10, 20, 40$, corresponding to $n = 64, 128, 256$, respectively. Alternatively, the feasible estimate $\bar{\nu}_8$ was derived from the OLS regression (34) with γ, $\gamma + \theta$, $\delta + \theta$ replaced by corresponding parametric estimates derived as in (36) and (40). We fixed the sets where the respective functions were optimized in the following way: $\mathcal{D} = [\widetilde{\delta + \theta} - 0.15, \widetilde{\delta + \theta} + 0.15]$, which in all cases contains the asymptotic 95% confidence interval $[\widetilde{\delta + \theta} - 1.96 s.e.(\widetilde{\delta + \theta}), \widetilde{\delta + \theta} + 1.96 s.e.(\widetilde{\delta + \theta})]$, where $s.e.(\widetilde{\delta + \theta}) = \pi/\sqrt{24m}$; $\mathcal{C} = [\gamma - 0.5, \gamma + 0.5]$, $\mathcal{B} = [\gamma + \theta - 1, \min(\gamma + \theta + 0.5, \widehat{\delta + \theta} - 0.05)]$. Note that $\mathcal{C}$ and $\mathcal{B}$ are infeasible sets, but we found them reasonably large. In particular, the upper bound $\widehat{\delta + \theta} - 0.05$ seems sensible because we found it unrealistic that a very small β (less than 0.05) could be detected.

Monte Carlo bias and standard deviations are reported in Tables 6–11. In Tables 6, 8, and 10 the inconsistency of $\bar{\nu}_1$ and $\bar{\nu}_2$ is clearly reflected, with very large negative (positive) biases related to positive (negative) values of θ. Their corresponding feasible D estimates $\bar{\nu}_5$ and $\bar{\nu}_6$ behave in a rather unsatisfactory way. In general, smallest biases correspond to the cases $(\gamma, \delta) = (0, 0.6)$, $(0.4, 0.8)$, $(0.7, 1)$, but they do not react in the appropriate direction as n increases (except for some cases with $\rho = -0.5$). On the contrary, for the cases with $\delta = 1.2, 2$, larger positive biases are reported, but in general, they decrease as n increases, usually very slowly however, indicating that a faster estimate of θ or larger sample sizes may be needed to obtain acceptable results for this class of estimates. Also, although their values are relatively similar, the NBLS appears to be inferior to the OLS. As expected, $\bar{\nu}_7$ performs extremely well, with very small biases sharing the sign of ρ and decreasing in absolute value as β increases. Larger biases are reported for its feasible version $\bar{\nu}_8$, although this estimate still works well when $\beta \geq 0.8$, whereas for the other cases, biases are large but react appropriately when n increases. Finally, it is worth mentioning that ρ has a rather different effect on $\bar{\nu}_8$ from the one reported in Robinson and Hualde (2003). Now, it seems that in comparison to the $\rho = 0$ situation, positive and negative correlation benefits and worsens the estimate, respectively.

Monte Carlo standard deviations are reported in Tables 7, 9, and 11. For the cases of strong cointegration the smallest values correspond to $\bar{\nu}_7$, closely followed by $\bar{\nu}_8$ when $(\gamma, \delta) = (0,2)$ and the inconsistent estimates $\bar{\nu}_1$, $\bar{\nu}_2$ otherwise. For the weak cointegration situation, $\bar{\nu}_1$ and $\bar{\nu}_2$ are best, followed by $\bar{\nu}_7$. In general, for $\beta \geq 0.8$, $\bar{\nu}_8$ beats $\bar{\nu}_5$ and $\bar{\nu}_6$, the opposite happening when $\beta < 0.8$, although in this case the values corresponding to $\bar{\nu}_8$ enjoy faster decreases as n increases, noting that these standard deviations are severely harmed by replications where the parametric estimates of $\gamma + \theta$ and $\delta + \theta$ are very close to each other. Standard deviations decrease as n increases for all estimates.

Finally, the last part of the Monte Carlo is devoted to motivating the correct use of graphical tools to detect the possible presence of strongly unbalanced cointegration. Thus, we generated pairs of time series of dimension $n = 1{,}000$ as in the previous two parts of the experiment, fixing $\psi_{12} = 0$ and concentrating on different values of γ, δ, θ. In Figures 1 and 2 we present the situation where $\delta = 1.4$, $\theta = 0.4$, and under the heading of cointegrated and no-cointegrated we denote situations where $\gamma = 0$ and $\gamma = \delta$ in (11), respectively. Similarly, Figures 3 and 4, 5 and 6, and 7 and 8 represent pairs of series generated as in the first case (but with different seed originating the white noise process), for the cases $(\delta, \theta) = (1, 0.8)$, $(0.8, 0.4)$, and $(0.6, 0.8)$, respectively, where in all situations $\gamma = 0$ or $\gamma = \delta$ corresponds to the cases of cointegration and noncointegration, respectively. Note that if $\gamma = 0$, processes y_t and $x_t(\theta)$ are $CI(\delta,0)$ cointegrated, whereas if $\gamma = \delta$, they are not cointegrated. Graphs on the left of each figure simply represent both time series as a function of time. Clearly, the main consequence of the gap between the integration orders is a different dimension in the series, most evident when $\theta = 0.8$. Hence, these figures simply suggest that the orders of integration of the observable series are different, but it is usually not possible to assess therefrom whether there is an intrinsic linkage between the series or not, it being hard to make a guess about the possible existence of unbalanced cointegration. This picture changes dramatically when focusing on the graphs on the right of each figure. These represent exactly the same series as on the corresponding left graphs but with the important difference that the smallest integration order series is drawn with respect to a different scale given on the right vertical axis. Here, it is evident that strong comovements exist (especially when $\theta = 0.4$) between the two unbalanced series in the case that cointegration exists, whereas, as expected, the cotrending does not appear when the series are not cointegrated. We admit that

TABLE 6. Monte Carlo bias, $\rho = 0$

			$\theta = 0.3$						$\theta = -0.3$					
n	γ	δ	$\bar{\nu}_1$	$\bar{\nu}_5$	$\bar{\nu}_2$	$\bar{\nu}_6$	$\bar{\nu}_7$	$\bar{\nu}_8$	$\bar{\nu}_1$	$\bar{\nu}_5$	$\bar{\nu}_2$	$\bar{\nu}_6$	$\bar{\nu}_7$	$\bar{\nu}_8$
64	0	0.6	−0.547	0.036	−0.574	0.189	−0.004	0.162	0.291	0.033	0.594	0.175	−0.004	0.153
	0	1.2	−0.616	0.661	−0.616	0.720	−0.001	0.001	1.23	0.716	1.28	0.837	−0.001	0.000
	0	2	−0.608	0.687	−0.607	0.686	0.000	0.000	1.49	1.02	1.49	1.02	0.000	0.000
	0.4	0.8	−0.597	0.069	−0.606	0.196	−0.006	0.366	0.628	0.113	0.844	0.234	−0.006	0.332
	0.4	1.2	−0.618	0.675	−0.619	0.749	−0.003	0.027	1.22	0.716	1.27	0.834	−0.003	0.031
	0.7	1	−0.623	0.290	−0.626	0.395	−0.009	0.417	0.944	0.338	1.05	0.473	−0.009	0.403
128	0	0.6	−0.628	0.153	−0.645	0.286	−0.001	0.037	0.380	0.136	0.733	0.246	−0.001	0.048
	0	1.2	−0.685	0.658	−0.685	0.670	0.000	0.003	1.74	0.816	1.78	0.849	0.000	0.003
	0	2	−0.677	0.536	−0.677	0.536	0.000	0.001	2.03	0.815	2.03	0.814	0.000	0.001
	0.4	0.8	−0.669	0.210	−0.674	0.282	−0.001	0.329	0.894	0.231	1.14	0.295	−0.001	0.305
	0.4	1.2	−0.686	0.709	−0.687	0.721	−0.001	0.014	1.73	0.862	1.77	0.898	−0.001	0.008
	0.7	1	−0.690	0.418	−0.691	0.454	0.000	0.480	1.38	0.559	1.49	0.627	0.000	0.492
256	0	0.6	−0.697	0.242	−0.707	0.358	0.000	0.007	0.449	0.201	0.846	0.294	0.000	0.007
	0	1.2	−0.746	0.595	−0.746	0.597	0.000	0.004	2.39	0.772	2.41	0.776	0.000	0.005
	0	2	−0.736	0.430	−0.736	0.430	0.000	0.000	2.74	0.606	2.74	0.606	0.000	0.000
	0.4	0.8	−0.731	0.297	−0.733	0.344	−0.001	0.089	1.18	0.301	1.47	0.342	−0.001	0.090
	0.4	1.2	−0.746	0.668	−0.746	0.671	0.000	0.002	2.38	0.851	2.41	0.857	0.000	0.002
	0.7	1	−0.745	0.491	−0.745	0.505	−0.002	0.334	1.93	0.641	2.04	0.666	−0.002	0.337

Note: $\bar{\nu}_1 = \bar{\nu}_0([n/2])$, $\bar{\nu}_2 = \bar{\nu}_0(m)$, $\bar{\nu}_5 = \bar{\nu}_{\hat{\theta}}([n/2])$, $\bar{\nu}_6 = \bar{\nu}_{\hat{\theta}}(m)$, $\bar{\nu}_7 = \bar{\nu}(\gamma, \gamma + \theta, \delta + \theta)$, $\bar{\nu}_8 = \bar{\nu}(\hat{\gamma}, \widehat{\gamma + \theta}, \widehat{\delta + \theta})$, where $m = 10, 20, 40$, corresponding to $n = 64, 128, 256$, respectively.

