Working Paper 98-41
Statistics and Econometrics Series 20
June 1998

Departamento de Estadística y Econometría
Universidad Carlos III de Madrid
Calle Madrid, 126
28903 Getafe (Spain)
Fax (341) 624-9849

FM-OLS ESTIMATION OF COINTEGRATING RELATINSHIPS AMONG NONSTATIONARY FRACTIONALLY INTEGRATED PROCESSES

Dolado, J.J. and Marmol F. *

Abstract

In this paper we address the issue of the efficient estimation of the cointegrating vector in linear regression models with variables that follow nonstationary fractionally integrated processes and with the equilibrium error evolving as a weakly stationary linear process. The estimation method herein considered is the FM-OLS first developed by Phillips and Hansen (1990). When $d>1$, a simple FM-OLS estimator is proposed which just entails correcting for the endogeneity bias. By contrast, when delaing with the case where $½<d<1$, we show that the limitng distribution of the OLS estimator is nonstandard and that the FM-OLS methodology cannot longer be implemented. On the other hand, the same conclusions apply when we consider a mixture of nonstationary fractionally integrated processes. Finally, we also study the consequences of applying the original semiparametric FM-OLS estimator for contegrated *I(1)* variables when the true order of integration lies in the nonstationary range and is different from the unit root case. Not surprisingly we find that, under those more general cases, the limitng distribution of the original FM-OLS estimator is no longer mixed normal, loosing its optimal properties.

Key Words

Nonstationary fractionally integrated processes; cointegration; fully modified estimation; misspecification.

*Department of Economics, Universidad Carlos III de Madrid, e-mail: dolado@eco.uc3m.es; Department of Statistics y Econometrics, Universidad Carlos III de Madrid, e-mail: fmarmol@est-econ.uc3m.es. We are very grateful to Manuel Arellano, Niels Haldrup, Sven Hylleberg, Grayham Mizon, Mark Salmon and Esnrique Sentana for helpful comments and suggestions. We wish to thank seminar audiences at Aarhun University, Aarhus, C.E.M.F.I, Madrid, European University Institute, Florence, European Meeting of the Econometric Society 1996, Istambul, Universitat Autònoma de Barcelona and Universidad Carlos III de Madrid. Usual disclaimers apply.

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

FM-OLS ESTIMATION OF COINTEGRATING RELATIONSHIPS AMONG NONSTATIONARY FRACTIONALLY INTEGRATED PROCESSES

JUAN J. DOLADO

Department of Economics, Universidad Carlos III *de Madrid*

and

FRANCESC MARMOL

Department of Statistics and Econometrics, Universidad Carlos III *de Madrid*

June 1998

Abstract: In this paper we address the issue of the efficient estimation of the cointegrating vector in linear regression models with variables that follow nonstationary fractionally integrated processes and with the equilibrium error evolving as a weakly stationary linear process. The estimation method herein considered is the FM-OLS first developed by Phillips and Hansen (1990). When $d > 1$, a simple FM-OLS estimator is proposed which just entails correcting for the endogeneity bias. By contrast, when dealing with the case where $\frac{1}{2} < d < 1$, we show that the limiting distribution of the OLS estimator is nonstandard and that the FM-OLS methodology cannot longer be implemented. On the other hand, the same conclusions apply when we consider a mixture of nonstationary fractionally integrated processes. Finally, we also study the consequences of applying the original semiparametric FM-OLS estimator for cointegrated $I(1)$ variables when the true order of integration lies in the nonstationary range and is different from the unit root case. Not surprisingly we find that, under those more general cases, the limiting distribution of the original FM-OLS estimator is no longer mixed normal, loosing its optimal properties.

Keywords: Nonstationary fractionally integrated processes; cointegration; fully modified estimation; misspecification.
J.E.L. Classification: C12, C15, C22.

We are very grateful to Manuel Arellano, Niels Haldrup, Sven Hylleberg, Grayham Mizon, Mark Salmon and Enrique Sentana for helpful comments and suggestions. We wish to thank seminar audiences at Aarhus University, Aarhus, C.E.M.F.I., Madrid, European University Institute, Florence, European Meeting of the Econometric Society 1996, Istambul, Universitat Autònoma de Barcelona and Universidad Carlos III de Madrid. Usual disclaimers apply. Address correspondence to: Juan J. Dolado, Department of Economics, Universidad Carlos III de Madrid, C/. Madrid 126, 28903 Getafe, Madrid, Spain. E-mail: dolado@eco.uc3m.es. Fax: (341) 6249313.

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, e.g., 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 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, 1998). The FM-OLS estimation with $I(2)$ processes has been recently developed by Chang and Phillips (1995).

Secondly, and most important, the analysis of higher (integer) order integrated processes is not the only way to generalize the results in the unit-root literature. Fractionally integrated processes have become popular with economic data, too, and the associated concept of fractional cointegration, correspondingly, has also become an important and relevant topic in applied time series analysis in recent years. See, for instance, Cheung and Lai (1993), Baillie and Bollerslev (1994), Booth and Tse (1995) and Dittmann (1998). All of them find evidence of fractional cointegration in their data.

In light of the above comments, this paper attempts to examine, from a theoretical point of view, the issue of the efficient estimation of the cointegrating vector in linear regression models with variables that follow nonstationary fractionally integrated processes and with the equilibrium error evolving as a weakly stationary linear process. For this, the paper is organized as follows. In Section 2 we introduce the relevant asymptotic theory and notation and derive the asymptotic distribution of the OLS estimator of the corresponding cointegrating vector. In Section 3 we study the behavior of the FM-OLS estimation method under the proposed fractional set-up. Section 4 extends the results obtained in the preceding sections to the multicointegrated case. Section 5 is concerned with a robustness analysis of the behavior of the original FM-OLS estimator for $I(1)$ variables, as formulated by Phillips and Hansen (1990), when the true order of integration of the variables is different from unity. Some concluding comments are provided in Section 6. Finally, proofs are 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 The Model and Underlying Assumptions

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

(1) $$y_t = (y_{1t}, y_{2t}^{'})^{'}$$

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

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

(3) $$\Delta^d y_{2t} = \varepsilon_{2t},\ t = 1,2,\ldots,T,$$

with $d \in D = \left\{x \in \Re \mid x > \frac{1}{2}, x \neq j + \frac{1}{2}, j = 1,2,\ldots\right\}$. The set D excludes the points $j + \frac{1}{2}, j = 1,2,\ldots$, in order to avoid problems of non invertibility. Further, deterministic components in (3), besides a constant term, are omitted for simplicity, without affecting the main results of the paper; c.f., see Marmol (1998) for the suitable modifications. With respect to the innovation sequence $\varepsilon_t = \left(\varepsilon_{1t}, \varepsilon_{2t}^{'}\right)'$, we shall assume that it satisfies the following general characterization.

<u>ASSUMPTION A.</u> *Let* $\varepsilon_t = \left(\varepsilon_{1t}, \varepsilon_{2t}^{'}\right)'$ *be generated by the linear process*

$$(4) \qquad \varepsilon_t = \sum_{j=0}^{\infty} C_j v_{t-j}, \quad v_t = 0 \text{ for } t \leq 0,$$

where the sequence of random vectors $v_t = \left(v_{1t}, v_{2t}^{'}\right)'$ *is i.i.d.* $(0, \Sigma)$ *with* $\Sigma > 0$, $E\left(v_t^{'} v_t \middle| \ldots, v_{t-2}, v_{t-1}\right) \leq c$ *(a.s.) for some constant* $c > 0$ *and the sequence of matrix coefficients* $\left\{C_j\right\}_{j=0}^{\infty}$ *is 1-summable, i.e.,* $\sum_{j=0}^{\infty} j\left\|C_j\right\| < \infty$.

Further, assume that $\max_i \sup_t E\left|v_{it}\right|^g < \infty$, *where*

(*i*) $g = 2$ *if* $d > \frac{3}{2}$,

(*ii*) $g = 4$ *if* $\frac{3}{4} \leq d < \frac{3}{2}$, *and*

(*iii*) $g = \dfrac{8(1-d)}{2d-1}$ *if* $\frac{1}{2} < d < \frac{3}{4}$.

