

# Trend stationarity versus long-range dependence in time series analysis



Francesc Marmol<sup>a,\*</sup>, Carlos Velasco<sup>b</sup>

<sup>a</sup>*Department of Statistics and Econometrics, Universidad Carlos III de Madrid,  
C/. Madrid 126, 28903 Getafe (Madrid), Spain*

<sup>b</sup>*Department of Statistics and Econometrics, Universidad Carlos III de Madrid,  
Avenida Universidad 30, 28911 Leganés (Madrid), Spain*

Received 13 January 1998; revised 17 May 2001; accepted 13 July 2001

---

## Abstract

Empirically, it is difficult to offer unequivocal judgment as to whether many real economic variables are fractionally integrated or trend stationary. The objective of this paper is to study the effects of spurious detrending of a nonstationary fractionally integrated  $NFI(d)$ ,  $d \in (\frac{1}{2}, \frac{3}{2})$ . With respect to the performance of the traditional least squares estimators and tests we prove that the estimated time trend coefficient is consistent but that the corresponding  $t$ -Student test diverges. We also analyze a local version in the frequency domain of least squares. We are able to show the consistency of this estimator and that, after conveniently adjusting variance estimates, its  $t$ -ratio has a well-defined but nonstandard limiting distribution. Nonetheless, in this latter case it is possible to obtain a set of critical values giving rise to the correct size for any given  $d \in (\frac{1}{2}, \frac{3}{2})$ . © 2002 Elsevier Science B.V. All rights reserved.

*JEL classification:* C22

*Keywords:* Trend stationarity; Long-range dependence; Spurious detrending; Local frequency domain estimation

---

---

\* Corresponding author.

*E-mail addresses:* fmarmol@est-econ.uc3m.es (F. Marmol), cavelas@est-econ.uc3m.es (C. Velasco).

# 1. Introduction

When a given time series  $y_t$  becomes weakly stationary after applying the filter  $\Delta^d = (1 - L)^d = \sum_{k=0}^{\infty} \Gamma(k - d)L^k / \Gamma(k + 1)\Gamma(-d)$ , with  $L$  being the backshift operator  $y_{t-k} = L^k y_t$ ,  $\Gamma(\cdot)$  standing for the gamma function and the degree of differentiation or *memory parameter*,  $d$ , being a real number, then the series is said to be *fractionally integrated of order  $d$* , denoted  $y_t \sim FI(d)$ . These processes have received increasing attention because of their ability to provide a natural and flexible characterization of the nonstationary and persistent characteristics of economic time series. See Beran (1994), Robinson (1994) and Baillie (1996) for overviews on fractionally integrated and related long memory processes.

A fractionally integrated process is nonstationary if  $d \geq \frac{1}{2}$ . In spite of being nonstationary, the process is mean-reverting with *transitory memory*, i.e., with any random shock having only a temporary influence on the series, if  $d < 1$ , in contrast with the case when  $d \geq 1$ , where the process is both nonstationary and not mean-reverting with *permanent memory*, i.e., with any random shock now having a permanent effect on the present and future path of the series.

However, it is not by large an easy task to discriminate whether a real economic time series is actually trend stationary or fractionally integrated. Empirically, macroeconomic variables appear to be fractionally integrated in the post-Second World War quarterly data, whilst for the historical data covering eighty or more years, it is difficult to offer unequivocal judgment as to whether many real economic variables are fractionally integrated or trend stationary. See, e.g., Chambers (1996).

The purpose of this paper is to analyze the effects of spuriously detrending a nonstationary fractionally integrated process (henceforth denoted *NFI*) of order  $d \in (\frac{1}{2}, \frac{3}{2})$ , the most empirically relevant range with nonstationary series. For this, in Section 2 of the paper we derive the asymptotic distributions of the traditional least squares statistics when regressing an *NFI* process of order  $d \in (\frac{1}{2}, \frac{3}{2})$  onto a constant and a linear time trend. In doing this, the asymptotic distributions of Durlauf and Phillips (1988), which regressed a difference stationary process onto a constant and a linear time trend, are embedded in our results. We prove the consistency and asymptotic normality (after a suitable normalization) of the estimated time trend coefficient and that the corresponding  $t$ -ratio diverges at the  $T^{1/2}$  rate, invalidating in this way any inferential procedure. In effect, some Monte Carlo evidence clearly shows that the null hypothesis of negligible linear trend is (wrongly) rejected in almost all cases in moderate to large samples and for any  $d \in (\frac{1}{2}, \frac{3}{2})$  when using standard normal critical values.

In Section 3 we explore a local version in the frequency domain of least squares. The estimator of the time trend coefficient happens to be

asymptotically equivalent to full band *OLS*. Furthermore, in this case the corresponding *t*-ratio turns out to have a well-defined but nonstandard limiting distribution after conveniently adjusting variance estimates. Experimental evidence using standard normal critical values corroborates theoretical findings but also shows that the null hypothesis of negligible linear trend is still (wrongly) rejected in about half of the occasions. The rejection percentage grows with  $d$  but to a large extent does not depend on the sample size. This fact lead us to construct the set of critical values for the variance-adjusted *t*-ratio, obtaining Monte Carlo sizes very close to the nominal ones for rather short time series. Finally, some concluding remarks are presented in Section 4. Proofs are given in the Appendix.

With respect to the notation employed, the symbols “ $\Rightarrow$ ”, “ $\xrightarrow{p}$ ” and “ $\equiv$ ” denote weak convergence, convergence in probability and equality in distribution, respectively, while  $[\cdot]$  denotes “integer part”. Stochastic processes such that  $y_\infty(r)$  with  $r \in [0, 1]$  are frequently written as  $y_\infty$ . Similarly, we write integrals with respect to Lebesgue measure such as  $\int_0^1 y_\infty(r) dr$  as  $\int y_\infty$ . The symbol  $\sum_{t=1}^T$  is denoted simply as  $\sum$  unless otherwise stated. Finally, all limits given in the paper are as the sample size  $T \rightarrow \infty$ .

## 2. Time trends vs. *NFI* processes

Let us initially consider the analysis of the following OLS regression:

$$y_t = \hat{\alpha} + \hat{\beta}t + res., \quad t = 1, 2, \dots, T, \quad (1)$$

where  $y_t$  is assumed to be a stationary time trend process,  $\hat{\beta} = a_T^{-1} [T \sum t y_t - \sum t \sum y_t]$  and  $\hat{\alpha} = a_T^{-1} [\sum t^2 \sum y_t - \sum t y_t \sum t]$ , with  $a_T = T \sum t^2 - (\sum t)^2$ .

