EFFICIENT ESTIMATION OF COINTEGRATING RELATIONSHIPS AMONG HIGHER ORDER AND FRACTIONALLY INTEGRATED PROCESSES

Juan J. Dolado and Francesc Marmol

Banco de España - Servicio de Estudios
Documento de Trabajo nº 9617

EFFICIENT ESTIMATION OF COINTEGRATING RELATIONSHIPS AMONG HIGHER ORDER AND FRACTIONALLY INTEGRATED PROCESSES

Juan J. Dolado (*)
and Francesc Marmol (**)

(*) Research Department, Bank of Spain.
(**) Department of Economics, Universitat Autònoma de Barcelona.

We are very grateful to Manuel Arellano, Grayham Mizon, Mark Salmon and Enrique Sentana for helpful comments and suggestions. We also wish to thank seminar audiences at the Bank of Spain, CEMFI, Madrid, European University Institute, Florence, Universitat Autònoma de Barcelona, Universidad de Alicante and Universitat Pompeu Fabra. Usual disclaimers apply.

Banco de España - Servicio de Estudios
Documento de Trabajo nº 9617

In publishing this series the Banco de España seeks to disseminate studies of interest that will help acquaint readers better with the Spanish economy.

The analyses, opinions and findings of these papers represent the views of their authors; they are not necessarily those of the Banco de España.

ISSN: 0213-2710
ISBN: 84-7793-490-8
Depósito legal: M. 20868-1996
Imprenta del Banco de España

Abstract: In this paper we address the issue of the efficient estimation of the cointegrating vector in linear regression models with variables that follow general (higher order and fractionally) integrated processes. We prove that, when the underlying processes are formed by higher order *I*(*d*) integrated processes, then a standard Fully-Modified (FM-OLS) estimation procedure as the one proposed by Phillips and Hansen (1990) only yields (asymptotically) efficient estimates of the cointegrating vector when $d = 1$. For $d = 2, 3, \ldots$, a simple FM-OLS estimator is proposed which just entails correcting for the endogeneity bias. When dealing with nonstationary $(d > 1/2)$ fractionally integrated *FI*(*d*) processes which are fractionally cointegrated, i.e., with the equilibrium error evolving as a $FI(\delta)$, with $d > \delta$, then the latter comment applies for all the assumed range of the memory parameter d *if* $\delta < 1/2$, in which case, we propose a ***fractional fully modified OLS*** estimator, denoted as FFM-OLS. Otherwise, the OLS estimator weakly converges to a random variable having a law that cannot be made gaussian with a FM-OLS procedure. Finally, we also study the consequences of applying the standard semiparametric FM-OLS estimator for cointegrated $I(1)$ variables when the true order of integration is *I*(*d*) or *FI*(*d*). We show that, under those more general cases, the limit distribution of the standard FM-OLS estimator is no longer mixed normal, loosing its optimal properties.

Keywords: Cointegration; fully modified estimation; higher order integrated processes; fractionally integrated processes; stochastic integral; misspecification.

J.E.L. Classification: C12, C15, C22.

1 Introduction

When estimating the cointegrating vector of linear regression models with $I(1)$ variables, it is well known that the OLS estimator in a static regression is found to be *super-consistent* $\left(\text{i.e., } O_p\left(T^{-1}\right)\right)$ under quite general assumptions, including endogeneity in the regressors and serial correlation in the innovations (see Stock, 1987). However, the performance of the OLS estimator is adversely affected by the existence of serial correlation and endogeneity biases that do not affect its consistency but introduce non-zero means and non-normalities in the limiting distribution of the standardized statistics, except in some special cases. Such biases can play an important role in finite samples, as shown in the simulations of Banerjee et al. (1986). To overcome these problems, Phillips and Hansen (1990) proposed a semi-parametric correction of the OLS estimator, denoted as Fully Modified estimator (henceforth FM-OLS), which is asymptotically equivalent to maximum likelihood and yields median-unbiased and asymptotically normal estimates, so that conventional techniques for inference are valid.

However, confining the analysis of efficient estimation in a single-equation framework to the case of $I(1)$ variables might be restrictive for at least two reasons. First, despite the fact that many economic time series are empirically characterized as $I(1)$ processes, there are other variables, especially nominal ones such as the price level or the money stock (in logarithms), that seem better described as $I(2)$ processes. These $I(2)$ variables and, in general, higher order $I(d)$ processes lead to new interesting problems such as the existence of multicointegrating or polynomially cointegrating relationships (see, e.g., Granger and Lee, 1989, 1990, Gregoir and Laroque, 1994 and Haldrup and Salmon, 1995). The FM-OLS estimation with $I(2)$ processes has been recently developed by Chang and Phillips (1995). Herein we address the issue of the efficient estimation in a single-equation framework in the general $I(d)$ case. Secondly, the analysis of higher order $I(d)$ processes is not the only way to generalize the results in the unit-root literature. Cointegration requires the equilibrium error to be mean-reverting. Yet, for the equilibrium error to have such a property it does not need to be $I(0)$ exactly. Fractionally integrated processes, denoted as $FI(d)$, also display mean-reversion in some cases. Therefore, the associated concept of fractional cointegration (see, for instance, Cheung and Lai, 1993 and Baillie and Bollerslev, 1994), by avoiding the knife-edged unit-root versus no unit-root distinction in the equilibrium error, allows for a wider range of mean-reversion than standard cointegration analysis. In light of the above comments, this paper also attempts to reexamine the issue of the efficient estimation of cointegrating vectors in the presence of fractionally integrated $FI(d)$ processes.

The paper is organized as follows. In Section 2, we study the behaviour of the FM-OLS estimation method when the data generating process (DGP) is assumed to be formed by higher order $I(d)$ processes that are cointegrated in such a way that they give rise to a stationary $I(0)$ error term, whereas Section 3 is devoted to exploring the derivation of a FM-OLS estimator when the DGP is formed by (nonstationary) $FI(d)$ processes and the equilibrium error is possibly mean-reverting but not necessarily $I(0)$. Section 4 is concerned with a robustness analysis of the behaviour of the standard FM-OLS estimator for $I(1)$ variables, as formulated by Phillips and Hansen (1990), when the true order of integration of the variables in the DGP is $I(d)$ or $FI(d)$. Some concluding comments are provided in Section 5. Finally, some technical material is gathered in the Appendix.

The notation follows Phillips and Hansen (1990). Therefore, the symbols "$\Rightarrow$", "$\xrightarrow{p}$" and "$\equiv$" denote weak convergence, convergence in probability and equality in distribution, respectively, $[\cdot]$ denotes "integer part" and the inequality ">0" denotes positive-definite when applied to matrices. Brownian motion $B(r)$, with $r \in [0,1]$, is frequently written as B for notational simplicity. Similarly, we write integrals with respect to Lebesgue measure such as $\int_0^1 B(r)dr$ more simply as $\int B$. The symbol $\sum_{t=1}^{T}$ is denoted simply as $\sum$. Vector Brownian motion with covariance matrix Ω is written $BM(\Omega)$. We use $\|A\|$ to represent the Euclidean norm $tr(A'A)^{1/2}$ of the matrix A. Finally, all limits given in the paper are as the sample size $T \to \infty$ unless otherwise stated.

2 FM-OLS Estimation with Higher-Order I(d) Processes

In this section we shall be working with an n-dimensional vector y_t partitioned as

$$y_t = \left(y_{1t}, y_{2t}'\right)' \tag{1}$$

where y_{1t} is a scalar and y_{2t} is an m-vector ($m+1=n$), and generated according to the system

$$y_{1t} = \alpha + \beta' y_{2t} + \varepsilon_{1t}, \tag{2}$$

$$\Delta^d y_{2t} = \varepsilon_{2t} \tag{3}$$

with $d = 1,2, ..,$and where the d initial values, $y_0, ..., y_{1-d}$, have been set equal to zero without loss of generality. The restriction to single-equation models is unimportant, the generalization to system estimation with known cointegrating rank being straightforward and thus omitted. Equally, the model can be easily extended to cover situations in which the elements of y_{2t} have different orders of integration and are possibly multicointegrated. Further deterministic components in (3), besides a constant term, are omitted for simplicity, without affecting the main results of the paper. With respect to the innovation sequence $\varepsilon_t = \left(\varepsilon_{1t}, \varepsilon_{2t}'\right)'$, we shall assume that it is stationary and ergodic with zero mean, finite covariance matrix $\Xi > 0$, continuous density matrix $f_{\varepsilon\varepsilon}(\lambda)$ and long-run covariance matrix $\Omega = 2\pi f_{\varepsilon\varepsilon}(0)$. We also require the partial sum process constructed from $\{\varepsilon_t\}_{t=1}^{\infty}$ to satisfy a multivariate invariance principle

$$T^{-1/2}\sum_{t=1}^{[Tr]} \varepsilon_t \Rightarrow B(r) \equiv BM(\Omega),$$

where $B(r)$, $r \in [0,1]$, is an n-dimensional Brownian motion with covariance matrix Ω assumed to be positive definite implying that the regressors y_{2t} are not allowed to be cointegrated among themselves. Let us partition Ω and $B(r)$ conformably with ε_t

$$\Omega = \begin{pmatrix} \omega_{11} & \omega_{21}' \\ \omega_{21} & \Omega_{22} \end{pmatrix} \qquad B = \begin{pmatrix} B_1 \\ B_2 \end{pmatrix},$$

and decompose the long-run covariance matrix Ω as

$$\Omega = \Xi + \Lambda + \Lambda'$$

where $\Xi = E(\varepsilon_0\varepsilon_0'), \Lambda = \sum_{k=1}^{\infty} E(\varepsilon_0\varepsilon_k')$, and define $\Delta = \Xi + \Lambda$. These matrices are again partitioned conformably with ε_t. Let $\hat{\alpha}$ and $\hat{\beta}$ be estimates based on OLS estimation of (2) with a sample of size T

$$\begin{pmatrix} \hat{\alpha} \\ \hat{\beta} \end{pmatrix} = \begin{pmatrix} T & \sum y_{2t}' \\ \sum y_{2t} & \sum y_{2t}y_{2t}' \end{pmatrix}^{-1} \begin{pmatrix} \sum y_{1t} \\ \sum y_{2t}y_{1t} \end{pmatrix} \tag{4}$$

so that the deviations of the OLS estimators in (4) from the population values α and β that describe the cointegrating relation (2) are given by the expression

$$\begin{pmatrix} \hat{\alpha}-\alpha \\ \hat{\beta}-\beta \end{pmatrix} = \begin{pmatrix} T & \sum y_{2t}' \\ \sum y_{2t} & \sum y_{2t}y_{2t}' \end{pmatrix}^{-1} \begin{pmatrix} \sum \varepsilon_{1t} \\ \sum y_{2t}\varepsilon_{1t} \end{pmatrix}. \tag{5}$$

Let us next define the weight matrix

$$\Im_T = diag\{T^{1/2}, T^d I_m\}, \tag{6}$$

which, in turn, implies that the OLS equation (5) could be rewritten as

$$\Im_T \begin{pmatrix} \hat{\alpha}-\alpha \\ \hat{\beta}-\beta \end{pmatrix} = \left(\Im_T^{-1} \begin{pmatrix} T & \sum y_{2t}' \\ \sum y_{2t} & \sum y_{2t}y_{2t}' \end{pmatrix} \Im_T^{-1} \right)^{-1} \Im_T^{-1} \begin{pmatrix} \sum \varepsilon_{1t} \\ \sum y_{2t}\varepsilon_{1t} \end{pmatrix}. \tag{7}$$

Now, following Sims et al. (1990) and applying the continuous mapping theorem, CMT henceforth, (see Billingsley, 1968), it is straightforward to prove the following result.