https://doi.org/10.1017/S0266466606060361 Published online by Cambridge University Press

TABLE 7. Monte Carlo standard deviation, $\rho = 0$

			$\theta = 0.3$						$\theta = -0.3$					
n	γ	δ	$\bar{\nu}_1$	$\bar{\nu}_5$	$\bar{\nu}_2$	$\bar{\nu}_6$	$\bar{\nu}_7$	$\bar{\nu}_8$	$\bar{\nu}_1$	$\bar{\nu}_5$	$\bar{\nu}_2$	$\bar{\nu}_6$	$\bar{\nu}_7$	$\bar{\nu}_8$
64	0	0.6	0.104	0.324	0.092	0.471	0.117	1.05	0.266	0.323	0.346	0.465	0.117	0.992
	0	1.2	0.056	1.12	0.056	1.26	0.025	0.086	0.446	1.36	0.428	1.71	0.025	0.080
	0	2	0.046	1.58	0.046	1.58	0.003	0.013	0.320	3.06	0.319	3.08	0.003	0.013
	0.4	0.8	0.119	0.531	0.118	0.700	0.212	1.80	0.456	0.510	0.504	0.669	0.212	1.70
	0.4	1.2	0.067	1.18	0.067	1.40	0.069	0.392	0.470	1.34	0.456	1.65	0.069	0.488
	0.7	1	0.179	0.881	0.181	1.15	0.305	2.17	0.703	0.970	0.745	1.29	0.305	2.13
128	0	0.6	0.086	0.272	0.075	0.400	0.066	0.268	0.274	0.278	0.364	0.380	0.066	0.359
	0	1.2	0.043	0.846	0.042	0.863	0.010	0.038	0.553	1.32	0.537	1.47	0.010	0.038
	0	2	0.038	0.785	0.038	0.784	0.001	0.005	0.374	1.44	0.374	1.44	0.001	0.006
	0.4	0.8	0.085	0.521	0.083	0.618	0.128	1.31	0.532	0.475	0.567	0.560	0.128	1.26
	0.4	1.2	0.046	0.907	0.046	0.924	0.035	0.212	0.564	1.35	0.549	1.52	0.035	0.112
	0.7	1	0.129	0.832	0.129	0.902	0.189	1.65	0.797	1.10	0.818	1.26	0.189	1.67
256	0	0.6	0.069	0.257	0.061	0.369	0.041	0.125	0.266	0.247	0.363	0.342	0.041	0.127
	0	1.2	0.036	0.591	0.036	0.594	0.004	0.022	0.686	0.933	0.674	0.940	0.004	0.021
	0	2	0.031	0.486	0.031	0.486	0.000	0.002	0.463	0.817	0.463	0.817	0.000	0.002
	0.4	0.8	0.060	0.504	0.058	0.567	0.086	0.572	0.603	0.445	0.633	0.496	0.086	0.566
	0.4	1.2	0.037	0.656	0.037	0.659	0.020	0.061	0.691	1.00	0.679	1.01	0.020	0.061
	0.7	1	0.083	0.710	0.083	0.731	0.130	1.21	0.881	1.01	0.886	1.07	0.130	1.21

Note: $\bar{\nu}_1 = \bar{\nu}_0([n/2])$, $\bar{\nu}_2 = \bar{\nu}_0(m)$, $\bar{\nu}_5 = \bar{\nu}_{\hat{\theta}}([n/2])$, $\bar{\nu}_6 = \bar{\nu}_{\hat{\theta}}(m)$, $\bar{\nu}_7 = \bar{\nu}(\gamma, \gamma + \theta, \delta + \theta)$, $\bar{\nu}_8 = \bar{\nu}(\hat{\gamma}, \widehat{\gamma + \theta}, \widehat{\delta + \theta})$, where $m = 10, 20, 40$, corresponding to $n = 64, 128, 256$, respectively.

TABLE 8. Monte Carlo bias, $\rho = 0.5$

			$\theta = 0.3$						$\theta = -0.3$					
n	γ	δ	$\bar{\nu}_1$	$\bar{\nu}_5$	$\bar{\nu}_2$	$\bar{\nu}_6$	$\bar{\nu}_7$	$\bar{\nu}_8$	$\bar{\nu}_1$	$\bar{\nu}_5$	$\bar{\nu}_2$	$\bar{\nu}_6$	$\bar{\nu}_7$	$\bar{\nu}_8$
64	0	0.6	−0.491	0.398	−0.543	0.512	0.001	0.046	0.696	0.406	0.899	0.501	0.001	0.026
	0	1.2	−0.614	0.787	−0.615	0.835	0.001	−0.006	1.28	0.865	1.30	0.968	0.001	−0.004
	0	2	−0.606	0.352	−0.606	0.351	0.000	−0.010	1.48	0.663	1.48	0.662	0.000	−0.010
	0.4	0.8	−0.527	0.459	−0.544	0.584	0.001	0.136	1.06	0.486	1.26	0.593	0.001	0.116
	0.4	1.2	−0.605	0.961	−0.606	1.04	0.001	−0.003	1.36	1.03	1.39	1.15	0.001	0.004
	0.7	1	−0.533	0.683	−0.538	0.779	0.000	0.185	1.48	0.773	1.60	0.937	0.000	0.113
128	0	0.6	−0.597	0.502	−0.628	0.587	0.003	0.020	0.777	0.484	1.02	0.538	0.003	0.013
	0	1.2	−0.685	0.810	−0.685	0.818	0.000	−0.007	1.77	0.991	1.79	1.02	0.000	−0.006
	0	2	−0.675	0.176	−0.674	0.176	0.000	−0.007	2.01	0.370	2.01	0.369	0.000	−0.006
	0.4	0.8	−0.628	0.541	−0.636	0.616	0.005	0.130	1.31	0.584	1.54	0.643	0.005	0.131
	0.4	1.2	−0.681	1.05	−0.681	1.07	0.001	0.005	1.84	1.24	1.87	1.28	0.001	0.008
	0.7	1	−0.630	0.829	−0.632	0.866	0.006	0.370	1.92	1.02	2.03	1.12	0.006	0.358
256	0	0.6	−0.681	0.584	−0.699	0.662	0.001	0.006	0.840	0.543	1.13	0.585	0.001	0.005
	0	1.2	−0.747	0.748	−0.747	0.749	0.000	−0.018	2.41	0.992	2.43	0.999	0.000	−0.018
	0	2	−0.736	0.069	−0.736	0.069	0.000	−0.003	2.74	0.226	2.73	0.226	0.000	−0.003
	0.4	0.8	−0.710	0.607	−0.713	0.646	0.003	0.039	1.59	0.619	1.85	0.648	0.003	0.040
	0.4	1.2	−0.745	1.05	−0.745	1.05	0.001	0.001	2.47	1.33	2.49	1.34	0.001	0.001
	0.7	1	−0.712	0.824	−0.712	0.835	0.004	0.217	2.44	1.05	2.56	1.08	0.004	0.220

Note: $\bar{\nu}_1 = \bar{\nu}_0([n/2])$, $\bar{\nu}_2 = \bar{\nu}_0(m)$, $\bar{\nu}_5 = \bar{\nu}_{\hat{\theta}}([n/2])$, $\bar{\nu}_6 = \bar{\nu}_{\hat{\theta}}(m)$, $\bar{\nu}_7 = \bar{\nu}(\gamma, \gamma + \theta, \delta + \theta)$, $\bar{\nu}_8 = \bar{\nu}(\hat{\gamma}, \widehat{\gamma + \theta}, \widehat{\delta + \theta})$, where $m = 10, 20, 40$, corresponding to $n = 64, 128, 256$, respectively.

https://doi.org/10.1017/S0266466606060361 Published online by Cambridge University Press

TABLE 9. Monte Carlo standard deviation, $\rho = 0.5$

			$\theta = 0.3$						$\theta = -0.3$					
n	γ	δ	$\bar{\nu}_1$	$\bar{\nu}_5$	$\bar{\nu}_2$	$\bar{\nu}_6$	$\bar{\nu}_7$	$\bar{\nu}_8$	$\bar{\nu}_1$	$\bar{\nu}_5$	$\bar{\nu}_2$	$\bar{\nu}_6$	$\bar{\nu}_7$	$\bar{\nu}_8$
64	0	0.6	0.135	0.365	0.103	0.529	0.101	0.670	0.222	0.347	0.300	0.496	0.101	0.727
	0	1.2	0.055	1.16	0.054	1.33	0.021	0.088	0.418	1.35	0.416	1.67	0.021	0.082
	0	2	0.047	1.47	0.047	1.47	0.003	0.016	0.305	2.65	0.304	2.66	0.003	0.015
	0.4	0.8	0.130	0.614	0.122	0.800	0.184	1.66	0.411	0.561	0.457	0.731	0.184	1.73
	0.4	1.2	0.068	1.25	0.067	1.50	0.059	0.263	0.422	1.39	0.419	1.73	0.059	0.242
	0.7	1	0.165	0.911	0.164	1.09	0.266	2.23	0.649	1.12	0.685	1.46	0.266	2.24
128	0	0.6	0.102	0.316	0.080	0.456	0.058	0.228	0.236	0.284	0.323	0.390	0.058	0.256
	0	1.2	0.043	0.893	0.043	0.915	0.009	0.037	0.526	1.43	0.521	1.52	0.009	0.037
	0	2	0.039	0.669	0.039	0.668	0.001	0.006	0.384	1.03	0.384	1.03	0.001	0.006
	0.4	0.8	0.092	0.600	0.086	0.724	0.113	0.852	0.480	0.518	0.513	0.622	0.113	0.792
	0.4	1.2	0.047	1.05	0.047	1.09	0.031	0.120	0.518	1.60	0.512	1.74	0.031	0.116
	0.7	1	0.116	0.886	0.116	0.965	0.168	1.62	0.731	1.26	0.750	1.53	0.168	1.54
256	0	0.6	0.078	0.286	0.065	0.415	0.034	0.123	0.232	0.249	0.330	0.331	0.034	0.122
	0	1.2	0.033	0.659	0.033	0.662	0.004	0.017	0.661	1.26	0.654	1.28	0.004	0.017
	0	2	0.030	0.379	0.030	0.379	0.000	0.003	0.443	0.733	0.443	0.732	0.000	0.003
	0.4	0.8	0.065	0.545	0.062	0.609	0.073	0.353	0.568	0.444	0.599	0.490	0.073	0.338
	0.4	1.2	0.034	0.829	0.034	0.837	0.016	0.061	0.649	1.49	0.641	1.53	0.016	0.061
	0.7	1	0.078	0.680	0.077	0.696	0.112	0.967	0.839	1.14	0.839	1.21	0.112	0.947

