TESTING $I(1)$ AGAINST $I(d)$ ALTERNATIVES WITH WALD TESTS IN THE PRESENCE OF DETERMINISTIC COMPONENTS

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, Universidad Pompeu Fabra.

December 21, 2006

Abstract

This paper analyses how to test $I(1)$ against $I(d)$, $d < 1$, in the presence of deterministic components in the DGP, by extending a Wald-type test, i.e., the (Efficient) Fractional Dickey-Fuller (EFDF) test, to this case. Tests of these hypotheses are important in many economic applications where it is crucial to distinguish between permanent and transitory shocks because $I(d)$ processes with $d < 1$ are mean-reverting. On top of it, the inclusion of deterministic components becomes a necessary addition in order to analyze most macroeconomic variables. We show how simple is the implementation of the EFDF in these situations and argue that, in general, has better properties than LM tests. Finally, an empirical application is provided where the EFDF approach allowing

*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 acknoledged. The usual disclaimer applies.

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 growth theories, and on which there has been some controversy.

JEL Clasification: C12 C22 O40

Keywords: Deterministic components, Dickey-Fuller test, Fractionally Dickey-Fuller test, Fractional processes, Long memory, Trends, Unit roots.

1. INTRODUCTION

It is well known that lack of power of unit root tests may lead to the wrong conclusion that a time series (y_t) is $I(1)$ when it happens to be a fractionally integrated $I(d)$ process with $0 \leq d < 1$. This mistake can have very serious consequences, particularly in the medium and long run. To mention only two: (i) shocks can be identified as permanent when in fact are mean reverting, and (ii) two series can be considered to be spuriously cointegrated (i.e., a concept introduced and analyzed in Gonzalo and Lee, 1998) when in fact they are independent at all leads and lags. 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. We first extend an existing Wald-type testing procedure for detecting a unit root against mean-reverting fractional alternatives in time series free of deterministic components to the more realistic case where they may exhibit a wide variety of trending behaviors. Secondly, we show that this test, apart from its simplicity, has better properties than other available tests in the literature with the same goal.

Specifically, we focus on a modification recently suggested by Lobato and Velasco (2005; LV hereafter) of the Fractional Dickey-Fuller (FDF) test proposed by Dolado, Gonzalo and Mayoral (2002, 2003; DGM hereafter) that achieves a slight improvement in efficiency over the latter. 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 $d \in [0, 0.5) \cup (0.5, 1)$.[1] Both the FDF

[1] Although the case where $d = 0.5$ was treated in DGM, it constitutes a discontinuity point in the analysis

and EFDF tests belong to the family of Wald tests and rely upon the DF approach. The underlying idea is to test for the statistical significance of the coefficient of the regressor in a possibly unbalanced regression where the dependent variable is the time series filtered under the null (Δy_t) and the regressor is some transformation of the series filtered under the alternative ($\Delta^d y_t$). Whereas DGM suggested choosing $\Delta^d y_{t-1}$ as the regressor, LV have shown that a more efficient test could be achieved by using the alternative regressor $z_{t-1}(d) = (1-d)^{-1}(\Delta^d - \Delta)y_t$.[2] As in the FDF procedure, the EFDF test is based upon the t-ratio, $t_\varphi(d)$, of the relevant coefficient on $z_{t-1}(d)$, φ. Thus, non-rejection of H_0: $\varphi = 0$ against H_A: $\varphi < 0$, implies that the process is $I(1)$, namely, $\Delta y_t = \varepsilon_t$ where ε_t are assumed to be i.i.d. Conversely, rejection of the null implies that the process is $I(d)$, with $d < 1$.

In order to compute the filtered regressors, an input value for d is needed. Both DGM and LV recommend to select this value using a T^κ-consistent estimate (with $\kappa > 0$) of the true integration order, d, and show that the limiting distribution of the resulting statistic is a $N(0,1)$.

The advantages of these Wald-type tests, in parallel with the DF approach, rely on their simplicity (e.g., they can be easily implemented in standard econometric software) and good finite sample performance. Further LV (2005, Theorem 1) have shown that, under a sequence of local alternatives approaching $H_0 : d = 1$ from below at a rate of $T^{-1/2}$ with Gaussian errors, the EFDF test is asymptotically equivalent to the uniformly most powerful invariant (UMPI) test and hence asymptotically equivalent under local alternatives to the LM test introduced by Robinson (1994). But interestingly, as discussed in section 3, an additional important advantage of both the FDF and EFDF tests is that their non-centrality

of fractionally integrated processes, splitting the class of $I(d)$ processes into stationary (for $d < 0.5$) and nonstationarity (for $d \geq 0.5$). Moreover the behaviour of $\{y_t\}$ differs between $d = 0.5$ and $d > 0.5$; cf. Liu (1998). For this reason, as is often the case in most of the literature, we ignore this possibility. To simplify the notation in the sequel, however, we will refer to the permissable range of d under the alternative as $0 \leq d < 1$.

[2] Both regressors can be constructed by applying the truncated binomial expansion of the filter $\Delta^d = (1-L)^d$ to y_t, 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.

parameters under H_A are smaller under non-local alternatives than that corresponding to the LM test, and hence they are more powerful under alternatives that are not local to the null. These asymptotic results are corroborated by Monte Carlo simulations that show the superiority in terms of power of the Wald-type tests versus the LM one. Finally, the Wald tests present the advantage of not requiring the correct specification of a parametric model, a useful property stemming from the possible choice of semiparametric estimators for the input value of d when performing the test.[3]

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 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 is neither considered by DGM nor by LV. Although we will consider other types of trends, we will focus mainly 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 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 correlation, it is sufficient to augment the set of regressors in the auxiliary regression described above with k lags of the dependent variable such that $k \uparrow \infty$ as $T \uparrow \infty$, and

[3] Although DGM proposed a $\sqrt{T}-$consistent estimator for the input value of d in the FDF test, LV (2006) have shown that a Gaussian semiparametric estimator, such as the one proposed by Velasco (1999) suffices to achieve consistency and asymptotic normality, a result which also holds for the EFDF test (see sections 2 and 3 below).

