A Simple Test of Long-memory vs. Structural Breaks in the Time Domain: What is what?

By Juan J. Dolado[a], Jesus Gonzalo[a], and Laura Mayoral[b] *

[a]Dept. of Economics, Universidad Carlos III de Madrid.

[b]Dept. of Economics, Universidat Pompeu Fabra

March 6, 2004

Abstract

This paper proposes a time-domain test of a process being $I(d)$, $0 < d < 1$, under the null against the alternative of being $I(0)$ with deterministic components subject to structural breaks at known or unknown dates. Denoting by $A_B(t)$ the different types of structural breaks in the deterministic component of a time series considered by Perron (1989), the test statistic proposed here is based on the t-ratio (or the infimum of a sequequence of t-ratios) of the estimated coefficient on y_{t-1} in an OLS regression of $\Delta^d y_t$ on $\Delta^d A_B(t)$ and y_{t-1}, possibly augmented by a suitable number of lags of $\Delta^d y_t$ to cater for autocorrelated errors. The statistic is labelled as the NFDF (new Fractional Dickey-Fuller) test since it is based on the same principles as the well-known Dickey-Fuller unit root test. Both its asymptotic behavior and finite sample properties are analyzed, and an empirical application is provided. The proposed NFDF test is computationally simple and presents a number of advantages over other available test statistics addressing a similar issue.

*We are grateful to MCYT for finacial support through grant SEC01-0890.

INTRODUCTION

The issue of distinguishing between a time-series process exhibiting long-range dependence (LRD) and one with short memory but suffering from structural shifts has been around for some time in the literature. The detection of LRD effects is often based on statistics of the underlying time series, such as the sample ACF, the periodogram, the R/S statistic, the rate of growth of the of variances of partial sums of the series, etc. In the statistical literature, however, it has been pointed out by some authors (see e.g., Bhattacharya *et al.* ,1983, and Teverosky and Taqqu ,1997), that statistics based on short memory perturbed by some kind of nonstationarity may display the same properties as those prescribed by LRD under alternative assumptions. In particular, the sample variance of aggregated time series and the R/S statistic may exhibit LRD type of behaviour when applied to short-memory processes affected by shifts in trends or in the mean. More recently, a similar observation regarding this identification problem has been made in the econometric literature when analyzing financial data. For example, Ding and Granger (1996) and Mikosch and Starica (1999) claim that the LRD behaviour detected in both the absolute squared log-returns of financial prices (bonds, exchange rates, options, etc.) may be well explained by changes in the parameters of one model to another over different subsamples due to significant events, such as the Great Depression of 1929, the oil-price shocks in the 1970s, the Black Monday of 1987 or the collapse of the EMS in 1992.

There does not exist a unique definition of LRD (see e.g., Beran, 1994, Baillie, 1986, and Brockwell and Davies, 1996). One possible way to define it for a stationary time series (y_t) is via the condition that $\lim_{j\to\infty}\sum_j \left|\rho_y(j)\right| = \infty$, where ρ_y denotes the ACF of sequence (y_t). Typically, for series exhibiting long-memory, this requires a hyperbolic decay of the autocorrelations instead of the standard exponential one. An equivalent form of expressing that property in the frequency domain is to require that the spectral density $f_y(\omega)$ of the sequence is asymptotically of order $L(\omega)\omega^{-d}$ for some $d > 0$ and a slowly varying function $L(.)$, as $\omega \uparrow 0$. Specifically, if, for some constants c_ρ and c_f, $\rho_y(j) \approx c_\rho\ j^{2d-1}$ for large j and $d \in (0, \frac{1}{2})$, $f_y(\omega) \approx c_f\ \omega^{-2d}$ for small frequencies ω, and the normalized partial sums of such a series converge to fractional

Brownian motion (fBM), then y_t is said to be fractionally integrated of order d, $I(d)$. An $I(d)$ process is defined as $\Delta^d y_t = \eta_t$, when $0 < d < 0.5$, $\Delta = 1 - L$ and η_t is an $I(0)$ error term, and as $(1 - L)y_t = \Delta^{1-d}\eta_t$, $0.5 \leq d < 1$ when $(1 - L)y_t = \Delta^{1-d}\eta_t$. Hence, fractional integration is a particular case of LRD.

To illustrate the above-mentioned source of confusion between long-memory and short memory processes subject to structural breaks, let us consider a simple example where y_t is generated by an $I(0)$ process subject to a break in its mean at date T_B

$$y_t = \alpha_1 + (\alpha_2 - \alpha_1)DU_t(\lambda) + u_t, \tag{1}$$

such that u_t is a stationary zero-mean process, $\lambda = T_B/T$ and $DU_t(\lambda) = \mathbf{1}(t > T_B)$, $1 < T_B < T$, is an indicator function.. Then, denoting the sample mean by $\overline{y}_T$, the sample autocovariances of the sequence (y_t) are given by

$$\widetilde{\gamma}_{T,y}(j) = \frac{1}{T}\sum_{t=1}^{T-j} y_t y_{t+j} - (\overline{y}_T)^2, \ j \in \aleph. \tag{2}$$

By the ergodic theorem it follows that for fixed $j \geq 0$, with $\lambda \in (0,1)$,as $T \uparrow \infty$

$$\widetilde{\gamma}_{T,y}(j) \to \gamma_u(j) + \lambda(1-\lambda)(\alpha_2 - \alpha_1)^2 \text{ a.s.} \tag{3}$$

From (3), even if the autocovariances $\gamma_u(j)$ decay to zero exponentially as $j \uparrow \infty$, the sample autocovariances, $\widetilde{\gamma}_y(j)$,approaches a positive constant given by the second term in (3), as long as $\alpha_2 \neq \alpha_1$, for longer lags, as if the process exhibited LRD. Note that this result can be easily generalized for multiple breaks in the mean (see Mikosch and Starica, 1999). In order to check the consequences of ignoring such a structural break, a small Monte Carlo experiment is performed by simulating 100 series of sample size $T = 20,000$ where y_t is generated according to (2), with $\lambda = 0.5$, $\alpha_1 = 0$, $\varepsilon_t \sim n.i.d.(0,1)$, and $(\alpha_2 - \alpha_1) = 0$ (no break), 0.2 (small break) and 0.5 (large break). Then, d is estimated, ignoring the break in the mean, by means of the Geweke and Porter-Hudak (GHP,1983) ´estimator at different frequencies $\omega_0 = 2\pi/g(T)$, including the popular choice in GPH estimation of $g(T) = T^{0.5}$. From the results shown in Table 1, it becomes clear that the estimates of d monotonically increase with the size of the break in the mean, giving the wrong impression that y_t

is $I(d)$ when clearly it is not. This is the source of confusion which has been stressed in the literature. This problem aggravates even more when the DGP contains a break in the trend. For example, using the same experiment with a DGP given by $y_t = \alpha_1 + \beta_1 DT(\lambda)_t + \varepsilon_t$, with $DT_t(\lambda) = (t - T^*)\mathbf{1}_{(T^*+1 \leq t \leq T)}$ and $\beta_1 = 0.1$ yields estimates of d in the range $(1.008, 1.031)$, depending on the choice of frequency, well in accord with the results of Perron (1989) about the lack of consistency of the DF test of a unit root in such a case.

Table 1

ESTIMATES OF d (DGP(1))

Frequency	$T^{0.5}$	$T^{0.45}$	$T^{0.4}$	$T^{0.35}$
$\alpha_2 - \alpha_1 = 0.0$	-0.004	-0.004	-0.003	-0.005
$\alpha_2 - \alpha_1 = 0.2$	0.150**	0.212**	0.298**	0.404**
$\alpha_2 - \alpha_1 = 0.5$	0.282**	0.3709**	0.477**	0.585**

****Rejection of the null hypothesis* $d = 0$ at 1% S.L.

Along the same argument, the previous results have been also pointed out by Granger and Hyung (1999) who propose an extreme version of the DGP in (1) where now y_t is assumed to be generated by

$$\begin{aligned} y_t &= m_t + \varepsilon_t, \\ \Delta m_t &= q_t \eta_t, \end{aligned} \tag{4}$$

with q_t following an *i.i.d.* binomial distribution such that $q_t = 1$ with probability p and $q_t = 0$ with probability $(1-p)$, and $\varepsilon_t \sim i.i.d(0, \sigma_\varepsilon^2)$, $\eta_t \sim i.i.d(0, \sigma_\eta^2)$. Then, $\text{var}(y_t) = tp\sigma_\eta^2 + \sigma_\varepsilon^2$, and it can be shown that the ACF verifies

$$\widetilde{\rho}_{T,y}(j) = \frac{\frac{c\sigma_\eta^2}{6}(1 - \frac{j}{T})(1 - 2\frac{j}{T})^2}{\frac{c\sigma_\eta^2}{6} + \sigma_\varepsilon^2}, \; j \in \aleph. \tag{5}$$

where $c = pT$ is the expected number of structural breaks in the sample period T. It is easy to check that if $0 < c < \infty$, then $\widetilde{\rho}_{T,y}(j) \to (1 + \frac{6\sigma_\varepsilon^2}{c\sigma_\eta^2})^{-1}$ as $T \uparrow \infty$ for fixed j, implying that the sample ACFs tend to stabilize around a positive value for long lags, again as if there were LRD