__Theorem__ 1. *Under the assumptions made on the disturbances, the OLS estimation of the conditional model* (2) *yields*

$$\begin{pmatrix} T^{1/2}(\hat{\alpha}-\alpha) \\ T^d(\hat{\beta}-\beta) \end{pmatrix} \Rightarrow \begin{pmatrix} 1 & \int (B_2^d)' \\ \int B_2^d & \int B_2^d (B_2^d)' \end{pmatrix}^{-1} \begin{pmatrix} B_1(1) \\ \Theta \end{pmatrix}, \tag{8}$$

where

$$\Theta = \begin{cases} \int B_2 dB_1 + \Delta_{21} & if\ d = 1 \\ \int B_2^d dB_1 & if\ d = 2,3,\ldots, \end{cases} \tag{9}$$

and $B^d(r)$ *denotes the* (d-1)*-fold integral of* $B(r)$ *recursively defined as* $B^d(r) = \int_0^r B^{d-1}(s)ds$, *with* $B^1(r) = B(r)$.

Note that the OLS estimator of the slope coefficient β in the cointegrating vector is $O_p(T^{-d})$ consistent. However, the presence of the nuisance parameters Δ_{21} and ω_{21} in the limiting OLS distribution prevents achieving an asymptotic mixture of normals. Indeed, Park and Phillips (1988, Lemma 5.1) proved that asymptotic gaussianity applies when variables are $CI(1,1)$ and $\omega_{21} = \Delta_{21} = 0$, i.e., the case when the conditioning variables are strictly exogenous. The same result can be extended to the general $CI(d,d)$ case by a straightforward application of their Lemma 5.1. This is a very convenient case, since, under asymptotic gaussianity, valid inference can be conducted using standard distributions.

When $d = 1$, the difference between the distribution derived in Theorem 1 and the convenient special case where asymptotic gaussianity applies is due to the presence of both Δ_{21} and ω_{21} nuisance parameters. On the one hand, $\omega_{21} \neq 0$ implies that B_1 and B_2 are not long-run independent giving rise to an endogeneity bias. On the other hand, $\Delta_{21} \neq 0$ causes the so-called *serial correlation or second-order bias* effect. Although none of these biases affect the consistency properties of the OLS estimator, they can be important in finite samples. In turn, when $d > 1$, Theorem 1 shows that the second-order bias term Δ_{21} is no longer present.

As is well known, in the case when $d = 1$, Phillips and Hansen (1990) have proposed a semi-parametric correction to the unadjusted OLS estimators, which eliminates the previous biases and achieve asymptotic gaussianity. This method, known as FM-OLS, is asymptotically equivalent to performing maximum likelihood estimation. In what follows, we will make use of the result in Theorem 1 to extend the FM-OLS estimation procedure to the more general case where variables are $I(d)$, $d = 1,2,..$

An important feature of the FM-OLS method is that it relies upon the use of a consistent estimator of the long-run covariance matrix Ω. While any consistent estimator of this matrix will produce the same asymptotic distributions, Phillips and Hansen (1990) were concerned with a specific class of kernel estimators. In particular, letting

$\hat{\varepsilon}_t = (\hat{\varepsilon}_{1t} \quad \hat{\varepsilon}_{2t})'$, with $\hat{\varepsilon}_{1t}$ being the least squares residual from (2), then the class of positive semidefinite kernel estimators of Ω they considered is given by

$$\hat{\Omega} = \sum_{j=-M}^{M} \ell(j/M) T^{-1} \sum_t \hat{\varepsilon}_{t-j} \hat{\varepsilon}_t', \tag{10}$$

where the kernel weights $\ell(\cdot)$ satisfy that for all $x \in \Re, |\ell(x)| \leq 1$ and $\ell(x) = \ell(-x)$, $\ell(0) = 1$, $\ell(x)$ is continuous at zero, for almost all $x \in \Re$ $\int_{\Re} |\ell(x)| dx < \infty$ and for all $\lambda \in \Re$, $\int_{-\infty}^{\infty} \ell(x) exp(-ix\lambda) \geq 0$. Kernels that satisfy these requirements include Truncated, Barlett, Parzen, Tuckey-Hanning and Quadratic Spectral kernels (e.g. see Hannan, 1970 and Priestley, 1981). Throughout this paper we shall assume the same class of kernel estimates. Equally, the following kernel-based estimator of the one-sided long-run covariance matrix can be defined as

$$\hat{\Delta} = \sum_{j=0}^{M} \ell(j/M) T^{-1} \sum_t \hat{\varepsilon}_{t-j} \hat{\varepsilon}_t', \tag{11}$$

Then, under some regularity conditions[1] on the bandwidth parameter, M, and the remaining assumptions on the disturbances it can be proved how the consistency of the kernel estimators of the long-run covariance matrices to their theoretical counterparts also holds for the general $I(d)$ case. For instance, if we assume the following bandwidth condition

$$M \to \infty \quad \text{as} \quad T \to \infty \quad \text{such that} \quad T^{-1/2} M \to 0. \tag{B0}$$

then we can prove the consistency of the term $\hat{\omega}_{21}$ to the corresponding theoretical counterpart as follows. Given that

$$\hat{\omega}_{21} = \sum_{j=-M}^{M} \ell(j/M) T^{-1} \sum \hat{\varepsilon}_{2,t-j} \hat{\varepsilon}_{1t} = \sum_{j=-M}^{M} \ell(j/M) T^{-1} \sum \varepsilon_{2,t-j} \hat{\varepsilon}_{1t}$$

$$= \sum_{j=-M}^{M} \ell(j/M) T^{-1} \sum \varepsilon_{2,t-j} \varepsilon_{1t} - \sum_{j=-M}^{M} \ell(j/M) T^{-1} \sum \varepsilon_{2,t-j} (\hat{\pi} - \pi)' x_t$$

$$= \wp_{1T} - \wp_{2T} \quad \text{(say)},$$

[1] We refer the reader to Andrews (1991), Chang and Phillips (1995) and Phillips (1991, 1995) for a detailed account of these regularity conditions.

where $\pi' = (\alpha,\ \beta')$ and $x_t' = (1,\ y_{2t}')$, then, from Andrews (1991), it follows that

$$M^{-1}T^{1/2}(\wp_{1T} - \omega_{21}) \xrightarrow{p} 0. \tag{B1}$$

As regards the $\wp_{2T}$ term, we have that

$$\left\| M^{-1}T^{1/2}\wp_{2T} \right\| \le M^{-1}\sum_{j=-M}^{M} \left| \ell\left(j/M\right) \right| \left\| T^{-1}\sum \varepsilon_{2,t-j} T^{1/2}(\hat{\pi} - \pi)x_t \right\|$$

$$\le \left(\int_{-\infty}^{\infty} \left| \ell(x) \right| dx \right) \left(T^{-1}\sum \varepsilon_{2t}'\varepsilon_{2t} \right)^{1/2} \Gamma^{1/2} = O_p(1). \tag{B2}$$

where $\Gamma = (\hat{\pi} - \pi)' \Im_T \Im_T^{-1} \sum x_t x_t' \Im_T^{-1} \Im_T (\hat{\pi} - \pi) = O_p(1)$. Thus, (B1) and (B2) imply that $M^{-1}T^{1/2}(\hat{\omega}_{21} - \omega_{21}) = O_p(1)$ and, from (B0), we finally get $\hat{\omega}_{21} \xrightarrow{p} \omega_{21}$.

Let us now consider the case where $d > 1$. From Theorem 1 we can see that the bias term Δ_{21} is no longer present. Therefore, in order to achieve asymptotic gaussianity we should only correct for the bias stemming from $\omega_{21} \neq 0$. So, let us define the *endogeneity bias-corrected* ε_{1t} disturbance

$$\varepsilon_{1t}^{+} = \varepsilon_{1t} - \omega_{12}\Omega_{22}^{-1}\Delta^d y_{2t} = \varepsilon_{1t} - \omega_{12}\Omega_{22}^{-1}\varepsilon_{2t},$$

which has zero coherence at the origin with ε_{2t}. In this case, we can write $(\varepsilon_{1t}^{+}\quad \varepsilon_{2t}')' = Q'(\varepsilon_{1t}\quad \varepsilon_{2t}')'$, where

$$Q = \begin{pmatrix} 1 & -\omega_{12}\Omega_{22}^{-1} \\ 0 & I_m \end{pmatrix} = \begin{pmatrix} Q_1' \\ Q_2' \end{pmatrix},$$

being Q_1' of dimension $(1 \times n)$ and Q_2' of dimension $(m \times n)$. Now subtracting $\omega_{12}\Omega_{22}^{-1}\Delta^d y_{2t}$ from both sides of (2), yields

$$y_{1t}^{+} = \alpha + \beta' y_{2t} + \varepsilon_{1t}^{+}, \tag{12}$$

where $y_{1t}^{+} = y_{1t} - \omega_{12}\Omega_{22}^{-1}\Delta^d y_{2t}$. In this case, the FM-OLS estimator equals the OLS estimator of the parameters in (12), yielding

$$\begin{pmatrix} \hat{\alpha}^{+} \\ \hat{\beta}^{+} \end{pmatrix} = \begin{pmatrix} T & \sum y_{2t}' \\ \sum y_{2t} & \sum y_{2t}y_{2t}' \end{pmatrix}^{-1} \begin{pmatrix} \sum y_{1t}^{+} \\ \sum y_{2t}y_{1t}^{+} \end{pmatrix},$$

or, proceeding as in equation (7),

$$\Im_T\begin{pmatrix}\hat{\alpha}^+ - \alpha\\ \hat{\beta}^+ - \beta\end{pmatrix} = \left(\Im_T^{-1}\begin{pmatrix}T & \sum y_{2t}'\\ \sum y_{2t} & \sum y_{2t}y_{2t}'\end{pmatrix}\Im_T^{-1}\right)^{-1}\Im_T^{-1}\begin{pmatrix}\sum \hat{\varepsilon}_{1t}^+\\ \sum y_{2t}\hat{\varepsilon}_{1t}^+\end{pmatrix},$$

where the corrected disturbance term ε_{1t}^+ has been replaced by $\hat{\varepsilon}_{1t}^+ = \varepsilon_{1t} - \hat{\omega}_{12}\hat{\Omega}_{22}^{-1}\varepsilon_{2t}$ in order to derive feasible FM-OLS estimators. Then, we have the following result.

__Theorem 2.__ *Under the assumptions made on the disturbances and on the kernel estimators, the FM-OLS estimation of the conditional model* (12) *yields*

$$\begin{pmatrix}T^{1/2}(\hat{\alpha}^+ - \alpha)\\ T^d(\hat{\beta}^+ - \beta)\end{pmatrix} \Rightarrow \begin{pmatrix}1 & \int (B_2^d)'\\ \int B_2^d & \int B_2^d(B_2^d)'\end{pmatrix}^{-1}\begin{pmatrix}B_1^+(1)\\ \int B_2^d dB_1^+\end{pmatrix} \tag{13}$$

$$\equiv \int_{\varsigma>0} N(0,\varsigma)dP(\varsigma),$$

where $B_1^+(r) \equiv BM(\omega_{11}^+)$, *with* $\omega_{11}^+ = \omega_{11} - \omega_{12}\Omega_{22}^{-1}\omega_{21}$, *and*

$$\varsigma = \begin{pmatrix}1 & \int (B_2^d)'\\ \int B_2^d & \int B_2^d(B_2^d)'\end{pmatrix}^{-1}.$$

PROOF. Define

$$\hat{Q} = \begin{pmatrix}1 & -\hat{\omega}_{12}\hat{\Omega}_{22}^{-1}\\ 0 & I_m\end{pmatrix}$$

and note that, under the assumptions made, $\hat{Q}' \xrightarrow{p} Q'$, so that

$$\begin{pmatrix}\hat{\varepsilon}_{1t}^+\\ \varepsilon_{2t}\end{pmatrix} = \hat{Q}'\begin{pmatrix}\varepsilon_{1t}\\ \varepsilon_{2t}\end{pmatrix} \xrightarrow{p} Q'\begin{pmatrix}\varepsilon_{1t}\\ \varepsilon_{2t}\end{pmatrix} = \begin{pmatrix}\varepsilon_{1t}^+\\ \varepsilon_{2t}\end{pmatrix}.$$

having a long-run covariance matrix given by

$$\Omega^+ = Q'\Omega Q = \begin{pmatrix}\omega_{11}^+ & 0'\\ 0 & \Omega_{22}\end{pmatrix},$$

where ω_{11}^+ has been defined in the text of the theorem.