Hence, throughout this paper, we shall allow ε_t be generated by the linear process (4). This general class of stationary $I(0)$ processes includes all stationary and invertible *ARMA* processes and is therefore of wide applicability. Further, Assumption A implies that the process ε_t is strictly stationary and ergodic with continuous spectral density given by

$$(5) \qquad f_{\varepsilon\varepsilon}(\lambda) = \frac{1}{2\pi}\left(\sum_{j=0}^{\infty} C_j \exp(ij\lambda)\right)\Sigma\left(\sum_{j=0}^{\infty} C_j \exp(ij\lambda)\right)^{*},$$

and long-run covariance matrix $\Omega = 2\pi f_{\varepsilon\varepsilon}(0)$.

Under Assumption A, the partial sum process constructed from $\{\varepsilon_t\}_{t=1}^{\infty}$ satisfies a multivariate invariance principle

$$(6) \qquad T^{-1/2}\sum_{t=1}^{[Tr]} \varepsilon_t \Rightarrow B(r) \equiv BM(\Omega),$$

(*c.f.* Phillips and Durlauf, 1986), 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 $\Omega = \Sigma + \Lambda + \Lambda'$, where $\Sigma = E(\varepsilon_0 \varepsilon_0'), \Lambda = \sum_{k=1}^{\infty} E(\varepsilon_0 \varepsilon_k')$, and define $\Delta = \Sigma + \Lambda$. These matrices are again partitioned conformably with ε_t.

Moreover, under Assumption A, the following results, recently proved by Dolado and Marmol (1998), also hold.

<u>THEOREM 1.</u> *Under Assumption A, as* $T \to \infty$,

$$(7) \qquad T^{1/2-d} y_{2,[Tr]} \Rightarrow B_2^d(r),$$

$$(8) \qquad T^{-d}\sum_{t=1}^{T} y_{2t}\varepsilon_{1t} \Rightarrow \int_0^1 B_2^d(r)dB_1(r) \qquad \text{when } d > 1,$$

$$(9) \qquad T^{-1}\sum_{t=1}^{T} y_{2t}\varepsilon_{1t} \Rightarrow \int_0^1 B_2(r)dB_1(r) + \Delta_{21} \qquad \text{when } d = 1,$$

$$(10) \qquad T^{-1}\sum_{t=1}^{T} y_{2t}\varepsilon_{1t} \xrightarrow{p} \Delta_{21}^d \qquad \text{when } d < 1,$$

where $\Delta_{21}^d \equiv \sum_{k=1}^{\infty} E(\Delta y_{2,0}\varepsilon_{1,k})$, $\Delta_{21}^1 = \Delta_{21}$ *and* $B_2^d(r) = \int_0^r \Gamma(d)^{-1}(r-s)^{d-1}dB_2(s)$, $B_2 \equiv BM(\Omega_{22})$.

Let $\hat{\alpha}$ and $\hat{\beta}$ be estimates based on OLS estimation of (2) with a sample of size T

$$(11)\quad \begin{pmatrix}\hat{\alpha}\\ \hat{\beta}\end{pmatrix}=\begin{pmatrix}T & \sum y_{2t}^{'}\\ \sum y_{2t} & \sum y_{2t}y_{2t}^{'}\end{pmatrix}^{-1}\begin{pmatrix}\sum y_{1t}\\ \sum y_{2t}y_{1t}\end{pmatrix}$$

so that the deviations of the OLS estimators in (11) from the population values α and β that describe the cointegrating relation (2) are given by the expression

$$(12)\quad \begin{pmatrix}\hat{\alpha}-\alpha\\ \hat{\beta}-\beta\end{pmatrix}=\begin{pmatrix}T & \sum y_{2t}^{'}\\ \sum y_{2t} & \sum y_{2t}y_{2t}^{'}\end{pmatrix}^{-1}\begin{pmatrix}\sum \varepsilon_{1t}\\ \sum y_{2t}\varepsilon_{1t}\end{pmatrix}.$$

Now, from expression (6) and Theorem 1 it is straightforward to prove the following result.

<u>THEOREM 2</u>. *Under Assumption A, the OLS estimation of the conditional model* (2) *yields*

$$(13)\quad \text{when } d>1,\quad \begin{pmatrix}T^{1/2}(\hat{\alpha}-\alpha)\\ T^{d}(\hat{\beta}-\beta)\end{pmatrix}\Rightarrow\begin{pmatrix}1 & \int (B_2^d)^{'}\\ \int B_2^d & \int B_2^d(B_2^d)^{'}\end{pmatrix}^{-1}\begin{pmatrix}B_1(1)\\ \int B_2^d dB_1\end{pmatrix},$$

$$(14)\quad \text{when } d=1,\quad \begin{pmatrix}T^{1/2}(\hat{\alpha}-\alpha)\\ T(\hat{\beta}-\beta)\end{pmatrix}\Rightarrow\begin{pmatrix}1 & \int (B_2)^{'}\\ \int B_2 & \int B_2(B_2)^{'}\end{pmatrix}^{-1}\begin{pmatrix}B_1(1)\\ \int B_2 dB_1+\Delta_{21}\end{pmatrix},$$

and

$$(15)\quad \text{when } d<1,\quad \begin{pmatrix}T^{d-1/2}(\hat{\alpha}-\alpha)\\ T^{2d-1}(\hat{\beta}-\beta)\end{pmatrix}\Rightarrow\begin{pmatrix}1 & \int (B_2^d)^{'}\\ \int B_2^d & \int B_2^d(B_2^d)^{'}\end{pmatrix}^{-1}\begin{pmatrix}0\\ \Delta_{21}^d\end{pmatrix}.$$

Note that the OLS estimator of the slope coefficient β in the cointegrating vector is $O_p(T^{-d})$ for $d\geq 1$ and $O_p(T^{1-2d})$ for $d<1$. Thus, for all $d\in D$, the OLS estimator is consistent, even though not always at super-consistent rates. In particular, when $\frac{1}{2}<d<\frac{3}{4}$ the rate of convergence is smaller than the standard $T^{1/2}$. On the other hand, for all $d\in D$, the presence of nuisance parameters in the limiting OLS distribution prevents achieving an asymptotic mixture of normals.

In the particular unit root case ($d=1$), these nuisance parameters are given by Δ_{21} and ω_{21}. 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. 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. This is a very convenient case, since, under asymptotic gaussianity, valid inference can be conducted using standard distributions.

In turn, when $d > 1$ Theorem 2 shows that the second-order bias is no longer present in the limiting OLS distribution. However, the endogeneity bias remains, preventing from achieving a mixture of normals. When $d < \frac{1}{2}$, the bias present is now of second-order. Again, the limiting OLS distribution is, thus, nonstandard.

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 results in Theorem 2 to extend their FM-OLS estimation procedure to the more general nonstationary fractional set-up herein analyzed.

3 Fractional FM-OLS Estimation

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 $\hat{\varepsilon}_t = \left(\hat{\varepsilon}_{1t}, \varepsilon_{2t}'\right)'$, 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

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

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 confine our analysis to the same class of kernel estimates. Equally, the following kernel-based estimator of the one-sided long-run covariance matrix can be defined as

(17) $$\hat{\Delta} = \sum_{j=0}^{M} \ell\left(j/M\right) T^{-1} \sum_{t} \hat{\varepsilon}_{t-j} \hat{\varepsilon}_t^{'} .$$

Then, under some regularity conditions[1] on the bandwidth parameter, M, and Assumption A 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 nonstationary fractionally integrated case. For instance, if we assume the following bandwidth condition,

ASSUMPTION B.

$$M \to \infty \quad \text{as} \quad T \to \infty \quad \textit{such that} \quad T^{-1/2} M \to 0,$$

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

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

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

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

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

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

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

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

(19) $$\leq \left(\int_{-\infty}^{\infty} \left| \ell(x) \right| dx \right) \left(T^{-1} \sum \varepsilon_{2t}^{'} \varepsilon_{2t} \right)^{1/2} \Gamma^{1/2} = O_p(1),$$

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

where $\Gamma = (\hat{\pi} - \pi)' \Im_T \Im_T^{-1} \sum x_t x_t' \Im_T^{-1} \Im_T (\hat{\pi} - \pi) = O_p(1)$ and $\Im_T = diag\{T^{1/2}, T^d I_m\}$. Thus, (18) and (19) imply that $M^{-1} T^{1/2}(\hat{\omega}_{21} - \omega_{21}) = O_p(1)$ and, under Assumption B we finally get $\hat{\omega}_{21} \xrightarrow{p} \omega_{21}$. In the same manner it can be proved that, under Assumption B, $\hat{\omega}_{12} \xrightarrow{p} \omega_{12}$, $\hat{\Omega}_{22} \xrightarrow{p} \Omega_{22}$ and $\hat{\Delta} \xrightarrow{p} \Delta$.