Note: $\bar{\nu}_1 = \bar{\nu}_0([n/2])$, $\bar{\nu}_2 = \bar{\nu}_0(m)$, $\bar{\nu}_5 = \bar{\nu}_{\hat{\theta}}([n/2])$, $\bar{\nu}_6 = \bar{\nu}_{\hat{\theta}}(m)$, $\bar{\nu}_7 = \bar{\nu}(\gamma, \gamma + \theta, \delta + \theta)$, $\bar{\nu}_8 = \bar{\nu}(\hat{\gamma}, \widehat{\gamma + \theta}, \widehat{\delta + \theta})$, where $m = 10, 20, 40$, corresponding to $n = 64, 128, 256$, respectively.

https://doi.org/10.1017/S0266466606060361 Published online by Cambridge University Press

Table 10. Monte Carlo bias, $\rho = -0.5$

			$\theta = 0.3$						$\theta = -0.3$					
n	γ	δ	$\bar{\nu}_1$	$\bar{\nu}_5$	$\bar{\nu}_2$	$\bar{\nu}_6$	$\bar{\nu}_7$	$\bar{\nu}_8$	$\bar{\nu}_1$	$\bar{\nu}_5$	$\bar{\nu}_2$	$\bar{\nu}_6$	$\bar{\nu}_7$	$\bar{\nu}_8$
64	0	0.6	−0.602	−0.274	−0.605	−0.100	0.000	0.331	−0.094	−0.270	0.318	−0.101	0.000	0.299
	0	1.2	−0.617	0.501	−0.617	0.560	0.000	0.040	1.22	0.573	1.29	0.692	0.000	0.029
	0	2	−0.605	0.898	−0.604	0.897	0.000	0.014	1.50	1.22	1.50	1.22	0.000	0.013
	0.4	0.8	−0.653	−0.228	−0.655	−0.116	0.000	0.562	0.242	−0.212	0.487	−0.098	0.000	0.547
	0.4	1.2	−0.627	0.346	−0.627	0.402	0.000	0.105	1.14	0.421	1.20	0.506	0.000	0.142
	0.7	1	−0.687	−0.058	−0.688	0.034	0.000	0.636	0.511	−0.024	0.622	0.093	0.000	0.630
128	0	0.6	−0.663	−0.159	−0.666	−0.004	−0.001	0.063	−0.005	−0.180	0.460	−0.046	−0.001	0.054
	0	1.2	−0.688	0.498	−0.688	0.511	0.000	0.018	1.76	0.721	1.81	0.749	0.000	0.016
	0	2	−0.677	0.779	−0.677	0.778	0.000	0.008	2.05	1.07	2.05	1.07	0.000	0.008
	0.4	0.8	−0.706	−0.133	−0.707	−0.066	−0.002	0.351	0.514	−0.126	0.791	−0.064	−0.002	0.355
	0.4	1.2	−0.693	0.320	−0.693	0.330	−0.001	0.022	1.69	0.512	1.74	0.531	−0.001	0.024
	0.7	1	−0.733	0.040	−0.734	0.071	−0.003	0.634	0.949	0.142	1.06	0.186	−0.003	0.587
256	0	0.6	−0.716	−0.074	−0.719	0.047	−0.001	0.007	0.065	−0.106	0.572	0.000	−0.001	0.012
	0	1.2	−0.747	0.469	−0.746	0.473	0.000	0.043	2.40	0.634	2.44	0.639	0.000	0.042
	0	2	−0.738	0.762	−0.738	0.761	0.000	0.004	2.75	0.962	2.75	0.962	0.000	0.004
	0.4	0.8	−0.754	−0.055	−0.754	−0.014	−0.003	0.103	0.796	−0.053	1.10	−0.015	−0.003	0.111
	0.4	1.2	−0.748	0.289	−0.748	0.291	0.000	0.004	2.35	0.423	2.38	0.426	0.000	0.002
	0.7	1	−0.773	0.076	−0.774	0.087	−0.006	0.319	1.45	0.171	1.56	0.187	−0.006	0.312

Note: $\bar{\nu}_1 = \bar{\nu}_0([n/2])$, $\bar{\nu}_2 = \bar{\nu}_0(m)$, $\bar{\nu}_5 = \bar{\nu}_{\hat{\theta}}([n/2])$, $\bar{\nu}_6 = \bar{\nu}_{\hat{\theta}}(m)$, $\bar{\nu}_7 = \bar{\nu}(\gamma, \gamma + \theta, \delta + \theta)$, $\bar{\nu}_8 = \bar{\nu}(\hat{\gamma}, \widehat{\gamma+\theta}, \widehat{\delta+\theta})$, where $m = 10, 20, 40$, corresponding to $n = 64, 128, 256$, respectively.

Table 11. Monte Carlo standard deviation, $\rho = -0.5$

			$\theta = 0.3$						$\theta = -0.3$					
n	γ	δ	$\bar{\nu}_1$	$\bar{\nu}_5$	$\bar{\nu}_2$	$\bar{\nu}_6$	$\bar{\nu}_7$	$\bar{\nu}_8$	$\bar{\nu}_1$	$\bar{\nu}_5$	$\bar{\nu}_2$	$\bar{\nu}_6$	$\bar{\nu}_7$	$\bar{\nu}_8$
64	0	0.6	0.074	0.270	0.076	0.408	0.096	1.55	0.292	0.272	0.362	0.397	0.096	1.53
	0	1.2	0.054	0.991	0.053	1.19	0.020	0.442	0.458	1.26	0.426	1.58	0.020	0.362
	0	2	0.047	1.27	0.047	1.27	0.002	0.016	0.299	2.35	0.298	2.35	0.002	0.016
	0.4	0.8	0.091	0.438	0.092	0.555	0.178	2.00	0.450	0.432	0.486	0.573	0.178	1.97
	0.4	1.2	0.056	1.01	0.056	1.21	0.056	0.850	0.485	1.24	0.460	1.47	0.056	0.979
	0.7	1	0.135	0.733	0.136	0.902	0.259	2.18	0.629	0.810	0.660	1.03	0.259	2.14
128	0	0.6	0.063	0.243	0.063	0.356	0.057	0.634	0.300	0.242	0.380	0.344	0.057	0.460
	0	1.2	0.042	0.787	0.042	0.823	0.009	0.042	0.552	1.30	0.527	1.38	0.009	0.041
	0	2	0.037	0.743	0.037	0.743	0.001	0.006	0.364	1.53	0.364	1.52	0.001	0.006
	0.4	0.8	0.067	0.424	0.067	0.510	0.112	1.32	0.530	0.386	0.555	0.468	0.112	1.37
	0.4	1.2	0.042	0.729	0.042	0.748	0.030	0.357	0.573	1.16	0.551	1.21	0.030	0.427
	0.7	1	0.102	0.652	0.102	0.708	0.167	1.78	0.723	0.800	0.737	0.885	0.167	1.75
256	0	0.6	0.057	0.223	0.056	0.304	0.036	0.208	0.303	0.220	0.392	0.289	0.036	0.217
	0	1.2	0.035	0.529	0.035	0.532	0.004	0.025	0.688	0.912	0.670	0.922	0.004	0.025
	0	2	0.029	0.521	0.029	0.521	0.000	0.003	0.457	0.967	0.457	0.967	0.000	0.003
	0.4	0.8	0.051	0.412	0.051	0.460	0.076	0.643	0.623	0.374	0.644	0.416	0.076	0.671
	0.4	1.2	0.035	0.499	0.035	0.501	0.017	0.069	0.704	0.785	0.687	0.790	0.017	0.066
	0.7	1	0.073	0.523	0.073	0.536	0.116	1.19	0.829	0.689	0.828	0.715	0.116	1.19

Note: $\bar{\nu}_1 = \bar{\nu}_0([n/2])$, $\bar{\nu}_2 = \bar{\nu}_0(m)$, $\bar{\nu}_5 = \bar{\nu}_{\hat{\theta}}([n/2])$, $\bar{\nu}_6 = \bar{\nu}_{\hat{\theta}}(m)$, $\bar{\nu}_7 = \bar{\nu}(\gamma, \gamma + \theta, \delta + \theta)$, $\bar{\nu}_8 = \bar{\nu}(\hat{\gamma}, \widehat{\gamma + \theta}, \widehat{\delta + \theta})$, where $m = 10, 20, 40$, corresponding to $n = 64, 128, 256$, respectively.