Being $\varepsilon_t^+ = (\varepsilon_{1t}^+ \quad \varepsilon_{2t}^{'})' = Q'(\varepsilon_{1t} \quad \varepsilon_{2t}^{'})'$ a finite linear combination of the original innovation vector, the CMT holds for the corrected innovations so that

$$T^{-1/2}\sum \varepsilon_t^+ = T^{-1/2}\sum Q'\varepsilon_t \Rightarrow B^+(r) \equiv Q'B(r) \equiv BM(\Omega^+).$$

Now, partitioning B^+ and Ω^+ conformably with ε_t^+, the first part of the theorem follows by the same arguments as in Theorem 1. With respect to the gaussian properties, they are implied by the fact that B_1^+ and $B_2^+ \equiv B_2$ are independent processes so that Lemma 5.1 in Park and Phillips (1988) applies when conditioning on the σ-field generated by these stochastic processes. ■

The limiting distribution obtained in this Theorem is now full ranked, median-unbiased and a mixture of normals. Both FM-OLS estimators $\hat{\alpha}^+$ and $\hat{\beta}^+$ are consistent and their limiting distributions are free of nuisance parameters. Hence, conventional asymptotic procedures for inference can be applied. For instance, consider the usual Wald form of the chi-squared test of q restrictions on the cointegrating slope coefficients of the form

$$H_0: \quad R\beta = r,$$

where R is a $(q \times m)$ known matrix such that $rank(R) = q$ and r is a $(q \times 1)$ known vector. Define the Wald statistic constructed from $\hat{\beta}^+$ by

$$\xi = \left(R\hat{\beta}^+ - r\right)'\left\{\hat{\omega}_{11}^+(0 \quad R)\begin{pmatrix} T & \sum y_{2t}^{'} \\ \sum y_{2t} & \sum y_{2t}y_{2t}^{'} \end{pmatrix}^{-1}\begin{pmatrix} 0' \\ R' \end{pmatrix}\right\}^{-1}\left(R\hat{\beta}^+ - r\right).$$

Therefore, we have that, under the null hypothesis, the Wald statistic can be rewritten as follows

$$\xi = \left[RT^d\left(\hat{\beta}^+ - \beta\right)\right]'\left(\hat{\omega}_{11}^+\right)^{-1}\left\{(0 \quad R)\left(\Im_T^{-1}\begin{pmatrix} T & \sum y_{2t}^{'} \\ \sum y_{2t} & \sum y_{2t}y_{2t}^{'} \end{pmatrix}\Im_T^{-1}\right)^{-1}\begin{pmatrix} 0' \\ R' \end{pmatrix}\right\}^{-1}\left[RT^d\left(\hat{\beta}^+ - \beta\right)\right]$$

so that from Theorem 2 it immediately follows that $\xi \Rightarrow \chi^2_{(q)}$, a chi-squared distribution with q degrees of freedom[2]. In the particular case when we wish to use a single coefficient test

$$H_0: \; \beta_i = \beta_i^0,$$

then we can construct the following modified t-statistic:

$$t_{\beta_i} = \frac{\hat{\beta}_i^+ - \beta_i^0}{\left(\hat{\omega}_{11}^+\right)^{1/2} Z_{ii}^{-1/2}} \equiv N(0,1),$$

where Z_{ii} denotes the iith-component of the second-moment matrix of the regressors .

Finally, note that, when $d = 1$, it can be easily proved that

$$T^{-1}\sum y_{2t}\hat{\varepsilon}_{1t}^+ \xrightarrow{p} T^{-1}\sum y_{2t}\varepsilon_{1t}^+ \Rightarrow \int B_2 dB_1^+ + \Delta_{21}^+,$$

where Δ_{21}^+ would be the corresponding submatrix of the corrected one-sided long-run covariance matrix

$$\Delta^+ = \sum_{k=0}^{\infty} E\left(\varepsilon_0^+ \varepsilon_k^+\right),$$

with $\varepsilon_t^+ = \left(\varepsilon_{1t}^+, \; \varepsilon_{2t}\right)'$. Therefore, in this case, efficient estimators of the cointegrating relationships should not only take account of the endogeneity bias, as when $d > 1$, but should also correct for the second-order bias term Δ_{21}^+. As in the previous analysis, derivation of a feasible FM-OLS estimator is based on the following (kernel-based) estimator of the $\hat{\Delta}_{21}^+$ term

$$\hat{\Delta}_{21}^+ = \sum_{j=0}^{M} \ell\left(j/_M\right) T^{-1} \sum \varepsilon_{2,t-j} \hat{\varepsilon}_{1t}^+,$$

so that the feasible FM-OLS estimator will be now

$$\begin{pmatrix} \hat{\alpha}^{++} - \alpha \\ \hat{\beta}^{++} - \beta \end{pmatrix} = \begin{pmatrix} T & \sum y_{2t}' \\ \sum y_{2t} & \sum y_{2t} y_{2t}' \end{pmatrix}^{-1} \begin{pmatrix} \sum \hat{\varepsilon}_{1t}^+ \\ \sum y_{2t}\hat{\varepsilon}_{1t}^+ - T\hat{\Delta}_{21}^+ \end{pmatrix}.$$

[2] When we consider multicointegrating relationships, it is convenient to restrict inference to tests of separable restrictions. Therefore, the restrictions matrix R must be block-diagonal across the components of β which are related with integrated processes of different orders. This is due to the fact that, in this case, the corresponding FM-OLS estimators will converge at different rates, implying the possibility of rank defficiencies. See Haldrup (1994) for more detailed comments.

This is the standard FM-OLS formula derived in the seminal paper by Phillips and Hansen (1990), which has the same mixed normal and parameter invariant limit distribution than we obtained in expression (13).

3 Efficient Estimation with Fractionally Integrated Processes

In this section we extend the previous results to the case where d is not an integer but a real number. In particular we will consider general autoregressive fractionally integrated moving average processes of order d, henceforth denoted $ARFIMA(p, d, q)$, defined as

$$\alpha(B)\Delta^d y_t = \zeta(B)\varepsilon_t,$$

with $\varepsilon_t \sim IID(0, \sigma^2)$, where $\alpha(B)$ and $\zeta(B)$ are autoregressive and moving average lag polynomials, respectively, with roots lying outside the unit circle. The memory parameter, d, is now consider to be any real number so that the fractional difference operator Δ^d can be expressed in terms of a Maclaurin expansion as

$$\Delta^d = (1-B)^d = \sum_{k=0}^{\infty} \frac{\Gamma(k-d)B^k}{\Gamma(k+1)\Gamma(-d)} = \sum_{k=0}^{\infty} \pi_k B^k, \quad \pi_k = \frac{k-1-d}{k}\pi_{k-1} \quad \pi_0 = 1,$$

with $\Gamma(\cdot)$ being the gamma function.

These processes have received an increasing attention because of their ability to provide a flexible and natural characterization of nonstationary behaviours, nesting the $I(d)$ model as a special and potentially restrictive case. It can be proved that the process is both stationary and invertible if $-1/2 < d < 1/2$. In spite of being nonstationary, it is mean-reverting with transitory memory if $d < 1$, in contrast with the case when $d \geq 1$, where the process is both nonstationary and not mean-reverting with permanent memory. Finally, it is stationary with short-memory if $d < 0$, whereas it is stationary with long-memory if $0 < d < 1/2$ and as such may be expected to be useful in modelling long-term persistence. When $d = 0$, the process is white noise, with zero correlations and constant spectral density; c.f., for instance, Granger and Joyeux (1980), Hosking (1981), Gourieroux and Monfort (1990), Brockwell and Davis (1991) and Cheung and Lai

(1993).[3] Moreover, these processes are not strong-mixing (e.g., Helson and Sarason, 1967 and Viano et al., 1995).

When working with nonstationary fractionally integrated processes, the standard notion of cointegration can be extended in a natural way and we can define a vector $y_t = (y_{1t}, y_{2t})'$ of *FI(d)* processes to be fractionally cointegrated of order b, denoted as *FCI(d,b)*, with $d \geq b > 0$, if there exists a linear combination between y_{1t} and y_{2t} which is fractionally integrated of order d-b. Cheung and Lai (1993), have argued that, under fractional cointegration, the OLS estimator of the corresponding cointegrating vector will be $O_p(T^{-b})$ consistent. Their proof of this result, however, is rather heuristic, without explicitly deriving the asymptotic OLS distribution. Thus, a preliminary step in this section will be to derive a formal proof of this claim.

For the sake of simplicity, we will concentrate on the following bivariate DGP, similar to that considered by Cheung and Lai (1993):

$$y_{1t} = \beta y_{2t} + u_{1t}, \tag{14a}$$

$$\alpha(B)\Delta^d y_{2t} = \zeta(B)\varepsilon_{2t}, \tag{14b}$$

$$\phi(B)\Delta^{d-b} u_{1t} = \vartheta(B)\varepsilon_{1t}, \tag{14c}$$

$$\begin{pmatrix} \varepsilon_{1t} \\ \varepsilon_{2t} \end{pmatrix} \sim WN\left\{\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \Xi \equiv \begin{pmatrix} \sigma_{11} & \sigma_{12} \\ & \sigma_{22} \end{pmatrix}\right\}, \tag{14d}$$

where $d > 1/2, d \geq b > 0$, $\phi(B)$ and $\alpha(B)$ are finite autoregressive polynomials and $\vartheta(B)$ and $\zeta(B)$ are finite moving average polynomials. All these polynomials in the lag operator have their roots outside the unit circle without sharing common roots to accomplish the identification requirements. Thus, from (14b)-(14d), we can derive the corresponding long-run covariance matrix

$$\Omega = \Phi(1)^{-1}\Theta(1)\Xi\Theta(1)'\Phi(1)^{-1}$$

[3] A process is said to have *permanent memory* if the effect of any random shock on the series has a permanent effect. Conversely, the process is said to have *transitory memory*, if the effect of any random shock on the series has only a temporary influence.

We say that a stationary process is *short-memory* if it has autocorrelations that decay at an exponential rate, whereas it is *long-memory* if its autocorrelations die out at the slower hyperbolic rate.

$$= \begin{pmatrix} \sigma_{11}\vartheta(1)^2\phi(1)^{-2} & \sigma_{12}\vartheta(1)\zeta(1)\phi(1)^{-1}\alpha(1)^{-1} \\ & \sigma_{22}\zeta(1)^2\alpha(1)^{-2} \end{pmatrix} \equiv \begin{pmatrix} \omega_{11} & \omega_{12} \\ & \omega_{22} \end{pmatrix}, \qquad (14e)$$

where $\Phi(B) = Diag\{\phi(B), \alpha(B)\}$ and $\Theta(B) = Diag\{\vartheta(B), \zeta(B)\}$ Let $\hat{\beta}$ be the OLS estimator of β in (14a), and denote by $\delta \equiv d - b \geq 0$ the memory parameter of the equilibrium error.

