

DOCUMENTOS DE ECONOMÍA Y FINANZAS INTERNACIONALES

Working Papers on International Economics and Finance

**DEFI 13-01
February 2013**

A non-linear approach with long range dependence based on Chebyshev polynomials

Juan Carlos Cuestas

Luis A. Gil-Alana



Asociación Española de Economía y Finanzas Internacionales
www.aeefi.com
ISSN: 1696-6376

A NON-LINEAR APPROACH WITH LONG RANGE DEPENDENCE BASED ON CHEBYSHEV POLYNOMIALS

Juan Carlos Cuestas*
University of Sheffield, UK

and

Luis A. Gil-Alana
University of Navarra, Pamplona, Spain

ABSTRACT

This paper examines the interaction between non-linear deterministic trends and long run dependence by means of employing Chebyshev time polynomials and assuming that the detrended series displays long memory with the pole or singularity in the spectrum occurring at one or more possibly non-zero frequencies. The combination of the non-linear structure with the long memory framework produces a model which is linear in parameters and therefore it permits the estimation of the deterministic terms by standard OLS-GLS methods. Moreover, the orthogonality property of Chebyshev's polynomials makes them especially attractive to approximate non-linear structures of data. We present a procedure which allows us to test (possibly fractional) orders of integration at various frequencies in the presence of the Chebyshev trends with no effect on the standard limit distribution of the method. Several Monte Carlo experiments are conducted and the results indicate that the method performs well, and an empirical application, using data of real exchange rates is also carried out at the end of the article.

Keywords: Chebyshev polynomials; long run dependence; fractional integration
JEL Classification: C22

*Corresponding author: Juan C. Cuestas, University of Sheffield, Department of Economics, 9 Mappin Street, Sheffield, S1 4 DT, United Kingdom, email: j.cuestas@sheffield.ac.uk, tel. +441142223410. The authors gratefully acknowledge Jenny Roberts and Kostas Mouratidis for comments on an earlier version. The usual disclaimer applies. Luis A. Gil-Alana gratefully acknowledges financial support from the Ministry of Education of Spain (ECO2011-2014 ECON Y FINANZAS, Spain) and from a Jeronimo de Ayanz project of the Government of Navarra. Juan Carlos Cuestas acknowledges the MINECO (Ministerio de Economía y Competitividad, Spain) research grant ECO2011-30260-C03-01.

1. Introduction

It is well known that a number of economic variables reflect cycles in their paths, which can be modeled as non-linear trends. However, non-linearities might make the estimation of models complex, in particular when trying to use functions which are not linear in parameters to approximate those trends. This paper deals with the analysis of long range dependence in the context of non-linear models. In particular, we employ the Chebyshev polynomials in time to describe the deterministic part of the model, and suppose that the detrended series displays long memory behavior. We use a general definition of long memory that allows the inclusion of one or more poles or singularities in the spectrum at various frequencies. Thus, we consider the standard case of $I(d, d > 0)$ behavior, but also other possibilities such as seasonal/cyclical long range dependence and multiple cyclical structures. This is particularly interesting for macroeconomic data with a high seasonal component or cyclical movement due to economic activity.

The main problem with the non-linear deterministic trends in the context of fractional integration is that the interaction of the two structures produces a model with a non-linear structure for the coefficients, implying that linear methods are invalid for the estimation of the parameters unless we impose certain a priori conditions on various coefficients (e.g., Caporale and Gil-Alana, 2007). Also, a misspecified deterministic component may affect the power of the tests for the order of integration of the variables (see Perron, 1989, amongst many others). Many authors such as Zivot and Andrews (1992), Lumsdaine and Papell (1997), Lee and Strazicich (2003) and Papell and Prodan (2006), *inter alia*, have proposed unit root tests incorporating structural breaks, so as to improve the performance of the tests. However, structural breaks may still not be a proper specification of the deterministic component. Changes can occur smoothly rather than suddenly. In this line, Ouliaris *et al.* (1989) proposed regular polynomials to

approximate deterministic components in the data generation process. However, as later pointed out by Bierens (1997), Chebyshev polynomials might be a better mathematical approximation of the time functions, since Chebyshev polynomials are bounded and orthogonal. Chebyshev polynomials are cosine functions of time, which according to Bierens (1997), can be very flexible to approximate deterministic trends. With respect to the long range dependence we use a very general framework that allows the incorporation of one or more integer or fractional orders of integration of arbitrary order anywhere on the unit circle in the complex plane. This will allow us the analysis of a great variety of model specifications, including for example seasonal and cyclical behaviors of any stationary or nonstationary degree. Also, given that the inference based on t-statistics remains valid under the fractional integration specification used, we propose a very simple way to choose the order of the Chebyshev polynomials based on a “general to specific” approach and using the statistical significance of the Chebyshev coefficients.

The structure of the paper is as follows: Section 2 describes the statistical model incorporating non-linear (Chebyshev) trends and long range dependence. Section 3 presents a testing procedure for the fractional differencing parameters that includes the estimation of the non-linear trend coefficients. Section 4 contains a simulation study. Section 5 is devoted to the empirical work that includes an application using real effective exchange rates for 40 industrialized countries, and its implications for the purchasing power parity (PPP) theory. Section 6 concludes the paper.

2. The statistical model

We consider the following model,

$$y_t = f(\theta; z_t) + x_t, \quad t = 1, 2, \dots, \quad (1)$$

where y_t is the observed time series, f is a general function that might be non-linear and depends on the unknown parameter vector of dimension m , θ , and z_t which is a vector of deterministic terms or weakly exogenous variables; finally, we suppose that the error term x_t can be described in terms of the following model,

$$\rho(L; d) x_t = u_t, \quad t = 1, 2, \dots, \quad (2)$$

with

$$\rho(L; d) = (1 - L)^{d_1} (1 + L)^{d_2} \prod_{j=3}^M (1 - 2 \cos w_r^{(j)} L + L^2)^{d_j} \quad (3)$$

and u_t assumed to be $I(0)$. For the purpose of the present work we define an $I(0)$ process as a covariance stationary process with a spectral density function that is positive and bounded at all frequencies in the spectrum. Thus, it includes for u_t in (2) stationary and invertible autoregressive and moving average (ARMA) processes. Coming back to (3), L is the backshift operator (i.e., $Lx_t = x_{t-1}$) and d is an $(M \times 1)$ vector of real values containing the fractional differencing parameters that correspond to different poles or singularities in the spectrum. We observe that this is a very general specification that includes many cases of interest such as the standard $I(d)$ models (in case of $d_j = 0$ for all $j \neq 1$, and $d_1 = d$); cyclical fractional models based on Gegenbauer processes (when $d_j = 0$ for all $j \neq 3$); seasonal models ($M = 3$ with $w_r^{(3)} = \pi$), etc. (See Section 3.1 below).

Given the above set-up we focus on the estimation and testing of the unknown parameters corresponding to the vectors d and θ referring respectively to the differencing parameters and the non-linear deterministic trend coefficients.

The main problem we face with this set-up is the interaction between the equations (1) and (2) in the context of a non-linear function f , and in particular, between the long memory polynomial ρ and the non-linear function f . Under many circumstances the combination of the two produces a non-linear model in parameters, which hinders

the task of estimating the parameter vector θ . This problem can be solved by using the Chebyshev time polynomials.