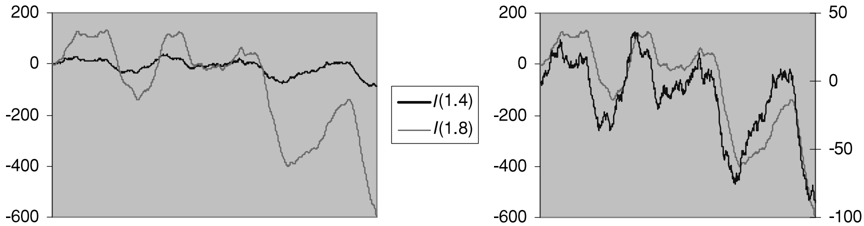

Figure 1. Strongly unbalanced cointegrated $I(1.8)$, $I(1.4)$ series.

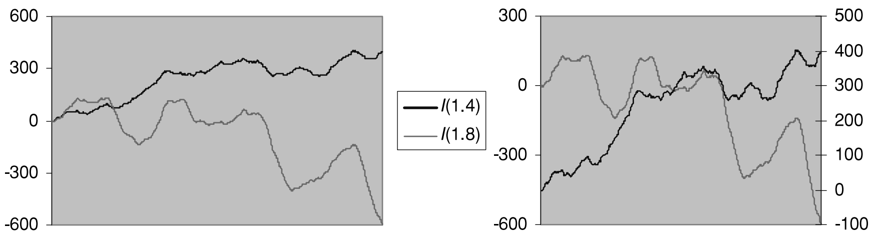

Figure 2. Noncointegrated $I(1.8)$, $I(1.4)$ series.

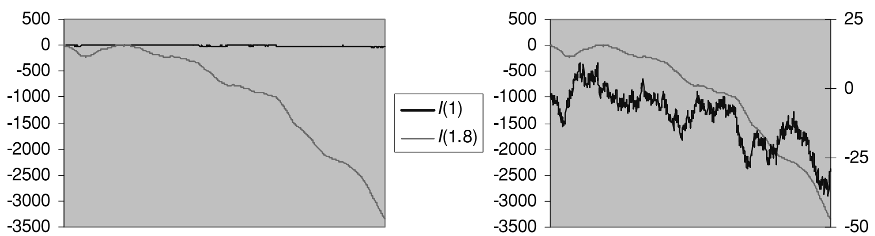

Figure 3. Strongly unbalanced cointegrated $I(1.8)$, $I(1)$ series.

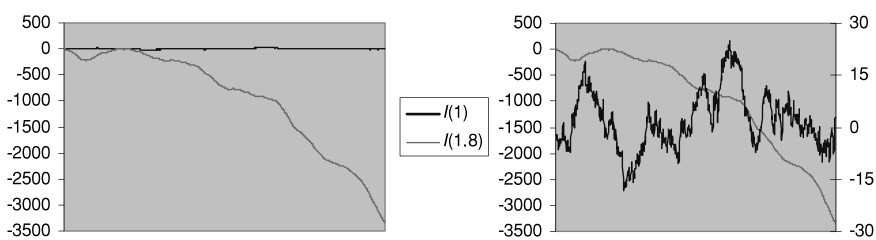

Figure 4. Noncointegrated $I(1.8)$, $I(1)$ series.

https://doi.org/10.1017/S0266466606060361 Published online by Cambridge University Press

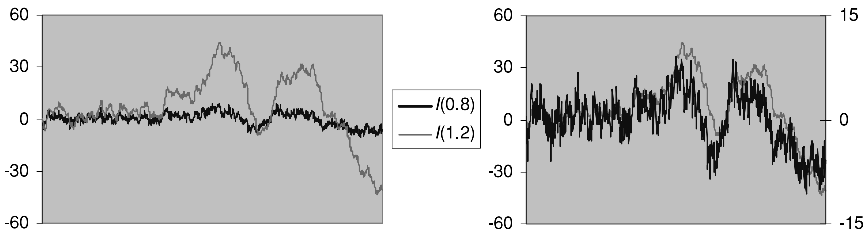

FIGURE 5. Strongly unbalanced cointegrated $I(1.2)$, $I(0.8)$ series.

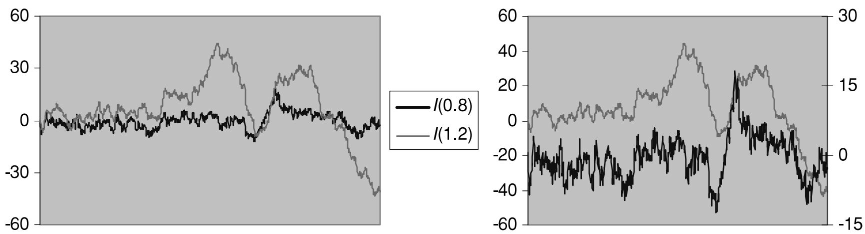

FIGURE 6. Noncointegrated $I(1.2)$, $I(0.8)$ series.

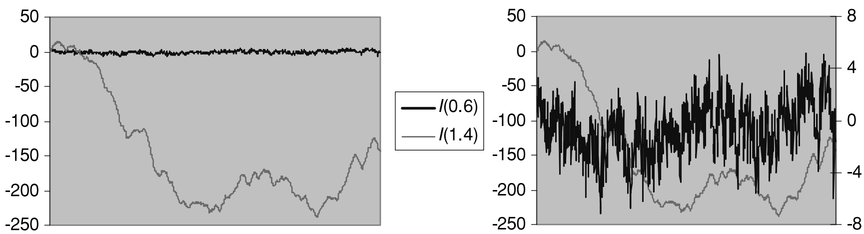

FIGURE 7. Strongly unbalanced cointegrated $I(1.4)$, $I(0.6)$ series.

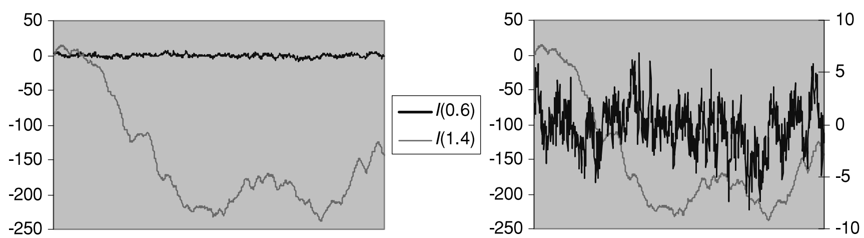

FIGURE 8. Noncointegrated $I(1.4)$, $I(0.6)$ series.

this is not a significative study, and undoubtedly further empirical and Monte Carlo investigation is needed, but our intention in showing these figures was simply to give the flavor of the possible existence of intrinsic linkages between series with different orders of integration and thus motivate empirical researchers to interpret real data following these lines.

REFERENCES

Abramowitz, M. & I. Stegun (1970) *Handbook of Mathematical Functions*. Dover.

Ahtola, J. & G.C. Tiao (1984) Parameter inference for a nearly nonstationary first-order autoregressive model. *Biometrika* 71, 263–272.

Akonom, J. & C. Gourieroux (1987) A Functional Central Limit Theorem for Fractional Processes. Preprint, CEPREMAP, Paris.

Backus, D., S. Foressi, & C. Telmer (1996) Affine models of currency pricing. NBER Working paper 5623.

Baillie, R.T. & T. Bollerslev (2000) The forward premium anomaly is not as bad as you think. *Journal of International Money and Finance* 19, 471–488.

Bekaert, G. (1996) The time variation of risk and return in foreign exchange markets: A general equilibrium approach. *Review of Financial Studies* 9, 427–470.

Bekaert, G., R.J. Hodrick, & D.A. Marshall (1997) The implications of first order risk aversion for asset market risk premiums. *Journal of Monetary Economics* 40, 3–39.

Bingham, N.H., C.M. Goldie, & J.L. Teugels (1989) *Regular Variation*. Cambridge University Press.

Brockwell, P.J. & R.A. Davis (1991) *Time Series: Theory and Methods*. Springer-Verlag.

Chan, N.H. & C.Z. Wei (1987) Asymptotic inference for nearly nonstationary AR(1) processes. *Annals of Statistics* 15, 1050–1063.

Cox, D.D. & I. Llatas (1991) Maximum likelihood type estimation for nearly nonstationary autoregressive time series. *Annals of Statistics* 19, 1109–1128.

Elliott, G. (1998) On the robustness of cointegration methods when regressors almost have unit roots. *Econometrica* 66, 149–158.

Engle, R.F. & C.W.J. Granger (1987) Cointegration and error correction: Representation, estimation and testing. *Econometrica* 55, 251–276.

Evans, G.B.A. & N.E. Savin (1981) Testing for unit roots, part 1. *Econometrica* 49, 753–779.

Geweke, J. & S. Porter-Hudak (1983) The estimation and application of long memory time series models. *Journal of Time Series Analysis* 4, 221–238.

Granger, C.W.J. & R. Joyeux (1980) An introduction to long-memory time series models and fractional differencing. *Journal of Time Series Analysis* 1, 15–29.

Hosking, J.R.M. (1981) Fractional differencing. *Biometrika* 68, 165–176.

Hualde, J. (2002) Testing for the Equality of Orders of Integration. Preprint, Universidad de Navarra.

Hualde, J. & P.M. Robinson (2006) Root-*n*-consistent estimation of weak fractional cointegration. *Journal of Econometrics*, forthcoming.

Hualde, J. & P.M. Robinson (2004a) Semiparametric Estimation of Fractional Cointegration. Preprint, Universidad de Navarra.