A further generalization of this process is the so-called error duration (ED) model proposed by Parke (1999) whereby

$$y_t = \sum_{s=-\infty}^{t} g_{s,t}\varepsilon_s, \tag{6}$$

where $g_{s,t} = \mathbf{1}(s \leq t \leq s+n_s)$ is now an indicator function for the event that a shock arising in period s survives until $s+n_s$. Assume that ε_s and $g_{s,t}$ are independent for all $t \geq s$ and that p_j is the probability of survival from s to $s+j$, i.e., $p_s = P(g_{s,s+j} = 1)$, such that $p_0 = 1$. Then, Parke (1999) shows that $\gamma_y(j) = \sigma_\varepsilon^2 \sum_{s=j}^{\infty} p_s$ and that var $(y_t) = \sigma_\varepsilon^2(1+\lambda)$ with $\lambda = \sum_{s=1}^{\infty} p_s$. Next, assuming that for some constant c_γ and $d > 0$, $\gamma_y(j)/c_\gamma j^{2d-1} \to 1$ as $j \uparrow \infty$, it can be easily checked that $\gamma_y(j) = O(T^{2d-1})$ yielding again a LRD property.

For data generating processes (DGP) such as (1) and (4), the property that the spectral density behaves as $\omega^{-\varsigma}, \varsigma > 0$, near zero, and therefore has a singularity at zero, holds. In effect, under (1), Mikosch and Starica (1999) show that for $\omega \uparrow 0$, $f_y(\omega) \approx f_u(\omega) + (T\omega^2)^{-1}(1-\cos(2\pi\lambda))(\alpha_1-\alpha_2)^2$, which explodes at zero if $\omega^2 = 2\pi T^{-\delta}$ with $\frac{1}{2} < \delta$. Likewise, if (4) holds, for $\omega \uparrow 0$, then $f_y(\omega) \approx f_\varepsilon(\omega) + c\omega^{-2}$, which explodes at frequency zero when $0 < c < \infty$.

In the same vein, Diebold and Inoue (2001) define LRD in terms of the growth rate of variances of partial sums, i.e., $\text{var}(S_T) = O(T^{2d+1})$ with $S_T = \sum_{t=1}^{T-j} y_t$ and $0 < d < 1$. For a DGP like (4), if $p = O(T^{2d-2})$ so that $c = pT = O(T^{2d-1})$, namely the expected number of breaks goes to zero as $T \uparrow \infty$, then $\text{var}(S_T) = O(T^{2d+1})$ as if it were an $I(d)$ process.

Nonetheless, Davidson and Sibbersten (2003) have recently pointed out that the normalized partial sums of (y_t) generated by the ED model in (6) do not converge to fBM. Hence, despite sharing some properties with I(d) processes, the ED model is not able to reproduce this key result of I(d) processes. The intuition behind this result lies in that $\Delta y_t = \varepsilon_t + \sum_{s=-\infty}^{t-1} \Delta g_{st}$, where $\Delta g_{st} = g_{s,t} - g_{s,t-1} = g_{st} - 1$, since survival from period s to period t clearly implies survival to $t-1$. Hence, even with Gaussian shocks, a linear representation like for an $I(d)$ process cannot be obtained since the number of nonzero terms for $s < t$ is a random variable with mean falling between zero (when $g_{st} = 1$) and -1 (when $g_{st} = 0$). However, Taqqu *et al.* (1997) have shown that cross-sectional aggregation of processes generated by (6), suitably normalized, does converge to fBM. In effect, choosing M independent copies of y_t, i.e., $y_t^{(1)}, ..., y_t^{(M)}$, and defining $Y_t^{(M)} = M^{-1/2} \sum_{i=1}^{M} y_t^{(i)}$ and $\sigma_T^2 = \sum_{l=1}^{T} \sum_{m=1}^{T} \gamma_{|l-m|} = E(\sum_{t=1}^{T} Y_t^{(M)})$, it follows that $Y_T^{(M)}(r) = \sigma_T^{-1} \sum_{t=1}^{[Tr]} Y_t^{(M)}, 0 \leq r \leq 1$ and $[x]$ being largest integer not exceeding x, converges in distribution to a fBM with $\sigma_T = O(T^{d+\frac{1}{2}})$.

Summing up, all the models described so far are nonlinear models capable of reproducing some observationally equivalent characteristics of I(d) processes , albeit *not all.* Given that among all the models sharing properties with long-memory processses, the ones having more impact on empirical research are those popularized by Perron, where the deterministic components of $I(0)$ processes are subject to structural breaks , our aim in this paper is to devise statistical procedures to distinguish them from long-memory ones. Thus, our testing strategy confronts directly an $I(d)$ series with an $I(0)$ series subject to occasional regime shifts. In parallel with Perron (1989) who uses suitably modified Dickey-Fuller (DF) tests for the $I(1)$ vs. $I(0)$ case in the presence of regime shifts, our starting point is the DF´s approach generalization proposed by Dolado, Gonzalo and Mayoral (DGM, 2002) to test $I(1)$ vs. $I(d)$, $0 \leq d < 1$, adjusted now to test $I(d)$ vs. $I(0)$. In DGM (2002) it was shown that if $d_1 < d_0$, where d_0 and d_1 denote the orders of integration of the series under the null and the alternative hypothesis, respectively, then an unbalanced OLS regression of the form $\Delta^{d_0} y_t = \phi \, \Delta^{d_1} y_{t-1} + \varepsilon_t$, or that regression augmented with a suitable number of lagged values of $\Delta^{d_1} y_t$ when the error term in the DGP is autocorrelated, yields a consistent test of $H_0 : d = d_0$ based on the t-ratio of $\widehat{\phi}_{ols}$. In the spirit of DF´s popular methodology such a test is denoted as a Fractional Dickey-Fuller (FDF) test. To operationalise the FDF test for unit roots , the regressor $\Delta^d y_{t-1}$ is constructed by applying the truncated binomial expansion of the filter $(1-L)^d$ to y_{t-1}, so that $\Delta^d y_t = \sum_0^{t-1} \pi_i(d) \; y_{t-i}$ where $\pi_i(d)$ is the i-th coefficient in that expansion. The degree of integration under the alternative hypothesis (d_1) can be taken to be known or, alternatively estimated with a $T^{\frac{1}{2}}$ -consistent estimator. Empirical applications of such a testing procedure can be found in DGM (2003a) and a generalization of the FDF test in the $I(1)$ vs. $I(d)$ case allowing for deterministic components (drift/

linear trend) under the maintained hypothesis is in DGM (2003b).

Following the previous developments, we propose in this paper a test of $I(d), 0 < d \leq 1$, vs. $I(0)$ *cum* structural breaks, namely $d_0 = d$ and $d_1 = 0$, along the lines of the well-known procedures proposed by Perron (1989) when the date of the break is taken to be *a priori* known, and the extensions of Banerjee et al. (1992) and Zivot and Andrews (1992) when it is assumed to be unknown. The test considered in this paper is will only be derived for a single break, in order to highlight its basic principles. However, extensions to more than one break, along the lines of Bai (1999) and Bai and Perron (1998), should not be too difficult to devise once the simple case is worked out. We leave this issue for further research. As in Perron (1989) the following shifts will be considered: a crash shift, a changing growth one, and a combination of both.

In order to avoid confusion with the FDF for unit roots, the test presented here for $I(d)$ vs. $I(0)$ processes will be henceforth denoted as the *new* FDF test (NFDF), which this time is based on the t-ratio of $\widehat{\phi}_{ols}$ in an OLS regression of the form $\Delta^{d_0} y_t = \Pi(L)A_B(t) + \phi\ y_{t-1} + \varepsilon_t$, where $\Pi(L) = \Delta^d - \phi L$ and $A_B(t)$ captures the different structural breaks defined above. Hence, as in the cases of the FDF test, the NFDF test is based on the principle of unbalanced regressions which underlies the popular DF approach.

