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

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

September 5, 2005

Abstract

This paper proposes a new time-domain test of a process being $I(d)$, $0 < d \leq 1$, under the null, against the alternative of being $I(0)$ with deterministic components subject to structural breaks at known or unknown dates, with the goal of disentangling the existing identification issue between long-memory and structural breaks. Denoting by $A_B(t)$ the different types of structural breaks in the deterministic components of a time series considered by Perron (1989), the test statistic proposed here is based on the t-ratio (or the infimum of a sequence of t-ratios) of the estimated coefficient on y_{t-1} in an OLS regression of $\Delta^d y_t$ on a simple transformation of the above-mentioned deterministic components and y_{t-1}, possibly augmented by a suitable number of lags of $\Delta^d y_t$ to account for serial correlation in the error terms. The case where $d = 1$ coincides with the Perron (1989) or the Zivot and Andrews (1992) approaches if the break date is known or unknown, respectively. The statistic is labelled as the SB-FDF (Structural Break-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 two empirical applications are provided.

1

1. INTRODUCTION

This paper proposes a new test for the null hypothesis that a time-series process exhibits long-range dependence (LRD) against the alternative that it has short memory, but suffers from structural shifts in its deterministic components. The problem of distinguishing between both types of processes 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 variances of partial sums of the series, etc. However, as pointed out some time ago in the applied probability literature, statistics based on short memory perturbed by some kind of nonstationarity may display similar properties as those prescribed by LRD under alternative assumptions (see e.g. Bhattacharya *et al.*, 1983, and Teverosky and Taqqu, 1997). In particular, this is the case of short-memory processes affected by shifts in trends or in the mean. In a certain sense, it can be thought that the inherent difficulty of this identification problem originates directly from the fact that some of the statistics used to detect LRD were originally proposed to detect the existence of structural breaks (see Naddler and Robbins, 1971). More recently, a similar issue has re-emerged in the econometric literature dealing with financial data. For example, Ding and Granger (1996), and Mikosch and Starica (2004) claim that the LRD behavior detected in both the absolute and the squared 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 or the Black Monday of 1987. On the contrary, Lobato and Savin (1998) conclude that the LRD found in the squared returns is genuine and, thus, not an spurious feature.

A useful starting point to pose this problem is to give some definitions of LRD (see e.g., Beran, 1994, Baillie, 1996, and Brockwell and Davies, 1991). In the time domain, LRD is defined for a stationary time series $\{y_t\}$ via the condition that $\lim_{j\to\infty} \sum_j |\rho_y(j)| = \infty$, 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 exponen-

tial decay. In the frequency domain, LRD requires that the spectral density $f_y(\omega)$ of the sequence is asymptotically of order $L(\omega)\omega^{-\nu}$ for some $\nu > 0$ and a slowly varying function $L(.)$, as $\omega \uparrow 0$. Both characterizations are not necessarily equivalent, but for fractionally integrated processes of order d (henceforth, $I(d)$), to which we restrict our attention in the rest of the paper when defining the null hypothesis. Specifically, y_t is said to be $I(d)$ 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, the variance of the partial sums of the series increases at the rate T^{1+2d}, and the normalized partial sums converge to fractional Brownian motion (fBM). Hence, fractional integration is a particular case of LRD.

In view of these properties, a simple way to illustrate the source of confusion between an $I(d)$ process and a short-memory one subject to structural breaks is to consider the following simple data generating process (DGP). Let y_t be generated by an $I(0)$ process whose mean is subject to a break at a known date T_B,

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

where u_t is a zero-mean $I(0)$ process with autocovariances $\gamma_u(j)$, $\lambda = T_B/T$ is the fraction of the sample where the break occurs, and $DU_t(\lambda) = \mathbf{1}(t > T_B)$, with $1 < T_B < T$, is an indicator function of the breaking date. Then, denoting the sample mean by $\overline{y}_T$, it is straightforward to show by means of the ergodic theorem that the sample autocovariances of the sequence $\{y_t\}_{t=1}^T$, given by

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

behave as follows when $T \uparrow \infty$,

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

for fixed $j \geq 1$ and $\lambda \in (0, 1)$. From (3), it can be observed that, even if the autocovariances $\gamma_u(j)$ decay to zero exponentially as $j \uparrow \infty$ for longer lags as u_t is $I(0)$, the sequence of sample autocovariances, $\widetilde{\gamma}_y(j)$, approaches a positive constant given by the second term in

3

(3) as long as $\alpha_2 \neq \alpha_1$. Thus, despite having a non-zero asymptote, the ACF of the process in (1) is bound to mimic the slow (hyperbolic) convergence to zero of LRD.[1] This presumption can be confirmed by performing a small Monte Carlo experiment. We simulate 1000 series of sample size $T = 20,000$, such that y_t is generated according to (1), with $\lambda = 0.5$, $\alpha_1 = 0$, $\varepsilon_t \sim n.i.d.(0, 1)$. Three cases are considered: $(\alpha_2 - \alpha_1) = 0$ (no break), 0.2 (small break) and 0.5 (large break). Then, in order to examine the consequences of ignoring the break in the mean in (1), we estimate the order of fractional integration, d, of the series by means of the well-known Geweke and Porter-Hudak (GHP, 1983) semiparametric 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 in Table 1, it becomes clear that the estimates of d increase monotonically with the size of the shift in the mean, giving the wrong impression that y_t is $I(d)$, $d > 0$ when it happens to be $I(0)$ (a more detailed explanation can be found in Perron and Qu, 2004). Additionally, Figure 1 depicts, for $T = 20,000$, the estimated ACFs of two processes. The first one is a process like (1), with $\lambda = 0.5$ and u_t being an $AR(1)$ process with a parameter equal to 0.7, while the second one is an $I(d)$ process with $d = 0.3$. As can be inspected, except for the first few autocorrelations, the ACFs behave very similarly in both cases. Thus, this type of result illustrates the source of confusion which has been stressed in the literature. The 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 $DT_t^*(\lambda) = (t - T_B)\mathbf{1}_{(T_B+1 \leq t \leq T)}$, $\beta_1 = 0.1$, and $\varepsilon_t \sim n.i.d(0, 1)$ yields estimates of d in the range (1.008, 1.0310), 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.

---

[1] This result has been recently generalized by Mikosch and Starica (2004) to the case of multiple breaks in the mean.
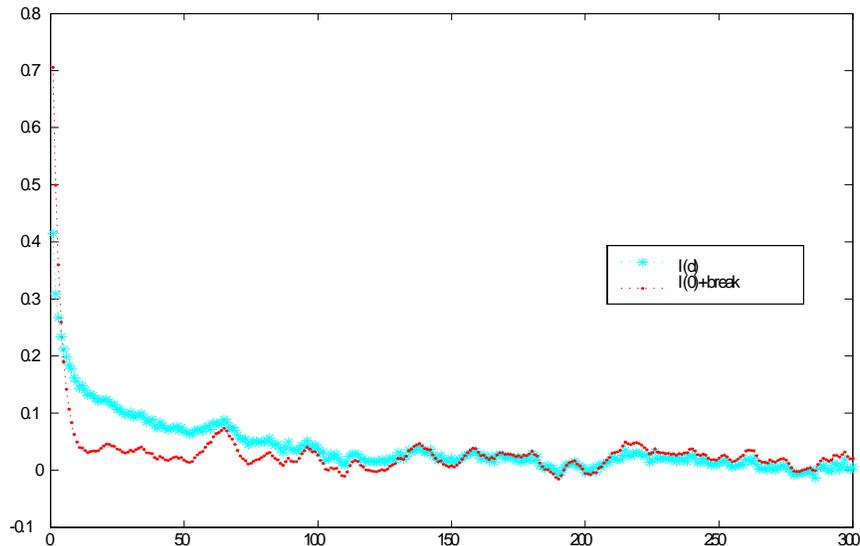
FIG. 1. Sample ACF of an I(d) and an $I(0)$+break processes.

Table 1

GPH 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% significance level (s.l.).

Along this way of reasoning, similar results about the possibility of confusing other types of nonlinear models with $I(d)$ processes have been derived very recently in slightly different frameworks to the one discussed in (1). First, there is Parke's (1999) error duration (ED) model, which considers the cumulation of a sequence of shocks that switch to 0 after a random delay that follows a power law distribution, so that if the delays were of infinite extent the process would be a random walk, and if of zero extent, an $i.i.d.$ process. Controlling the probability that a shock survives for k periods, p_k, to decrease with k at the

5

rate $p_k = s^{2d-2}$ for $d \in (0, 1]$, Parke shows that the ED model generates a process with the same autocovariance structure as an $I(d)$ process, i.e., $\gamma_y(j) = O(T^{2d-1})$ for large j. Secondly, there are Granger and Hyung 's (1999) and Diebold and Inoue's (2001) models which consider processes that are stationary and short memory, but exhibit periodic regime shifts, i.e., random changes in the mean of the series. For example, consider a DGP with $y_t = m_t + \varepsilon_t$ and $\Delta m_t = q_t \eta_t$, where $\Delta = (1 - L)$, and such that q_t follows an $i.i.d.$ binomial distribution where $q_t = 1$ with probability p and $q_t = 0$ with probability $(1 - p)$, and ε_t and η_t are independent $i.i.d.$ processes. Then, by assuming that the regime-shift probability p declines at a certain rate as the sample size increases, i.e., $p = O(T^{2d-2})$, it can be shown that the variance of the partial sums in this model will be related to the sample size in just the same way as in an $I(d)$ process, i.e., increasing at the rate T^{1+2d}. Therefore, the message to be drawn from all these works is that modelers face a hazard of mis-identification when the incidence of structural shifts is linked to the sample size in a particular $ad-hoc$ way.

On the whole, the class of models described above are nonlinear models capable of re-producing *some* observationally equivalent characteristics of $I(d)$ processes, albeit *not all*.[2] Moreover, they do so as long as the incidence of the shifts is related to the sample size in the specific fashion described earlier, which may be too restrictive in practice. For this reason, we consider more relevant for practitioners to develop a test statistic which helps to distinguish an $I(d)$ process from a short-memory process subject to a small number of breaks, in the spirit of DGP (1). Since among the $I(0)$ processes subject to breaks, the ones having more impact on empirical research are those popularized by Perron (1989), which are tested against $I(1)$ processes as the null, our main contribution in this paper is to extend Perron's testing approach to the more general setting of $I(d)$ processes, with $d \in (0, 1]$, instead of $d = 1$. For the most part, our analysis in the sequel will focus on the case of a single break, although we briefly conjecture about how to deal with processes containing more breaks.