$k^3/T \uparrow 0$, as in Said and Dickey (1984), leading 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 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 test. An empirical application dealing with testing the possibility that long GNP *per capita* series for several OECD countries may follow mean-reverting $I(d)$ processes (supporting the hypothesis of beta-convergence) instead of $I(1)$ (no convergence), serves to illustrate our proposed methodology.

The rest of the paper is structured as follows. Sections 2 analyzes the derivation of invariant EFDF tests when the null hypothesis is a random walk with or without deterministic components. Section 3 focuses on the comparison of both FDF and EFDF tests with the LM tests discussed above. Section 4 discusses an empirical application of the previous tests. 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 an (asymptotically) stationary process when $d < 0.5$, and of a non-stationary (truncated) process when $d > 0.5$. 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 and are defined as

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

The indicator function is denoted by $1_{(.)}$. Finally, $\stackrel{w}{\rightarrow}$ and $\stackrel{p}{\rightarrow}$ denote weak convergence and convergence in probability, respectively.

2. DEFINITION OF THE INVARIANT EFDF TEST

2.1 The i.i.d. case

Like in Robinson (1994), we assume that the process y_t 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 $u_t = \Delta^{-d}\epsilon_t 1_{t>0}$ is a purely stochastic $I(d)$ process and ϵ_t is an i.i.d random variable. For the case where $\mu(t) \equiv 0$,[4] DGM introduced a Wald-type (FDF) test for testing the null hypothesis of $H_0 : d = 1$ versus a simple alternative $\mathrm{H}_A : d = d_A < 1$ or a composite alternative $d < 1$, based on the t-statistic associated to the hypothesis $\phi = 0$ in the regression

$$\Delta y_t = \phi \Delta^{d_1} y_{t-1} + \varepsilon_t. \tag{2}$$

They show that if the value d_1 was chosen using a $\sqrt{T}-$ consistent estimator of d, the asymptotic distribution of the resulting t-statistic, $t_\phi(d_1)$ is $N(0,1)$. Furthermore, they also show that, in spite of not being locally optimal (as Robinson's LM test is), its finite sample performance is more satisfactory except when considering local alternatives with gaussian errors.

Recently, LV (2005) 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, again assuming that $\mu(t) \equiv 0$ (or known). More specifically, they propose to compute the t-statistic, $t_\varphi(d_1)$, associated to the null hypothesis $\varphi = 0$ in the regression

$$\Delta y_t = \varphi z_{t-1}(d_1) + \epsilon_t, \tag{3}$$

where $z_{t-1}(d_1)$, with d_1 being an input value for d, is defined as

$$z_{t-1}(d_1) = \frac{\left(\Delta^{d_1-1} - 1\right)}{(1-d_1)} \Delta y_t.$$

[4] 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.

A similar model was first proposed by Granger (1986) in the context of testing for cointegration with multivariate series, a modification of which has been recently considered by Johansen (2005). Notice that it serves to test the null $H_0 : d = 1$ against the alternative H_A: $0 \leq d < 1$. When $\varphi = 0$, the model becomes a random walk, i.e., $\Delta y_t = \epsilon_t$, while under the alternative, if the input value d_1 is chosen such that $d_1 = d$, then $\varphi = -(1-d)$ and the process becomes $\Delta^d y_t = \epsilon_t$.

In the rest of this section, we extend the previous testing approach to the more realistic case where $\mu(t) \neq 0$ is considered to be unknown and examine how this deterministic term should be taken into account to carrying out the test. We will concentrate on the general case where a composite alternative hypothesis is considered (i.e., $H_A : d < 1$) in a regression model that includes the regressor $z_{t-1}(d_1)$ to perform the test.

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 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 easy to show that the stochastic component in y_t dominates the deterministic term when T is large. Hence, the term $\mu(t)$ has no effect on either 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 1 *(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 for testing* $\varphi = 0$ *in* (3) *(denoted* *by EFDF$_\mu$ test), where the input of $z_{t-1}(\widehat{d}_1)$ is a $T^{\kappa}-$consistent estimator of d_1, for some $d_1 > 0.5$ with $\kappa > 0$, are given by,*

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

$$t_{\varphi}(\widehat{d}_1) \overset{w}{\to} N(0, 1).$$

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

$$t_{\varphi}(\widehat{d}_1) \overset{w}{\to} N(-\gamma h(d_1), 1),$$

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, the test based on $t_{\varphi}(\widehat{d}_1)$ is consistent.

LV (2005) show that the function $h(.)$ achieves a global maximum at $d_1 = 1$ where $h(1) = \sqrt{\pi^2/6}$, and that $h(1)$ equals the noncentrality parameter of the locally optimal Robinson's LM test. Hence, if a T^{κ}-consistent estimator of d is used as input of $z_{t-1}(d)$, the EFDF test is locally asymptotically equivalent to the LM test even in the case where the DGP contains a deterministic term, $\mu(t)$, verifying Condition A. 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, 2006) but also, less restrictively, some semiparametric estimators of d as LV (2005, 2006) have shown. Among the latter class, the estimators proposed by Shimotsu (2006) and Velasco (1999) represent good choices since both allow for the existence of deterministic components.