In 1988, Durlauf and Phillips considered the estimation of this model when in fact the true data generating process (*DGP*) for  $y_t$  is a difference stationary process,  $\Delta y_t = \varepsilon_t$ , with the partial sums of the  $\{\varepsilon_t\}$  sequence assumed to satisfy a functional central limit theorem (*FCLT*) of the type discussed in Phillips (1987), allowing for weak dependence and some heterogeneity over time.

Under this *DGP*, Durlauf and Phillips (1988), found the exact convergence rates of the *OLS* estimates in model (1). In particular, they proved that the estimated time trend coefficient is  $T^{1/2}$ -consistent, while  $\hat{\alpha}$  diverges at the  $T^{1/2}$  rate as the corresponding significance ( $H_0: \alpha = 0, \beta = 0$ ) *t*-Student tests, rejecting, thus, with probability one, as  $T$  grows large, the null hypothesis of no significance.

In this section, we will extend the Durlauf and Phillips (1988) findings from the assumption that the true *DGP* for  $y_t$  is a difference stationary process to the more general case where the true *DGP* is assumed to be composed by an *NFI* process of order  $d \in (\frac{1}{2}, \frac{3}{2})$ .

*Assumption A.*

- (i)  $\Delta^d y_t = u_t 1\{t > 0\}$ ,  $u_t = c(L)\varepsilon_t$ ,  $d \in (\frac{1}{2}, \frac{3}{2})$ ,  $c(L) = \sum_{j=0}^{\infty} c_j L^j$ ,  $c_0 = 1$ ,  $c(1) \neq 0$ ,  $\sum_{j=0}^{\infty} j|c_j| < \infty$ .
- (ii)  $\varepsilon_t$  is i.i.d. with zero mean, constant variance  $\sigma_\varepsilon^2$ .
- (iii)  $E|\varepsilon_t|^q < \infty$ ,  $q > \max\{2, 2/(2d - 1)\}$ .

Under Assumption A, Akonom and Gouriéroux (1987) and Marinucci and Robinson (2000) prove that, asymptotically,

$$T^{1/2-d} y_{[Tr]} \Rightarrow c \int_0^r (r-s)^{d-1} dW(s) \quad [\equiv y_\infty(r), \text{ say}], \quad (2)$$

where  $c = \sigma c(1)/\Gamma(d)$ ,  $\sigma^2 = E(\varepsilon_1^2) + 2 \sum_{j=2}^{\infty} E(\varepsilon_1 \varepsilon_j)$ , and  $W(r)$  is a standard Brownian motion. Expression (2) is an *FCLT* for *NFI*( $d$ ) processes that applies to a large class of long memory processes including the well-known (nonstationary) *ARFIMA*( $p, d, q$ ) processes.

It is worth noting that  $y_\infty(r)$  is a Gaussian process with almost surely continuous sample paths and nonstationary increments that does not correspond with the most standard version of the fractional Brownian motion as originally introduced by Mandelbrot and Van Ness (1968) (see also Samorodnitsky and Taqqu, 1994), namely,

$$\begin{aligned} B_H(r) &= \frac{1}{A(H)} \int_{-\infty}^0 \{(r-s)^{H-1/2} - (-s)^{H-1/2}\} dB(s) \\ &\quad + \frac{1}{A(H)} \int_0^r (r-s)^{H-1/2} dB(s), \end{aligned}$$

where  $0 < H < 1$ ,  $r \geq 0$  and

$$A(H) = \left\{ \frac{1}{2H} + \int_0^\infty \{(1+s)^{H-1/2} - s^{H-1/2}\}^2 ds \right\}^{1/2}.$$

In particular, letting  $W_H(r) = (2H)^{1/2} \int_0^r (r-s)^{H-1/2} dW(s)$ ,  $r \geq 0$  and  $H = d - \frac{1}{2}$  so that  $y_\infty(r) = c(2H)^{-1/2} W_H(r)$ , then Marinucci and Robinson (1999) show that for  $r \geq 0$ ,

$$\begin{aligned} B_H(r) &= \frac{1}{A(H)} \left\{ \frac{1}{(2H)^{1/2}} W_H(r) + \int_{-\infty}^0 \{(r-s)^{H-1/2} \right. \\ &\quad \left. - (-s)^{H-1/2}\} dB(s) \right\}, \quad (3) \end{aligned}$$

and  $B_H(r)$  is composed of two independent components, one of them a scaled  $W_H(r)$ .

Table 1

Percentage of rejections (5% size, two-sided  $t$ -test, standard critical values) of the null  $H_0: \beta = 0$  in model (1)<sup>a</sup>

$d$	$T = 100$	$T = 250$	$T = 1000$
0.8	80.88	87.96	93.49
1	87.60	92.02	95.93
1.2	91.81	94.69	97.01

<sup>a</sup>True model:  $\Delta^d y_t = \varepsilon_t$ ,  $\varepsilon_t \sim N(0, 1)$ . 10,000 replications. OLS estimation.

*Theorem 1. Under Assumption A,*

$$T^{1/2-d} \hat{\alpha} \Rightarrow -6 \int (r - 2/3) y_\infty(r) dr, \quad (4)$$

$$T^{3/2-d} \hat{\beta} \Rightarrow 12 \int (r - 1/2) y_\infty(r) dr, \quad (5)$$

$$T^{-1/2} t_\beta \Rightarrow \left\{ \frac{12 \{ \int r y_\infty - 0.5 \int y_\infty \}^2}{\int y_\infty^2 - \{ \int y_\infty \}^2 - 12 \{ \int (r - 1/2) y_\infty \}^2} \right\}^{1/2}, \quad (6)$$

$$T^{-1/2} t_\alpha \Rightarrow \left\{ \frac{9 \{ \int (r - 2/3) y_\infty \}^2}{\int y_\infty^2 - \{ \int y_\infty \}^2 - 12 \{ \int (r - 1/2) y_\infty \}^2} \right\}^{1/2}. \quad (7)$$

Thus, when the true *DGP* is assumed to be an *NFI*( $d$ ) process, we find that the constant term  $\hat{\alpha}$  in regression (1) diverges in distribution at a rate,  $T^{d-1/2}$ , that depends on the memory parameter  $d$ . From expression (5) we can see that the estimated time trend coefficient  $\hat{\beta}$  has a well-defined limiting distribution upon suitable standardization given by  $T^{3/2-d}$ . Consequently, this estimator will be consistent for  $d \in (\frac{1}{2}, \frac{3}{2})$ .