In order to derive the asymptotic distribution of the OLS estimator, it is convenient to distinguish between the following cases:

3.1. *Case* $\delta > 1/2$.

First, consider the situation where the equilibrium error is nonstationary, i.e., $\delta > 1/2$ (but possibly mean-reverting if $\delta < 1$). In this case, let us follow the approach developed by Akonom and Gourieroux (1988) and Gourieroux et al. (1989). In this way, and defining $z_t = (u_{1t}, y_{2t})'$, we have that

$$D_T z_{[Tr]} \Rightarrow z_\infty(r) \equiv \int_0^r E(r-s)dB(s), \qquad (15a)$$

and

$$F_T\left(\sum z_t z_t'\right)F_T \Rightarrow \int z_\infty z_\infty', \qquad (15b)$$

where $D_T = Diag\{T^{1/2-\delta}, T^{1/2-d}\}$, $E(r-s) = Diag\{\Gamma(\delta)^{-1}(r-s)^{\delta-1}, \Gamma(d)^{-1}(r-s)^{d-1}\}$ and $F_T = Diag\{T^{-\delta}, T^{-d}\}$. For this convergence to hold, we must assume (Akonom and Gourieroux, 1988) that the $\varepsilon_t = (\varepsilon_{1t}, \varepsilon_{2t})'$ sequence has moments of order strictly greater than $max\{2,\ d - 1/2\}$. Here, $B(r) = (B_1(r), B_2(r))'$ is a 2-dimensional Brownian motion with long-run covariance matrix Ω.

Therefore, from (15b), we have that

$$T^{-2d}\sum y_{2t}^2 \Rightarrow \int y_{2\infty}^2 \qquad T^{-d-\delta}\sum y_{2t}u_{1t} \Rightarrow \int y_{2\infty}u_{1\infty},$$

where $y_{2\infty}$ and $u_{1\infty}$ are fractional Brownian motions defined as follows

$$y_{2\infty}(r) = \Gamma(d)^{-1}\int_0^r (r-s)^{d-1} dB_2(s),$$

and

$$u_{1\infty}(r) = \Gamma(\delta)^{-1}\int_0^r (r-s)^{\delta-1} dB_1(s)$$

Thus, using these results and the CMT, we get

$$T^{b}(\hat{\beta}-\beta) = \frac{T^{-d-\delta}\sum y_{2t}u_{1t}}{T^{-2d}\sum y_{2t}^2} \Rightarrow \frac{\int y_{2\infty}u_{1\infty}}{\int y_{2\infty}^2}. \tag{16}$$

Consequently, the OLS estimator is $O_p(T^{-b})$ consistent, as argued by Cheung and Lai (1993), and its limiting distribution is given in (16). In order to examine its properties, let us proceed by using the Choleski decomposition of the long-run covariance matrix given by

$$\Omega = LL', \tag{17}$$

with

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

which implies that

$$\begin{pmatrix} B_1(r) \\ B_2(r) \end{pmatrix} = L\begin{pmatrix} W_1(r) \\ W_2(r) \end{pmatrix}$$

where $W(r) = (W_1(r),\ W_2(r))' \equiv BM(I_2)$, so that

$$y_{2\infty}(r) = l_{22}\Gamma(d)^{-1}\int_0^r (r-s)^{d-1} dW_2(s) = l_{22}\tilde{y}_{2\infty}(r), \tag{18a}$$

and

$$u_{1\infty}(r) = \frac{l_{11}}{\Gamma(\delta)}\int_0^r (r-s)^{\delta-1} dW_1(s) + \frac{l_{12}}{\Gamma(\delta)}\int_0^r (r-s)^{\delta-1} dW_2(s) \equiv l_{11}\ddot{u}_{1\infty}(r) + l_{12}\ddot{y}_{2\infty}(r), \tag{18b}$$

from which expression (16) becomes

$$T^{b}(\hat{\beta}-\beta) \equiv \frac{l_{11}\int \tilde{y}_{2\infty}\ddot{u}_{1\infty}}{l_{22}\int \tilde{y}_{2\infty}^2} + \frac{l_{12}\int \tilde{y}_{2\infty}\ddot{y}_{2\infty}}{l_{22}\int \tilde{y}_{2\infty}^2}. \tag{19a}$$

Thus, in this case, the OLS distribution is no longer mixed normal, even assuming strict exogeneity ($l_{12} = 0$). This is so since, in this latter case, the limiting distribution of the OLS estimator becomes

$$T^{b}\left(\hat{\beta}-\beta\right) \equiv \frac{l_{11}\int \tilde{y}_{2\infty}\ddot{u}_{1\infty}}{l_{22}\int \tilde{y}_{2\infty}^{2}}, \tag{19b}$$

which is a ratio of two Lebesgue integrals of fractional Brownian motions and, consequently, cannot be expressed as a mixture of normals. In particular, notice that, when $d = 1$ and $\delta = 1$ (so that $b = 0$), expression (19b) reduces to that found by Phillips (1986, Theorem 1) in his classical study on spurious regressions among independent random walks, and that when $d = \delta = 1, 2, \ldots$, we get the higher order generalization of the spurious phenomenon analyzed by Marmol (1995, 1996).

Equally, in the general case when $l_{12} \neq 0$, it is straightforward to show that the customary OLS t-statistic has the following limiting distribution:

$$T^{-1/2}t_{\beta} \Rightarrow \frac{\int y_{2\infty}u_{1\infty}}{\left\{\int y_{2\infty}^{2}\int u_{1\infty}^{2} - \left(\int y_{2\infty}u_{1\infty}\right)^{2}\right\}^{1/2}},$$

so that we have proved the important result that the t-statistic *always diverges* (*at the rate* $O_p\left(T^{1/2}\right)$) when the disturbance is a *nonstationary* fractionally integrated process, irrespectively of the existence or not of any cointegrating relationship.

Lastly, notice that expression (19b) would be the argument, a, which minimizes the following (continuous time) least squares criteria:

$$\int\left[\ddot{u}_{1\infty}(r) - a\tilde{y}_{2\infty}(r)\right]^{2}dr.$$

In summary, we have shown that, is spite of achieving consistency, in order to get standard limiting distributions, mean-reversion is a necessary but not a sufficient condition. In fact, as shown below, for the OLS estimator to have a mixed normal limiting distribution, we need that the memory parameter of the equilibrium error lies within the stationary range.

3.2. Case $0 \leq \delta < 1/2$.

Let us now assume that $0 \leq \delta < 1/2$, with the equilibrium error, u_{1t}, being a stationary and invertible process, so that $\sum y_{2t} u_{1t}$ corresponds to the sample cross moment between a nonstationary and a stationary fractionally integrated processes. In this case, it is more convenient (see Appendix A) to apply a different invariance principle than that applied in the subsection 3.1. Consequently, letting $\sigma_{uT}^2 = \omega_{11}^{-1} \operatorname{var}\left(\sum_{t=1}^{T} u_{1t}\right)$ and assuming that $E|\varepsilon_{1t}|^f < \infty$, for $f \geq max\{4, -8\delta/(1+2\delta)\}$, it can be proved (see Appendix A) that

$$T^{-1-2\delta} \sigma_{uT}^2 \xrightarrow{p} \frac{1}{\Gamma(1+\delta)^2 (2\delta+1)}, \quad \left(\equiv \theta_u^2, \text{ say}\right),$$

and

$$\sigma_{uT}^{-1} \sum_{j=1}^{[Tr]} u_{1j} \Rightarrow u_{1\infty}^{*}(r) \equiv \frac{1}{\Gamma(1+\delta)} \int_0^r (r-s)^{\delta} dB_1(s).$$

In turn, the fractional Brownian motion $u_{1\infty}^{*}$, by applying the Choleski decomposition in (17), can be decomposed into two independent components given by

$$u_{1\infty}^{*}(r) = \frac{l_{11}}{\Gamma(1+\delta)} \int_0^r (r-s)^{\delta} dW_1(s) + \frac{l_{12}}{\Gamma(1+\delta)} \int_0^r (r-s)^{\delta} dW_2(s) \equiv l_{11} \tilde{u}_{1\infty}(r) + l_{12} \tilde{u}_{1\infty}(r),$$

say. Now, notice that the y_{2t} process can be reparametrized in the following manner

$$\Delta^q y_{2t} = u_{2t} \tag{20a}$$

$$\alpha(B) \Delta^e u_{2t} = \zeta(B) \varepsilon_{2t}, \tag{20b}$$

with $d = q + e$ and where $q = 1, 2, \ldots$ is an integer number such that $e \in (-1/2, 1/2)$, i.e., with the u_{2t} series in (20b) being a stationary fractionally integrated process. Furthermore, we need to assume that the restriction $\delta + e > 0$ holds (see Appendix B).

In this case, and denoting $\sigma_{2T}^2 = \omega_{22}^{-1} \operatorname{var}\left(\sum_{t=1}^{T} u_{2t}\right)$ and $\theta_2^2 \equiv \Gamma(1+e)^{-2}(2e+1)^{-1}$, it is straightforward, from Appendix B, to prove the following result

$$T^{-d-\delta} \sum y_{2t} u_{1t} \Rightarrow \theta_2 \theta_u \int u_{2\infty}^{*q} du_{1\infty}^{*} \tag{21}$$

where $u_{2\infty}^{*q}$ stands for the $(q-1)$-fold integral of $u_{2\infty}^{*}$ recursively defined as $u_{2\infty}^{*q}(r) = \int_0^r u_{2\infty}^{*q-1}(s)ds$, and where

$$u_{2\infty}^{*}(r) \equiv \frac{1}{\Gamma(1+e)}\int_0^r (r-s)^e dB_2(s) = \frac{l_{22}}{\Gamma(1+e)}\int_0^r (r-s)^e dW_2(s) \equiv l_{22}\tilde{u}_{2\infty}^{*}. \quad (22)$$

In the same way, given that

$$T^{-2d}\sum y_{2t}^2 = T^{-2q-2e}\sum y_{2t}^2 = \left(T^{-1-2e}\sigma_{2T}^2\right)\left(T^{-1}\sum\left(T^{1-q}\sigma_{2T}^{-1}y_{2t}\right)^2\right),$$

we can apply the CMT and get

$$T^{-2d}\sum y_{2t}^2 \Rightarrow \theta_2^2\int\left(u_{2\infty}^{*q}\right)^2. \quad (23)$$

Hence, (21), (23) and the CMT yield

$$T^{d-\delta}\left(\hat{\beta}-\beta\right) = T^{b}\left(\hat{\beta}-\beta\right) = \frac{T^{-d-\delta}\sum y_{2t}u_{1t}}{T^{-2d}\sum y_{2t}^2} \Rightarrow \frac{\theta_u\int u_{2\infty}^{*q}du_{1\infty}^{*}}{\theta_2\int\left(u_{2\infty}^{*q}\right)^2}. \quad (24)$$

Therefore, except when the exogeneity assumption $\omega_{12}=0$ holds, we can see from expression (24) that the OLS estimation is consistent but that does not lead to a mixed normal asymptotic distribution due to the lack of independence of the fractional Brownian motions $u_{2\infty}^{*}$ and $u_{1\infty}^{*}$. Yet, we can proceed as in the previous section and define the following long-run bias-corrected u_{1t} equilibrium error

$$\hat{u}_{1t}^{+} = u_{1t} - \hat{\omega}_{12}\hat{\omega}_{22}^{-1}\Delta^{b}y_{2t}. \quad (25)$$