Let us now consider the case where $d > 1$. From Theorem 2 we have that, in order to achieve asymptotic gaussianity, we should only correct for the bias stemming from $\omega_{21} \neq 0$. For this, let us define the *endogeneity bias-corrected* ε_{1t} disturbance

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

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

(21) $y_{1t}^{+} = \alpha + \beta' y_{2t} + \varepsilon_{1t}^{+}$,

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 (21), yielding

$$(22) \quad \begin{pmatrix} \hat{\alpha}^{+} \\ \hat{\beta}^{+} \end{pmatrix} = \begin{pmatrix} T & \sum y_{2t}' \\ \sum y_{2t} & \sum y_{2t} y_{2t}' \end{pmatrix}^{-1} \begin{pmatrix} \sum y_{1t}^{+} \\ \sum y_{2t} y_{1t}^{+} \end{pmatrix},$$

or

$$(23) \quad \begin{pmatrix} \hat{\alpha}^{+} - \alpha \\ \hat{\beta}^{+} - \beta \end{pmatrix} = \begin{pmatrix} T & \sum y_{2t}' \\ \sum y_{2t} & \sum y_{2t} y_{2t}' \end{pmatrix}^{-1} \begin{pmatrix} \sum \hat{\varepsilon}_{1t}^{+} \\ \sum y_{2t} \hat{\varepsilon}_{1t}^{+} \end{pmatrix},$$

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.

<u>THEOREM 3</u>. *Under Assumptions A and B, when* $d > 1$ *the FM-OLS estimation of the conditional model* (21) *yields*

(24) $$\begin{pmatrix} T^{1/2}(\hat{\alpha}^+ - \alpha) \\ T^d(\hat{\beta}^+ - \beta) \end{pmatrix} \Rightarrow \begin{pmatrix} 1 & \int (B_2^d)' \\ \int B_2^d & \int B_2^d (B_2^d)' \end{pmatrix}^{-1} \begin{pmatrix} B_1^+(1) \\ \int B_2^d dB_1^+ \end{pmatrix} \equiv \int_{\varsigma>0} N(0,\varsigma) dP(\varsigma),$$

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

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

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: $R\beta = r$, 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

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

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

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

so that from Theorem 3 it immediately follows that $\xi \Rightarrow \chi^2_{(q)}$, a chi-squared distribution with q degrees of freedom. In the particular case where we wish to use a single coefficient test H_0: $\beta_i = \beta_i^0$, then we can construct the following modified t-statistic:

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

where Z_{ii} denotes the iith-component of the second-moment matrix of the regressors. Expression (24) was first obtained by Chang and Phillips (1995) for the case $d = 2$.

On the other hand, 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

$$(25)\qquad \begin{pmatrix} \hat{\alpha}^{++} - \alpha \\ \hat{\beta}^{++} - \beta \end{pmatrix} = \begin{pmatrix} T & \sum y_{2t}^{'} \\ \sum y_{2t} & \sum y_{2t} y_{2t}^{'} \end{pmatrix}^{-1} \begin{pmatrix} \sum \hat{\varepsilon}_{1t}^{+} \\ \sum y_{2t}\hat{\varepsilon}_{1t}^{+} - T\hat{\Delta}_{21}^{+} \end{pmatrix}.$$

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 (24) when $d = 1$. The reader is referred to this paper for further details.

Lastly, consider the case $d < \frac{1}{2}$. From expression (15) we have that the limiting OLS distribution appears only affected by second-order biases, and note that a kernel correction of the Δ_{21}^{d} term would lead in this case to a degenerate limiting distribution. Thus, when $d < \frac{1}{2}$ and we allow the perturbations to be both contemporary and serially correlated, there is no endogeneity effects in the limiting distribution because the signal from the fractionally integrated regressors is weak relative to the effects of the induced serial correlation.

Remark. Consider without loss of generality the $m = 1$ case, so that expression (15) becomes

$$(26)\qquad T^{2d-1}\left(\hat{\beta} - \beta\right) = \frac{T^{-1}\sum\left(y_{2t} - \bar{y}_2\right)\varepsilon_{1t}}{T^{-2d}\sum\left(y_{2t} - \bar{y}_2\right)^2} \Rightarrow \frac{\Delta_{21}^{d}}{\int\left(\bar{B}_2^{d}\right)^2},$$

where $\bar{y}_2 = T^{-1}\sum y_{2t}$ and $\bar{B}_2^d = B_2^d - \int B_2^d$ denotes a demeaned Brownian motion. This suggest the following modification of the standard OLS estimator:

$$(27)\ \tilde{\beta} = \frac{\sum (y_{2t} - \bar{y}_2) y_{1t}}{\hat{\Delta}_{21}^d \sum (y_{2t} - \bar{y}_2)^2},$$

where $\hat{\Delta}_{21}^d = \sum_{j=0}^{M} \ell(j/M) T^{-1} \sum \Delta y_{2,t-j} \hat{\varepsilon}_{1t}$. Under Assumptions A and B, it is straightforward to prove that

$$(28)\ T^{2d-1}\left(\tilde{\beta} - \beta\right) \Rightarrow \frac{1}{\int \left(\bar{B}_2^d\right)^2},$$

which is free of nuisance parameters, so that the critical values can be obtained for each d by Monte Carlo simulations. Note the similarity of this limiting distribution apart from the fractional nature of the Brownian motion in (28) with the demeaned $\hat{P}_u$ variance ratio test proposed by Phillips and Ouliaris (1990) and with the demeaned R_T^{μ} modified Sargan-Bhargava statistic reported by Stock (1994) in the unit root case.

Finally, it is worth noting a restrictive but important case. When the y_{2t} series are strictly exogenous for β, then it follows from Dolado and Marmol (1998) that the OLS limiting distribution in the conditional model (2) becomes

$$(29)\quad \begin{pmatrix} T^{1/2}(\hat{\alpha} - \alpha) \\ T^{d}\left(\hat{\beta} - \beta\right) \end{pmatrix} \Rightarrow \begin{pmatrix} 1 & \int \left(B_2^d\right)' \\ \int B_2^d & \int B_2^d \left(B_2^d\right)' \end{pmatrix}^{-1} \begin{pmatrix} B_1(1) \\ \int B_2^d dB_1 \end{pmatrix}$$

for all $d > \frac{1}{2}$, where now B_1 and B_2 are independent Brownian motions so that (29) is a mixture of normals.

4 FM-OLS Estimation in Multicointegrated Systems

Consider now the following DGP:

$$(30)\quad y_{0t} = \alpha + \beta_1' y_{1t} + \beta_2' y_{2t} + \varepsilon_{0t},$$

where y_{0t} is a scalar and y_{1t} and y_{2t} are m_1- and m_2-dimensional $(m_1 + m_2 + 1 = n)$, respectively, and generated according to $\Delta^{d_1} y_{1t} = \varepsilon_{1t}$ and $\Delta^{d_2} y_{2t} = \varepsilon_{2t}$ where $d_2 > d_1$. In this section, we will consider the situation in which y_{0t} and y_{2t} are $CI(d_2 . d_2 - d_1)$

with cointegrating vector $\left(1, -\beta_2'\right)'$ and where the resulting error $y_{0t} - \beta_2' y_{2t}$ cointegrates with y_{1t}, having a fully cointegrated system such that ε_{0t} be stationary.

