SPURIOUS REGRESSION THEORY WITH NONSTATIONARY FRACTIONALLY INTEGRATED PROCESSES

Francesc Marmol

97-03

WORKING PAPERS

Universidad Carlos III de Madrid

Working Paper 97-03
Statistics and Econometrics Series 03
January 1997

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

SPURIOUS REGRESSION THEORY WITH NONSTATIONARY FRACTIONALLY INTEGRATED PROCESSES.

Francesc Marmol*

Abstract

This paper develops an analytical study of the asymptotic distributions obtained when we run linear regressions in the levels of nonstationary fractionally integrated FI(d) processes, that are spuriously related in a multivariate single-equation setting which allows for the existence of cointegrating relationships and quite general deterministic components. In doing this, the analytical studies of Phillips (1986), haldrup (1994) and Marmol (1995, 1996) are embedded in our results.

Keywords:
Spurious regressions; nonstationary fractionally integrated processes; deterministic trends.

*Marmol, Department of Statistics and Econometrics, Universidad Carlos III de Madrid; This is a much revised version of an earlier paper which was previously circulated under the title "Spurious Fractional Cointegration". I am very grateful to Juan J. Dolado, Cecilia García-Peñalosa, Stéphane Gregoir and Howard Petith for helpful comments and suggestions. Thanks also to various seminar audiences. The associate editor and two anonymous referees deserve considerable credit for their help in preparing this manuscript. Thanks also to the Generalitat de Catalunya for financial support. The usual disclaimers apply.

Spurious Regression Theory with Nonstationary Fractionally Integrated Processes

Francesc Marmol[+]

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

and

G.R.E.M.A.Q.-I.D.E.I., Université des Sciences Sociales de Toulouse

November, 1996

Abstract: This paper develops an analytical study of the asymptotic distributions obtained when we run linear regressions in the levels of nonstationary fractionally integrated $FI(d)$ processes, that are spuriously related in a multivariate single-equation setting which allows for the existence of cointegrating relationships and quite general deterministic components. In doing this, the analytical studies of Phillips (1986), Haldrup (1994) and Marmol (1995, 1996) are embedded in our results.

Keywords: Spurious regressions; nonstationary fractionally integrated processes; deterministic trends.

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

---

1. Introduction

It is a well-known fact since Yule (1926) and Granger and Newbold (1974) that regressions involving nonstationary series can induce spurious correlations despite the absence of any correction between the underlying series. The first analytical study of these *spurious regressions* was undertaken by Phillips (1986) who derives the limiting

---

[+]**Forthcoming in Journal of Econometrics.** This is a much revised version of an earlier paper which was previously circulated under the title "Spurious Fractional Cointegration". I am very grateful to Juan J. Dolado, Cecilia García-Peñalosa, Stéphane Gregoir and Howard Petith for helpful comments and suggestions. Thanks also to various seminar audiences. The associate editor and two anonymous referees deserve considerable credit for their help in preparing this manuscript. Thanks also to the Generalitat de Catalunya for financial support. The usual disclaimers apply.

[*] *Address for correspondence:* Department of Statistics and Econometrics, Universidad Carlos III de Madrid, C/. Madrid 126, 28903 Getafe, Madrid, SPAIN.

distributions of various OLS statistics when dealing with regressions entailing quite general independently distributed integrated of order one $I(1)$ processes. He proved that the distributions of the conventional regression statistics were quite different from those derived under the assumption of stationarity. Using a similar set-up, Marmol (1995) showed how to extend his results to the general $I(d)$ case, with d being an integer number, whereas Haldrup (1994), Hassler (1996) and Marmol (1996) show that the spurious problem also appears in situations where the underlying processes are allowed to have different orders of integration.

A common feature of the above mentioned papers is that they assume that the relevant processes become stationary after taking some number of integer differences. This assumption, however, can introduce discontinuities which often have drawbacks for the interpretation of the statistical estimation methods (e.g., the consequences of under or overdifferencing) or for their properties (e.g., the identification step). These problems can be solved by working with the so-called fractional processes or long-memory models (Granger and Joyeux, 1980, Hosking, 1981), where the memory parameter, d, is assumed to be a real rather than an integer number. This class of processes have proved useful for macro-modelling, nesting the integrated processes as a special, and potentially restrictive, case. They are naturally introduced when we consider the aggregation of heterogeneous time series (Granger, 1980, Gonçalves and Gourieroux, 1987). Moreover, by allowing a rich range of spectral behaviour near the origin, they can provide superior approximations to the Wold representations of many economic time series (Granger, 1966). See Baillie (1996) for a review of the growing literature of econometric work on fractional processes and their applications in economics and finance.

In this paper we tackle the question of spurious regressions under the fractional hypothesis in a multivariate single-equation set-up, which allows for the existence of cointegrating relationships among a set of nonstationary fractionally integrated processes, possibly with maintained deterministic components. We analyze how the use of the fractional methodology proves to be helpful in order to deal with the spurious phenomenon. The plan of the paper is as follows. In Section 2, we introduce the models of interest and the relevant asymptotic theory with nonstationary fractionally integrated processes. Section 3 collects the results obtained on spurious regressions with fractional processes, while Section 4 concludes. Proofs are gathered in the Appendix.

A word on notation. We use the symbols "$\Rightarrow$", "$\xrightarrow{p}$" and "$\equiv$" to signify weak convergence, convergence in probability and equality in distribution, respectively. Stochastic processes such as $B(r)$, $0 \leq r \leq 1$, are written as B for notational simplicity. Similarly, we write integrals with respect to Lebesgue measure such as $\int_0^1 B(s)ds$ more simply as $\int B$. The symbol $\sum_{t=1}^{T}$ is denoted simply as $\sum$. Following the standard stochastic order of magnitude notation, we write $A_T \equiv O_p(1)$ to signify that the random variable A_T has a well-defined limiting distribution. Finally, all limits given in the paper are taken as the sample size $T \to \infty$.

2. The models and the relevant asymptotic theory

When a given series, y_t, becomes stationary after differenciating d times and the degree of differentiation, d, is not an integer but a real number, then the series is said to be *fractionally integrated*, denoted $FI(d)$, and written as

$$\Delta^d y_t = \varepsilon_t,$$

where the equilibrium error, ε_t, is a weak stationary and invertible process and where the fractional difference operator Δ^d can be expressed in terms of a Maclaurin expansion as

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

with $\Gamma(\cdot)$ being the gamma function. It can be proved that the process is stationary and invertible when $d \in (-1/2, 1/2)$; nonstationary but mean-reverting, i.e., returning to its equilibrium or long-run behaviour after any random shock, when $d < 1$, and nonstationary mean-averting, when $d \geq 1$. Furthermore, it is stationary with short-memory when $d \leq 0$, and stationary with long-memory when $0 < d < 1/2$ [1] (e.g. Hosking,

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

1981, Cheung and Lai, 1993 and Baillie, 1996). We will assume throughout the paper that the relevant fractionally integrated processes have memory parameters lying within the nonstationary range.