Hualde, J. & P.M. Robinson (2004b) Gaussian Pseudo-Maximum Likelihood Estimation of Fractional Time Series Models. Preprint, Universidad de Navarra.

Jansson, M. & N. Haldrup (2002) Regression theory for nearly cointegrated time series. *Econometric Theory* 18, 1309–1335.

Jeganathan, P. (1999) On asymptotic inference in cointegrated time series with fractionally integrated errors. *Econometric Theory* 15, 583–621.

Kim, C.S. & P.C.B. Phillips (2000) Fully Modified Estimation of Fractional Cointegration Models. Preprint, Yale University.

Lobato, I.N. & P.M. Robinson (1996) Averaged periodogram estimation of long memory. *Journal of Econometrics* 73, 303–324.

Marinucci, D. & P.M. Robinson (1999) Alternative forms of fractional Brownian motion. *Journal of Statistical Planning and Inference* 80, 111–122.

Marinucci, D. & P.M. Robinson (2000) Weak convergence of multivariate fractional processes. *Stochastic Processes and Their Applications* 86, 103–120.

Marinucci, D. & P.M. Robinson (2001) Semiparametric fractional cointegration analysis. *Journal of Econometrics* 105, 225–247.

Marmol, F. (1998) Spurious regression theory with nonstationary fractionally integrated processes. *Journal of Econometrics* 84, 232–250.

Marmol, F. & C. Velasco (2004) Consistent testing of cointegrating relationships. *Econometrica* 72, 1809–1844.

Maynard, A. & P.C.B. Phillips (2001) Rethinking an old empirical puzzle: Econometric evidence on the forward discount anomaly. *Journal of Applied Econometrics* 16, 671–708.

Park, J.Y., S. Ouliaris, & B. Choi (1988) Spurious Regressions and Tests for Cointegration. CAE Working paper 88–07, Cornell University.

Phillips, P.C.B. (1986) Understanding spurious regressions in econometrics. *Journal of Econometrics* 33, 311–340.

Phillips, P.C.B. (1987) Towards a unified asymptotic theory for autoregression. *Biometrika* 74, 535–547.

Phillips, P.C.B. (1988) Regression theory for near-integrated time series. *Econometrica* 56, 1021–1043.

Phillips, P.C.B. (1991) Optimal inference in cointegrated systems. *Econometrica* 59, 283–306.

Phillips, P.C.B. & S.N. Durlauf (1986) Multiple time series regression with integrated processes. *Review of Economic Studies* 53, 473–495.

Robinson, P.M. (1994) Semiparametric analysis of long-memory time series. *Annals of Statistics* 22, 515–539.

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) The distance between rival nonstationary fractional processes. *Journal of Econometrics* 128, 283–300.

Robinson, P.M. & J. Hualde (2003) Cointegration in fractional systems with unknown integration orders. *Econometrica* 71, 1727–1766.

Robinson, P.M. & D. Marinucci (2001) Narrow-band analysis of nonstationary processes. *Annals of Statistics* 29, 947–986.

Robinson, P.M. & D. Marinucci (2003) Semiparametric frequency-domain analysis of fractional cointegration. In P.M. Robinson (ed.), *Time Series with Long Memory*, pp. 334–373. Oxford University Press.

Robinson, P.M. & Y. Yajima (2002) Determination of cointegrating rank in fractional systems. *Journal of Econometrics* 106, 217–241.

Silveira, G. (1991) Contributions to strong approximations in time series with applications in nonparametric statistics and functional central limit theorems. Ph.D. thesis, University of London.

Velasco, C. (1999a) Non-stationary log-periodogram regression. *Journal of Econometrics* 91, 325–371.

Velasco, C. (1999b) Gaussian semiparametric estimation of non-stationary time series. *Journal of Time Series Analysis* 20, 87–127.

Velasco, C. & P.M. Robinson (2000) Whittle pseudo-maximum likelihood estimation for nonstationary time series. *Journal of the American Statistical Association* 95, 1229–1243.

Zygmund, A. (1977) *Trigonometric Series*. Cambridge University Press.

APPENDIX A: Characterization of Limiting Distributions

THEOREM BCU.

(i) If $\gamma + \delta > 1$

$$p_i(n) = n^{\beta}, \qquad \Xi_i(\gamma,\delta) = \xi' B(1)\Psi^{1/2}\int_0^1 W(r;\delta)W'(r;\gamma)\,dr\Psi^{1/2}B'(1)\zeta, \qquad \gamma > \tfrac{1}{2}, \qquad i = 1,2,$$

$$p_i(n) = n^{\beta}, \qquad \Xi_i(\gamma,\delta) = \xi' B(1)\Psi^{1/2}\int_0^1 W(r;\delta)\,dW'(r)\Psi^{1/2}B'(1)\zeta, \qquad \gamma = 0, \qquad i = 1,2; \qquad \textbf{(A.1)}$$

(ii) if $\gamma + \delta = 1$

$$p_1(n) = n^{2\delta-1}/\log n, \qquad \Xi_1(\gamma,\delta) = 2h_{12}(0)\sin(\delta\pi), \qquad \gamma > 0,$$

$$p_1(n) = n, \qquad \Xi_1(\gamma,\delta) = \xi' B(1)\Psi^{1/2}\int_0^1 W(r)\,dW'(r)\Psi^{1/2}B'(1)\zeta + \sum_{j=0}^{\infty} E(u_{10}v_{2,-j}), \qquad \gamma = 0,$$

$$p_2(n) = n^{2\delta-1}/\log m, \qquad \Xi_2(\gamma,\delta) = \Xi_1(\gamma,\delta), \qquad \gamma > 0,$$

$$p_2(n) = n, \qquad \Xi_2(\gamma,\delta) = \xi' B(1)\Psi^{1/2}\int_0^1 W(r)\,dW'(r)\Psi^{1/2}B'(1)\zeta + \pi h_{12}(0), \qquad \gamma = 0;$$

(iii) if $\gamma + \delta < 1$

$$p_1(n) = n^{2\delta-1}, \qquad \Xi_1(\gamma,\delta) = \int_{-\pi}^{\pi} a(\delta;\lambda)a(\gamma;-\lambda)h_{12}(\lambda)\,d\lambda,$$

$$p_2(n) = n^{\beta}m^{\gamma+\delta-1}, \qquad \Xi_2(\gamma,\delta) = 2(2\pi)^{1-\gamma-\delta}h_{12}(0)\frac{\cos(\beta\pi/2)}{1-\gamma-\delta}.$$

THEOREM ABCU.

(i) If $\gamma + \delta > 1$

$$q_i(n) = n^{\beta}, \qquad F_i(\gamma,\delta,\nu,\eta) = \frac{\Xi_i(\gamma,\delta)}{\Upsilon(\delta)}, \qquad \frac{h_n}{n^{\beta}\log n} \to \infty, \qquad i = 1,2,$$

$$q_i(n) = n^{\beta}, \qquad F_i(\gamma,\delta,\nu,\eta) = \frac{\Xi_i(\gamma,\delta)}{\Upsilon(\delta)} - \nu\eta, \qquad h_n \sim n^{\beta}\log n, \qquad i = 1,2,$$

$$q_i(n) = h_n/\log n, \qquad F_i(\gamma,\delta,\nu,\eta) = -\nu\eta, \qquad \frac{h_n}{n^{\beta}\log n} \to 0, \qquad i = 1,2,$$

https://doi.org/10.1017/S0266466606060361 Published online by Cambridge University Press

where for any two sequences a_n, b_n, $\sim$ *denotes that* $a_n/b_n \to 1$ *as* $n \to \infty$*;*

(ii) *if* $\gamma + \delta = 1$

$$q_1(n) = n^{2\delta-1}/1_{\log n}, \qquad F_1(\gamma,\delta,\nu,\eta) = \frac{\Xi_1(\gamma,\delta)}{\Upsilon(\delta)}, \qquad \frac{h_n}{n^{2\delta-1}\log n 1_{\log^{-1} n}} \to \infty,$$

$$q_1(n) = n^{2\delta-1}/1_{\log n}, \qquad F_1(\gamma,\delta,\nu,\eta) = \frac{\Xi_1(\gamma,\delta)}{\Upsilon(\delta)} - \nu\eta,$$

$$h_n \sim n^{2\delta-1}\log n 1_{\log^{-1} n},$$

$$q_1(n) = h_n/\log n, \qquad F_1(\gamma,\delta,\nu,\eta) = -\nu\eta, \qquad \frac{h_n}{n^{2\delta-1}\log n 1_{\log^{-1} n}} \to 0,$$

$$q_2(n) = n^{2\delta-1}/1_{\log m}, \qquad F_2(\gamma,\delta,\nu,\eta) = \frac{\Xi_2(\gamma,\delta)}{\Upsilon(\delta)}, \qquad \frac{h_n}{n^{2\delta-1}\log n 1_{\log^{-1} m}} \to \infty,$$

$$q_2(n) = n^{2\delta-1}/1_{\log m}, \qquad F_2(\gamma,\delta,\nu,\eta) = \frac{\Xi_2(\gamma,\delta)}{\Upsilon(\delta)} - \nu\eta,$$

$$h_n \sim n^{2\delta-1}\log n 1_{\log^{-1} m},$$

$$q_2(n) = h_n/\log n, \qquad F_2(\gamma,\delta,\nu,\eta) = -\nu\eta, \qquad \frac{h_n}{n^{2\delta-1}\log n 1_{\log^{-1} m}} \to 0,$$