In parallel with Perron (1989) who uses suitably modified Dickey-Fuller (DF) tests for

---

[2] Davidson and Sibbersten (2003) have recently demonstrated that the normalized sequence partial sums of $\{y_t\}$ generated by the ED and the other periodic regime shift models do not converge to fBM.

the $I(1)$ vs. $I(0)$ case in the presence of regime shifts, our strategy lies upon generalizing the Fractional Dickey-Fuller (FDF) approach proposed by Dolado, Gonzalo and Mayoral (DGM, 2002, 2004) to test $I(1)$ vs. $I(d)$, $0 \leq d < 1$, but now suitably modified to test $I(d)$ vs. $I(0)$ *cum* structural breaks. In DGM (2002) it was shown that if the order of fractional integration under the alternative is $0 \leq d < 1$ and no deterministic components are present, then an unbalanced OLS regression of the form $\Delta y_t = \phi \, \Delta^d y_{t-1} + \varepsilon_t$, yields a consistent test of $H_0 : d = 1$ against $H_1 : d \in [0, 1)$ based on the t-ratio of $\widehat{\phi}_{ols}$ in the previous regression model.[3] If the error term in the DGP is autocorrelated, then the regression should be augmented with a suitable number of lagged values of Δy_t. The degree of integration under the alternative hypothesis (d_1) can be taken to be known (in a simple alternative) or estimated with a $T^{\frac{1}{2}}$-consistent estimator (in a composite alternative).[4] If deterministic components, $\mu(t)$, are present under the null (say, a constant or a linear trend), DGM (2004) derive a FDF test now based on the regression $\Delta y_t = H(L)\mu(t) + \phi \Delta^d y_{t-1} + \varepsilon_t$, where $H(L) = \Delta - \phi \Delta^d L$.

Despite focusing on the test of $I(1)$ vs. $I(d)$ processes, DGM (2002) show that their results could be generalized to the case where the degrees of fractional integration under the null (d_0) and the alternative (d_1) verify the inequality $d_1 < d_0$. Accordingly, we propose in this paper a new test of $I(d)$ vs. $I(0)$ *cum* structural breaks, namely $d_0 = d \in (0, 0.5) \cup (0.5, 1]$[5] 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.

---

[3]To operationalise the FDF test , the regressor $\Delta^d y_{t-1}$ is constructed by applying the truncated binomial expansion of the filter $(1-L)^d$ to y_t, 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 given by $\pi_i = \Gamma(i - d) / \Gamma(-d) \Gamma(i + 1)$ with $\Gamma(.)$ the Gamma function.

[4]Empirical applications of such a testing procedure can be found in DGM (2003), whereas 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 has been developed in DGM (2004).

[5]Although the case of $d = 0.5$ was treated in DGM (2002), it constitutes a discontinuity point in the analysis of $I(d)$ processes; cf. Liu (1998). For this reason, as is often the case in the literature, we exclude this possibility in our analysis. Nonetheless, to simplify notation in the sequel, we will refer to the permissable range of d under the null as $0 < d \leq 1$.

To avoid confusion with the FDF for unit roots, the test presented hereafter will be denoted as the *Structural Break* FDF test (SB-FDF henceforth). It 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 different structural breaks. As in Perron (1989), we will consider the following possibilities: a crash shift, a changing growth shift, and a combination of both.[6]

At this stage it should be stressed that a test that extends Perron's DF testing approach of $I(1)$ vs. $I(0)$ *cum* structural breaks to a null of $I(d)$ with $d \in (0, 1]$ can be very useful in order to improve the power of the DF test when the true process is $I(d)$ but $d < 1$. In such an instance, Sowell (1990) and Krämer (1998) have shown that the DF and ADF tests are consistent, i.e. they reject the null of $I(1)$ with probability one as the sample size tends to infinity. However, the simulation results in Diebold and Rudebusch (1991) indicate that the power against fractional alternatives in finite samples can be low in some cases . To obtain further evidence on this issue, we report in Table 2 the rejection frequencies of Zivot and Andrews (1992) generalization of Perron's (1989) DF test in the case where the true DGP is an $I(d)$ process with $0 < d < 1$, and $I(1)$ is tested against $I(0)$ *cum* a changing growth shift, assuming that T_B is unknown. The number of replications is $5,000$ and $\varepsilon_t \sim n.i.d \, (0, 1)$. As can be observed, when $T = 100$, the rejection frequencies of the I(1) null hypothesis are high for values of d up to 0.7 but then, as one would expect, falls drastically as d gets closer to unity. For $T = 400$, this reduction in power occurs when d is above 0.8. In light of our previous discussion about the confound of LRD and structural breaks, these results seem to imply that, when the true series is $I(d)$, for low and moderate values of d, one is bound to find structural breaks "too often", whereas for high values of d, the false null of $I(1)$ will hardly be rejected. Thus, this evidence supports the need of a test where the null

---

[6]Note, however, that extensions to more than one break, along the lines of Bai and Perron (1998) and Bai (1999) should not be too difficult to devise once the simple case of a single break is is worked out. Some discussion on this case can be found in Section 4.

is $I(d)$.

<div align="center">

TABLE 2

POWER OF DF TEST OF I(1) VS. I(0)+BREAKS

Regression Model: $\Delta y_t = \alpha + \beta t + \gamma DT_t^*(0.5) + \phi y_{t-1} + \varepsilon_t$

</div>

DGP: $\Delta^{d_0} y_t = \varepsilon_t$; $\varepsilon_t \sim n.i.d\ (0, 1)$								
Sample size/d_0	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
T=100	100%	100%	99.9%	99.7%	94.4%	73.0%	38.0%	14.2%	4.0%
T=400	100%	100%	100%	100%	100%	99.9%	93.5%	49.9%	8.9%

Lastly, to place our SB-FDF test in the existing literature, it is convenient to differentiate between several characterizations of the process governing the evolution of a time series under the null and alternative hypotheses. These different characterizations depend upon the degree of integration of the series and the potential existence of structural breaks in its deterministic components. Specifically, four possibilities can be considered: (I) $I(0)$ and no structural breaks; (II) $I(0)$ and structural breaks; (III) $I(d)$ and no structural breaks, and (IV) $I(d)$ and structural breaks. The cases (I) against (II), and (I) against (III) have been extensively analyzed in the literature; cf. for the former see Perron (2005), and for the latter see Robinson (2003). The case (III) against (IV), when $0 < d < 0.5$, has also been treated, among others, by Hidalgo and Robinson (1996) when the break date is assumed to be known, and by Lazarova (2003) who extends their analysis to the case where the break date is considered to be unknown. Using a frequency-domain approach, Hidalgo and Robinson (1996) derive a test statistic whose limiting distribution under the null of no break is chi-squared for the former case, whereas in the case of an unknown break critical values need to be obtained by bootstrap methods (see Lazarova, 2003). Sibbertsen and Venetis (2004) have also studied the case where there are only shifts in mean, the break date is considered to be unknown, and $0 < d < 0.5$. Their test is based upon the (squared) difference between the GPH estimator of d and its tapered version, which converges to zero under the null of no break and diverges under the alternative of a break. It has a limiting chi-squared distribution if the errors are gaussian; otherwise, again bootstrap methods need to be used

<div align="center">9</div>

to obtain critical values. The case (II) versus (IV) has been treated by Giraitis, Kokoszka and Leipus (2001) who analyze how large the size of the mean shift in an I(0) series has to be in order to be mistaken with a stationary long-memory process $(0 < d < 0.5)$ in a given sample.

In comparison to all these previous works, the case considered in this paper is (III) against (II) which surprisingly has received much less attention in the literature. This is the case that can really help in solving the identification issue between LRD and SB which can be extremely relevant for at least three reasons: (i) shock identification (persistent vs. transitory), (ii) forecasting (do we need a long-history of the time series or only a short past will be of much use in forecasting?), and (iii) detection of spurious fractional cointegration (see Gonzalo and Lee, 1998). Indeed, Giraitis, Kokoszka and Leipus (2001) and the comments by Robinson to Lobato and Savin (1998) have stressed the importance of constructing a proper test statistic to distinguish between these alternative specifications. However, so far, the only attempt in this direction that we know of is Mayoral (2004a) whose approach relies upon a LR test in the time domain, but which is can only be applied to non-stationary processes under the null. Thus, to the best of our knowledge, there is not yet any available test in the literature for testing $I(d)$ versus $I(0)$ *cum* structural breaks allowing for both stationary and non-stationary processes under the null-hypothesis. This is the goal of the proposed SB-FDF test, which additionally presents the advantage of not requiring a correct specification of a parametric model and other distributional assumptions, besides being computationally simple since it is based on an OLS regression.

The rest of the paper is organized as follows. In Section 2 we derive the properties of the FDF test for $I(d)$ vs. $I(0)$ in the presence of deterministic components, like a constant or a linear trend, but without considering breaks yet, to next discuss the effects on this test of ignoring structural breaks in means or slopes when they exist.[7] Given that, as discussed above, the power of this test can be severely affected by parameter changes in

---

[7]Note that the FDF tests proposed in DGM (2002, 2004) refer to the case of $I(1)$ vs. $I(d)$. Thus, this section extends our previous results to the new setup of $I(d)$ vs. $I(0)$, with $d \in (0, 1]$.

the deterministic component of the process, the SB-FDF test of $I(d)$ vs $I(0)$ with a *single* structural break at a known or an unknown date is introduced in Section 3, where both its limiting and finite-sample properties are analyzed in detail. Section 4 contains a brief discussion of how to modify the test to account for autocorrelated disturbances, in the spirit of the ADF test, leading to the SB-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 two empirical applications; the first application deals with long U.S GNP series, where the stochastic or deterministic nature of their trending components has engendered some controversy in the literature; the second application centers on the behavior of the absolute values and squares of financial log-returns series, which has also been subject to some dispute. Finally, Section 6 concludes. Appendix A gathers the proofs of theorems and lemmae while Appendix B contains the tables of critical values for the cases where the limiting distributions are non-standard.

2. THE FDF TEST FOR $I(D)$ VS. $I(0)$

2.1 Preliminaries

Before considering the case of structural breaks, it is convenient to start by analyzing the problem of testing $I(d)$, with $0 < d \leq 1$, against trend stationarity, i.e. $d = 0$, along the lines of the FDF framework for the general case of $I(d_0)$ vs. $I(d_1)$ processes. 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 provide the foundation for the treatment of 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 $(\mu + \beta t)$,

$$y_t = \mu + \frac{\varepsilon_t 1_{(t>0)}}{\Delta^{d_0} - \phi L},\tag{4}$$

$$y_t = \mu + \beta t + \frac{\varepsilon_t 1_{(t>0)}}{\Delta^{d_0} - \phi L}, \tag{5}$$

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

$$\Delta^{d_0} y_t = \alpha + \Delta^{d_0} \delta + \phi y_{t-1} + \varepsilon_t, \tag{6}$$

$$\Delta^{d_0} y_t = \alpha + \Delta^{d_0} \delta + \gamma t + \Delta^{d_0-1} \varphi + \phi y_{t-1} + \varepsilon_t, \tag{7}$$

where $\alpha = -\phi\mu$, $\delta = \mu$, $\gamma = -\phi\beta$ and $\varphi = \beta$. For simplicity, hereafter we write $\varepsilon_t 1_{(t>0)} = \varepsilon_t$. Under H_0, when $\phi = 0$, $\Delta^{d_0} y_t = \Delta^{d_0}\mu + \varepsilon_t$ in (6) and $\Delta^{d_0} y_t = \mu\Delta^{d_0} + \beta\Delta^{d_0-1} + \varepsilon_t$ in (7).[8] Thus, $E(\Delta^d y_t) = \Delta^d \mu$ and $E(\Delta^d y_t) = \Delta^d(\mu + \beta t)$, respectively. Note that $\Delta^d \mu = \mu \sum_{i=0}^{t-1} \pi_i(d)$ and $\beta\Delta^{d-1} = \beta \sum_{i=0}^{t-1} \pi_i(d-1)$ 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 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 demonstrated in DGM (2004), $\tau(.)$ is a concave function for values of $d < 0$, being the function less steep the smaller (in absolute value) 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. Consequently, under H_1, y_t is $I(0)$ and admits the representation

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

Computing the trends $\tau_t(\xi)$, $\xi = d$ or $d - 1$ in (6) or (7) does not entail any difficulty since it only depends on d_0, which is known under H_0. As discussed earlier, the case where

---

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

$d_0 = 1$ and a linear trend is allowed under the alternative of $d_1 = d$, $0 < d < 1$, has been analyzed in DGM (2004) where it is shown that the FDF test is (numerically) invariant to the values of μ and β in the DGP.

