Testing Fractional Unit Roots with Non-linear Smooth Break Approximations using Fourier functions

Gil-Alana, Luis A. and Yaya, OlaOluwa S

University of Navarra, Pamplona, Spain, University of Ibadan, Ibadan, Nigeria

16 November 2018

Testing Fractional Unit Roots with Non-linear Smooth Break Approximations using Fourier functions

Luis A. Gil-Alana[1,2]   OlaOluwa S. Yaya[3]
[1]Faculty of Economics and NCID, University of Navarra, Pamplona, Spain
[3]Economic and Financial Statistics Unit, Department of Statistics, University of Ibadan, Ibadan, Nigeria

Abstract

In this paper we present a testing procedure for fractional orders of integration in the context of non-linear terms approximated by Fourier functions. The procedure is a natural extension of the linear method proposed in Robinson (1994) and similar to the one proposed in Cuestas and Gil-Alana (2016) based on Chebyshev polynomials in time. The test statistic has an asymptotic standard normal distribution and several Monte Carlo experiments conducted in the paper show that it performs well in finite samples. Various applications using real life time series, such as US unemployment rates, US GNP and Purchasing Power Parity (PPP) of G7 countries are presented at the end of the paper.

Keyword: Fractional unit root; Chebyshev polynomial; Monte Carlo simulation; Nonlinearity; Smooth break; Fourier transform

JEL Classification: C15, C12, C22

Corresponding author: OlaOluwa S. Yaya
Dept of Statistics
University of Ibadan,
Ibadan, Nigeria

Email: os.yaya@ui.edu.ng; o.s.olaoluwa@gmail.com

2 The first-named author gratefully acknowledges financial support from the Ministerio de Economía y Competitividad (ECO2017-85503-R).

1. Introduction

The classical Dickey-Fuller (DF, Dickey and Fuller, 1979) unit root test has been the foundation for various unit root tests in the literature, recently. The DF test is developed on the assumption that the underlying time series process is linear and the hypothesis of a unit root [I(1)] is tested against the alternative of series being I(0) in such a linear time series framework. Meanwhile, in reality, it is uncommon to have linear time structured series, particularly in economic and financial research where time series often possess anomalies such as regime switch, breaks, jumps and heteroscedasticity. Thus, the classical DF unit root test (and other unit root tests such as Phillips and Perron, 1988; Elliot et al., 1996; etc.) lack power in this regard against, for example, trend stationary alternatives (De Jong et al., 1992), structural breaks (Campbell and Perron, 1996), regime switching (Nelson et al., 2001), etc. Moreover, the restriction of I(1)/I(0) alternatives is also too restrictive since time series could be I(d) where d is any real number, either an integer value or a fraction, and this explanation has more meaningful policy implications in economics and finance. For example, $d = 0$ implies (covariance) stationarity, $d = 1$ implies nonstationarity. For the fractional case, $-0.5 < d < 0.5$ means the series is stationary and invertible, $0 < d < 0.5$ implies the series is stationary mean reverting and possesses long memory, and lastly, $0.5 \leq d < 1$ is mean reverting nonstationary range for the series. These make fractional unit root testing more appealing compared to classical unit root testing. In fact, many authors such as Diebold and Rudebusch (1991); Hasslers and Wolters (1994) and Lee and Schmidt (1996) have shown the extremely low power of standard unit root methods under this type of alternative. Thus, there is the need to embrace fractional unit root testing in both linear and non-linear frameworks.

The first attempt to introduce fractional integration in DF-type models is in the fractional testing procedure of Dolado, Gonzalo and Mayoral (2002). The testing model does not include an augmentation component that controls the whitening of the random process in the model. Other non-linear unit root tests also applied the DF model. Due to the appealing interpretation of the results in the I(d) case, it is straightforward to combine the Fourier function in sine and cosine functions of time (see Enders and Lee, 2012a; 2012b) in a fractional testing framework. The Fourier function can capture breaks of unknown form even if the time series do not show any periodic behaviour. The approach works better than when using dummy variables for describing breaks, irrespective of whether the breaks are instantaneous or smooth. The polynomial allows for smooth trends similar to those of the smooth transition regression dynamics in Granger and Teräsvirta (1993). Thus, the realization of non-linear unit root testing will mimic the expectation of Leybourne, Newbold and Vougas (1998) and Kapetanios, Shin and Snell (2003), which are the first non-linear unit root methods in asymmetric and symmetric smooth transition regression frameworks, respectively.