As regards the hypothesis testing, from expressions (6) and (7) we find that the distributions of both traditional  $t$ -Student tests diverge at a rate,  $T^{1/2}$ , that does not depend on  $d$ . Consequently, the consistency of  $\hat{\beta}$  to the true structural coefficient of zero does not translate into desirable properties for these conventional significance tests. This latter result mirrors the asymptotic behavior of the  $t$ -Student statistic in the general spurious framework studied by Marmol (1998), and is clearly reflected in Table 1, where the percent of rejections of the null  $H_0: \beta = 0$  in model (1) is shown when the true model is assumed to be a nonstationary fractional white noise,  $\Delta^d y_t = \varepsilon_t$ ,  $\varepsilon_t \sim N(0, 1)$ .

On the other hand, it is worth noting that  $\hat{\alpha}$  and  $\hat{\beta}$ , suitably standardized, have Gaussian limiting distributions. In effect, since  $\int y_\infty$  and  $\int r y_\infty$  are

Gaussian processes, so is  $\int(r-a)y_\infty = \int r y_\infty - a \int y_\infty \sim N(0, c^2 \Theta_a)$ , where

$$\begin{aligned} \Theta_a &= \int_0^1 \int_0^1 \int_0^r rs(r-\xi)^{d-1}(s-\xi)^{d-1} d\xi dr ds \\ &\quad + a^2 \int_0^1 \int_0^1 \int_0^r (r-\xi)^{d-1}(s-\xi)^{d-1} d\xi dr ds \\ &\quad - 2a \int_0^1 \int_0^1 \int_0^r r(r-\xi)^{d-1}(s-\xi)^{d-1} d\xi dr ds, \end{aligned} \quad (8)$$

implying from expressions (4) and (5) that

$$T^{1/2-d} \hat{\alpha} \equiv N(0, c^2 \Theta^\alpha), \quad (9)$$

and

$$T^{3/2-d} \hat{\beta} \equiv N(0, c^2 \Theta^\beta), \quad (10)$$

where  $\Theta^\alpha = 36\Theta_{2/3}$ ,  $\Theta^\beta = 144\Theta_{1/2}$ , and with  $\Theta_a$  as defined in expression (8). In the particular case where  $d = 1$ ,  $\Theta_a$  reduces to  $(\frac{1}{60})(8 + 20a^2 - 25a)$  and (9) and (10) become

$$T^{1/2} \hat{\alpha} \equiv N\left(0, \frac{\sigma^2 c(1)^2 2}{15}\right), \quad (11)$$

and

$$T^{1/2} \hat{\beta} \equiv N\left(0, \frac{\sigma^2 c(1)^2 6}{5}\right), \quad (12)$$

as proved by Durlauf and Phillips (1988), Theorem 2.1.

*Remark 1.* The issue of spurious detrending arises after the work of Nelson and Plosser (1982) suggesting that most macroeconomic time series are best modeled as difference stationary  $I(1)$  processes; hence the interest in detecting the consequences of *OLS* linear detrending of time series characterized by stochastic trends. In the same manner, though, it is not difficult to prove that if the memory parameter  $d$  is allowed to be greater than  $\frac{3}{2}$  and we are interested in studying the consequences of spurious detrending, then the order of the fitted time polynomial must match the memory parameter.

For instance, assume now that visual inspection suggests that a time series might be either  $NFI(d)$  with  $d \in (\frac{3}{2}, \frac{5}{2})$  or stationary around a quadratic trend, and suppose that an empirical researcher decides in favor of the latter hypothesis so that one fits a quadratic trend,

$$y_t = \hat{\alpha} + \hat{\beta}t + \hat{\gamma}t^2 + \text{res.}, \quad (13)$$

and then works with the regression residuals. If the time series is truly an  $NFI(d)$  process with  $d \in (\frac{3}{2}, \frac{5}{2})$ , then, following the same steps as in

Theorem 1, it is not difficult to prove that  $\hat{\alpha}$  and  $\hat{\beta}$  diverge while  $T^{5/2-d}\hat{\gamma}$  has a nondegenerate limiting distribution and all the  $t$ -ratios diverge.

*Remark 2.* Many time series clearly are not fractionally integrated processes about a zero mean. See, for instance, Hassler and Wolters (1995). The linear trend has widely been used to model the growing tendency in many economic series, on the ground that it represents some constant growth rate. In this respect, Durlauf and Phillips (1988) considered the effect of estimating model (1) when, in fact, the true model was a difference stationary process with drift,  $\Delta y_t = \mu + \eta_t$ ,  $\eta_t \sim I(0)$ , showing that the regression theory for this model was identical to the driftless case.

The intuition behind this result is that a difference stationary process with drift can be converted into a time trend plus a driftless difference stationary process

$$\Delta y_t = \mu + \eta_t \Leftrightarrow y_t = \Delta^{-1}\mu + \Delta^{-1}\eta_t = \mu t + \tilde{y}_t,$$

assuming zero initial conditions,  $y_0 = 0$ . Consequently, in the difference stationary case, regressing levels  $y_t$  against a time trend will generate results identical to those obtained in Theorem 1 where  $\mu = 0$ .

The same results will be obtained in the fractional case if we assume that  $y_t$  has been generated according to

$$y_t = \alpha + \beta t + y_t^0 = \gamma' \zeta_t + y_t^0$$

where  $\gamma = (\alpha, \beta)'$ ,  $\zeta_t = (1, t)'$  and  $y_t^0$  is an  $NFI(d)$  process generated under Assumption A. In this case, least squares estimation of  $\gamma$  in (1) yields

$$\hat{\gamma} - \gamma = \left( \sum \zeta_t \zeta_t' \right)^{-1} \left( \sum \zeta_t y_t^0 \right)$$

and by defining the diagonal matrix  $\mathfrak{J}_T = \text{diag}\{1, T\}$  jointly with the *CMT* and expression (2), we find that

$$T^{1/2-d} \mathfrak{J}_T^{-1} (\hat{\gamma} - \gamma) \Rightarrow \left( \int \tau \tau' \right)^{-1} \left( \int \tau y_\infty \right),$$

where  $\tau = (1, r)'$  so that

$$T^{1/2-d} (\hat{\alpha} - \alpha) \Rightarrow -6 \int_0^1 (r - 2/3) y_\infty(r) dr$$

and

$$T^{3/2-d} (\hat{\beta} - \beta) \Rightarrow 12 \int_0^1 (r - 1/2) y_\infty(r) dr$$

as in Theorem 1. Consequently, if we only erroneously ignore stochastic regressors in model (1) but we correctly specified the deterministic components, then the regression theory is identical to the case where  $y_t$  is assumed to be generated without these nonstochastic trends.

### 3. Frequency domain estimation

The spurious problem can also be seen from a frequency domain point of view since there is a complete equivalence between a regression model among variables expressed in the time domain and a regression model in the frequency domain using the Fourier transform of such variables.