As in Section 2, we shall require the partial sum of the error sequence $\varepsilon_t = \left(\varepsilon_{0t}, \varepsilon_{1t}', \varepsilon_{2t}'\right)'$ to satisfy Assumption A and the multivariate invariance principle

$$(31) \qquad T^{-1/2} \sum_{t=1}^{[Tr]} \varepsilon_t \Rightarrow B(r) \equiv BM(\Omega),$$

with long-run covariance matrix Ω partitioned conformably with ε_t as

$$(32) \qquad \Omega = \begin{pmatrix} \omega_{00} & \omega_{10}' & \omega_{20}' \\ \omega_{10} & \Omega_{11} & \Omega_{21}' \\ \omega_{20} & \Omega_{21} & \Omega_{22} \end{pmatrix} = \Sigma + \Lambda + \Lambda' = \Delta + \Lambda',$$

where we shall assume that Ω_{11} and Ω_{22} are positive definite so that we do not allow for cointegrating relationships between the respective groups of variables. Equally, partition B, Δ and Σ conformably with the disturbance terms and denote by $\hat{\alpha}$, $\hat{\beta}_1$ and $\hat{\beta}_2$ the OLS estimates of the parameters of interest in (30).

<u>THEOREM 4.</u> *Under Assumption* 1,

(*i*) *when* $d_1 \geq 1$,

$$(33) \qquad \begin{pmatrix} T^{1/2}(\hat{\alpha} - \alpha) \\ T^{d_1}\left(\hat{\beta}_1 - \beta_1\right) \\ T^{d_2}\left(\hat{\beta}_2 - \beta_2\right) \end{pmatrix} \Rightarrow \begin{pmatrix} 1 & \int \left(B_1^{d_1}\right)' & \int \left(B_2^{d_2}\right)' \\ \int B_1^{d_1} & \int B_1^{d_1}\left(B_1^{d_1}\right)' & \int B_1^{d_1}\left(B_2^{d_2}\right)' \\ \int B_2^{d_2} & \int B_2^{d_2}\left(B_1^{d_1}\right)' & \int B_2^{d_2}\left(B_2^{d_2}\right)' \end{pmatrix}^{-1} \begin{pmatrix} B_0(1) \\ \Theta \\ \int B_2^{d_2} dB_0 \end{pmatrix},$$

where

$$\Theta = \begin{cases} \int B_1 dB_0 + \Delta_{10} & d_1 = 1 \qquad (\equiv \Theta^1, \text{say}) \\ \int B_1^{d_1} dB_0 & d_1 > 1 \qquad (\equiv \Theta^{d_1}, \text{say}), \end{cases} \quad \textit{and}$$

(*ii*) when $\frac{1}{2} < d_1 < 1$,

$$(34) \qquad \begin{pmatrix} T^{d_1 - 1/2}(\hat{\alpha} - \alpha) \\ T^{2d_1 - 1}\left(\hat{\beta}_1 - \beta_1\right) \\ T^{d_2 + d_1 - 1}\left(\hat{\beta}_2 - \beta_2\right) \end{pmatrix} \Rightarrow \begin{pmatrix} 1 & \int \left(B_1^{d_1}\right)' & \int \left(B_2^{d_2}\right)' \\ \int B_1^{d_1} & \int B_1^{d_1}\left(B_1^{d_1}\right)' & \int B_1^{d_1}\left(B_2^{d_2}\right)' \\ \int B_2^{d_2} & \int B_2^{d_2}\left(B_1^{d_1}\right)' & \int B_2^{d_2}\left(B_2^{d_2}\right)' \end{pmatrix}^{-1} \begin{pmatrix} 0 \\ \Delta_{10}^{d_1} \\ 0 \end{pmatrix},$$

where $\Delta_{10}^{d_1} = \sum_{j=0}^{\infty} E\left(\Delta y_{1,0} \varepsilon_{0,k}\right)$.

Consider first the case where $d_1 \geq 1$. Note that, as in Theorem 2, the OLS estimate of the cointegrating vector is consistent irrespectively of the possible serial correlation of the error term, ε_t, and that the presence of a drift term, Δ_{10}, induces a bias in the limiting distribution of $T^{d_1}\left(\hat{\beta}_1 - \beta_1\right)$ when $d_1 = 1$, due to the fact that the random variable Θ^1 appearing in its limiting distribution would have a non-zero mean. Equally, the presence of the nuisance parameters Δ_{10}, ω_{10} and ω_{20} implies that the OLS estimate have an asymptotic distribution that is not mixed normal and parameter invariant. These results have been proved in the $d_2 = 2, d_1 = 1$ case by Park and Phillips (1989) and Haldrup (1994).

Given that the OLS estimator of the cointegrating vector in (14), in spite of being consistent, has an asymptotic distribution that is generally nonstandard and is plagued with nuisance parameters causing second-order bias effects in finite samples, one can argue as in Section 3 and propose a FM-OLS estimation procedure. This FM-OLS estimator will make use of first-stage (kernel-based) estimates of the long-run covariance matrix Ω

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

where now $\hat{\varepsilon}_t = \left(\hat{\varepsilon}_{0t}, \varepsilon_{1t}, \varepsilon_{2t}\right)'$, being $\hat{\varepsilon}_{0t}$ is the least squares residual from the OLS estimation of (30), where we define a kernel-based consistent estimate of the one-sided long-run covariance matrix Δ by

(36) $$\hat{\Delta} = \sum_{j=0}^{M} \ell\left(j/M\right) T^{-1} \sum_t \hat{\varepsilon}_{t-j} \hat{\varepsilon}_t' .$$

The consistency of these kernel-based estimates in the multicointegrated model (30) can be proved in the same manner as in Section 3, under the assumptions made on the disturbances and if Assumption B holds.

In this sense, when $d_1 > 1$, we can see from (33) that the second-order bias term Δ_{10} disappears, so that we only need to correct the OLS estimation of (30) for the simultaneity bias. In order to perform this correction, let us define the *bias-corrected* disturbances/residuals

(37) $$\varepsilon_{0t}^{+} = \varepsilon_{0t} - \omega_{0*} \Omega_{**}^{-1} \varepsilon_{*t},$$

(38) $\hat{\varepsilon}_{0t}^{+} = \varepsilon_{0t} - \hat{\omega}_{0*}\hat{\Omega}_{**}^{-1}\varepsilon_{*t}$,

so that $(\varepsilon_{0t}^{+}, \varepsilon_{*t}^{'})' = Q'(\varepsilon_{0t}, \varepsilon_{*t}^{'})'$ and $(\hat{\varepsilon}_{0t}^{+}, \varepsilon_{*t}^{'})' = \hat{Q}'(\varepsilon_{0t}, \varepsilon_{*t}^{'})'$, where

(39) $Q' = \begin{pmatrix} 1 & -\omega_{0*}\Omega_{**}^{-1} \\ 0 & I_m \end{pmatrix}$,

and $\hat{Q}'$ is the kernel-based consistent counterpart of Q'. Notice that $\hat{Q}' \xrightarrow{p} Q'$. Here, we use the subscript "*" to signify elements corresponding to "1" and "2" are taken together.

These corrected perturbation terms now have a long-run covariance matrix given by

(40) $\Omega^{+} = Q'\Omega Q = \begin{pmatrix} \omega_{00}^{+} & 0 \\ 0 & \Omega_{**} \end{pmatrix}$,

with $\omega_{00}^{+} = \omega_{00} - \omega_{0*}\Omega_{**}^{-1}\omega_{*0}$. Substracting $\omega_{0*}\Omega_{**}^{-1}\varepsilon_{*t}$ from both sides of (30) we get

(41) $y_{0t}^{+} = \alpha + \beta_1^{'} y_{1t} + \beta_2^{'} y_{2t} + \varepsilon_{0t}^{+}$,

so that a feasible FM-OLS estimate of the cointegrated relation (30) can be formulated as follows

$$(41) \quad \begin{pmatrix} \hat{\alpha}^{+} - \alpha \\ \hat{\beta}_1^{+} - \beta_1 \\ \hat{\beta}_2^{+} - \beta_2 \end{pmatrix} = \left[\begin{Bmatrix} T & \sum y_{1t}^{'} & \sum y_{2t}^{'} \\ \sum y_{1t} & \sum y_{1t} y_{1t}^{'} & \sum y_{1t} y_{2t}^{'} \\ \sum y_{2t} & \sum y_{2t} y_{1t}^{'} & \sum y_{2t} y_{2t}^{'} \end{Bmatrix} \right]^{-1} \begin{pmatrix} \sum \hat{\varepsilon}_{0t}^{+} \\ \sum y_{1t}\hat{\varepsilon}_{0t}^{+} \\ \sum y_{2t}\hat{\varepsilon}_{0t}^{+} \end{pmatrix}.$$