The advantages of the NFDF test, in line with those of the FDF test, rely on its simplicity (time-domain instead of frequency- domain) and on its good performance in finite samples both in terms of size and power. Specifically, it has several advantages over some other statistical procesures available in the literature to address a similar issue. For example, some authors like Choi and Zivot (2002) estimate d from the residuals of an OLS projection the original series on a set a potential structural breaks whose unknown dates are determined by means of the sequential Bai and Perron ´s (1998) procedure. The problem with this approach is that the limiting distribution of the estimate of d obtained from the residuals is different from that obtained by GPH (1983) and is most likely not to be invariant to the values of the deterministic components and the choices of the breaking dates. By contrast, this difficulty does not arise if one uses Robinson´s (1994) LM test as Gil-Alaña and Robinson (1997) do since working under the null implies that the value d_0 is known when tested against

$d_0+\theta$. Thus, $\Delta^{d_0} y_t$ can be regressed on $\Delta^{d_0} A(t)$ and a test with $N(0,1)$ limiting distribution be performed on those residuals. The problem, however, is that, being an LM test, there is no simple alternative making it impossible to reject an $I(d)$ process in favour of an $I(0)$ plus structural breaks. More recently, Mayoral (2004) has proposed a LR test of I(d) vs. I(0) subject to potential structural breaks which is an UMPI test under a sequence of local alternatives. Nonetheless, the problem with Mayoral´s test is that deviations from gaussianity in the innovations are bound to affect its good power properties. Hence, the NFDF test, despite not being UMPI, presents the advantage of not requiring a correct specification of a parametric model and other distributional assumptions, besides being computationally simple.

The rest of the paper is organized as follows. Section 2 derives the properties of the NFDF test in the presence of deterministic components like a constant or a linear trend and discusses the effects of ignoring structural breaks in means or slopes. Given that power can be severely affected under those circumstances, a NFDF test of I(d) vs. I(0) with a single structural break at a known or an unknown date is derived in Section 3 where both its limiting and finite-sample properties are discussed at length. Section 4 contains a brief discussion of how to modify the test to cater for autocorrelated disturbances, in the spirit of the ADF and AFDF test (where "A" stands for augmented versions of the test-statistics) and conjectures on how to generalize the testing strategy to multiple breaks rather than a single one. Section 5 contains an empirical application using long U.S GNP and GNP per capita series for which there has been quite a lot of controversy in the literature about the stochastic or deterministic nature of their trending components. Finally, Section 6 concludes. An Appendix (to be completed) gathers the proofs of theorems and lemmae.

THE NFDF TEST FOR LONG MEMORY VERSUS TREND STATIONARITY

Preliminaries

Before considering the main goal of the paper, it is convenient to analyze the problem of testing for fractional integration, i.e. $I(d)$ with $0 < d < 1$, against trend stationarity, i.e. $d = 0$, within the NFDF framework. The motivation for doing this is twofold. First, taking an $I(d)$ process as a generalization of the unit root parameterization, the question of whether the trend is better represented as a stochastic or a deterministic component arises on the same grounds as in the $I(1)$ case. And, secondly, the analysis in this subsection will serve as the basis for the general case where non-stationarity can arise due to the presence of structural breaks.

Under the alternative hypothesis, H_1,we consider processes with an unknown mean μ or a linear trend $(\mu + \beta t)$

$$y_t = \mu + \frac{\varepsilon_t 1_{(t>0)}}{\Delta^{d_0} - \phi L}, \tag{7}$$

$$y_t = \mu + \beta t + \frac{\varepsilon_t 1_{(t>0)}}{\Delta^{d_0} - \phi L} \tag{8}$$

where, ε_t is assumed to be $i.i.d.(0,\, \sigma_\varepsilon^2\ )$ and $d_0 \in (0, 1]$. Hence, under H_1,

$$\Delta^{d_0} y_t = \alpha + \Delta^{d_0}\delta + \phi y_{t-1} + \varepsilon_t \tag{9}$$

$$\Delta^{d_0} y_t = \alpha + \Delta^{d_0}\delta + \gamma t + \varphi\Delta^{d_0 - 1} + \phi y_{t-1} + \varepsilon_t \tag{10}$$

where $\alpha = -\phi\mu$, $\delta = \mu$, $\gamma = -\phi\beta$ and $\varphi = \beta$. For simplicity, hereafter we will write $\varepsilon_t 1_{(t>0)} = \varepsilon_t$. Under H_0, when $\phi = 0$, $\Delta^{d_0} y_t = \mu\Delta^{d_0} + \varepsilon_t$ in (9) and $\Delta^{d_0} y_t = \mu\Delta^{d_0} + \beta\Delta^{d_0 - 1} + \varepsilon_t$ in (11 (10).[1] Thus $E(\Delta^d y_t) = \Delta^d \mu$ and $E(\Delta^d y_t) = \Delta^d(\mu + \beta t)$, respectively. Note that $\mu\Delta^d = \mu \sum_{i=0}^{t-1} \pi_i(d)$ and $\beta\Delta^{d-1} = \mu \sum_{i=0}^{t-1} \pi_i(d-1)$ where the sequence $\{\pi_i(\xi)_{i=0}^{\infty}\}$ comes from the expansion of $(1 - L)^\xi$ in powers of L and the coefficients are defined as $\pi_i(\xi) = \Gamma(i - \xi)/[\Gamma(-\xi)\Gamma\,(i + 1)]$. In the sequel, we use the notation $\tau_t(\xi) = \sum_{i=1}^{t-1} \pi_i(\xi)$. Also note that $\tau_t(d)$ for $d < 0$ induces a deterministic

[1] Note that $\Delta^d t = \Delta^d \Delta^{-1} = \Delta^{d-1}$, after suitable truncation.

trend which is less steep than a linear trend and coincides with it when $d = -1$ since $\tau_t(-1) = \sum_{i=0}^{t-1} \pi_i(-1) = t$. As shown in DGM (2003, Figure 1) for values of $d < 0$, $\tau(.)$ is a concave function, being less steep the smaller d is. Under H_1, the polynomial $\Pi(z) = \left((1-z)^d - \phi z\right)$ has absolutely summable coefficients and verifies $\Pi(0) = 1$ and $\Pi(1) = -\phi \neq 0$. All the roots of the polynomial are outside the unit circle if $-2^d < \phi < 0$. As in the DF framework, this condition excludes explosive processes. By contrast, under H_1, with the previous restriction on ϕ, y_t is $I(0)$ and admits the representation

$$\begin{aligned} y_t &= \mu + u_t, \text{ or } y_t = \mu + \beta t + u_t, \\ u_t &= \Lambda(L)\varepsilon_t, \ \Lambda(L) = \Pi(L)^{-1} \end{aligned}$$

Computing the trends $\tau_t(\xi)$, $\xi = d$ or $d-1$ in (9) or (10) does not entail any difficulty since only depends on d_0, which is a known parameter since it is the integration order under H_0. This case, also considering the presence of a linear trend under the null of $d_0 = 1$ against the alternative of $d_1 = d$, $0 < d < 1$, has been analyzed at length in DGM (2003) where it is shown that the FDF test is (numerically) invariant to the value of μ in the DGP. Note that an alternative computational strategy for the NFDF test arises from simple use of the Frisch-Waugh Theorem implying that the test statistics can also be computed in a two-step procedure as follows. First, regress $\Delta^d y_t$ and y_{t-1} on a constant term and $\tau_t(d_0)$ in (9) and on those terms plus $\tau_t(d_0 - 1)$ in (10), in order to obtain the residuals denoted $\Delta^d \widehat{u}_t$ and $\widehat{u}_{t-1}$, respectively. Secondly, compute the t-ratio of the estimated coefficient in the regression of $\Delta^d \widehat{u}_t$ on $\widehat{u}_{t-1}$.

Next, we derive the corresponding result for $H_0 : d_0 = d, 0 < d \leq 1$,vs. $H_1 : d_1 = 0$. The following theorem summarizes the main result.

Theorem 1 *Under the null hypothesis that y_t is an $I(d)$ process as defined in* (7) *or* (8) *with $\phi = 0$, the OLS coefficient associated to ϕ in regression model* (9), $\hat{\phi}_{ols}^{\mu}$, *or* (10), $\hat{\phi}_{ols}^{\tau}$, *respectively is a consistent estimator of $\phi = 0$ and converges at a rate T^d if $0.5 < d \leq 1$ and at the usual rate $T^{1/2}$ when $0 < d < 0.5$. The asymptotic distribution* *of the associated* t − *statistic,* $t_{\hat{\phi}^i_{ols}}$, $i = \{\mu, \tau\}$ *is given by*

$$t_{\hat{\phi}^i_{ols}} \xrightarrow{w} \frac{\int_0^1 B_d^i(r)\, dB(r)}{\left(\int_0^1 \left(B_d^i(r)\right)^2 d(r)\right)^{1/2}}, \ \textit{if}\ 0.5 < d \leq 1,$$

and

$$t_{\hat{\phi}^i_{ols}} \xrightarrow{w} N(0,1), \ \textit{if}\ 0 < d < 0.5.$$

where $B_d^i(r)$ *,*$i = \{\mu, \tau\}$ *is a "detrended" fBM, appropiately defined in the Appendix.*

The intuition for this result is similar to the one offered by DGM (2002) for the $I(1)$ vs. $I(d)$ case with $0 < d < 1$. When d_0 (= 1 in that case) and d_1 are close, then the asymptotic distribution is asymptotically normal whereas it is a functional of fBM when both parameters are far apart. Hence, since in our case $d_0 = d$, $0 < d \leq 1$, and $d_1 = 0$, asymptotic normality arises when $0 < d \leq 0.5$. Also note that $d_0 = 1$ renders the standard DF limiting distribution