Trending processes are characterized by low frequency behavior, so the periodogram of nonstationary series shows very high power at low frequencies, and hence, it is quite plausible to get meaningful regressions among these variables, even if they are in fact independent of each other. Due to this dominance of low frequencies, regressions can be carried out in only a narrow band around zero frequency neglecting high frequency behavior. See Robinson and Marinucci (1998, 2000) for local regressions involving fractionally integrated time series. In this case, studentization of the regression coefficients should also consider the differential contribution across frequencies unlike the standard least squares variance estimate for uncorrelated observations.

In order to study these possibilities, in this section we consider again the estimation of model (1), where  $y_t$  is an *NFI* process of order  $d \in (\frac{1}{2}, \frac{3}{2})$ , and define the class of local estimates  $\hat{\beta}_N$ ,

$$\hat{\beta}_N = \left( \sum_{j=1}^N I_{tt}(\lambda_j) \right)^{-1} \sum_{j=1}^N \Re I_{ty}(\lambda_j), \quad (14)$$

where  $1 \leq N \leq T/2$ ,  $\Re$  stands for the real part,  $\lambda_j = 2\pi j/T$  are the Fourier frequencies and the (cross) periodogram is defined as

$$I_{ab}(\lambda_j) = w_a(\lambda_j)w_b(-\lambda_j), \quad w_a(\lambda_j) = (2\pi T)^{-1/2} \sum_{t=1}^T a_t \exp(it\lambda_j).$$

Then,  $\hat{\beta}_{[T/2]} = \hat{\beta}_{T-1} = \hat{\beta}$  (assuming  $T$  is odd) by symmetry, but if  $N \ll T/2$ , then we are only using information at low frequencies. On the other hand, dropping the zero frequency  $\lambda_0$  accounts for estimation with intercept by mean correction, while a symmetric sum around  $j=0$  would also eliminate the contribution of the imaginary parts of the cross-periodogram  $I_{ty}(\lambda_j)$ . Thus, we can set  $\hat{\alpha}_N = \bar{y} - \hat{\beta}_N \bar{t}$  to estimate the intercept. Furthermore, along this section we define  $\hat{u}_t = y_t - \hat{\alpha}_N - \hat{\beta}_N t$ . In the next theorem we estimate the contribution of high frequencies in parameter estimation.

*Theorem 2. Under assumption A, as  $T, N \rightarrow \infty, N \leq T/2$ ,*

$$T^{3/2-d}(\hat{\beta}_N - \hat{\beta}) = O_p \left( N^{-\min\{d,1\}} \right), \quad (15)$$

$$T^{1/2-d}(\hat{\alpha}_N - \hat{\alpha}) = O_p \left( N^{-\min\{d,1\}} \right), \quad (16)$$



Robinson and Marinucci (2000) have shown for a more general class of regressors that this contribution is indeed  $o_p(1)$  as long as the number of frequencies  $N$  increases with  $T$  as slowly as desired, as in our Theorem 2, which additionally provides a probability rate. From Theorem 2, it turns out that expressions (4)–(7) also hold for the statistics defined with  $\hat{\beta}_N$  and  $\hat{\alpha}_N$ . No apparent gains to alleviate the spurious problem can then be achieved by using the restricted version in the frequency domain of least squares herein proposed, at least without adjusting variance estimates.

In this sense, the natural estimate of the variance of  $\hat{\beta}_N$  in the frequency domain is given by

$$\hat{V}_N = \frac{1}{2} \left( \sum_{j=1}^N I_{tt}(\lambda_j) \right)^{-2} \sum_{j=1}^N I_{tt}(\lambda_j) I_{\hat{\beta}\hat{\beta}}(\lambda_j),$$

which assumes that the cross-periodogram ordinates are asymptotically uncorrelated at different frequencies, the  $\frac{1}{2}$  factor showing up by considering only real parts in  $\hat{\beta}_N$ .

*Theorem 3.* Under assumption A, as  $T, N \rightarrow \infty, N \leq T/2$ ,

$$T^{3-2d} \hat{V}_{[T/2]} \Rightarrow 288 \int_0^1 h(s) \gamma_\infty^R(s) ds, \quad (17)$$

$$T^{3-2d} (\hat{V}_{[T/2]} - \hat{V}_N) = O_p(N^{-1}), \quad (18)$$

$$t_N := \frac{\hat{\beta}_N}{(\hat{V}_N)^{1/2}} \Rightarrow \frac{\int_0^1 (r - \frac{1}{2}) y_\infty(r) dr}{(2 \int_0^1 h(s) \gamma_\infty^R(s) ds)^{1/2}}, \quad (19)$$

with  $h(s) = g(s) + 2g(1-s)$ ,  $g(x) = (\frac{1}{12}) + (x^3/6) - (x/4)$  and

$$\gamma_\infty^R(x) = \int_0^{1-x} \{y_\infty^D(r) - (r - 0.5)\beta_\infty\} \{y_\infty^D(r+x) - (r+x - 0.5)\beta_\infty\} dr,$$

where  $y_\infty^D(r) = y_\infty(r) - \int_0^1 y_\infty(r) dr$  and  $\beta_\infty = \int_0^1 (r - 0.5) y_\infty(r) dr$ .

As in Theorem 2, from (18) for variance estimation it is enough to asymptotically consider a small band around zero frequency, but the  $t$ -ratios are now  $O_p(1)$  and have a nondegenerated distribution. In fact, using (8) and denoting by  $\sigma(y_\infty(r))$  the  $\sigma$ -algebra generated by  $y_\infty(r)$ , the conditioned distribution of the local  $t$ -ratio in expression (19) has a mixed normal distribution given by

$$N\left(0, \frac{c^2 \Theta_{1/2}}{2 \int_0^1 h(s) \gamma_\infty^R(s) ds}\right).$$

Table 2

Percentage of rejections (5% size, two-sided  $t$ -test based on  $t_N$ , standard normal critical values) of the null  $H_0: \beta = 0$  in model (1)<sup>a</sup>

Sample size	Bandwidth	$d = 0.8$	$d = 1$	$d = 1.2$
$T = 100$	$N = T^{0.3} = 4$	43.44	54.95	64.50
	$N = T^{0.5} = 10$	46.72	56.83	65.89
	$N = T/2 = 50$	47.95	57.99	67.11
$T = 250$	$N = T^{0.3} = 5$	45.36	55.44	64.54
	$N = T^{0.5} = 16$	46.95	57.10	66.24
	$N = T/2 = 125$	47.74	57.91	67.12
$T = 1000$	$N = T^{0.3} = 8$	42.62	51.92	60.75
	$N = T^{0.5} = 32$	43.76	53.54	62.72
	$N = T/2 = 500$	44.31	54.11	63.40

<sup>a</sup>True model:  $\Delta^d y_t = \varepsilon_t$ ,  $\varepsilon_t \sim N(0, 1)$ . 10,000 replications. Frequency domain estimation and studentization.