Next, we derive the corresponding result for $H_0 : d_0 = d, d \in (0, 1]$ 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 defined as in (4) or (5) with $\phi = 0$, the OLS coefficient associated to ϕ in regression model (6), $\hat{\phi}_{ols}^\mu$, or (7), $\hat{\phi}_{ols}^\tau$, 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 $t - statistic$, $t_{\hat{\phi}_{ols}}^i$, $i = \{\mu, \tau\}$ is given by*

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

and

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

where $\xrightarrow{w}$ denotes weak convergence and $B_d^i(r)$ with $i = \{\mu, \tau\}$ is the L_2 projection residual from the continuous time regressions[9] $B_d(r) = \hat{\alpha}_0 + \hat{\alpha}_2 r^{-d} + B_d^\mu(r)$ and $B_d(r) = \hat{\alpha}_0 + \hat{\alpha}_1 r^{-d} + \hat{\alpha}_2 r^{1-d} + \hat{\alpha}_3 r + B_d^\tau(r)$, respectively.

The intuition for this result is similar to the one offered by DGM (2002) in the case of $I(1)$ vs. $I(d)$ processes with $0 < d < 1$. The different nature of the limiting distributions depend on the distance between d_0 and d_1. When d_0 ($= 1$ in the DGM 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 $0 < d < 0.5$. Also note that $d_0 = 1$ renders the standard DF limiting distribution.

---

[9]The parameters $\hat{\alpha}_0$, $\hat{\alpha}_1$, $\hat{\alpha}_2$, and $\hat{\alpha}_3$ solve $\min_{\alpha_0, \alpha_1} \int_0^1 \left| B_d(r) - \hat{\alpha}_0 - \hat{\alpha}_1 r^{-d} \right|^2 dr$ and $\min_{\alpha_0, \alpha_1, \alpha_2, \alpha_3} \int_0^1 \left| B_d(r) - \hat{\alpha}_0 - \hat{\alpha}_1 r^{-d} + \hat{\alpha}_2 r^{1-d} + \hat{\alpha}_3 r \right|^2 dr$ for the constant and constant and trend cases, respectively.

The finite-sample critical values of the cases considered in Theorem 1 are presented in Appendix B (Tables B1 and B2). Three sample sizes are considered, $T = 100$, 400 and $1,000$, and the number of replications is $10,000$. Table B1 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 $\varepsilon_t \sim n.i.d\,(0,\,1)$, when (6) is considered to be the regression model. Table B2, in turn, offers the corresponding critical values when (7) 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 for the three significance levels reported below are -1.28, -1.64 and -2.33, respectively,) when $0 < d < 0.5$, particularly for $T \geq 400$. However, for $d > 0.5$ the critical values start to differ drastically from those of a normal distribution, increasing in absolute value as d gets larger.

As for power, Table 3 reports the rejection rates at the 5% level (using the effective sizes in Table B2) of the FDF test in (7) in a similar Monte Carlo experiment to the one above, where now the DGP is an i.i.d. process cum a linear trend, i.e., $y_t = \alpha + \beta t + \varepsilon_t$, with $\alpha = 0.1$, $\beta = 0.5$. The main finding is that, except for low values of d and $T = 100$ where rejection rates of the null still reach 55%, the test turns out to be very powerful in all the other cases.

TABLE 3

POWER (CORRECTED SIZE: 5%)

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

DGP: $y_t = \alpha + \beta t + \varepsilon_t$; $\alpha = 0.1$, $\beta = 0.5$, $\varepsilon_t \sim n.i.d\,(0,\,1)$

$d_0/$ Sample size	$T = 100$	$T = 400$	$T = 1000$
0.2	54.9%	98.9%	100%
0.4	98.4%	100%	100%
0.7	100%	100%	100%
0.9	100%	100%	100%
1.0	100%	100%	100%

2.2 The effects of structural breaks on the FDF test of $I(d)$ vs. $I(0)$

Following Perron's (1989) analysis in his Theorem 1, our next step is to assess the effects on the FDF tests for $I(d)$ vs. $I(0)$ of ignoring the presence in the DGP of a shift in the mean of the series or a shift in the slope of the linear trend. Let us first consider the consequences of performing the FDF test with an *invariant* mean, as the one discussed above, when the DGP contains a break in the mean. Thus, y_t is assumed to be generated by,

$$DGP\ 1:\ y_t = \mu_0 + \zeta_0 DU_t(\lambda) + \varepsilon_t, \tag{8}$$

where $\varepsilon_t \sim iid(0, \sigma_\varepsilon^2)$ and $DU_t(\lambda) = \mathbf{1}_{(T_B+1 \leq t \leq T)}$. Ignoring the break in the mean, the SB-FDF test will be based on regression (6) which is repeated for convenience

$$\Delta^d y_t = \alpha + \delta \tau_t(d) + \phi y_{t-1} + \varepsilon_t. \tag{9}$$

Then, the following theory holds.

Theorem 2 *If y_t is given by DGP 1 in (8) and regression model (9) is used to estimate ϕ, when $T_B = \lambda T$ for all T and $0 < \lambda < 1$, then, as $T \to \infty$, it follows that,*

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

$$\hat{\phi}_{ols} \xrightarrow{p} \frac{d\sigma^2}{[\zeta_0^2 \lambda(1-\lambda) + \sigma_\varepsilon^2]},\ \text{if } 0.5\ <\ d \leq 1,$$

and,

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

where $\xrightarrow{p}$ denotes convergence in probability,

$$D\left(d, \sigma_\varepsilon^2\right) = C_1^2(d)\left[\zeta_0 \lambda(1-\lambda) + \sigma_\varepsilon^2\right] - C_2\{\zeta_0^2(2 - \lambda - \lambda^{1-d})\left(2\lambda - \lambda^{1-d} - 1\right) + \sigma_\varepsilon^2\},$$

and

$$C_1(d) = \Gamma(2-d),\ C_2(d) = \Gamma^2(1-d)(1-2d).$$

Theorem 2 shows that, under the crash hypothesis, the limit depends on the size of relative shift in the mean, ζ_0. Note that if $\lambda = 0$ or $\lambda = 1$, i.e., when there is no break, then $\widehat{\phi}_{ols} \xrightarrow{p} -d$. This result is quite intuitive since, being $y_t \sim I(0)$ under DGP 1, the covariance between $\Delta^d y_t$ and y_{t-1} is $\pi_1(d) = -d$. Further, for $d = 1$, the expression in part (a) in Theorem 1 of Perron (1989) is recovered, generalizing therefore his results to the more general case of $d \in (0, 1]$. Moreover, the fact that $\widehat{\phi}_{ols}$ converges to a finite negative number implies that $T^{1/2}\widehat{\phi}_{ols}$ for $d \in (0, 0.5)$, $T^d\widehat{\phi}_{ols}$ for $d \in (0.5, 1)$ and the corresponding t-ratios in each case will diverge to $-\infty$. Thus the FDF test for $I(d)$ vs. $I(0)$ would eventually reject the null hypothesis of $d = d_0$, $0 < d_0 < 1$, when it happens to be false. Notice, however, that, as in Perron's analysis, the power of the FDF test will be decreasing in the distance between the null and the alternative, namely as d gets closer to its true zero value, and in the size of the break, namely as ζ_0 gets larger relative to σ_ε^2.

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

$$DGP\ 2: \ y_t = \mu_0 + \beta_0 t + \psi_0 DT_t^*(\lambda) + \varepsilon_t, \tag{10}$$

where $DT_t^*(\lambda) = (t - T_B)\mathbf{1}_{(T_B+1 \leq t \leq T)}$. The FDF test is again implemented ignoring the breaking trend, that is, it is computed in the regression,

$$\Delta^d y_t = \alpha + \gamma t + \delta \tau_t(d) + \varphi \tau_t(d-1) + \phi y_{t-1} + \varepsilon_t. \tag{11}$$

Theorem 3 *If y_t is given by DGP 2 in (10) and regression model (11) is used to estimate ϕ, when $T_B = \lambda T$ for all T and $0 < \lambda < 1$, then, as $T \to \infty$, it follows that*

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

and,

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

In contrast to the result in Theorem 2, the FDF is unambiguously inconsistent when a breaking trend is ignored. The intuition behind this result, which again extends part (b)

of Theorem 1 in Perron (1989) to our more general setting, is that $\widehat{\phi}_{ols}$ is $O_p(T^{-d})$ with a positive limiting constant term for $d \in (0, 1]$ and that the sample s.d $(\widehat{\phi}_{ols})$ is $O_p(T^{-1/2})$, implying that the t-ratio is $O_p(T^{1/2-d})$. Therefore, it will tend to 0, for $d \in (0.5, 1]$, and to $+\infty$, for $d \in (0, 0.5)$.

In sum, the FDF test for $I(d)$ vs. $I(0)$ without consideration of structural shifts is not consistent against breaking trends and, despite being consistent against a break 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 FDF test that could distinguish an $I(d)$ process from a process being $I(0)$ around deterministic terms subject to structural breaks.

3. THE SB-FDF TEST OF $I(D)$ VS. $I(0)$ WITH STRUCTURAL BREAKS

Given the above considerations, we now proceed to derive the SB-FDF invariant test for $I(d)$ vs. $I(0)$ allowing for structural breaks under H_1. To account for structural breaks, we consider the following variant of (5) as the maintained hypothesis,

$$y_t = A_B(t) + \frac{a_t 1(t > 0)}{\Delta^d - \phi L}, \tag{12}$$

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. In line with Perron (1989) and Zivot and Andrews (1992), the null hypothesis of $I(d)$ will be implied by a value of $\phi = 0$ whereas $\phi < 0$ means that the process is $I(0)$. In line with these papers, three definitions of $A_B(t)$ are considered,

$$Case \ A: \ A_B^A(t) = \mu_0 + (\mu_1 - \mu_0)DU_t(\lambda), \tag{13}$$

$$Case \ B: \ A_B^B(t) = \mu_0 + \beta_0 t + (\beta_1 - \beta_0)DT_t^*(\lambda), \tag{14}$$

$$Case \ C: \ A_B^C(t) = \mu_0 + \beta_0 t + (\mu_1 - \mu_0)DU_t(\lambda) + (\beta_1 - \beta_0)DT_t(\lambda). \tag{15}$$

17

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 $DT_t(\lambda) = t\mathbf{1}_{(T_B+1\leq t\leq T)}$ with $\lambda = T_B/T$.