The finite-sample properties of the NFDF test for this particular case are presented in Tables 2a, b and 3. Three sample sizes are considered, $T = 100, 400$ and $1,000$, and the number of replications is $10,000$. Table 2a gathers the corresponding critical values for the case where the DGP is a pure $I(d)$ without drift (since the test is invariant to the value of μ), i.e, $\Delta^d y_t = \varepsilon_t$ with $\varepsilon_t \sim N(0,1)$, when (9), or the two-step procedure, is considered as the regression model. Table 2b, in turn, offers the corresponding critical values when (10) is taken to be the regression model. As can be observed, the empirical critical values are close to those of a standardized $N(0,1)$ (whose critical values are -1.28, -1.64 and -2.33, respectively, for the three significance levels reported below) when $0 < d \leq 0.5$, particularly for $T \geq 400$. However, for $d > 0.5$ the critical values start to differ drastically from those of a normal distribution, increasing as d gets larger. As for power, Table 3 reports the rejection rates at the 5% level of the NFDF in (11) when the data are generated according to the DGP: $y_t = \alpha + \beta t + \varepsilon_t$ with $\alpha = 0.1, \beta = 0.5$. Except for low values of d and $T = 100$, where power still reaches 55%, the test is very powerful in all the remaining cases.

TABLE 2a

Critical Values

Percentage points of the asymptotic distribution of $t^{\mu}_{\phi_{ols}}$

	$T = 100$			$T = 400$			$T = 1000$		
d_0/ S.L.	90%	95%	99%	90%	95%	99%	90%	95%	99%
0.1	-1.547	-1.894	-1.891	-1.326	-1.698	-2.397	-1.308	-1.668	-2.352
0.2	-1.567	-1.9497	-1.983	-1.367	-1.814	-2.420	-1.350	-1.727	-2.390
0.3	-1.640	-2.003	-1.991	-1.439	-1.832	-2.520	-1.407	-1.784	-2.432
0.4	-1.683	-2.132	-2.132	-1.573	-1.862	-2.578	-1.432	-1.805	-2.512
0.5	-1.712	-2.137	-2.173	-1.670	-1.921	-2.607	-1.573	-1.8758	-2.586
0.6	-2.641	-2.201	-2.546	-2.075	-2.407	-3.099	-2.028	-2.382	-3.004
0.7	-2.769	-2.364	-2.720	-2.252	-2.577	-3.208	-2.217	-2.540	-3.180
0.8	-2.804	-2.50	-2.837	-2.394	-2.689	-3.320	-2.397	-2.710	-3.326
0.9	-2.812	-2.599	-2.929	-2.551	-2.857	-3.497	-2.485	-2.784	-3.351

TABLE 2b

Critical Values

Percentage points of the asymptotic distribution of $t^{\tau}_{\phi_{ols}}$

	$T = 100$			$T = 400$			$T = 1000$		
d_0/ S.L	90%	95%	99%	90%	95%	99%	90%	95%	99%
0.1	-1.567	-1.913	-2.788	-1.368	-1.733	-2.442	-1.2289	-1.638	-2.383
0.2	-1.616	-1.957	-2.815	-1.719	-1.797	-2.470	-1.589	-1.648	-2.404
0.3	-2.049	-2.096	-2.845	-1.853	-1.801	-2.528	-1.767	-1.677	-2.429
0.4	-2.138	-2.166	-2.897	-2.051	-1.847	-2.678	-1.795	-1.747	-2.487
0.5	-2.465	-2.531	-3.521	-2.284	-2.631	-2.913	-2.231	-2.163	-2.641
0.6	-2.694	-3.021	-3.658	-2.560	-2.894	-3.560	-2.488	-2.800	-3.407
0.7	-2.935	-3.257	-3.895	-2.824	-3.131	-3.764	-2.773	-3.086	-3.750
0.8	-3.159	-3.480	-4.087	-3.067	-3.367	-3.921	-3.011	-3.320	-3.930
0.9	-3.366	-3.700	-4.390	-3.291	-3.590	-4.143	-3.250	-3.553	-4.094

TABLE 3

POWER (NOMINAL SIZE: 5%)

R.M.: $\Delta^{d_0} y_t = \alpha + \delta\tau_t(d) + \gamma t + \varphi\tau_t(d-1) + \phi y_{t-1} + \varepsilon_t$

	DGP: $y_t = \alpha + \beta t + \varepsilon_t$;		
d_0/ sig. lev.	$T = 100$	$T = 400$	$T = 400$
0.2	54.9%	98.9%	100%
0.4	98.4%	100%	100%
0.5	99.6%	100%	100%
0.7	100%	100%	100%
0.9	100%	100%	100%
1.0	100%	100%	100%

The effects of structural breaks on the NFDF test

To assess the effects of the presence of a drift in the level of the series or a shift in the slope of the trend on the NFDF tests for $I(d)$ vs. $I(0)$, let us consider first the consequences of performing the NFDF test in (10) when there is a break in the mean, so that y_t is generated by

$$DGP\ 1: \ y_t = \mu + \zeta DU_t(\lambda) + \varepsilon_t, \tag{11}$$

where $\varepsilon_t \sim iid(0, \sigma_\varepsilon^2)$ and $DU_t(\lambda) = \mathbf{1}_{(T_B+1 \leq t \leq T)}$. According to (9), where this kind of break is ignored, the NFDF test is based on the following regression model that we repeat for convenience

$$\Delta^d y_t = \alpha + \delta\tau_t(d) + \phi y_{t-1} + \varepsilon_t. \tag{12}$$

Then, the following theorem holds

Theorem 2 *If y_t is given by DGP 1 and model* (12) *is used to estimate ϕ,then for $0 < \lambda < 1$ it follows that*

$$\widehat{\phi}_{ols} \xrightarrow{p} -\frac{d\sigma_\varepsilon^2[C_1^2(d) - C_2(d)]}{D(d, \sigma_\varepsilon^2)}, \ \textit{if}\ 0 < d < 0.5,$$

$$\widehat{\phi}_{ols} \xrightarrow{p} -\frac{d\sigma_\varepsilon^2}{[\zeta^2\lambda(1-\lambda)+\sigma_\varepsilon^2]}, \text{ if } 0.5 < d \leq 1,$$

and

$$t_{\widehat{\phi}_{ols}} \xrightarrow{p} -\infty, \text{ if } 0 < d \leq 1$$

with ,

$$D(d,\sigma_\varepsilon^2) = C_1^2(d)[\zeta\lambda(1-\lambda)+\sigma_\varepsilon^2] - C_2\{\zeta^2[1-\lambda+(1-\lambda^{1-d})(2\lambda-\lambda^{1-d}-1)]+\sigma_\varepsilon^2\}$$

and

$$C_1(d) = (d-1)\Gamma(-d) \text{ , } C_2(d) = \Gamma^2(-d)(2d-1)$$

Thus, Theorem 2 shows that, under the crash hypothesis the limit depends on the size of relative shift in the mean, ζ. Note that if $\lambda = 0$ or1, i.e., there is no break, $\widehat{\phi}_{ols} \xrightarrow{p} -d$, which makes sense under DGP 1, since being $y_t \sim I(0)$, the covariance between $\Delta^d y_t$ and y_{t-1} is $\pi_1(d) = -d$. Further, for $d = 1$, it yields expression (a) in Theorem 1 of Perron (1989). Hence, $\widehat{\phi}_{ols}$ converges to a finite negative number, and hence $T^{1/2}\widehat{\phi}_{ols}$, for $d \in (0, 0.5)$ and $T^d\widehat{\phi}_{ols}$, for $d \in (0.5, 1)$ and the corresponding t-ratios in each case diverge to $-\infty$. Thus the NFDF test would eventually reject the null hypothesis of $d = d_0$, $0 < d_0 < 1$, when it happens to be false. Notice, however, that the power of the NFDF will be decreasing with the distance between the null and the alternative, namely as d gets closer to its true zero value.

Next consider the case where there is (continuous) break in the slope of the linear trend, such that y_t is generated by

$$DGP\ 2: \ y_t = \mu_0 + \beta_0 t + \psi_0 DT_t^*(\lambda) + \varepsilon_t, \tag{13}$$

where $DT_t^*(\lambda) = (t - T_B)\mathbf{1}_{(T_B+1 \leq t \leq T)}$, whilst the NFDF test is implemented according to model (10), which does not account for the breaking trend, namely

$$\Delta^d y_t = \alpha + \gamma t + \delta\tau_t(d) + \varphi\tau_t(d-1) + \phi y_{t-1} + \varepsilon_t, \tag{14}$$

Theorem 3 *If* y_t *is given by DGP 2 and model* (12) *is used to estimate* ϕ*, then for* $0 < \lambda < 1$ *it follows that*

$$t_{\widehat{\phi}_{ols}} \xrightarrow{p} +\infty \ \ if \, 0 < d < 0.5$$

and

$$t_{\widehat{\phi}_{ols}} \xrightarrow{p} 0, \ \ if \, 0.5 < d \leq 1$$

Hence, when ignoring a breaking trend, the FDF is unambiguously inconsistent. The intuition behind this result, which generalizes Theorem 1 in Perron (1989, part b), is that $\widehat{\phi}_{ols}$ is $O_p(T^{-d})$ with a positive limiting constant term for $d \in (0, 1]$ and that s.d $(\widehat{\phi}_{ols})$ is $T^{-1/2}$, implying that the t-ratio is $O_p(T^{1/2-d})$. Therefore, it will tend to zero for $1 \geq d > 0.5$ and to $+\infty$ for $0 < d < 0.5$.