Alternatively, we could use time domain estimates that take into account the correlation at all lags, e.g.

$$\tilde{V}_T = T \left( \sum_{t=1}^T (t - \bar{t})^2 \right)^{-2} \sum_{j=1-T}^{T-1} \hat{\gamma}_{\hat{u}\hat{u}}(j) \hat{\gamma}_{\hat{u}}(j),$$

where  $\hat{\gamma}_{\hat{u}\hat{u}}(j) = T^{-1} \sum_{t=1}^{T-|j|} \hat{u}_t \hat{u}_{t+|j|}$ ,  $\bar{u} = 0$ , and  $\hat{\gamma}_{\hat{u}}(j)$  are the usual sample autocovariances at lag  $j$ . Smoothed versions of this estimate are common for studentization in the econometric literature (see Robinson and Velasco, 1997, for a review). Smoothed estimates include only small lag autocovariances for consistent estimation in stationary environments, though Robinson (1998) shows that  $\tilde{V}_T$  may remain consistent without smoothing when the sample autocovariances decay to zero fast enough. In our setup, thus, the conclusions of Theorem 3 remain valid for the  $t$ -ratio  $\tilde{t}_N := \hat{\beta}_N / \sqrt{\tilde{V}_N}$  just replacing  $h(s)$  by  $g(s)$  in the asymptotic distribution of (19).

To analyze the properties of  $t_N$  we perform a complementary Monte Carlo analysis with the same time series that produced Table 1. We consider the following choices of the bandwidth number  $N$ :  $N = T/2$  (for *OLS* estimation with a full-band studentization),  $N = \lceil T^{0.3} \rceil$  and  $N = \lceil T^{0.5} \rceil$ .

From Table 2, we read that the rejection probabilities of  $t_N$  compared to standard normal critical values increase with  $d$  but they are fairly stable with  $N$  given  $d$ , in agreement with our asymptotic results. These rejection probabilities, in turn, are inappropriately large, though they are clearly smaller than those obtained when using the conventional least squares  $t$ -test.

As noted above, the results in Table 2 are obtained for a two-sided test using the  $\pm 2$  standard 5% normal critical values. However, since  $t_N$  has a

Table 3  
Distribution of  $t_N^a$

Quantiles	$d = 0.6$	0.8	1	1.2	1.4
0.50	0.022	0.047	0.068	0.039	0.019
0.90	3.080	4.037	5.474	7.391	9.752
0.95	4.278	5.695	7.894	10.841	14.370
0.975	5.448	7.433	10.281	14.919	19.358
0.995	8.086	11.485	17.451	25.123	33.239

<sup>a</sup>True model:  $\Delta^d y_t = \varepsilon_t$ ,  $\varepsilon_t \sim N(0, 1)$ .  $T = 1000$ ; 10,000 replications.

well-defined limiting distribution, by simulation we obtained the corresponding critical values for some values of  $d$  in the interval  $(\frac{1}{2}, \frac{3}{2})$  assuming that the true model is given by a nonstationary fractional white noise  $\Delta^d Y_t = \varepsilon_t$  with  $\varepsilon_t \sim N(0, 1)$ . According to expression (19), the distribution of  $t_N$  is symmetric around zero. So, in Table 3 we only present some of the most empirically relevant upper quantiles.

According to Table 3, the distribution of  $t_N$  has an increasing dispersion the larger the memory parameter  $d$  is. Moreover, we find that a quadratic relationship between the critical values and the memory parameters  $d$  included in Table 3 fits the data of Table 3 (adjusted  $R^2$  greater than 0.99 in all cases) very well, so that one can easily obtain the critical value for any  $d \in (\frac{1}{2}, \frac{3}{2})$  by (quadratic) interpolation. For instance, the regressions for the  $\alpha = 1\%$ , 5% and 10% critical values ( $c_\alpha$ ) for *two-sided tests* are given by

$$c_1 = 5.409 - 7.819d + 19.896d^2, \quad (20)$$

$$c_5 = 4.836 - 6.263d + 11.958d^2, \quad (21)$$

and

$$c_{10} = 4.119 - 5.092d + 8.879d^2. \quad (22)$$

So, for example, the approximated 5% critical value for testing the null  $H_0: \beta = 0$  against  $H: \beta \neq 0$  in model (1) by means of  $t_N$  when  $d = 0.7$  is 6.311 according to Eq. (21). The effect of this critical value estimation procedure is expected to be negligible in inference compared to sample variability.

To assess the performance in finite samples of  $t_N$  compared with the customary least squares  $t$ -ratio,  $t_\beta$ , we conducted a small Monte Carlo experiment. Table 4 collects the percentage of rejections of the null  $H_0: \beta = 0$  against  $H_1: \beta \neq 0$  in model (1) when the true model is  $\Delta^d y_t = \varepsilon_t$ ,  $\varepsilon_t \sim N(0, 1)$  for some values of  $d$  and for a small sample of size  $T = 100$ . We considered four critical values:  $\pm 2$  refers to the standard normal 5% size critical value, whereas 1%, 5% and 10% refer to the critical values obtained by means of expressions (20)–(22), respectively. The upper part of Table 4 shows the percentage of rejections obtained when using the least squares variance, whereas

Table 4

Percentage of rejections of the null  $H_0: \beta = 0$  in model (1)<sup>a</sup>

Nominal sizes	$d = 0.7$	0.9	1.1	1.3
<i>OLS studentization</i>				
5%- $N(0, 1)$	75.98	84.56	89.97	93.02
	76.37	84.64	90.08	93.04
	77.19	85.06	90.43	93.08
10%	45.68	52.11	55.73	58.64
	46.54	52.51	56.05	58.86
	47.83	53.20	56.39	59.10
5%	33.70	39.84	44.20	47.07
	34.65	40.30	44.35	47.36
	36.10	40.97	45.09	47.49
1%	14.56	17.98	21.39	23.93
	15.24	18.32	21.54	24.14
	16.35	19.04	21.90	24.37
<i>Frequency domain studentization</i>				
5%- $N(0, 1)$	42.51	52.96	62.87	71.08
	41.06	51.68	61.71	69.91
	40.12	50.09	59.77	68.36
10%	9.76	10.41	10.57	10.42
	9.33	9.57	9.89	9.69
	9.46	9.02	9.18	8.89
5%	4.85	4.59	5.10	5.16
	4.61	4.28	4.82	4.82
	4.86	4.20	4.38	4.35
1%	1.06	0.88	1.01	0.97
	1.04	0.72	0.92	0.83
	1.22	0.72	0.94	0.76

<sup>a</sup>True model:  $\Delta^d y_t = \varepsilon_t$ ,  $\varepsilon_t \sim N(0, 1)$ . 10,000 replications.  $T = 100$ . Numbers in the table are explained in the main text. 5%- $N(0, 1)$  refers to the (approximate) 5% critical value of the standard normal distribution in a two-sided test, i.e.,  $\pm 2$ .

the lower part of Table 4 delivers the percentage of rejections obtained when using  $\hat{V}_N$ . Finally, each cell in Table 4 has three numbers giving (from top to the bottom) the percentage of rejections when  $N = T/2$  (customary least squares  $t$ -ratio)  $N = 10$  and 4.