where $1_a = 1(\gamma = 0) + a1(\gamma > 0)$*;*

(iii) *if* $\gamma + \delta < 1$

$$q_1(n) = n^{2\delta-1}, \qquad F_1(\gamma,\delta,\nu,\eta) = \frac{\Xi_1(\gamma,\delta)}{\Upsilon(\delta)}, \qquad \frac{h_n}{n^{2\delta-1}\log n} \to \infty,$$

$$q_1(n) = n^{2\delta-1}, \qquad F_1(\gamma,\delta,\nu,\eta) = \frac{\Xi_1(\gamma,\delta)}{\Upsilon(\delta)} - \nu\eta, \qquad h_n \sim n^{2\delta-1}\log n,$$

$$q_1(n) = h_n/\log n, \qquad F_1(\gamma,\delta,\nu,\eta) = -\nu\eta, \qquad \frac{h_n}{n^{2\delta-1}\log n} \to 0,$$

$$q_2(n) = n^{\beta}m^{\gamma+\delta-1}, \qquad F_2(\gamma,\delta,\nu,\eta) = \frac{\Xi_2(\gamma,\delta)}{\Upsilon(\delta)}, \qquad \frac{h_n}{n^{\beta}m^{\gamma+\delta-1}\log n} \to \infty,$$

$$q_2(n) = n^{\beta}m^{\gamma+\delta-1}, \qquad F_2(\gamma,\delta,\nu,\eta) = \frac{\Xi_2(\gamma,\delta)}{\Upsilon(\delta)} - \nu\eta,$$

$$h_n \sim n^{\beta}m^{\gamma+\delta-1}\log n,$$

$$q_2(n) = h_n/\log n, \qquad F_2(\gamma,\delta,\nu,\eta) = -\nu\eta, \qquad \frac{h_n}{n^{\beta}m^{\gamma+\delta-1}\log n} \to 0.$$

THEOREM UU.

(i) If $2\delta + \theta > 1, \quad \delta + \theta > \frac{1}{2}$

$$r_i(n) = n^{\theta}, \quad \Theta_i(\theta,\delta) = \zeta' A(1)\Sigma^{1/2}\int_0^1 W(r;\delta)W'(r;\delta+\theta)\,dr\Sigma^{1/2}A'(1)\xi,$$

$$\Pi_i(\delta+\theta) = \xi' A(1)\Sigma^{1/2}\int_0^1 W(r;\delta+\theta)W'(r;\delta+\theta)\,dr\Sigma^{1/2}A'(1)\xi, \qquad i = 1,2;$$

(ii) if $2\delta + \theta > 1,\ 0 \le \delta + \theta < \frac{1}{2}$

$$r_1(n) = n^{1-(2\delta+\theta)}, \quad \Pi_1(\delta+\theta) = \int_{-\pi}^{\pi} |1-e^{i\lambda}|^{-2(\delta+\theta)} f_{22}(\lambda)\,d\lambda, \qquad \textbf{(A.2)}$$

$$r_2(n) = n^{\theta} m^{1-2(\delta+\theta)}, \quad \Pi_2(\delta+\theta) = \frac{2f_{22}(0)}{(2\pi)^{2(\delta+\theta)}(1-2(\delta+\theta))}, \qquad \textbf{(A.3)}$$

$$\Theta_i(\theta,\delta) = \zeta' A(1)\Sigma^{1/2}\int_0^1 W(r;\delta)\,dW'(r)\Sigma^{1/2}A'(1)\xi,$$

$$\delta + \theta = 0, \qquad i = 1,2; \qquad \textbf{(A.4)}$$

(iii) if $2\delta + \theta = 1,\ 0 < \delta + \theta < \frac{1}{2}$

$$r_1(n) = \log^{-1} n, \quad \Pi_1(\delta+\theta) = \int_{-\pi}^{\pi} |1-e^{i\lambda}|^{-2(\delta+\theta)} f_{22}(\lambda)\,d\lambda,$$

$$r_2(n) = \log^{-1} m \left(\frac{m}{n}\right)^{1-2(\delta+\theta)}, \quad \Pi_2(\delta+\theta) = \frac{2f_{22}(0)}{(2\pi)^{2(\delta+\theta)}(1-2(\delta+\theta))},$$

$$\Theta_i(\theta,\delta) = 2f_{12}(0)\sin(\delta\pi), \qquad i = 1,2;$$

(iv) if $2\delta + \theta = 1,\ \delta + \theta = 0$ *(or equivalently* $\delta = 1,\ \theta = -\delta$*)*

$$r_1(n) = 1, \quad \Pi_1(\delta+\theta) = \int_{-\pi}^{\pi} f_{22}(\lambda)\,d\lambda,$$

$$\Theta_1(\theta,\delta) = \zeta' A(1)\Sigma^{1/2}\int_0^1 W(r)\,dW'(r)\Sigma^{1/2}A'(1)\xi + \sum_{j=0}^{\infty} E(v_{20}v_{1,-j}),$$

$$r_2(n) = mn^{-1}, \quad \Pi_2(\delta+\theta) = 2f_{22}(0),$$

$$\Theta_2(\theta,\delta) = \zeta' A(1)\Sigma^{1/2}\int_0^1 W(r)\,dW'(r)\Sigma^{1/2}A'(1)\xi + \pi f_{12}(0);$$

https://doi.org/10.1017/S0266466606060361 Published online by Cambridge University Press

(v) if $2\delta + \theta < 1$, $\delta + \theta \geq 0$

$$r_1(n) = 1, \qquad \Pi_1(\delta + \theta) = \int_{-\pi}^{\pi} |1 - e^{i\lambda}|^{-2(\delta+\theta)} f_{22}(\lambda)\, d\lambda,$$

$$\Theta_1(\theta,\delta) = \int_{-\pi}^{\pi} a(\delta;\lambda) a(\delta + \theta;-\lambda) f_{12}(\lambda)\, d\lambda,$$

$$r_2(n) = \left(\frac{m}{n}\right)^{-\theta}, \qquad \Pi_2(\delta + \theta) = \frac{2f_{22}(0)}{(2\pi)^{2(\delta+\theta)}(1 - 2(\delta + \theta))},$$

$$\Theta_2(\theta,\delta) = \frac{2(2\pi)^{1-(2\delta+\theta)} f_{12}(0)\cos(\theta\pi/2)}{1 - (2\delta + \theta)}.$$

Alternative Representation of the Limiting Distributions of Theorem UU. Under the same conditions of Theorem UU, assuming $\omega_{12} \neq 0$, *as* $n \to \infty$,

$$r_i(n)\bar{\nu}_i \Rightarrow \nu \Psi_i(\delta,\theta), \qquad i = 1,2,$$

where $r_i(n)$ are the normalizing sequences given in Theorem UU and denoting $g_{ij}(\lambda) = f_{ij}(\lambda)/f_{ij}(0)$, $i,j = 1,2$,

(i) if $2\delta + \theta > 1$, $\delta + \theta > \frac{1}{2}$

$$\Psi_i(\delta,\theta) = \frac{(\rho^{-1}(1-\rho^2)^{1/2}, 1)\int_0^1 W(r;\delta)W'(r;\delta + \theta)\, dr\xi}{\xi' \int_0^1 W(r;\delta + \theta)W'(r;\delta + \theta)\, dr\xi}, \qquad i = 1,2;$$

(ii) if $2\delta + \theta > 1$, $\delta + \theta = 0$ (or equivalently $\delta > 1$, $\theta = -\delta$)

$$\Psi_1(\delta,\theta) = \frac{2\pi(\rho^{-1}(1-\rho^2)^{1/2}, 1)\int_0^1 W(r;\delta)\, dW'(r)\xi}{\int_{-\pi}^{\pi} g_{22}(\lambda)\, d\lambda},$$

$$\Psi_2(\delta,\theta) = \pi(\rho^{-1}(1-\rho^2)^{1/2}, 1)\int_0^1 W(r;\delta)\, dW'(r)\xi;$$

(iii) if $2\delta + \theta = 1$, $0 < \delta + \theta < \frac{1}{2}$

$$\Psi_1(\delta,\theta) = \frac{2\sin(\delta\pi)}{\int_{-\pi}^{\pi} |1 - e^{i\lambda}|^{-2(\delta+\theta)} g_{22}(\lambda)\, d\lambda},$$

$$\Psi_2(\delta,\theta) = (2\pi)^{2(\delta+\theta)}(1 - 2(\delta + \theta))\sin(\delta\pi);$$

https://doi.org/10.1017/S0266466606060361 Published online by Cambridge University Press

(iv) if $2\delta + \theta = 1$, $\delta + \theta = 0$ (or equivalently $\delta = 1$, $\theta = -\delta$)

$$\Psi_1(\delta,\theta) = \frac{2\pi(\rho^{-1}(1-\rho^2)^{1/2}, 1)\int_0^1 W(r)\,dW'(r)\xi + f_{12}^{-1}(0)\sum_{j=0}^{\infty} E(v_{20}v_{1,-j})}{\int_{-\pi}^{\pi} g_{22}(\lambda)\,d\lambda},$$

$$\Psi_2(\delta,\theta) = \pi(\rho^{-1}(1-\rho^2)^{1/2}, 1)\int_0^1 W(r)\,dW'(r)\xi + \frac{\pi}{2};$$

(v) if $2\delta + \theta < 1$, $\delta + \theta \geq 0$

$$\Psi_1(\delta,\theta) = \frac{\int_{-\pi}^{\pi} a(\delta;\lambda)a(\delta+\theta;-\lambda)g_{12}(\lambda)\,d\lambda}{\int_{-\pi}^{\pi} |1-e^{i\lambda}|^{-2(\delta+\theta)} g_{22}(\lambda)\,d\lambda},$$

$$\Psi_2(\delta,\theta) = \frac{(2\pi)^{1+\theta}(1-2(\delta+\theta))\cos(\theta\pi/2)}{1-(2\delta+\theta)}.$$