In sum, the NFDF test without consideration of structural breaks is not consistent against breaking trends and, though consistent against a beak in the mean, its power is likely to be reduced if such a break is large. Hence, there is a need for alternative forms of the NFDF test that could distinguish an $I(d)$ process from a process being $I(0)$ around deterministic terms subject to structural breaks.

THE NFDF TEST OF I(D) VS. I(0) WITH STRUCTURAL BREAKS

In line with the above considerations, we now proceed to derive the NFDF invariant test for $I(d)$ vs. $I(0)$ allowing for structural breaks under H_1. To do so, it seems convenient to consider the most common definition of (possibly) non-stationary $I(d)$ processes used, among others, by Beran (1995), Velasco and Robinson (2000) and Mayoral (2003), which is as follows. Consider the ARFIMA(p, d_0, q) process for y_t, $t = 1, 2, ..., T$ which can be written as

$$\Phi_0 (L) \Delta^{\varphi_0} (\Delta^{m_0} y_t - \mu_0) = \Theta_0 (L) \varepsilon_t. \qquad (15)$$

where the memory parameter, d_0, belongs to the closed interval $[\nabla_1, \nabla_2]$, with $-0.5 < \nabla_1 < \nabla_2 < \infty$. Notice that d_0 can be interpreted as the sum of an integer and a fractional part such that $d_0 = m_0 + \varphi_0$. On the one hand, the integer $m_0 = [d_0 + 1/2]$, where [.] denotes integer part, is the number of times that y_t must be differenced to achieve stationarity (therefore $m_0 \geq 0$). On the other, the parameter φ_0, the fractional part, lies in the interval $(-0.5, 0.5)$, in such a way that, for a given d_0, $\varphi_0 = d_0 - [d_0 + 1/2]$. Consequently, once the process y_t is differenced m_0 times, the differenced process is a stationary fractionally integrated process with integration order φ_0. For $m_0 = 0$, μ_0 is the expected value of the stationary process y_t and for $m_0 \geq 1$, $\mu_0 \neq 0$ implies a deterministic polynomial trend. In particular $m_0 = 1$ implies a linear time trend (i.e $\mu_0 t$).

To account to structural breaks, we consider the following maintained hypothesis

$$y_t = A_B(t) + \frac{a_t 1(t > 0)}{\Delta^d - \phi L}, \tag{16}$$

where $A_B(t)$ is a linear deterministic trend function that may contain breaks at unknown dates (in principle, just a single break at date T_B would be considered) and a_t is a stationary $I(0)$ process. From the above arguments, it can be easily shown that if $\phi = 0$, then y_t is an $I(d)$ process with $0 < d < 1$, while, if $\phi < 1$, then the resulting process would be $I(0)$ subject to structural breaks. Note that, under the null, the model can be written as $\Delta^d[y_t - A_B(t)] = a_t$ or $\Delta^{d-1}[\Delta y_t - \mu_0] = a_t$ which corresponds to the ARFIMA family defined in (16) with $\mu_0 = \Delta A_B(t)$ and $\Phi_0(L) = \Theta_0(L) = 1$

In common with Perron (1989) and Zivot and Andrews (1992), three definitions of $A_B(t)$ are considered

$$\textit{Case A}: \ A_B^A(t) = \mu_0 + (\mu_1 - \mu_0) DU_t(\lambda) \tag{17}$$

$$\textit{Case B}: \ A_B^B(t) = \mu_0 + \beta_0 t + (\beta_1 - \beta_0) DT_t^*(\lambda) \tag{18}$$

$$\textit{Case C}: \ A_B^C(t) = \mu_0 + \beta_0 t + (\mu_1 - \mu_0) DU_t(\lambda) + (\beta_1 - \beta_0) DT_t(\lambda) \tag{19}$$

Case A corresponds to the *crash* hypothesis, case B to the *changing growth* hypothesis and case C to a combination of both. The dummy variables $DU_t(\lambda)$ and $DT_t^*(\lambda)$are defined as before, and $DT_t(\lambda) = tT_B)\mathbf{1}_{(T_B+1 \leq t \leq T)}$, where $\lambda = T_B/T$. For the time being, let us assume that the break date T_B is taken to be known a priori.

Then, the NFDF test of $I(d)$ vs. $I(0)$ in the presence of structural breaks is based on the t- ratio on the coefficient ϕ in the regression model

$$\Delta^d y_t = \Delta^d A_B^i(t) - \phi A_B^i(t-1) + \phi y_{t-1} + a_t, i = A, B, C. \tag{20}$$

As before, the NFDF test can be computed by first obtaining the residuals from regressions of $\Delta^d y_t$ and y_{t-1} on $\Delta^d A_B^i(t)$ and $A_B^i(t-1)$, denoted as $\Delta^d \widehat{u}_t$ and $\widehat{u}_{t-1}$ respectively, and next computing the t-ratio of $\widehat{\phi}_{ols}$in the regression of $\Delta^d \widehat{u}_t$ on $\widehat{u}_{t-1}$. It is easy to check that under $H_1 : \phi < 0$, y_t is $I(0)$ subject to the regime shifts defined by $A_B^i(t)$ whilst under $H_0 : \phi = 0$, it is $I(d)$ such that $E[\Delta^d(y_t - A_B^i(t))] = 0$. Moreover, the NFDF test is invariant to the values of μ_0, μ_1, β_0 and $\beta_1$under H_0. Using similar arguments to those in Theorem 1, the following theory holds.

Theorem 4 *Let y_t be a process generated as in* (16) *with possibly $\mu_0 = \mu_1 = \beta_0 = \beta_1 = 0$. Then, under the null hypothesis of $\phi = 0$, the OLS estimator associated to ϕ in regression model* (20) *is consistent. The asymptotic distribution of the associated $t - ratio$ is given by*

$$t^i_{\widehat{\phi}(\lambda)} \xrightarrow{w} \frac{\int_0^1 B_d^{*i}(\lambda, r)\, dB(r)}{\left(\int_0^1 B_d^{*i}(\lambda, r)^2\, d(r)\right)^{1/2}} \quad if\, d \in (0.5, 1]$$

$$t^i_{\widehat{\phi}(\lambda)} \xrightarrow{w} B^*(1) \equiv N(0,1) \quad if\, d \in (0, 0.5)$$

where $B_d^{*i}(.)$ is the projection residual from the corresponding continuous time regression associated to models $i = \{A, B, C\}$.

Although previously we assumed that the date of the break was known, it seems more plausible, along the lines of the arguments exposed in the Introduction, to assume that T_B is unknown a priori. Hence, following the approach in Banerjee et al. (1992) and Zivot and Andrews (1992), an extension of the previous procedure is to estimate this breakpoint in such a way that gives the highest weight to the alternative $I(0)$ alternative. The estimation scheme will therefore consist in choosing the breakpoint that gives the least favorable result for the null hypothesis of $I(d)$ using the NFDF test described above in each of the three cases, $i = A, B, C$. Thus, the $t - statistic$ on $\hat{\phi}^{i}_{ols}$, $t_{\widehat{\phi}(\lambda)}$, is computed for several values of $\lambda \in \Lambda = (2/T, (T-1)/T)$ and then the infimum value would be chosen to run the test. The test would be then to reject the null hypothesis when

$$\inf_{\lambda \in \Lambda} t_{\widehat{\phi}(\lambda)} > k^{i}_{\inf,\alpha},$$

where $k^{i}_{\inf,\alpha}$ is a critical value to be provided below. Under theese conditions, the following theory holds.