The Chebyshev time polynomials $P_{i,T}(t)$ are defined by:

$$P_{0,T}(t) = 1,$$

$$P_{i,T}(t) = \sqrt{2} \cos(i\pi(t-0.5)/T), \quad t = 1, 2, \dots, T; \quad i = 1, 2, \dots \quad (4)$$

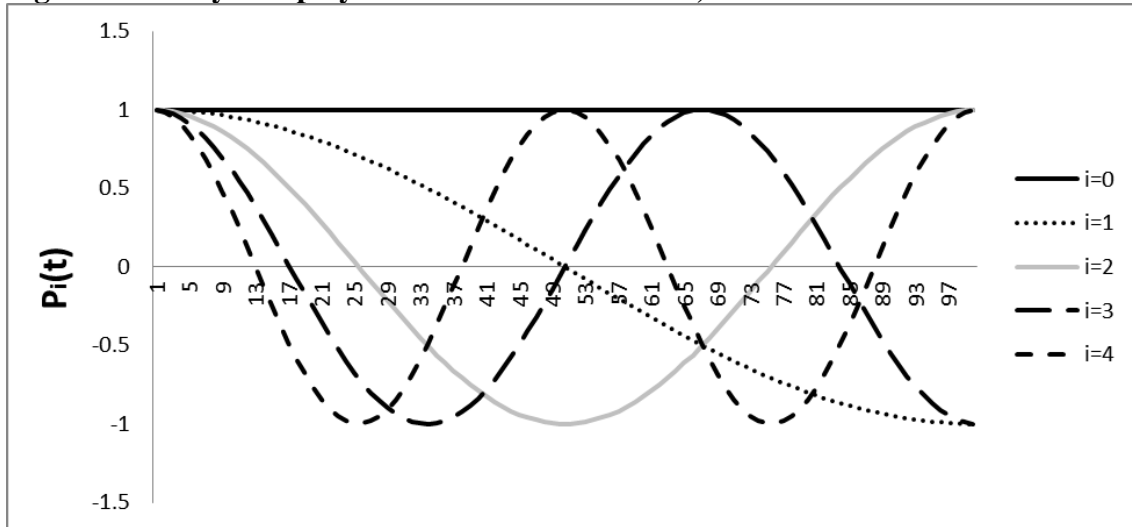
See Hamming (1973) and Smyth (1998) for a detailed description of these polynomials. Bierens (1997) uses them in the context of unit root testing. The latter author proposes several unit root tests, which account for a drift and a unit root under the null hypothesis, and stationarity around a linear or non-linear trend under the alternative. Hence, within the analysis of the order of integration of the variables, Bierens (1997) unit root tests, allow us to test whether the process is linear or non-linear trend stationary. In addition Bierens and Martins (2010) propose the use of Chebyshev polynomials in the framework of time-varying cointegrating parameters. There are several advantages in using these polynomials; first, since they are orthogonal, it avoids the problem of near collinearity in the regressors matrix in comparison with using regular time polynomials. Second, according to Bierens (1997) and Tomasevic *et al.* (2009), it is possible to approximate highly nonlinear trends with a rather low degree of polynomials. Finally, given their particular shape, there are good to approximate cyclical behavior. See Appendix A.1 for a brief discussion of the Chebyshev polynomials.

Figure 1, depicts the shape of the Chebyshev polynomials in (4) as a function of t , for different i .¹ Hence, for $i=0$, we have a drift, and for $i=1$ we have nearly a linear trend. For $i>2$, the trend mimics cycles, which increase their frequency as i increases.

¹ Note that the polynomials have been normalised to 1.

Given this type of shape, these polynomials are a good way to approximate economic cycles, which are present in many variables, with a rather low i .

Figure 1: Chebyshev polynomials as a function of t , for different i .



Across the present paper we employ Chebyshev polynomials to describe the deterministic trend. Thus, we can replace (1) by

$$y_t = \sum_{i=0}^m \theta_i P_{iT}(t) + x_t, \quad t = 1, 2, \dots, \quad (5)$$

with m indicating the order of the Chebyshev polynomial, and x_t following the model given by the equations (2) and (3). As earlier mentioned, if $m = 0$ the model contains an intercept, if $m = 1$ it adds a linear trend, and if $m > 1$ the model becomes non-linear, and the higher m is the less linear the approximated deterministic component becomes. Any combination of values for i might be valid to approximate the deterministic component of the data generation process (DGP). Given the shape of these polynomials, one can approximate economic cycles with a rather low m , releasing degrees of freedom. An issue that immediately arises here is the determination of the optimal choice for m . However, as will be argued below, standard t -statistics will remain valid under the specification given by (5), (2) and (3) noting that the error term is $I(0)$ by definition. The choice of m will, then, depend on the significance of the Chebyshev coefficients in

the joint specification of the model, for a particular choice of the (possibly *ARMA*) model selected of the $I(0)$ disturbances.

3. The procedure

The method proposed in this paper is a slight modification of Robinson (1994). He considers the same set-up as in (1) and (2) with f in (1) based exclusively on the linear form $\theta' z_t$, testing the null hypothesis:

$$H_o : d = d_o, \quad (6)$$

for any real vector value d_o . Under H_o , and using the two equations,

$$y_t^* = \theta' z_t^* + u_t, \quad t = 1, 2, \dots, \quad (7)$$

where $y_t^* = \rho(L; d_o) y_t$, $z_t^* = \rho(L; d_o) z_t$, and the symbol ' indicating transposition.

Then, given the linear structure of the above relationship and the $I(0)$ nature of the error term u_t , the coefficients in (7) can be estimated by standard ordinary least square/generalized least square (OLS/GLS) methods.² The same happens in our approach, whereby f contains the Chebyshev polynomials, noting that in spite of the non-linear structure, the relation is linear in parameters. Thus, combining equations (2) and (5) we get

$$y_t^* = \theta' P_T^*(t) = \sum_{i=0}^m \theta_i P_{iT}^*(t) + u_t, \quad t = 1, 2, \dots, \quad (8)$$

where

$$P_{iT}^*(t) = \rho(L; d_o) P_{iT}(t),$$

² Although Robinson (1994) focuses exclusively on the linear case, he argues (page 1421) that “ (...) undoubtedly a non-linear regression will also leave our limit distributions unchanged, under standard regularity conditions.” These conditions are briefly explained in Appendix A.2.

which can also be expressed in terms of (7) as in Robinson (1994) ($P_T^*(t) = z_t^*$), and then, using OLS/GLS methods, under the null hypothesis (6), the residuals are

$$\hat{u}_t = y_t^* - \sum_{i=0}^m \hat{\theta}_i P_{iT}^*(t);$$

with

$$\hat{\theta} = \left(\sum_{t=1}^T P_T^*(t) P_T^*(t)' \right)^{-1} \left(\sum_{t=1}^T P_T^*(t) y_t^* \right),$$

and $P_T^*(t)$ as the $(m \times 1)$ vector of transformed Chebyshev polynomials. Based on the above residuals \hat{u}_t , we estimate the variance,

$$\hat{\sigma}^2(\tau) = \frac{2\pi}{T} \sum_{j=1}^T g(\lambda_j; \hat{\tau})^{-1} I_{\hat{u}}(\lambda_j); \quad \lambda_j = 2\pi j/T, \quad (10)$$

where $I_{\hat{u}}(\lambda_j)$ is the periodogram of \hat{u}_t ; g is a function related with the spectral density of u_t (i.e., $s.d.f.(u_t) = (\sigma^2/2\pi)g(\lambda_j; \tau)$); and the nuisance parameter τ is estimated, for example, by $\hat{\tau} = \arg \min_{\tau \in T^*} \sigma^2(\tau)$, where T^* is a suitable subset of the R^q Euclidean space.

The test statistic, which is clearly based on Robinson (1994), for testing H_o (6) in (5), (2) and (3) uses the Lagrange Multiplier (LM) principle, and is given by

$$\hat{R} = \frac{T}{\hat{\sigma}^4} \hat{a}' \hat{A}^{-1} \hat{a}, \quad (11)$$

where T is the sample size, and

$$\hat{a} = \frac{-2\pi}{T} \sum_j^* \psi(\lambda_j) g(\lambda_j; \hat{\tau})^{-1} I_{\hat{u}}(\lambda_j),$$

$$\hat{A} = \frac{2}{T} \left(\sum_j^* \psi(\lambda_j) \psi(\lambda_j)' - \sum_j^* \psi(\lambda_j) \hat{\varepsilon}(\lambda_j)' \left(\sum_j^* \hat{\varepsilon}(\lambda_j) \hat{\varepsilon}(\lambda_j)' \right) \sum_j^* \hat{\varepsilon}(\lambda_j) \psi(\lambda_j)' \right)$$

with

$$\psi(\lambda_j) = \operatorname{Re} \left(\frac{\partial}{\partial d} \log \rho(e^{i\lambda_j}; d) \right), \quad \text{and} \quad \hat{\varepsilon}(\lambda_j) = \left. \frac{\partial}{\partial \tau} \log g(\lambda_j; \tau) \right|_{\tau = \hat{\tau}},$$

and the sum over * above refers to all the bounded discrete frequencies in the spectrum.

Under very mild regularity conditions,³ it can be shown that as in Robinson (1994):

$$\hat{R} \rightarrow_d \chi_M^2 \quad \text{as} \quad T \rightarrow \infty, \quad (12)$$

and, based on Gaussianity of u_t , it can also be shown the Pitman efficiency theory of the test against local departures from the null. That means that if we direct the test against local alternatives of form:

$$H_a : d = d_o + \delta T^{-1/2},$$

where δ is a non-null parameter vector, $\hat{R} \rightarrow_d \chi_M^2(\Lambda)$ as $T \rightarrow \infty$, indicating a non-central chi-squared distribution with non-centrality parameter which is optimal under Gaussianity of u_t .