All the experimental findings in Table 4 are in agreement with the theoretical results developed in this paper, namely, that standard *OLS* inference remains invalid when spuriously detrending an *NFI* process and that, in order to improve the chance of finding nonsense relationships, one can resort to computationally convenient local versions in the frequency domain of the *OLS* methodology whenever the studentization takes into account the dependence structure at the relevant frequencies and with the appropriate set of

critical values. In this latter case, from Table 4 we learn that the empirical sizes of the test are very close to the nominal ones.

#### 4. Concluding remarks

The aim of this paper has been to provide a generalization of the available results on the behavior of difference weakly stationary processes which are misspecified as trend stationary time series to the more general framework of misspecified *NFI* processes of order  $d \in (\frac{1}{2}, \frac{3}{2})$ .

Several conclusions can be drawn from our study. First, the usual least squares estimated time trend coefficient converges in probability to zero. Second, the corresponding conventional *t*-Student statistics diverge at the  $T^{1/2}$  rate, independent of  $d$ . Consequently, they will, with probability one, as  $T$  grows large, reject the null hypothesis of no significance. Hence, standard *OLS* inference remains invalid, as in the difference stationary ( $d = 1$ ) case. Third, in order to improve the chance of finding nonsense relationships, there are no gains from using restricted versions in the frequency domain of the *OLS* methodology unless studentization takes into account the dependence structure at the relevant frequencies, and even in this latter case, standard inference remains invalid, unless we use corrected critical values.

When one is interested in testing whether the process of interest is trend stationary or *NFI* of known order  $d_0 \in (\frac{1}{2}, \frac{3}{2})$ , one can use  $t_N$  along with the corresponding critical values obtained from expressions (20)–(22). This is especially relevant for the unit root case where  $d_0 = 1$ . In this case the critical values can be readily obtained from Table 3.

For other values of  $d \in (\frac{1}{2}, \frac{3}{2})$ , the studentization methodology herein developed can be combined with simulated critical values together with consistent estimates of  $d$ . These can be obtained through semiparametric estimation based either on higher-order tapered data, which is simultaneously robust to linear trends and to nonstationarity (cf. Velasco, 1999a, b) or on residuals, as in Hassler et al. (2000), though these authors used a different definition of nonstationary fractionally integrated processes (cf. Marinucci and Robinson, 1999).

#### Acknowledgements

We are greatly indebted to Professor Peter M. Robinson and two anonymous referees for helpful comments and suggestions. F. Marmol's research supported by DGICYT grants PB98-0026 and SEC2001-0890. C. Velasco's research supported by DGICYT grants PB98-0025 and BEC2001-1270.

## Appendix. Mathematical proofs

*Proof of Theorem 1.* Using (2) and the continuous mapping theorem (CMT), we find that

$$\begin{aligned} & T^{1/2-d} \hat{\alpha} \\ &= (T^{-4} a_T)^{-1} \left[ T^{-3} \sum t^2 T^{-1/2-d} \sum y_t - T^{-3/2-d} \sum t y_t T^{-2} \sum t \right] \\ &\Rightarrow 12 \left[ \frac{1}{3} \int y_\infty - \frac{1}{2} \int r y_\infty \right] = -6 \int \left( r - \frac{2}{3} \right) y_\infty. \end{aligned} \quad (\text{A.1})$$

and

$$\begin{aligned} & T^{3/2-d} \hat{\beta} = (T^{-4} a_T)^{-1} \left[ T^{-3/2-d} \sum t y_t - T^{-2} \sum t T^{-1/2-d} \sum y_t \right] \\ &\Rightarrow 12 \left[ \int r y_\infty - \frac{1}{2} \int y_\infty \right] = 12 \int \left( r - \frac{1}{2} \right) y_\infty. \end{aligned} \quad (\text{A.2})$$

Consider now the  $t$ -Student statistics,

$$t_{\hat{\beta}}^2 = \frac{\hat{\beta}^2}{T^{-1} \sum \hat{u}_t^2 / \sum (t - \bar{t})^2}, \quad t_{\hat{\alpha}}^2 = \frac{\hat{\alpha}^2}{T^{-1} \sum \hat{u}_t^2 \sum t^2 / T \sum (t - \bar{t})^2}.$$

Using (2) and the CMT, we get the following results:

$$\begin{aligned} & T^{-2d} \sum (y_t - \bar{y})^2 = T^{-2d} \sum y_t^2 - \left( T^{-1/2-d} \sum y_t \right)^2 \\ &\quad \Rightarrow \int y_\infty^2 - \left\{ \int y_\infty \right\}^2, \end{aligned} \quad (\text{A.3})$$

$$T^{-3/2-d} \sum (y_t - \bar{y})(t - \bar{t}) \Rightarrow \int \left( r - \frac{1}{2} \right) y_\infty, \quad (\text{A.4})$$

and

$$\begin{aligned} & T^{-2d} \sum \hat{u}_t^2 = T^{-2d} \sum (y_t - \bar{y})^2 - T^{3/2-d} \hat{\beta} T^{-3/2-d} \sum (y_t - \bar{y})(t - \bar{t}) \\ &\quad \Rightarrow \int y_\infty^2 - \left\{ \int y_\infty \right\}^2 - 12 \left\{ \int \left( r - \frac{1}{2} \right) y_\infty \right\}^2, \end{aligned} \quad (\text{A.5})$$

where  $\bar{y} = T^{-1} \sum y_t$  and  $\bar{t} = T^{-1} \sum t$ .