Theorem 5 *Let y_t be a process generated as in* (16). *Then, under the null hypothesis of $\phi = 0$, the OLS estimator associated to ϕ in regression model* (20) *is consistent. Let Λ be a closed subset of* $(0,1)$. *Then, the asymptotic distribution of the associated $t - statistic$ associated to ϕ is given by,*

$$\inf_{\lambda \in \Lambda} t^{i}_{\widehat{\phi}(\lambda)} \xrightarrow{w} \inf_{\lambda \in \Lambda} \frac{\int_0^1 B^{*i}_d(\lambda, r)\, dB(r)}{\left(\int_0^1 B^{*i}_d(\lambda, r)^2\, d(r)\right)^{1/2}} \quad \text{if } d \in (0.5, 1],$$

$$\inf_{\lambda \in \Lambda} t^{i}_{\widehat{\phi}(\lambda)} \xrightarrow{w} \inf_{\lambda \in \Lambda} B^{*}(1) \quad \text{if } d \in (0, 0.5).$$

To generate critical values of the *inf* NFDF t-ratio test, a pure $I(d_0)$ process with $\varepsilon_t \sim n.i.d.(0,1)$ has simulated $10,000$ times, whereas the three regression models (A, B and C) have been considered for samples of size $T = 100, 400, 1000$. Tables 4a, b, c, report the corresponding critical values. Note that they are larger than the critical values of the NFDF test reported in Tables 2a, b when considering the left tail. Hence, one should expect a loss in power.

TABLE 4a

Critical Values

Percentage points of the asymptotic distribution of $t^A_{\hat{\phi}(\lambda)}$

	T=100			T=400			T=1000		
d_0 / S.L	90%	95%	99%	90%	95%	99%	90%	95%	99%
0.1	-2.056	-2.427	-3.075	-1.739	-2.100	-2.807	-1.599	-1.975	-2.698
0.2	-2.271	-2.630	-3.349	-1.936	-2.297	-2.955	-1.738	-2.115	-2.827
0.3	-2.443	-2.784	-3.499	-2.119	-2.459	-3.085	-1.989	-2.334	-2.992
0.4	-2.668	-2.989	-3.645	-2.387	-2.726	-3.450	-2.236	-2.593	-3.188
0.5	-2.964	-3.315	-3.965	-2.688	-3.006	-3.653	-2.582	-2.895	-3.551
0.6	-3.236	-3.532	-4.161	-2.999	-3.342	-4.009	-2.918	-2.918	-3.219
0.7	-3.519	-3.847	-4.484	-3.331	-3.634	-4.221	-3.241	-3.241	-3.538
0.8	-3.761	-4.069	-4.692	-3.602	-3.875	-4.437	-3.561	-3.561	-3.861
0.9	-3.978	-4.266	-4.852	-3.870	-4.137	-4.613	-3.784	-3.784	-4.043

TABLE 4b

Critical Values

Percentage points of the asymptotic distribution of $t^B_{\hat{\phi}(\lambda)}$

	T=100			T=400			T=1000		
d_0/ S.L.	90%	95%	99%	90%	95%	99%	90%	95%	99%
0.1	-2.448	-2.797	-3.504	-1.950	-2.333	-3.016	-1.758	-2.123	-2.867
0.2	-2.681	-3.032	-3.713	-2.2001	-2.567	-3.195	-1.9474	-2.303	-2.984
0.3	-2.893	-3.244	-3.950	-2.432	-2.771	-3.405	-2.238	-2.577	-3.241
0.4	-3.178	-3.520	-4.176	-2.754	-3.113	-3.790	-2.554	-2.883	-3.505
0.5	-3.522	-3.872	-4.514	-3.1158	-3.452	-4.076	-2.946	3.274	-3.860
0.6	-3.848	-4.156	-4.801	-3.519	-3.855	-4.530	-3.379	-3.682	-4.254
0.7	-4.209	-4.532	-5.196	-3.936	-4.239	-4.788	-3.815	-4.105	-4.693
0.8	-4.540	-4.858	-5.494	-4.298	-4.577	-5.069	-4.239	-4.525	-5.090
0.9	-4.892	-5.198	-5.808	-4.627	-4.901	-5.406	-4.579	-4.859	-5.410

TABLE 4c

CRITICAL VALUES

PERCENTAGE POINTS OF THE ASYMPTOTIC DISTRIBUTION OF $t^{C}_{\widehat{\phi}(\lambda)}$

	T=100			T=400			T=1000		
d_0 /S.L..	90%	95%	99%	90%	95%	99%	90%	95%	99%
0.1	-2.449	-2.800	-3.508	-1.951	-2.333	-3.016	-1.758	-2.129	-2.867
0.2	-2.683	-3.032	-3.7070	-2.201	-2.568	-3.200	-1.946	-2.303	-2.984
0.3	-2.895	-3.250	-3.962	-2.429	-2.770	-3.406	-2.238	-2.577	-3.241
0.4	-3.179	-3.524	-4.176	-2.755	-3.112	-3.788	-2.554	-2.881	-3.506
0.5	-3.525	-3.873	-4.522	-3.115	-3.453	-4.075	-2.946	3.273	-3.861
0.6	-3.848	-4.151	-4.797	-3.519	-3.856	-4.529	-3.379	-3.682	-4.253
0.7	-4.209	-4.533	-5.196	-3.938	-4.239	-4.789	-3.815	-4.106	-4.693
0.8	-4.540	-4.8580	-5.494	-4.298	-4.577	-5.069	-4.238	-4.525	-5.090
0.9	-4.892	-5.197	-5.809	-4.628	-4.901	-5.406	-4.579	-4.859	-5.410

In order to examine the power of the test, we have generated 1000 replications of DGP 2 with sample sizes T=100, 400, where and $\lambda = 0.5$, that is a changing growth model with a break in the middle of the sample ($\lambda = 0.5$). Both regression models B and C have been estimated. Rejection rates are reported in Table 5. An important characteristic to check is whether power increases with the distance between the alternative and the null hypotheses. Interestingly, this is not the case here, since power is non-monotonic, first increasing and then decreasing, and attains a maximum around values of d close to 0.6. From a technical viewpoint , the reason behind this result is that the values of the statistics are, in general, monotonically decreasing in d but the critical values decrease faster and therefore, power deteriorates. From an intuitive viewpoint, what happens is that the trend functions of an $I(1)$ process with drift and a $I(0)$ process with a linear trend are much more similar than those of the latter process and an $I(d)$ process with a value of d around 0.6. Hence with large values d power decreases. This has a very interesting implication for empirical work, namely that a test with null of $I(d)$ with an intermediate value d in $(0, 1]$, as in the NFDF test, has much more power than a test based on the null of an $I(1)$.

TABLE 5

Power NFDF, FI(d) Structural breaks

DGP: $y_t = \mu_0 + \beta_0 t + \psi_0 DT_t^*(\lambda) + \varepsilon_t$; $\mu_0 = 1$; $\beta_0 = 0.5; \sigma_\varepsilon = 1$

	$\psi_0 = 0.1$				$\psi_0 = 0.2$			
d_0 / S.L	Model B		Model C		Model B		Model C	
T	T=100	T=400	T=100	T=400	T=100	T=400	T=100	T=400
0.1	6.0%	54.8%	6.0%	56.0%	13.2%	100%	12.4%	99.4%
0.3	9.4%	98.2%	9.2%	98.0%	52.0%	100%	50.4%	100%
0.6	4.8%	98.8%	4.8%	98.0%	56.8%	100%	56.6%	100%
0.7	5.0%	71.2%	5.0%	71.2%	24.2%	100%	24.2%	100%
0.9	5.0%	7.9%	2.0%	4.8%	8.2%	100%	8.0%	100%

AUGMENTED NFDF TEST AND MULTIPLE BREAKS

The limiting distributions derived above are valid for the case where the innovations are *i.i.d.* and no extra terms are added in the regression equations. If some autocorrelation structure or heterogeneous distributions are allowed in the innovation process, then the asymptotic distributions will depend on some nuisance parameters. To solve the nuisance-parameter dependency, two approaches have been tipically employed in the literature. One is the non-parametric approach proposed by Phillips and Perron (1987) which is based on finding consistent estimators for the nuisance parameters. The other, which is the one we follow here, is the well-known parametric approach proposed by Dickey and Fuller (1981) which consists of adding a suitable number of lags of $\Delta^d y_t$ to the set of regressors (see DGM, 2002). As Zivot and Andrews (1992) point out, a formal proof of the limiting distributions when the assumption of *i.i.d.* disturbances is relaxed is likely to be very involved. However, along the lines of the proof for AFDF test in Theorem 7 of DGM (2002), we conjecture that if the DGP is $\Delta^d y_t = u_t 1_{(t>0)}$ and u_t follows an invertible and stationary ARMA (p,q) process $\alpha_p(L)u_t = \beta_q(L)\varepsilon_t$ with $E\left|\varepsilon_t\right|^{4+\delta} < \infty$ for some $\delta > 0$, then the inf NFDF test based on the t-ratio of $\widehat{\phi}_{ols}$ in regression models like (18)-(20) augmented with k lags of $\Delta^d y_t$

will have the same limiting distributions as in Theorem 5 above and will be consistent when $T \to \infty$ and $k \to \infty$, as long as $k^3/T \to 0$. Hence, the augmented NFDF test (denoted as ANFDF) will be based on the regression model