3.1 Simple particular cases

In this section, we simplify the functional form of the above test statistic for some particular cases of interest.

a) White noise u_t

If we suppose that the disturbances are white noise, then, the spectral density function of u_t is simply $\sigma^2/2\pi$, and therefore, $g \equiv I$. Also, $\hat{\varepsilon}(\lambda_j) = 0$. Then, \hat{R} in (11) simplifies, with

$$\hat{a} = \frac{-2\pi}{T} \sum_j^* \psi(\lambda_j) I_{\hat{a}}(\lambda_j), \quad \text{and} \quad \hat{A} = \frac{2}{T} \left(\sum_j^* \psi(\lambda_j) \psi(\lambda_j)' \right).$$

³ These conditions only include moments up to a second order. (See Appendix A.2).

b) The case of the standard $I(d)$ model

A very standard case examined in the literature is the one corresponding to $\rho(L;d) = (1-L)^d$. These processes are called fractionally integrated or $I(d)$; they were introduced by Granger (1980), Granger and Joyeux (1981) and Hosking (1981), and have been widely employed in empirical works in the last twenty years to describe the dynamics of many economic and financial time series (Diebold and Rudebusch, 1989; Sowell, 1992; Gil-Alana and Robinson, 1997; etc.).⁴

In this context, $M = 1$, and the test statistic also substantially simplifies, with $\psi(\lambda_j) = \log \left| 2 \sin \frac{\lambda_j}{2} \right|$, implying that

$$\hat{A} = \frac{2}{T} \sum_{j=1}^{T-1} \left(\log \left| 2 \sin \frac{\lambda_j}{2} \right| \right)^2,$$

which can be asymptotically approximated by $\pi^2/6$.

Note here that the fact that d can be any real value allows us to consider the case of $I(1)$ processes, but also other structures such as the “ $1/f$ noise” model ($d = 1/2$), widely employed in many areas such as physics, hydrology and traffic flow (Marinari, Parisi, Ruelle and Windey, 1983; Wolf, 1997; Ninness, 1998; Eliazar and Klafter, 2010), and the “ $1/f^{d/2}$ noise” ($d = 1/4$) examined for example in Fox and Taqqu (1985) and Kaulakys (2000).

c) The case of a cyclical $I(d)$ model

In the previous case, the spectral density function is unbounded at the long run or zero frequency. However, the pole or singularity in the spectrum may occur at a non-zero

⁴ Earlier Robinson (1978) and Granger (1980) theoretically justified these processes by means of aggregation of heterogeneous processes.

frequency. In such a case we can consider $\rho(L; d) = (1 - 2\cos w_r L + L^2)^d$, with $w_r = 2\pi r/T$, $r = T/s$, and thus s will indicate the number of time periods per cycle, while r refers to the frequency that has a pole or singularity in the spectrum of the series. Gray *et al.* (1989, 1994) showed that this polynomial can be expressed in terms of the Gegenbauer polynomial, such that, denoting $\mu = \cos w_r$, for all $d \neq 0$,

$$(1 - 2\mu L + L^2)^{-d} = \sum_{j=0}^{\infty} C_{j,d}(\mu) L^j,$$

where $C_{j,d}(\mu)$ are orthogonal Gegenbauer polynomial coefficients defined recursively as:

$$C_{0,d}(\mu) = 1, \quad C_{1,d}(\mu) = 2\mu d,$$

$$C_{j,d}(\mu) = 2\mu \left(\frac{d-1}{j} + 1 \right) C_{j-1,d}(\mu) - \left(2 \frac{d-1}{j} + 1 \right) C_{j-2,d}(\mu), \quad j = 2, 3, \dots,$$

(see Magnus *et al.*, 1966, Rainville, 1960, etc. for further details on Gegenbauer polynomials). This type of process was introduced by Andel (1986) and subsequently analysed by Gray, Zhang and Woodward (1989, 1994), Chung (1996a,b), Gil-Alana (2001) and Dalla and Hidalgo (2005) among many others.

In this case, M is also equal to 1, and

$$\psi(\lambda_j) = \log (\cos \lambda_j - \cos w_r).$$

d) The case of multiple cycles

We can also study the case of processes that contain multiple poles or singularities in the spectrum. In these cases, $\rho(L; d) = \prod_{u=1}^M (1 - 2\cos w_r^{(u)} L + L^2)^{d_u}$. These processes

were introduced by Giraitis and Leipus (1995), Woodward *et al.* (1998), Ferrara and

Guegan (2001), and Sadek and Khotanzad (2004) among others. One special case here is the seasonal $I(d)$ model that, using a very simple specification may be expressed as

$$(1 - L^s)^d x_t = u_t, \quad t = 1, 2, \dots,$$

s indicating the number of time periods per year. Thus, for example, for quarterly data, $s = 4$, and it is a particular case of d with $M = 3$, and $w_r^{(u)} = 0, \pi/2$ and π respectively for $(u) = 1, 2$ and 3 . These processes were introduced by Porter-Hudak (1990) and have been subsequently examined by Ray (1993), Sutcliffe (1994) and Gil-Alana and Robinson (2001) and others.

If $s = 4$ and $\rho(L; d) = (1 - L^4)^d$, then $M = 1$,⁵ and $\psi(\lambda_j)$ becomes:

$$\psi(\lambda_j) = \log \left| 2 \sin \frac{\lambda_j}{2} \right| + \log \left(2 \cos \frac{\lambda_j}{2} \right) + \log |2 \cos \lambda_j|,$$

and allowing for a greater degree of generality, we can consider the case of different orders of integration at each frequency, so that $\rho(L; d) = (1-L)^{d_1} (1+L)^{d_2} (1-L^2)^{d_3}$. In this case, $M = 3$ and $\psi(\lambda_j)$ becomes a (3×1) vector of form:

$$\left(\log \left| 2 \sin \frac{\lambda_j}{2} \right|; \log \left(2 \cos \frac{\lambda_j}{2} \right); \log |2 \cos \lambda_j| \right)'$$

(See Gil-Alana and Robinson, 2001, Ooms and Franses, 2001).

e) **The case of Bloomfield (1973) disturbances**

Finally, we can suppose that the disturbances u_t follow a non-parametric approach due to Bloomfield (1973). This model does not provide an explicit formula for the error term, but it is implicitly determined by its spectral density function, which is given by

⁵ Note that M refers to the dimension of d , while m indicates the order of the Chebyshev polynomials.

$$f(\lambda_j; \tau) = \frac{\sigma^2}{2\pi} \exp\left(2 \sum_{r=1}^N \tau_r \cos(\lambda_j r)\right), \quad (13)$$

where N indicates the number of parameters required to describe the short run dynamics. Bloomfield (1973) showed that the logarithm of an estimated spectral density function is often found to be a fairly well behaved function and thus can be approximated by a truncated Fourier series. He showed that (13) approximates the spectral density of an $ARMA(p, q)$ process well when p and q are small values, which is usually the case for most economic time series. Like the stationary AR model, this has exponentially decaying autocorrelations and thus, using this specification, one does not need to rely on as many parameters as in the case of $ARMA$ processes. Moreover, it accommodates extremely well in the context of the testing procedure presented above. Thus, formulae for Newton-type iterations for estimating the τ_j are very simple (involving no matrix inversion), updating formulae when N is increased is also simple, and we can replace \hat{A} in the functional form of the test statistic in (11) by the population quantity:

$$\sum_{l=N-1}^{\infty} l^{-2} = \frac{\pi^2}{6} - \sum_{l=1}^N l^{-2},$$

which indeed is constant with respect to the τ_j .⁶

Other simple specifications can also be obtained from the general formula in (11). If u_t is $ARMA(p, q)$ of the form:

⁶ See Gil-Alana (2004) for an explanation of the accommodation of the model of Bloomfield (1973) in the context of fractional integration, and more in particular, in the context of the tests of Robinson (1994).

$$\left(1 - \sum_{k=1}^p \phi_k L^k\right) u_t = \left(1 + \sum_{k=1}^q \theta_k L^k\right) \varepsilon_t, \quad t = 1, 2, \dots,$$

then, the function g becomes:

$$\left| \frac{\left(1 + \sum_{k=1}^q \theta_k e^{i\lambda_j k}\right)}{\left(1 - \sum_{k=1}^p \phi_k e^{i\lambda_j k}\right)} \right|^2,$$

and in the simplest $AR(p)$ case,

$$g(\lambda_j; \tau) = \left| 1 - \sum_{k=1}^p \phi_k e^{i\lambda_j k} \right|^{-2},$$

and thus, implying that the l^{th} element of $\varepsilon(\lambda_j)$ becomes:

$$\varepsilon(\lambda_j) = \left(2 \left[\cos(l \lambda_j) - \sum_{k=1}^p \phi_k \cos((l-k)\lambda_j) \right] g(\lambda_j; \tau) \right).$$