<u>THEOREM 5</u>. *In the multicointegrated model* (30), *under Assumptions A and B, then,*

$$(42) \quad \begin{pmatrix} T^{1/2}(\hat{\alpha}^{+} - \alpha) \\ T^{d_1}(\hat{\beta}_1^{+} - \beta_1) \\ T^{d_2}(\hat{\beta}_2^{+} - \beta_2) \end{pmatrix} \Rightarrow \begin{pmatrix} 1 & \int (B_1^{d_1})' & \int (B_2^{d_2})' \\ \int B_1^{d_1} & \int B_1^{d_1}(B_1^{d_1})' & \int B_1^{d_1}(B_2^{d_2})' \\ \int B_2^{d_2} & \int B_2^{d_2}(B_1^{d_1})' & \int B_2^{d_2}(B_2^{d_2})' \end{pmatrix}^{-1} \begin{pmatrix} B_0^{+}(1) \\ \int B_2^{d_1} dB_0^{+} \\ \int B_2^{d_2} dB_0^{+} \end{pmatrix}$$

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

where

$$\varsigma = \begin{pmatrix} 1 & \int (B_1^{d_1})' & \int (B_2^{d_2})' \\ \int B_1^{d_1} & \int B_1^{d_1}(B_1^{d_1})' & \int B_1^{d_1}(B_2^{d_2})' \\ \int B_2^{d_2} & \int B_2^{d_2}(B_1^{d_1})' & \int B_2^{d_2}(B_2^{d_2})' \end{pmatrix}^{-1},$$

and where $B_0^{+}(r) \equiv BM(\omega_{00}^{+})$.

When $d_1 = 1$, it can be easily proved that

(43) $$T^{-1}\sum y_{1t}\hat{\varepsilon}_{0t}^{+} \xrightarrow{p} T^{-1}\sum y_{1t}\varepsilon_{0t}^{+} \Rightarrow \int B_1 dB_0^{+} + \Delta_{10}^{+},$$

where Δ_{10}^{+} is the conformably part of the one-sided long-run corrected matrix

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

where $\varepsilon_t^{+} = \left(\varepsilon_{0t}^{+}, \left(\varepsilon_{*t}^{+}\right)'\right)' = Q'\varepsilon_t = \left(\varepsilon_{0t}^{+}, \varepsilon_{*t}'\right)'$ so that $\Delta_{10}^{+} = \sum_{k=0}^{\infty} E\left(\varepsilon_{10}\varepsilon_{0k}^{+}\right)$.

To take account of this nuisance parameter and hence, to be able to derive a feasible FM-OLS estimator in the $d_1 = 1$ case, define the following (kernel-based) estimator:

(44) $$\hat{\Delta}_{10}^{+} = \sum_{j=0}^{M} \ell\left(j/M\right) T^{-1}\sum \varepsilon_{1,t-j}\hat{\varepsilon}_{0t}^{+}.$$

A similar reasoning to that made in Section 3 can be applied in this case to show the consistency of the $\hat{\Delta}_{10}^{+}$ estimator to its theoretical counterpart. The feasible FM-OLS will be now given by

(45) $$\begin{pmatrix} \hat{\alpha}^{++} - \alpha \\ \hat{\beta}_1^{++} - \beta_1 \\ \hat{\beta}_2^{++} - \beta_2 \end{pmatrix} = \begin{pmatrix} T & \sum y_{1t}' & \sum y_{2t}' \\ \sum y_{1t} & \sum y_{1t}y_{1t}' & \sum y_{1t}y_{2t}' \\ \sum y_{2t} & \sum y_{2t}y_{1t}' & \sum y_{2t}y_{2t}' \end{pmatrix}^{-1} \begin{pmatrix} \sum \hat{\varepsilon}_{0t}^{+} \\ \sum y_{1t}\hat{\varepsilon}_{0t}^{+} - T\hat{\Delta}_{10}^{+} \\ \sum y_{2t}\hat{\varepsilon}_{0t}^{+} \end{pmatrix}.$$

Therefore, reasoning as in the proof of Theorem 5 and taking account of the consistency of the $\hat{\Delta}_{10}^{+}$ estimator, it is straightforward to prove that, under Assumptions A and B, when $d_1 = 1$, then

(46) $$\begin{pmatrix} T^{1/2}\left(\hat{\alpha}^{++} - \alpha\right) \\ T\left(\hat{\beta}_1^{++} - \beta_1\right) \\ T^{d_2}\left(\hat{\beta}_2^{++} - \beta_2\right) \end{pmatrix} \Rightarrow \begin{pmatrix} 1 & \int (B_1)' & \int (B_2^{d_2})' \\ \int B_1 & \int B_1 (B_1)' & \int B_1 (B_2^{d_2})' \\ \int B_2^{d_2} & \int B_2^{d_2}(B_1)' & \int B_2^{d_2}(B_2^{d_2})' \end{pmatrix}^{-1} \begin{pmatrix} B_0^{+}(1) \\ \int B_1 dB_0^{+} \\ \int B_2^{d_2} dB_0^{+} \end{pmatrix}.$$

Remark. Notice from the definition of the corrected one-sided long-run covariance Δ_{10}^{+} that

$$\Delta_{10}^{+} = \sum_{k=0}^{\infty} E\left(\varepsilon_{10}\varepsilon_{0k}^{+}\right) = \sum_{k=0}^{\infty} E\left(\varepsilon_{10}\left[\varepsilon_{0k} - \omega_{0*}\Omega_{**}^{-1}\varepsilon_{*k}\right]\right) =$$

$$\sum_{k=0}^{\infty} E\begin{pmatrix} \varepsilon_{10}\varepsilon_{0k} \\ \varepsilon_{10}\varepsilon_{*k}' \end{pmatrix} J' = \left(\Delta_{10} \quad \Delta_{1*}\right) J' = \Delta_{1\bullet} J',$$

where $J = \left(1 \quad -\omega_{0*}\Omega_{**}^{-1}\right)$ and the subscript "•" signifies "0" and "∗" taken together. This in turn allows us to rewrite equation (45) in the following manner:

$$(47) \quad \begin{pmatrix} \hat{\alpha}^{++} \\ \hat{\beta}_1^{++} \\ \hat{\beta}_2^{++} \end{pmatrix} = \begin{pmatrix} T & \sum y_{1t}' & \sum y_{2t}' \\ \sum y_{1t} & \sum y_{1t} y_{1t}' & \sum y_{1t} y_{2t}' \\ \sum y_{2t} & \sum y_{2t} y_{1t}' & \sum y_{2t} y_{2t}' \end{pmatrix}^{-1} \left(\begin{pmatrix} \sum \hat{y}_{0t}^{+} \\ \sum y_{1t} \hat{y}_{0t}^{+} \\ \sum y_{2t} \hat{y}_{0t}^{+} \end{pmatrix} - T \begin{pmatrix} 0 \\ \hat{\Delta}_{1\bullet} \hat{J}' \\ 0' \end{pmatrix} \right),$$

where $\hat{J}$ and $\hat{\Delta}_{1\bullet}$ are constructed from the corresponding parts of (35) and (36), respectively. This is the standard FM-OLS format as presented in the seminal paper by Phillips and Hansen (1990), and derived for the particular case where $d_2 = 2, d_1 = 1$ by Chang and Phillips (1995).

