WALD TESTS OF $I(1)$ AGAINST $I(d)$ ALTERNATIVES: SOME NEW PROPERTIES AND AN EXTENSION TO PROCESSES WITH TRENDING COMPONENTS

By Juan J. Dolado[a], Jesus Gonzalo[a], and Laura Mayoral[b] *
[a]Dept. of Economics, Universidad Carlos III de Madrid.
[b]Institut d ´Análisi Económica (CSIC).

June, 25, 2007

Abstract

This paper analyses the power properties, under fixed alternatives, of a Wald-type test, i.e., the (Efficient) Fractional Dickey-Fuller (EFDF) test of $I(1)$ against $I(d)$, $d < 1$, relative to LM tests. Further, it extends the implementation of the EFDF test to the presence of deterministic trending components in the DGP. Tests of these hypotheses are important in many macroeconomic applications where it is crucial to distinguish between permanent and transitory shocks because shocks die out in $I(d)$ processes with $d < 1$. We show how simple is the implementation of the EFDF in these situations and argue that, under fixed alternatives, it has better power properties than LM tests. Finally, an empirical application is provided where the EFDF approach allowing for deterministic components is used to test for long-memory in the GDP p.c. of several OECD countries, an issue that has important consequences to discriminate between alternative growth theories.

JEL Clasification: C12 C22 O40

Keywords: Deterministic components, Fractional processes, Power, Unit roots.

*Corresponding E-mail:jesus.gonzalo@uc3m.es. We are grateful to Claudio Michelacci and Bart Verspagen for making the data available to us, and to Javier Hidalgo, Javier Hualde, Francesc Mármol, Peter Robinson, Carlos Velasco and participants in seminars at CREST (Paris), Ente Luigi Einaudi (Rome), ECARES (Brussels), and World Congress of the Econometric Society (2005) for useful comments on preliminary drafts of this paper. Financial support from the Spanish Ministry of Education through grants SEC2003-04429 and SEC2003-04476 and also from the Barcelona Economics Program of CREA is acknowledged. The usual disclaimer applies.

1. INTRODUCTION

It is well known that tests of $I(1)$ vs. $I(0)$ processes reject very occasionally the null that a time series $\{y_t\}$ is $I(1)$ when the true DGP is a fractionally integrated, $I(d)$, process, specially if $0.5 < d < 1$. This issue can have serious consequences for studies focusing on the medium and long- run properties of the variables of interest. To mention only two: (i) shocks could be identified as permanent when in fact they die out eventually, and (ii) two series could be considered as spuriously cointegrated (i.e., a concept introduced and analyzed in Gonzalo and Lee, 1998) when they are independent at all leads and lags. Further, these mistakes are more likely to occur in the presence of deterministic components as, for example, in the case of trending economic variables.

In view of this problem, the goal of this paper is twofold. First, we discuss the power behavior of a recently proposed Wald test of $I(1)$ vs. $I(d)$, $d \in [0, 1)$ relative to the one achieved by well-known LM tests. In particular, we derive new analytical results regarding the non-centrality parameters of LM and Wald tests under *fixed* alternatives and show that the latter tends to be more powerful than the former even though they are asymptotically equivalent under *local* alternatives. Secondly, we extend this Wald-type testing procedure, originally derived for processes without deterministic components, to the more realistic case where they are present.

Specifically, we focus on a modification of the Fractional Dickey-Fuller (FDF) test by Dolado, Gonzalo and Mayoral (2002; DGM hereafter) recently proposed by Lobato and Velasco (2007a; LV hereafter) which achieves an improvement in efficiency over the former. This test, henceforth denoted as the EFDF (efficient FDF) test, generalizes the traditional DF test of $I(1)$ against $I(0)$ processes without deterministic components to the broader framework of testing $I(1)$ against $I(d)$ with $0 \leq d < 1$. The FDF and EFDF tests belong to the family of Wald tests and rely upon the principle underlying the popular Dickey-Fuller (DF) approach. The idea is to test for the statistical significance of the slope coefficient, φ, in a (possibly) unbalanced regression where the dependent variable and the regressor are filtered in such a way that the resulting processes are $I(0)$ under the null and the alternative hypothesis, respectively. [1] Both DGM and LV set Δy_t as the dependent variable, where $\Delta = (1 - L)$.[2] As regards the regressor, whereas DGM

[1] In the DF setup, these filters are $\Delta = (1 - L)$ and $\Delta^0 = 1$, so that the regressand and regressor are Δy_t and y_{t-1}, respectively.

[2] As shown in DGM (2002), both regressors can be constructed by applying the truncated version of the

choose $\Delta^d y_{t-1}$, LV show that $z_{t-1}(d) = (1-d)^{-1}(\Delta^{d-1} - 1)\Delta y_t$ improves the efficiency of the test. These tests belong to the Wald family because the coefficient φ is linearly related to the parameter of interest, d. Like in the FDF procedure, the EFDF test is based upon the t-ratio, t_φ, of the OLS estimate of φ. Thus, non-rejection of H_0: $\varphi = 0$ against H_1: $\varphi < 0$, implies that the process is $I(1)$ and, conversely, rejection of the null implies that the process is $I(d)$.

In order to compute either $\Delta^d y_{t-1}$ or $z_{t-1}(d)$, an input value for d is required. One could either consider a (known) simple alternative, $H_A : d = d_A < 1$ or, more realistically, a composite one, $H_1 : d < 1$. In the latter case, which is the one we focus on here, both DGM and LV show that it suffices to use a T^κ-consistent estimate (with $\kappa > 0$) of the true integration order, and show that the limiting distribution of the resulting statistic is a $N(0,1)$.

Under a sequence of local alternatives approaching $H_0 : d = 1$ from below at a rate of $T^{-1/2}$, LV (2007a, Theorem 1) prove that, with Gaussianity, the EFDF test is asymptotically equivalent to the uniformly most powerful invariant (UMPI) test, i.e., the LM test introduced by Robinson (1991, 1994) and later adapted by Tanaka (1999) for the time domain. Our first contribution in this paper is to analyze the properties of Wald and LM tests in the case where the alternative is fixed. Our findings here point out that, in line with the standard result about the better power properties of Wald tests relative to LM tests (see Engle, 1984), the former fare better than the latter in this dimension. Moreover, when compared to other tests of $I(1)$ vs. $I(d)$ which rely on direct inference about semiparametric estimators of d, the EFDF test exhibits in general better power properties, under a correct specification of the stationary short-run dynamics of the error term in the auxiliary regression. This is due to the fact that the semiparametric estimation procedures often imply larger confidence intervals of the memory parameter, in exchange with less restrictive assumptions on the error term.[3] By contrast, the combination of a wide range of semiparametric estimators for the input value of d with an auxiliary parametric regression, as the one discussed above, yields a parametric rate for the Wald tests.[4] Thus, in a sense, the Wald tests combine the favorable features of both approaches to improve power while reducing

binomial expansion of the filter $(1-L)^d$ in the lag operator L to y_t $(t = 0, 1, ...)$, so that $\Delta_+^d y_t = \sum_{i=0}^{t-1} \pi_i(d)$ y_{t-i}, where $\pi_i(d)$ is the i-th coefficient in that expansion, defined at the end of this Introduction. In the sequel, we will refer to this truncated filter simply as Δ^d.

[3]See, e.g., Velasco (1999), Robinson (2003), Abadir et al. (2005), Shimotsu and Phillips (2005) and Shimotsu (2006).

[4]LV (2007b) have shown that a Gaussian semiparametric estimator, such as the one proposed by Velasco (1999) suffices to achieve consistency and asymptotic normality in the analyzed Wald tests (see sections 2 and 3 below).

the danger of misspecifying short-run dynamics.

Following the development of the unit root tests in the past, where the canonical zero-mean AR(1) model was subsequently augmented with deterministic components (including drifts, and linear, nonlinear and broken trends), our second contribution in this paper is to investigate how to implement this Wald test when some deterministic components are considered in the DGP, a case which was neither considered by LV nor by DGM. Although we will analyze other types of trends, we will mainly focus on the role of a *linear trend* since many (macro) economic time series exhibit this type of trending behavior in their levels. Our main result is that, in contrast with what happens with most tests for $I(1)$ against $I(0)$, the EFDF test remains being efficient in the presence of deterministic components and it maintains the same asymptotic distribution, insofar as they are correctly filtered. In this respect, this result mimics the one found for LM tests when deterministic components are present; cf. Robinson (1994), Tanaka (1999) and Gil-Alaña and Robinson (1997).

Lastly, we wish to stress that, despite focusing the previous discussion on the case where the error term in the DGP is $i.i.d$, the asymptotic results obtained here remain valid when the disturbance is allowed to be autocorrelated, as it happens in the (augmented) DF case (ADF henceforth). In this respect, DGM (2002, Theorems 6 and 7) have proved that, in order to remove the autocorrelation, it is sufficient to augment the set of regressors in the auxiliary regression of the FDF test with k lags of the dependent variable such that $k \uparrow \infty$ as $T \uparrow \infty$, and $k^3/T \uparrow 0$, as in Said and Dickey (1984). This leads to the augmented FDF (AFDF) test. As regards the EFDF test, we conjecture that a similar result holds, although we will confine our discussion below, as in LV (2007a), to the case of finite-lag autoregressive processes. The procedure based on the EFDF test turns out to be much simpler than accounting for serial correlation in the LM framework. Further, we point out that the two-step procedure proposed by LV (2007a) can be simplified to a feasible linear single-step estimation approach. An empirical application dealing with testing the possibility that long GNP *per capita* series for several OECD countries may follow nonstationary $I(d)$ processes, yet with shocks that die out (supporting the hypothesis of beta-convergence) instead of $I(1)$ (no convergence), illustrates our proposed methodology.

The rest of the paper is structured as follows. Sections 2 briefly overviews the properties of the EFDF test when the process is a driftless random walk under the null and derives new results about the power of this test relative to the LM test under fixed alternatives. Section 3 extends the previous results to the case where the process contains trending deterministic components (e.g., a linear trend), considering both the case of i.i.d. and autocorrelated errors. Section 4 discusses an empirical application of the previous test. Finally, Section 5 draws some concluding remarks.

Proofs of the theorems are collected in the Appendix.