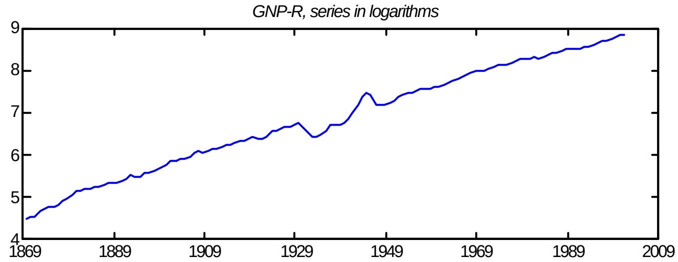
$$\Delta^d y_t = \Delta^d A_B^i(t) - \phi A_B^i(t-1) + \phi y_{t-1} + \sum_{j=1}^{k} \Delta^d y_{t-j+} + a_t, i = A, B, C. \qquad (21)$$

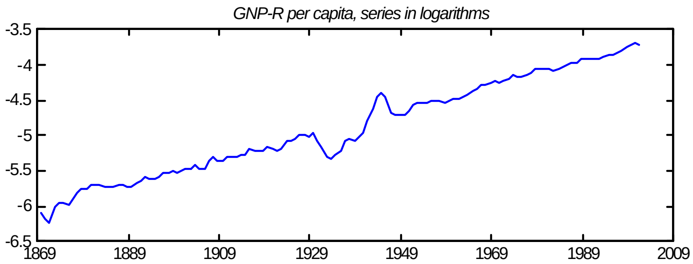
A generalization of the previous results to multiple breaks can be done along the lines of the procedure devised by Bai and Perron (1998). In their framework there m possible breaks affecting the mean and the trend slope and they suggest the following procedure to select the number of breaks. Letting *sup* $\mathrm{F}_T(l)$ be the F- statistic of no structural break ($l = 0$) vs. k breaks ($k \leq m$), they consider two statistics to test the null of no breaks gainst an unknown number of breaks given some specific bound on the maximum number of shifts considered. The first one is the double maximum statistic ($\mathrm{UD}_{\max}$) where $\mathrm{UD}_{\max} = \max_{1 \leq k \leq m} supF_T(l)$ while the second one is $\sup F_t(l + 1/l)$ which test the null of l breaks against the alternative of $l + 1$ breaks. In practice, they advise to use a sequential procedure based upon testing first for one break and if rejected for a second one, etc., using the sequence of $\sup F_t(l+1/l)$ statistics. Therefore, our proposal is to use such a procedure to determine $\lambda_{1,...,}\lambda_k$,in the $\mathrm{A}_B(t)$ terms in (18)-(20). By continuity of the sup function and tightness of the probability measures associated with $t_{\widehat{\phi}_{ols}}$, we conjecture that a similar result to that obtained in Theorem 5 will hold as well, this time with the sup of a suitable functional of fBM. Derivation of these results and computation of the corresponding critical values exceeeds the scope of this paper but is definitely in our future research agenda.

AN EMPIRICAL APPLICATION

In order to provide some empirical illustrations of how the NFDF test can be applied in practice, we consider some long series of U.S. real GNP and real GNP per capita which basically correspond to the same data set used in Diebold and Senhaji (1996) (DS henceforth) in their interesting discussion on whether GNP data is informative enough to distinguish between trend stationarity (T-ST) and first-difference stationarity (D-ST). The data are annual and range from 1869 to 2001 giving rise to a sample of 133 observations where the last 8 observations have been added to DS´s riginal sample ending in 1995; cf. Mayoral, 2004, for a detailed discussion of the construction of the series. As in DS, the series have been obtained from the two alternative sources which differ in their pre-1929 values but are identical afterwards. These correspond to the historical annual real GNP series constructed by Balke and Gordon (1989) and Romer (1989), so that the series are denoted as GNP-BG and GNP-R series, respectively. In order to convert them in per capita (PC) terms, they have been divided by the total population residing in theU.S (in thousands of people) obtained from the *Historical Statistics of the United States* (1869-1970, Table A-7) and the *Census Bureau´s Current Population Reports* (Series P-25, 1971-2001). All the series are logged.

According to DS´s analysis, there is conclusive evidence in favour of T-ST and against D-ST. To achieve this conclusion, DS follow Rudebusch (1993)´s bootstrap approach in computing the best-fitting T-ST and D-ST models for each of the four series. Then, they compute the exact finite sample distribution of the t-ratios of the lagged GNP level in an augmented Dickey-Fuller (ADF) test for a unit root when the best-fitting T-ST D-ST models are used as the DGPs. Their main finding is that the p-value of such the ADF test was very small under the D-ST model but quite large under the T-ST model, implying that the sample value of the ADF test was very unlikely under the latter model. Nonetheless, as DS acknowledge, rejecting the null does not mean that the alternative is a good characterization of the data. Indeed, Mayoral (2004) has pointed out that if the same exercise is done with the KPSS test, then the null of TS-T is also rejected in all four series. This inconclusive outcome leads Mayoral (2004) to conjecture that, since both the I(0) and I(1) null hypotheses are rejected, it may be the case that the right process is an $I(d)$, $0 < d < 1$, for which she finds favourable evidence using the FDF test of $I(1)$ vs. $I(d)$, which rejects the null , and a LR test of $I(d)$ vs.$I(0)$, which does not reject the null, in both cases with values of d in the range 0.6- 0.7. However, by observing Figures 1 and 2, where both logged GNP and logged GNP-PC are depicted (only the GNP-R series are shown since they are not too different from the GNP-BG series), one could consider as well that a reasonable conjecture is that the data are generated by a T-ST process subject to some structural breaks. Hence, this example provides a nice illustration of the usefulness of the NFDF test proposed here, since there is some mixed evidence about the data being generated either by an $I(d)$ process or by an $I(0)$ *cum* structural breaks alternative.





In Tables 6 and 7, we report the t-ratios of the NFDF test constructed according to either (20) or (21) where up to three lags of $\Delta^d y_t$ have been included as additional regressors in order to account for residual correlation. Table 6 presents the results obtained from the GNP-BG and GNP-BG-PC series whereas Table 7 presents the corresponding results for the GNP-R and GNP-R-PC series. In both instances the critical values are those reported in Tables 4b,c for T=100. Values of d in the non-stationary (albeit mean-reverting) range $(0.5, 1)$ have been considered to construct $\Delta^d y_t$ and $\Delta^d y_t A_B^i(t), i = A, B, C$. In view of the series, the most appropiate model would be either model B or C which account for the upward trending behaviour. Hence results for model A are not reported. As it can be observed, except for the case where there are no lags, in most instances, the null of I(d) is often rejected at the 5% level (significant values marked with an asterisk) in favour of a changing growth model with a breaking date located around 1939 coinciding with the beginning of

World War II.

TABLE 6

NFDF and ANFDF Tests

GNP-BG PC

Model	B				C			
Lags/d	0.6	0.7	0.8	0.9	0.6	0.7	0.8	0.9
0	1.93	0.48	-0.88	-2.16	0.84	-0.60	-2.00	-3.35
1	-5.38*	-5.34*	-5.39*	-5.51*	-6.35*	-6.40*	-6.57*	-6.81*
2	-4.55*	-4.74*	-4.93*	-5.19*	-4.94*	-5.10*	-5.34*	-5.65*
3	-3.96	-4.34	-4.51	-4.81	-4.86*	-5.03*	-5.24*	-5.50*
				GNP-BG				
0	-2.65	-0.17	-1.45	-2.65	0.07	-1.31	-2.72	-4.05
1	-4.91*	-4.34	-4.59	-4.91	-6.21*	-6.31*	-6.50*	-6.77*
2	-5.29*	-4.68*	-4.97*	-5.29*	-6.05*	-6.31*	-6.58*	-6.87*
3	-4.63*	-3.96	-4.28	-4.63	-5.07*	-5.33*	-5.62*	-5.93*

TABLE 7

NFDF and ANFDF Tests

GNP-R PC

Model	B				C			
Lags/d.	0.6	0.7	0.8	0.9	0.6	0.7	0.8	0.9
0	1.21	-0.21	-1.53	-2.73	0.29	-1.12	-2.48	-3.79
1	-4.73*	-4.84*	-5.04*	-5.32*	-5.75*	-5.92*	-6.23*	-6.67*
2	-4.35*	-4.46	-4.64	-4.90	-4.64*	-4.72*	-4.95*	-5.27*
3	-3.80	-4.04	-4.30	-4.57	-4.83*	-4.96*	-5.13*	-5.36*
				GNP-R				
0	-1.65	-0.27	-1.05	-2.35	0.47	-1.31	-2.72	-4.05
1	-4.71*	-4.84*	-5.09*	-5.26*	-7.21*	-7.31*	-7.50*	-7.77*
2	-4.39*	-4.65*	-4.97*	-5.32*	-5.65*	-5.97*	-6.28*	-6.87*
3	-3.73	-4.02	-4.38	-4.65	-5.67*	-5.83*	-6.02*	-6.33*

CONCLUSIONS

In this paper we provide a simple test of the null hypothesis of a process being $I(d), d \in (0,1)$ against the alternative of being $I(0)$ with deterministic terms subject to structural changes at known or unknown dates. The test, denotes as New Fractional Dickey-Fuller (NFDF) test is a time-domain one and performs well in finite samples, both in terms of power and size. Denoting by $A_B(t)$ the different types of structural breaks considered by Perron (1989), the NFDF test is based on the t-ratio of the coefficient on y_{t-1}in an OLS regression of $\Delta^d y_t$ on $\Delta^d A_B(t)$ and y_{t-1},plus a suitable number of lags of $\Delta^d y_t$ to cater for autocorrelated errors. Interestingly, power is maximized for intermediate values of d which when the deterministic components of the process under the null and the alternative differ the most. Hence, gains in power relative to the conventional DF tests proposed by Perron (1998), for known breaking date, and Banerjee et al (1992) and Zivot and Andrews (1992), for unknown breaking date, can be sustantial. An empirical application of the test to long U.S real GNP and GNP per capita series rejects the null of fractional integration in favour of a changing growth model with a break around World War II.