Now, consider an n-dimensional time series $\{y_{0t}, \xi_t, y_{1t}', y_{2t}'\}'$ generated according to

$$y_{0t} = \gamma_0' \xi_t + y_{0t}^0 , \tag{1}$$

$$y_{1t} = \gamma_1' \xi_t + y_{1t}^0, \qquad \Delta^d y_{1t}^0 = \varepsilon_{1t}, \tag{2}$$

$$y_{2t} = \gamma_2' \xi_t + y_{2t}^0, \qquad \Delta^p y_{2t}^0 = \varepsilon_{2t}, \tag{3}$$

and

$$y_{0t}^0 = \beta_1 y_{1t}^0 + \beta_2 y_{2t}^0 + u_t , \tag{4}$$

where ξ_t is an m_0-dimensional deterministic sequence of general form, y_{1t}^0 and y_{2t}^0 are m_1- and m_2-dimensional $(m_0 + m_1 + m_2 = m)$ nonstationary fractionally integrated stochastic processes of order d and p, respectively, with $p > d > 1/2$, and the one-dimensional $(m + 1 = n)$ time series y_{0t}^0 is generally fractionally integrated of order p. Here, $\gamma_i, i = 0,1,2$, are coefficients of the associated deterministic components as they are defined in ξ_t. Assume, without loss of generality, that the fractionally integrated processes y_{0t}^0, y_{1t}^0 and y_{2t}^0 have initial conditions equal to zero for $t \le 0$.

Throughout this paper, we will postulate that the components of the vector $y_t^0 = \left(y_{0t}^0, \left(y_{1t}^0\right)', \left(y_{2t}^0\right)'\right)'$ are nonstationary fractionally integrated processes, possibly with different memory parameters, $\left(d_0, d_{11}, \ldots, d_{1m_1}, d_{21}, \ldots, d_{2m_2}\right)'$, that can be divided in two groups, $\left(\underline{d_0, d_{21}, \ldots, d_{2m_2}}, \underline{d_{11}, \ldots, d_{1m_1}}\right)'$, according to some characteristic of interest and such that $p = d_0 = d_{21} =, \ldots, = d_{2m_2}$ and $d = d_{11} =, \ldots, = d_{1m_1}$, with $p > d$. For instance, one can be interested in separate those variables which are mean-reverting from those variables which manifest a mean-averting behaviour. Within each of these groups, one can reasonably assume that the processes, at least in the short-run, evolve similarly.

Using (1)-(4), we have

$$y_{0t} = \beta_0' \xi_t + \beta_1' y_{1t} + \beta_2' y_{2t} + u_t = \beta' x_t + u_t, \tag{5}$$

with $\beta_0 = \gamma_0' - \beta_1\gamma_1' - \beta_2\gamma_2'$. This set-up is similar to that considered by Haldrup (1994) in the $p = 2, d = 1$ particular case, and also includes the framework of Phillips (1986) if $d = p = 1$, Marmol (1995) if $d = p = 1,2,3,\ldots$and Marmol (1996) if $d \neq p = 1,2,3\ldots$, $\gamma_1 = \gamma_2 = 0$, $\xi_t = 1$ and we do not allow for multicointegrating relationships in the conditional model (5).

The fractionally integrated processes have received increasing attention due to their ability to provide a better description of nonstationary series, allowing for more parsimonious models. In particular, they seem to be well adapted to the study of the cointegrating properties of a number of series. In this sense, the notion of fractional cointegration, where we allow the equilibrium error to evolve like a (stationary or not, mean-reverting or not) fractionally integrated process (see, e.g., Cheung and Lai, 1993), avoids the knife-edged unit-root vs. no-unit-root distinction in the equilibrium error and therefore it permits a wider range of mean-reversion behaviour than the standard cointegration analysis.

In this paper we will consider two possible cases of interest from the point of view of the study of the spurious regressions. The first one is the case where the equilibrium error, u_t, is $FI(p)$, i.e., such that $\Delta^p u_t = v_t$ is stationary. In this case, there is no fractional cointegration. We will refer to it as the *spurious case*. Second, if u_t is $FI(d)$ so that $\Delta^d u_t = v_t$ is stationary, then y_{0t} and y_{2t} will be fractionally cointegrated *FCI(p,p-d)* processes, i.e., such that the equilibrium error follows a fractionally integrated process with cointegrating vector $(1, -\beta_2)'$, and such that the process $y_{0t}^0 - \beta_2 y_{2t}^0$ is a *FI(d)* process, that does not fractionally cointegrate with y_{1t}^0. Let us denote this second situation as *partially spurious case*. Both cases were considered by Haldrup (1994) in the $p = 2$, $d = 1$ particular case.

In order to analyze the asymptotic properties of the above mentioned processes, we adopt the methodology developed in Akonom and Gourieroux (1988) and Gourieroux et al. (1989), and follow Haldrup's (1994) notation and procedures as close as possible.

Assume that the error sequence $w_t = (v_t, \varepsilon_{1t}', \varepsilon_{2t}')'$ is composed by (weak) stationary processes with zero mean and having moment of order r strictly greater than

$max\{p-1/2,\ 2\}$. Under this assumptions, a functional central limit theorem for nonstationary fractionally integrated processes due to Akonom and Gourieroux (1988) holds, so that, by letting by $f_t = \left(u_t, (y_{1t}^0)', (y_{2t}^0)'\right)'$, it can be shown that

$$D_T^{-1} f_{[Tr]} \Rightarrow B_\infty(r) \equiv \int_0^r E(r-s)dB(s), \tag{6}$$

where

$$D_T = diag\left\{T^{\delta-1/2},\ T^{d-\frac{1}{2}} I_{m_1}, T^{p-\frac{1}{2}} I_{m_2}\right\},$$

$$E(r-s) = diag\left\{\Gamma(\delta)^{-1}(r-s)^{\delta-1},\ \Gamma(d)^{-1}(r-s)^{d-1} I_{m_1}, \Gamma(p)^{-1}(r-s)^{p-1} I_{m_2}\right\},$$

with $r,s \in [0,1]$, $r \neq s$, and where $\delta = p,d$, according to whether we are in the spurious or in the partially spurious case, respectively. Here, $B(r) = \left(B_0(r), B_1(r)', B_2(r)'\right)'$ denotes an $(m_1 + m_2 + 1)$-dimensional Brownian motion with long-run covariance matrix Ω given by

$$\Omega = \begin{pmatrix} \omega_{00} & \omega_{01} & \omega_{02} \\ \omega_{10} & \Omega_{11} & \Omega_{12} \\ \omega_{20} & \Omega_{21} & \Omega_{22} \end{pmatrix},$$

partitioned conformably with w_t, and where we assume that the diagonal submatrices Ω_{11} and Ω_{22} are positive definite such that y_{1t}^0 and y_{2t}^0 are not allowed to be individually fractionally cointegrated.