Initially, let us assume that the break date T_B is known a priori and that $a_t = \varepsilon_t$ where ε_t is an $i.i.d(0, \sigma^2)$ process. Then, the SB-FDF test of $I(d)$ vs. $I(0)$ in the presence of structural breaks is based on the t-ratio of the coefficient ϕ in the regression model,

$$\Delta^d y_t = \Delta^d A_B^i(t) - \phi A_B^i(t-1) + \phi y_{t-1} + \varepsilon_t, \ i = A, \ B, \ \text{and} \ C. \tag{16}$$

As above, the SB-FDF test is invariant to the values of μ_0, μ_1, β_0 and β_1 under H_0. Following the discussion in section 2.1, it is easy to check that under $H_1 : \phi < 0$, y_t is $I(0)$ and is subject to the regime shifts defined by $A_B^i(t)$. Conversely, under $H_0 : \phi = 0$, the process is $I(d)$ with $E[\Delta^d(y_t - A_B^i(t))] = 0$. Using similar arguments to those employed in Theorem 1, the following theory holds.

Theorem 4 *Let y_t be a process generated as in (12) with $a_t = \varepsilon_t \sim i.i.d\left(0, \sigma^2\right)$. Then, when $T_B = \lambda T$ for all T and $0 < \lambda < 1$, under the null hypothesis of $\phi = 0$, the OLS estimator associated to ϕ in regression model (16), as $T \rightarrow \infty$, is consistent. The asymptotic distribution of the associated $t-ratio$ is given by,*

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

$$t_{\widehat{\phi}(\lambda)}^i \xrightarrow{w} N(0,1) \ if \ d \ \in \ (0, 0.5),$$

where $B_d^{*i}(.,.)$ is the L_2 projection residual from the corresponding continuous time regressions associated to models $i = \{A, B, \text{and } C\}$, defined in Appendix A.

Although it was previously assumed that the date of the break T_B is known, in general this might not be the case. Thus, we explore in what follows how to implement the SB-FDF test under this more realistic situation.

For that, we follow the approach in Banerjee et al. (1992) and Zivot and Andrews (1992) by assuming that, under H_0, no break occurs, that is, $\phi = 0$, $\mu_0 = \mu_1$ and $\beta_0 = \beta_1$, implying

that (12) can be written as,

$$\Delta^d \left(y_t - \mu_0 \right) = a_t, \; if \; i = A, \; t = 1, \; 2, ..., \tag{17}$$

or,

$$\Delta^d \left(y_t - \mu_0 - \beta_0 t \right) = a_t, \; if \; i = \{B \text{ and } C\}, \; t = 1, \; 2, ..., \tag{18}$$

where a_t is an $I(0)$ process. As discussed before, under the alternative hypothesis, the process is $I(0)$ and may contain a single break in (some of) the parameters associated to the deterministic components, that occurs at an unknown time $T_B = \lambda T, \; \lambda \in \Lambda \subset (0, \; 1)$. Then, the potential break point under H_1 will be estimated in such a way that gives the highest weight to the $I(0)$ alternative. The estimation strategy will therefore consist in choosing the break point that gives the least favorable result for the null hypothesis of $I(d)$ using the SB-FDF test in (16) for each of the three cases, $i = A$, B, and C. The $t-statistic$ on $\hat{\phi}_{ols}^i$, $t_{\widehat{\phi}(\lambda)}$, is computed for the values of $\lambda \in \Lambda = (0.15, 0.85)$, following Andrews' (1993) choice of range, and then the infimum (most negative) value would be chosen to run the test. Thus, the test would reject the null hypothesis when

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

where $k_{\text{inf},\alpha}^i$ is a critical value to be provided in Appendix B. Under these conditions, the following theory holds.

Theorem 5 *Let y_t be a process generated as in (17) or (18) with $a_t \equiv \varepsilon_t \sim i.i.d \left(0, \; \sigma^2 \right)$ and possibly $\mu_0 = \beta_0 = 0$. Let Λ be a closed subset of $(0, \; 1)$. Then, under the null hypothesis of $\phi = 0$, the asymptotic distribution of the $t-statistic$ associated to ϕ in regression model (16), (with $A_B^A(t)$ if y_t is generated by (17) and $A_B^B(t)$ or $A_B^C(t)$ if y_t is generated in (18)) is given by,*

$$\inf_{\lambda \in \Lambda} t_{\widehat{\phi}(\lambda)}^i \xrightarrow{w} \inf_{\lambda \in \Lambda} \frac{\int_0^1 B_d^{*i} \left(\lambda, r \right) dB \left(r \right)}{\left(\int_0^1 B_d^{*i} \left(\lambda, r \right)^2 d \left(r \right) \right)^{1/2}} \; if \; d \in (0.5, 1], \; i = \{A, \; B \text{ and } C\}$$

and,

$$\inf_{\lambda \in \Lambda} t_{\widehat{\phi}(\lambda)}^i \xrightarrow{w} N(0, 1) \; if \; d \in (0, \; 0.5).$$

19

To compute critical values of the inf SB-FDF t-ratio test, a pure $I(d_0)$ process with $\varepsilon_t \sim n.i.d.(0, 1)$ has been simulated $10,000$ times, whereas the three regression models (A, B, and C) have been considered for samples of size $T = 100, 400, 1000$. Notice hat the sequence $\{t^i_{\hat{\phi}(\lambda)}\}$ are normally distributed and perfectly correlated. Hence, as shown in the proof of Theorem 5, the inf of this sequence corresponds to a N(0, 1) as well. However, in finite samples, there are several asymptotically negligible terms which depend on the product $(1 - \lambda)^{-1}T^{-2\delta}$, $0 < \delta < 1$, which may be sizeable for sufficiently large (small) values of λ (δ) at a given T (see equation A.2 in the Appendix). Tables B3, B4 and B5 in Appendix B report the corresponding critical values which, due to the presence of those terms, are larger (in absolute value) than the critical values of the SB-FDF test reported in Tables B1 and B2 when considering the left tail. Even for $T = 1000$, the critical values for $d \in (0, 0.5)$ are to the left of those of a $N(0, 1)$ and, in unreported simulations, we found that they slowly converge to them for sample sizes with around 5000 observations. Thus, for smaller sample sizes, we advice to use the effective critical values rather than the nominal ones.

In order to examine the power of the test, we have generated 5000 replications of DGP 2 (10) with sample sizes $T = 100$ and 400, where $\lambda = 0.5$, i.e., a changing growth model with a break in the middle of the sample. Both regression models B and C, in (16), have been estimated. Rejection rates are reported in Table 4 where, in order to compute the size-corrected power, the corresponding critical values in Tables B4 and B5 have been used. The power is very high except when $d = 0.1$ and $T = 100$ but even in this case it is close to 50% when $T = 400$.[10]

---

[10]In DGM (2002, 2004), we provide both Monte-Carlo and analytical results showing that the FDF test for $I(1)$ vs. $I(d)$ processes has better size-corrected power, except for very local alternatives (and even in this case the loss in power is small), than other available time-domain tests in the literature, like Tanaka's (1999) LM test. Unreported calculations (available upon request), show that this is also the case for our new $I(d)$ vs. $I(0)$ testing setup.

TABLE 4

POWER inf SB-FDF, I(d) vs. I(0) + S.B.

Regression Model: $\Delta^d y_t = \Delta^d A^i_B(t) - \phi A^i_B(t-1) + \phi y_{t-1} + \varepsilon_t$, $i = \{B, C\}$

DGP: $y_t = \mu_0 + \beta_0 t + \psi_0 DT^*_t(\lambda) + \varepsilon_t$; $\mu_0 = 1$, $\beta_0 = 0.5$, $\lambda = 0.5$, $\varepsilon_t \sim n.i.d\,(0, 1)$

| | $\psi_0 = 0.1$ | | | | $\psi_0 = 0.2$ | | | |
| | RModel B | | RModel C | | RModel B | | RModel C | |
d_0/Sample size	T=100	T=400	T=100	T=400	T=100	T=400	T=100	T=400
0.1	15.6%	50.3%	14.5%	48.9%	13.7%	48.9%	12.4%	47.4%
0.3	67.4%	74.9%	26.2%	73.5%	66.3%	73.4%	62.9%	71.3%
0.6	99.7%	100%	68.8%	100%	99.8%	100%	99.6%	100%
0.7	100%	100%	98.0%	100%	100%	100%	100%	100%
0.9	100%	100%	99.9%	100%	100%	100%	100%	100%

4. AUGMENTED SB-FDF 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 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 typically employed in the literature. One is the non-parametric approach proposed by Phillips and Perron (1988) 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 the AFDF test in Theorem 7 of DGM (2002), it can be conjectured 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)

21

process $\alpha_p(L)u_t = \beta_q(L)\varepsilon_t$ with $E |\varepsilon_t|^{4+\delta} < \infty$ for some $\delta > 0$, then the *inf* SB-FDF test based on the t-ratio of $\widehat{\phi}_{ols}$ in (19) 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* SB-FDF test (denoted as SB-AFDF) will be based on the regression model,

$$\Delta^d y_t = \Delta^d A_B^i(t) - \phi A_B^i(t-1) + \phi y_{t-1} + \sum_{j=1}^{k} \varsigma_j \Delta^d y_{t-j} + a_t, \quad i = A, B, \text{ and } C. \quad (19)$$

Although a generalization of the previous results to multiple breaks is not considered in this paper, it is likely that this extension can be done following the same reasoning as in the procedure devised by Bai and Perron (1998). In their framework, where there are m possible breaks affecting the mean and the trend slope, they suggest the following procedure to select the number of breaks. Letting *sup* $F_T(l)$ 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 against an unknown number of breaks given some specific bound on the maximum number of shifts considered. The first one is the double maximum statistic (UD$_{\text{max}}$) where UD$_{\text{max}} = \max_{1 \leq k \leq m} supF_T(l)$ while the second one is $supF_t(l + 1/l)$ 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 $supF_t(l + 1/l)$ statistics. Therefore, our proposal is to use such a procedure to determine $\lambda_1, ..., \lambda_k$ in the $A_B(t)$ terms in (16). 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 would 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 exceeds the scope of this paper but they are in our future research agenda.

5. EMPIRICAL APPLICATIONS

In order to provide some empirical illustrations of how easy the SB-FDF test can be used in practice, we consider the following two applications.