REFERENCES

Bai, J. and P. Perron (1998) "Estimating and testing linear models with multiple structural changes" *Econometrica*, 66, 47-78.

Baillie, R.T. (1996): "Long memory processes and fractional integration in economics and finance," *Journal of Econometrics*, 73, 15-131.

Balke, N. and R. J. Gordon (1989): "The estimation of prewar gross national product: methodology and new evidence," *Journal of Political Economy*, 97, 38-92.

Banerjee, A., R. Lumsdaine and J. H. Stock (1992): "Recursive and sequential tests for a unit root and structural breaks in long annual GNP series," *Journal of Business and Economic Statistics*, 10, 271-287.

Beran, J.(1996): *Statistics for long memory processes.* New York. Chapman and Hall.

BHATACHARYA, R.N., GUPTA, V.K., E: WAYMIRE (1983): " The Hurst effect under trends". *Journal of Applied Probability*, 20, 649-662.

BROCKWELL, P.J. AND R.A. DAVIS. (1996): LIntroduction to time series and forecasting. Springer, New York.

CHOI, K AND E. ZIVOT (2002) "Long memory and structural changes in the forward discount: An empirical investigation". University of Wisconsin (mimeo)

DAVIDSON, J. AND P. SIBBERTSEN (2003): "Generating schemes for long memory processes: Regimes, Aggregation and Linearity," University of Cardiff (mimeo).

DICKEY, D. A. AND W.A. FULLER (1981): "Likelihood ratio tests for autoregressive time series with a unit root," *Econometrica,* 49, 1057-1072.

DIEBOLD, F AND A. INOUE (2001) " Long memory and regime switching", *Journal of Econometrics*, 105, 131-159.

DIEBOLD, F. AND S. SHENHAJI (1996): " The uncertain unit root in Real GNP: Comment," *American Economic Review*, 86, 1291-98.

DING, Z. AND C.W.J. GRANGER (1996): " Modelling volatility persistence of speculative returns: A new approach", *Journal of Econometrics*, 73, 185-215.

DOLADO,J.,GONZALO, J. AND L.MAYORAL (2002)," A fractional Dickey-Fuller test for unit roots", *Econometrica*, 70, 1963-2006.

DOLADO, GONZALO, J. AND L.MAYORAL, (2003a), " Long range dependence in Spanish political polll series". *Journal of Applied Econometrics*, (mimeo).

DOLADO,J. GONZALO, J. AND L. MAYORAL, (2003b), "Testing for a unit root against fractional alternatives in the presence of a maintained trend", Universidad Carlos III (mimeo).

GIL-ALAÑA, L. AND P.M. ROBINSON (1997): " Testing of unit roots and other nonstationaty hypotheses in macoeconomic time series". *Journal of Econometrics*, 80, 241-268.

GRANGER, C.W.J. AND N. HYUNG (1999) " Occasional structural breaks and long memory". University of San Diego (mimeo).

GEWEKE, J. AND S. PORTER-HUDAK (1983) "The estimation and application of long memory time series models," *Journal of Time Series Analysis*, 4, 15-39.

MARINUCCI, D., AND P. ROBINSON (1999): "Alternative forms of Brownian motion," *Journal of Statistical Planning and Inference*, 80, 11-122.

MAYORAL, L. (2004) " Further evidence on the uncertain (fractional) unit root of real GNP".Universitart Pompeu Fabra (mimeo).

PARKE, W.R. (1999): " What is fractional integration ?" *Review of Economics and Statistics*, 81, 632-638.

PERRON, P.(1989): " The grat crash, the oil price shock, and the unit root hypothesis". *Econometrica*, 57, 1361-1401.

ROBINSON, P.M.(1994): "Efficient tests on nonstationarity hypotheses". *Journal of the american Economic Association*, 89, 1420-1437.

ROMER, C.D. (1989): "The prewar businness cycle reconsidered: New estimates of Gross National Product, 1869-1908," *Journal of Political Economy*, 97, 1-37.

RUDEBUSCH, G.D. (1993): " The uncertain unit root in real GNP," *American Economic Review,* 83, 264-272.

TEVEROSKY, V.AND M. TAQQU (1997): " Testing for long range dependence in the presence of shifting means or a slowly declining trend, using a variance-type estimator". *Journal of Time Series Analysis*, 18, 279-304.

TAQQU, M., WILLINGER, W. AND R. SHERMAN (1997) " Proof of a fundamental result in self-similar traffic modeling". *Computing Communication Review,* 27, 5-23.

ZIVOT E. AND D.W.K ANDREWS (1992) "Further evidence on the Great Crash, the Oil-Price Shock, and the Unit-Root Hypothesis", *Journal of Business and Economic Statistics*, 10, 3, pp.251-270.

APPENDIX

Proof of Theorem 1

The proof of consistency of $\hat{\phi}^{i}_{ols}$ is identical to that of Theorem 1 in DGM (2003).

With respect to the asymptotic distributions, consider first the case $0.5 < d \leq 1$, where the process is a non-stationary $FI(d)$ under the null hypothesis. Following Phillips (1988) define $\mathrm{B}^{i}(r)$ to be the stochastic process on [0,1] that is the projection residual in $L_2[0,1]$ of a fractional brownian motion projected onto the subspace generated by the following 1) $i = \mu : 1$, r^{-d} and 2) $i = \tau : 1$, r^{-d}, r^{1-d} and r. That is,

$$B_d(r) = \hat{\alpha}_0 + \hat{\alpha}_2 r^{-d} + B_d^{\mu}(r),$$

and,

$$B_d(r) = \hat{\alpha}_0 + \hat{\alpha}_2 r^{-d} + \hat{\alpha}_3 r^{1-d} + \hat{\alpha}_4 r + B_d^{\tau}(r),$$

where $B_d(r)$ is Type-I fBM as defined in Marinucci and Robinson (1999). Then, a straightforward application of the Frisch-Waugh Theorem provides for the desired result.

The case where $0 \leq d < 0.5$ is similar to that consider in DGM (2003) and therefore is omitted.■

Proof of Theorem 2

The result is obtained from using the weighting matrix $\Upsilon_T = diag(T^{-1/2}, T^{d-1/2}, T^{-1/2})$ in the vector of OLS estimators of $\theta = (\alpha, \delta, \phi)'$ in model (13) such that $\widehat{\theta} = \Upsilon_T^{-1}[\Upsilon_T^{-1} X'X \Upsilon_T^{-1}]^{-1}\Upsilon_T^{-}$ with $x_t = (1, \tau_t(d), y_{t-1})'$ and $z_t = \Delta^d y_t$ such that $\mu = 0$ in DGP1 (due to the invariance) and the following set of results:

$\sum \tau_t = \frac{1}{C_1(d)} O(T^{1-d})$, $\sum \tau_t^2 = \frac{1}{C_2(d)} O(\mathrm{T}^{1-2d})$ if $d \in (0, 0.5)$ and$= O(1)$ if $d \in [0.5,1]$, $\sum y_{t-1} = \delta(1-\lambda) O(T)$, $\sum y_{t-1}^2 = \frac{[\sigma_\varepsilon^2 + \delta^2(1-\lambda)]}{C_1(d)} O(T)$, $\sum \tau_t y_{t-1} = \frac{\delta(1-\lambda^{1-d})}{C_1(d)} O(T^{1-d})$,

$\sum \Delta^d y_t = \frac{1-\lambda^{1-d}}{C_1(d)} O(T^{1-d})$, $\sum \tau_t \Delta^d y_t = \frac{\delta[1-\lambda^{1-2d}]}{C_2(d)} O(T^{1-2d})$ if $d \in (0, 0.5)$ and$= O(1)$ if $d \in [0.5,1]$, $\sum y_{t-1} \Delta^d y_t = -T d \sigma_\varepsilon^2$

Proof of Theorem 3

Similar to Theorem 2

Proof of Theorem 4 (to be added)

Proof of Theorem 5

To be added