In the sequel, the definition of a $I(d)$ process that we will adopt is that of a non-stationary (truncated) process when $1 > d > 0.5$. Those definitions are similar to those used in, e.g., Robinson (1994) and Tanaka (1999) and are summarized in Appendix A of DGM. Moreover, the following conventional notation is adopted throughout the paper: $\Gamma(.)$ denotes the gamma function, $\{\pi_i(d)\}$ represents the sequence of coefficients associated to the expansion of Δ^d in powers of L , defined as

$$\pi_i(d) = \frac{\Gamma(i-d)}{\Gamma(-d)\,\Gamma(i+1)}.$$

The indicator function is denoted by $1_{(.)}$. Finally, $\xrightarrow{w}$ denotes weak convergence in $D[0,1]$ endowed with the Skorohod J_1 topology, and $\xrightarrow{p}$ means convergence in probability.

2. THE EFDF TEST

2.1 Definitions

Like in Robinson (1994), we consider a process y_t that is generated by an additive model, namely as the sum of a deterministic component, $\mu(t)$, and an $I(d)$ component, u_t, so that

$$y_t = \mu(t) + u_t, \tag{1}$$

where $u_t = \Delta^{-d}\varepsilon_t 1_{t>0}$ is a purely stochastic $I(d)$ process and ε_t is a zero-mean i.i.d. random variable.

For the case where $\mu(t) \equiv 0$,[5] DGM introduced a Wald-type (FDF) test for testing the null hypothesis of $H_0 : d = 1$ against the composite alternative $H_1 : 0 \leq d < 1$, based on the t-statistic associated to the hypothesis $\phi = 0$ in the OLS regression

$$\Delta y_t = \phi \Delta^{\widehat{d}_T} y_{t-1} + \upsilon_t. \tag{2}$$

If the value d is chosen under a composite H_1 using a $\sqrt{T}-$ consistent estimator of d, $\widehat{d}_T$, DGM prove that the asymptotic distribution of the resulting t-statistic, t_ϕ is $N(0,1)$. They

[5] Alternatively, $\mu(t)$ could be considered to be known. In this case, the same arguments go through after substracting it from y_t to obtain a purely stochastic process.

provide Monte Carlo evidence showing that, despite not being locally optimal (as Robinson's LM test is), its finite sample performance is more satisfactory than the LM test, except when local alternatives with gaussian errors are considered. In this case, as will be discussed below, the asymptotic relative efficiency of the FDF test is 0.8 (see LV, 2007b).

Recently, LV (2007a) have proposed the EFDF test based on a modification of regression (2) that permits to achieve higher efficiency while keeping the good finite-sample properties of Wald tests under the assumption of $\mu(t) \equiv 0$ (or known). More specifically, their proposal is to compute the t-statistic, t_φ, associated to the null hypothesis $\varphi = 0$ in the regression

$$\Delta y_t = \varphi z_{t-1}(d) + \varepsilon_t, \tag{3}$$

where $z_{t-1}(d)$ is defined as[6]

$$z_{t-1}(d) = \frac{\left(\Delta^{d-1} - 1\right)}{(1-d)} \Delta y_t.$$

Note that, when $\varphi = 0$, the model becomes a random walk, $\Delta y_t = \varepsilon_t$, while, under the alternative, $\varphi = (d-1) < 0$, it becomes a pure fractional process, $\Delta^d y_t = \varepsilon_t$. The reason why the EFDF test is more efficient than the FDF test is that, under H_1, the regression model considered in (2) with known d can be written as $\Delta y_t = \Delta^{1-d}\varepsilon_t = \varepsilon_t + (d-1)\varepsilon_{t-1} + 0.5d(d-1)\varepsilon_{t-2} + ... = \phi\Delta^d y_{t-1} + \varepsilon_t + 0.5d(d-1)\varepsilon_{t-2} + ...$ with $\phi = d-1$. Thus, the error term υ_t in (2) is autocorrelated. Although OLS provides a consistent estimator of ϕ, since the regressor $\Delta y_{t-1} = \varepsilon_{t-1}$ is not in υ_t, it is not the most efficient one. By contrast, the regression model used in the EFDF test does not suffer from this problem since, by construction, yields an i.i.d. error term.

Theorem 1 in LV (2007a), which we reproduce below for completeness, establishes the asymptotic properties of t_φ.

Theorem 1 *Under the assumption that the DGP is given by $y_t = \Delta^{-d}\epsilon_t 1_{(t>0)}$, where $d \leq 1$, ε_t is i.i.d. with finite fourth moment, the asymptotic properties of the t-statistic t_ϕ for testing $\varphi = 0$ in (3), where the input of $z_{t-1}(\widehat{d_T})$ is a $T^\kappa-$consistent estimator of d, for some $d > 0.5$ with $\kappa > 0$, are given by,*

a) Under the null hypothesis ($d = 1$),

$$t_\varphi(\widehat{d_T}) \xrightarrow{w} N(0,\ 1).$$

[6] A similar model was first proposed by Granger (1986) in the more general context of testing for cointegration with multivariate series, a modification of which has been recently considered by Johansen (2005).

b) Under local alternatives, $(d = 1 - \gamma/\sqrt{T})$,

$$t_{\varphi}(\widehat{d}_T) \xrightarrow{w} N\left(-\gamma h\left(d\right),\ 1\right),$$

where $h(\varrho) = \Sigma_{j=1}^{\infty} j^{-1}\pi_j(\varrho - 1)/\sqrt{\Sigma_{j=1}^{\infty}\pi_j(\varrho-1)^2}$, $0.5 < \varrho < 1$.

c) Under fixed alternatives $(d < 1)$, *the test based on* $t_{\varphi}(\widehat{d}_T)$ *is consistent.*

LV (2007a) show that the function $h(.)$ achieves a global maximum at $d = 1$ where $h(1) = \sqrt{\pi^2/6}$, and that $h\left(1\right)$ equals the noncentrality parameter of the locally optimal Robinson's LM test (see subsection 2.2 below). Moreover, if a T^{κ}-consistent estimator of d is used as input of $z_{t-1}(d)$ with $\kappa > 0$, the asymptotic distribution of the EFDF test under the null is invariant to the use of an estimated input $\widehat{d}_T$ (see, Robinson and Hualde, 2003, Assumption 3). A power-rate consistent estimate of d can be easily obtained by applying any parametric $\sqrt{T}$-consistent estimator of this quantity (such as Beran, 1995, Velasco and Robinson, 2000 or Mayoral, 2007) but also, much less restrictively, by using some semiparametric estimators of d, as LV (2007a, b) have shown. Among the latter class, the estimators proposed by Abadir et al. (2005), Shimotsu (2006) and Velasco (1999) represent good choices since they also allow for the existence of the deterministic components considered in section 3.

2.2 Power comparisons under fixed alternatives

As discussed in the Introduction, the closer competitor to the Wald (FDF and EFDF) tests is the LM test proposed by Robinson (1991, 1994) in the frequency domain, subsequently extended by Tanaka (1999) to the time domain. In this section we discuss the power properties of the three competing tests under the case of fixed alternatives.[7]

We start with the LM test, denoted as LM_T, which considers the null hypothesis of $\theta = 0$ against the alternative $\theta \neq 0$ for the DGP $\Delta^{d_0+\theta}y_t = \varepsilon_t$. In line with the hypotheses considered in this paper, we will focus on the particular case where $d_0 = 1$ and $-1 \leq \theta < 0$. Assuming that $\varepsilon_t \sim N(0, \sigma^2)$, the score-LM test is computed as

$$LM_T = \sqrt{\frac{6}{\pi^2}}T^{1/2}\sum_{j=1}^{T-1} j^{-1}\widehat{\rho}_j \xrightarrow{w} N\left(0,1\right), \tag{4}$$

where $\widehat{\rho}_j = \sum_{t=j+1}^{T}\Delta y_t\ \Delta y_{t-j}/\sum_{t=1}^{T}(\Delta y_{t-j})^2$ (see Robinson, 1991 and Tanaka, 1999).

[7] The available results in the literature only establish the consistency of the Wald and LM test under fixed alternatives. Yet, they do not derive the non-centrality parameters as we do below.

Breitung and Hassler (2002) have shown that an alternative simpler way to compute the score-LM test is as the t-ratio (t_λ) of $\widehat{\lambda}_{ols}$ in the regression

$$\Delta y_t = \lambda x^*_{t-1} + e_t, \tag{5}$$

where $x^*_{t-1} = \sum_{j=1}^{t-1} j^{-1}\Delta y_{t-j}$. Intuitively, since $t_\lambda = \sum(\Delta y_t x^*_{t-1})/\widehat{\sigma}_e(\sum(x^*_{t-1})^2)^{1/2}$ and, under $H_0 : \theta = 0$, $\widehat{\sigma}_e$ tends to σ and plim $T^{-1}\sum(x^*_{t-1})^2 = \pi^2/6$, then t_λ has the same limiting distribution as LM_T.

As mentioned earlier, under a sequence of local alternatives of the type $\theta = \gamma T^{-1/2}$ with $\gamma > 0$, the LM_T (or t_λ) is the UMPI test. Under a sequence of local alternatives, LV (2007a) have shown that the EFDF is asymptotically equivalent to the UMPI. This is so because when d tends to 1, the indetermination 0/0 in the filter $\left(\Delta^{d_1-1} - 1\right)/(1-d_1)\rfloor_{d_1=1}$ used in the EFDF test is easily solved by L' Hôpital rule yielding the same linear filter as in the LM test, namely $-ln(1-L) = \Sigma_{j=1}^{\infty} j^{-1}L^j$, so that the test becomes asymptotically equivalent to the LM_T and t_λ tests. Hence, as stated in Theorem 1 above, when $\mu(t) \equiv 0$ (or known) and $d = 1 - \gamma T^{-1/2}$, the limiting distribution of the the EFDF test is identical to that of the LM test, i.e., $N(-\gamma h(d), 1)$ where $h(.)$ is $\pi/\sqrt{6}$ for $d = 1$. DGM (2002, Theorem 3) in turn obtained that the corresponding distribution of the FDF under local alternatives test is $N(-\gamma, 1)$. Hence, the asymptotic efficiency of the FDF test relative to the LM and EFDF tests is 0.78 ($\simeq \sqrt{6}/\pi$).