For the ith component, ξ_{it}, of the deterministic sequence, ξ_t, we assume, as in Haldrup (1994), that there exists a number a_i and a function $\zeta_i(r)$, such that $\xi_{T_i}(r) = T^{-a_i}\xi_{i,[Tr]} \Rightarrow \zeta_i(r)$, where $\xi_{T_i}(r)$ and $\zeta_i(r)$ both are defined on $[0,1]$ and are bounded, such that $\zeta(r) = \left\{\zeta_1(r), \zeta_2(r), \ldots, \zeta_{m_0}(r)\right\}$ be linearly independent to ensure the nonsingularity of $\int \zeta\zeta'$.

Now, define the $m_0 \times m_0$ and $m \times m$ diagonal matrices D_{0T} and $\Im_T$, respectively, by

$$D_{0T} = diag\left\{T^{a_1}, T^{a_2}, \ldots, T^{a_{m_0}}\right\},$$

and

$$\Im_T = diag\{D_{0T}, T^{d-1/2} I_{m_1}, T^{p-1/2} I_{m_2}\}, \tag{7}$$

and let $z_t = (\xi_t, y_{1t}^{\sigma}, y_{2t}^{\sigma})'$ be an m-dimensional vector collecting the deterministic sequence and the fractional stochastic trends, so that z_t and x_t are connected through a one-to-one mapping given by the relation

$$x_t = G' z_t \tag{8}$$

with the $m \times m$ matrix G being

$$G = \begin{pmatrix} I_{m_0} & \gamma_1 & \gamma_2 \\ 0 & I_{m_1} & 0 \\ 0 & 0 & I_{m_2} \end{pmatrix}.$$

In this case, by using the functional central limit theorem (6), the weight matrix (7) and relation (8), we obtain the following weak convergence result,

$$\Im_T^{-1} z_t = \Im_T^{-1} (G')^{-1} x_t \Rightarrow B_* \equiv (\zeta', B_1^{d'}, B_2^{p'})', \tag{9}$$

and where $B_i^m, i = 0,1,2, \ m = d, p,$ is a fractional Brownian motion defined as

$$B_i^m(r) = \frac{1}{\Gamma(m)} \int_0^r (r-s)^{m-1} dB_i(s).$$

3. Spurious regressions under the null of fractional processes

Let us now address the phenomenon of spurious regressions under the assumption that the time series of interest are generated according to the data generating process (DGP, henceforth) given by equations (1)-(4), with an equilibrium error which is assumed to be generated either as a $FI(p)$ process (*spurious case*) or as a $FI(d)$ process (*partially spurious case*). To be more explicit, we are concerned with the following DGP's.

$$\underbrace{y_{0t}^0}_{FI(p)} = \beta_1 \underbrace{y_{1t}^0}_{FI(d)} + \beta_2 \underbrace{y_{2t}^0}_{FI(p)} + \underbrace{u_t}_{FI(p)} \qquad (\textit{spurious case}), \tag{10a}$$

and

$$\underbrace{y_{0t}^{0}}_{FI(p)} = \beta_1 \underbrace{y_{1t}^{0}}_{FI(d)} + \beta_2 \underbrace{y_{2t}^{0}}_{FI(p)} + \underbrace{u_t}_{FI(d)} \qquad \textit{(partially spurious case)}. \qquad (10b)$$

To start with, let us consider the analysis of the following linear regression model

$$y_{0t} = \hat{\beta}_0 \xi_t + \hat{\beta}_1 y_{1t} + \hat{\beta}_2 y_{2t} + \hat{u}_t = \hat{\beta}' x_t + \hat{u}_t, \qquad (11)$$

with $\hat{\beta} \equiv \left(\hat{\beta}_0, \hat{\beta}_1, \hat{\beta}_2\right)'$ denoting the OLS estimators. In the same manner, let us denote by DW and R^2 the usual Durbin-Watson statistic and the coefficient of multiple correlation, respectively. These statistics are defined as

$$DW = \frac{\sum (\hat{u}_t - \hat{u}_{t-1})^2}{\sum \hat{u}_t^2}, \qquad (12)$$

and

$$R^2 = 1 - \frac{\sum \hat{u}_t^2}{\sum (y_{0t} - \bar{y}_0)^2}. \qquad (13)$$

In order to test hypotheses of the form $H_0: \Re\beta = r$, where $\Re$ is a $(q \times m)$ restriction matrix and r is a known vector with dimension $(q \times 1)$, without allowing for cross restrictions between fractionally integrated processes of different orders (see Haldrup, 1994), we will use the standard F-test statistic, denoted and constructed as

$$F_\beta = q^{-1} \frac{\left(\Re\hat{\beta} - r\right)' \left[\Re\left(\sum x_t x_t'\right)^{-1} \Re'\right]^{-1} \left(\Re\hat{\beta} - r\right)}{T^{-1} \sum \hat{u}_t^2}. \qquad (14)$$

Lastly, let $t_{\beta_{ij}}$ denote the customary t-test statistic under the null hypothesis $H_0: \beta_{ij} = 0$, for $i = 0,1,2$, i.e., the coefficients associated with the deterministic, y_{1t} and y_{2t} processes, respectively, with $j = 1,2,\ldots,m_i$.

<u>*Theorem*</u>. *Assume true the DGP* (1)-(5), *and consider the conditional regression model* (11). *Then, asymptotically,*

(*a*) *in the spurious case, i.e, under* (10a),

(*i*) $T^{1/2-p} \Im_T G\left(\hat{\beta} - \beta\right) \equiv O_p(1)$;

(*ii*) $R^2 \equiv O_p(1)$, *assuming that* y_{0t} *is not dominated by higher-order deterministic trends*;

(*iii*) $T^{-1}F_\beta \equiv O_p(1)$;

(*iv*) $T^{-1/2}t_{\beta_{ij}} \equiv O_p(1)$, *for* $i = 0,1,2$, *and* $j = 1,2,\ldots,m_i$;

(*v*) $T^{2d-1}DW \equiv O_p(1)$, *if* $d < 3/2$, *and*

$T^2 DW \equiv O_p(1)$, *if* $d > 3/2$,

(*b*) *whereas, in the partially spurious case, i.e., under* (10b),

(*i*) $T^{1/2-d}\mathfrak{I}_T G(\hat{\beta} - \beta) \equiv O_p(1)$;

(*ii*) $T^{2(p-d)}(R^2 - 1) \equiv O_p(1)$, *assuming that* y_{0t} *is not dominated by higher-order deterministic trends;*

(*iii*) $T^{-1}F_\beta \equiv O_p(1)$;

(*iv*) $T^{-1/2}t_{\beta_{ij}} \equiv O_p(1)$, *for* $i = 0,1,2$, *and* $j = 1,2,\ldots,m_i$;

(*v*) $T^{2d-1}DW \equiv O_p(1)$, *if* $d < 3/2$, *and*

$T^2 DW \equiv O_p(1)$, *if* $d > 3/2$.