Now, from (4), (5), (A.5) and the CMT, it is straightforward to show that

$$\begin{aligned} & T^{-1} t_{\hat{\beta}}^2 = \frac{T^{3-2d} \hat{\beta}^2}{T^{-2d} \sum \hat{u}_t^2 / T^{-3} \sum (t - \bar{t})^2} \\ &\quad \Rightarrow \frac{12 \{ \int r y_\infty - 0.5 \int y_\infty \}^2}{\int y_\infty^2 - \{ \int y_\infty \}^2 - 12 \{ \int (r - \frac{1}{2}) y_\infty \}^2}, \end{aligned}$$

and

$$\begin{aligned} T^{-1}t_\alpha^2 &= \frac{T^{1-2d} \hat{\alpha}^2}{T^{-2d} \sum \hat{u}_t^2 T^{-3} \sum t^2/T^{-3} \sum (t-\bar{t})^2} \\ &\Rightarrow \frac{9\{\int(r-\frac{2}{3})y_\infty\}^2}{\int y_\infty^2 - \{\int y_\infty\}^2 - 12\{\int(r-\frac{1}{2})y_\infty\}^2}. \quad \square \end{aligned}$$

*Proof of Theorem 2.* We have for  $T^* = [T/2]$  assuming  $T$  is odd (otherwise set  $T^* = T/2 - 1$  and include periodograms at frequency  $j = T/2$  in all sums with weight  $\frac{1}{2}$  so that all arguments go through with minor modifications),

$$\hat{\beta}_N = \left( \sum_{j=1}^{T^*} I_u(\lambda_j) - \sum_{j=N+1}^{T^*} I_u(\lambda_j) \right)^{-1} \left( \sum_{j=1}^{T^*} \Re I_{ty}(\lambda_j) - \sum_{j=N+1}^{T^*} \Re I_{ty}(\lambda_j) \right). \quad (\text{A.6})$$

Now, as  $T \rightarrow \infty$ ,

$$\sum_{j=1}^{T^*} I_u(\lambda_j) = \frac{1}{2} \sum_{j=1}^{T^*} I_{t-\bar{t}, t-\bar{t}}(\lambda_j) = \pi \sum_{t=1}^{T^*} (t-\bar{t})^2 \asymp \frac{\pi}{12} T^3, \quad (\text{A.7})$$

with the symbol  $\asymp$  denoting asymptotic equivalence, i.e.,  $a_T \asymp b_T$  if  $a_T/b_T \rightarrow 1$  as  $T$  grows arbitrarily large, whereas

$$\sum_{t=1}^T t \exp(it\lambda_j) = \frac{-e^{i\lambda_j T}}{1 - e^{i\lambda_j}}, \quad I_u(\lambda_j) = \frac{T}{2\pi} |1 - e^{i\lambda_j}|^{-2} = O(T^3 j^{-2}), \quad (\text{A.8})$$

$1 \leq j \leq T^*$ , as  $\lambda_j \rightarrow 0$ , so

$$\sum_{j=N+1}^{T^*} I_u(\lambda_j) = O\left( \sum_{j=N+1}^T T^3 j^{-2} \right) = O(T^3 N^{-1}) = o(T^3)$$

as  $N \rightarrow \infty$ .

On the other hand, as  $T \rightarrow \infty$ ,

$$\sum_{j=1}^{T^*} \Re I_{ty}(\lambda_j) = \frac{1}{2} \sum_{j=1}^T I_{t-\bar{t}, y}(\lambda_j) = \pi \sum_{t=1}^T (t-\bar{t}) y_t = O_p(T^{3/2+d}) \quad (\text{A.9})$$

by (A.4), while  $\sum_{t=N+1}^{T^*} \Re I_{ty}(\lambda_j)$  has zero mean and variance equal to the real part of

$$\begin{aligned} &\frac{1}{2} \sum_{j=N+1}^{T^*} \sum_{k=N+1}^{T^*} \{w_t(-\lambda_j) w_t(\lambda_k) E[w_y(\lambda_j) w_y(-\lambda_k)] \\ &\quad + w_t(\lambda_j) w_t(\lambda_k) E[w_y(-\lambda_j) w_y(-\lambda_k)]\}. \end{aligned} \quad (\text{A.10})$$

From Proposition 4.1 of Robinson and Marinucci (1998), we obtain  $E[I_{yy}(\lambda_j)] = O(\lambda_j^{-2d})$  if  $\frac{1}{2} < d < 1$  and  $O(\lambda_j^{-2d} j^{2(d-1)})$  if  $1 \leq d < \frac{3}{2}$ . Then, applying Cauchy inequality, it follows that  $E|w_y(\lambda_j)w_y(-\lambda_k)| = O(T^{2d}(jk)^{-\min\{d,1\}})$ , for  $1 \leq j, k \leq T^*$  so that, as  $N \rightarrow \infty$  and using (A.8), (A.10) is

$$\begin{aligned} & O\left(\sum_{j=N+1}^T \sum_{k=N+1}^T T^3(jk)^{-1} T^{2d}(jk)^{-\min\{d,1\}}\right) \\ &= O(T^{3+2d} N^{-2\min\{d,1\}}). \end{aligned} \quad (\text{A.11})$$

Therefore, from (A.6), (A.7), (A.9) and (A.11),

$$\begin{aligned} \hat{\beta}_N &= \left( \sum_{j=1}^{T^*} \Re I_{ly}(\lambda_j) [1 + O_p(N^{-\min\{1,d\}})] \right) \\ &\quad \left( \sum_{j=1}^{T^*} I_u(\lambda_j) \right)^{-1} [1 + O_p(N^{-1})] \\ &= \hat{\beta} [1 + O_p(N^{-\min\{d,1\}})], \end{aligned}$$

as  $N \rightarrow \infty$ . Finally, the result for  $\hat{\alpha}_N$  follows from that for  $\hat{\beta}_N$ .  $\square$

*Proof of Theorem 3.* First, we have

$$\begin{aligned} \hat{V}_{T^*} &= \left( \sum_{j=1}^{T-1} I_u(\lambda_j) \right)^{-2} \sum_{j=1}^{T-1} I_u(\lambda_j) I_{\hat{u}\hat{u}}(\lambda_j) \\ &= T \left( \sum_{t=1}^T (t - \bar{t})^2 \right)^{-2} \sum_{j=1}^{T-1} \hat{\gamma}_{\hat{u}\hat{u}}(j) \{ \hat{\gamma}_{\hat{u}}(j) + 2\hat{\gamma}_{\hat{u}}(T - |j|) \}, \end{aligned}$$

using the fact that  $\sum_{t=1}^T \exp(it\lambda_j) = 0$ ,  $j \neq 0 \pmod{T}$ . Then

$$\hat{V}_{T^*} = 2T \left( \sum_{t=1}^T (t - \bar{t})^2 \right)^{-2} \sum_{j=1}^{T-1} \hat{\gamma}_{\hat{u}\hat{u}}(j) \{ \hat{\gamma}_{\hat{u}}(j) + 2\hat{\gamma}_{\hat{u}}(T - j) \} + O_p(T^{2d-4}), \quad (\text{A.12})$$

since from (A.2) we find that  $T(\sum_{t=1}^T (t - \bar{t}))^{-2} \propto (12)^{-2} T^{-5} = O(T^{-5})$ ,  $\hat{\gamma}_{\hat{u}}(0) = O(T^2)$ , and

$$\hat{\gamma}_{\hat{u}\hat{u}}(0) = \hat{\gamma}_{yy}(0) - T^{-1} \sum_{t=1}^T \{ 2\hat{\beta}(y_t - \bar{y})(t - \bar{t}) - \hat{\beta}^2(t - \bar{t})^2 \} = O_p(T^{2d-1}),$$

using (A.3) and Theorems 1 and 2.