Remark. Given that the constructed FM-OLS estimator is a mixture of normals, we can construct conventional Wald statistics and tests restrictions on the cointegrating vector as in Section 3. Nevertheless, when multicointegration is present, we must take account of the fact that $\hat{\beta}_1^{++}$ and $\hat{\beta}_2^{++}$ converge at differing rates implying the possibility of rank defficiencies. Thus, it is convenient to restrict inference to tests of separable restrictions. In particular, this implies that in order to test a null hypothesis of the form H_0: $R\beta = r$, where $\beta' = \left(\beta_1', \; \beta_2'\right)$, the matrix of restrictions R must be block-diagonal across the components of β which are of different orders. See Park and Phillips (1988, 1989), Phillips and Hansen (1990), Hansen (1992) and Haldrup (1994) for more details and comments. Therefore, we must consider a hypothesis test involving q restrictions on β, of the form H_0: $R_{\beta_1}\beta_1 + R_{\beta_2}\beta_2 = r$, where $R = diag\left\{R_{\beta_1}, \; R_{\beta_2}\right\}$ and where R_{β_1} and R_{β_2} are $\left(q \times m_1\right)$ and $\left(q \times m_2\right)$ known matrices, respectively, describing the restrictions. After taking account of the peculiar form of the restrictions matrix, the construction of the Wald test should follow the same lines as in Section 3.

Finally, let us be concerned with the case where $\frac{1}{2} < d_1 < 1$. From (34) two comments arise. First, the OLS estimators remain consistent. Second, as in (15), the serial correlation of the y_{1t} series with the innovation errors of (30) is so strong that it annihilates any endogeneity bias in the (34) OLS system, and prevents the use of any fully-modified correction in order to get mixture of normals.

On the other hand, as in Section 3, if we assume, in turn, that a strictly exogeneity assumption of the underlying series with respect to the parameters of interest holds (i.e., $\Delta_{10} = \omega_{10} = \omega_{20} = 0$), then the OLS estimator of the cointegrating vector in (30) will be a mixture of normals and, hence, standard inferential results will apply. This result was proved by Haldrup (1994) for the $d_2 = 2, d_1 = 1$ particular case.

5 Some Misspecification Analysis

In this last section, we shall briefly investigate the consequences of applying the original FM-OLS estimator, efficient when the relevant processes are $I(1)$ and the equilibrium error is $I(0)$, when in fact the data generating process is composed by nonstationary fractionally integrated processes with $d \in D - \{1\}$. For convenience, let us rewrite the necessary steps to construct such an estimator where, in order to avoid excessive notation, we shall assume that $\alpha = 0$ in (2).

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

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

with $\hat{\varepsilon}_{1t}^* = \varepsilon_{1t} - \hat{\omega}_{1\Delta}\hat{\Omega}_{\Delta\Delta}^{-1}\Delta y_{2t}$, $\hat{\Delta}_{\Delta 1}^* = \hat{\Delta}_{\Delta 1} - \hat{\Delta}_{\Delta\Delta}\hat{\Omega}_{\Delta\Delta}^{-1}\hat{\omega}_{\Delta 1}$, and where the (kernel-based) estimators of the long-run covariances are constructed as $\hat{\omega}_{ab} = \sum_{j=-M}^{M} \ell\left(j/M\right)\hat{\gamma}_{ab}(j)$ and $\hat{\Delta}_{ab} = \sum_{j=0}^{M} \ell\left(j/M\right)\hat{\gamma}_{ab}(j)$, 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}.

<u>THEOREM 6</u>. *Under Assumptions A and B, then*

(*i*) *when* $d \geq 2$,

(49) $$T^d\left(\hat{\beta}^* - \beta\right) \Rightarrow \left[\int B_2^d \left(B_2^d\right)'\right]^{-1}\left[\int B_2^d dB_1 - \left(\int B_2^d \left(B_2^{d-1}\right)'\right) \times \left\{\upsilon_1 \int B_2^{d-1}\left(B_2^{d-1}\right)'\right\}^{-1}\left\{\upsilon_1 \int B_2^{d-1} dB_1 + \zeta_{21}\right\}\right],$$

(*ii*) *when* $\frac{3}{2} < d < 2$,

(50) $$T^{2(d-1)}\left(\hat{\beta}^* - \beta\right) \Rightarrow \left[\int B_2^d \left(B_2^d\right)'\right]^{-1}\left[\left(-\int B_2^d \left(B_2^{d-1}\right)'\right)\left\{\upsilon_1\left[\int B_2^d \left(B_2^{d-1}\right)'\right]\right\}^{-1}\omega_{21}^d\right],$$

(*iii*) *when* $1 < d < \frac{3}{2}$,

(51) $T\left(\hat{\beta}^{*}-\beta\right) \Rightarrow\left[\int B_{2}^{d}\left(B_{2}^{d}\right)^{\prime}\right]^{-1}\left(-\int B_{2}^{d}\left(dB_{2}^{d-1}\right)^{\prime} \Omega_{22}^{-1} \omega_{21}\right)$, *and*

(*iv*) *when* $\frac{1}{2}<d<1$,

(52) $T^{2d-1}\left(\hat{\beta}^{*}-\beta\right) \Rightarrow\left[\int B_{2}^{d}\left(B_{2}^{d}\right)^{\prime}\right]^{-1}\left(\Delta_{21}^{d}-\Delta_{22}^{d} \Omega_{22}^{-1} \omega_{21}\right)$.

Chang (1993), Phillips and Chang (1994) and Harris (1996) first considered the issue of possible misspecification in using the original FM-OLS estimator when in fact the series were $I(2)$, showing that the limit theory for such misspecified estimator were nonstandard and depending on nuisance parameters. Theorem 6 shows how the same comments extend for the rest of values of $d \in D-\{1\}$. Not surprisingly, when the conditions under which it was derived do no hold, the original FM-OLS estimator of β remains consistent (at different rates) but it looses its efficiency properties.

6 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 case where the regressors are assumed to be composed by nonstationary fractionally integrated processes, cointegrated in such a way that the innovation errors are $I(0)$.

Several conclusions can be drawn from our study. First, when $d>1$, a FM-OLS estimator exists which does not need to correct for any serial correlation bias, but only for possible endogeneity bias. This estimator has a nuisance parameter-free mixed normal limiting distribution. Second, when $\frac{1}{2}<d<1$, this optimality can not be achieved by that family of semi-parametric corrections. The OLS estimator is consistent, free of nuisance parameter (after some modifications), but with a nonstandard limiting distribution.

Third, the same comments apply in the multicointegrated case herein analyzed. Fourth, from our misspecification analysis we deduce that even very small deviations from the $d=1$ case prevents the original FM-OLS estimator from achieving its optimal properties. In view of this lack of robustness, explicit account of the fractional hypothesis seems to conform the most suitable solution.

The next step in our analysis is the study, by means of the fully-modified methodology, of the case where the fractional order of the processes as well as the cointegrating dimension are unknown. Moreover, the case where the assumption where the equilibrium error is $I(0)$ is relaxed, and become fractionally integrated, $FI(\delta)$, with $d > \delta$, is also of great interest[2]. All these extensions are currently under investigation.

Appendix

PROOF OF THEOREM 2. Define the weight matrix

$$\text{(A.1)} \quad \Im_T = diag\{T^{1/2}, T^d I_m\},$$

which, in turn, implies that the OLS system (12) can be rewritten as

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

Using (6) and (7) jointly with the continuous mapping theorem (CMT), it is direct to show that

$$\text{(A.3)} \quad \left(\Im_T^{-1} \begin{pmatrix} T & \sum y_{2t}' \\ \sum y_{2t} & \sum y_{2t} y_{2t}' \end{pmatrix} \Im_T^{-1} \right)^{-1} \Rightarrow \begin{pmatrix} 1 & \int B_2^{d'} \\ \int B_2^d & \int B_2^d B_2^{d'} \end{pmatrix}^{-1} \equiv \Pi_d^{-1}, \text{ say,}$$

for all $d \in D$, and where Π_d is positive definite (a.s.).

On the other hand, when $d > 1$, it follows from (6) and (8) that

$$\text{(A.4)} \quad \Im_T^{-1} \begin{pmatrix} \sum \varepsilon_{1t} \\ \sum y_{2t} \varepsilon_{1t} \end{pmatrix} \Rightarrow \begin{pmatrix} B_1(1) \\ \int B_2^d dB_1 \end{pmatrix},$$

and (A.2)-(A.4) jointly with the CMT yield expression (13).