From our Theorem, the following comments are in order. First, notice, from (*a.i*) and (*b.i*), that the order of the least squares regression coefficients will differ across fractionally integrated processes of different orders. In particular, $\hat{\beta}_1$ will diverge at the rate $O_p(T^{p-d})$ in the spurious case, while it will be nondegenerate in the partially spurious case. Conversely, $\hat{\beta}_2$ will have a well-defined limiting distribution in the spurious case, and it will be a consistent estimator of its corresponding theoretical counterpart at the rate $O_p(T^{d-p})$ whenever we are in the partially spurious case. Notwithstanding, notice that the rate of convergence will depend on the difference $p-d$. This implies, in particular, that if $p-d \approx 0$, then $\hat{\beta}_2$ will converge to its theoretical counterpart but only for fairly large samples, and the spurious and the partially spurious cases will have in practical terms almost the same effects. On the other hand, when $p-d \geq 1$ (as in Haldrup, 1994), we get superconsistency.

Consider now the OLS estimators of the deterministic coefficients. Writing out expression $T^{1/2-\delta}\Im_T G(\hat{\beta}-\beta)$, where $\delta = p, d$, according to whether we are in the spurious or in the partially spurious case, respectively, then, for the ith component of the deterministic regression coefficient, we obtain

$$T^{1/2-\delta+a_i}\left(\hat{\beta}_{0i}-\beta_{0i}\right)+T^{1/2-\delta+a_i}\gamma_{i1}\left(\hat{\beta}_1-\beta_1\right)+T^{1/2-\delta+a_i}\gamma_{i2}\left(\hat{\beta}_2-\beta_2\right)\equiv O_p(1),$$

where γ_{i1} and γ_{i2} denote the ith rows of the γ_1 and γ_2 matrices, respectively. This expression is identical to equation (15) in Haldrup (1994). As he notes, it is no longer clear in general which is the order of the deterministic components since it depends on γ_1 and γ_2. In the particular case when $\gamma_1 = \gamma_2 = 0$, then $\hat{\beta}_{0i} \equiv O_p\left(T^{\delta-1/2-a_i}\right)$. Lastly, notice that, even in this particular case, δ can be greater, equal or less than $1/2+a_i$.

Secondly, as regards the coefficient of multiple correlation, we have, from (*b.i*), that it has a well-defined limiting distribution in the spurious case. Conversely, it can be seen from (*b.ii*) that in the partially spurious case it tends to one in probability, having a well-defined limiting distribution of order $O_p\left(T^{2(d-p)}\right)$, which depends on the difference between the underlying memory parameters. Notice that, as in Haldrup (1994), we have assumed that y_{0t} is not dominated by higher-order deterministic trends.

If we now focus our attention on the inferential results, we can observe, from (*a.iii*) and (*b.iii*), that the F-statistic diverges, as it has a nondegenerate limiting distribution of order $O_p(T)$, irrespectively of whether we deal with the spurious case or with the partially spurious case. Therefore, with probability one, this statistic will reject any proposed null hypothesis. With regard to the individual t-statistics, from parts (*a.iv*) and (*b.iv*) we observe that they also diverge at the order $O_p\left(T^{1/2}\right)$, in both cases. This claim remains true even for the deterministic components. Notice, lastly, that the divergence rate of the F-statistic is greater than the divergence rate of the individual t-tests, so that, we might expect a greater rejection rate for the former than for the latter.

Lastly, consider the performance of the Durbin-Watson statistic. Haldrup (1994) showed that, in the particular case where $p=2$ and $d=1$, $DW \xrightarrow{p} 0$ at the order

$O_p(T)$, regardless of whether we are in the spurious or in the partially spurious case. In parts $(a.v)$ and $(b.v)$ we have proved that $DW \xrightarrow{p} 0$ for all values of $p, d > 1/2$ irrespectively again of whether we are in the spurious or the partially spurious case. Hence, the behaviour of this statistic continues to provide a useful way of discriminating between spurious and genuine regressions in the fractional case too. Notice, as well, that it converges to zero at the rate $O_p\left(T^{1-2d}\right)$ if $d < 3/2$ and at the rate $O_p\left(T^{-2}\right)$ if $d > 3/2$. In this sense, Haldrup (1994) has already noted in an Appendix that if no $I(1)$ variables are present in the system, then $DW \equiv O_p\left(T^{-2}\right)$ (in the spurious case). Herein we showed how this claim is only true in the case where $d > 3/2$.

As a last comment, Granger and Newbold (1974) suggested treating any regression for which $R^2 > DW$ as one which is likely to be spurious. The reason is that this inequality could be interpreted as a sign of lack of any equilibrium relationship among the variables in the regression, i.e., the presence of *unbalanced* or *inconsistent* regressions[2], which in turn implies a nonstationary error term and hence strong autocorrelation in the regression residuals. From the results obtained in our Theorem, we can see how this rule continues to apply for the general set-up that we consider in this paper. Consequently, a test based on this statistic may have poor power properties in small samples, given that in the case $d < 3/2$, the order of convergence to zero of the DW statistic is almost negligible as d approaches $1/2$.

Conclusions

This paper has studied the behaviour of spurious regressions in a single-equation multivariate set-up, that allows for the presence of quite general deterministic components as well as for the possibility of multicointegrating relationships, and where the underlying processes of interest are assumed to be nonstationary fractionally

[2] The term *unbalanced* seems to have many different meanings in the literature. As employed in the main text, models with regressors of different orders of integration are potentially *unbalanced* if there is no cointegration among the variables. Johansen (1988), however, uses the notion in a different context with variables of different orders of integration, providing some rather technical conditions to be fulfilled in order for *unbalancing* to occur. I am grateful to one referee for pointing out this difference to me.

integrated processes. We further assumed that these processes have memory parameters which can conveniently be separated in two groups, d and p, according to some characteristic of interest to the practitioner. In this way, we encompass the set-ups considered by Phillips (1986), Haldrup (1994) and Marmol (1995, 1996) that were concerned with integrated frameworks.