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 tests described above are unfeasible. Nevertheless, it is still possible to obtain a feasible test with the same asymptotic properties as those described in Theorem 1 if a consistent estimate of $\mu(t)$ is substracted from the original processes. All the coefficients of $\mu(t)$ but the constant term, can be consistently estimated by OLS in a regression, under the null, of

Δy_t on $\Delta\mu(t)$. For instance, consider the case where the DGP contains a linear time trend, that is,

$$y_t = \alpha + \beta t + \Delta^{-d}\epsilon_t, \tag{4}$$

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_A. Under H_0, $\hat{\beta}$ is a $T^{1/2}$ -consistent estimator of β whereas under H_A, $\hat{\beta}$ is $T^{3/2-d}$-consistent (see Hosking 1996, Theorem 8). Notice that, under $H_A : d < 1$, it holds that $3/2 - d > 0.5$. Hence, the following theory holds.

Theorem 2 *(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,* $t_\varphi(\widehat{d_1})$*, for testing* $\varphi = 0$ *in the regression*

$$\widetilde{\Delta y_t} = \varphi \widetilde{z_{t-1}}\left(\hat{d}_1\right) + e_t \tag{5}$$

(denoted by EFDF$_\tau$ *test), where the input* $\hat{d}_1$ *of* $\widetilde{z_{t-1}}\left(\hat{d}_1\right)$ *is a* $T^\kappa-$ *consistent estimator of* $d_1 > 0.5$ *with* $\kappa > 0$*,* $\widetilde{\Delta y_t} = \Delta y_t - \Delta\hat{\mu}(t)$*,* $\widetilde{z_{t-1}}\left(\hat{d}_1\right) = \frac{\left(\Delta^{\hat{d}_1 - 1} - 1\right)}{(1-\hat{d}_1)}(\Delta y_t - \Delta\hat{\mu}(t))$*, 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 for testing* $\varphi = 0$ *in (*5*) are the same as those described 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.

2.2 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 $|z| \leq 1$. For the case where $\mu(t) \equiv 0$ (or known), LV recommended to apply a two-step procedure that allows one to obtain efficient tests also in the autocorrelated case. In the first step, the coefficients of $\Phi_p(L)$ are estimated (under H_A) by OLS in the equation

$$\Delta^{\widehat{d}_1} y_t = \sum_{t=1}^{p} \phi_j \Delta^{\widehat{d}_1} y_{t-j} + a_t, \tag{6}$$

where $\widehat{d}_1$ is a T^{κ}-consistent estimator of d_1 such that 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

$$\Delta y_t = \varphi[\widehat{\Phi}_p(L) z_{t-1}(\hat{d}_1)] + \sum_{j=1}^{p} \phi_j \Delta y_{t-j} + v_t, \tag{7}$$

where $\widehat{\Phi}_p(L)$ is the estimator from the first step, and $\widehat{d}_1$ denotes the same estimated input used in that step as well. As LV (Theorem 2) have shown, the $t_{\varphi}(\widehat{d}_1)$ statistic associated to φ in this augmented regression is still both normally distributed and locally optimal. The test will be denoted by AEFDF (augmented EFDF) test in the following.

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 (7) can be implemented and the asymptotic properties stated in LV (2005, 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 (6) and (7). The coefficients of $\mu(t)$ can be estimated by OLS under the null in a similar way as that described in the previous section. Next, regressions (6) and (7) can be computed after conveniently substracting the estimated deterministic regressors. For instance, if the DGP is defined as in (4), a consistent estimator of β can be obtained (after taking first differences) by computing 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 2.1. Then, regression (6) simply becomes

$$\Delta^{\widehat{d}_1}(y_t - \hat{\beta}t) = [1 - \Phi_p(L)]\Delta^{\widehat{d}_1}(y_t - \hat{\beta}t) + a_t,$$

whereas regression (7) would be

$$\widetilde{\Delta y_t} = \varphi[\widehat{\Phi}_p(L)\widetilde{z_{t-1}}\left(\hat{d}_1\right)] + \sum_{t=1}^{p} \phi_j \widetilde{\Delta y_{t-j}} + v_t, \tag{8}$$

and $\widetilde{\Delta y_t} = \Delta y_t - \hat{\beta}$ and $\widetilde{z_{t-1}}\left(\hat{d}_1\right) = \frac{\left(\Delta^{\hat{d}_1 - 1} - 1\right)}{\left(1 - \hat{d}_1\right)}(\Delta y_t - \hat{\beta})$. In the case where the DGP contains a quadratic term, Δy_t should be regressed on a constant and a time trend and a procedure similar to the one described above should be implemented.

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

Theorem 3 *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 $d \leq 1$, 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 for testing $\varphi = 0$ in (7) or (8) for $\mu(t)$ satisfying condition A or B, respectively, using a $T^{\kappa}-$consistent estimator of d_1, for some $d_1 > 0.5$ with $\kappa > 0$, are given by*

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

$$t_{\varphi}(\widehat{d_1}) \xrightarrow{w} N(0,1).$$

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

$$t_{\varphi}(\widehat{d_1}) \xrightarrow{w} N(-\gamma\omega, 1).$$

c) Under fixed alternatives $(d < 1)$, the test based on is $t_{\varphi}(\widehat{d_1})$ is consistent.

If $\hat{d}_1$ is a consistent estimator of d, then LV (2005) have shown that,

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

$\varkappa = (\varkappa_1, ..., \varkappa_p)'$ with $\varkappa_k = \sum_{j=k}^{\infty} j^{-1} c_{j-k}$, $k = 1, ..., p$, where c_j 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. Notice that the use of semiparametric estimators for d_1 is very convenient here, since one does not need to care about a parametric specification of the autocorrelation in the error terms.