5.1 Real GNP.—

The first application deals with the log. of a long series of U.S. real GNP, y_t, which basically corresponds to the same data set used in Diebold and Senhadji (1996) (DS henceforth) in their 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 original sample ending in 1995; cf. Mayoral (2004b), for a detailed discussion of the construction of the series. Since this series is based on the historical annual real GNP series constructed by Balke and Gordon (1989), it is denoted as GNP-BG.[11]

According to DS's analysis, there is conclusive evidence in favor 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 the series. Then, they compute the exact finite sample distribution of the t-ratios of the lagged (log. of) GNP-BG 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 the ADF test is very small under the D-ST model but quite large under the T-ST model, providing overwhelming support in favour of 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 (2004b) has pointed out that when the same exercise is done with the well-known KPSS test, where the null is TS-T, it is also rejected in both series. This inconcluive outcome leads this author 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 $I(d)$, $0 < d < 1$. Considering values of d in the range 0.6-0.7, she finds favorable evidence for an $I(d)$ 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. Nonetheless, from inspection of Figure 2, where the log. of GNP-BG is displayed, one could as well conjecture that the data are generated by a T-ST process subject to some

---

[11] Following DS (1996), we have also implemented the SB-FDF test to a long series of U.S. real GNP based on the historical annual series of Romer (1989). The results, available upon request, are not reported since they are very similar to those displayed in Table 5.
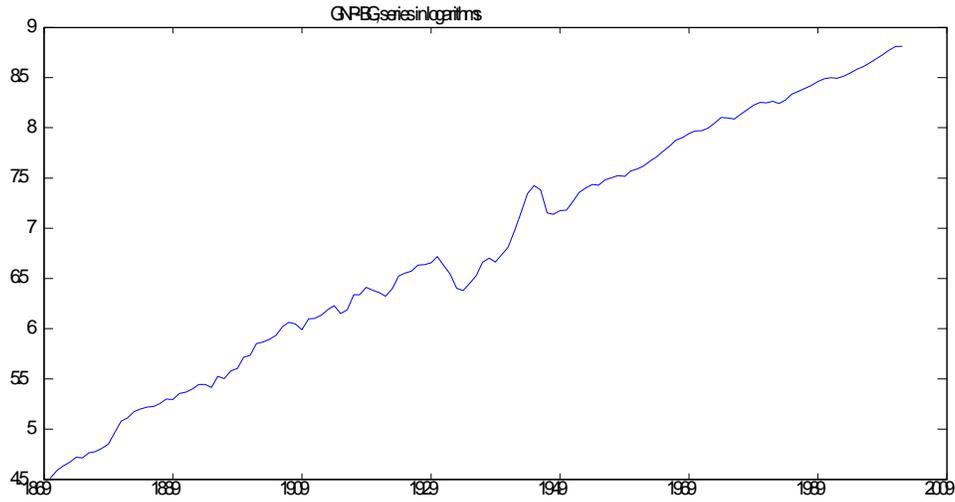
Fɪɢ. 2. Plot of (logged) real GNP-BG series.

structural breaks as in Perron (1989). Hence, given the mixed evidence about the data being generated either by an $I(d)$ process or by an $I(0)$ process *cum* structural breaks, this example provides a good illustration of the usefulness of the SB-FDF test.

In Table 5, we report the t-ratios of the *inf* SB-AFDF test constructed according to (19) where up to three lags of $\Delta^d y_t$ have been included as additional regressors in order to account for residual correlation. The critical values that have been employed are those reported in Tables B4, B5 for $T = 100$. Values of d in the non-stationary (albeit mean-reverting) range (0.5, 1) have been used to construct $\Delta^d y_t$ and $\Delta^d A_B^i(t)$, $i = B$, and C, since the trending behaviour of the series precludes the use of model A which does not include a linear trend in the maintained hypothesis.

24

TABLE 5

inf SB-AFDF Tests

GNP-BG series

Model	Model B				Model C			
Lags/d	0.6	0.7	0.8	0.9	0.6	0.7	0.8	0.9
0	1.594	0.257	-0.963	-2.088	1.598	0.268	-0.946	-2.067
1	-3.660†	-3.792†	-3.976†	-4.216†	-3.622	-3.755†	-3.940†	-4.181
2	-3.384	-3.719	-4.047	-4.376	-3.335†	-3.667	-3.993	-4.320†
3	-2.587	-2.939	-3.003	-3.311	-2.564	-2.910	3.260	-3.620
C.V. (95%)	-3.85	-4.25	-4.54	-4.88	-4.15	-4.53	-4.86	-5.20

†number of lags chosen by the AIC criterion. *Rejection at the 95% s.l.

As can be inspected, the null of $I(d)$ cannot be rejected at the conventional significance level. Interestingly, this result against T-ST is reinforced by application of the conventional Zivot and Andrews' (1992) *inf* test for the null of $I(1)$, with two lags, which yields values of -4.23 (model B) and -5.03 (model C) against 5% critical values of -4.42 and -5.08, respectively. However, concluding that the series is $I(1)$ may not be correct (see Table 2), given the results of the SB-AFDF test which does not reject the null for large values of d, yet below unity.

In sum, the evidence provided by the SB-FDF test seems to point out that the GNP-BG series behaves as an $I(d)$ process with a large d. This result is qualitatively consistent with other empirical investigations of fractional processes in GNP, such as Sowell (1992), on the basis that GNP is obtained by aggregating heterogeneously persistent sectorial value added which, according to Granger's (1980) aggregation argument, yields long memory. In the same direction, recently Michelacci (2004) has shown that if Gibrat's law fails and small firms grow faster than big firms, as empirical evidence suggests, aggregate output should exhibit a fractional order of integration.

5.2 Financial returns.—

25

The second application deals with the (absolute values and squared) financial returns series, r_t, obtained from the Standard & Poor's 500 composite stock index over the period January 2, 1953 to October 10, 1977 (the number of observations is 6217). This series has been modelled in several papers where it has been argued that shifts in the unconditional variance of an ARCH or GARCH model may induce the typical ACF of a long memory process (see, i*nter alia*, Ding et al., 1996, Lobato and Savin (1998), and Mikosch and Starica, 2004). Notice that, although the sample considered in these papers are quite longer than ours, we restrict our sample to 1953-1977 because Mikosch and Starica (2004) claim that there is a single structural break in the constant term of a GARCH (1,1) model over that period, as a consequence of the first oil crisis in 1973.[12] Since our proposed approach refers to a single shift, we deem this an appropriate choice of sample size. Further, the end of the sample has been chosen so that the potential break date in included in the range $\lambda \in [0.15, 0.85]$ used to implement the *inf* SB-AFDF test. From Figure 3, where the stock returns are depicted, it becomes clear that the series experiences a higher variance after 1973. Figures 4a and b shows the ACF for the absolutes values, $|r_t|$, and the squares, r_t^2, respectively, which display the typical plateau for longer lags, as if LRD were present. The results of performing the *inf* SB-AFDF test for an unknown break date in version A (no trend) of model (19) are shown in Table 6, where the number of lags is $k = 15$ (further lags were insignificant), and only values of $d < 0.5$ have been used, in agreement with the estimates of d obtained from the estimation of ARFIMA models applied to both series. Since the sample size in this case is large enough to use the N(0, 1) approximation for the limiting distribution, a 5% critical value of -1.64 is used. The main finding is that the null hypothesis of $I(d)$ cannot be rejected for moderate values of d, in the range $0.1 - 0.4$ for $|r_t|$, and $0.1 - 0.3$ for r_t^2, which contain the estimated values of d for this data set.[13] Note that in the case of $|r_t|$, the null of $d = 0.4$ could be almost rejected at the 10% level (with a c.v. of

---

[12] For example, the estimates of the GARCH (1,1) model $\sigma_t^2 = \alpha_0 + \alpha_1 \sigma_{t-1}^2 + \beta_1 h_{t-1}^2$ yield $\alpha_0 = .325 x 10^{-6}$, $\alpha_1 = 0.150$, $\beta_1 = 0.600$ for T=1953-1972 and $\alpha_0 = .1.40 x 10^{-5}$, $\alpha_1 = 0.150$, $\beta_1 = 0.600$ for T=1953-1977.

[13] The estimated values of d using the GPH approach with $g(T) = T^{0.5}$, are $\widehat{d} = 0.4005$ for the absolute values, and $\widehat{d} = 0.3241$ for the squares of the returns.
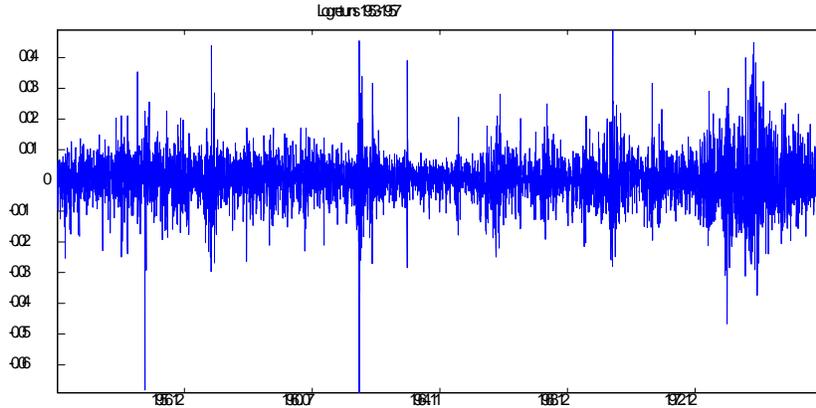
Fɪɢ. 3. Plot of S&P 500 daily returns, 1953-1977.

-1.28) whereas it is clearly rejected in the case of r_t^2 at the 5% level.[14] Hence, overall , the SB-AFDF test yields some evidence in favor of these transformation of the stock returns behaving during 1953-1977 as $I(d)$ processes with a low degree of fractional integration. Of course, fractional integration and structural breaks can cohabit (for an empirical illustration see Choi and Zivot, 2005), yet its consideration is beyond the scope of this application.

TABLE 6

inf SB-AFDF ᴛᴇsᴛs

	S&P 500 ᴅᴀᴛᴀ 1953-1977			
d	0.1	0.2	0.3	0.4
$|r_t|$	2.561	1.549	0.294	−1.214
r_t^2	1.063	0.037	-1.216	-2.680*

*Rejection at the 95% s.l.; $r_t = \log P_t - \log P_{t-1}$.

[14]This conclusion is reinforced when we perform the test for the nonstationary null of $d = 0.6$. With a 5% c.v. of -2.54 (see the left panel in Table B3), the values of the test for the absolute values and the squares are -4.63 and -6.15, respectively. Thus the null is strongly rejected.
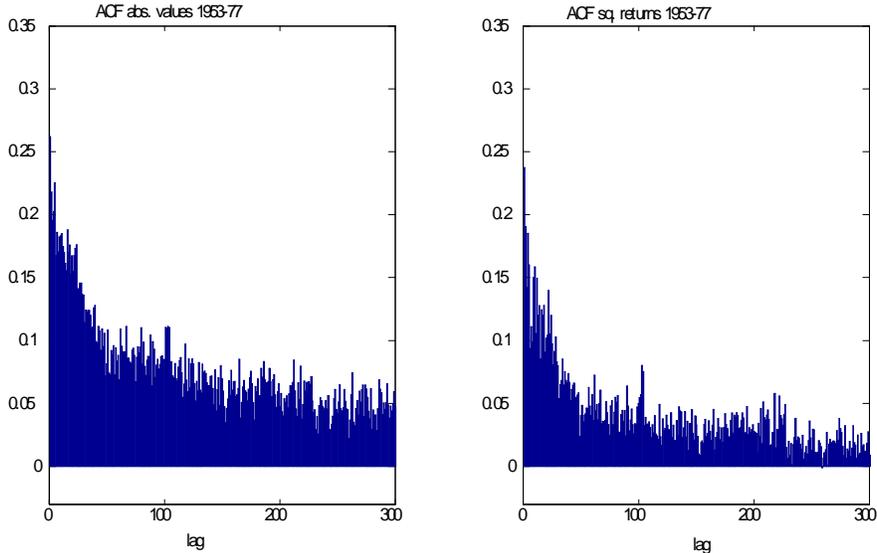
FIG. 4. Sample ACF abs. values and squares of S&P 500 returns.