In this set-up, we have proved that when dealing with spuriously related nonstationary fractionally integrated processes, standard OLS inference remains invalid. More precisely, our results show that the F- and t-tests diverge, which implies that rejections grow with the size of the sample, despite the lack of relation among the relevant processes. We, then, find that the coefficient of multiple correlation has a well-defined limiting distribution in the spurious case, while it converges in probability to unity when the presence of multicointegrating relationships is allowed, although at a rate that depends on the difference between the memory parameters p,d. With regards to the least squares regression coefficients, we have proved that they have orders which differ across fractionally integrated processes of different orders. They can converge, diverge or even achieve a nondegenerate limiting distribution, depending on the difference between the actual memory parameters, and on whether we are in the spurious or in the partially spurious case. Lastly, the DW statistic converges to zero in probability in all cases, providing one useful way of discriminating between spurious and genuine regressions.

Appendix: Proof of the theorem

<u>*OLS coefficient estimates.*</u> Given that, by definition,

$$\hat{\beta}-\beta=\left(\sum x_t x_t'\right)^{-1}\left(\sum x_t u_t\right),$$

$$\Leftrightarrow T^{1/2-\delta}\Im_T G\left(\hat{\beta}-\beta\right)=\Im_T G\left(T^{-1}\sum x_t x_t'\right)^{-1}G'\Im_T\Im_T^{-1}G'^{-1}T^{-1/2-\delta}\sum x_t u_t, \quad \text{(A1)}$$

then, it follows from (8), (9) in the main text and the continuous mapping theorem (CMT, henceforth) that

$$\Im_T G\left(T^{-1}\sum x_t x_t'\right)^{-1}G'\Im_T \Rightarrow \left(\int B_\bullet B_\bullet'\right)^{-1}\equiv\Pi^{-1} \quad \text{(say)}, \qquad \text{(A2)}$$

with the $(m \times m)$ matrix Π being positive definite (a.s.). Equally, from (6), (8), (9) and the CMT, it follows that

$$\Im_T^{-1} G'^{-1} T^{-1/2-\delta} \sum x_t u_t \Rightarrow \int B_* B_0^{\delta} \equiv \Theta_{\delta} \quad \text{(say)}, \tag{A3}$$

where

$$\delta = \begin{cases} p & \text{under (10a)} \\ d & \text{under (10b)}. \end{cases}$$

Therefore, from (A2), (A3) and the CMT, expression (A1) yields

$$T^{1/2-\delta} \Im_T G(\hat{\beta} - \beta) \Rightarrow \Pi^{-1} \Theta_{\delta}. \tag{A4}$$

<u>*Coefficient of multiple correlation.*</u> As in Haldrup (1994), assume for the sake of simplicity, that y_{0t} is not dominated by higher-order trends. In this case,

$$T^{1/2-p} y_{0t} = T^{1/2-p} \beta_2 y_{2t}^0 + T^{1/2-p} u_t + o_p(1),$$

so that, under (10a), we have that

$$T^{1/2-p} y_{0t} \Rightarrow \beta_2 B_2^p + B_0^p,$$

from which it follows that

$$T^{-2p} \sum (y_{0t} - \bar{y}_0)^2 \Rightarrow \beta_2 \left[\int B_2^p (B_2^p)' - \int B_2^p \int (B_2^p)' \right] \beta_2 +$$

$$+ \int (B_0^p)^2 - \left(\int B_0^p \right)^2 + 2\beta_2 \int B_2^p B_0^p - 2\beta_2 \int B_2^p \int B_0^p, \tag{A5}$$

whereas, under (10b),

$$T^{1/2-p} y_{0t} \Rightarrow \beta_2 B_2^p,$$

$$T^{-2p} \sum (y_{0t} - \bar{y}_0)^2 \Rightarrow \beta_2 \left[\int B_2^p (B_2^p)' - \int B_2^p \int (B_2^p)' \right] \beta_2. \tag{A6}$$

With respect to the residual, $\hat{u}_t$, given that, by definition, $\hat{u}_t^2 = u_t^2 - (\hat{\beta} - \beta)' x_t$, then

$$T^{-2\delta} \sum \hat{u}_t^2 = T^{-2\delta} \sum u_t^2 + T^{1/2-\delta} (\hat{\beta} - \beta)' G' \Im_T \Im_T^{-1} (G')^{-1} T^{-1} \sum x_t x_t' G^{-1} \Im_T^{-1} T^{1/2-\delta} \Im_T G (\hat{\beta} - \beta)$$

$$- 2T^{1/2-\delta} (\hat{\beta} - \beta)' G' \Im_T \Im_T^{-1} (G')^{-1} T^{-1/2-\delta} \sum x_t u_t.$$

Hence, using (A2)-(A4) and the CMT, it can be deduced that

$$T^{-2\delta} \sum \hat{u}_t^2 \Rightarrow \int (B_0^{\delta})^2 - \Theta_{\delta}' \Pi^{-1} \Theta_{\delta}. \tag{A7}$$

Lastly, from (A5), (A7) and the CMT, we have that, under (10a),

$$R^2 \Rightarrow 1 - \frac{\int (B_0^p)^2 - \Theta_p' \Pi^{-1} \Theta_p}{\beta_2 \left[\int B_2^p (B_2^p)' - \int B_2^p \int (B_2^p)' \right] \beta_2 + \int (B_0^p)^2 - \left(\int B_0^p \right)^2 + 2\beta_2 \int B_2^p B_0^p - 2\beta_2 \int B_2^p \int B_0^p}$$

whereas, from (A6), (A7) and the CMT, under (10b), we get

$$T^{2(p-d)}\left(R^2 - 1\right) \Rightarrow -\frac{\int (B_0^d)^2 - \Theta_d' \Pi^{-1} \Theta_d}{\beta_2 \left[\int B_2^p (B_2^p)' - \int B_2^p \int (B_2^p)' \right] \beta_2}.$$

F-test statistic. Equation (14) in the main text can be rewritten as

$$T^{-1} F_\beta = q^{-1} \frac{\left(\hat{\beta} - \beta\right)' T^{1/2-\delta} G' \Im_T \Im_T^{-1} (G')^{-1} \Re' \left[\Re \left(T^{-1} \sum x_t x_t' \right)^{-1} \Re' \right]^{-1}}{T^{-2\delta} \sum \hat{u}_t^2}$$

$$\times \frac{\Re G^{-1} \Im_T^{-1} T^{1/2-\delta} \Im_T G \left(\hat{\beta} - \beta\right)}{T^{-2\delta} \sum \hat{u}_t^2},$$

in which case, (A2), (A4), (A7) and the CMT, yield

$$T^{-1} F_\beta \Rightarrow \frac{\Theta_\delta' \Pi^{-1} \Re' \left[\Re \Pi^{-1} \Re' \right]^{-1} \Re \Pi^{-1} \Theta_\delta}{\int (B_0^\delta)^2 - \Theta_\delta' \Pi^{-1} \Theta_\delta}.$$

t-test statistic. By definition,

$$t_{\beta_{ij}} = \frac{\hat{\beta}_{ij}}{s_{\beta_{ij}}},$$

where $s_{\beta_{ij}}^2 = \hat{\sigma}_u^2 e_{ij}' \left\{ \sum x_t x_t' \right\}^{-1} e_{ij}$, and $\hat{\sigma}_u^2 = T^{-1} \sum \hat{u}_t^2$ stands for the estimator of the variance of the error term. Here, e_{ij} denotes an m-dimensional vector such that

$$e_{ij} = \begin{cases} 1 & \text{for its } ij\text{th component} \\ 0 & \text{otherwise,} \end{cases}$$

so that $e_{ij}' \beta = \beta_{ij}$.