4. A simulation experiment

In this section we briefly examine the finite sample behavior of some simple versions of the tests by means of Monte Carlo simulations. All calculations were carried out using Fortran and the programs are available from the authors upon request. Given the variety of cases and the number of possibilities covered by the tests, we concentrate on some simple cases, widely employed in the literature such as the case of standard $I(d)$ processes with the singularity or pole in the spectrum occurring at the long run or zero frequency. In particular, we consider the following DGP:

$$y_t = \sum_{i=0}^m \theta_i P_{iT}(t) + x_t, \quad (1 - L)^d x_t = u_t, \quad (14)$$

with $m = 3$ to justify some degree of non-linear behavior, and u_t as a white noise process with mean zero and variance 1. Also, for simplicity, we suppose that $\theta_i = 1$ for

all i , and take d in (14) equal to 0, 0.25, 0.50, 0.75 and 1, thus, including stationary and nonstationary hypotheses. Note that one additional advantage of our testing approach is that it is valid for any real fractional differencing parameter d , including thus stationary ($d < 0.5$) and nonstationary ($d \geq 0.5$) hypotheses. We generate Gaussian series using the routines GASDEV and RAN3 of Press, Flannery, Teukolsky and Vetterling (1986), for different sample sizes $T = 50, 100, 300$ and 500 , taking 10,000 replications for each case, and present the results for a nominal size of 5%.

Based on the model given by (14) we test the null hypothesis (6) for different d_o -values. However, noting that in this context $M = 1$, we can consider one-sided alternatives such as $H_a: d > d_o$ or $d < d_o$, and then, consider the test statistic:

$$\hat{r} = \sqrt{\hat{R}} = \frac{\sqrt{T} \hat{a}}{\sqrt{\hat{A}} \hat{\sigma}^2}, \quad (15)$$

which is asymptotically distributed as

$$\hat{r} \rightarrow_d N(0,1) \quad \text{as } T \rightarrow \infty, \quad (16)$$

See Robinson (1994). Thus, an approximate one-sided $100\alpha\%$ -level of (6) against the alternative $d > d_o$ is given by the rule:

$$\text{“Reject } H_o \text{ if } \hat{r} > z_\alpha \text{”},$$

where the probability that a standard normal variate exceeds z_α is α . In the same way, an approximate one-sided $100\alpha\%$ -level of (6) against the alternative $d < d_o$ is given by the rule:

$$\text{“Reject } H_o \text{ if } \hat{r} < -z_\alpha \text{”}.$$

We examine the size and the power properties of the test in the case of the model given by (14) with $d = 1$ and look in Table 1 at the rejection frequencies of \hat{r} in (15) with $d_o = 0, 0.25, 0.5, 0.75, 1, 1.25, 1.5, 1.75$ and 2 . Thus, the values corresponding to $d_o = 1$ will indicate the size of the test. We see in this table that the sizes of the tests are

clearly biased if the sample size is small. Thus, for example, if $T = 50$ and the tests are directed against $d > d_o$, the size is 0.018 ; however, when directed against $d < d_o$, it becomes much higher than the nominal size of 0.050 (0.109); however, as the sample size increases the values tend to approximate to the 5% level, which is consistent with the asymptotic nature of the tests. If we focus now on the rejection frequencies, we observe that the higher sizes observed in the case of $d < d_o$ also produce higher rejection probabilities in all cases compared with the case of alternatives with $d < 1$. Nevertheless, for departures higher than 0.5 even with small sample sizes, the tests behave fairly well, and if $T \geq 300$ the probabilities are very close to 1 in all cases. Note that the null consists of a unit root with Chebyshev time polynomials, so the test performs well even in strong nonstationary contexts. Performing the experiment with θ -coefficients different from 1 , and also with other values of d lead to essentially the same conclusions implying that the test performs relatively well if the sample size is large enough.

Table 1: Rejection frequencies against one-sided alternatives with Gaussian u_t

	d_o	$T = 50$	$T = 100$	$T = 300$	$T = 500$
$H_a: d > d_o$	0.00	0.788	0.907	1.000	1.000
	0.25	0.519	0.788	0.903	0.999
	0.50	0.308	0.554	0.702	0.945
	0.75	0.103	0.341	0.671	0.893
	1.00	0.018	0.027	0.039	0.047
$H_a: d < d_o$	1.00	0.109	0.088	0.075	0.056
	1.25	0.608	0.701	0.855	0.939
	1.50	0.771	0.886	0.996	0.998
	1.75	0.983	1.000	1.000	1.000
	2.00	1.000	1.000	1.000	1.000

Note: The nominal size is 5% . In bold the size of tests.

Table 2: Rejection frequencies against one-sided alternatives with t_3 -distributed u_t

	d_o	$T = 50$	$T = 100$	$T = 300$	$T = 500$
	0.00	0.793	0.914	1.000	1.000

$H_a: d > d_o$	0.25	0.520	0.793	0.955	1.000
	0.50	0.311	0.570	0.724	0.946
	0.75	0.107	0.344	0.683	0.894
	1.00	0.022	0.034	0.040	0.047
$H_a: d < d_o$	1.00	0.101	0.088	0.069	0.055
	1.25	0.603	0.693	0.831	0.917
	1.50	0.747	0.877	0.974	0.981
	1.75	0.979	0.992	1.000	1.000
	2.00	1.000	1.000	1.000	1.000

Note: The nominal size is 5%. In bold the size of tests.

Next we perform a similar experiment in non-Gaussian contexts. For this purpose, we examine the same null model as in Table 1 but assuming now that the disturbances are t -Student distributed with 3 degrees of freedom. This distribution is interesting because it just satisfies the second moment condition required in the test, its third moments not existing. The results, displayed in Table 2, are competitive with the Gaussian ones, with the sizes being closer to the nominal one of 5% in practically all cases. If we focus on the rejection frequencies, they tend to be slightly larger for values of $d_o < 1$, and lower when $d_o > 1$ compared with Table 1. Very similar results were obtained if weak autocorrelation ($AR(1)$ and $AR(2)$) is permitted for the $I(0)$ disturbances term, and the same applies for other values of d in (14).

5. An empirical application

In this section we apply the fractional integration tests developed in this paper to examine the mean reversion of real exchange rates and purchasing power parity (PPP). The absolute version of the PPP theory postulates that the price levels in two different countries should converge when measured in the same currency, so as to equalize the purchasing power of the currencies. This, therefore, implies that the real exchange rate,

defined as the ratio of prices in both places, translated to a common currency using the nominal exchange rate, should converge to 1. However, it is well known within the literature that the absolute version of the PPP hypothesis may be too restrictive. Hence, a less restrictive version of PPP is the relative PPP hypothesis, which implies that prices in common currency may converge to a constant different from 1. This relative version of the PPP implies then that what is actually expected in the long run is that the real exchange rate should be reverting to a constant, which may be different from 1. The intuition behind this is related to the fact that because of the existence of trade barriers, transport costs, and different measures of price indices, there may be a gap between price levels in different countries. Hence, on average, changes in real exchange rates should be zero, according to the relative version of the PPP theory.

In view of the above comments, testing for mean reversion becomes of paramount importance when testing for the empirical validity of the PPP theory, which at the same time, can be seen as a measure of the degree of over/under-valuation of the currencies, and it is used as a base for a number of macroeconomic models, i.e. the Dornbusch model. However, real exchange rate convergence, on average, to a constant along time may not be very realistic, in particular when countries experience different levels of economic growth and productivity gains, as well as, when countries suffer from changes in economic fundamentals, which may indeed change the equilibrium value of real exchange rates. For instance, the well known dynamic Penn effect and the Balassa-Samuelson effect, may induce deterministic trends in the data (see Lothian and Taylor, 2000, among others), and the existence of structural changes, may, in addition, induce changes in those trends. Hence, the importance of controlling for non-linear deterministic trends when testing for real exchange rate mean reversion. In a recent contribution, Cushman (2008) tests for the PPP hypothesis using the Bierens (1997) unit

root tests for bilateral exchange rates. He finds evidence to support that real exchange rates may in fact contain non-linear trends. However, it is not possible to test for the significance of these trends, unless the null is rejected.