3. WALD VS. LM TESTS

As discussed in the Introduction, the closest competitor to the Wald (FDF and EFDF) tests is the LM test proposed by Robinson (1994) in the frequency domain, subsequently extended by Tanaka (1999) to the time domain. In this section we will discuss the power of the three competing tests.

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 $\Delta^{d_0+\theta}[y_t - \mu(t)] = \varepsilon_t$. Thus, 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 log-likelihood function can be written as

$$L\ (\theta, \sigma) = -\frac{T}{2}\ln(\sigma^2) - \frac{1}{2\sigma^2}\sum_{t=1}^{T}[(1-L)^{1+\theta}y_t]^2. \tag{9}$$

Then, taking the derivative of the log-likelihood function w.r.t. θ, evaluated at $\theta = 0$, and making use of the result $\sum_{j=1}^{\infty} j^{-2} = \pi^2/6$, yields the following score-LM test (see Robinson, 1994 and Tanaka, 1999)

$$LM_T = \sqrt{\frac{6}{\pi^2}}T^{1/2}\sum_{j=1}^{T-1} j^{-1}\widehat{\rho}_j \xrightarrow{w} N(0,1), \tag{10}$$

where $\widehat{\rho}_j = \sum_{t=j+1}^{T}\widetilde{\Delta y_t}\widetilde{\Delta y_{t-j}} \,/\sum_{t=1}^{T}(\widetilde{\Delta y_t})^2$, and $\widetilde{\Delta y_t}$ are the OLS residuals from regressing Δy_t on $\Delta\mu(t)$. Therefore, if just a constant term is considered, then $\widetilde{\Delta y_t} = \Delta y_t$; likewise, with a linear trend, $\widetilde{\Delta y_t} = \Delta y_t - \overline{\Delta y}$ where $\overline{\Delta y}$ denotes the sample mean of Δy_t.

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

$$\widetilde{\Delta y_t} = \lambda x^*_{t-1} + e_t, \tag{11}$$

where $x^*_{t-1} = \sum_{j=1}^{t-1} j^{-1}\widetilde{\Delta y_{t-j}}$. Intuitively, since $t_\lambda = \sum(\Delta\widetilde{y}_t x^*_{t-1})/\widehat{\sigma}_e(\sum(x^*_{t-1})^2)^{1/2}$ 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.

Tanaka (1999) has proved that, under a sequence of local alternatives of the type $\theta = T^{-1/2}\gamma$ with $\gamma > 0$, LM_T (or t_λ) is the UMPI test. However, as LV have pointed out, when $\widehat{d}_1$ tends to 1 in the EFDF test, the indetermination 0/0 in the filter $\left(\Delta^{d_1-1} - 1\right)/(1 - d_1)\rfloor_{d_1=1}$ 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. In the case where $\mu(t) \equiv 0$ (or known) and $0.5 < d < 1$, under the sequence of local alternatives described above in Theorem 1, LV (2005) have shown that the limiting distribution of the LM test is $N(-\gamma h(d),\, 1)$ where $h(.)$ is $\pi^2/6$ when $d = 1$. On the other hand, DGM (2002, Theorem 3) obtained that the corresponding distribution of the FDF under local alternatives test is $N(-\gamma,\, 1)$. Since $\pi^2/6 > 1$ ($\simeq 1.25$), close to the null, the asymptotic efficiency of the FDF test relative to the LM and EFDF tests is 0.80 ($\simeq 1/1.25$).

Since the results above on local alternatives are well known, we focus in the rest of this section on the case of fixed alternatives, where 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 $d \in (0,\, 1)$ and where, for simplicity, we assume that there are not deterministic regressors and that the true value of d is used to compute the FDF and EFDF tests. Hence, $\Delta y_t = \Delta^{-b}\varepsilon_t$ where $b = d - 1 < 0$. Then, the following result holds.

Theorem 4 *If $\Delta^d y_t = \varepsilon_t$ with $d \in (0, 1)$, the t-statistics associated to the EFDF and FDF tests, denoted as $t_\varphi(d)$ and $t_\phi(d)$, respectively, verify,*

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

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

while, under the same DGP, the LM test defined in (10) *satisfies that,*

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

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 above-mentioned non-centrality parameters for $d \in (0,1)$.[5] It is seen that 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, although the LM test was devised to be the UIMP test for local alternatives, it performs much 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 (2005) arguments. Finally, the results in Theorem 4 are asymptotic and, as will be shown below, for realistic sample sizes, the rejection rates of the Wald tests under the alternative are also larger than those of the LM test, except in cases where d is very close to unity and the error term is normally distributed, where the EFDF and LM tests behave similarly. Thus, for fixed alternatives, with approximately $d < 0.90$, the EFDF (and the FDF) test is bound to exhibit much higher power than the LM test.

[5] Notice that Theorem 4 excludes the point $d = 0$. For $d = 0$, it is easy to show that $c_{EFDF}(0) = c_{FDF}(0) = -1$. As for C_{LM}, notice that $\Delta y_t = \Delta\varepsilon_t$ and therefore the only non-zero correlation is $\rho_1 = -0.5$. Thus $c_{LM}(0) = -0.5\sqrt{6/\pi^2} \simeq -0.39$.