With respect to $\hat{\sigma}_u^2$, notice that, by using equation (A7), we get

$$T^{1-2\delta} \hat{\sigma}_u^2 = T^{-2\delta} \sum \hat{u}_t^2 \Rightarrow \int (B_0^\delta)^2 - \Theta_\delta' \Pi^{-1} \Theta_\delta \quad \left(\equiv \sigma_{\delta\infty}^2, \text{ say} \right). \qquad \text{(A8)}$$

Let us transform the expression of $s^2_{\beta_{ij}}$ in the following suitable form:

$$Ts^2_{\beta_{ij}} = \hat{\sigma}^2_u e'_{ij} G^{-1}\Im_T^{-1}\Im_T G\left\{T^{-1}\sum x_t x_t'\right\}^{-1} G'\Im_T\Im_T^{-1}(G')^{-1} e_{ij}, \tag{A9}$$

where, by construction, we have that

$$G^{-1}\Im_T^{-1} = \begin{pmatrix} D_{0T}^{-1} & -T^{1/2-d}\gamma_1 & -T^{1/2-p}\gamma_2 \\ 0 & T^{1/2-d}I_{m_1} & 0 \\ 0 & 0 & T^{1/2-p}I_{m_2} \end{pmatrix}. \tag{A10}$$

Consider now the case $i = 1$, $j = 1, 2, \ldots, m_1$. From (A10) we have that

$$e'_{1j}G^{-1}\Im_T^{-1} = T^{1/2-d}e'_{1j}.$$

Therefore, under (10a), equation (A9) becomes

$$T^{1-2(p-d)}s^2_{\beta_{1j}} = T^{1-2p}\hat{\sigma}^2_u e'_{1j}\Im_T G\left\{T^{-1}\sum x_t x_t'\right\}^{-1} G'\Im_T e_{1j},$$

so that

$$T^{1-2(p-d)}s^2_{\beta_{1j}} \Rightarrow \sigma^2_{p\infty} e'_{1j}\Pi^{-1}e_{1j}, \tag{A11}$$

by using (A2), (A8) and the CMT.

Likewise, under (10b), equation (A9) yields

$$Ts^2_{\beta_{1j}} \Rightarrow \sigma^2_{d\infty} e'_{1j}\Pi^{-1}e_{1j}, \tag{A12}$$

employing the same arguments as in (A11).

Under the null hypothesis $H_0 : \beta_{1j} = 0$, we have, by construction, that $\hat{\beta}_{1j} = e'_{1j}(\hat{\beta} - \beta)$.

Thus, using result (A4), (A10) and the CMT, it follows that

$$T^{d-p}\hat{\beta}_{1j} \Rightarrow e'_{1j}\Pi^{-1}\Theta_p, \tag{A13}$$

under (10a), while

$$\hat{\beta}_{1j} \Rightarrow e'_{1j}\Pi^{-1}\Theta_d, \tag{A14}$$

if we assume (10b). Therefore, compiling all of the above results, we have that, under (10a),

$$T^{-1/2}t_{\beta_{1j}} \Rightarrow \frac{e'_{1j}\Pi^{-1}\Theta_p}{\sigma_{p\infty}\left(e'_{1j}\Pi^{-1}e_{1j}\right)^{1/2}},$$

combining (A11), (A13) and the CMT, whilst, under (10b), from (A12), (A14) and the CMT, similarly, we get

$$T^{-1/2} t_{\beta_{1j}} \Rightarrow \frac{e_{1j}' \Pi^{-1} \Theta_d}{\sigma_{d\infty} \left(e_{1j}' \Pi^{-1} e_{1j} \right)^{1/2}}.$$

The case where $i = 2, j = 1,2,\ldots,m_2$ is enterely analogous to the one above. Therefore, performing the same steps as in the $i = 1, j = 1,2,\ldots,m_1$ case, it is straightforward to show that, under (10a),

$$T^{-1/2} t_{\beta_{2j}} \Rightarrow \frac{e_{2j}' \Pi^{-1} \Theta_p}{\sigma_{p\infty} \left(e_{2j}' \Pi^{-1} e_{2j} \right)^{1/2}},$$

whereas, under (10b),

$$T^{-1/2} t_{\beta_{2j}} \Rightarrow \frac{e_{2j}' \Pi^{-1} \Theta_d}{\sigma_{d\infty} \left(e_{2j}' \Pi^{-1} e_{2j} \right)^{1/2}}.$$

Lastly, consider the case where $i = 0, j = 1,2,\ldots,m_0$, i.e., the t-ratio for the deterministic trends. In this case, it follows that

$$e_{0j}' G^{-1} \Im_T^{-1} = \left(\underbrace{0,\ldots,T^{-a_j},\ldots,0}_{m_0 \text{ times}}, -T^{1/2-d} \gamma_{j1}, -T^{1/2-p} \gamma_{j2} \right) \equiv \left(T^{-a_j} \left(e_{0j}^0 \right)', -T^{1/2-d} \gamma_{j1}, -T^{1/2-p} \gamma_{j2} \right)$$

where now e_{0j}^0 denotes an m_0-dimensional vector such that

$$e_{0j}^0 = \begin{cases} 1 & \text{for its } 0j\text{th component} \\ 0 & \text{otherwise,} \end{cases}$$

and where γ_{j1} and γ_{j2} stand for the jth row of the matrices γ_1 and γ_2, respectively. Now assume $a_j > d - 1/2$. This, in turn, entails

$$e_{0j}' G^{-1} \Im_T^{-1} = T^{1/2-d} \left(T^{d-a_j-1/2} \left(e_{0j}^0 \right)', -\gamma_{j1}, -T^{d-p} \gamma_{j2} \right).$$