Our newly developed fractional integration testing procedure, taking into account Chebyshev polynomials to approximate non-linear deterministic trends, solves these problems with the flexibility of having non-integer orders of integration. Given that the residuals of the auxiliary regression are $I(0)$ stationary by assumption, t -statistics are valid to test for the significance of the non-linear trends. This novelty solves the problem of choosing the order of the Chebyshev polynomials, which was not clearly defined by Bierens (1997). Thus, we can start from a fairly general degree of non-linearity (e.g. $m = 3$) and check the t -values of the estimated coefficients, removing those which are found statistically insignificant. We stop with the model with all significant coefficients.

The data used in the empirical application are real effective exchange rates against each country's 27 main trade partners, downloaded from *Eurostat* (code *ert_eff_ic_q*) for 40 countries, with different degrees of economic integration and development. We have used quarterly data from 1994:Q1 until 2011:Q3.

Across this section we consider the following model,

$$y_t = \sum_{i=0}^m \theta_i P_{iT}(t) + x_t, \quad (1 - L)^d x_t = u_t, \quad (17)$$

assuming that u_t is a white noise process.

Table 3 displays the estimates of d and the 95% confidence bands of the non-rejection values of d for the cases of $m = 0, 1, 2$ and 3. Higher values of m lead to non-significant coefficients for θ_i ($i > 3$) in all cases. These estimates were obtained using the Whittle function in the frequency domain and they coincide with the values of d_o that produce the lowest statistics in absolute value when using our (one-sided) testing

approach with a fine grid of d_o -values (with 0.001 increments). We observe in this table that the values

Table 3: Estimates of d based on white noise disturbances

Series	$m = 0$	$m = 1$	$m = 2$	$m = 3$
AUSTRIA	1.055 (0.907, 1.272)	1.048 (0.907, 1.260)	0.971 (0.786, 1.224)	0.929 (0.711, 1.197)
AUSTRALIA	1.205 (1.007, 1.501)	1.209 (1.021, 1.493)	1.202 (0.997, 1.497)	1.199 (0.984, 1.493)
BELGIUM	1.206 (1.069, 1.395)	1.203 (1.068, 1.391)	1.137 (0.955, 1.355)	1.123 (0.939, 1.337)
BRAZIL	1.114 (0.961, 1.356)	1.103 (0.952, 1.341)	1.029 (0.832, 1.317)	0.986 (0.746, 1.303)
BULGARY	0.914 (0.743, 1.261)	0.948 (0.737, 1.263)	0.947 (0.734, 1.266)	0.821 (0.451, 1.222)
CANADA	1.136 (0.895, 1.465)	1.133 (0.884, 1.462)	1.133 (0.882, 1.461)	1.121 (0.863, 1.461)
CHINA	1.179 (1.025, 1.426)	1.168 (1.022, 1.409)	1.160 (1.016, 1.407)	0.953 (0.694, 1.317)
CYPRUS	0.658 (0.568, 0.806)	0.602 (0.478, 0.784)	0.503 (0.347, 0.725)	0.429 (0.242, 0.688)
CZECK REP.	1.003 (0.806., 1.389)	1.049 (0.822, 1.392)	1.041 (0.802, 1.394)	0.972 (0.633, 1.384)
DENMARK	1.058 (0.861, 1.323)	1.062 (0.860, 1.327)	1.048 (0.847, 1.316)	1.048 (0.830, 1.322)
ESTONIA	1.439 (1.274, 1.681)	1.443 (1.293, 1.667)	1.443 (1.295, 1.673)	1.399 (1.252, 1.617)
FINLAND	1.202 (0.993, 1.495)	1.190 (0.974, 1.486)	1.179 (0.953, 1.486)	1.176 (0.951, 1.473)
FRANCE	0.907 (0.821, 1.043)	0.859 (0.748, 1.016)	0.721 (0.556, 0.933)	0.664 (0.453, 0.897)
GERMANY	1.072 (0.941, 1.255)	1.072 (0.940, 1.255)	0.985 (0.825, 1.200)	0.935 (0.747, 1.164)
GREECE	0.774 (0.661, 0.933)	0.800 (0.685, 0.954)	0.722 (0.569, 0.904)	0.701 (0.543, 0.897)
HONG-KONG	1.206 (1.067, 1.425)	1.187 (1.032, 1.414)	1.158 (1.002, 1.396)	0.987 (0.741, 1.293)
HUNGARY	0.909 (0.753, 1.344)	0.759 (0.427, 1.307)	0.755 (0.405, 1.307)	0.738 (0.384, 1.307)
IRELAND	1.195 (1.035, 1.437)	1.149 (0.972, 1.403)	1.148 (0.973, 1.404)	1.114 (0.907, 1.396)
ITALY	1.062 (0.833, 1.352)	1.062 (0.825, 1.349)	1.059 (0.817, 1.343)	1.050 (0.807, 1.331)
JAPAN	1.067 (0.904, 1.311)	1.056 (0.884, 1.305)	1.056 (0.881, 1.306)	1.027 (0.851, 1.283)
LATVIA	1.326 (1.167, 1.553)	1.336 (1.193, 1.554)	1.333 (1.192, 1.554)	1.193 (0.992, 1.462)
LITUANIA	1.146 (1.013, 1.322)	1.184 (1.081, 1.333)	1.146 (1.037, 1.306)	0.941 (0.766, 1.177)
MALTA	0.694 (0.584, 0.938)	0.738 (0.606, 0.975)	0.676 (0.523, 0.920)	0.309 (0.071, 0.693)

MEXICO	1.151 (1.003, 1.352)	1.150 (1.003, 1.351)	1.116 (0.952, 1.334)	1.097 (0.932, 1.308)
NETHERLANDS	1.088 (0.932, 1.304)	1.082 (0.924, 1.301)	1.081 (0.922, 1.303)	1.034 (0.861, 1.262)
NORWAY	0.944 (0.729, 1.222)	0.952 (0.741, 1.222)	0.949 (0.731, 1.229)	0.943 (0.722, 1.221)
NEW ZEELAND	1.274 (1.039, 1.624)	1.265 (1.044, 1.603)	1.264 (1.034, 1.603)	1.255 (1.033, 1.566)
POLAND	1.029 (0.804, 1.354)	1.034 (0.823, 1.346)	1.023 (0.796, 1.346)	0.997 (0.741, 1.331)
PORTUGAL	1.051 (0.872, 1.292)	1.039 (0.855, 1.288)	1.039 (0.853, 1.288)	0.984 (0.762, 1.244)
ROMANIA	1.145 (0.931, 1.433)	1.134 (0.917, 1.433)	1.133 (0.917, 1.435)	1.129 (0.906, 1.422)
RUSSIAN FED.	1.245 (1.022, 1.533)	1.249 (1.042, 1.553)	1.242 (1.027, 1.556)	1.242 (1.027, 1.554)
SOUTH KOREA	1.094 (0.861, 1.411)	1.096 (0.873, 1.417)	1.096 (0.876, 1.417)	1.083 (0.846, 1.407)
SLOVAKIA	1.137 (0.984, 1.417)	1.107 (0.944, 1.395)	0.992 (0.722, 1.366)	0.980 (0.692, 1.354)
SLOVENIA	1.342 (1.037, 1.755)	1.337 (1.072, 1.744)	1.342 (1.077, 1.711)	1.332 (1.054, 1.591)
SPAIN	0.907 (0.813, 1.047)	0.859 (0.744, 1.016)	0.721 (0.554, 0.933)	0.664 (0.459, 0.899)
SWEDEN	1.063 (0.854, 1.376)	1.044 (0.807, 1.377)	1.044 (0.803, 1.364)	1.042 (0.797, 1.382)
SWITZERLAND	1.252 (1.096, 1.463)	1.207 (1.076, 1.398)	1.208 (1.073, 1.384)	1.207 (1.066, 1.373)
TURKEY	0.824 (0.643, 1.308)	0.677 (0.318, 1.269)	0.678 (0.317, 1.263)	0.648 (0.207, 1.255)
U.K.	1.227 (1.082, 1.433)	1.228 (1.087, 1.444)	1.151 (0.956, 1.405)	1.132 (0.933, 1.388)
U.S.A	1.212 (1.047, 1.346)	1.213 (1.045, 1.467)	1.172 (0.983, 1.444)	1.119 (0.911, 1.393)

Note: In bold, evidence of mean reversion ($d < 1$). In brackets we display the confidence intervals at the 95%.

of d are very similar across the different values for m , in general, observing a slight reduction in the degree of integration as we increase m .⁷ We also notice that most of the estimates of d are within the unit root interval and some of them are even significantly above 1. The only evidence of mean reversion (i.e. d significantly below 1) is obtained for the cases of Cyprus, Greece and Malta (for all values of m) and for France and Spain

⁷ This might indicate a degree of competition between the non-linear structure due to the Chebyshev polynomials and the $I(d)$ framework in describing the structure of the series.