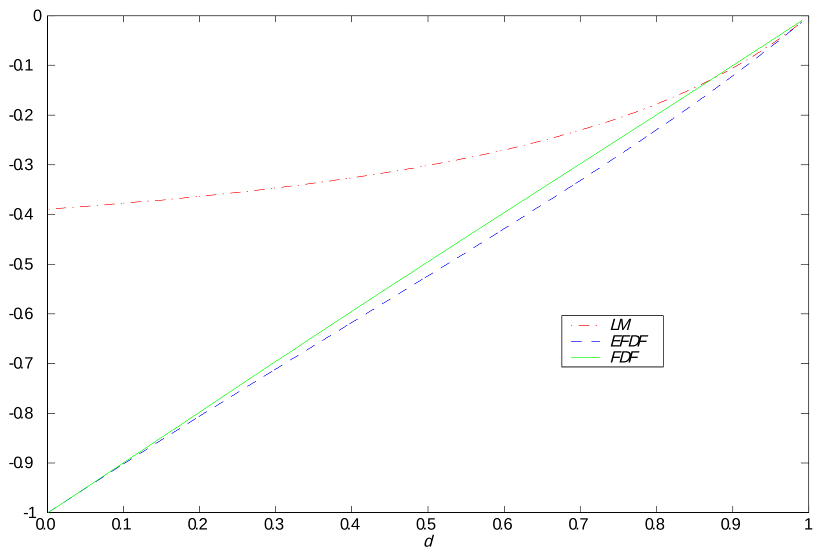

FIG 1. Non-centrality parameters of LM and Wald tests

Monte-Carlo evidence in favor of the EFDF and FDF tests has been provided by LV and DGM, respectively, when deterministic components are absent. 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 and the LM test. The DGP, $y_t = \alpha + \beta t + \Delta^{-d}\varepsilon_t$, is simulated 10,000 times, with $\varepsilon_t \sim i.i.d\ N(0,1)$, $d = 1 - \gamma/T^{1/2}$ for $\gamma = \{0,\ 0.5,\ 1.0,\ 2.0 \text{ and } 5.0\}$, $T = \{25,\ 50,\ 100,\ 400\}$, and the input value for d, (d_1), has been estimated using Shimotsu's (2006) exact local Whittle estimator. The figures corresponding to $EFDF_\mu$ and LM_μ are obtained by setting $(\alpha = 1,\ \beta = 0)$ whereas those for EFDF_τ and LM_τ are computed by setting $(\alpha = 1,\ \beta = 1)$. To compute the EFDF_μ and EFDF_τ tests, regression models (3) and (5) have been used, respectively. As can be observed, for the smaller sample sizes (when $\gamma = 0$) the LM test are slightly undersized whereas the EFDF test is slightly over-sized, especially when we allow for a trend. However, the most relevant finding is that, in general (using effective sizes), the EFDF has larger power than the LM tests, in accord with the result derived in Theorem 4 above.

[**Table 1 about here**]

Table 2, in turn, reports the size and the (size-adjusted) power when the errors are autocorrelated, so that the DGP is $\Delta^d y_t = \varepsilon_t/(1-0.6L)$, for several values of $d = 1 - \gamma/T^{1/2}$, for the same values of γ and $T = \{100,\, 400\}$. In this case, the AEFDF test clearly outperforms the LM tests. Lastly, we briefly report some results on the consequences of having departures from Gaussianity in the distribution of ε_t in the above-mentioned DGP. For example, when the errors follow an *i.i.d.* zero-mean standardized $\chi^2(1)$ distribution, the power of the EFDF test, for $d = 0.8$, 0.9 and $T = 100$, is 62.9 % and 30.5 % whereas the corresponding rejection frequencies of Tanaka's LM_T are 52.8 % and 17.2 %, respectively. Thus, the EFDF test also fares better than the LM test in the presence on non-Gaussian errors.

[**Table 2 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 $d < 1$. 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 in an interesting paper 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 values of d in the nonstationary range ($0.5 < d < 1$), and that for all values of $d \in [0,\, 1)$ there is "mean reversion", in the sense that the effects of shocks die out. Notice that the $I(d)$ nature of the GDP p.c. series 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 and Joyeux (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 a log-periodogram estimator of d due to Robinson (1995), they 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 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 10 observations to initialize the series.

The previous results have been recently criticized by Silverberg and Verspagen (2001) on the following grounds. First, they are critical with the procedure used by MZ of filtering out a deterministic linear-in-logs trend and then using the difference $(1-L)^{1/2}$ on the residuals. Second, they criticize the use of the Geweke and Porter-Hudak (GPH) semi-parametric estimation procedure as modified by Robinson, which suffers from serious small-sample bias. Instead, they propose to use the first-difference filter, $(1-L)$, to remove the trend, and then employ the nonparametric FGN estimator due to Beran (1994) and the Sowell's (1992) parametric ML estimator of ARFIMA models to tackle short-memory contamination in the estimation of d. Using the non-parametric FGN estimator they 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 in an extended sample of 25 countries.

To shed light on this controversy, we apply the AEFDF test developed in Section 2.2 to the logged GDP p.c. of a subset of thirteen of the main OECD countries, listed in Table 3, where the estimated intercept and its standard deviation in the regression $\Delta y_t = \beta + u_t$ is reported.[6] As can be inspected, the mean (average GDP p.c. growth rate) is always highly

[6] Maddison's (2004) dataset has been employed in this case, which adds 9 observations to the data

significant making it convenient to use a model which allows for a linear trend, as in (4), as the maintained hypothesis. Indeed, when the ADF and the Phillips-Perron (P-P) unit root tests (not reported) were computed using a constant and a time trend in the regression model, the $I(1)$ null hypothesis could not be rejected in most cases. Further, the KPPS test, which takes $I(0)$ as the null, also yielded rejection in about 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 3 about here**]

Pre-estimation of d using Shimotsu's (2006) nonparametric approach allows one to estimate a value of d for each country. Taking into account that the standard error (s.e.) of this estimator is $\sqrt{1/4m}$ with $m = T^{0.65}$, with a sample size of $T = 134$, happens to be $s.e. = 0.102$ in all cases. The estimated values of d are always in the non-stationary range. Notice that for 12 out of the 13 countries the value $d = 1$ is included in an appropriate confidence interval, yielding similar results to those in Silverberg and Verspagen (2001). Nevertheless, using the AEFDF test with the above-mentioned estimated input value, $\widehat{d}_1$, the first column of Table 3 shows strong rejections of H_0: $d = 1$ in 6 out of the 13 countries.[7] The intuition for this higher rejection rate is the higher power of the EFDF test relatively to pure semiparametric tests which yield wider confidence intervals. Thus, our results for these countries seem to favor nonstationary, albeit mean-reverting, values of d, more in line with MZ. 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

considered by MZ.