In the rest of this section, we analyze the case with fixed alternatives where, to our knowledge, results are new. In particular, we derive the non-centrality parameters of the three above-mentioned tests under an $I(d)$ alternative where the DGP is assumed to be $\Delta^d y_t = \varepsilon_t$ with $d \in (0, 1)$. Hence, $\Delta y_t = \Delta^{-b}\varepsilon_t$ where $b = d - 1 < 0$. Then, the following result holds.

Theorem 2 *If $\Delta^d y_t = \varepsilon_t$ with $d \in [0, 1)$, the t-statistics associated to the EFDF and FDF tests, denoted as t_φ and t_ϕ, respectively, verify,*

$$T^{-1/2} t_\varphi \xrightarrow{p} -\left(\frac{\Gamma(3-2d)}{\Gamma^2(2-d)} - 1\right)^{1/2} \equiv c_{EFDF}(d),$$

$$T^{-1/2} t_\phi \xrightarrow{p} -\frac{(1-d)\Gamma(2-d)}{[\Gamma(3-2d) - (d-1)^2\Gamma^2(2-d)]^{1/2}} \equiv c_{FDF}(d),$$

while, under the same DGP, the LM test defined in (4) *satisfies that,*

$$T^{-1/2} LM_T \xrightarrow{p} -\sqrt{\frac{6}{\pi^2}} \frac{\Gamma(2-d)}{(1-d)\Gamma(d-2)} \sum_{j=1}^{\infty} \frac{\Gamma(j+d-1)}{j\Gamma(j+2-d)} \equiv c_{LM}(d),$$

where c_{EFDF} (d), c_{FDF} (d) and $c_{LM}(d)$ denote the non-centrality parameter under the fixed alternative $d \in (0,1)$ of the EFDF, FDF and LM tests, respectively.

Figure 1 displays the three non-centrality parameters for $d \in (0,1)$. As expected, the EFDF and the LM tests behave similarly for values of d very close to the null hypothesis, whereas the FDF test is slightly less powerful for these local alternatives. Nevertheless, despite being devised to be the UIMP test for local alternatives, the LM test performs significantly worse than both Wald-type tests when the alternative is not local. The EFDF tests performs slightly better than the FDF test in line with LV's (2007a) arguments about efficiency. The intuition for the worse power performance of the LM test is that there does not exist any value for λ in (5) that makes e_t both i.i.d. and independent of the regressor for fixed alternatives, implying that x^*_{t-1} does not maximize the correlation with Δy_t. Although the results in Theorem 2 are asymptotic, for realistic sample sizes the rejection rates of the Wald tests under the alternative are also larger than those of the LM test, except when d is very close to unity and the error term is normally distributed, in which case the EFDF and LM tests behave similarly. Thus, for fixed alternatives with, approximately, $d < 0.9$, the EFDF (and maybe the FDF) test is bound to exhibit higher power than the LM test.[8]

As regards semiparametric estimators, both the Fully Extended Local Whittle (FELW, see Abadir et al., 2005) and the Exact Local Whittle estimators (ELW, see Shimotsu, 2006) verify the asymptotic property $\sqrt{m}(\widehat{d}_T - d) \xrightarrow{w} N\left(0, \frac{1}{4}\right)$ for $m = o(T^{\frac{4}{5}})$. Test statistics for unit roots are based on $\tau_d = 2\sqrt{m}(\widehat{d}_T - 1) \xrightarrow{w} N(0,1)$. Therefore, their rate of divergence under $H_1 : d < 1$ is the nonparametric rate $O_p(\sqrt{m})$ which is smaller than the $O_p(\sqrt{T})$ parametric rate achieved by the Wald test.

[8]Note, however, that in the case of the FDF test, despite having a more negative non-centrality parameter than the LM test, its smaller efficiency (see the discussion after (5)) manifested by a larger variance in the denominator of the t-ratio under H_1, could imply lower power in some cases.

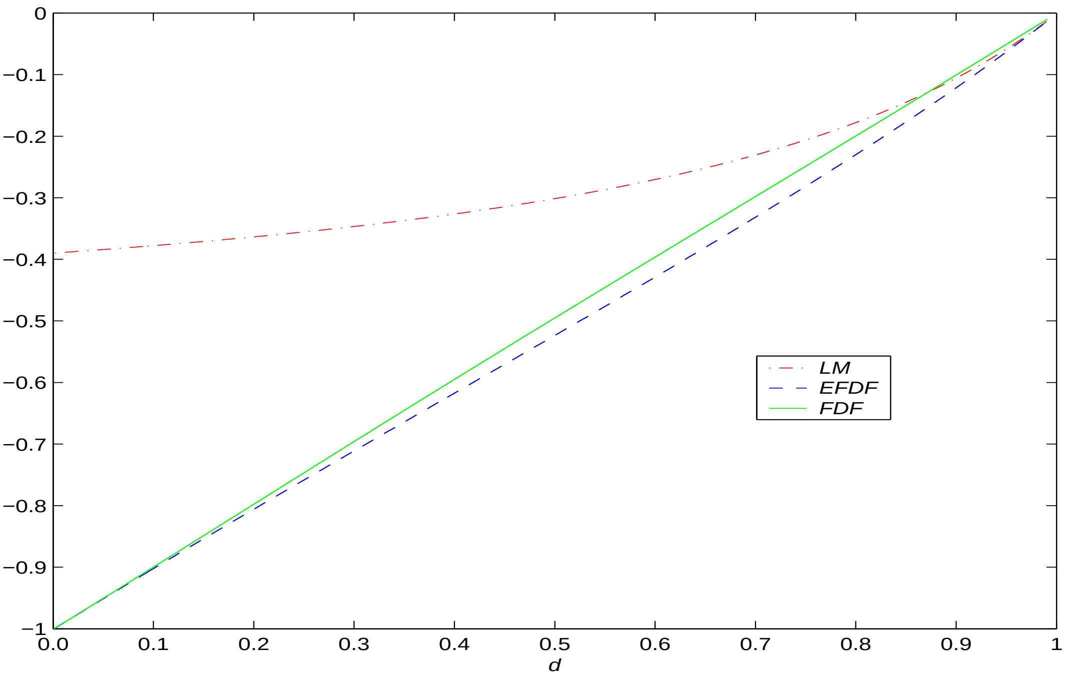

FIG 1. Non-centrality parameters of LM and Wald tests

3. THE EFDF TEST FOR TRENDING I(D) PROCESSES

3.1 i.i.d. case

In this section, we extend the EFDF testing approach to the more realistic case where $\mu(t) \neq 0$ and unknown. Our goal is to examine how this (unknown) deterministic term should be taken into account when implementing the test.

Following Elliott et al. (1996), we consider two different types of $\mu(t)$.

Slowly Evolving Deterministic component

Condition A. (Slowly evolving trend). The deterministic component $\mu(t)$ verifies

$$\mu(t) = O(t^{\delta}),\ \ \delta < 0.5.$$

Condition A is immediately satisfied if $\mu(t)$ is a constant term but also holds for a variety of time functions, such as slowly increasing trends, (e.g., t^{δ}, $\delta < 0.5$ or $\log t$).

In this case, it is straightforward to show that the stochastic component in y_t dominates the deterministic term when T is large. Hence, $\mu(t)$ has no effect either on the asymptotic distribution of the t-ratio statistic or on the efficiency properties of the test in the absence of $\mu(t)$. Therefore, one can proceed to run regression (3) ignoring the presence of these slowly evolving trends.

The following theorem presents the properties of the EFDF test when the DGP is given by (1) and $\mu(t)$ verifies Condition A.

Theorem 3 *(Slowly evolving trends) Under the assumption that the DGP is given by* $y_t = \mu(t) + \Delta^{-d}\epsilon_t 1_{(t>0)}$, *where* $d \leq 1$, ϵ_t *is i.i.d. with finite fourth moment, and* $\mu(t)$ *verifies Condition A, the asymptotic properties of the t-statistic* t_φ *for testing* $\varphi = 0$ *in (*3*) (denoted by EFDF*$_\mu$ *test), where the input of* $z_{t-1}(\widehat{d_T})$ *is a* $T^\kappa-$*consistent estimator of* d, *for some* $d > 0.5$ *with* $\kappa > 0$, *are identical to those stated in Theorem 1.*

Evolving Deterministic Components

Condition B. (Evolving trend). $\mu(t)$ *is a polynomial in* t *of known order.*

Under Condition B, the DGP is allowed to contain trending regressors in the form of polynomials (of known order) of t. Hence, when the coefficients of $\mu(t)$ are unknown, the test described above are unfeasible. Nevertheless, it is still possible to obtain a feasible test with the same asymptotic properties as in Theorem 1 if a consistent estimate of $\mu(t)$ is removed from the original process. Indeed, under H_0, the relevant coefficients of $\mu(t)$ can be consistently estimated by OLS in a regression of Δy_t on $\Delta\mu(t)$. For instance, consider the case where the DGP contains a linear time trend, that is,

$$y_t = \alpha + \beta t + \Delta^{-d}\epsilon_t, \tag{6}$$

which, under $H_0 : d = 1$, corresponds to the popular random walk with drift case. Taking first differences, it follows that $\Delta y_t = \beta + \Delta^{1-d}\varepsilon_t$. The OLS estimate of β, $\hat{\beta}$, (i.e., the sample mean of Δy_t) is consistent under both H_0 and H_1. In effect, under H_0, $\hat{\beta}$ is a $T^{1/2}$ -consistent estimator of β whereas, under H_1, it is $T^{3/2-d}$-consistent with $3/2 - d > 0.5$ (see Hosking 1996, Theorem 8). Hence, the following theory holds.