Table 4: Estimates of the Chebyshev polynomials in the case of $m = 3$

Series	θ_0	θ_1	θ_2	θ_3
AUSTRIA	99.016 (39.39)	2.373 (1.68)	1.575 (2.01)	0.707 (1.31)
AUSTRALIA	107.251 (3.10)	-3.701 (-0.17)	4.266 (0.49)	-1.872 (-0.35)
BELGIUM	99.301 (20.94)	0.610 (0.21)	2.286 (1.81)	0.665 (0.83)
BRAZIL	102.362 (3.09)	6.446 (0.33)	18.811 (1.90)	-8.191 (-1.23)
BULGARY	114.558 (11.12)	-27.761 (-4.75)	-1.164 (-0.32)	-5.309 (-2.06)
CANADA	126.198 (4.82)	-7.224 (-0.45)	-1.048 (-0.15)	-2.934 (-0.66)
CHINA	90.971 (6.95)	0.391 (0.05)	-3.134 (-0.77)	-9.764 (-3.56)
CYPRUS	104.742 (130.10)	-2.909 (-5.77)	1.089 (2.60)	0.613 (1.70)
CZECK REP.	115.654 (9.51)	-21.675 (-3.03)	1.994 (0.54)	-3.420 (-1.37)
DENMARK	99.238 (21.86)	-0.806 (-0.29)	0.802 (0.62)	0.020 (0.02)
ESTONIA	67.849 (2.65)	-0.254 (-0.38)	0.651 (0.12)	-2.962 (-0.96)
FINLAND	93.879 (9.22)	2.774 (0.43)	1.587 (0.61)	0.490 (0.30)
FRANCE	104.771 (91.91)	-4.643 (-7.21)	1.555 (3.40)	0.545 (1.67)
GERMANY	97.290 (26.80)	4.052 (1.91)	2.630 (2.31)	1.188 (1.72)
GREECE	99.848 (43.86)	-3.694 (-2.88)	1.791 (2.03)	-0.690 (-1.02)
HONG-KONG	87.038 (6.70)	12.242 (1.90)	-6.114 (-1.87)	-8.055 (-3.09)
HUNGARY	120.065 (18.02)	-17.302 (-4.61)	-0.560 (-0.22)	1.031 (0.55)
IRELAND	107.079 (13.75)	-5.871 (-1.23)	0.080 (0.03)	1.548 (1.16)
ITALY	103.153 (13.02)	-2.892 (-0.60)	-0.518 (-0.23)	-0.810 (-0.55)
JAPAN	102.040 (4.28)	10.698 (0.75)	-0.514 (-0.07)	-4.812 (-1.06)
LATVIA	92.662 (5.18)	-13.363 (-1.20)	-1.255 (-0.28)	-7.690 (-2.78)
LITUANIA	99.697 (11.62)	-21.677 (-4.33)	-6.914 (-2.58)	-7.910 (-4.32)
MALTA	101.467 (136.29)	-5.077 (-9.78)	-2.017 (-4.42)	-2.416 (-5.88)
MEXICO	141.923 (3.47)	-2.875 (-0.11)	-13.782 (-1.24)	-7.252 (-1.01)
NETHERLANDS	102.410 (21.65)	-1.291 (-0.45)	0.254 (0.18)	1.339 (1.50)

NORWAY	104.085 (11.81)	-1.934 (-0.37)	0.741 (0.27)	-0.730 (-0.39)
NEW ZEELAND	87.099 (2.01)	7.085 (0.26)	2.264 (0.22)	3.277 (0.52)
POLAND	108.376 (5.03)	-11.451 (-0.90)	-3.565 (-0.55)	-3.853 (-0.90)
PORTUGAL	101.431 (34.79)	-3.058 (-1.78)	-0.064 (-0.07)	0.906 (1.54)
ROMANIA	119.977 (3.29)	-22.595 (-1.01)	-1.933 (-0.20)	-1.976 (-0.32)
RUSSIAN FED.	99.158 (0.93)	-13.008 (-0.19)	12.455 (0.49)	-1.812 (-0.11)
SOUTH KOREA	110.881 (3.08)	6.752 (0.31)	-1.080 (-0.10)	3.805 (0.57)
SLOVAKIA	131.870 (10.30)	-34.771 (-4.61)	7.797 (2.02)	-1.482 (-0.57)
SLOVENIA	86.649 (5.25)	-0.551 (-0.05)	0.370 (0.09)	-0.693 (-0.32)
SPAIN	104.771 (91.91)	-4.643 (-7.21)	1.554 (3.40)	0.545 (1.73)
SWEDEN	95.991 (8.10)	4.633 (0.65)	0.266 (0.07)	-0.599 (-0.27)
SWITZERLAND	93.915 (4.87)	8.133 (0.67)	0.222 (0.04)	-0.632 (-0.21)
TURKEY	111.308 (11.74)	-16.485 (-3.07)	0.170 (0.04)	-1.649 (-0.54)
U.K.	97.291 (5.72)	2.831 (0.27)	-8.440 (-1.88)	-2.380 (-0.84)
U.S.A	112.153 (4.76)	1.777 (0.12)	-8.525 (-1.35)	-5.725 (-1.43)

Note: In bold, significant coefficients at the 5% level. *T*-statistics are given in brackets.

if $m = 2$ or 3 , i.e. assuming the existence of non-linearities. The results from Table 3 also point out that it is possible to reduce the order of integration of the variable by increasing artificially the order of the Chebyshev polynomials, m . This is consistent with other works that show that fractional integration and nonlinearities are issues which are intimately related (Diebold and Inoue, 2001; Granger and Hyung, 2004; etc.). Given that, as aforementioned, inference based on *t*-statistics remain valid, this approach makes much easier the selection of the appropriate deterministic component.

Next we examine the deterministic terms in more detail, checking if the Chebyshev coefficients are statistically significant for the selected estimates of d . The results are presented in Table 4. We notice several cases where non-linearities are

present. Based on these significant terms, we selected the appropriate model for each series, and the summary of the results (based only on the significant Chebyshev coefficients at the 5% level) are reported in Table 5. We see that strong evidence of non-linearities (with the two non-linear coefficients statistically significantly different from zero) is obtained for the cases of Cyprus, France, Malta, Spain, Germany, Hong-Kong and Lithuania. In the first four cases, the unit root hypothesis is rejected in favour of mean reversion, while in the remaining three cases, though the estimated values of d are smaller than 1, the unit root cannot be rejected. Evidence of non-linearity with significant θ_2 -coefficient is observed for Austria, Greece and Slovakia, the unit root being rejected in favor of mean reversion in the case of Greece. Also, for some countries only one of the two non-linear coefficients is significant, such as China (with only θ_3 being statistically significant, and an estimate of d of 0.979) as well as Bulgaria and Latvia (with d equal to 0.827 and 1.197 respectively), and also, Belgium, Brazil and the UK (with θ_2 significant but not θ_3) and the unit root being not rejected. For the remaining cases, only an intercept or a linear trend is required.

We also conducted the analysis based on weakly autocorrelated errors. We tried both seasonal and non-seasonal autoregressions and the results, not displayed, indicate that though quantitatively there are some differences when computing the results based on autocorrelated errors qualitatively the same conclusions hold, since the number of cases corresponding to “mean reversion”, “unit roots” or “explosive roots” affect exactly to the same series as in the case of white noise errors.