THEOREM UCD.

$$s_i(n) = p_{i-2}(n), \qquad G_i(\gamma,\delta,\nu,\Lambda) = \Xi_{i-2}(\gamma,\delta)/\Upsilon(\delta), \qquad i = 3,4;$$

(i) If $\gamma + \delta > 1$

$$s_i(n) = n^\beta, \qquad G_i(\gamma,\delta,\nu,\Lambda) = \frac{\Xi_i(\gamma,\delta)}{\Upsilon(\delta)}, \qquad \frac{g_n}{n^\beta \log n} \to \infty, \qquad i = 5,6,$$

$$s_i(n) = n^\beta, \qquad G_i(\gamma,\delta,\nu,\Lambda) = \frac{\Xi_i(\gamma,\delta)}{\Upsilon(\delta)} + \nu\Lambda, \qquad g_n \sim n^\beta \log n, \qquad i = 5,6,$$

$$s_i(n) = g_n/\log n, \qquad G_i(\gamma,\delta,\nu,\Lambda) = \nu\Lambda, \qquad \frac{g_n}{n^\beta \log n} \to 0, \qquad i = 5,6;$$

(ii) if $\gamma + \delta = 1$

$$s_5(n) = n^{2\delta-1}/1_{\log n}, \qquad G_5(\gamma,\delta,\nu,\Lambda) = \frac{\Xi_1(\gamma,\delta)}{\Upsilon(\delta)}, \qquad \frac{g_n}{n^{2\delta-1}\log n 1_{\log^{-1} n}} \to \infty,$$

$$s_5(n) = n^{2\delta-1}/1_{\log n}, \qquad G_5(\gamma,\delta,\nu,\Lambda) = \frac{\Xi_1(\gamma,\delta)}{\Upsilon(\delta)} + \nu\Lambda,$$

$$g_n \sim n^{2\delta-1}\log n 1_{\log^{-1} n},$$

$$s_5(n) = g_n/\log n, \qquad G_5(\gamma,\delta,\nu,\Lambda) = \nu\Lambda, \qquad \frac{g_n}{n^{2\delta-1}\log n 1_{\log^{-1} n}} \to 0,$$

$$s_6(n) = n^{2\delta-1}/1_{\log m}, \qquad G_6(\gamma,\delta,\nu,\Lambda) = \frac{\Xi_2(\gamma,\delta)}{\Upsilon(\delta)}, \qquad \frac{g_n}{n^{2\delta-1}\log n 1_{\log^{-1} m}} \to \infty,$$

$$s_6(n) = n^{2\delta-1}/1_{\log m}, \qquad G_6(\gamma,\delta,\nu,\Lambda) = \frac{\Xi_2(\gamma,\delta)}{\Upsilon(\delta)} + \nu\Lambda,$$

$$g_n \sim n^{2\delta-1}\log n 1_{\log^{-1} m},$$

$$s_6(n) = g_n/\log n, \qquad G_6(\gamma,\delta,\nu,\Lambda) = \nu\Lambda, \qquad \frac{g_n}{n^{2\delta-1}\log n 1_{\log^{-1} m}} \to 0,$$

where $1_a = 1(\gamma = 0) + a1(\gamma > 0)$*;*

(iii) if $\gamma + \delta < 1$

$$s_5(n) = n^{2\delta-1}, \qquad G_5(\gamma,\delta,\nu,\Lambda) = \frac{\Xi_1(\gamma,\delta)}{\Upsilon(\delta)}, \qquad \frac{g_n}{n^{2\delta-1}\log n} \to \infty,$$

$$s_5(n) = n^{2\delta-1}, \qquad G_5(\gamma,\delta,\nu,\Lambda) = \frac{\Xi_1(\gamma,\delta)}{\Upsilon(\delta)} + \nu\Lambda, \qquad g_n \sim n^{2\delta-1}\log n,$$

$$s_5(n) = g_n/\log n, \qquad G_5(\gamma,\delta,\nu,\Lambda) = \nu\Lambda, \qquad \frac{g_n}{n^{2\delta-1}\log n} \to 0,$$

$$s_6(n) = n^{\beta} m^{\gamma+\delta-1}, \qquad G_6(\gamma,\delta,\nu,\Lambda) = \frac{\Xi_2(\gamma,\delta)}{\Upsilon(\delta)}, \qquad \frac{g_n}{n^{\beta} m^{\gamma+\delta-1}\log n} \to \infty,$$

$$s_6(n) = n^{\beta} m^{\gamma+\delta-1}, \qquad G_6(\gamma,\delta,\nu,\Lambda) = \frac{\Xi_2(\gamma,\delta)}{\Upsilon(\delta)} + \nu\Lambda,$$

$$g_n \sim n^{\beta} m^{\gamma+\delta-1}\log n,$$

$$s_6(n) = g_n/\log n, \qquad G_6(\gamma,\delta,\nu,\Lambda) = \nu\Lambda, \qquad \frac{g_n}{n^{\beta} m^{\gamma+\delta-1}\log n} \to 0.$$

APPENDIX B: Proofs of Theorems

Proof of Theorem BNCUD. First, we show (28). Now

$$\bar{\nu}_5 - \bar{\nu}_1 = \frac{\sum_{t=1}^{n} y_t(x_t(\hat{\theta}) - x_t)\sum_{t=1}^{n} x_t^2 - \sum_{t=1}^{n} x_t y_t \sum_{t=1}^{n}(x_t^2(\hat{\theta}) - x_t^2)}{\sum_{t=1}^{n} x_t^2(\hat{\theta})\sum_{t=1}^{n} x_t^2},$$

https://doi.org/10.1017/S0266466606060361 Published online by Cambridge University Press

so that, in view of Theorems 4.4 and 5.1 of Robinson and Marinucci (2001), (28) holds on showing

$$\sum_{t=1}^{n} y_t(x_t(\hat{\theta}) - x_t) = O_p(n^{2\delta} g_n^{-1} \log n), \tag{B.1}$$

$$\sum_{t=1}^{n} (x_t^2(\hat{\theta}) - x_t^2) = O_p(n^{2\delta} g_n^{-1} \log n). \tag{B.2}$$

First, as in Robinson and Hualde (2003), by Taylor's theorem, for some $R > 1$, the left side of (B.1) is

$$\sum_{t=1}^{n} y_t \sum_{j=0}^{t-1} \sum_{r=1}^{R-1} \frac{(-\hat{\theta})^r}{r!} a_j^{(r)}(\delta) v_{2,t-j} + \sum_{t=1}^{n} y_t \sum_{j=0}^{t-1} \frac{(-\hat{\theta})^R}{R!} a_j^{(R)}(\bar{\delta}) v_{2,t-j}, \tag{B.3}$$

where

$$a_j^{(r)}(b) = \frac{d^r}{db^r} a_j(b)$$

and $\bar{\delta}$ is an intermediate point between δ and $\delta - \hat{\theta}$. Now, in view of Lemmas D.1 and D.5 of Robinson and Hualde (2003) and Assumption PE, choosing R large enough, the dominant term in (B.3) is the first one, so that (B.1) follows from minor modifications of Theorems 4.4 and 5.1 of Robinson and Marinucci (2001), where the only difference is that the weights $a_j^{(r)}(\delta)$ are not covered by those of Robinson and Marinucci (2001) because of the presence of log factors but they just contribute the log factor in (B.1) (see, e.g., Bingham, Goldie, and Teugels, 1989, Cor. 1.7.3). The proof of (B.2) is identical to that of (B.1), to conclude for (28). Finally, (29) follows from minor modifications of Propositions 4.1 and 4.2 of Robinson and Marinucci (2003). ■

Proof of Theorem ABCU. We give only the proof for $\bar{\nu}_1$, the proof for $\bar{\nu}_2$ being almost identical in view of Propositions 4.1 and 4.2 of Robinson and Marinucci (2003). Now

$$\bar{\nu}_1 - \nu = \frac{\sum_{t=1}^{n} u_{1t}(-\gamma) v_{2t}(-\delta)}{\sum_{t=1}^{n} v_{2t}^2(-\delta)} + \nu \frac{\sum_{t=1}^{n} (v_{2t}(-\delta) - x_{t,n}) x_{t,n}}{\sum_{t=1}^{n} x_{t,n}^2} \tag{B.4}$$

$$+ \frac{\sum_{t=1}^{n} u_{1t}(-\gamma) x_{t,n}}{\sum_{t=1}^{n} x_{t,n}^2} - \frac{\sum_{t=1}^{n} u_{1t}(-\gamma) v_{2t}(-\delta)}{\sum_{t=1}^{n} v_{2t}^2(-\delta)}. \tag{B.5}$$

The first term on the right-hand side of (B.4) is the usual one appearing with strictly balanced orders, whose asymptotic behavior was discussed in Theorem BCU. Next, as in the proof of Theorem BNCUD, by Taylor's theorem and (32), it is readily seen that the dominant term on the second term on the right-hand side of (B.4) is

https://doi.org/10.1017/S0266466606060361 Published online by Cambridge University Press

$$-\nu\theta_n \frac{\sum_{t=1}^{n} v_{2t}(-\delta)\sum_{j=0}^{t-1} a_j^{(1)}(\delta)v_{2,t-j}}{\sum_{t=1}^{n} v_{2t}^2(-\delta)}.$$