Next, for any  $x \in [0, 1]$  with  $r + x \in [0, 1]$  and setting  $j = [Tx]$ , we find that

$$\lim_{T \rightarrow \infty} \hat{\gamma}_t(j) = g(x),$$

and applying the *CMT*,

$$\begin{aligned} T^{1-2d} \hat{\gamma}_{yy}(j) &= T^{-2d} \left\{ \sum_{t=1}^{T-|j|} y_t y_{t+|j|} - \bar{y} \sum_{t=1}^{T-|j|} (y_{t+|j|} + y_t) + [1 - |j|] \bar{y}^2 \right\} \\ &\Rightarrow \int_0^{1-x} y_\infty(r) y_\infty(r+x) dr + (1-x) \left( \int_0^1 y_\infty(r) dr \right)^2 \\ &\quad - \int_0^1 y_\infty(r) dr \int_0^{1-x} (y_\infty(r+x) + y_\infty(r)) dr = \gamma_\infty(x). \end{aligned}$$

Therefore,

$$\begin{aligned} T^{1-2d} \hat{\gamma}_{\hat{u}\hat{u}}(j) &= T^{1-2d} \hat{\gamma}_{yy}(j) - T^{-2d} \sum_{t=1}^{T-|j|} \{ \hat{\beta}[(y_{t+|j|} - \bar{y})(t - \bar{t}) \\ &\quad + (y_t - \bar{y})(t + |j| - \bar{t})] - (t - \bar{t})(t + |j| - \bar{t}) \} \\ &\Rightarrow \int_0^{1-x} \{ y_\infty(r) + (r - 0.5)(r + x - 0.5) \} dr \\ &\quad - \beta_\infty \int_0^{1-x} \left\{ \left[ y_\infty(r+x) - \int_0^1 y_\infty(z) dz \right] [r - 0.5] \right. \\ &\quad \left. + \left[ y_\infty(r) - \int_0^1 y_\infty(z) dz \right] [r + x - 0.5] \right\} dr = \gamma_\infty^R(x). \end{aligned}$$

Then using (A.12) and the definitions of  $h(x)$  and  $g(x)$ , it follows that  $T^{3-2d} \hat{\gamma}_T^*$ , has the same asymptotic distributions as

$$288T^{-1} \sum_{j=1}^{T-1} h(j/T) T^{1-2d} \hat{\gamma}_{\hat{u}\hat{u}}(j) \Rightarrow 288 \int_0^1 h(x) \gamma_\infty^R(x) dx,$$

by the *CMT*. The distribution of the  $t$ -ratio in (19) follows now from Theorems 1 and 2, while (18) can be shown using the method of the proof of Theorem 2.  $\square$

## References

Akonon, J., Gourieroux, C., 1987. A functional central limit theorem for fractional processes. Discussion Paper 8801, CEPREMAP, Paris.

- Baillie, R., 1996. Long memory processes and fractional integration in econometrics. *Journal of Econometrics* 73, 5–59.
- Beran, J., 1994. *Statistics for Long Memory Processes*. Chapman & Hall, New York.
- Chambers, M.J., 1996. Fractional integration, trend stationarity and difference stationarity. Evidence from some U.K. macroeconomic time series. *Economics Letters* 50, 19–24.
- Durlauf, S.N., Phillips, P.C.B., 1988. Trends versus random walks in time series analysis. *Econometrica* 56, 1333–1354.
- Hassler, U., Marmol, F., Velasco, C., 2000. Residual log-periodogram inference for long-run relationships. Preprint, Universidad Carlos III de Madrid.
- Hassler, U., Wolters, J., 1995. Long memory in inflation rates: Empirical evidence. *Journal of Business and Economic Statistics* 13, 37–45.
- Mandelbrot, B.B., Van Ness, J.W., 1968. Fractional brownian motions, fractional noises and applications. *SIAM Review* 10, 422–437.
- Marinucci, D., Robinson, P.M., 1999. Alternative forms of fractional brownian motion. *Journal of Statistical Planning and Inference* 80, 111–122.
- Marinucci, D., Robinson, P.M., 2000. Weak convergence of multivariate fractional processes. *Stochastic Processes and their Applications* 86, 103–120.
- Marmol, F., 1998. Spurious regression theory with nonstationary fractionally integrated processes. *Journal of Econometrics* 84, 232–250.
- Nelson, C.R., Plosser, C.I., 1982. Trends and random walks in macroeconomic time series. *Journal of Monetary Economics* 10, 139–162.
- Phillips, P.C.B., 1987. Time series regression with a unit root. *Econometrica* 55, 277–301.
- Robinson, P.M., 1994. Time series with strong dependence. In: Sims, C.A. (Ed.), *Advances in Econometrics*, Vol. 1. Cambridge University Press, Cambridge, pp. 47–95.
- Robinson, P.M., 1998. Inference-without-smoothing in the presence of nonparametric autocorrelation. *Econometrica* 66, 1163–1182.
- Robinson, P.M., Marinucci, D., 1998. Semiparametric frequency domain analysis of fractional cointegration. STICERD Discussion Paper EM-348, LSE.
- Robinson, P.M., Marinucci, D., 2000. The averaged periodogram for nonstationary vector time series. *Statistical Inference for Stochastic Processes*, Vol. 3, 149–160.
- Robinson, P.M., Velasco, C., 1997. Autocorrelation-robust inference. In: Maddala, G.S., Rao, C.R. (Eds.), *Handbook of Statistics*, Vol. 15 (Robust Inference). North-Holland, Amsterdam, pp. 267–298.
- Samorodnitsky, G., Taqqu, M.S., 1994. *Stable Non-Gaussian Random Processes*. Chapman & Hall, New York.
- Velasco, C., 1999a. Non-stationary log-periodogram regression. *Journal of Econometrics* 91, 325–371.
- Velasco, C., 1999b. Gaussian semiparametric estimation of non-stationary time series. *Journal of Time Series Analysis* 20, 87–127.