With this correction, we have that

$$\Delta^{\delta}\hat{u}_{1t}^{+} = \Delta^{\delta}u_{1t} - \hat{\omega}_{12}\hat{\omega}_{22}^{-1}\Delta^{d}y_{2t} = v_{1t} - \hat{\omega}_{12}\hat{\omega}_{22}^{-1}v_{2t} = \hat{v}_{1t}^{+},$$

where $v_{1t} = \phi(B)^{-1}\vartheta(B)\varepsilon_{1t}$ and $v_{2t} = \alpha(B)^{-1}\zeta(B)\varepsilon_{2t}$, so that

$$\text{var}\begin{pmatrix}\hat{v}_{1t}^{+}\\ v_{2t}\end{pmatrix} \xrightarrow{p} \text{var}\begin{pmatrix}v_{1t}^{+}\\ v_{2t}\end{pmatrix} = \Omega^{+} = \begin{pmatrix}\omega_{11}^{+} & 0\\ 0 & \omega_{22}\end{pmatrix},$$

where $\omega_{11}^{+} = \omega_{11} - \omega_{12}^2\omega_{22}^{-1}$. Consequently, a (feasible) *fractional fully modified* OLS estimator, denoted FFM-OLS, will be given by

$$\hat{\beta}^{+} = \frac{\sum y_{2t}\hat{y}_{1t}^{+}}{\sum y_{2t}^2}, \quad (26)$$

where $\hat{y}_{1t}^{+} = y_{1t} - \hat{\omega}_{12}\hat{\omega}_{22}^{-1}\Delta^{b}y_{2t}$, whose limiting distribution is

$$T^{b}\left(\hat{\beta}^{+}-\beta\right) \Rightarrow \frac{\theta_u \int u_{2\infty}^{*q} du_{1\infty}^{+}}{\theta_2 \int \left(u_{2\infty}^{*q}\right)^2}, \tag{27}$$

with

$$u_{1\infty}^{+}(r) \equiv \frac{1}{\Gamma(1+\delta)} \int_0^r (r-s)^{\delta} dB_1^{+}(s) = \frac{\left(\omega_{11}^{+}\right)^{1/2}}{\Gamma(1+\delta)} \int_0^r (r-s)^{\delta} dW_1(s) \equiv \left(\omega_{11}^{+}\right)^{1/2} \tilde{u}_{1\infty}^{*},$$

and $B_1^{+} \equiv BM\left(\omega_{11}^{+}\right)$. Next, by construction, it can be shown that the limiting distribution given in (27) is a $\left(O_p\left(T^{-b}\right)\right.$ consistent) mixture of normal variables (see Appendix B), and conventional inference analysis can be conducted in a standard way. In effect, let us define the following modified t-statistic:

$$t_{\beta} = \frac{\hat{\beta}^{+}-\beta}{\hat{\rho}_{\hat{\beta}}^{+}}, \tag{28}$$

with $\hat{\rho}_{\hat{\beta}} = \left(\hat{\omega}_{11}^{+}\right)^{1/2}\left(\sum y_{2t}^2\right)^{-1/2}$. In this case, we get

$$t_{\beta}^{+} \equiv T^{-\delta}\theta_u^{-1} t_{\beta} \Rightarrow \frac{\int u_{2\infty}^{*q} d\tilde{u}_{1\infty}^{*}}{\left(\int \left(u_{2\infty}^{*q}\right)^2\right)^{1/2}} \equiv N(0,1). \tag{29}$$

Consequently, from expression (29), we can deduce that only when $\delta = 0$, so that the equilibrium error u_{1t} is a standard *ARMA* process, the t-statistic t_{β} will have a well-defined standard gaussian distribution. Otherwise, when $0 < \delta < 1/2$, t_{β} will *diverge* as the sample size tends to infinity, over-rejecting the null hypothesis. Nevertheless, in the latter case, and assuming that the orders of integration δ and e are known, we can always define the standardized t-statistic t_{β}^{+} which is (asymptotically) distributed as $N(0,1)$.

4 Some Misspecification Analysis

In this last section, we want to investigate the consequences of applying the standard FM-OLS estimator, efficient when the underlying processes are $I(1)$ and the equilibrium error is $I(0)$, to series whose DGP departs from the previous assumptions.

4.1. *Higher order integrated processes*

To start with, let us consider first the same framework as in Section 2, i.e., a DGP composed by $I(d)$ processes, $d > 1$ with an $I(0)$ equilibrium error. Define

$$\overline{B}^d = B^d - \int B^d,$$

to be the demeaned Brownian motion, so that $\int \overline{B}^d(\overline{B}^d)' = \int B^d(B^d)' - \int B^d(\int B^d)'$, and $\overline{y}_{2t} = y_{2t} - (T^{-1}\sum y_{2t})i_m$, where i_m is an m-dimensional unitary vector. With this notation, we know, from Theorem 1, that the distribution of the OLS estimator of β in the cointegrating relation is given by

$$T^d(\hat{\beta} - \beta) \Rightarrow \left[\int \overline{B}_2^d(\int \overline{B}_2^d)'\right](\int \overline{B}_2^d dB_1), \tag{30}$$

whereas in Theorem 2 we proved that we can construct a FM-OLS estimator of β yielding an optimal mixed normal limiting distribution, whose expression is

$$T^d(\hat{\beta}^+ - \beta) \Rightarrow \left[\int \overline{B}_2^d(\int \overline{B}_2^d)'\right]^{-1}(\int \overline{B}_2^d dB_1^+).$$

Now an interesting exercise would be to examine the behaviour of the standard Phillips and Hansen's FM-OLS estimator under the previous DGP. For convenience, let us rewrite the necessary steps to construct such an estimator:

$$\hat{\beta}^* = \left(\sum \overline{y}_{2t}\overline{y}_{2t}'\right)^{-1}\left(\sum \overline{y}_{2t}\overline{y}_{1t}^* - T\hat{\Delta}_{\Delta 1}^*\right).$$

$$\Leftrightarrow T^d(\hat{\beta}^* - \beta) = \left(T^{-2d}\sum \overline{y}_{2t}\overline{y}_{2t}'\right)^{-1}\left(T^{-d}\sum \overline{y}_{2t}\hat{\varepsilon}_{1t}^* - T^{1-d}\hat{\Delta}_{\Delta 1}^*\right)$$

with

$$\hat{\varepsilon}_{1t}^* = \varepsilon_{1t} - \hat{\omega}_{1\Delta}\hat{\Omega}_{\Delta\Delta}^{-1}\Delta\tilde{y}_{2t},$$
$$\hat{\Delta}_{\Delta 1}^* = \hat{\Delta}_{\Delta 1} - \hat{\Delta}_{\Delta\Delta}\hat{\Omega}_{\Delta\Delta}^{-1}\hat{\omega}_{\Delta 1},$$

with the (kernel-based) estimators of the long-run covariances constructed as follows

$$\hat{\omega}_{ab} = \sum_{j=-M}^{M} \ell(j/M)\hat{\gamma}_{ab}(j)$$

and

$$\hat{\Delta}_{ab} = \sum_{j=0}^{M} \ell(j/M)\hat{\gamma}_{ab}(j),$$

where $\hat{\gamma}_{ab}(j) = T^{-1}\sum a_{t-j}b_t'$ for any pair of time series a_t and b_t, the symbol Δ as sub-index meaning Δy_{2t}. With this notation, a standard application of Sims et al.'s (1990) results, (30) and the CMT, yields the following results

$$T^{2-d}\hat{\gamma}_{1\Delta}(j) \Rightarrow \int dB_1(\overline{B}_2^{d-1})' - \left[\int dB_1(\overline{B}_2^{d})'\right]\left[\int \overline{B}_2^{d}(\overline{B}_2^{d})'\right]^{-1}\left[\int \overline{B}_2^{d}(\overline{B}_2^{d-1})'\right] + \Lambda_{12}(j), \quad (31a)$$

where

$$\Lambda_{12}(j) = \begin{cases} \sum_{k=1}^{\infty} E\left(\varepsilon_{10}\varepsilon_{2,k+j}'\right) & \text{if } d = 2 \\ 0 & \text{otherwise} \end{cases},$$

$$T^{3-2d}\hat{\gamma}_{\Delta\Delta}(j) = T^{2-2d}\sum \Delta\overline{y}_{2,t-j}\Delta\overline{y}_{2t}' \Rightarrow \int \overline{B}_2^{d-1}(\overline{B}_2^{d-1})', \quad (31b)$$

$$T^{2-d}\hat{\gamma}_{\Delta 1}(j) \Rightarrow \int \overline{B}_2^{d-1}dB_1 - \left[\int \overline{B}_2^{d-1}(\overline{B}_2^{d})'\right]\left[\int \overline{B}_2^{d}(\overline{B}_2^{d})'\right]^{-1}\int \overline{B}_2^{d}dB_1 + \Lambda_{21}(j), \quad (31c)$$

with

$$\Lambda_{21}(j) = \begin{cases} \sum_{k=1}^{\infty} E\left(\varepsilon_{20}\varepsilon_{1,k+j}\right) & \text{if } d = 2 \\ 0 & \text{otherwise} \end{cases},$$

and

$$T^{-2d}\sum \overline{y}_{2t}\overline{y}_{2t}' \Rightarrow \int \overline{B}_2^{d}(\overline{B}_2^{d})'. \quad (31d)$$

From (31a)-(31d), we get

$$M^{-1}T^{2-d}\hat{\omega}_{1\Delta} \Rightarrow \upsilon_1\left[\int dB_1(\overline{B}_2^{d-1})' - \left[\int dB_1(\overline{B}_2^{d})'\right]\left[\int \overline{B}_2^{d}(\overline{B}_2^{d})'\right]^{-1}\left[\int \overline{B}_2^{d}(\overline{B}_2^{d-1})'\right]\right] + \underbrace{\omega_{12}}_{\text{if } d=2}, \quad (32a)$$

(with $\omega_{12} = 0$ if $d > 2$),

$$M^{-1}T^{3-2d}\hat{\Omega}_{\Delta\Delta} \Rightarrow \upsilon_1\left[\int \overline{B}_2^{d-1}(\overline{B}_2^{d-1})'\right], \quad (32b)$$

$$M^{-1}T^{3-2d}\hat{\Delta}_{\Delta\Delta} \Rightarrow \upsilon_0\left[\int \overline{B}_2^{d-1}(\overline{B}_2^{d-1})'\right], \quad (32c)$$

$$M^{-1}T^{2-d}\hat{\omega}_{\Delta 1} \Rightarrow \upsilon_1\left[\int \overline{B}_2^{d-1}dB_1 - \left[\int \overline{B}_2^{d-1}(\overline{B}_2^{d})'\right]\left[\int \overline{B}_2^{d}(\overline{B}_2^{d})'\right]^{-1}\int \overline{B}_2^{d}dB_1\right] + \underbrace{\omega_{21}}_{\text{if } d=2}, \quad (32d)$$

and

$$M^{-1}T^{2-d}\hat{\Delta}_{\Delta 1} \Rightarrow \upsilon_0\left[\int \overline{B}_2^{d-1}dB_1 - \left[\int \overline{B}_2^{d-1}(\overline{B}_2^{d})'\right]\left[\int \overline{B}_2^{d}(\overline{B}_2^{d})'\right]^{-1}\int \overline{B}_2^{d}dB_1\right] + \underbrace{\Delta_{21}}_{\text{if } d=2}, \quad (32e)$$

where $\upsilon_1 = \int_{-1}^{1} \ell(x)dx$ and $\upsilon_0 = \int_0^1 \ell(x)dx$.