6. CONCLUSIONS

One of the recent identification issues in time series is the difficulty in distinguishing long-memory from structural breaks. This can be extremely relevant for at least three reasons: (i) shock identification (persistent vs. transitory), (ii) forecasting (do we need a long-history of the time series or only a short past will be of much use in forecasting?), and (iii) detection of spurious fractional cointegration. In order to contribute to solve this empirical identification problem, 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, denoted as *Structural Break* Fractional Dickey-Fuller (SB-FDF) test, is a time-domain one, performs fairly well in finite samples in terms of power, and it is easy to implement since it relies on a time-series OLS regression. Denoting by $A_B(t)$ the different types of a single structural break considered by Perron (1989), the SB-FDF test is based on the t-ratio of the coefficient on y_{t-1} in an OLS regression of $\Delta^d y_t$ on $\Delta^d A_B(t)$, $A_B(t-1)$ and y_{t-1}. A suitable number

of lags of $\Delta^d y_t$ may be added to account for serially correlated errors. When $d \in (0.5, 1]$ its asymptotic distribution is nonstandard and critical values are simulated. By contrast, when $d \in (0, 0.5)$, it is asymptotically normally distributed. In future research we plan to extend the proposed testing approach to multiple breaks along the lines discussed in Section 4 of this paper, as well as to test $I(d_0)$ versus $I(d_1)$ *cum* structural breaks with $d_0 > d_1$.

REFERENCES

ANDREWS, D. W. K. (1993): "Test for parameter instability and structural break with unknown change point," *Econometrica*, 61, 139-165.

BAI, J. (1999): "Likelihood ratio tests for multiple structural changes," *Journal of Econometrics*, 91, 299-323.

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.(1994): *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 (1991): *Introduction to time series and forecasting.* Springer, New York.

CHOI, K. AND E. ZIVOT (2005): "Long memory and structural breaks in the forward discount: An empirical investigation," *Journal of International Money and Finance* (forthcoming).

DAVIDSON, J. (1994): *Stochastic Limit Theory.* New York: Oxford University Press.

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 G. RUDEBUSCH (1991): "On the power of the Dickey-Fuller tests

against fractional alternatives," *Economic Letters*, 35, 155-160.

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

DIEBOLD, F. AND S. SENHADJI (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, J., GONZALO, J. AND L.MAYORAL, (2003): "Long range dependence in Spanish political poll series," *Journal of Applied Econometrics*, Vol. 18, No. 2, 137-155.

DOLADO,J. GONZALO, J. AND L. MAYORAL, (2004): "Testing I(1) vs. I(d) alternatives in the presence of deterministic components," Universidad Carlos III (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.

GIRAITIS, L., KOKOSZKA, P. AND LEIPUS, R. (2001): "Testing for long memory in the presence of a general trend," *Journal of Applied Probability*, 38, 1033-1054.

GONZALO, J. AND T. LEE (1998): "Pitfalls in testing long run relationships," *Journal of Econometrics*, 86, 129-154.

GRANGER, C.W.J. (1980): "Long memory relationships and the aggregation of dynamic models," *Journal of Econometrics*, 14, 227-238.

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

HIDALGO, J. AND P. ROBINSON (1996): "Testing for structural change in a long memory environment," *Journal of Econometrics*, 73, 261-284.

HOSKING, J.R.M. (1996): "Asymptotic distributions of the sample mean, autocovariances, and autocorrelations of long memory time series," *Journal of Econometrics*, 73, 261-284.

KRAMER, W. (1998): "Fractional integration and the augmented Dickey-Fuller test,"

Economics Letters, 61, 269-272.

LAZAROVA, S. (2003): "Testing for structural change in regression with long memory processes," *Journal of Econometrics* (forthcoming).

LIU, M. (1998): "Asymptotics of nonstationary fractionally integrated series," *Econometric Theory*, 14, 641-662.

LOBATO, I.N. AND N.E. SAVIN. (1998): "Real and spurious long memory properties of stock market data," *Journal of Business and Economic Statistics*, 16, 261-268.

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

MARMOL, F. AND C. VELASCO (2002): "Trend stationarity versus long-range dependence in time series analysis," *Journal of Econometrics*, 108, 25-42.

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

MAYORAL, L. (2004b): "Is the observed persistence spurious? A test for fractional integration versus structural breaks," Universidad Pompeu Fabra (mimeo).

MICHELACCI, C. (2004): "Cross-Sectional heterogeneity and the persistence of aggregate fluctuations," *Journal of Monetary Economics*, 51, 1321-1352.

MIKOSCH, T. AND C. STARICA (2004): "Nonstationarities in financial time series, long range dependence and the IGARCH model," *Review of Economics and Statistics*, 86, 378-390.

NADDLER, J. AND N.B. ROBBINS (1971): "Some characteristics of Page's two-sided procedure for detecting a change in a location parameter," *The Annals of Mathematical Statistics*, 42, 538-551.

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

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

PERRON, P. (2005): "Dealing with Structural Breaks," in T. C. Mills and K. Patterson (editors), *Palgrave Handbook of Econometrics* (vol 1: *Econometric Theory*). Palgrave

MacMillan.

PERRON, P. AND Z. QU (2004): "An analytical evaluation of the log-periodogram estimate in the presence of level shift and its implications for stock returns volatility," Boston University (mimeo).

PHILLIPS, P.C.B AND P. PERRON (1988): "Testing for a unit root in time series regressions," *Biometrika*, 75, 335-345.

ROBINSON, P.M. (2003): *Time Series with Long Memory.* Edited by P. Robinson. Oxford University Press.

ROMER, C.D. (1989): "The prewar business 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.

SIBBERTSEN, P. AND I. VENETIS (2004): "Distinguishing between long-range dependence and deterministic trends," Mimeo.

SOWELL, F. B. (1990): "The fractional unit root distribution," *Econometrica*, 58, 495-505.

SOWELL, F. B. (1992): "Modelling long-run behavior with the fractional ARIMA model," *Journal of Monetary Economics, 29*, 277-302.

TANAKA, K. (1999): "The nonstationary fractional unit root," *Econometric Theory*, 15, 549-582.

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.

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, 251-270.

APPENDIX A

Proof of Theorem 1

The proof of consistency of $\hat{\phi}^i_{ols}$ is analogous to that of Theorem 1 in DGM (2004) and therefore is omitted.

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. Define $B^i_d(r)$ 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) $i = \mu : \left(\begin{array}{cc} 1, & r^{-d} \end{array} \right)$ and 2) $i = \tau : \left(\begin{array}{cccc} 1, & r^{-d}, & r^{1-d}, & r \end{array} \right)$. That is,

$$B_d(r) = \hat{\alpha}_0 + \hat{\alpha}_1 r^{-d} + B^\mu_d(r),$$

and,

$$B_d(r) = \hat{\alpha}_0 + \hat{\alpha}_1 r^{-d} + \hat{\alpha}_2 r^{1-d} + \hat{\alpha}_3 r + B^\tau_d(r),$$

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

The case where $0 \leq d < 0.5$ is similar to that consider in DGM (2004) and thus is omitted for the sake of brevity. ∎

Proof of Theorem 2

The result is obtained from using the weighting matrix $\Upsilon_T = diag(T^{1/2}, T^{1/2-d}, T^{1/2})$ if $d \in (0, 0.5)$, and $\Upsilon_T = diag(T^{1/2}, 1, T^{1/2})$ if $d \in (0.5, 1]$, in the vector of OLS estimators of $\theta = (\alpha, \delta, \phi)'$ in model (9) such that $\hat{\theta} = \Upsilon_T^{-1}[\Upsilon_T^{-1}X'X\Upsilon_T^{-1}]^{-1}\Upsilon_T^{-1}X'z$, where the rows of the matrix X are given by $x_t = (1, \tau_t(d), y_{t-1})$ while the elements of the vector z are defined as $z_t = \Delta^d y_t$. Assuming $\mu_0 = 0$ in (8) (due to the invariance of the limiting distribution of the test to the value of μ_0 in DGP 1) then the following set of results hold (with the omitted sums limits going from 2 to T):

1. $\lim_{T \to \infty} \frac{\sum \tau_t}{T^{1-d}} = \frac{1}{C_1(d)}$,

2. $\lim_{T \to \infty} \frac{\sum \tau_t^2}{T^{1-2d}} \sum \tau_t^2 = \frac{1}{C_2(d)}$, if $d \in (0, 0.5)$ and $\lim \sum \tau_t^2 = O(1)$ if $d \in (0.5, 1]$,

3. $\frac{\sum y_{t-1}}{T} \xrightarrow{p} \zeta_0(1-\lambda)$,

4. $\frac{\sum y_{t-1}^2}{T} \xrightarrow{p} \frac{[\sigma_\varepsilon^2 + \zeta_0^2(1-\lambda)]}{C_1(d)}$,

5. $\frac{\sum \tau_t y_{t-1}}{T^{1-d}} \xrightarrow{p} \frac{\zeta_0(1-\lambda^{1-d})}{C_1(d)}$,

6. $\frac{\sum \Delta^d y_t}{T^{1-d}} \xrightarrow{p} \frac{1-\lambda^{1-d}}{C_1(d)}$,

7. $\frac{\sum \tau_t \Delta^d y_t}{T} = \frac{\zeta_0[1-\lambda^{1-2d}]}{C_2(d)} O_p(T^{1-2d})$ if $d \in (0, 0.5)$ and

   $\sum \tau_t \Delta^d y_t = O_p(1)$ if $d \in (0.5, 1]$,

8. $\frac{\sum y_{t-1} \Delta^d y_t}{T} \xrightarrow{p} -d\sigma_\varepsilon^2$.

To obtain the expressions for $C_1(d)$ and $C_2(d)$, notice that the j-th coefficient in the binomial expansion of $(1-L)^d$ is $\pi_j(d) = \frac{\Gamma(j-d)}{\Gamma(-d)\Gamma(j+1)} \sim \frac{1}{\Gamma(-d)} j^{-(d+1)}$. Also notice that since $\sum_{i=0}^{\infty} \pi_j(d) = 0$ for any $d > 0$, then $\tau_t = \sum_{i=0}^{t-1} \pi_j(d) = -\sum_{i=t}^{\infty} \pi_j(d) = \frac{1}{\Gamma(-d)} t^{-d}$ (see Davidson, 1994, p.32). Hence,

$\sum \tau_t \simeq \frac{1}{\Gamma(-d)} \sum_{s=2}^{T} t^{-d} \simeq \frac{1}{\Gamma(-d)} \cdot \frac{1}{(-d)(1-d)} T^{1-d}$, where $\Gamma(-d)(-d)(1-d) = \Gamma(2-d) \equiv C_1(d)$.

Likewise, for $d \in (0, 0.5)$,

$\sum \tau_t^2 \simeq \frac{1}{\Gamma^2(-d)} \sum_{s=2}^{T} t^{-2d} \simeq \frac{1}{\Gamma^2(-d)} \cdot \frac{1}{(-d)^2(1-2d)} T^{1-2d}$, where $\Gamma^2(-d)(-d)^2(1-2d) = \Gamma^2(1-d)(1-2d) \equiv C_2(d)$.