Theorem 4 *(Evolving trends) Under the assumption that the DGP is given by* $y_t = \mu(t) + \Delta^{-d}\epsilon_t 1_{(t>0)}$, *where* $d \leq 1$, ϵ_t *is i.i.d. with finite fourth moment, and* $\mu(t)$ *satisfies Condition B, the asymptotic properties of the t- statistic* t_φ *for testing* $\varphi = 0$ *in the regression*

$$\widetilde{\Delta y_t} = \varphi \widetilde{z_{t-1}}\left(\hat{d}_T\right) + e_t \tag{7}$$

(denoted by EFDF$_\tau$ test), where the input $\hat{d}_T$ of $\widetilde{z_{t-1}}\left(\hat{d}_T\right)$ is a $T^{\kappa}-$ consistent estimator of $d > 0.5$ with $\kappa > 0$, $\widetilde{\Delta y_t} = \Delta y_t - \Delta\hat{\mu}(t)$, $\widetilde{z_{t-1}}\left(\hat{d}_T\right) = \frac{\left(\Delta^{\hat{d}_T - 1} - 1\right)}{\left(1 - \hat{d}_T\right)}(\Delta y_t - \Delta\hat{\mu}(t))$, and the coefficients of $\Delta\hat{\mu}(t)$ are estimated by an OLS regression of Δy_t on $\Delta\mu(t)$, then the asymptotic properties of the t-statistic t_φ for testing $\varphi = 0$ in (7) are identical to those stated in Theorem 1.

As mentioned above, Shimotsu's (2006) semiparametric estimator provides power rate consistent estimators of $d \leq 1$ for the case where the DGP contains a linear or a quadratic trend whereas Velasco's (1999) estimator is invariant to a linear (and possibly higher order) time trend.

3.2 Serial correlation case: The invariant AEFDF test

Next, we generalize the DGP considered in (1) by assuming that u_t follows an stationary linear $AR(p)$ process, namely, $\Phi_p(L)u_t = \epsilon_t 1_{t>0}$ where $\Phi_p(L) = 1 - \phi_1 L - ...\phi_p L^p$ with $\Phi_p(z) \neq 0$ for $|z| \leq 1$. This motivates the following nonlinear regression model

$$\Delta y_t = \varphi[\Phi_p(L)z_{t-1}(d)] + \sum_{j=1}^{p} \phi_j \Delta y_{t-j} + \epsilon_t, \tag{8}$$

which is similar to (3), except for the inclusion of the lags of Δy_t and for the filter $\Phi_p(L)$ in the regressor whose significance is tested. Estimation of this model is complicated because of the nonlinearity in the parameters φ and $\Phi = (\phi_1, ..., \phi_p)$. Compared with the white noise case, the practical problem arises because the vector Φ is unknown and therefore the regressor $[\Phi_p(L)z_{t-1}(d)]$ is unfeasible. For this reason LV (2007a) recommended to apply a two-step procedure that allows one to obtain efficient tests also with autocorrelated errors.

3.2.1 Two-step procedure.—

For the case where $\mu(t) \equiv 0$ (or known), LV (2007a) implement the two step procedure as follows. In the first step, the coefficients of $\Phi_p(L)$ are estimated (under H_1) by OLS in the equation

$$\Delta^{\widehat{d}_T} y_t = \sum_{t=1}^{p} \phi_j \Delta^{\widehat{d}_T} y_{t-j} + a_t, \tag{9}$$

where $\widehat{d_T}$ satisfies the conditions stated in Theorem 1. The estimator of $\Phi_p(L)$ is consistent with a convergence rate which depends on the rate κ. Second, estimate by OLS the equation

$$\Delta y_t = \varphi[\widehat{\Phi}_p(L) z_{t-1}(\hat{d}_T)] + \sum_{j=1}^{p} \phi_j \Delta y_{t-j} + v_t, \tag{10}$$

where $\widehat{\Phi}_p(L)$ is the estimator from the first step, and $\widehat{d_T}$ denotes the same estimated input used in that step as well. As LV (Theorem 2) have shown, the t_φ statistic in this augmented regression is still both normally distributed and locally optimal. The test will be denoted by AEFDF (augmented EFDF) test in the sequel.

For the case where the coefficients of $\mu(t)$ are considered to be unknown, a similar procedure as that described in section 2.1 can be implemented and efficient tests will still be obtained.

If $\mu(t)$ is a slowly moving trend satisfying Condition A, the test based on regression (10) can be implemented and the asymptotic properties stated in LV (2007a, Theorem 2) still hold through. For the case where $\mu(t)$ satisfies Condition B, in order to maintain the good properties of the test, it is necessary to substract these terms from the original variables prior to computing regressions (9) and (10). As discussed before, the coefficients of $\mu(t)$ can be estimated by OLS under the null and regressions (9) and (10) can be computed after conveniently substracting the estimated deterministic regressors. For instance, if the DGP is defined as in (6), a consistent estimator of β is obtained from the OLS estimator of a regression of Δy_t on a constant term. Clearly, this estimator has the same properties in this case as those described in Section 3.1. Then, regression (9) simply becomes

$$\Delta^{\widehat{d_T}}(y_t - \hat{\beta}t) = [1 - \Phi_p(L)]\Delta^{\widehat{d_T}}(y_t - \hat{\beta}t) + a_t,$$

whereas regression (10) would be

$$\widetilde{\Delta y_t} = \varphi[\widehat{\Phi}_p(L)\widetilde{z_{t-1}}\left(\hat{d}_T\right)] + \sum_{t=1}^{p} \phi_j \widetilde{\Delta y_{t-j}} + v_t, \tag{11}$$

and $\widetilde{\Delta y_t} = \Delta y_t - \hat{\beta}$ and $\widetilde{z_{t-1}}\left(\hat{d}_T\right) = \frac{\left(\Delta^{\hat{d}_T - 1} - 1\right)}{\left(1 - \hat{d}_T\right)}(\Delta y_t - \hat{\beta})$. In the case where $\mu(t)$ in the DGP contains a quadratic term, Δy_t should be regressed on a constant and a linear time trend and so forth for higher-order time trends.

The following theorem states the properties of the AEFDF test in the more general case where short term autocorrelation is present.

Theorem 5 *Under the assumption that the DGP is an ARFIMA$(p, d, 0)$ process defined as $\Phi_p(L)\,\Delta^d(y_t - \mu(t)) = \epsilon_t 1_{t>0}$, where $d \in (0.5, 1)$, ϵ_t is i.i.d. with finite fourth moment and $\Phi_p(L)$ has all its roots outside the unit circle, the asymptotic properties of the t-ratio t_φ for testing $\varphi = 0$ in (10) or (11) for $\mu(t)$ satisfying condition A or B, respectively, using a $T^\kappa-$consistent estimator of d, $\widehat{d}_T$, for some $d > 0.5$ with $\kappa > 0$, are given by*

a) Under the null $(d = 1)$

$$t_\varphi(\widehat{d}_T) \xrightarrow{w} N(0, 1).$$

b) Under local alternatives ($d = 1 - \gamma/\sqrt{T}$, $\gamma > 0$)

$$t_\varphi(\widehat{d}_T) \xrightarrow{w} N(-\gamma\omega, 1).$$

c) Under fixed alternatives $(d < 1)$, the test based on is $t_\varphi(\widehat{d}_1)$ is consistent.

If $\hat{d}_T$ verifies the previous conditions, then LV (2007a) have shown that,

$$\omega^2 = \frac{\pi^2}{6} - \varkappa' \Psi^{-1} \varkappa$$

where $\varkappa = (\varkappa_1, ..., \varkappa_p)'$ with $\varkappa_k = \sum_{j=k}^{\infty} j^{-1} c_{j-k}$, $k = 1, ..., p$, such that c_j ´s are the coefficients of L^j in the expansion of $1/\Phi(L)$, and where $\Psi = [\Psi_{k,j}]$, $\Psi_{k,j} = \sum_{t=0}^{\infty} c_t c_{t+|k-j|}$, $k, j = 1, ..., p$, denotes the Fisher information matrix for $\Phi(L)$ under Gaussianity. Note that ω^2 is identical to the drift of the of the limiting distribution of the LM test under local alternatives (see Tanaka, 1999). Notice also that the use of semiparametric estimators for d is very convenient here, since one can be agnostic about a parametric specification of the autocorrelation in the error terms when estimating the input value of d.

3.2.2 Single-step procedure.—

Despite the fact that LV (2007a) proposed the above-mentioned two-step procedure to account for autocorrelated errors, it is interesting to notice that a feasible linear single-step procedure can also be applied with the same properties. In effect, under H_1 the process for the detrended variable, $\widetilde{\Delta y_t}$, would be

$$\Phi_p(L)\widetilde{\Delta^d y_t} = \varepsilon_t, \tag{12}$$

such that adding and substracting the process under H_0, $\Phi_p(L)\widetilde{\Delta y_t}$, it becomes

$$\Phi_p(L)\widetilde{\Delta y_t} = \Phi_p(L)[1-\Delta^{d-1}]\widetilde{\Delta y_t} + \varepsilon_t,$$

or

$$\widetilde{\Delta y_t} = \varphi[\Phi_p(L)\widetilde{z_{t-1}}(d)] + [1-\Phi_p(L)]\widetilde{\Delta y_t} + \varepsilon_t. \tag{13}$$

The one-step method we propose is based on the following decomposition of the lag polynomial $\Phi_p(L)$

$$\Phi_p(L) = \Phi_p(1) + \frac{1}{\Delta^{d-1}-1}\Phi_p^*(L), \tag{14}$$

where the polynomial $\Phi_p^*(L)$ is defined by equating (13) to the standard polynomial decomposition

$$\Phi_p(L) = \Phi_p(1) + \Delta\widetilde{\Phi}_p(L). \tag{15}$$

Hence,

$$\Phi_p^*(L) = (\Delta^d - \Delta)\widetilde{\Phi}_p(L) = \Delta^d\widetilde{\Phi}_p(L) - [\widetilde{\Phi}_p(L) - \widetilde{\Phi}_p(1)]. \tag{16}$$

Substitution of (13) into (12), using (15) and noticing that $\varphi = d-1$ and $\widetilde{z_{t-1}}(d) = \frac{\Delta^{d-1}-1}{1-d}\widetilde{\Delta y_t}$, yields

$$\widetilde{\Delta y_t} = \varphi\widetilde{z_{t-1}}(d) - \frac{\widetilde{\Phi}_p(L)}{\Phi_p(1)}\widetilde{\Delta^{1+d}y_t} + \frac{1}{\Phi_p(1)}\varepsilon_t. \tag{17}$$

In order to have only predetermined variables on the right hand side of (16), let us rewrite $\widetilde{\Phi}_p(L)$ as $(\widetilde{\Phi}_p(L) - \widetilde{\Phi}_p(0)) + \widetilde{\Phi}_p(0)$ and Δ^{1+d} as $[\Delta^d - 1 + 1]\Delta$. Then, some simple algebra yields

$$\widetilde{\Delta y_t} = \varphi\frac{\Phi_p(1)}{D}\widetilde{z_{t-1}}(d) - \frac{\widetilde{\Phi}_p(L)}{D}[\Delta^d - 1]\widetilde{\Delta y_t} - \frac{\widetilde{\Phi}_p(L) - \widetilde{\Phi}_p(0)}{D} + \frac{1}{D}\varepsilon_t. \tag{18}$$

where $D = \Phi_p(1) + \widetilde{\Phi}_p(0)$. Notice that the second and third regressors are predetermined since $(\Delta^d - 1)$ and $(\widetilde{\Phi}_p(L) - \widetilde{\Phi}_p(0))$ do not include contemporaneous values of $\widetilde{\Delta y_t}$. Hence, in this model the AEFDF test can be done in a single step. For example, in the case of an AR(1) disturbance, i.e., $\Phi_1(L) = 1 - \phi L$, we have that $\Phi_1(1) = 1 - \phi$ and $\widetilde{\Phi}_1(L) = \widetilde{\Phi}_1(0) = \phi$, so that (17) becomes

$$\widetilde{\Delta y_t} = \varphi(1-\phi)\widetilde{z_{t-1}}(d) - \phi[\Delta^d - 1]\widetilde{\Delta y_t} + \varepsilon_t, \tag{19}$$

which, when implemented with a $T^{\kappa}-$consistent estimator of d, $\widehat{d}_T$, gives rise to a t_{φ} test with the same properties as the test based on two steps.