Using (32a)-(32e), yields

$$T^{-d}\sum \overline{y}_{2t}\hat{\varepsilon}_{1t}^{*} = T^{-d}\sum \overline{y}_{2t}\varepsilon_{1t} - T^{1-2d}\sum \overline{y}_{2t}\Delta\overline{y}_{2t}'\left(M^{-1}T^{3-2d}\hat{\Omega}_{\Delta\Delta}\right)^{-1}M^{-1}T^{2-d}\hat{\omega}_{\Delta 1}$$

$$\Rightarrow \int \overline{B}_2^{d}dB_1 - \left(\int \overline{B}_2^{d}(\overline{B}_2^{d-1})'\right)\left\{\upsilon_1\left[\int \overline{B}_2^{d-1}(\overline{B}_2^{d-1})'\right]\right\}^{-1} \times$$

$$\left\{\upsilon_1\left[\int \overline{B}_2^{d-1}dB_1 - \left[\int \overline{B}_2^{d-1}(\overline{B}_2^{d})'\right]\left[\int \overline{B}_2^{d}(\overline{B}_2^{d})'\right]^{-1}\int \overline{B}_2^{d}dB_1\right] + \underbrace{\omega_{21}}_{\text{if } d=2}\right\}, \quad (33a)$$

and

$$T^{1-d}\hat{\Delta}_{\Delta 1}^{*} = M^{-1}T^{2-d}\hat{\Delta}_{\Delta 1}MT^{-1} - M^{-1}T^{3-2d}\hat{\Delta}_{\Delta\Delta}MT^{-1}\left(M^{-1}T^{3-2d}\hat{\Omega}_{\Delta\Delta}\right)^{-1}M^{-1}T^{2-d}\hat{\omega}_{\Delta 1} \xrightarrow{p} 0.$$

(33b)

Lastly, (33a) and (33b) together imply

$$T^{d}\left(\hat{\beta}^{*} - \beta\right) \Rightarrow \left[\int \overline{B}_2^{d}(\overline{B}_2^{d})'\right]^{-1}\left\{\int \overline{B}_2^{d}dB_1 - \left(\int \overline{B}_2^{d}(\overline{B}_2^{d-1})'\right)\left\{\upsilon_1\left[\int \overline{B}_2^{d-1}(\overline{B}_2^{d-1})'\right]\right\}^{-1} \times\right.$$

$$\left.\upsilon_1\left[\int \overline{B}_2^{d-1}dB_1 - \left[\int \overline{B}_2^{d-1}(\overline{B}_2^{d})'\right]\left[\int \overline{B}_2^{d}(\overline{B}_2^{d})'\right]^{-1}\int \overline{B}_2^{d}dB_1\right] + \underbrace{\omega_{21}}_{\text{if } d=2}\right\},$$

which obviously is not mixed-normal. Therefore, even though the standard FM-OLS estimator of β remains consistent in this more general case, it looses its efficiency properties. Notice also that a second-order bias term, reflected now by ω_{21}, is also present when $d = 2$.

4.2. *Fractionally integrated processes*

Next, let us consider the fractional case. From the analysis in Section 3, we showed that only in the case where the equilibrium error evolves as a stationary and invertible fractional process, we can construct an efficient FM-OLS estimator. Hence, we shall only be concerned with the following bivariate DGP: Let $y_{1t}, y_{2t} \sim FI(d)$, $d > 1/2$, be fractionally cointegrated, i.e.,

$$y_{1t} = \beta y_{2t} + u_{1t},$$

where $u_{1t} \sim FI(\delta)$, $0 \leq \delta < 1/2$.

What happens if we use the standard FM-OLS estimator in this case? Note that this estimator is constructed under the assumption that $y_{1t}, y_{2t} \sim I(1)$ *and* that $u_{1t} \sim I(0)$. Therefore, the use of FM-OLS estimator now implies two sources of error, and not just one as in the misspecification analysis with $I(d), d > 1$, processes. Namely, as in previous case, the first source of error is that the variables are assumed to be $I(1)$ rather than *FI(d)*. However, in this case, there is a second source of error stemming from the assumption that the equilibrium error is taken to be $I(0)$ rather than $FI(\delta)$.

In order to derive the limiting distribution of the standard FM-OLS estimator under the assumption of fractional processes, let us define

$$\hat{y}_{1t}^{*} = y_{1t} - \hat{\omega}_{u\Delta}\hat{\omega}_{\Delta\Delta}^{-1}\Delta y_{2t}$$

so that

$$\hat{u}_{1t}^{*} = u_{1t} - \hat{\omega}_{u\Delta}\hat{\omega}_{\Delta\Delta}^{-1}\Delta y_{2t}.$$

Equally, define

$$\hat{\Delta}_{\Delta u}^{*} = \hat{\Delta}_{\Delta u} - \hat{\omega}_{u\Delta}\hat{\omega}_{\Delta\Delta}^{-1}\hat{\Delta}_{\Delta\Delta}.$$

The (feasible) standard FM-OLS would be given by the expression

$$\hat{\beta}^{*} = \frac{\sum \hat{y}_{1t}^{*} y_{2t} - T\hat{\Delta}_{\Delta u}^{*}}{\sum y_{2t}^{2}}$$

$$\Leftrightarrow \left(\hat{\beta}^{*} - \beta\right) = \frac{\sum \hat{u}_{1t}^{*} y_{2t} - T\hat{\Delta}_{\Delta u}^{*}}{\sum y_{2t}^{2}}.$$

As regards the numerator, we need to derive the asymptotic behaviour of the following sample correlations:

$$\hat{\gamma}_{u\Delta}(j) = T^{-1}\sum \hat{u}_{1,t-j}\Delta y_{2t} = T^{-1}\sum u_{1,t-j}\Delta y_{2t} - T^{-1}\left(\hat{\beta} - \beta\right)\sum y_{2,t-j}\Delta y_{2t}$$

and

$$\hat{\gamma}_{\Delta\Delta}(j) = T^{-1}\sum \Delta y_{2,t-j}\Delta y_{2t},$$

which, in turn, will depend on the order of the memory parameter d. Therefore, in view of the previous results, let us consider the following cases.

4.2.1. *Case* $d \geq 3/2$.

First, assume that $d \geq 3/2$ so that $q \geq 2$. In this range, we have that Δy_{2t} is a nonstationary fractionally integrated process, in which case, the following results hold

$$T^{1-2d}\sum y_{2,t-j}\Delta y_{2t} \Rightarrow \theta_2^2 \int u_{2\infty}^{*q}\Delta u_{2\infty}^{*q-1} \tag{34a}$$

$$T^{-d-\delta}\sum u_{1,t-j}y_{2t} \Rightarrow \theta_2\theta_u \int u_{2\infty}^{*q}du_{1\infty}^{*} \tag{34b}$$

$$T^{-d-\delta+1}\sum u_{1,t-j}\Delta y_{2t} \Rightarrow \theta_2\theta_u \int \Delta u_{2\infty}^{*q-1}du_{1\infty}^{*}, \tag{34c}$$

so that

$$T^{2-d-\delta}\hat{\gamma}_{u\Delta}(j) \Rightarrow \theta_2\theta_u\left[\int \Delta u_{2\infty}^{*q-1}du_{1\infty}^{*} - \frac{\int u_{2\infty}^{*q}du_{1\infty}^{*}\int u_{2\infty}^{*q}\Delta u_{2\infty}^{*q-1}}{\int\left(u_{2\infty}^{*q}\right)^2}\right] \equiv \eta \text{ (say)}, \tag{35a}$$

and

$$T^{3-2d}\hat{\gamma}_{\Delta\Delta}(j) \Rightarrow \theta_2^2\int\left(\Delta u_{2\infty}^{*q-1}\right)^2, \tag{35b}$$

with

$$\Delta u_{2\infty}^{*} = \frac{1}{\Gamma(e)}\int_0^r (r-s)^{e-1}dB_2(s).$$

Expressions (35a)-(35b), in turn, imply that

$$M^{-1}T^{2-d-\delta}\hat{\omega}_{u\Delta} \Rightarrow \upsilon_1\eta \tag{36a}$$

$$M^{-1}T^{2-d-\delta}\hat{\Delta}_{u\Delta} \Rightarrow \upsilon_0\eta \tag{36b}$$

$$M^{-1}T^{3-2d}\hat{\omega}_{\Delta\Delta} \Rightarrow \upsilon_1\theta_2^2\int\left(\Delta u_{2\infty}^{*q-1}\right)^2 \tag{36c}$$

and

$$M^{-1}T^{3-2d}\hat{\Delta}_{\Delta\Delta} \Rightarrow \upsilon_0\theta_2^2\int\left(\Delta u_{2\infty}^{*q-1}\right)^2. \tag{36d}$$

From (36a)-(36d) we get

$$\begin{aligned}T^{-d-\delta}\sum\hat{u}_{1t}^{*}y_{2t} &= T^{-d-\delta}\sum u_{1t}y_{2t} - T^{-d-\delta}\hat{\omega}_{u\Delta}\hat{\omega}_{\Delta\Delta}^{-1}\sum\Delta y_{2t}y_{2t}\\ &= T^{-d-\delta}\sum u_{1t}y_{2t} - M^{-1}T^{2-d-\delta}\hat{\omega}_{u\Delta}\left(M^{-1}T^{3-2d}\hat{\omega}_{\Delta\Delta}\right)^{-1}T^{1-2d}\sum\Delta y_{2t}y_{2t}\end{aligned}$$

$$\Rightarrow \theta_2\theta_u \int u_{2\infty}^{*q} du_{1\infty}^{*} - \eta \frac{\int u_{2\infty}^{*q} \Delta u_{2\infty}^{*q-1}}{\int \left(\Delta u_{2\infty}^{*q-1}\right)^2}, \tag{37a}$$

and

$$T^{-d-\delta}\left(T\hat{\Delta}_{\Delta u}^{*}\right) = MT^{-1}M^{-1}T^{1-d-\delta}\left(T\hat{\Delta}_{\Delta u}\right) - MT^{-1}M^{-1}T^{2-d-\delta}\hat{\omega}_{u\Delta}\left(M^{-1}T^{3-2d}\hat{\omega}_{\Delta\Delta}\right)^{-1}M^{-1}T^{3-2d}\hat{\Delta}_{\Delta\Delta}$$

so that

$$T^{-d-\delta}\left(T\hat{\Delta}_{\Delta u}^{*}\right) \xrightarrow{p} 0. \tag{37b}$$

Lastly, from (23) and (37a)-(37b), we obtain

$$T^{b}\left(\hat{\beta}^{*} - \beta\right) \Rightarrow \frac{\theta_u \int u_{2\infty}^{*q} du_{1\infty}^{*}}{\theta_2 \int \left(u_{2\infty}^{*q}\right)^2} - \eta \frac{\int u_{2\infty}^{*q} \Delta u_{2\infty}^{*q-1}}{\theta_2^2 \int \left(u_{2\infty}^{*q}\right)^2 \int \left(\Delta u_{2\infty}^{*q-1}\right)^2}. \tag{38}$$

4.2.2. *Case* $d < 3/2$.

Next, let us be concerned with the case where $d < 3/2$. Notice that the particular $\{d = 1, \delta = 0\}$ DGP considered by Phillips and Hansen (1990) is not included in what follows. For instance, given that, when $d = 1$, then $q = 1$ and $e = 0$, so that δ must be assumed to be positive given the restriction $\delta + e > 0$. Notwithstanding, this difference it covers DGP's arbitrarily close to the previous one.