[7] When the estimated value of d_1 was bigger than one, a value of $\hat{d}_1 = 1$ was employed to run the test.

memory versus structural breaks, and (ii) 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 4 about here]

5. CONCLUSIONS

This paper has developed simple regression statistics for detecting the presence of a unit root in time-series data against the alternative of mean-reverting fractional processes allowing for a wide variety of deterministic terms, $\mu(t)$, in the DGP by using a Wald test based on the EFDF testing approach. Three main findings have been obtained. *First*, if $\mu(t)$ is slowly evolving trend (including just a constant term), then the EFDF test ignoring $\mu(t)$ can be implemented without losing any of its optimal asymptotic properties. *Secondly*, if $\mu(t)$ is a polynomial in t of known order but unknown coefficients, then these properties remain if one runs the EFDF test on the residuals of the regression of Δy_t on $\mu(t)$ under the null of $d = 1$. *Thirdly,* we provide new theoretical results regarding the gains in power, under fixed alternatives, of applying the EFDF (and FDF) test instead of conventional LM tests. 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, yet mean-reverting, $I(d)$ process serves to illustrate the usefulness and simplicity of the testing approach proposed here.

Useful extensions of the present paper's setup that are under current investigation by the authors include testing fractional integration versus $I(0)$ allowing for structural breaks (see Dolado, Gonzalo and Mayoral, 2005), 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

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 (2003),"Long memory in the Spanish political opinion polls," *Journal of Applied Econometrics*, 18, 137-155

Dolado, J., Gonzalo, J. and L. Mayoral (2005), "Structural breaks vs. long memory: What is what?" Universidad Carlos III, Madrid, Mimeo.

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. and R. Joyeux (1980), "An introduction to long-range time series models and fractional integration," *Journal of Time Series Analysis*, 1, 15-30.

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 (2005), “Efficient Wald Tests for fractional Unit roots,” forthcoming, *Econometrica*.

Lobato, I. and C. Velasco (2006), “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,” forthcoming, *The Econometrics Journal*.

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.

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,” Mimeo.

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 1

In order to prove part (a), we consider first the case where $d_1 \in (0.5, 1)$ is a fixed number and then, the proof is extended to the stochastic case. In the general case where $\mu(t)$ is different from zero, the t-statistic on the coefficient φ from the simple regression of Δy_t on z_{t-1} is given by,

$$t_{\varphi}(d_1, \mu(t)) = \frac{\sum_{t=2}^{T} \Delta y_t z_{t-1}(d_1)}{\hat{S}_T(d_1)\sqrt{\sum_{t=2}^{T}(z_{t-1}(d_1))}}, \tag{A1}$$

where $\hat{S}_T^2(d_1) = T^{-1}\sum_{t=2}^{T}(\Delta y_t - \widehat{\varphi} z_{t-1}(d_1))^2$. We now show that the asymptotic distribution of $(A1)$ 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 (2005), we now prove that, for $d_1 \neq 1$,

$$t_{\varphi_{ols}}(d_1, \mu(t)) - t_{\varphi_{ols}}(d_1, \mu(t) \equiv 0) = o_p(1),$$

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 $t_{\varphi}(d_1, \mu(t))$, which is the numerator, since the analysis of the denominator is similar but simpler. Under H_0, the numerator of (A1), multiplied by $T^{-1/2}(1-d_1)^{-1}$, is given by,

$$T^{-1/2}(1-d_1)^{-1}\sum_{t=2}^{T}\Delta y_t z_{t-1}(d_1) = T^{-1/2}\sum_{t=2}^{T}(\Delta\mu(t) + \varepsilon_t)\left(\left(\Delta^{d_1} - \Delta\right)\mu(t) + \left(\Delta^{d_1-1} - 1\right)\varepsilon_t\right)$$

$$= T^{-1/2}\left(\sum_{t=2}^{T}\varepsilon_t\left(\Delta^{d_1-1} - 1\right)\varepsilon_t + \sum_{t=2}^{T}\left(\Delta\mu(t)(\Delta^{d_1} - \Delta)\mu(t)\right) + \right. \tag{A2}$$

$$\left.\sum_{t=2}^{T}\Delta\mu(t)\left(\Delta^{d-1} - 1\right)\varepsilon_t + \sum_{t=2}^{T}\varepsilon_t(\Delta^{d_1} - \Delta)\mu(t)\right). \tag{A3}$$

We now show that if $\mu(t) = t^{\delta}$, $\delta \in [0, 0.5)$ all the terms in (A2) and (A3) but the first, $\left(T^{-1/2}\sum_{t=2}^{T}\varepsilon_t\left(\Delta^{d_1-1} - 1\right)\varepsilon_t\right)$, 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 (A2) verifies that,

$$\begin{aligned} T^{-1/2}\left(\sum_{t=2}^{T}\Delta\mu(t)\,\Delta^{d_1}\mu(t) - \sum_{t=2}^{T}(\Delta\mu(t))^2\right) &\approx T^{-1/2}\left(\sum_{t=2}^{T}t^{2\delta-d_1-1} - \sum_{t=2}^{T}t^{2(\delta-1)}\right) \\ &= T^{-1/2}\left(O\left(T^{2\delta-d_1}\right) - O(1)\right) \to 0 \quad \text{(A4)} \end{aligned}$$

if $d_1 > 0.5$ and $\delta < 0.5$.