3.3 Power comparison in finite samples

Monte-Carlo evidence in favour of the EFDF and FDF tests can be found in LV (2007a) and DGM (2002), respectively, for the case where $\mu(t) \equiv 0$. In what follows, we provide some additional simulations when $\mu(t) = \alpha + \beta t$. Table 1 presents the rejection frequencies for local alternatives at the 5% level of the EFDF, LM and Shimotsu´s ELW tests (denoted as ELW_{μ} and ELW_{τ}, respectively). The DGP is $y_t = \alpha + \beta t + \Delta^{-d}\varepsilon_t$, which has been simulated $10,000$ times with $\varepsilon_t \sim n.i.d\ (0,1)$, $d = 1 - \gamma/T^{1/2}$ for $\gamma = \{0,\ 0.5,\ 1.0,\ 2.0 \text{ and } 5.0\ \}$ and $T = \{100,\ 400\}$. Shimotsu's (2006) ELW estimator has been used for the input value of d, $\widehat{d}_T$. The figures corresponding to EFDF_{μ}, LM_{μ} and ELW_{μ} are obtained by setting $\alpha = 1$, $\beta = 0$, whereas those for EFDF_{τ}, LM_{τ} and ELW_{τ} pertain to $\alpha = 1$, $\beta = 0.2$. As can be observed, for the smaller sample sizes (when $\gamma = 0$) the LM test is slightly under-sized whereas the EFDF and ELW test are slightly over-sized, specially when we allow for a linear trend. For this reason, we compute size-adjusted power for $\gamma > 0$. The most relevant finding is that, as expected, both EFDF and LM tests have similar power for the two smaller values of γ whereas the former has larger power for $\gamma = 2$ and 5, with improvements up to five percentage points in some instances. In turn, the ELW test behaves somewhat similarly to the other two tests for $\gamma = 0.5$ and 1.0, whilst it loses quite a lot of power for the larger values of γ.

[**Table 1 about here**]

Table 2 reports the (size-adjusted) power when the errors are autocorrelated. The DGP is now $\Delta^d y_t = \varepsilon_t/(1 - 0.2L)$, for several values of $d = 1 - \gamma/T^{1/2}$, using the the same values of γ and $T = \{100,\ 400\}$. The AEFDF test is implemented using model (18). Although for this AR(1) error term, power is lower than in the i.i.d. case, the power comparison across the three tests is similar to the one above, with the AEFDF test performing better for the larger values of γ. Lastly, in Table 3, we briefly report some results on the consequences of having departures from Gaussianity in the distribution of ε_t. We simulate the same DGP as in Table 1, except

$\gamma = 5$, but with errors following an *i.i.d.* (demeaned) $\chi^2(1)$ distribution rather than $n.i.d\ (0,1)$. The reported results correspond to the τ-version of the tests, reaching similar conclusions to the ones discussed earlier.[9]

[**Table 2 about here**]

[**Table 3 about here**]

4. EMPIRICAL ILLUSTRATION

An interesting application of the theoretical results applied above is to examine whether the time-series of GDP per capita of several OECD countries behave as $I(d)$ processes with $d \in (0.5, 1)$. These are series which are clearly trending upwards and therefore provide nice examples of the role of deterministic terms in the use of the EFDF test. As pointed out by Michelacci and Zaffaroni (2000; henceforth, MZ), such a long-memory behavior could well explain the seemingly contradictory results obtained in the literature on growth and convergence. The puzzling result is that a unit root cannot be rejected in (the log of) those series and yet a 2% rate convergence rate to a steady-state level (approximated by a linear trend) is typically found in most empirical exercises testing the so-called unconditional *beta-* convergence hypothesis (see Barro and Sala i Martín, 1995 and Jones, 1995). The explanation offered by MZ to this puzzle relies upon two well-known results in the literature on long-memory processes, namely that standard unit root tests have low power against fractional values of d in the nonstationary range, and that for all values of $d \in [0, 1)$ the effects of shocks die out. Notice that consideration of GDP p.c as an $I(d)$ process may be very reasonable since GDP is obtained as the aggregation of value-added in a wide range of productive sectors which are likely to have different persistence properties (see Lo and Haubrich, 2001). Thus, the aggregation argument popularized by Granger (1980) applies strongly to this case.

Using Maddison's (1995) data set of annual GDP per capita series for 16 OECD countries during the period 1870-1994 and the log-periodogram estimator of d due to Robinson (1995), MZ find that in most countries the order of fractional integration is in the interval (0.5, 1), theoretically compatible with the 2% rate of convergence found in the literature of beta-convergence and, therefore, validating in this way their explanation of the puzzle. Since that estimation

[9] Similar conclusions also hold when the error tem in the DGP follow a Student´s t distribution with 5 d.f.

procedure is restricted to the range of $I(d)$ processes with finite variance, namely, $|d| < 1/2$, MZs proceed by first detrending the data and then applying the truncated filter $(1-L)^{1/2}$ to the residuals, discarding the first ten observations to initialize the series.

The previous results have been criticized by Silverberg and Verspagen (2001) on the grounds that the use of the Geweke and Porter-Hudak (GPH) semi-parametric estimation procedure, as modified by Robinson, suffers from serious small-sample bias. Instead, they propose to use the first-difference filter, $(1-L)$, to remove the trend, and then employ both Beran´s (1994) nonparametric estimator and Sowell's (1992) parametric ML estimator of ARFIMA models to tackle short-memory contamination in the estimation of d. By using these estimation procedures, Silverberg and Verspagen (2001) find, in stark contrast to MZ 's results, that d tends to be either not significantly different from unity or significantly above unity for most countries.

To shed light on this controversy, we apply the AEFDF test developed in Section 3.2 to the logged GDP p.c. of a subset of thirteen of the main OECD countries, listed in Table 4, where (under the null) the estimated intercept and its (Newey-West robust) standard deviation in the regression $\Delta y_t = \beta + u_t$ is reported.[10] As can be inspected, the mean (average GDP p.c. growth rate) is always highly significant making it convenient to use a model which allows for a linear trend, as in (6), as the maintained hypothesis. Indeed, when the ADF and the Phillips-Perron (P-P) unit root tests (not reported) were computed using Elliott et al. (1996)´s efficient GLS detrending procedure, the $I(1)$ null hypothesis could not be rejected in most cases[11]. The KPPS test, which takes $I(0)$ as the null, also yielded rejection in more than half of the cases, confirming the high persistence of the series. Thus it seems clear that the levels of the series have a linear trend and that deviations from such a trend are likely to be nonstationary. In addition, since there were clear signs of autocorrelation in u_t, an AEFDF test was applied to the series. The number of lags of the dependent variable was chosen according to the AIC with a maximum lag of length $k = 5$.

[**Table 4 about here**]

Pre-estimation of d using Shimotsu's (2006) nonparametric approach allows one to estimate a value of d for each country. The estimated values of d are always in the non-stationary range. Taking into account that the standard error (s.e.) of this estimator is $\sqrt{1/4m}$ with $m = T^{0.65}$,

[10] Maddison's (2004) dataset has been employed in this case, which adds 9 observations to the data considered by MZ.

[11] The only exceptions are Canada, Germany and the US with p-values of 0.045, 0.049 and 0.040, respectively.

with a sample size of $T = 134$, it takes a value of 0.102 in all cases. Using this s.e., the value $d = 1$ is included in an appropriate confidence interval of 12 out of the 13 countries, yielding similar results to those in Silverberg and Verspagen (2001). Nevertheless, using the AEFDF test with the above-mentioned estimated input value, $\widehat{d}_T$, the first column of Table 5 shows strong rejections of H_0: $d = 1$ in 6 out of the 13 countries.[12] As discussed earlier, the intuition for this higher rejection rate is the higher power of the EFDF test relative to pure semiparametric tests which yield wider confidence intervals. Thus, our results in almost half of the countries seem to favor nonstationary $I(d)$ processes with $d < 1$, in line with MZ´s conclusions. As Jones (1995) first suggested, this evidence is inconsistent with endogenous growth theories for which permanent changes in certain policy variables have permanent effects on the rate of economic growth. We are aware that a definitely conclusion on this issue requires a deeper data analysis in at least two directions: (i) testing long memory versus structural breaks, and (ii) deriving a panel version of the proposed EFDF test. Both directions are being under current investigation by the authors (for the former, see Dolado, Gonzalo and Mayoral, 2005).