When $d < 3/2$, we have that, $q = 1$ and the Δy_{2t} process is a stationary fractionally integrated process. Consequently, under some suitable regularity conditions (see footnote 1), the following consistency results hold for the family of kernel-based estimators considered in this paper,

$$\hat{\omega}_{\Delta\Delta} \xrightarrow{p} \omega_{22} \equiv \sum_{k=-\infty}^{\infty} E\left(u_{20}u_{2k}\right), \tag{39a}$$

$$\hat{\omega}_{u\Delta} \xrightarrow{p} \omega_{12} \equiv \sum_{k=-\infty}^{\infty} E\left(u_{10}u_{2k}\right), \tag{39b}$$

$$\hat{\Delta}_{\Delta\Delta} \xrightarrow{p} \Delta_{22} \equiv \sum_{k=0}^{\infty} E\left(u_{20}u_{2k}\right), \tag{39c}$$

and

$$\hat{\Delta}_{u\Delta} \xrightarrow{p} \Delta_{12} \equiv \sum_{k=0}^{\infty} E\left(u_{10}u_{2k}\right). \tag{39d}$$

All these quantities exist as proved in Appendix C. Therefore,

$$\hat{\Delta}^{*}_{\Delta u} \xrightarrow{p} \Delta_{21} - \omega_{12}\omega_{22}^{-1}\Delta_{22} \equiv \Delta^{*}_{21}. \tag{40}$$

With respect to the cross sample moment $\sum y_{2t}u_{2t}$, if the restriction $e > 0$ holds, it follows from Appendix B that

$$T^{1-2e}\sum y_{2t}u_{2t} \Rightarrow \theta_2^2 \int u^{*}_{2\infty}du^{*}_{2\infty}. \tag{41}$$

Now, by using equations (21), (40) and (41), it is straightforward to prove that the standard FM-OLS estimator fails to achieve a mixed normal limiting distribution within the $d < 3/2$ range as well. In particular, and due to its importance in practice, let us consider the case where $1 < d$. In this case, when $\delta < e$, it can be proved that,

$$\left(\hat{\beta}^{*} - \beta\right) \Rightarrow \frac{-\omega_{12}\omega_{22}^{-1}\int u^{*}_{2\infty}du^{*}_{2\infty}}{\int \left(u^{*}_{2\infty}\right)^2}, \tag{42}$$

whereas, when $\delta = e > 0$, it follows that

$$\left(\hat{\beta}^{*} - \beta\right) \Rightarrow \frac{\int u^{*}_{2\infty}du^{*}_{1\infty} - \omega_{12}\omega_{22}^{-1}\int u^{*}_{2\infty}du^{*}_{2\infty}}{\int \left(u^{*}_{2\infty}\right)^2}. \tag{43}$$

Lastly, when $\delta > e > 0$, we get

$$T^{e-\delta}\left(\hat{\beta}^{*} - \beta\right) \Rightarrow \frac{\theta_u \int u^{*}_{2\infty}du^{*}_{1\infty}}{\theta_2 \int \left(u^{*}_{2\infty}\right)^2}. \tag{44}$$

Hence, (42), (43) and (44) show that the standard FM-OLS estimator fails to achieve a normal mixture limiting distribution even in the case where y_{1t} and y_{2t} are both $FI(1 \pm \varepsilon_1)$ processes with a stationary $FI(\delta)$ disturbance, with $\delta = \pm\varepsilon_2$ and where $\varepsilon_1, \varepsilon_2$ denote real numbers arbitrarily close to zero. Thus, the most important implication of the previous analysis would be the lack of robustness of the standard FM-OLS as derived by Phillips and Hansen (1990) to deviations from the $\{d = 1, \delta = 0\}$ case.

5 Conclusions

In this paper we have generalized the available results on the efficient estimation of cointegrating vectors in a single-equation framework with $I(1)$ variables, to more general cases including both higher order $I(d)$ and fractionally integrated $FI(d)$ processes.

Several conclusions can be drawn from our study. First, when considering the case of cointegration among higher order $CI(d,d)$ processes, $d = 2,3,\ldots$,a FM-OLS estimator exists which does not need to correct for any serial correlation bias, but only for possible endogeneity bias. Indeed, if the standard FM-OLS estimator is implemented in this case, then its limiting distribution is no longer a mixture of nonnals. Second, when analyzing the case of fractional $FCI(d,\delta)$ cointegration, a FM-OLS estimator exists only when $0 \leq \delta < 1/2$. Thus, mean-reversion, i.e., $\delta < 1$, is not sufficient to achieve asymptotic normality. As in the previous case, deviations from the standard DGP where $d = 1$ and $\delta = 0$, as considered by Phillips and Hansen (1990), prevents the standard FM-OLS estimator from achieving its optimal properties. In view of this lack of robustness, the FFM-OLS proposed in this paper, which explicitly takes account of the fractional hypothesis, may constitute a relevant alternative efficient estimator.

Appendix

A. Weak convergence of fractional processes

A property of the fractionally integrated series

$$\Delta^{\delta} u_{1t} = v_{1t}, \tag{A1}$$

$$\phi(B) v_{1t} = \vartheta(B) \varepsilon_{1t}, \tag{A2}$$

$$\varepsilon_{1t} \sim i.i.d.\left(0, \sigma_{\varepsilon}^{2}\right), \tag{A3}$$

is the dependence on the memory parameter, δ, of the growth of the variance of the partial sums. In particular, this implies that the distribution theory used in subsection 3.2. that requires the variance of the partial sums to grow at a linear rate, is not general enough to deal with stationary fractionally integrated series. More precisely, when $\phi(B) = \vartheta(B) = 1$, Sowell (1990) proved that $\operatorname{var}\left(\sum u_{1t}\right) = O_p\left(T^{1+2\delta}\right)$, so that, by assuming $E|\varepsilon_{1t}|^f < \infty$, for $f \geq max\{4, -8\delta/(1+2\delta)\}$, the following functional central limit theorem holds

$$\mathrm{var}\left(\sum u_{1t}\right)^{-1/2} \sum_{j=1}^{[Tr]} u_{1j} \Rightarrow u_{1\infty}^{*}(r) \equiv \frac{1}{\Gamma(1+\delta)} \int_0^r (r-s)^{\delta} dB_1(s). \tag{A4}$$

If we now assume the presence of autoregressive and moving average terms, then, the complexity of the autocovariances of a stationary fractionally integrated process in this general case (see Sowell, 1992 and Chung, 1995) motivates the following rather simple procedure in order to find the rate of convergence of the variance of the partial sums in the general case.

For this, define the following process

$$\Delta x_t = u_{1t}, \tag{A5}$$

and assume $x_0 = 0$, so that $x_t = \sum_{j=1}^{t} u_{1j}$ and $\mathrm{var}\left(\sum u_{1t}\right) = \mathrm{var}(x_T)$.

In doing this, equation (A1) can be equivalently rewritten as follows

$$\Delta^d x_t = v_{1t}, \tag{A6}$$

where $d = 1+\delta \geq 1/2$, so that x_t would be a nonstationary fractionally integrated process. In this case, from Gourieroux and Monfort (1990) it follows that

$$x_T \Rightarrow \frac{\omega_{11}}{\Gamma(d)} \int (1-s)^{d-1} dW_1(s).$$

Therefore,

$$\sigma_{uT}^2 = \omega_{11}^{-1} \mathrm{var}\left(\sum u_{1t}\right) = \omega_{11}^{-1} \mathrm{var}(x_T) = \frac{1}{\Gamma(d)} \int (1-s)^{2d-2} ds = \frac{T^{2d-1}}{\Gamma^2(d)(2d-1)},$$

so that $\mathrm{var}\left(\sum u_{1t}\right) = O_p\left(T^{2d-1}\right) = O_p\left(T^{1+2\delta}\right)$ and the functional central limit theorem (A4) continues to hold in the general ARMA framework.

B. Weak convergence of fractional processes to stochastic integrals

In this section we address the problem of the convergence of the partial sum

$$\sum y_{2,t-1} u_{1t} = \sum \left(\sum_{j=1}^{t-1} u_{2j} \right) \left(\sum_{j=1}^{t} u_{1j} - \sum_{j=1}^{t-1} u_{1j} \right), \tag{B1}$$

where, employing the same notation as in the main text, it is initially assumed that y_{2t} is a nonstationary fractionally integrated process of order d such that $q = 1$:

$$\Delta y_{2t} = u_{2t}, \tag{B2}$$

$$\alpha(B)\Delta^{e} u_{2t} = \zeta(B)\varepsilon_{2t}. \tag{B3}$$

Let $X_{2T}(t) = \sigma_{2T}^{-1}\sum_{j=1}^{t-1} u_{2j}$ and $X_{uT}(t) = \sigma_{uT}^{-1}\sum_{j=1}^{t-1} u_{1j}$ so that equation (B1) becomes

$$\sigma_{2T}\sigma_{uT}\sum X_{2T}(t)\left[X_{uT}(t+1) - X_{uT}(t)\right]. \tag{B4}$$

Now, given that $\left(X_{uT}(t), X_{2T}(t)\right) \Rightarrow \left(u_{1\infty}^{*}, u_{2\infty}^{*}\right)$, and noting that both u_{1t} and u_{2t} are covariance stationary fractionally integrated processes and then $O_p(T)$, following Chan and Terrin (1995), we have the following result:

$$\sum X_{2T}(t)\left[X_{uT}(t+1) - X_{uT}(t)\right] \Rightarrow \int u_{2\infty}^{*} du_{1\infty}^{*}. \tag{B5}$$

Therefore,

$$T^{-d-\delta}\sum y_{2,t-1}u_{1t} = T^{-1/2-e}\sigma_{2T}T^{-1/2-\delta}\sigma_{uT}\sum X_{2T}(t)\left[X_{uT}(t+1) - X_{uT}(t)\right]$$

$$\Rightarrow \theta_2\theta_u\int u_{2\infty}^{*} du_{1\infty}^{*}. \tag{B6}$$

When $q > 1$, we can proceed as in Sims et al. (1990), yielding

$$T^{-d-\delta}\sum y_{2,t-1}u_{1t} \Rightarrow \theta_2\theta_u\int u_{2\infty}^{*q} du_{1\infty}^{*}, \tag{B7}$$

where $u_{2\infty}^{*q}$ stands for the $(q-1)$-fold integral of $u_{2\infty}^{*}$ recursively defined as $u_{2\infty}^{*q}(r) = \int_0^r u_{2\infty}^{*q-1}(s)ds$.

Notice that the term in equation (B6) is a stochastic integral with respect to a fractional Wiener process that cannot be defined in the usual Îto sense, because fractional Brownian motion is not a semimartingale. Nonetheless, it can be properly defined by using its spectral representation as a double Wiener-Îto integral as defined by Major (1981) that exists in the L_2 sense *provided that* $\delta + e > 0$. We refer the reader to Chan and Terrin (1995) and to Comte and Renault (1996) for more detailed comments.

Lastly, lemma 5.1. in Park and Phillips (1988) was conceived assuming that the stochastic processes of interest, B and W, were independent Brownian motions. In such a case, they proved that, by conditioning on the sigma-field generating B, then

$$\int BdW \equiv N(0, \int B^2).$$

This result is suitable of generalization to include the fractional case. In this way, if $u_{1\infty}$ and $u_{2\infty}$ denote two independent fractional brownian motions, then, it is straightforward to prove that $\int u_{2\infty} du_{1\infty} \equiv N\left(0, \int u_{2\infty}^2\right)$.