With respect to the first term in (A3),

$$T^{-1/2}E\left(\sum_{t=2}^{T}\Delta t^{\delta}\left(\Delta^{d_1-1}-1\right)\varepsilon_t\right) = 0, \quad \text{(A5)}$$

and

$$T^{-1}Var\left(\sum_{t=2}^{T}\Delta t^{\delta}\left(\Delta^{d_1-1}-1\right)\varepsilon_t\right) \approx T^{-1}\left(\sigma_{\varepsilon}^2 + \sigma_{\Delta^{d-1}\varepsilon}^2\right)\sum_{t=2}^{T}t^{2(\delta-1)} \to 0, \quad \text{(A6)}$$

where $\sigma^2_{\Delta^{d_1-1}\varepsilon}$ denotes the variance of the stationary fractionally integrated process $\Delta^{d_1-1}\varepsilon_t$. Expressions (A5) and (A6) imply that $\sum_{t=2}^{T}\Delta t^{\delta}\left(\Delta^{d_1-1}-1\right)\varepsilon_t \xrightarrow{p} 0$. The same type of argument can be used to show that the second term in (A3) also converges to zero. Therefore, for $d_1 \neq 1$, it follows that

$$(1-d_1)^{-1}T^{-1/2}\sum_{t=2}^{T}\Delta y_t z_{t-1}(d_1) = (1-d_1)^{-1}T^{-1/2}\sum_{t=2}^{T}\varepsilon_t\left(\Delta^{d_1-1}-1\right)\varepsilon_t + o_p(1), \quad \text{(A7)}$$

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_1 is a fixed number $\in (0.5, 1)$. Considering an stochastic input for d_1 amounts to show that

$$t_{\varphi}(d_1, \mu(t)) - t_{\varphi_{ols}}\left(\hat{d}_1, \mu(t)\right) = o_p(1),$$

where $\hat{d}_1$ 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 estimated d_1 converge to zero.

Hence, the numerator of $t_{\varphi_{ols}}(d_1,\ \mu(t)) - t_{\varphi_{ols}}\left(\hat{d}_1,\ \mu(t)\right)$ can be written as

$$(d_1-1)^{-1}T^{-1/2}\left(\sum_{t=2}^{T}\varepsilon_t\left(\Delta^{d_1-1}-1\right)\varepsilon_t-\sum_{t=2}^{T}\varepsilon_t\left(\Delta^{\hat{d}_1-1}-1\right)\varepsilon_t\right)+o_p(1),$$

and LV (2005, 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 (2005). Finally, using the results in DGM and LV, it is straightforward to prove the consistency of the test under fixed alternatives.

Proof of Theorem 2

We start by analyzing the case where the input of z_{t-1}, d_1, is fixed. We now show that under $H_0 : d = 1$, $t_{\varphi}(d_1, \mu(t)=0) - t_{\varphi}(d_1, \hat{\mu}(t)) \xrightarrow{p} 0$, where in this case $t_{\varphi}(d_1, \hat{\mu}(t))$ is given by,

$$t_{\varphi}(d_1,\hat{\mu}(t))=\frac{\sum_{t=2}^{T}\widetilde{\Delta y_t}\widetilde{z_{t-1}}(d_1)}{\hat{S}_T(d_1)\sqrt{\sum_{t=2}^{T}\left(\widetilde{z_{t-1}}(d_1)\right)}},$$

where $\widetilde{\Delta y_t} = (\Delta y_t - \Delta\hat{\mu}(t))$, $\widetilde{z_{t-1}}(d_1) = (1-d_1)^{-1}\left(\Delta^{d_1-1}-1\right)(\Delta y_t - \Delta\hat{\mu}(t))$ and $\hat{S}_T^2(d_1) = T^{-1}\sum_{t=2}^{T}\left(\widetilde{\Delta y_t} - \hat{\varphi}\widetilde{z_{t-1}}(d_1)\right)^2$ and $\mu(t)$ satisfies condition B.

For simplicity, we consider the DGP

$$y_t = \alpha + \beta t + \Delta^{-d}\varepsilon_t, d \le 1, \tag{A8}$$

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 $(A8)$, $\hat{\beta}$ is a $T^{3/2-d}$-consistent estimator of β (see Hosking, 1996). As in Theorem 1, we analyze the numerator of the t-statistic for testing $\varphi = 0$ in (5) since the analysis of the denominator is similar but simpler.

The numerator of $t_{\varphi}(d_1, \hat{\mu}(t))$ multiplied by $(1-d_1)$ is given by,

$$T^{-1/2}(1-d_1)\sum_{t=2}^{T}\widetilde{\Delta y_t}\widetilde{z_{t-1}} = T^{-1/2}\sum_{t=2}^{T}\varepsilon_t\left(\Delta^{d_1-1}-1)\varepsilon_t\right)+T^{-1/2}A_t,$$

where

$$T^{-1/2}A_t = T^{-1/2}\left(\beta-\hat{\beta}\right)\left(\sum\left(\Delta^{d_1-1}-1)\varepsilon_t\right)+\left(\beta-\hat{\beta}\right)\sum_{t=2}^{T}\tau_t(d_1)+\left(\sum_{t=2}^{T}(\tau_t(d_1)-1)\varepsilon_t\right)\right),$$