Our results pinpoint a few economic insights. We first observe that in many cases structural breaks in the form of non-linear trends are present in the data. Second, for a number of countries, for instance the Czech Republic and Hungary, a linear trend is enough to approximate the data. This implies that the Balassa-Samuelson effect might

Table 5: Summary results based on the selected model for each series

Series	d	θ_0	θ_1	θ_2	θ_3
AUSTRIA	0.971 (UR)	99.839	2.464	1.578	---
AUSTRALIA	1.205 (AB)	105.363	---	---	---
BELGIUM	1.138 (UR)	101.068	---	2.307	---
BRAZIL	1.043 (UR)	99.724	---	18.879	---
BULGARIA	0.827 (UR)	113.038	-27.794	---	-5.316
CANADA	1.136 (UR)	110.355	---	---	---
CHINA	0.979 (UR)	87.019	---	---	-9.745
CYPRUS	0.429 (MR)	104.742	-2.909	1.089	0.613
CZECK REP.	1.049 (UR)	112.905	-21.855	---	---
DENMARK	1.058 (UR)	99.257	---	---	---
ESTONIA	1.439 (AB)	54.584	---	---	---
FINLAND	1.202 (UR)	100.749	---	---	---
FRANCE	0.664 (MR)	104.771	-4.643	1.555	0.545
GERMANY	0.935 (UR)	97.290	4.052	2.630	1.188
GREECE	0.722 (MR)	99.245	-3.901	1.781	---
HONG KONG	0.987 (UR)	87.038	12.242	-6.114	-8.055
HUNGARY	0.759 (UR)	120.406	-17.042	---	---
IRELAND	1.195 (AB)	101.113	---	---	---
ITALY	1.062 (UR)	97.160	---	---	---
JAPAN	1.067 (UR)	109.704	---	---	---
LATVIA	1.197 (AB)	90.952	-13.409	---	-7.681
LITHUANIA	0.941 (UR)	99.697	-21.677	-6.914	-7.910
MALTA	0.309 (MR)	101.46	-5.077	-2.017	-2.416
MEXICO	1.151 (AB)	108.404	---	---	---
NETHERLANDS	1.088 (UR)	102.809	---	---	---
NORWAY	0.944 (UR)	101.390	---	---	---
NEW ZEELAND	1.274 (AB)	105.200	---	---	---
POLAND	1.029 (UR)	81.641	---	---	---
PORTUGAL	1.039 (UR)	102.489	-2.993	---	---
ROMANIA	1.145 (UR)	82.272	---	---	---
RUSSIAN FED.	1.245 (AB)	95.927	---	---	---
SOUTH KOREA	1.094 (UR)	124.409	---	---	---
SLOVAKIA	0.992 (UR)	129.793	-34.799	7.812	---
SLOVENIA	1.342 (AB)	85.759	---	---	---
SPAIN	0.664 (MR)	104.771	-4.643	1.554	0.545
SWEDEN	1.063 (UR)	102.116	---	---	---
SWITZERLAND	1.252 (AB)	104.935	---	---	---
TURKEY	0.677 (UR)	110.839	-16.902	---	---
U.K.	1.152 (UR)	98.004	---	-8.481	---
U.S.A.	1.212 (AB)	94.615	---	---	---

Note: MR means Mean Reversion ($d < 1$) and AB refers to the cases where d is significantly greater than 1. UR means that it contains the unit root case (i.e. $d = 1$).

be present, which makes economic sense given the process of catching-up with Western Europe during the transition period from communism to market economies. Finally, that in all cases of mean reversion, it occurs along with structural breaks. Comparing our results to those by Cushman (2008), although the results are not directly comparable, we can say that we find evidence of mean reversion using a lower order for the Chebyshev polynomials.

5. Concluding comments

In this paper we have examined a model that incorporates Chebyshev polynomials in time in the context of long range dependence. For the latter we use a very general expression that permits us to examine stationary and nonstationary hypotheses with one or more unit or fractional degrees of integration with the singularities in the spectrum occurring at zero and non-zero frequencies. The main advantage of this model is that combining the two structures (non-linear Chebyshev polynomials and fractional integration) leads to a new model that is linear in parameters, permitting the estimation of the Chebyshev polynomials in a very simple way. Moreover, we describe a testing procedure, originally proposed by Robinson (1994) for the linear case that displays several advantages in the present context. Thus, it allows us to test any real vector value for the differencing parameters, including stationary and nonstationary hypotheses; the incorporation of the Chebyshev polynomials allows its estimation with a straightforward method, including the use of the significance of the coefficients throughout standard t -values. The limit distribution of the procedure is standard chi-

squared distributed, and several Monte Carlo experiments conducted in the paper show it performs well even with small samples. A small empirical application based on this approach and using real effective exchange rates is also conducted in the paper. In our empirical application for real exchange rates, we find that it is possible to reduce the order of the polynomials in comparison with previous studies and that non-linear trends tend to be significant when explaining the evolution of real exchange rates in a number of countries. This is the case in sixteen out of the forty countries analysed. Hence, we show that our proposed test could be an effective way to estimate trends/cycles along with the orders of integration of variables alike.

A Appendices

A.1 Chebyshev polynomials

The Chebyshev polynomials used by Bierens (1997), Bierens and Martins (2010), and in the present paper, are the ones of the “first kind”. They are a special case of the Gegenbauer polynomials. They are defined as:

$$T_n(x) = \cos(n \cos^{-1} x).$$

These polynomials satisfy the Weierstrass theorem and the minimax criterion, which basically establishes that the polynomials can be approximated at the set of zeros occurring when

$$x = \cos \left[\frac{\pi(i-0.5)}{n} \right]$$

for $i=0, 1, \dots, n$, defining the order of the polynomials, and they only exist in the interval $[-1, 1]$.

Chebyshev polynomials are orthogonal with respect to the weighting function $w(x)=(1-x^2)^{0.5}$, implying that

$$\int_{-1}^1 w(x)T_i(x)T_j(x)d(x) = \begin{cases} 0.5\pi\delta_{ij} & \text{For } i \neq 0, j \neq 0 \\ \pi & \text{For } i=j=0 \end{cases}$$

with δ_{ij} being the Kronecker delta. Given that the polynomials are orthogonal, the terms of any expansion of them will be independent. This is a nice property for regression modelling, since it eliminates any collinearity in the regressors matrix. This implies that the estimated coefficients of any nonlinear approximation using Chebyshev polynomials are independent on the order of the polynomials. Hence, the estimation of redundant parameters does not affect the rest. Additionally, the standard errors of the estimated parameters of an approximation using these polynomials are the same.

We define $f(x)$ as a function to be approximated by non-linear terms using regular polynomials,

$$f(x) = \sum_{n=0}^{\infty} \alpha_n x^n + \varepsilon.$$

One of the interesting properties for mathematical modelling is the fact that it is possible to reduce the degree of polynomials by means of “economization of power series”, which implies to rewrite the series of regular polynomials as Chebyshev polynomials,

$$\sum_{n=0}^{\infty} \alpha_n x^n = \sum_{n=0}^{\infty} \beta_n T_n(x),$$

for $n=0,1,2,\dots$. Since $T_n(x)$ belongs to the interval $[-1,1]$, the number of terms which can be eliminated from the approximation depends on β_n .

A.2 Classes of processes in Robinson (1994)

Theorems 1 – 4 in Robinson (1994) presents the null and local limit distributions of Robinson’s (1994) tests under white noise and weak autocorrelated disturbances. In these theorems he refers to classes F, G and H (and K and L), which are defined in what follows.

Class F:

F is the class of sequences $\{V_t, t = 0, \pm 1, \dots\}$ of stationary random variables satisfying $E(V_t | B_{t-1}) = 0$ and $E(V_t^2 | B_{t-1}) = \sigma^2$ almost surely, where $0 < \sigma^2 < \infty$ and B_t is the σ field of events generated by $V_s, s \leq t$.

This condition imposes a martingale difference assumption on the white noise u_t in (2) that is substantially weaker than the Gaussianity used in motivating the test statistic. In the case of autocorrelated errors, this condition is extended to Class K.

Class K:

K is the class of sequences satisfying

$$u_t = \sum_{j=0}^{\infty} \alpha_j \varepsilon_{t-j}, \quad t = 0, \pm 1, \dots, \text{ where } \{\varepsilon_t, t = 0, \pm 1, \dots\} \text{ in } F \text{ and } \sum_{j=1}^{\infty} j^{1/2} |\alpha_j| < \infty.$$

These two conditions are clearly satisfied in our set-up since u_t satisfies the same properties as in Robinson (1994).

Class G:

G is the class of $(k \times l)$ vector sequences $\{z_t, t = 0, \pm 1, \dots\}$ such that $z_t = 0$ for $t \leq 0$, and

$$D = \sum_{t=1}^T z_t^* z_t^{*T} \text{ is positive definite for sufficiently large } T.$$

This condition imposes a mild lack of multicollinearity assumption on z_t^* that is clearly satisfied by our specification with the Chebyshev polynomials in time, since they are orthogonal.