Unit root tests based on smooth transition regression are expected to lack power since the implementation requires the estimation of a considerable number of autoregressive parameters. Also, the framework assumes a single gradual break with a known date and functional form (Enders and Lee, 2012a), though the break dates and the number of breaks are unlikely to be known.[1] Thus, such tests are not powerful, as many parameters need to be estimated. Also, Becker, Ender and Hurn (2004) stated that if the number of breaks in a time series is unknown, a test of structural change that uses even a single trigonometric component can have greater power than the Bai and Perron (2003) multiple structural breaks test.[2] Around such a single frequency, fractional frequencies can be obtained in order to obtain a good Fourier approximation just as in the Fractional Frequency Flexible Fourier Function (FFFFF) with multiple smooth breaks for the DF unit root test proposed in Omay (2015).

[1] Becker, Enders and Lee (2006), Enders and Lee (2012a,b) and Rodrigues and Taylor (2012) suggested using a single frequency component for the Fourier function in the detection of smooth breaks since higher frequencies lead to over-filtration.

[2] The unit root test proposed by Becker, Enders and Lee (2006) and Ender and Lee (2012a) is used when the break dates and the precise form of the breaks are unknown.

This author found that the small sample properties of the FFFFF-DF perform better than that of its integer frequency Fourier counterpart.

In this paper, we propose a fractional unit root test with non-linear smooth trend functions using the Fourier series approximation. The testing procedure is based on fractional integration, implementing Robinson's (1994) framework and extending it to the non-linear case for this particular approximation.

The rest of the paper is structured as follows: Section 2 describes the model and the test statistic. Section 3 is devoted to the analysis of the statistical properties of the test statistic in a finite sample experiment by using Monte Carlo simulations. Section 4 contains various empirical applications, while Section 5 concludes the paper.

2. Model and Test Statistic

2.1 Non-linear trends

We consider the model,

$$y_t = f(t) + x_t, \quad t = 1, 2, \ldots, \tag{1}$$

where y_t is the observed (univariate) time series, $f(t)$ is the smooth trend Fourier function or other smooth polynomials or functions in time t; x_t is the fractionally integrated process, given by,

$$(1-L)^d x_t = u_t, \quad t = 1, 2, \ldots \tag{2}$$

where d is the fractional unit root parameter; L is the lag operator (i.e., $L^k x_t = x_{t-k}$), and u_t is the white noise process, though it can be easily extended to any type of I(0) process, including, thus, the stationary and invertible ARMA-class of models. Initially, we assume that the function $f(t)$ is of the following form:

$$f(t) = \alpha + \beta t + \sum_{k=1}^{n} \lambda_k \sin(2\pi kt/T) + \sum_{k=1}^{n} \gamma_k \cos(2\pi kt/T); \quad n \le T/2; \; t = 1,2,... \tag{3}$$

where α is the intercept, and β is the coefficient of the linear trend, t; λ_k and γ_k measure, respectively, the amplitude and displacement of the sinusoidal component of the deterministic term; n is the number of frequencies in the approximation, k is a particular frequency set initially equal to 1, 2, …. n, and T is the number of observations. Clearly, if $\lambda_k = \gamma_k = 0$ for all k, the process is linear in time, and traditional fractional unit root tests apply here (see, Robinson, 1994; Tanaka, 1999; Dolado et al., 2002; etc.). The significance of at least one of (λ_k, γ_k) implies a non-linear unit root. The presence of many frequency components involves degrees of freedom which could lead to overfitting and loss of power if the correct value of k is not precisely determined. In this sense, the value of k may not necessarily be restricted to an integer value, and fractional numbers can also be considered as in Omay (2015) and Omay, Gupta and Bonaccolto (2017). Then, equation (3) can be replaced by:

$$f(t) = \alpha + \beta t + \sum_{k=1}^{n} \lambda_k \sin(2\pi j_k t / T) + \sum_{k=1}^{n} \lambda_k \cos(2\pi j_k t / T), \; n \le T/2, \quad t = 1,2,..., \tag{4}$$