When $d = 1$, from (6), (9) and CMT we obtain

$$\text{(A.5)} \quad \Im_T^{-1} \begin{pmatrix} \sum \varepsilon_{1t} \\ \sum y_{2t} \varepsilon_{1t} \end{pmatrix} \Rightarrow \begin{pmatrix} B_1(1) \\ \int B_2 dB_1 + \Delta_{21} \end{pmatrix},$$

which jointly with (A.2)-(A.3) yields expression (14). Finally, when $d < 1$ we have that

$$\text{(A.6)} \quad T^{d-1} \Im_T^{-1} \begin{pmatrix} \sum \varepsilon_{1t} \\ \sum y_{2t} \varepsilon_{1t} \end{pmatrix} \xrightarrow{p} \begin{pmatrix} 0 \\ \Delta_{21}^d \end{pmatrix},$$

and expression (16) now follows from (A.2), (A.3), (A.6) and the CMT in a direct way. ■

PROOF OF THEOREM 3. First rewrite (22) as

---

2 A preliminary version of this paper (see Dolado and Marmol, 1996) contains results on this case. Robinson and Marinucci (1998), in independent work (but acknowledging our previous research) present some more general results using frequency domain least squares (FDLS) estimators.

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

Now define

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

and note that, under Assumption B, $\hat{Q}'\xrightarrow{p}Q'$, so that

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

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 $\varepsilon_t^+=(\varepsilon_{1t}^+\ \ \varepsilon_{2t}')'=Q'(\varepsilon_{1t}\ \ \varepsilon_{2t}')'$ 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 2. With respect to the gaussian properties, they are implied by the fact that B_1^+ and $B_2^+\equiv B_2$ are independent Brownian motions so that Lemma 5.1 in Park and Phillips (1988) applies when conditioning on the σ-field generated by these stochastic processes. ■

PROOF OF THEOREM 4. The proof follows in a straightforward manner by defining the weight matrix $\Im_T=diag\{T^{1/2},T^{d_1}I_{m_1},T^{d_2}I_{m_2}\}$, rewriting the deviations of the OLS estimators from their corresponding population values as

$$(A.8)\quad \Im_T\begin{pmatrix}\hat{\alpha}-\alpha\\ \hat{\beta}_1-\beta_1\\ \hat{\beta}_2-\beta_2\end{pmatrix}=\left[\Im_T^{-1}\begin{Bmatrix}T & \sum y_{1t}' & \sum y_{2t}'\\ \sum y_{1t} & \sum y_{1t}y_{1t}' & \sum y_{1t}y_{2t}'\\ \sum y_{2t} & \sum y_{2t}y_{1t}' & \sum y_{2t}y_{2t}'\end{Bmatrix}\Im_T^{-1}\right]^{-1}\Im_T^{-1}\begin{pmatrix}\sum\varepsilon_{0t}\\ \sum y_{1t}\varepsilon_{0t}\\ \sum y_{2t}\varepsilon_{0t}\end{pmatrix},$$

and by proceeding as in the proof of Theorem 2. ■

PROOF OF THEOREM 5. The proof of this result follows the same lines as the proof of Theorem 3. Given that $\hat{Q}'\xrightarrow{p}Q'$, then $\hat{\varepsilon}_t^+\xrightarrow{p}\varepsilon_t^+=Q'\varepsilon_t$. This in turn implies, using the CMT theorem and the assumptions made on the perturbation terms, that

$T^{-1/2}\sum \hat{\varepsilon}_t^+ \xrightarrow{p} T^{-1/2}\sum \varepsilon_t^+ \Rightarrow B^+(\cdot) \equiv BM(\Omega^+)$. Notice from the fact $\varepsilon_t^+ = Q'\varepsilon_t$ that $B^+(\cdot) = \left(B_0^+,\ (B_*^+)'\right)' = Q'B(\cdot) = \left(B_0^+,\ B_*'\right)'$, and the result follows using the same arguments as in Theorem 3. ■

PROOF OF THEOREM 6. Consider first the behavior of the standard Phillips and Hansen's FM-OLS estimator for $I(1)$ processes when in fact the underlying series have memory parameter $d \geq 2$. Then, by applying the results in Theorem 1 and the CMT, we obtain

$$T^{-2d}\sum y_{2t}y_{2t}' \Rightarrow \int B_2^d\left(B_2^d\right)',$$

$$T^{3-2d}\hat{\gamma}_{\Delta\Delta}(j) \Rightarrow \int B_2^{d-1}\left(B_2^{d-1}\right)',$$

$$T^{2-d}\hat{\gamma}_{\Delta 1}(j) \Rightarrow \int B_2^{d-1}dB_1 + \Lambda_{21}(j),$$

with

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

so that, following Phillips (1991),

$$M^{-1}T^{3-2d}\hat{\Omega}_{\Delta\Delta} \Rightarrow \upsilon_1\left[\int B_2^{d-1}\left(B_2^{d-1}\right)'\right],$$

$$M^{-1}T^{3-2d}\hat{\Delta}_{\Delta\Delta} \Rightarrow \upsilon_0\left[\int B_2^{d-1}\left(B_2^{d-1}\right)'\right],$$

$$M^{-1}T^{2-d}\hat{\omega}_{\Delta 1} \Rightarrow \upsilon_1\left[\int B_2^{d-1}dB_1\right] + \zeta_{21},$$

$$M^{-1}T^{2-d}\hat{\Delta}_{\Delta 1} \Rightarrow \upsilon_0\left[\int B_2^{d-1}dB_1\right] + \tau_{21},$$

where $\zeta_{21} = \begin{cases} \omega_{21} & \text{if } d = 2 \\ 0 & \text{if } d > 2 \end{cases}$, $\tau_{21} = \begin{cases} \Delta_{21} & \text{if } d = 2 \\ 0 & \text{if } d > 2 \end{cases}$, $\upsilon_1 = \int_{-1}^{1} \ell(x)dx$ and $\upsilon_0 = \int_0^1 \ell(x)dx$.

Consequently,

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

$$\Rightarrow \int B_2^d dB_1 - \left(\int B_2^d\left(B_2^{d-1}\right)'\right)\left\{\upsilon_1\int B_2^{d-1}\left(B_2^{d-1}\right)'\right\}^{-1}\left\{\upsilon_1\int B_2^{d-1}dB_1 + \zeta_{21}\right\}$$

and

$$T^{1-d}\hat{\Delta}_{\Delta 1}^* = M^{-1}T^{2-d}\hat{\Delta}_{\Delta 1}MT^{-1} - M^{-1}T^{3-2d} \times$$

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

Finally, we have that

$$T^{d}\left(\hat{\beta}^{*}-\beta\right) \Rightarrow\left[\int B_{2}^{d}\left(B_{2}^{d}\right)^{'}\right]^{-1}\left[\int B_{2}^{d} dB_{1}-\left(\int B_{2}^{d}\left(B_{2}^{d-1}\right)^{'}\right) \times\right.$$

$$\left.\left\{\upsilon_{1} \int B_{2}^{d-1}\left(B_{2}^{d-1}\right)^{'}\right\}^{-1}\left\{\upsilon_{1} \int B_{2}^{d-1} dB_{1}+\zeta_{21}\right\}\right].$$

Assume now that $\frac{3}{2}<d<2$. Then, we get

$$T^{3-2d} \hat{\gamma}_{\Delta\Delta}(j) \Rightarrow \int B_{2}^{d-1}\left(B_{2}^{d-1}\right)^{'}$$

and

$$\hat{\gamma}_{\Delta 1}(j) \xrightarrow{p} \sum_{k=0}^{\infty} E\left(\Delta y_{2,0} \varepsilon_{1,k+j}\right) \equiv \Delta_{21}^{d}(j), \text{ say,}$$

so that $M^{-1} T^{3-2d} \hat{\Omega}_{\Delta\Delta} \Rightarrow \upsilon_{1}\left[\int B_{2}^{d-1}\left(B_{2}^{d-1}\right)^{'}\right]$, $M^{-1} \hat{\omega}_{\Delta 1} \xrightarrow{p} \sum_{k=-\infty}^{\infty} E\left(\Delta y_{2,0} \varepsilon_{1,k+j}\right) \equiv \omega_{21}^{d}$,