Finally, denoting the elements of the matrix $A = [\Upsilon_T^{-1} X' X \Upsilon_T^{-1}]$ by a_{ij} $(i, j = 1, 2, 3)$, its determinant by $det(A)$ and the element of the vector $\Upsilon_T^{-1} X' z$ by b_i $(i = 1, 2, 3)$, notice that $\lim_{T\to\infty} \widehat{\phi}_{ols} = (a_{22} - a_{12}^2) b_3 / det(A)$ if $d \in (0, 0.5)$ and $\lim_{T\to\infty} \widehat{\phi}_{ols} = a_{22} b_3 / det(A)$ if $d \in (0.5, 1]$. Substitution of the corresponding limiting expressions above yields the required result.∎

Proof of Theorem 3

Similar to Theorem 2, using the weighting matrix $\Upsilon_T = diag(T^{1/2}, T^{3/2}, T^{1/2-d}, T^{3/2-d}, T^{1/2})$ if $d \in (0, 0.5)$, and $\Upsilon_T = diag(T^{1/2}, T^{3/2}, 1, T^{3/2-d}, T^{1/2})$ if $d \in (0.5, 1]$, in the vector of OLS estimators of $\theta = (\alpha, \gamma, \delta, \varphi, \phi)'$ in model (11) such that

$$\widehat{\theta} = \Upsilon_T^{-1} [\Upsilon_T^{-1} X' X \Upsilon_T^{-1}]^{-1} \Upsilon_T^{-1} X' z,$$

where the rows of the matrix X are given now by $x_t = (1, t, \tau_t(d), \tau_t(d-1), y_{t-1})$ while the elements of the vector z are $z_t = \Delta^d y_t$, under the assumption that $\mu_0 = \beta_0 = 0$ in (10) (due to the invariance of the limiting distribution of the test to the value of μ_0 and β_0 in DGP 2).■

Proof of Theorem 4

The proof of this result is similar to that of Theorem 1 where, in this case, the corresponding 'detrended' fractional Brownian motion are obtained in the continuous time regressions defined as,

$$B_{d_0}^A (r) = \hat{\alpha}_0 + \hat{\alpha}_1 du(\lambda, r) + \hat{\alpha}_2 r^{-d} + \hat{\alpha}_3 r^{-d} du(\lambda, r) + B_{d_0}^{*A}(r, \lambda),$$

$$B_{d_0}^B (r) = \hat{\alpha}_0 + \hat{\alpha}_1 r + \hat{\alpha}_2 dt^*(\lambda, r) + \hat{\alpha}_3 r^{-d} + \hat{\alpha}_4 r^{d-1} + \hat{\alpha}_5 r^{-d} dt^*(\lambda, r) r^{-d} + B_{d_0}^{*B}(r, \omega),$$

and,

$$\begin{aligned} B_{d_0}^C (r) &= \hat{\alpha}_0 + \hat{\alpha}_1 r + \hat{\alpha}_2 du(\lambda, r) + \hat{\alpha}_3 dt^*(\lambda, r) + \hat{\alpha}_4 r^{-d} + \hat{\alpha}_5 r^{d-1} \\ &\quad + \hat{\alpha}_6 r^{-d} du(\lambda, r) + \hat{\alpha}_7 r^{-d} dt^*(\lambda, r) + B_{d_0}^{*C}(r, \omega), \end{aligned}$$

for models A, B and C, respectively, where $du(\lambda, r) = 1$ if $r > \lambda$ and 0 otherwise, $dt^* = r - \lambda$ if $r > \lambda$ and 0 otherwise.■

Proof of Theorem 5

1) Case $0.5 < d \leq 1$

The proof of this theorem is constructed along the lines of that of Theorem 1 in Zivot and Andrews (1992) (Z&A henceforth). They present an alternative approach to the traditional 'fidi plus tightness' method based on first showing that a set of relevant variables jointly converge and then using the Continuous Mapping theorem to complete the proof.

Following the notation in Z&A, let us define $z_{tT}^i(\lambda)$ for $i = \{A, B,$ and $C\}$ as the vector that contains the deterministic components for each model under the alternative hypothesis that depends explicitly on the break fraction and the sample size. For instance if $i = A$, $z_{tT}^A(\lambda)' = \left(1, \ DU_t(\lambda), \ \tau_t(d_0), \ (\tau_t(d_0) DU_t(\lambda)) \right)$. We will also need a rescaled version of the deterministic regressors, $\tilde{z}_T^i(\omega, r) = \delta_T^i z_{[Tr]T}(\lambda)$, where δ_T^i is a diagonal matrix of weights. The test statistics of interest is,

$$\inf_{\lambda \in \Lambda} t^i_{\phi(\lambda)} = \inf_{\lambda \in \Lambda} \frac{T^{-d} \sum_{i=2}^{T} \left(\Delta^{d_0} y^i_{t-1} (\lambda) \varepsilon_t \right)}{\left(T^{-2d} \sum_{i=1}^{T} \left(y^i_{t-1} (\lambda) \right)^2 \right)^{1/2} s(\lambda)}, \text{ for } i = \{A, \ B \text{ and } C\}, \qquad (A.1)$$

where $y^i_t = y_t - z^i_{tT}(\lambda)' \left(\sum_{s=1}^{T} z^i_{sT}(\lambda) z^i_{sT}(\lambda)' \right)^{-1} \sum_{s=1}^{T} z_{sT}(\lambda) y_s$,

$$\Delta^d y^i_t = \Delta^d y_t - z^i_{tT}(\lambda)' \left(\sum_{s=1}^{T} z^i_{sT}(\lambda) z^i_{sT}(\lambda)' \right)^{-1} \sum_{s=1}^{T} z_{sT}(\lambda) \Delta^d y_s \, for \, i = \{A, B, \text{ and } C\},$$

and $s^2(\lambda)$ is the usual estimator of the residual variance (see Z&A for its exact definition). Henceforth, only Model A will be considered where, for brevity, the superscript i is dropped. Proofs for the other models {B and C} are analogous and, therefore, are omitted.

The statistic in (A.1) can be rewritten as a functional g of X_T, $\tilde{z}_T$, $T^{1/2-d} \sum \tilde{z}_T \varepsilon_t$, σ^2 and s^2 plus an asymptotically negligible term, (see equations (A.1) and (A.2) in Z&A), where

$$X_T(r) = T^{1/2-d} \sigma^{-1} \sum_{i=0}^{[Tr]} \pi_i(-d) \varepsilon_{[Tr]-i}, \ \ (j-1) < r < (j+1) \text{ for } j = 1, ..., \ T.$$

By expression $(A.5)$ in Z&A,

$$
\begin{aligned}
T^{-2d} \sum_{i=2}^{T} \left(y^i_{t-1}(\omega) \right)^2 &= \int_0^1 \{ \sigma X_T(r) - \tilde{z}_T(\omega, r)' \left(\int_0^1 \tilde{z}_T(\omega, s)' \tilde{z}_T(\omega, s)' \, ds \right)^{-1} \\
&\quad \times \left(\int_0^1 \tilde{z}_T(\omega, s)' \sigma X_T(s)' \, ds \right) \}^2 dr + o_{p\lambda}(1) \\
&= H_1[\sigma X_T, \tilde{z}_T](\omega) + o_{p\lambda}(1),
\end{aligned}
$$

and by $(A.6)$ in Z&A,

$$T^{-d} \sum_{i=2}^{T} y^i_{t-1}(\lambda) \varepsilon_t = H_2[\sigma X_T, \tilde{z}_T, T^{1/2-d} \sum \tilde{z}_T \varepsilon_t](\lambda) + o_{p\lambda}(1),$$

$$T^{-2d_0} X_T(.) \xrightarrow{w} B_{d_0}(.),$$

where the symbol $o_{p\lambda}(1)$ denotes any random variable $\vartheta(\lambda)$ such that $\sup_{\lambda \in \Lambda} |\vartheta(\lambda)| \xrightarrow{p} 0$.

Since the limiting distribution of $\tilde{z}_T(.,.)$ is degenerate, it follows that $(X_T(.), \tilde{z}_T(.,.))$ converge weakly to $(B_d(.), z(.,.))$.[15]

Lemmas A.1-A.4 in $Z\&A$ guarantee that the processes X_T, $\tilde{z}_T$, $T^{1/2-d}\sum \tilde{z}_T\varepsilon_t$, σ^2 and s^2 jointly converge and that the functional g is continuous. The final result follows from the continuity of a composition of continuous functions and the CMT.

2) Case $0 < d < 0.5$

Let us start considering Case A: DGP: $\Delta^d y_t = \varepsilon_t$ and regression model,

$$\Delta^d y_t = \alpha DU_t(\lambda) + \phi y_{t-1} + \varepsilon_t,$$

with $\lambda \in [\lambda_0, \lambda_1] = \Lambda$, $0 < \lambda_0 < \lambda_1 < 1$, and ε_t are i.i.d $(0, 1)$ (we assume known variance for simplicity). Then, the t-ratio of $\widehat{\phi}_{ols}$ when the break date is at a fraction λ $(= T_B/T)$ of the sample size T, denoted in short as $t_{\widehat{\phi}(\lambda)}$ is given by,

$$
t_{\widehat{\phi}(\lambda)} = \left\{ \frac{\sum_1^T \varepsilon_t y_{t-1}}{T^{1/2}} - \frac{\sum_{r+1}^T \varepsilon_t}{T^{1/2}} \frac{\sum_{T_B}^T y_{t-1}}{T^{d+1/2}} \frac{1}{1-\lambda} \frac{1}{T^{1/2-d}} \right\} \times
$$
$$
\left\{ \frac{\sum_1^T y_{t-1}^2}{T} - \left(\frac{\sum_{T_B}^T y_{t-1}}{T^{d+1/2}} \right)^2 \frac{1}{1-\lambda} \cdot \frac{1}{T^{1-2d}} \right\}^{-1/2}.
$$

The following results in Hosking (1996) and DGM (2002) will be used:

(i) $T^{-1/2}\sum_2^T \varepsilon_t y_{t-1} \overset{w}{\to} N[0, \Gamma(1-2d) / \Gamma^2(1-d)]$,

(ii) $T^{-(d+1/2)}\sum_{T_B}^T y_{t-1} \overset{w}{\to} N[0, \Gamma(1-2d) / \Gamma(1-d)\Gamma(1+d)(1+2d)]$,

(iii)$T^{-1}\sum_2^T y_{t-1}^2 \overset{p}{\to} \Gamma(1-2d) / \Gamma^2(1-d)$.