Consequently, proceeding as in the former cases, it follows that, under (10a),

$$T^{1-2(p-d)} s^2_{\beta_{0j}} \Rightarrow \sigma^2_{p\infty}\left(0', -\gamma_{j1}, 0'\right)\Pi^{-1}\begin{pmatrix} 0 \\ -\gamma'_{j1} \\ 0 \end{pmatrix} \equiv \sigma^2_{p\infty}\eta'_1\Pi^{-1}\eta_1 \quad \text{(say)}, \qquad \text{(A15)}$$

and

$$T^{d-p}\hat{\beta}_{0j} \Rightarrow \eta'_1\Pi^{-1}\Theta_p, \qquad \text{(A16)}$$

while, under (10b),

$$Ts^2_{\beta_{0j}} \Rightarrow \sigma^2_{d\infty}\eta'_1\Pi^{-1}\eta_1, \qquad \text{(A17)}$$

and

$$\hat{\beta}_{0j} \Rightarrow \eta'_1\Pi^{-1}\Theta_d. \qquad \text{(A18)}$$

Indeed, when $a_j < d - 1/2$,

$$e'_{0j}G^{-1}\Im_T^{-1} = T^{-a_j}\left(\left(e^0_{0j}\right)', -T^{a_j+1/2-d}\gamma_{j1}, -T^{a_j+1/2-p}\gamma_{j2}\right),$$

which entails that, under (10a),

$$T^{2-2(p-a_j)} s^2_{\beta_{0j}} \Rightarrow \sigma^2_{p\infty}\left(\left(e^0_{0j}\right)', 0', 0'\right)\Pi^{-1}\begin{pmatrix} e^0_{0j} \\ 0 \\ 0 \end{pmatrix} \equiv \sigma^2_{p\infty}\eta'_0\Pi^{-1}\eta_0 \quad \text{(say)}, \qquad \text{(A19)}$$

and

$$T^{a_j+1/2-p}\hat{\beta}_{0j} \Rightarrow \eta'_0\Pi^{-1}\Theta_p, \qquad \text{(A20)}$$

whereas, under (10b),

$$T^{2-2(d-a_j)} s^2_{\beta_{0j}} \Rightarrow \sigma^2_{d\infty}\eta'_0\Pi^{-1}\eta_0 \qquad \text{(A21)}$$

and

$$T^{a_j+1/2-d}\hat{\beta}_{0j} \Rightarrow \eta'_0\Theta_d. \qquad \text{(A22)}$$

Therefore, equations (A15)-(A22) and the CMT finally lead to the following expression:

$$T^{-1/2}t_{\beta_{0j}} \Rightarrow \begin{cases} \dfrac{\eta'_1\Pi^{-1}\Theta_p}{\sigma_{p\infty}\left(\eta'_1\Pi^{-1}\eta_1\right)^{1/2}}, & \text{under (10a)} \\ \dfrac{\eta'_1\Pi^{-1}\Theta_d}{\sigma_{d\infty}\left(\eta'_1\Pi^{-1}\eta_1\right)^{1/2}}, & \text{under (10b),} \end{cases}$$

if $a_j > d-1/2$, while

$$T^{-1/2} t_{\beta_{0j}} \Rightarrow \begin{cases} \dfrac{\eta_0' \Pi^{-1} \Theta_p}{\sigma_{p\infty}\left(\eta_0' \Pi^{-1} \eta_0\right)^{1/2}}, & \text{under (10a)} \\ \dfrac{\eta_0' \Pi^{-1} \Theta_d}{\sigma_{d\infty}\left(\eta_0' \Pi^{-1} \eta_0\right)^{1/2}}, & \text{under (10b),} \end{cases}$$

for the case where $a_j < d-1/2$.

<u>*Durbin-Watson statistic.*</u> As in Haldrup (1994), assume by simplicity that no deterministic terms are present in z_t so that

$$x_t = z_t = \left((y_{1t}^0)', (y_{2t}^0)' \right)' = \left(y_{1t}', y_{2t}' \right)'.$$

In this case, the numerator of the Durbin-Watson statistic can be rewritten as follows

$$\sum (\Delta \hat{u}_t)^2 = \sum (\Delta u_t)^2 + (\hat{\beta} - \beta)' \sum \Delta x_t (\Delta x_t)' (\hat{\beta} - \beta) - 2(\hat{\beta} - \beta)' \sum \Delta x_t \Delta u_t .$$

Now, let us be concerned with the case where $1/2 < d < p < 3/2$. Under this assumption, Δy_{1t}, Δy_{2t} and Δu_t would be *stationary fractionally integrated processes* of orders $d-1, p-1$ and $\delta - 1$, respectively. By applying the weak law of large numbers, it follows that

$$T^{-1} \sum (\Delta u_t)^2 \xrightarrow{p} \operatorname{var}(\Delta u_t) \qquad \left(\equiv \Xi_{00}^{\delta}, \text{say}\right). \tag{A23}$$

In the same manner,

$$T^{-1} \sum \Delta y_{it} \left(\Delta y_{jt}\right)' \xrightarrow{p} \Xi_{ij}, \qquad i,j = 1,2, \tag{A24}$$

and

$$T^{-1} \sum \Delta y_{it} \Delta u_t \xrightarrow{p} \Xi_{i0}^{\delta}, \qquad i = 1,2. \tag{A25}$$

Therefore, under (10b), we have, using (A4), (A23)-(A25) and the CMT, that

$$T^{-1} \sum (\Delta \hat{u}_t)^2 \Rightarrow \Xi_{00}^{d} + \Theta_d' \Pi^{-1} \begin{pmatrix} \Xi_{11} & 0 \\ 0 & 0 \end{pmatrix} \Pi^{-1} \Theta_d - 2\Theta_d' \Pi^{-1} \begin{pmatrix} \Xi_{10}^{d} \\ 0 \end{pmatrix}$$

which, in turn, implies that

$$T^{2d-1}DW \Rightarrow \frac{\Xi_{00}^{d} + \Theta_d^{'}\Pi^{-1}\begin{pmatrix}\Xi_{11} & 0\\ 0 & 0\end{pmatrix}\Pi^{-1}\Theta_d - 2\Theta_d^{'}\Pi^{-1}\begin{pmatrix}\Xi_{10}^{d}\\ 0\end{pmatrix}}{\sigma_{d\infty}^{2}}.$$

Indeed, under model (10a), it is straightforward to prove that

$$T^{-1+2d-2p}\sum(\Delta\hat{u}_t)^2 \Rightarrow \Theta_p^{'}\Pi^{-1}\begin{pmatrix}\Xi_{11} & 0\\ 0 & 0\end{pmatrix}\Pi^{-1}\Theta_p,$$

and thus

$$T^{2d-1}DW \Rightarrow \frac{\Theta_p^{'}\Pi^{-1}\begin{pmatrix}\Xi_{11} & 0\\ 0 & 0\end{pmatrix}\Pi^{-1}\Theta_p}{\sigma_{p\infty}^{2}}.$$