$M^{-1} T^{3-2d} \hat{\Delta}_{\Delta\Delta} \Rightarrow \upsilon_{0}\left[\int B_{2}^{d-1}\left(B_{2}^{d-1}\right)^{'}\right]$ and $M^{-1} \hat{\Delta}_{\Delta 1} \xrightarrow{p} \sum_{k=0}^{\infty} E\left(\Delta y_{2,0} \varepsilon_{1,k+j}\right) \equiv \Delta_{21}^{d}$,

proceeding as in Phillips (1991), so that now

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

$$\Rightarrow-\left(\int B_{2}^{d}\left(B_{2}^{d-1}\right)^{'}\right)\left\{\upsilon_{1}\left[\int B_{2}^{d-1}\left(B_{2}^{d-1}\right)^{'}\right]\right\}^{-1} \omega_{21}^{d}$$

and $M^{-1} \hat{\Delta}_{\Delta 1}^{*}=M^{-1} \hat{\Delta}_{\Delta 1}-M^{-1} T^{3-2d} \hat{\Delta}_{\Delta\Delta}\left(M^{-1} T^{3-2d} \hat{\Omega}_{\Delta\Delta}\right)^{-1} M^{-1} \hat{\omega}_{\Delta 1} \xrightarrow{p} \Delta_{21}^{d}-\upsilon_{0} \upsilon_{1}^{-1} \omega_{21}^{d}$

and thus,

$$T^{2d-2}\left(\hat{\beta}^{*}-\beta\right)=\left(T^{-2d} \sum y_{2t} y_{2t}^{'}\right)^{-1}\left(T^{-2} \sum y_{2t} \hat{\varepsilon}_{1t}^{*}-T^{-3} M M^{-1} \Delta_{\Delta 1}^{*}\right)$$

$$\Rightarrow\left[\int B_{2}^{d}\left(B_{2}^{d}\right)^{'}\right]^{-1}\left[\left(-\int B_{2}^{d}\left(B_{2}^{d-1}\right)^{'}\right)\left\{\upsilon_{1}\left[\int B_{2}^{d}\left(B_{2}^{d-1}\right)^{'}\right]\right\}^{-1} \omega_{21}^{d}\right].$$

Third, assume now that the true DGP is composed by *NFI* processes with $1<d<\frac{3}{2}$. Then we have $\hat{\gamma}_{\Delta\Delta}(j) \xrightarrow{p} E\left(\Delta y_{2,t-j} \Delta y_{2t}^{'}\right)$ and $\hat{\gamma}_{\Delta 1}(j) \xrightarrow{p} E\left(\Delta y_{2,t-j} \varepsilon_{1t}\right)$, so that $\hat{\Omega}_{\Delta\Delta} \xrightarrow{p} \Omega_{22}$, $\hat{\omega}_{\Delta 1} \xrightarrow{p} \omega_{21}$, $\hat{\Delta}_{\Delta 1} \xrightarrow{p} \Delta_{21}$ and $\hat{\Delta}_{\Delta\Delta} \xrightarrow{p} \Delta_{22}$. Consequently, now

$$T^{1-2d} \sum y_{2t} \hat{\varepsilon}_{1t}^{*}=T^{1-2d} \sum y_{2t} \varepsilon_{1t}-T^{1-2d} \sum y_{2t} \Delta y_{2t}^{'} \hat{\Omega}_{\Delta\Delta}^{-1} \hat{\omega}_{\Delta 1} \Rightarrow-\int B_{2}^{d}\left(dB_{2}^{d-1}\right)^{'} \Omega_{22}^{-1} \omega_{21}$$

and $\hat{\Delta}^*_{\Delta 1} \xrightarrow{p} \Delta_{21} - \Delta_{22}\Omega_{22}^{-1}\omega_{21}$, giving rise to

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

$$\Rightarrow \left[\int B_2^d\left(B_2^d\right)'\right]^{-1}\left(-\int B_2^d\left(dB_2^{d-1}\right)'\Omega_{22}^{-1}\omega_{21}\right).$$

Finally, consider the case where $d < 1$. Thus, $\hat{\gamma}_{\Delta\Delta}(j) \xrightarrow{p} E\left(\Delta y_{2,t-j}\Delta y_{2t}'\right)$ and $\hat{\gamma}_{\Delta 1}(j) \xrightarrow{p} E\left(\Delta y_{2,t-j}\varepsilon_{1t}\right)$, and then, $\hat{\Omega}_{\Delta\Delta} \xrightarrow{p} \Omega_{22}$, $\hat{\omega}_{\Delta 1} \xrightarrow{p} \omega_{21}$, $\hat{\Delta}_{\Delta 1} \xrightarrow{p} \Delta_{21}$ and $\hat{\Delta}_{\Delta\Delta} \xrightarrow{p} \Delta_{22}$.

Therefore,

$$T^{-1}\sum y_{2t}\hat{\varepsilon}_{1t}^* = T^{-1}\sum y_{2t}\varepsilon_{1t} - T^{-1}\sum y_{2t}\Delta y_{2t}'\hat{\Omega}_{\Delta\Delta}^{-1}\hat{\omega}_{\Delta 1} \xrightarrow{p} \Delta_{21}^d - \Delta_{22}^d\Omega_{22}^{-1}\omega_{21}$$

where $\Delta_{22}^d \equiv \sum_{k=1}^{\infty} E\left(\Delta y_{20}\Delta y_{2k}\right)$, and $\hat{\Delta}^*_{\Delta 1} \xrightarrow{p} \Delta_{21} - \Delta_{22}\Omega_{22}^{-1}\omega_{21}$, meaning that

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

$$\Rightarrow \left[\int B_2^d\left(B_2^d\right)'\right]^{-1}\left(\Delta_{21}^d - \Delta_{22}^d\Omega_{22}^{-1}\omega_{21}\right),$$

and proving the theorem. ■

References

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.

BOOTH, G. and Y. TSE (1995), "Long Memory in Interest Rate Future Markets: A Fractional Cointegration Analysis", *The Journal of Future Markets* 5, 573-584.

CHANG, Y. (1993), "Fully Modified Estimation of Cointegrated Systems with *I*(2) Processes", *mimeo*, Yale University.

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.

DITTMANN, I. (1998), "Fractional Cointegration of Preferred and Common Stocks", *Discussion Papers in Economics* 98-03, Dortmund University.

DOLADO, J.J. and F. MARMOL (1996), "Efficient Estimation of Cointegrating Relationships Among Higher Order and Fractionally Integrated Processes", *Working Paper* 96-17, Banco de España.

DOLADO, J.J. and F. MARMOL (1998), "Asymptotic Inference with Nonstationary Fractionally Integrated Processes", *mimeo*, Universidad Carlos III de Madrid.

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 (1998), "Representations of $I(2)$ Cointegrated Systems Using the Smith-McMillan Form", *Journal of Econometrics* 84, 303-325.

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

HANSEN, B. (1992), "Efficient Estimation and Testing of Cointegrating Vectors in the Presence of Deterministic Trends", *Journal of Econometrics* 53, 87-121.

HARRIS, D. (1996), "Fully Modified Least Squares in $I(2)$ Regression", *Econometric Theory* 12, 201-204.

MARMOL, F. (1998), "Spurious Regression with Nonstationary Fractionally Integrated Processes", *Journal of Econometrics* 84, 232-250.

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

PARK, J.Y. and P.C.B. PHILLIPS (1989), "Statistical Inference in Regressions with Integrated Processes: Part 2", *Econometric Theory* 5, 95-131.

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, 413-435, 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 Y. CHANG (1994), "Fully Modified Least Squares in *I*(2) Regression", *Econometric Theory* 10, 967.

PHILLIPS, P.C.B. and S. N. DURLAUF (1986), "Multiple Time Series with Integrated Variables", *Review of Economic Studies* 53, 473-496.

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.

PHILLIPS, P.C.B. and S. OULIARIS (1990), "Asymptotic Properties of Residual Based Tests for Cointegration", *Econometrica* 58, 165-194.

PRIESTLEY, M.B. (1981), *Spectral Analysis and Time Series*, volumes I and II, New

ROBINSON, P. and D. MARINUCCI (1998), "Semiparametric Frequency Domain Analysis of Fractional Cointegration", LSE/STICERD *Discussion Paper* EM/98/348.

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

STOCK, J. (1994), "Unit Roots, Structural Breaks and Trends", in *Handbook of Econometrics, Vol.* IV, ed. by R.F. Engle and D.L. McFadden, 2739-2841, New York: Elsevier.