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

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

The same strategy can be used to show that the denominator of $t_\varphi(d_1, \hat{\mu}(t))$ equals the denominator of $t_\varphi(d_1, \mu(t) = 0)$ plus some terms that go to zero in probability. This implies that $t_\varphi(d_1, \hat{\mu}(t)) \xrightarrow{w} N(0,1)$. When d_1 is replaced by a $\mathrm{T}^{\kappa}-$consistent estimator, with $\kappa > 0$, if $t_{\varphi_{ols}}(d_1, \hat{\mu}(t)) - t_{\varphi_{ols}}\left(\hat{d}_1, \hat{\mu}(t)\right) = o_p(1)$, then the asymptotic distribution corresponding to $t_\varphi\left(\hat{d}_1, \mu(t)\right)$ would be the same as that of $t_\varphi(d_1, \mu(t))$. Following the same steps as above, it is straight forward to show that $T^{-1/2} A_t\left(\hat{d}_1\right)$ tends to zero. Then, the numerator of $(1-d_1)\left(t_\varphi(d_1, \mu(t)) - t_\varphi\left(\hat{d}_1, \mu(t)\right)\right)$ can be written as,

$$(d_1 - 1)^{-1} T^{-1/2} \left(\sum_{t=2}^{T} \varepsilon_t \left(\Delta^{d_1 - 1} - 1\right) \varepsilon_t - \sum_{t=2}^{T} \varepsilon_t \left(\Delta^{\widehat{d}_1 - 1} - 1\right) \varepsilon_t \right) + o_p(1),$$

and LV (2005) have shown that this expression tends to zero. Similar results can be easily obtained for the denominator. Hence, $t_\varphi\left(\hat{d}_1, \hat{\mu}(t)\right) \xrightarrow{w} N(0,1)$.

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(2005). Likewise, using the results in DGM and LV,the proof of the consistency of the test under fixed alternatives is straightforward.■

Proof of Theorem 3

The proof of this theorem can be easily constructed along the lines of Appendix 2 in LV (2005) and Theorems 1 and 2 above and, therefore, is omitted.■

Proof of Theorem 4

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,

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

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 $\rho_j = [\Gamma(j+b)\,\Gamma(1-b)/(\Gamma(j-b+1)\,\Gamma(b))]$, which for large values of j can be approximated by $[\Gamma(1-b)/\Gamma(b)]j^{2b-1}$. 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

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

whereas the two terms in the denominator converge to

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

and

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

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

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

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

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

By the LLN, the numerator tends to $(d-1)\,\sigma^2$. With respect to the denominator, we have that $T^{-1}\sum(\Delta y_t)^2 \xrightarrow{p} \sigma^2\Gamma(3-2d)/(\Gamma(2-d))^2$ and $\hat{\phi} \xrightarrow{p} (d-1)$. Combining these results, yields

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

Finally, by the LLN the LM test defined in (10), 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 $I\left(d-1\right)$ process. Using the formula of the autocorrelations given above, yields

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

TABLES

TABLE 1

Size and Power of EFDF and LM tests, 5% Level, Local Alternatives

DGP: $y_t = \alpha + \beta t + \Delta^{-d}\varepsilon_t$, $d = 1 - \gamma/\sqrt{T}$; $\alpha = 1$; $\varepsilon_t \sim N(0,1)$

	$EFDF_\mu$ ($\beta = 0$)				$EFDF_\tau$ ($\beta = 1$)				LM_μ ($\beta = 0$)				$LM\tau$ ($\beta = 1$)			
γ/ T	25	50	100	400	25	50	100	400	25	50	100	400	25	50	100	400
0	0.07	0.06	0.06	0.05	0.08	0.08	0.08	0.07	0.03	0.03	0.04	0.06	0.03	0.03	0.04	0.06
0.5	0.14	0.16	0.16	0.16	0.17	0.22	0.33	0.24	0.03	0.07	0.11	0.14	0.04	0.07	0.11	0.14
1	0.27	0.32	0.35	0.36	0.22	0.33	0.49	0.44	0.05	0.163	0.24	0.32	0.05	0.14	0.21	0.29
2	0.62	0.78	0.79	0.81	0.39	0.57	0.84	0.86	0.17	0.426	0.62	0.76	0.15	0.37	0.54	0.67
5	0.80	0.98	0.99	1.00	0.65	0.86	1.00	1.00	0.73	0.98	0.99	1.00	0.73	0.98	0.99	1.00

TABLE 2

Size and Power(*) of AEFDF and LM Tests, 5% Level

DGP: $y_t = \alpha + \beta t + \Delta^{-d}\epsilon_t/\left(1 - \phi L\right)$; $d = 1 - \gamma/\sqrt{T}$; $\phi = 0.6$; $\alpha = 1$; $\varepsilon_t \sim N(0,1)$

	$EFDF_\mu$ ($\beta = 0$)		$EFDF_\tau$ ($\beta = 1$)		LM_μ ($\beta = 0$)		LM_τ ($\beta = 1$)	
γ/ T	100	400	100	400	100	400	100	400
0	0.05	0.05	0.05	0.05	0.05	0.05	0.05	0.05
0.5	0.07	0.07	0.07	0.06	0.06	0.05	0.05	0.06
1	0.09	0.10	0.09	0.09	0.06	0.06	0.05	0.07
2	0.13	0.20	0.12	0.18	0.08	0.13	0.06	0.13
5	0.44	0.61	0.43	0.50	0.43	0.49	0.26	0.40

(*) Size-adjusted power.

TABLE 3

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	0.022	0.06
UK	0.0143	0.003
USA	0.0186	0.005
Spain	0.0199	0.005
Sweden	0.0193	0.005

TABLE 4

AEFDF Test

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

Country	$t_{\varphi}(\widehat{d}_1)$	$\widehat{d}_1$ $(s.e. = 0.10)$
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	-1.67*	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.