where $j_k = 1, 2, \ldots, n$ can be fractional values and $j_k = k$ in the case of integer numbers. As noted in Becker, Enders and Hurn (2004), structural shifts can be captured by the low frequency components of a series since breaks shift the spectral density function towards zero, though the higher frequency components of the series are likely to be associated with parameter variation. Also, n might be large in order to be viewed as an application of Gallant's (1981) Flexible Fourier Form (FFF) to modelling *f(t)*.[3] In the case of a single frequency component, i.e., n = 1,

$$f(t) = \alpha + \beta t + \lambda_1 \sin(2\pi t/T) + \gamma_1 \cos(2\pi t/T); \quad n \le T/2; \; t = 1,2,... \tag{5}$$

[3] See other papers such as Davies (1987), Gallant and Sonza (1991) and Bierens (1997).

Similarly to Schmidt and Phillips (1992), it is straightforward to devise a score-based test statistic with limiting distribution that depends on the frequency, *k*. In the fractional case, to determine the best frequency, the data-driven procedure of Davies (1987) can be used. This is a grid search-regression method based on (5) where frequencies are set from say 0.2 to 5 with an increment of 0.2 as well. Then, we obtain a specific $j_l\left(i=1,\ldots,k\right)$ corresponding to the minimum sum of square residuals in (5). Since j_l is unknown in a real sense, it is suggested in Enders and Lee (2012a,b) that the precise value of j_l be estimated with the least squares method, and the estimator be used in the subsequent analysis as in the fractional integration framework in our own case.

As mentioned in Leybourne, Newbold and Vougas (1998) and Kapetanios, Shin and Snell (2003) and Enders and Lee (2012a), unit root tests based on Logistic Smooth Transition Autoregression (LSTAR) and Exponential Smooth Transition Autoregression (ESTAR) breaks lack power compared to smooth breaks induced by flexible Fourier polynomials. Also, Fourier polynomials allow for increments in the number of frequencies in order to control for higher non-linearities.

By using (3), (4) or (5) in the framework of fractional integration in (2) using Robinson (1994) described below, we can obtain, simultaneously the fractional unit root order of integration and the parameters of the smooth Fourier function, along with the rest of the parameters in the model.

2.2 The set-up in Robinson (1994)

The Robinson (1994) fractional integration framework considers the following model,

$$y_t = \beta^T z_t + x_t, \quad t=1,2,..., \tag{6}$$

$$\rho(L;\theta)x_t = u_t, \quad t=1,2,..., \tag{7}$$

where y_t is the observed time series; β is a (kx1) vector of unknown parameters, and z_t is a k-vector of exogenous regressors or deterministic terms that may include for example, a linear time trend in case of $z_t = (1, t)^T$; x_t is described by means of equation (7), where the function ρ adopts the following form:

$$\rho(L;\, d)\, x_t = (1 - L)^{d_1}\, (1 + L)^{d_2} \prod_{j=3}^{h} (1 - 2\cos w_r L + L^2)^{d_j} \tag{8}$$

for any given h, for any given distinct real numbers w_j, j = 3, …. h on the integral (0, π), and for any given real numbers, d_j, j = 1, 2, …., h.

The specification in (8) is a very general formulation that permits us to consider one or more integer or fractional roots of arbitrary order anywhere on the unit circle in the complex plane. If the (hx1) vector d is formed by integer values, we face with many models of interest, such as the standard unit root case (when h = 1 and $d_1 = 1$), the I(2) case as in Johansen et al. (2010) (if h = 1, and $d_1 = 2$), seasonal unit roots of the form advocated by Dickey, Hasza and Fuller (DHF, 1984) and Hyllerberg, Engle, Granger and Yoo (1990), (if h = 3, $w_r = 0$, $d_1 = d_2 = d_3 = 1$), cyclical I(1) behaviour such as in Bierens (2001) and Kunst (2001) (if $d_1 = d_2 = 0$, and $d_3 = 1$), etc. However, d can also include fractional values, and then, the function in (7) and (8) permits us to study the classical I(d) models as in Baillie (1996), Gil-Alana and Robinson (1997), Abritti et al., 2016, and others (if h = 1 and d is fractional); seasonal I(d) models (Gil-Alana and Robinson (2001), Gil-Alana, 2002; Bisognin and Lopes, 2009, etc.) (if h = 3 and d_1, d_2 and d_3 are fractional) and cyclical I(d) (Gil-Alana, 2001) (with $d_1 = d_2 = 0$, and d_3 fractional).[44]