C. Covariance between two stationary fractionally integrated processes.

Consider the following stationary fractionally integrated processes

$$\Delta^{\delta} u_{1t} = v_{1t}$$
$$\Delta^{e} u_{2t} = v_{2t},$$

$$\operatorname{var}\begin{pmatrix} v_{1t} \\ v_{2t} \end{pmatrix} = \begin{pmatrix} \omega_{11} & \omega_{12} \\ \omega_{21} & \omega_{22} \end{pmatrix}.$$

By definition,

$$u_{1t} = \sum_{j=0}^{\infty} \frac{\Gamma(j+\delta)}{\Gamma(j+1)\Gamma(\delta)} v_{1,t-j}$$

and

$$u_{2t} = \sum_{i=0}^{\infty} \frac{\Gamma(i+e)}{\Gamma(i+1)\Gamma(e)} v_{2,t-i},$$

so that we have

$$E\left(u_{2t} u_{1,t+h}\right) = E\left(\sum_{i=0}^{\infty} \frac{\Gamma(i+e)}{\Gamma(i+1)\Gamma(e)} v_{2,t-i}\right)\left(\sum_{j=0}^{\infty} \frac{\Gamma(j+h+\delta)}{\Gamma(j+h+1)\Gamma(\delta)} v_{1,t-j-h}\right)$$

$$\propto \sum_{j=0}^{\infty} \frac{\Gamma(j+e)}{\Gamma(j+1)\Gamma(e)} \frac{\Gamma(j+h+\delta)}{\Gamma(j+h+1)\Gamma(\delta)}$$

$$= \frac{1}{\Gamma(e)\Gamma(\delta)} \sum_{j=0}^{\infty} \frac{\Gamma(j+e)\Gamma(j+h+\delta)}{\Gamma(j+1)\ \Gamma(j+h+1)}.$$

Now, given the definition of a hypergeometric function (see Abramowitz and Stegun, 1965)

$$F(a,b,c,z) = \frac{\Gamma(c)}{\Gamma(a)\Gamma(b)} \sum_{j=0}^{\infty} \frac{\Gamma(a+j)\Gamma(b+j)}{\Gamma(c+j)} \frac{z^j}{j!},$$

and the well-known relations

$$F(a,b,c,1)=\frac{\Gamma(c)\Gamma(c-a-b)}{\Gamma(c-a)\Gamma(c-b)}$$

and

$$\frac{\Gamma(j+x)}{\Gamma(j+y)}\sim j^{x-y}$$

as $j \to \infty$ (Sheppard's formula), it is straightforward (letting $a=e$, $b=h+\delta$ and $c=h+1$) to show that

$$E\left(u_{2t}u_{1,t+h}\right) \propto h^{\delta+e-1}\frac{\Gamma(1+e-\delta)}{\Gamma(1-\delta)\Gamma(\delta)}.$$

Therefore,

$$\sum_{k=1}^{\infty} E\left(u_{20}u_{1k}\right) \propto \frac{\Gamma(1+e-\delta)}{\Gamma(1-\delta)\Gamma(\delta)}\sum_{k=1}^{\infty} k^{\delta+e-1}. \tag{C1}$$

From the theory of infinite series, it is known that $\sum_{j=1}^{\infty} j^{s}$ converges for $s<1$ and otherwise diverges. Hence, given that, by assumption, $\delta+e<1$, it follows that $s=\delta+e-1<1$ so that (C1) exists.

References

Abramowitz, M. and I. Stegun (1965), *Handbook of Mathematical Functions*, New York: Dover.

Akonom, J. and C. Gourieroux (1988), "A Functional Limit Theorem for Fractional Processes", Working Paper # 8801, CEPREMAP.

Andrews, D.W.K. (1991), "Heteroskedasticity and Autocorrelation Consistent Covariance Matrix Estimation", *Econometrica* 59, 817-858.

Baillie, R.T. and T. Bollerslev (1994), "Cointegration, Fractional Cointegration and Exchange Rate Dynamics", *Journal of Finance* 49, 737-745.

Banerjee, A., Dolado, J.J., Hendry, D.F. and Smith, G.W. (1986), "Exploring Equilibrium Relationships in Econometrics through Static Models: Some Monte Carlo Evidence", *Oxford Bulletin of Economics and Statistics* 48, 253-277.

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

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

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

Chang, Y. and P.C.B. Phillips (1995), "Time Series Regression with Mixtures of Integrated Processes", *Econometric Theory* 11, 1033-1094.

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

Chung, C. (1995), "A Note on Calculating the Autocovariances of the Fractionally Integrated ARMA Models", *Economics Letters* 45, 293-297.

Comte, F. and E. Renault (1996), "Long Memory Continuous Time Models", forthcoming in *Journal of Econometrics*.

Gouriéroux, C., Maurel, F. and A. Monfort (1989), "Least Squares and Fractionally Integrated Regressors", Working Paper # 8913, INSEE.

Gourieroux, C. and A. Monfort (1990), *Séries Temporelles et Modèles Dynamiques*, Paris: Economica.

Granger, C.W.J. and R. Joyeux (1980), "An Introduction to Long-Memory Time Series Models and Fractional Differencing", *Journal of Time Series Analysis* 1, 15-39.

Granger, C. and T. Lee (1989), "Investigation of Production, Sales and Inventory Relationships Using Multicointegration and Non-Symmetric Error Correction Models", *Journal of Applied Econometrics* 4, 145-159.

Granger, C.W.J. and T.H. Lee (1990), "Multicointegration", in Rhodes, G.F. and T.H. Fomby (eds.), *Advances in Econometrics: Cointegration, Spurious Regressions and Unit Roots* 8, 71-84, New York: JAI Press.

Gregoir, S. and G. Laroque (1994), "Polynomial Cointegration: Estimation and Tests", *Journal of Econometrics* 63, 183-214.

Haldrup, N. (1994), "The Asymptotics of Single-Equation Cointegration Regressions with I(1) and I(2) Variables", *Journal of Econometrics* 63, 153-181.

Haldrup, N. and M. Salmon (1995), "Representations of I(2) Cointegrated Systems Using the Smith-McMillan Form", *mimeo*, University of Aarhus.

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

Helson, H. and D. Sarason (1967), "Past and Future", *Mathematical Scandinavia* 21, 5-16.

Hosking, J.R.M. (1981), "Fractional Differencing", *Biometrika* 68, 165-176.

Major, P. (1981), *Multiple Wiener-Îto Integrals*, Lecture Notes in Mathematics 849, New York: Springer-Verlag.

Marmol, F. (1995), "Spurious Regressions between I(d) Processes", *Journal of Time Series Analysis* 16, 313-321.

Marmol, F. (1996), "Nonsense Regressions between Integrated Processes of Different Orders", forthcoming in *Oxford Bulletin of Economics and Statistics.*

Park, J.Y. and P.C.B. Phillips (1988), "Statistical Inference in Regressions with Integrated Processes: Part 1", *Econometric Theory* 4, 468-497.

Phillips, P.C.B. (1986), "Understanding Spurious Regressions in Econometrics", *Journal of Econometrics* 33, 311-340.

Phillips, P.C.B. (1991), "Spectral Regression for Cointegrated Time Series", in

Nonparametric and Semiparametric Methods in Economics and Statistics, ed. by W. Barnett, J. Powell and G. Tauchen, New York: Cambridge University Press.

Phillips, P.C.B. (1995), "Fully Modified Least Squares and Vector Autoregression", *Econometrica* 63, 1023-1078.

Phillips, P.C.B. and B.E. Hansen (1990), "Statistical Inference in Instrumental Variable Regression with I(1) Variables", *Review of Economic Studies* 57, 99-125.

Priestley, M.B. (1981), *Spectral Analysis and Time Series*, volumes I and II, New York: Academic Press.

Sims, C.A., Stock, J.H. and M.W. Watson (1990), "Inference in Linear Time Series with Some Unit Roots", *Econometrica* 58, 113-44.

Sowell, F.B. (1990), " Fractional Unit Root Distribution", *Econometrica* 58, 495-506.

Sowell, F.B. (1992), "Maximum Likelihood Estimation of Stationary Univariate Fractionally Integrated Time Series Models", *Journal of Econometrics* 53, 165-188.

Stock, J. (1987), "Asymptotic Properties of Least Squares Estimators of Cointegrating Vectors", *Econometrica* 55, 1035-1056.

Viano, M.C., Deniau, C. and G. Oppenheim (1995), "Long-Range Dependence and Mixing for Discrete Time Fractional Processes", *Journal of Time Series Analysis* 16, 323-338.

WORKING PAPERS (1)

9525 **Aurora Alejano y Juan M.ª Peñalosa:** La integración financiera de la economía española: efectos sobre los mercados financieros y la política monetaria.

9526 **Ramón Gómez Salvador y Juan J. Dolado:** Creación y destrucción de empleo en España: un análisis descriptivo con datos de la CBBE.

9527 **Santiago Fernández de Lis y Javier Santillán:** Regímenes cambiarios e integración monetaria en Europa.

9528 **Gabriel Quirós:** Mercados financieros alemanes.

9529 **Juan Ayuso Huertas:** Is there a trade-off between exchange rate risk and interest rate risk? (The Spanish original of this publication has the same number.)

9530 **Fernando Restoy:** Determinantes de la curva de rendimientos: hipótesis expectacional y primas de riesgo.

9531 **Juan Ayuso and María Pérez Jurado:** Devaluations and depreciation expectations in the EMS.

9532 **Paul Scbulstad and Ángel Serrat:** An Empirical Examination of a Multilateral Target Zone Model.

9601 **Juan Ayuso, Soledad Núñez and María Pérez-Jurado:** Volatility in Spanish financial markets: The recent experience.

9602 **Javier Andrés e Ignacio Hernando:** ¿Cómo afecta la inflación al crecimiento económico? Evidencia para los países de la OCDE.

9603 **Barbara Dluhosch:** On the fate of newcomers in the European Union: Lessons from the Spanish experience.

9604 **Santiago Fernández de Lis:** Classifications of Central Banks by Autonomy: A comparative analysis.

9605 **M.ª Cruz Manzano Frías y Sofía Galmés Belmonte:** Credit Institutions' Price Policies and Type of Customer: Impact on the Monetary Transmission Mechanism. (The Spanish original of this publication has the same number.)

9606 **Malte Krüger:** Speculation, Hedging and Intermediation in the Foreign Exchange Market.

9607 **Agustín Maravall:** Short-Term Analysis of Macroeconomic Time Series.

9608 **Agustín Maravall and Christophe Planas:** Estimation Error and the Specification of Unobserved Component Models.

9609 **Agustín Maravall:** Unobserved Components in Economic Time Series.

9610 **Matthew B. Canzoneri, Behzad Diba and Gwen Eudey:** Trends in European Productivity and Real Exchange Rates.

9611 **Francisco Alonso, Jorge Martínez Pagés y María Pérez Jurado:** Weighted Monetary Aggregates: an Empirical Approach. (The Spanish original of this publication has the same number.)

9612 **Agustín Maravall and Daniel Peña:** Missing Observations and Additive Outliers in Time Series Models.

9613 **Juan Ayuso and Juan L. Vega:** An empirical analysis of the peseta's exchange rate dynamics.

9614 **Juan Ayuso Huertas:** Un análisis empírico de los tipos de interés reales *ex-ante* en España.

9615 **Enrique Alberola Ila:** Optimal exchange rate targets and macroeconomic stabilization.

9616 **A. Jorge Padilla, Samuel Bentolila and Juan J. Dolado:** Wage bargaining in industries with market power.

9617 **Juan J. Dolado and Francesc Marmol:** Efficient estimation of cointegrating relationships among higher order and fractionally integrated processes.

(1) Previously published Working Papers are listed in the Banco de España publications catalogue.

Queries should be addressed to: Banco de España
Sección de Publicaciones. Negociado de Distribución y Gestión
Telephone: 338 51 80
Alcalá, 50. 28014 Madrid