We want to compare $t_{\widehat{\phi}(\lambda)}$ with,

$$
\tau = \frac{T^{-1/2}\sum_2^T \varepsilon_t y_{t-1}}{\left\{ T^{-1}\sum_2^T y_{t-1}^2 \right\}^{-1/2}},
$$

which, from (i) and (iii), converges in distribution to a standard normal. For this consider the function,

$$
f_T(x, z) = \frac{a_T - \frac{1}{1-\lambda}xz}{(b_T - x^2 \frac{1}{1-\lambda} \frac{1}{T^{2\delta_1}})^{1/2}},
$$

_____

[15]The uniform metric is used in the first term whereas the hybrid uniform/L^2 d^* metric is used in the second (see Z&A for further details).

where $a_T = T^{-1/2} \sum_2^T \varepsilon_t y_{t-1}$, $b_T = T^{-1} \sum_2^T y_{t-1}^2 \geq 0$, and $\delta_0 + \delta_1 = 1 - 2d > 0$ with $0 < \delta_0$, $\delta_1 < 1$. The domain is $-T^{-\delta_1}[b_T(1 - \lambda)]^{1/2} < x < T^{-\delta_1}[b_T(1 - \lambda)]^{1/2}$. A simple two-dimensional mean value expansion yields,

$$f_T(x, z) - f_T(0, 0) = f'_{xT}(\widetilde{x}, \widetilde{z}) \, x + f'_{zT}(\widetilde{x}, \widetilde{z}) \, z,$$

where $(\widetilde{x}, \widetilde{z}) = (\vartheta x, \vartheta z)$, $0 < \vartheta < 1$, and the partial derivatives are given by,

$$f'_{xT}(x, z) = \frac{\frac{a_T}{(1-\lambda)T^{2\delta_1}}x - \frac{b_T}{1-\lambda}z}{(b_T - x^2 \frac{1}{1-\lambda} \frac{1}{T^{2\delta_1}})^{3/2}},$$

$$f'_{zT}(x, z) = \frac{-\frac{1}{1-\lambda}x}{(b_T - x^2 \frac{1}{1-\lambda} \frac{1}{T^{2\delta_1}})^{1/2}}.$$

Note that,

$$t_{\widehat{\phi}(\lambda)} = f_T(T^{-\delta_0} \frac{\sum_1^{T_B} y_{t-1}}{T^{d+1/2}}, T^{-\delta_1} \frac{\sum_1^{T_B} \varepsilon_t}{T^{1/2}}),$$

and that $\tau = f_T(0, 0)$. Hence,

$$\sup_{\lambda \in \Lambda} \left| t_{\widehat{\phi}(\lambda)} - \tau \right| \leq \sup_{\lambda \in \Lambda} \left| f'_{xT}(\widetilde{x}, \widetilde{z}) \right| \left| T^{-\delta_0} \frac{\sum_{r+1}^T y_{t-1}}{T^{d+1/2}} \right| + \sup_{\lambda \in \Lambda} \left| f'_{zT}(x, z) \right| \left| T^{-\delta_1} \frac{\sum_{r+1}^T \varepsilon_t}{T^{1/2}} \right|.$$

Now, $T^{-(d+1/2)} \sum_1^{[T\lambda]} y_{t-1}$ converges to fBM while $T^{-1/2} \sum_1^{[T\lambda]} \varepsilon_t$ converges to BM. Therefore, $\sup_{\lambda \in \Lambda} \left| T^{-\delta_0} \frac{\sum_{T_B}^T y_{t-1}}{T^{d+1/2}} \right|$ and $\sup_{\lambda \in \Lambda} \left| T^{-\delta_1} \frac{\sum_{T_B}^T \varepsilon_t}{T^{1/2}} \right|$ are $o_p(1)$. Observe further that,

$$
\begin{aligned}
\sup_{\lambda \in \Lambda} \left| f'_{xT}(\widetilde{x}, \widetilde{z}) \right| &= \sup_{\lambda \in \Lambda} \left| \frac{\frac{a_T}{(1-\lambda)T^{2\delta_1}}x - \frac{b_T}{1-\lambda}z}{(b_T - x^2 \frac{1}{1-\lambda} \frac{1}{T^{2\delta_1}})^{3/2}} \right| \quad &\text{(A.2)} \\[2mm]
&\leq \frac{\sup_{\lambda \in \Lambda} \left| \frac{a_T}{(1-\lambda)T^{2\delta_1}} \right| |x| + \left| \frac{b_T}{1-\lambda} \right| |z|}{\inf_{\lambda \in \Lambda} \left| (b_T - x^2 \frac{1}{1-\lambda} \frac{1}{T^{2\delta_1}})^{3/2} \right|} \\[2mm]
&\leq \frac{\left| \frac{a_T}{(1-\lambda_1)T^{2\delta_1}} \right| \sup_{\lambda \in \Lambda} |x| + \left| \frac{b_T}{1-\lambda_1} \right| \sup_{\lambda \in \Lambda} |z|}{\left| (b_T - \frac{1}{1-\lambda_1} \frac{1}{T^{2\delta_1}} \sup_{\lambda \in \Lambda} x^2)^{3/2} \right|}.
\end{aligned}
$$

From (ii) and (iii), note that $b_T \xrightarrow{p} \Gamma(1 - 2d)/\Gamma^2(1 - d)$, and that,

$$\sup_{\lambda \in \Lambda} T^{-\delta_0} \left(\frac{\sum_1^{T_B} y_{t-1}}{T^{d+1/2}} \right)^2 \xrightarrow{p} 0, \sup_{\lambda \in \Lambda} T^{-\delta_1} \left(\frac{\sum_1^{T_B} \varepsilon_t}{T^{1/2}} \right) \xrightarrow{p} 0. \tag{A.3}$$

39

Hence the denominator in (A.2) is bounded away from zero with probability approaching one. In view of (A.3), the numerator in (A.2) is bounded in probability. Now, both results imply,

$$\sup_{\lambda \in \Lambda} \left| t_{\widehat{\phi}(\lambda)} - \tau \right| \xrightarrow{p} 0.$$

Consequently, $\inf_{\lambda \in \Lambda} t_{\widehat{\phi}(\lambda)}$ has the same asymptotic distribution as $\inf_{\lambda \in \Lambda} \tau$ which of course is the distribution of τ.

The proofs for Cases B-C follow along similar lines by considering the following regression model,

$$\Delta^d y_t = \alpha DT_t^*(\lambda) + \phi y_{t-1} + \varepsilon_t,$$

where $DT_t^*(\lambda) = (t - T_B)\mathbf{1}_{(T_B+1 \leq t \leq T)}$ and the rest of assumptions are as above. Then, the t-ratio of $\widehat{\phi}_{ols}$ when the break date is at a fraction $\lambda \ (= T_B/T)$ of the sample size T, is given by,

$$
t_{\widehat{\phi}(\lambda)} = \left\{ \frac{\sum_1^T \varepsilon_t y_{t-1}}{T^{1/2}} - \frac{\sum_{T_B}^T t^* \varepsilon_t}{T^{3/2}} \frac{\sum_{T_B}^T t^* y_{t-1}}{T^{d+3/2}} \frac{3}{(1-\lambda)^3} \frac{1}{T^{1/2-d}} \right\}.
$$
$$
\left\{ \frac{\sum_1^T y_{t-1}^2}{T} - \left(\frac{\sum_{T_B}^T t^* y_{t-1}}{T^{d+3/2}} \right)^2 \frac{3}{(1-\lambda)^3} \cdot \frac{1}{T^{1-2d}} \right\}^{-1/2},
$$

where $t^* = t - T_B$. Use of the result $T^{-(d+3/2)} \sum_{T_B}^T t^* y_{t-1} = O_p(1)$ (non-degenerate) has been made (see Marmol and Velasco, 2002). Then, considering the function,

$$f_T(x, z) = \frac{a_T - \frac{3}{(1-\lambda)^3} xz}{(b_T - x^2 \frac{3}{(1-\lambda)^3} \frac{1}{T^{2\delta_1}})^{1/2}},$$

where a_T and b_T are defined as above, the proof is identical to the one in the previous case. ∎

APPENDIX B

Tables B1 and B2 gather the critical values of the distributions of FDF tests for testing $I(d)$ vs. $I(0)$ + deterministic components presented in section 2.1, whereas Tables B3- B5 present the corresponding critical values for the SB-FDF tests presented in section 3, for testing $I(d)$ vs. $I(0)$ + structural breaks with $\lambda \in [0.15, 0.85]$.

TABLE B1

CRITICAL VALUES FOR THE FDF TEST OF I(d) VS. I(0)+

DETERMINISTIC COMPONENTS BASED ON ESTIMATED $t^{\mu}_{\phi_{ols}}$ IN (6)

Sample Size	$T = 100$			$T = 400$			$T = 1000$		
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
0.6	-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 B2

CRITICAL VALUES FOR THE FDF TEST OF I(d) VS. I(0)+

DETERMINISTIC COMPONENTS BASED ON ESTIMATED $t^{\tau}_{\phi_{ols}}$ IN (7)

Sample Size	$T = 100$			$T = 400$			$T = 1000$		
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	-1.229	-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
0.6	-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 B3

CRITICAL VALUES THE SB-FDF TEST OF I(d) VS. I(0)+ BREAKS

BASED ON ESTIMATED POINTS OF THE $inf\ t^{A}_{\hat{\phi}(\lambda)}$ IN (16)

Sample Size	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
0.6	-3.236	-3.532	-4.161	-2.999	-3.342	-4.009	-2.545	-2.918	-3.219
0.7	-3.519	-3.847	-4.484	-3.331	-3.634	-4.221	-2.911	-3.241	-3.538
0.8	-3.761	-4.069	-4.692	-3.602	-3.875	-4.437	-3.325	-3.561	-3.861
0.9	-3.978	-4.266	-4.852	-3.870	-4.137	-4.613	-3.638	-3.784	-4.043

TABLE B4

CRITICAL VALUES THE SB-FDF TEST OF I(d) VS. I(0) + BREAKS

BASED ON ESTIMATED POINTS OF THE $inf\ t^B_{\widehat{\phi}(\lambda)}$ IN (16)

Sample Size	T=100			T=400			T=1000		
d_0/ s.l.	90%	95%	99%	90%	95%	99%	90%	95%	99%
0.1	-2.251	-2.601	-3.269	-1.833	-2.201	-2.901	-1.664	-2.044	-2.769
0.2	-2.447	-2.792	-3.463	-2.055	-2.417	-3.044	-1.846	-2.198	-2.864
0.3	-2.648	-3.003	-3.657	-2.267	-2.614	-3.266	-2.116	-2.455	-3.103
0.4	-2.929	-3.256	-3.913	-2.574	-2.918	-3.628	-2.402	-2.739	-3.393
0.6	-3.556	-3.853	-4.514	-3.331	-3.649	-4.300	-3.22d	-3.534	-4.131
0.7	-3.937	-4.249	-4.803	-3.728	-4.026	-4.642	-3.652	-3.959	-4.536
0.8	-4.252	-4.544	-5.191	-4.086	-4.390	-4.904	-4.087	-4.353	-4.923
0.9	-4.587	-4.882	-5.474	-4.458	-4.707	-5.213	-4.442	-4.507	-5.201

TABLE B5

CRITICAL VALUES THE SB-FDF TEST OF I(d) VS. I(0) + BREAKS

BASED ON ESTIMATED POINTS OF THE $inf\ t^C_{\widehat{\phi}(\lambda)}$ IN (16)

Sample Size	T=100			T=400			T=1000		
d_0 /s.l.	90%	95%	99%	90%	95%	99%	90%	95%	99%
0.1	-2.449	-2.810	-3.448	-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
0.6	-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