[**Table 5 about here**]

5. CONCLUSIONS

This paper provides new theoretical results regarding gains in power, under fixed alternatives, of applying a Wald test instead of the conventional LM test for detecting the presence of a unit root in time-series data against the alternative of $I(d)$, $d < 1$, possibly allowing for a wide variety of deterministic terms in the DGP. The Wald test is based on the EFDF testing approach (see LV, 2007a). Four main findings have been obtained. *First*, though the EFDF test is asymptotically equivalent to the LM test under local alternatives, it has larger power under fixed alternatives. This gain in power relative to the LM test may also hold for other Wald tests, like the FDF test (see DGM, 2002) which are less efficient than the EFDF test. *Secondly*, if $\mu(t)$ is slowly evolving trend (e.g, including just a constant term), then the EFDF test ignoring $\mu(t)$ can be implemented without losing any of its optimal asymptotic properties. *Thirdly*, if $\mu(t)$ is a

[12] When the estimated value of d was larger than unity, a value of $\hat{d}_T = 1$ was employed as an input to run the test.

polynomial in t of known order but unknown coefficients, then these properties remain identical if one runs the EFDF test on the OLS residuals of the regression of Δy_t on $\mu(t)$ under the null of $d = 1$. And, *fourthly,* under the presence of serial correlation, we show that the EFDF test can be performed in a feasible linear single- step instead of the two- step procedure proposed by LV (2007a). An empirical application regarding the issue of whether deviations from a trend of GDP p.c. in a variety of countries follow an $I(1)$ or a nonstationary $I(d)$ where shocks die out illustrates the usefulness and simplicity of the testing approach proposed here.

Interesting extensions under current investigation by the authors include testing fractional integration versus $I(0)$ allowing for structural breaks (see Dolado, Gonzalo and Mayoral, 2007), testing for cointegration between two $I(d)$ series which have a non-zero drift and where a constant term or a linear trend is included in the regression model and finally, an extension of this framework to panel data.

REFERENCES

Abadir, K., Distaso, W. and L. Giraitis (2005), "Semiparametric estimation for trending I(d) and related processes," mimeo.

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

Barro, R. J. and X. Sala i Martín (1995), *Economic Growth .* McGraw-Hill, New York.

Beran, J. (1994), *Statistics for Long Memory Processes*, New York: Chapman & Hall.

Beran, J. (1995), "Maximum likelihood estimation of the differencing parameter for invertible and short and long memory autoregressive integrated moving average models," *Journal of the Royal Statistical Society*, 57, 659-672.

Breitung, J. and U. Hassler (2002), "Inference on the cointegrated rank of in fractionally integrated processes," *Journal of Econometrics,* 110, 167-185.

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

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 (2007), "Structural breaks vs. long memory: What is what?" Universidad Carlos III, Madrid, Mimeo.

Elliott, G., Rothenberg, T. and J. Stock (1996), "Efficient tests for an autoregressive unit root," *Econometrica*, 64, 813-836.

Engle, R.F. (1984) "Wald, Likelihood Ratio and Lagrange Multiplier Tests in Econometrics," in Z. Griliches and R. Intrilligator (eds) *Handbook of Econometrics*, vol II, 75-826, Amsterdam: North Holland.

Gil-Alaña, L. A. and P. Robinson (1997), "Testing unit roots and other statistical hypothesis in macroeconomic time series," *Journal of Econometrics*, 80, 241-268.

Gonzalo, J. and T. Lee (1998), "Pitfalls in testing for 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. (1986), "Developments in the study of cointegrated economic variables", *Oxford Bulletin of Economics and Statistics*, 48, 213-228.

Hosking, J.R.M., (1996) "Asymptotic distributions of the sample mean, autocovariances, and autocorrelations of long-memory time series", *Journal of Econometrics,* Vol. 73, (1), 261-284.

Johansen, S. (2005), "A representation theory for a class of vector autoregressive models for fractional processes"

Jones, C. (1995), "Time series tests of endogenous growth models," *Quarterly Journal of Economics*, 110, 495-525.

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

Lo, A.W. and J.G. Haubrich (2001), "The sources and nature of long-term dependence in the business cycle," *Economic Review* 37, 15-30.

Lobato, I. and C. Velasco (2007a), "Efficient Wald tests for fractional unit roots," *Econometrica,* 75, 575-589.

Lobato, I. and C. Velasco (2007b), "Optimal fractional Dickey-Fuller tests for unit roots," *Econometrics Journal,* 9, 492-510.

Maddison, A. (2004), *The world economy: historical statistics,* Paris: OECD.

Maddison, A. (1995), *Monitoring the World Economy, 1820-1992*, Paris: OECD.

Mayoral L. (2006), "A new minimum distance estimation procedure of ARFIMA processes," *The Econometrics Journal, Vol. 10.*

Michelacci, C. and P. Zaffaroni, P.(2000), "Fractional beta convergence," *Journal of Monetary Economics*, 45, 129-153.

Robinson, P.M. (1994), "Efficient tests of nonstationary hypotheses," *Journal of the American Statistical Association, 89,* 1420-1437.

Robinson, P. M. (1995), "Log-periodogram of time series with long-range dependence,"*Annals of Statistics* , 23, 1048-1072.

Robinson, P. M. and J. Hualde (2003), "Cointegration in fractional systems with unknown integration orders," *Econometrica* , 71, 1727-1766.

Said, S. and D. Dickey (1984), "Testing for unit roots in autoregressive moving average models of unknown order," *Biometrika*, 71, 599-608.

Shimotsu, K. (2006), "Exact local Whittle estimation of fractional integration with unknown mean and time trend," QED Working Paper 1061.

Silverberg, G. and B. Verspagen (2001), "A note on Michelacci and Zaffaroni, long memory and time series of economic growth," University of Maastricht, Mimeo.

Sowell, F.B.(1992), "Maximum likelihood estimation of stationary univariate fractionally-integrated time-series models," *Journal of Econometrics,* 53, 165-188.

Tanaka, K. (1999), "The nonstationary fractional unit root," *Econometric Theory*, 15, 249-264.

Velasco, C. (1999), "Non-stationary log-periodogram regression," *Journal of Econometrics*, 91, 325-371.

Velasco C. and P. M. Robinson (2000), "Whittle pseudo-maximum likelihood estimation for nonstationary time series," *Journal of the American Association,* 95 **,** 1229-1243.

APPENDIX

Proof of Theorem 2

Let us first consider the case where the true value of d is used to compute the test. In this case, under the alternative hypothesis of $\Delta^d y_t = \varepsilon_t$ with $\varepsilon_t \sim i.i.d.(0,\, \sigma^2)$, the $t_\varphi(d)$ statistic associated to the coefficient of $z_{t-1}(d)$, in the regression of Δy_t on $z_{t-1}(d)$ can be written as,

$$T^{-1/2} t_\varphi(d) = \frac{\sum_{t=2}^{T} \Delta y_t z_{t-1}(d)/T}{\left(\left(\sum_{t=2}^{T} \left(\Delta y_t - \widehat{\varphi} z_{t-1}(d)\right)^2 /T\right)\left(\sum_{t=2}^{T} z_{t-1}^2(d)/T\right)\right)^{1/2}}.$$

Using the results collected in Baillie (1996) stating that, if $\Delta^b y_t = \varepsilon_t$ with $b > -1$, then the variance (γ_0) and the autocorrelation of order j (ρ_j) of y_t satisfy $\gamma_0 = \sigma^2 \Gamma(1-2b)/\Gamma^2(1-b)$ and $\rho_j = [\Gamma\,(j+b)\,(1-b)/\,(\Gamma\,(j-b+1)\,\Gamma(b))]$. In the previous case, where $\Delta y_t \sim I\,(d-1)$ (hence $b = d-1$), it is easy to check that the numerator of $T^{-1/2}\ t_\varphi(d)$ converges in probability to

$$\frac{\sum_{t=2}^{T} \Delta y_t z_{t-1}(d)}{T} = \frac{\sum_{t=2}^{T} (\Delta^{1-d}\varepsilon_t)(\varepsilon_t - \Delta^{1-d}\varepsilon_t)}{(1-d)T} \xrightarrow{p} \frac{\sigma^2}{1-d}[1 - \frac{\Gamma(3-2d)}{\Gamma^2(2-d)}],$$

whereas the two terms in the denominator converge to

$$\frac{\sum_{t=2}^{T} z_{t-1}^2(d)}{T} = \frac{\sum_{t=2}^{T} (\varepsilon_t - \Delta^{1-d}\varepsilon_t)^2}{(1-d)^2 T} \xrightarrow{p} \frac{\sigma^2}{(1-d)^2}[\frac{\Gamma(3-2d)}{\Gamma^2(2-d)} - 1],$$

and

$$\frac{\sum_{t=2}^{T} \left(\Delta y_t - \widehat{\varphi} z_{t-1}(d)\right)^2}{T} \xrightarrow{p} \sigma^2.$$

Replacing the previous limits in the expression for $T^{-1/2} t_\varphi(d)$ yields

$$T^{-1/2} t_\varphi(d) \xrightarrow{p} -\left(\frac{\Gamma(3-2d)}{\Gamma^2(2-d)} - 1\right)^{1/2} \equiv c_{EFDF}(d). \tag{A1}$$

Next, we examine the case where a $T^\kappa -$ consistent estimator of d, $\hat{d}_T$, for some $d > 0.5$ with $\kappa > 0$, is employed to construct the test. In this case, provided

$$T^{-1/2} t_\phi(d) - T^{-1/2} t_\phi(\hat{d}_T) = o_p\,(1)\,, \tag{A2}$$

the limit of $T^{-1/2} t_\phi(\hat{d}_T)$ would also be given by expression $(A1)$. Following LV, we consider the most critical component in this expression, i.e., the numerator of the difference in $(A2)$, given by

$$T^{-1}\left(\sum_{t=1}^{T}\Delta y_t z_{t-1}\left(d\right)-\sum_{t=1}^{T}\Delta y_t z_{t-1}\left(\hat{d}_T\right)\right).$$

Proceeding as Robinson and Hualde (2003), we just need to show that expression

$$T^{-1}\left(\sum_{t=2}^{T}\left(\Delta^{1-d}\varepsilon_t\right)\varepsilon_t-\sum_{t=2}^{T}\left(\Delta^{1-\hat{d}_T}\varepsilon_t\right)\varepsilon_t\right) \tag{A3}$$

tends to zero in probability. It is straightforward to see that

$$\frac{\sum_{t=1}^{T}\left(\Delta^{1-d}\varepsilon_t\right)\varepsilon_t}{T}=\frac{\sum_{t=1}^{T}\left(\varepsilon_t+\pi_1\left(1-d\right)\varepsilon_{t-1}+\pi_2\left(1-d\right)\varepsilon_{t-2}+...+\pi_{t-1}\left(1-d\right)\varepsilon_1\right)\varepsilon_t}{T}\xrightarrow{p}\sigma^2.$$

since all cross-products tend to zero in probability. As for the second term in $(A3)$, it can be written as

$$\frac{\sum\varepsilon_t^2}{T}+T^{-1}\sum_{t=1}^{T}\left(\sum_{i=1}^{t-1}\sum_{j=1}^{t-1}\pi_i\left(1-d\right)\pi_j\left(\hat{d}_T-d\right)\varepsilon_{t-i}\varepsilon_{t-j}\right),$$

where the first term tends to σ^2. By applying similar steps to those considered in LV (2007a, expressions (26)-(28) in appendix 1), it is easy to show that the second term tends to zero in probability. Hence, it follows that $(A3)$ tends to zero in probability and the desired result follows.