Let us now consider the case $d > 3/2$, i.e., with Δy_{1t}, Δy_{2t} and Δu_t being *nonstationary fractionally integrated processes* of orders $d-1, p-1$ and $\delta-1$, respectively. In this case, we get

$$T^{2-2\delta}\sum(\Delta u_t)^2 \Rightarrow \int\left(B_0^{\delta-1}\right)^2, \tag{A26}$$

$$T^{1/2}\Im_T^{-1}\sum\Delta x_t(\Delta x_t)^{'}T^{1/2}\Im_T^{-1} \Rightarrow \begin{pmatrix}\int B_1^{d-1}\left(B_1^{d-1}\right)^{'} & \int B_1^{d-1}\left(B_1^{p-1}\right)^{'}\\ \int B_1^{p-1}\left(B_1^{d-1}\right)^{'} & \int B_1^{p-1}\left(B_1^{p-1}\right)^{'}\end{pmatrix} \quad (\equiv \Lambda, \text{ say}), \tag{A27}$$

and

$$T^{3/2-\delta}\Im_T^{-1}\sum\Delta x_t(\Delta u_t)^{'} \Rightarrow \begin{pmatrix}\int B_1^{d-1}(B_0^{\delta-1})^{'}\\ \int B_2^{p-1}\left(B_0^{\delta-1}\right)^{'}\end{pmatrix} \quad (\equiv \Phi_\delta, \text{ say}). \tag{A28}$$

Then, using (A4), (A8), (A26)-(A28) and the CMT, we obtain

$$T^2 DW \Rightarrow \frac{\int\left(B_0^{\delta-1}\right)^2 + \Theta_\delta^{'}\Pi^{-1}\Lambda\Pi^{-1}\Theta_\delta - 2\Theta_\delta^{'}\Pi^{-1}\Phi_\delta}{\sigma_{\delta\infty}^{2}}.$$

Lastly, consider the case $p > 3/2, 1/2 < d < 3/2$. Now, Δy_{1t} will be a *stationary* fractionally integrated process of order $d-1$, Δy_{2t} a *nonstationary* fractionally integrated process of order $p-1$, so that

$$T^{2-2p}\sum\Delta y_{2t}(\Delta y_{2t})^{'} \Rightarrow \int B_2^{p-1}\left(B_2^{p-1}\right)^{'}, \tag{A29}$$

and Δu_t a *stationary* (*nonstationary*) fractionally integrated process of order $d-1$, if $\delta = d$ (of order p-1, if $\delta = p$).

Let us, first, deal with the subcase $\delta = p$, and consider the weak convergence of the cross-moment $(m_2 \times m_1)$ dimensional matrix

$$\sum \Delta y_{2t} \Delta y_{1t}^{'} = \left(\sum \Delta y_{2t}^{i} \Delta y_{1t}^{j}\right)\Big|_{\substack{i=1,\dots,m_2 \\ j=1,\dots,m_1}} .$$

Following Dolado and Marmol (1996), it can be proved that

$$T^{2-p-d} \sum \Delta y_{2t}^{i} \Delta y_{1t}^{j} \equiv O_p(1), \tag{A30}$$

and

$$T^{2-p-d} \sum \Delta y_{1t}^{j} \Delta u_t \equiv O_p(1). \tag{A31}$$

Hence, (A4), (A8), (A26), (A30)-(A31) and the CMT entail the following

$$T^{2d-1} DW \Rightarrow \frac{\Theta_p^{'} \Pi^{-1} \begin{pmatrix} \Xi_{11} & 0 \\ 0 & 0 \end{pmatrix} \Pi^{-1} \Theta_p}{\sigma_{p\infty}^2}.$$

Lastly, in the subcase where $\delta = d$, then

$$T^{2-p-d} \sum \Delta y_{2t}^{i} \Delta u_t \equiv O_p(1), \tag{A32}$$

and (A4), (A8), (A26), (A30)-(A32) and the CMT lead to the expression

$$T^{2d-1} DW \Rightarrow \frac{\Xi_{00}^{d} + \Theta_d^{'} \Pi^{-1} \begin{pmatrix} \Xi_{11} & 0 \\ 0 & 0 \end{pmatrix} \Pi^{-1} \Theta_d - 2\Theta_d^{'} \Pi^{-1} \begin{pmatrix} \Xi_{10}^{d} \\ 0 \end{pmatrix}}{\sigma_{d\infty}^2}.$$

Q.E.D.

References

Akonom, J. and C. Gourieroux, 1988, A functional limit theorem for fractional processes, Working Paper # 8801, CEPREMAP.

Baillie, R., 1996, Long memory processes and fractional integration in econometrics, Journal of Econometrics 73, 5-59.

Cheung, Y. and K. Lai, 1993, A fractional cointegration analysis of purchasing power parity, Journal of Business & Economics Statistics 11, 103-112.

Dolado, J. and F. Marmol, 1996, Efficient estimation of cointegrating relationships among higher order and fractionally integrated processes, Working Paper # 9617, Bank of Spain.

Engle, R.F. and C.W.J. Granger, 1987, Cointegration and error correction. Representation, estimation, and testing, Econometrica 55, 251-276.

Gonçalves, E. and C. Gourieroux, 1987, Agrégation de processus autorégressifs d'ordre 1, Annales d'Economie et de Statistique 12, 127-147.

Gourieroux, C., Maurel, F. and A. Monfort, 1989, Least squares and fractionally integrated regressors, Working Paper # 8913, INSEE.

Granger, C.W.J., 1966, The typical spectral shape of an economic variable, Econometrica 34, 150-161.

Granger, C.W.J., 1980, Long memory relationships and the aggregation of dynamic models, Journal of Econometrics 14, 227-238.

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, 15-29.

Granger, C.W.J. and P. Newbold, 1974, Spurious regressions in econometrics, Journal of Econometrics 2, 111-120.

Haldrup, N., 1994, The asymptotics of single-equation cointegration regressions with I(1) and I(2) variables, Journal of Econometrics 63, 153-181.

Hassler, U., 1996, Spurious regressions when stationary regressors are included, Economics Letters 50, 25-31.

Hosking, J., 1981, Fractional differencing, Biometrika 68, 165-176.

Johansen, S., 1988, The mathematical structure of error correction models, Contemporary Mathematics 80, 359-386.

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, Oxford Bulletin of Economics and Statistics 58, 525-536.

Phillips, P.C.B., 1986, Understanding spurious regressions in econometrics, Journal of Econometrics 33, 311-340.

Yule, G.U., 1926, Why we sometimes get nonsense correlations between time series? A study in sampling and the nature of time-series, Journal of the Royal Statistical Society 89, 1-64.