Taking derivatives in (2),

$$\sum_{j=0}^{t-1} a_j^{(1)}(\delta)v_{2,t-j} = \sum_{j=1}^{t-1} \psi(j+\delta)a_j(\delta)v_{2,t-j} - \psi(\delta)\sum_{j=1}^{t-1} a_j(\delta)v_{2,t-j}, \tag{B.6}$$

where $\psi(x)$ is the digamma function

$$\psi(x) = \frac{d}{dx}\log\Gamma(x).$$

By Taylor's theorem, the first term on the right-hand side of (B.6) is equal to

$$\sum_{j=1}^{t-1}\sum_{r=0}^{R-1} \frac{\psi^{(r)}(j)}{r!}\delta^r a_j(\delta)v_{2,t-j} + \sum_{j=1}^{t-1} \frac{\psi^{(R)}(j+\bar{\delta})}{R!}\delta^R a_j(\delta)v_{2,t-j}, \tag{B.7}$$

where $\psi^{(r)}(x)$ represents the rth derivative of the digamma function

$$\psi^{(r)}(x) = \frac{d^r}{dx^r}\psi(x),$$

with the convention $\psi^{(0)}(x) = \psi(x)$ and $j < j+\bar{\delta} < j+\delta$. Noting that as in the proof of Lemma D.1 of Robinson and Hualde (2003), for $l \geq 1$,

$$|\psi^{(l)}(x)| \leq K(1+x)^{-l},$$

the first term in the expansion (B.7) is the dominant one. By (6.3.21) in Abramowitz and Stegun (1970, p. 259), for $j > 0$

$$\psi(j) = \log j - \frac{1}{2j} - 2\int_0^\infty \frac{tdt}{(t^2+j^2)(e^{2\pi t}-1)}. \tag{B.8}$$

The absolute value of the third term is bounded by

$$K\int_0^\infty \frac{tdt}{(t^2+j^2)e^{2\pi t}} \leq \frac{K}{j}\int_0^\infty \frac{dt}{e^{2\pi t}} \leq \frac{K}{j},$$

implying from (B.8) that

$$\psi(j) = \log j + O(j^{-1}).$$

Thus, the first term in (B.6) is

$$\sum_{j=1}^{t-1} \log j a_j(\delta)v_{2,t-j} + O_p\left(\sum_{j=1}^{t-1} j^{-1}|a_j(\delta)v_{2,t-j}|\right). \tag{B.9}$$

By Marinucci and Robinson (2000), the first term in (B.9) is $O_p(t^{\delta-1/2}\log t)$, whereas, by Stirling's approximation, the second one is $O_p(1+\log t 1(\delta=1)+t^{\delta-1}1(\delta>1))$. In addition, by Marinucci and Robinson (2000) the second term in (B.6) is $O_p(t^{\delta-1/2})$, so that by Marinucci and Robinson (2000) and the continuous mapping theorem

$$\frac{\sum_{t=1}^{n} v_{2t}(-\delta)\sum_{j=1}^{t-1} a_j^{(1)}(\delta)v_{2,t-j}}{\log n\sum_{t=1}^{n} v_{2t}^2(-\delta)} \to_p 1.$$

Finally, noting that due to (32), (B.5) is in all cases of smaller order than the first term on the right-hand side of (B.4), we conclude the proof of the theorem. ■

Proof of Theorem UU. First, (i) follows from direct application of Theorem 1 of Marinucci and Robinson (2000), the continuous mapping theorem, and Theorems 4.4 and 5.1 of Robinson and Marinucci (2001). Next, under (ii), x_t is asymptotically stationary, and defining

$$\tilde{x}_t = \sum_{j=0}^{\infty} a_j(\delta+\theta)v_{2,t-j},$$

which is that covariance stationary version of x_t, it is straightforward to show that under our conditions

$$n^{-1}\sum_{t=1}^{n}\tilde{x}_t^2 \to_p E(\tilde{x}_t^2) = \int_{-\pi}^{\pi}|1-e^{i\lambda}|^{-2(\delta+\theta)}f_{22}(\lambda)\,d\lambda$$

(see, e.g., Hualde and Robinson, 2006). By the Cauchy–Schwarz inequality, the proof of (A.2) would be complete on showing

$$\sum_{t=1}^{n}(\tilde{x}_t - x_t)^2 = o_p(n). \tag{B.10}$$

The expectation on the left-hand side of (B.10) is

$$\sum_{t=1}^{n}\int_{-\pi}^{\pi}\sum_{j=t}^{\infty}\sum_{k=t}^{\infty} a_j(\delta+\theta)a_k(\delta+\theta)e^{i(j-k)\lambda}f_{22}(\lambda)\,d\lambda$$

$$\leq K\sum_{t=1}^{n}\int_{-\pi}^{\pi}\left|\sum_{j=t}^{\infty} a_j(\delta+\theta)e^{ij\lambda}\right|^2 \leq K\sum_{t=1}^{n}\sum_{j=t}^{\infty} j^{2(\delta+\theta)-2} \leq Kn^{2(\delta+\theta)},$$

as $\delta+\theta<\frac{1}{2}$, so (B.10) holds, to complete the proof of (A.2). Next, (A.3) follows on showing

$$\frac{1}{n^{2(\delta+\theta)}m^{1-2(\delta+\theta)}}\sum_{j=0}^{m} c_j I_x(\lambda_j) \to_p \frac{2f_{22}(0)}{(2\pi)^{2(\delta+\theta)}(1-2(\delta+\theta))}.$$

Now,

$$\sum_{j=0}^{m} c_j I_x(\lambda_j) = 2\sum_{j=1}^{m} I_{\tilde{x}}(\lambda_j) + 2\sum_{j=1}^{m} (I_x(\lambda_j) - I_{\tilde{x}}(\lambda_j)) + I_x(0). \tag{B.11}$$

Related to the first term, by Theorem 1 of Robinson (1994),

$$\frac{\dfrac{2\pi}{n}\displaystyle\sum_{j=1}^{m} I_{\tilde{x}}(\lambda_j)}{\displaystyle\int_0^{\lambda_m} |1-e^{i\lambda}|^{-2(\delta+\theta)} f_{22}(\lambda)\,d\lambda} \to_p 1,$$

where

$$\int_0^{\lambda_m} |1-e^{i\lambda}|^{-2(\delta+\theta)} f_{22}(\lambda)\,d\lambda \sim \frac{f_{22}(0)\lambda_m^{1-2(\delta+\theta)}}{1-2(\delta+\theta)},$$

where the symbol $\sim$ indicates here that the ratio of left- and right-hand sides tends to 1 as λ tends to 0. Thus,

$$\frac{2}{n^{2(\delta+\theta)} m^{1-2(\delta+\theta)}} \sum_{j=1}^{m} I_{\tilde{x}}(\lambda_j) \to_p \frac{2f_{22}(0)}{(2\pi)^{2(\delta+\theta)}(1-2(\delta+\theta))}, \tag{B.12}$$

noting that when $\delta+\theta=0$, (B.12) reflects the standard result for $I(0)$ processes (see, e.g., Brockwell and Davis, 1991, Thm. 10.4.1),

$$m^{-1}\sum_{j=1}^{m} I_{\tilde{x}}(\lambda_j) \to_p f_{22}(0).$$

Next, the second term in (B.11) is equal to

$$2\sum_{j=1}^{m} |w_x(\lambda_j) - w_{\tilde{x}}(\lambda_j)|^2 + 4\,\mathrm{Re}\left\{\sum_{j=1}^{m} (w_x(\lambda_j) - w_{\tilde{x}}(\lambda_j)) w_{\tilde{x}}(-\lambda_j)\right\}. \tag{B.13}$$

By the Cauchy–Schwarz inequality, the second term is bounded by

$$K\left(\sum_{j=1}^{m} |w_x(\lambda_j) - w_{\tilde{x}}(\lambda_j)|^2 \sum_{j=1}^{m} I_{\tilde{x}}(\lambda_j)\right)^{1/2}.$$

By previous arguments

$$\sum_{j=1}^{m} I_{\tilde{x}}(\lambda_j) = O_p(n^{2(\delta+\theta)} m^{1-2(\delta+\theta)}),$$

and by Robinson (2005),

$$\sum_{j=1}^{m} E|w_x(\lambda_j) - w_{\tilde{x}}(\lambda_j)|^2 \le K \sum_{j=1}^{m} j^{-1}\lambda_j^{-2(\delta+\theta)} \le Kn^{2(\delta+\theta)} \sum_{j=1}^{m} j^{-1-2(\delta+\theta)}$$

$$\le K \log m, \qquad \delta + \theta = 0,$$

$$\le Kn^{2(\delta+\theta)}, \qquad \delta + \theta > 0,$$

so the second term of (B.13) is $O_p(m^{1/2}\log^{1/2} m)$ if $\delta + \theta = 0$, or $O_p(n^{2(\delta+\theta)} m^{1/2-(\delta+\theta)})$ if $\delta + \theta > 0$, so that it is of smaller order than the first on the right-hand side of (B.11). Similarly, the first term of (B.13) is of smaller order. Finally, the third term on the right-hand side of (B.11) is

$$\frac{1}{2\pi n}\left(\sum_{t=1}^{n} x_t\right)^2 = O_p(n^{2(\delta+\theta)}),$$

by Marinucci and Robinson (2000), also of smaller order, to conclude the proof of (A.3).

Finally, (A.4) holds as (A.1), and the results corresponding to (iii), (iv), and (v) are straightforward applications of previous arguments and results in Robinson and Marinucci (2001). ■