Likewise, the FDF test is based on the t-ratio

$$T^{-1/2}t_{\phi}(\tilde{d}_T)=\frac{\sum\Delta y_t\Delta^{\tilde{d}_T}y_{t-1}/T}{\left(\left(\sum\left(\Delta y_t-\widehat{\phi}\Delta^{\tilde{d}_T}y_{t-1}\right)^2/T\right)\left(\sum(\Delta^{\tilde{d}_T}y_{t-1})^2/T\right)\right)^{1/2}}. \tag{A4}$$

As before, when the true value of d is used as input then, by the Law of Large Numbers (LLN), the numerator tends to $(d-1)\,\sigma^2$. With respect to the denominator, we have that $T^{-1}\sum\left(\Delta y_t\right)^2\xrightarrow{p}\sigma^2\Gamma(3-2d)/\left(\Gamma(2-d)\right)^2$ and $\hat{\phi}\xrightarrow{p}(d-1)$. Combining these results, yields

$$T^{-1/2}t_{\hat{\phi}}(d)\xrightarrow{p}\frac{(d-1)\Gamma(2-d)}{[\Gamma(3-2d)-(d-1)^2\Gamma^2(2-d)]^{1/2}}\equiv c_{FDF}(d). \tag{A5}$$

If a consistent estimate of d, $\hat{d}_T$ is employed to run the test, a similar strategy to that followed above can be used to show that $t_{\hat{\phi}}(d)$ also converges to $(A5)$.

Finally, by the LLN the LM test defined in (4), multiplied by $T^{-1/2}$, satisfies that,

$$T^{-1/2}LM_T\xrightarrow{p}\sqrt{\frac{6}{\pi^2}}\sum_{k=1}^{T-1}\frac{1}{k}\rho_k,$$

where ρ_k is the (population) correlation function of a pure $I(d-1)$ process. Using the formula for the autocorrelations given above, yields

$$T^{-1/2} LM_T \xrightarrow{p} \sqrt{\frac{6}{\pi^2} \frac{\Gamma(2-d)}{\Gamma(d-1)} \sum_{j=1}^{\infty} \frac{\Gamma(j+d-1)}{j\Gamma(j-d+2)}} \equiv c_{LM}(d). \blacksquare$$

Proof of Theorem 3

We consider first the case where $d \in (0.5, 1)$ is a fixed number and then extend the proof to case where it is stochastic. In the general case where $\mu(t)$ is different from zero, the t-statistic on the coefficient φ from the OLS regression of Δy_t on z_{t-1} is a function of $\mu(t)$ given by,

$$t_{\varphi}(d, \mu(t)) = \frac{\sum_{t=2}^{T} \Delta y_t z_{t-1}(d)}{\hat{S}_T(d) \sqrt{\sum_{t=2}^{T} (z_{t-1}(d))}}, \tag{A6}$$

where $\hat{S}_T^2(d) = T^{-1} \sum_{t=2}^{T} (\Delta y_t - \widehat{\varphi} z_{t-1}(d))^2$. We now show that the asymptotic distribution of $(A6)$ for the case where $\mu(t)$ satisfies Condition A is the same as in the case where $\mu(t) \equiv 0$. Following the same strategy as LV (2007a), we now prove that, for $d \neq 1$,

$$t_{\varphi}(d, \mu(t)) - t_{\varphi}(d, \mu(t) \equiv 0) = o_p(1),$$

which implies that the test computed ignoring the fact that the DGP contains slowly evolving trends has the same asymptotic properties as in the case where $\mu(t) \equiv 0$.

As in LV, we just analyze the most critical component of $t_{\varphi}(d, \mu(t))$, which is the numerator, since the analysis of the denominator is similar but simpler. Under H_0, the numerator of $(A6)$, multiplied by $T^{-1/2}(1-d)^{-1}$, is given by,

$$T^{-1/2}(1-d)^{-1} \sum_{t=2}^{T} \Delta y_t z_{t-1}(d) = T^{-1/2} \sum_{t=2}^{T} (\Delta\mu(t) + \varepsilon_t) \left(\left(\Delta^d - \Delta \right) \mu(t) + \left(\Delta^{d-1} - 1 \right) \varepsilon_t \right)$$

$$= T^{-1/2} \left(\sum_{t=2}^{T} \varepsilon_t \left(\Delta^{d-1} - 1 \right) \varepsilon_t + \sum_{t=2}^{T} \left(\Delta\mu(t) (\Delta^d - \Delta) \mu(t) \right) + \right. \tag{A7}$$

$$\left. \sum_{t=2}^{T} \Delta\mu(t) \left(\Delta^{d-1} - 1 \right) \varepsilon_t + \sum_{t=2}^{T} \varepsilon_t (\Delta^d - \Delta) \mu(t) \right). \tag{A8}$$

We now show that if $\mu(t) = t^{\delta}$, $\delta \in [0, 0.5)$ all the terms in (A7) and (A8) but the first, $\left(T^{-1/2} \sum_{t=2}^{T} \varepsilon_t \left(\Delta^{d-1} - 1 \right) \varepsilon_t \right)$, converge to zero. Any other specification of $\mu(t)$ satisfying Condition A can be dealt with analogously.

To prove this, notice that the terms t^{δ} and $\Delta^{-\delta}1_{(t>0)}$ are of the same order of magnitude. This is because $\Delta^{-\delta}1_{(t>0)} = \sum_{i=0}^{t-1} \pi_i(-\delta) \approx c\sum_{i=0}^{t-1} i^{\delta-1} = O(t^{\delta})$ (see Davidson, 1994, Theorem 2-27), where c is a constant and the coefficients $\pi_i(-\delta)$ are defined at the end of the Introduction.

The second term in (A7) verifies that,

$$\begin{aligned} T^{-1/2}\left(\sum_{t=2}^{T} \Delta\mu(t)\,\Delta^{d}\mu(t) - \sum_{t=2}^{T}(\Delta\mu(t))^{2}\right) &\approx T^{-1/2}\left(\sum_{t=2}^{T} t^{2\delta-d-1} - \sum_{t=2}^{T} t^{2(\delta-1)}\right) \\ &= T^{-1/2}\left(O\left(T^{2\delta-d}\right) - O(1)\right) \to 0, \quad \text{(A9)} \end{aligned}$$

if $d > 0.5$ and $\delta < 0.5$.

With respect to the first term in (A8),

$$T^{-1/2}E\left(\sum_{t=2}^{T} \Delta t^{\delta}\left(\Delta^{d-1} - 1\right)\varepsilon_t\right) = 0, \quad \text{(A10)}$$

and

$$T^{-1}Var\left(\sum_{t=2}^{T} \Delta t^{\delta}\left(\Delta^{d-1} - 1\right)\varepsilon_t\right) \approx T^{-1}\left(\sigma_{\varepsilon}^{2} + \sigma_{\Delta^{d-1}\varepsilon}^{2}\right)\sum_{t=2}^{T} t^{2(\delta-1)} \to 0, \quad \text{(A11)}$$

where $\sigma^2_{\Delta^{d-1}\varepsilon}$ denotes the variance of the stationary fractionally integrated process $\Delta^{d-1}\varepsilon_t$. Expressions (A10) and (A11) imply that $\sum_{t=2}^{T} \Delta t^{\delta}\left(\Delta^{d-1} - 1\right)\varepsilon_t \xrightarrow{p} 0$. The same type of argument can be used to show that the second term in (A8) also converges to zero. Therefore, for $d \neq 1$, it follows that

$$(1-d)^{-1}T^{-1/2}\sum_{t=2}^{T} \Delta y_t z_{t-1}(d) = (1-d)^{-1}T^{-1/2}\sum_{t=2}^{T} \varepsilon_t\left(\Delta^{d-1} - 1\right)\varepsilon_t + o_p(1), \quad \text{(A12)}$$

which in turn implies that the distribution for the case where the DGP contains slowly evolving trends is the same as that obtained with $\mu(t) = 0$ for the case where d is a fixed number $\in (0.5, 1)$. Considering an stochastic input for $\widehat{d_T}$ amounts to show that

$$t_{\varphi}(d, \mu(t)) - t_{\varphi_{ols}}\left(\hat{d}_T, \mu(t)\right) = o_p(1),$$

where $\hat{d}_T$ satisfies the conditions stated in Theorem 1. It is easy to show, following the same strategy as above, that the last three terms computed with the estimated input $\hat{d}_T$ converge to zero. Hence, the numerator of $t_{\varphi}(d,\ \mu(t)) - t_{\varphi}\left(\hat{d}_T,\ \mu(t)\right)$can be written as

$$(d-1)^{-1}T^{-1/2}\left(\sum_{t=2}^{T} \varepsilon_t\left(\Delta^{d-1} - 1\right)\varepsilon_t - \sum_{t=2}^{T} \varepsilon_t\left(\Delta^{\widehat{d}_T} - 1\right)\varepsilon_t\right) + o_p(1),$$

and LV (2007a, Appendix 1) have shown that the first term of this expression also tends to zero.