Class H:

H is the class of functions $\rho(z; d)$ such that $\rho(0; d) = 1$ for all d , and $\psi(\lambda)$ as given in Section 3 has finitely many poles $\rho_l, l = 1, \dots, r$, on $(-\pi, \pi]$ such that $\|\psi(\lambda)\|$ is monotonically increasing as $\lambda \rightarrow \rho_l$ and as $\lambda \rightarrow \rho_{l+}$ for $l = 1, \dots, r$, and there exist disjoint intervals $S_l, l = 1, \dots, r$, such that

$$\bigcup_{l=1}^r S_l \in (-\pi, \pi], \quad \rho_l \in S_l; \quad \rho_l \notin S_k \text{ for } l \neq k,$$

and

$$\sup_{\lambda \in S_k - (\rho_k - \delta, \rho_k + \delta)} \left\{ |\lambda - \rho_k| \left\| \psi(\lambda) - \psi\left(\lambda \pm \frac{1}{2}\delta\right) \right\| \right\} = O(\delta^\tau) \text{ as } \delta \rightarrow 0,$$

for $k = 1, \dots, r$ and some $\eta > 1/2$, where $\|x\|$ denotes the Euclidean norm. This is a technical condition that is required to justify approximating integrals by sums. The same happens with the Class L (not presented) and referred to the function g . These conditions are also satisfied in our set-up which is built up on the same function ρ as in Robinson (1994).

References

- Andel, J. (1986) Long memory time series models, *Kybernetika* 22, 105-123.
- Bierens, H.J. (1997) Testing the unit root with drift hypothesis against nonlinear trend stationarity with an application to the US price level and interest rate, *Journal of Econometrics* 81, 29-64.
- Bierens, H. J. and L. F. Martins (2012) Time varying cointegration, *Econometric Theory* 26, 1453-1490.
- Bloomfield, P. (1973) An exponential model in the spectrum of a scalar time series, *Biometrika* 60, 217-226.
- Caporale, G.M. and L.A. Gil-Alana, 2007, Non-linearities and fractional integration in the US unemployment rate, *Oxford Bulletin of Economics and Statistics* 69, 4, 544-608.
- Chung, C.-F. (1996a) A generalized fractionally integrated autoregressive moving-average process, *Journal of Time Series Analysis* 17, 111-140.
- Chung, C.-F. (1996b) Estimating a generalized long memory process, *Journal of Econometrics* 73, 237-259.
- Cushman, D. O. (2008) Real exchange rates may have nonlinear trends, *International Journal of Finance and Economics* 13, 158-173.
- Dalla, V. and J. Hidalgo (2005) A parametric bootstrap test for cycles, *Journal of Econometrics* 129, 219-261.
- Diebold, F.X. and A. Inoue (2001) Long memory and regime switching. *Journal of Econometrics* 105, 131-159.
- Diebold, F. X. and G. D. Rudebusch (1989) Long memory and persistence in aggregate output, *Journal of Monetary Economics* 24, 189-209.
- Eliazar, I. and J. Klafter (2010) Universal generation of $1/f$ noises, *Physical Review E*, 81, 021109.

Ferrara, L. and D. Guegan (2001), Forecasting with k-factor Gegenbauer processes: Theory and Applications, *Journal of Forecasting* 20, 581-601.

Fox, R. and M.S. Taqqu (1985) Noncentral limit Theorems for Quadratic Forms in random variables having long range dependence, *Annals of Probability* 13, 428-446.

Gil-Alana, L. A. (2001) Testing stochastic cycles in macroeconomic time series, *Journal of Time Series Analysis* 22, 411-430.

Gil-Alana, L. A. (2004) The use of Bloomfield (1973) model as an approximation to ARMA processes in the context of fractional integration, *Mathematical and Computer Modelling* 39, 429-436.

Gil-Alana, L. A. and P. M. Robinson (1997) Testing of unit roots and other nonstationary hypotheses in macroeconomic time series, *Journal of Econometrics* 80, 241-268.

Gil-Alana, L.A. and P.M. Robinson (2001) Testing of seasonal fractional integration in the UK and Japanese consumption and income, *Journal of Applied Econometrics* 16, 95-114.

Giraitis, L. and P. Leipus (1995) A generalized fractionally differencing approach in long memory modelling, *Lithuanian Mathematical Journal* 35, 65-81.

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

Granger, C.W.J. and N. Hyung (2004) Occasional structural breaks and long memory with an application to the S&P 500 absolute stock returns, *Journal of Empirical Finance* 11, 399-421.

Granger, C. W. J. and R. Joyeux (1980) An introduction to long memory time series and fractional differencing, *Journal of Econometrics* 16, 121-130.

- Gray, H. L., N. F. Zhang and W. A. Woodward (1989) On generalized fractional processes, *Journal of Time Series Analysis* 10, 233-257.
- Gray, H.L., N. F. Zhang and W. A. Woodward (1994) On generalized fractional processes. A correction, *Journal of Time Series Analysis* 15, 561-562.
- Hamming, R. W. (1973) *Numerical Methods for Scientists and Engineers*, Dover.
- Hosking, J. R. M. (1981) Fractional differencing, *Biometrika* 68, 165-176.
- Kaulakys, B. (2000) On the inherent origin of $1/f$ noise, *Lithuanian Journal of Physics* 40, 281-286.
- Lee, J. and M. C. Strazicich (2003) Minimum LM unit root test with two structural breaks, *Review of Economics and Statistics* 85, 1082-1089.
- Lothian, J. And M. P. Taylor (2000) Purchasing power parity over two centuries: strengthening the case for real exchange rate stability. A reply to Cuddington and Liang, *Journal of International Money and Finance* 19, 759-764.
- Lumsdaine, L. and D. Papell (1997) Multiple trend breaks and the unit-root hypothesis, *Review of Economics and Statistics* 79, 212-218.
- Magnus, W., F. Oberhettinger and R. P. Soni (1966) *Formulas and theorems for the special functions of mathematical physics*, Springer, Berlin.
- Marinari, E. G. Parisi, D. Ruelle and P. Windley (1983) On the interpretation of $1/f$ noise, *Communications in Mathematical Physics* 89, 1-12.
- Ninness, B. (1998) Estimation of $1/f$ noise, *IEEE Transactions of Information Theory* 44, 1, 32-46.
- Ooms, M. And P.H. Franses (2001), A seasonal periodic long memory model for monthly river flows, *Environmental Modelling and Software* 16, 559-569.

- Ouliaris, S., J. Y. Park and P. C. B. Phillips (1989) Testing for a unit root in the presence of a maintained trend. In Ray, B. (Ed.) *Advances in Econometrics and Modelling*. Kluwer, Dordrecht, 6-28.
- Papell, D. and R. Prodan (2006) Additional evidence of long-run purchasing power parity with restricted structural change, *Journal of Money, Credit and Banking* 38, 1329-1349.
- Perron, P. (1989) The great crash, the oil price shock and the unit root hypothesis, *Econometrica* 57, 1361-1402.
- Porter-Hudak, S. (1990) An application of the seasonal fractionally differenced model to the monetary aggregate, *Journal of the American Statistical Association* 85, 338-344.
- Press, W. H., B. P. Flannery, S. A. Teukolsky and W. T. Wetterling (1986) *Numerical recipes. The art of scientific computing*, Cambridge University Press, Cambridge.
- Rainville, E.D. (1960) *Special functions*, MacMillan, New York.
- Ray, B.K. (1993) Long range forecasting of IBM product revenues using a seasonal fractionally differenced ARMA model, *International Journal of Forecasting* 9, 255-269.
- Robinson, P.M. (1978) Statistical inference for a random coefficient autoregressive model, *Scandinavian Journal of Statistics*, 5, 163–168.
- Robinson, P.M. (1994) Efficient tests of nonstationary hypotheses, *Journal of the American Statistical Association* 89, 1420-1437.
- Sadek, N. and A. Khotanzad (2004) K-factor Gegenbauer ARMA process for network traffic simulation, *Computers and Communications* 2, 963-968.
- Sowell, F. (1992) Modelling long run behavior with the fractional ARIMA model, *Journal of Monetary Economics* 29, 277-302.
- Sutcliffe, A. (1994) Time series forecasting using fractional differencing, *Journal of Forecasting* 13, 383-393.

Tomasevic, N., M. Tomasevic and T. Stanivuk (2009) Regression analysis and approximation by means of Chebishev polynomials, *Informatologia*, 42, 166-172.

Wolf, M., (1997) $1/f$ noise in the distribution of prime numbers, *Physica A* 241, 493-499.

Woodward, W. A., Q. C. Cheng and H. L. Ray (1998) A k-factor Gamma long memory model, *Journal of Time Series Analysis* 19, 485-504.

Zivot, E. and D. W. K. Andrews (1992) Further evidence on the great crash, the oil price shock, and the unit root hypothesis, *Journal of Business and Economic Statistics* 10, 251-270.