The case where $d = 1 - \gamma/\sqrt{T}$ can be solved in an analogous fashion, taking into account the derivations reported in Appendix 1 of LV (2007a). Finally, using the results in DGM and LV, it is straightforward to prove the consistency of the test under fixed alternatives.$\blacksquare$

Proof of Theorem 4

We start, as before, by analyzing the case where the input of z_{t-1}, d, is fixed. We now show that under $H_0 : d = 1$, $t_{\varphi}(d, \mu(t) = 0) - t_{\varphi}(d, \hat{\mu}(t)) \xrightarrow{p} 0$, where in this case $t_{\varphi}(d, \hat{\mu}(t))$ is given by,

$$t_{\varphi}(d, \hat{\mu}(t)) = \frac{\sum_{t=2}^{T} \widetilde{\Delta y_t} \widetilde{z_{t-1}}(d)}{\hat{S}_T(d) \sqrt{\sum_{t=2}^{T} (\widetilde{z_{t-1}}(d))}},$$

where $\widetilde{\Delta y_t} = (\Delta y_t - \Delta \hat{\mu}(t))$, $\widetilde{z_{t-1}}(d) = (1-d)^{-1} \left(\Delta^{d-1} - 1\right) (\Delta y_t - \Delta \hat{\mu}(t))$ and $\hat{S}_T^2(d) = T^{-1} \sum_{t=2}^{T} \left(\widetilde{\Delta y_t} - \hat{\varphi} \widetilde{z_{t-1}}(d)\right)^2$ and $\mu(t)$ satisfies condition B.

For simplicity, we consider the DGP with a linear trend

$$y_t = \alpha + \beta t + \Delta^{-d} \varepsilon_t, d \leq 1, \tag{A13}$$

since any other polynomial of t can be handled accordingly. Let $\hat{\beta}$ be the OLS estimate of β, computed after taking first differences in (A8). Then, $\hat{\beta} = \overline{\Delta y_t}$, where $\overline{\Delta y_t}$ is the sample mean of Δy_t. Notice that under $(A13)$, $\hat{\beta}$ is a $T^{3/2-d}$-consistent estimator of β (see Hosking, 1996). As in Theorem 2, we analyze the numerator of t_{φ} since the analysis of the denominator is similar but simpler.

The numerator of $t_{\varphi}(d, \hat{\mu}(t))$ multiplied by $(1-d)$ is given by,

$$T^{-1/2}(1-d) \sum_{t=2}^{T} \widetilde{\Delta y_t} \widetilde{z_{t-1}} = T^{-1/2} \sum_{t=2}^{T} \varepsilon_t \left(\Delta^{d-1} - 1)\varepsilon_t\right) + T^{-1/2} A_t,$$

where

$$T^{-1/2} A_t = T^{-1/2} \left(\beta - \hat{\beta}\right) \left(\sum \left(\Delta^{d-1} - 1)\varepsilon_t\right) + \left(\beta - \hat{\beta}\right) \sum_{t=2}^{T} \tau_t(d) + \left(\sum_{t=2}^{T} (\tau_t(d) - 1)\varepsilon_t\right)\right),$$

with $\tau_t(\varrho) = \sum_{i=0}^{t-1} \pi_i(\varrho)$ and the coefficients $\pi_i(\varrho)$ are defined at the end of the Introduction. It is easy to check that, under H_0,

$$T^{-1/2} A_t(d_1) = O_p\left(T^{-1}\right) \left(o_p(T) + O_p\left(T^{-1/2}\right) O\left(T^{1-d}\right) + O_p\left(T^{1/2}\right)\right) \xrightarrow{p} 0.$$

The same strategy can be used to show that the denominator of $t_{\varphi}(d, \hat{\mu}(t))$ equals the denominator of $t_{\varphi}(d, \mu(t) = 0)$ plus some terms that go to zero in probability. This implies that

$t_{\varphi}\left(d,\hat{\mu}\left(t\right)\right) \xrightarrow{w} N\left(0,1\right).$ When d is replaced by $\widehat{d_T}$, if $t_{\varphi}\left(d,\hat{\mu}\left(t\right)\right)-t_{\varphi}\left(\hat{d}_T,\hat{\mu}\left(t\right)\right)=o_p\left(1\right)$, then the asymptotic distribution corresponding to $t_{\varphi}\left(\hat{d}_T,\mu\left(t\right)\right)$ would be the same as that of $t_{\varphi}\left(d,\mu\left(t\right)\right)$. Following the same steps as above, it is straightforward to show that $T^{-1/2}A_t\left(\hat{d}_T\right)$ tends to zero. Then, the numerator of $\left(1-d\right)\left(t_{\varphi}\left(d,\mu\left(t\right)\right)-t_{\varphi}\left(\hat{d}_T,\mu\left(t\right)\right)\right)$ can be written as,

$$\left(d-1\right)^{-1}T^{-1/2}\left(\sum_{t=2}^{T}\varepsilon_t\left(\Delta^{d-1}-1\right)\varepsilon_t-\sum_{t=2}^{T}\varepsilon_t\left(\Delta^{\widehat{d}_T-1}-1\right)\varepsilon_t\right)+o_p\left(1\right),$$

and LV (2007a) have shown that this expression tends to zero under the conditions stated in Theorem 1. Similar results can be easily obtained for the denominator. Hence, $t_{\varphi}\left(\hat{d}_T,\hat{\mu}\left(t\right)\right) \xrightarrow{w} N\left(0,1\right).$

Again, the case where $d=1-\gamma/\sqrt{T}$ can be solved in a similar manner, taking into account the derivations reported in Appendix 1 of LV(2007a). Likewise, using the results in DGM and LV, the proof of the consistency of the test under fixed alternatives is straightforward.■

Proof of Theorem 5

The proof of this theorem can be easily constructed along the lines of Appendix 2 in LV (2007a) and Theorems 2 and 3 above. Therefore, it is omitted.■

TABLES

TABLE 1

Size and Power(*) of EFDF, LM and ELW tests, 5% S.L.

DGP: $y_t = \alpha + \beta t + \Delta^{-d}\varepsilon_t$

$d = 1 - \gamma/\sqrt{T}$; $\alpha = 1$; $\varepsilon_t \sim N(0,1)$

	EFDF$_\mu$, (β=0)		EFDF$_\tau$, (β=0.2)		LM$_\mu$, (β=0)		LM$_\tau$, (β=0.2)		ELW$_\mu$, (β=0)		ELW$_\tau$, (β=0.2)	
$\gamma\backslash T$	100	400	100	400	100	400	100	400	100	400	100	400
0	0.061	0.045	0.074	0.087	0.031	0.029	0.049	0.046	0.071	0.075	0.072	0.074
0.5	0.116	0.195	0.123	0.164	0.111	0.190	0.100	0.162	0.091	0.090	0.086	0.096
1	0.287	0.378	0.252	0.328	0.273	0.369	0.237	0.324	0.150	0.158	0.125	0.154
2	0.728	0.834	0.649	0.774	0.681	0.803	0.612	0.748	0.361	0.353	0.301	0.339
5	1.000	1.000	1.000	1.000	0.980	0.991	0.951	0.962	0.953	0.932	0.909	0.940

(*) Size-adjusted power

TABLE 2

Size and Power(*) of AEFDF, LM and ELW Tests, 5% S.L.

DGP: $y_t = \alpha + \beta t + \Delta^{-d}\epsilon_t/\left(1 - \phi L\right)$

$d = 1 - \gamma/\sqrt{T}$; $\phi = 0.2$; $\alpha = 1$; $\beta = 0.2$; $\phi = 0.2$; $\varepsilon_t \sim N(0,1)$

	AEFDF$_\mu$, (β=0)		AEFDF$_\tau$, (β=0.2)		LM$_\mu$, (β=0)		LM$_\tau$, (β=0.2)		ELW$_\mu$, (β=0)		ELW$_\tau$, (β=0.2)	
$\gamma\backslash$T	100	400	100	400	100	400	100	400	100	400	100	400
0	0.061	0.066	0.075	0.065	0.033	0.031	0.029	0.050	0.058	0.056	0.057	0.053
0.5	0.093	0.091	0.075	0.085	0.079	0.077	0.070	0.079	0.097	0.072	0.083	0.070
1	0.139	0.154	0.108	0.155	0.112	0.151	0.091	0.134	0.139	0.121	0.121	0.118
2	0.328	0.352	0.237	0.342	0.285	0.338	0.224	0.314	0.321	0.278	0.312	0.291
5	0.952	0.973	0.840	0.942	0.891	0.947	0.811	0.914	0.910	0.898	0.814	0.890

(*) Size-adjusted power.

TABLE 3

POWER(*) OF EFDF, LM AND ELW TESTS, 5% S.L.

DGP: $y_t = \alpha + \beta t + \Delta^{-d}\varepsilon_t$,

$d = 1 - \gamma/\sqrt{T}$; $\alpha = 1$; $\beta = 0.2$; $\varepsilon_t \sim$ (demeaned) χ^2_1

	$EFDF_\tau$		LM_τ		ELW_τ	
$\gamma/\ T$	100	400	100	400	100	400
0.5	0.151	0.150	0.127	0.143	0.088	0.093
1	0.321	0.354	0.272	0.331	0.166	0.172
2	0.737	0.806	0.704	0.765	0.376	0.387

(*) Size-adjusted power

TABLE 4

ESTIMATES OF $\widehat{\beta}$ AND ROBUST S.E$(\widehat{\beta})$ IN $\Delta y_t = \beta + u_t$

Country	Mean	Robust s.e.
Australia	0.0148	0.004
Belgium	0.015	0.005
Canada	0.0195	0.005
Denmark	0.0184	0.008
France	0.0185	0.006
Germany	0.0176	0.007
Italy	0.0192	0.006
Netherlands	0.0154	0.006
Norway	0.0221	0.006
UK	0.0143	0.003
USA	0.0186	0.005
Spain	0.0199	0.005
Sweden	0.0193	0.005

TABLE 5

AEFDF Test

$H_0 : I(1)\ vs.\ H_A\!:\ d < 1$

Country	$t_\varphi(\widehat{d}_T)$	$\widehat{d}_T$ $(s.e. = 0.10)$
Australia	-1.02	1.10
Belgium	-0.74	0.98
Canada	-2.58*	0.80
Denmark	-0.72	0.99
France	-1.82*	1.08
Germany	-1.94*	0.83
Italy	-0.18	0.98
Netherlands	-1.76*	0.92
Norway	-1.03	0.98
UK	-1.94*	0.87
USA	-3.50*	0.63
Spain	-0.17	1.18
Sweden	-0.07	1.12

Note.- (*) denotes 5%-rejection.