In Robinson (1994), he tested the null hypothesis:

$$d = d_o \tag{9}$$

[44] Multiple (fractional) cyclical structures (Ferrara and Guegan, 2001, Sadek and Khotanzad, 2004, and Gil-Alana (2007) can also be examined if $d_1 = d_2 = 0$ with j > 3.

for any real-value vector d_o, in the set-up given by (6) – (8), and he showed that his test statistic follows asymptotically a χ^2_h-distribution. Moreover, he also showed that his test is the most efficient one in the Pitman sense against local departures of the null (see, Robinson, 1994).[55]

For the purpose of the present work, we simplify the model in Robinson (1994) and assume that h = 1 in (8), so that we have a scalar order of integration d. Moreover, in most of the empirical applications conducted based on Robinson (1994), z_t in (6) is supposed to be a linear time trend, i.e., $z_t = (1, t)^T$.[66] Thus, under the null hypothesis (9), Robinson's (1994) set-up becomes:

$$y_t = \beta_0 + \beta_1 t + x_t, \quad (1-L)^{d_o} x_t = u_t, \quad t = 1, 2, ..., \tag{10}$$

where the errors are assumed to be white noise or weakly autocorrelated processes in turn. The above two equations can be written in a single one as

$$y_t^* = \beta_0 1_t^* + \beta_1 t_t^* + u_t, \quad t = 1, 2, ..., \tag{11}$$

where $y_t^* = (1-L)^{d_o} y_t$: $1_t^* = (1-L)^{d_o} 1_t$: $t_t^* = (1-L)^{d_o} t_t$, and given that u_t is I(0) by construction, we can estimate β_0 and β_1 in (10) (or (11)) by standard OLS methods, such that the estimated residuals are then

$$\hat{u}_t = y_t^* - \hat{\beta}_0 1_t^* + \hat{\beta}_{11} t_t^*, \quad t = 1, 2, ...,$$

and the functional form of the test statistic is given by:

$$\hat{R} = \frac{T}{\hat{\sigma}^4} \hat{a}' \hat{A}^{-1} \hat{a}, \tag{12}$$

where T is the sample size, and

[55] That means that if the test is directed against local alternatives of the form: H_a: $d = d_o + \delta/(T)^{0.5}$, the limit distribution is $\chi^2(\omega)$ with a non-centrality parameter ω that is optimal under Gaussianity of u_t.

[66] See Gil-Alana (2005), Gil-Alana and Moreno (2012), Gil-Alana and Hjuibens (2018), etc.

$$\hat{a} = \frac{-2\pi}{T} \sum_j^{*} \psi(\lambda_j) g_u(\lambda_j;\hat{\tau})^{-1} I(\lambda_j) \quad ; \qquad \hat{\sigma}^2 = \sigma^2(\hat{\tau}) = \frac{2\pi}{T} \sum_{j=1}^{T-1} g_u(\lambda_j;\hat{\tau})^{-1} I(\lambda_j) \quad ,$$

$$\hat{A} = \frac{2}{T}\left(\sum_j^{*} \psi(\lambda_j)\psi(\lambda_j)' - \sum_j^{*} \psi(\lambda_j)\hat{\varepsilon}(\lambda_j)' \left(\sum_j^{*} \hat{\varepsilon}(\lambda_j)\hat{\varepsilon}(\lambda_j)' \right)^{-1} \sum_j^{*} \hat{\varepsilon}(\lambda_j)\psi(\lambda_j)' \right);$$

$$\psi(\lambda_j) = \log\left| 2 \sin\frac{\lambda_j}{2} \right|; \qquad \hat{\varepsilon}(\lambda_j) = \frac{\partial}{\partial \tau} \log g_u(\lambda_j;\hat{\tau}),$$

where $\lambda_j = 2\pi j/T$, and the summation in * in the above equations is over all frequencies which are bounded in the spectrum.[77] $I(\lambda_j)$ is the periodogram of $\hat{u}_t$, and $\hat{\tau} = \arg\min_{\tau \in T^*} \sigma^2(\tau)$, with T^* as a suitable subset of the R^q Euclidean space. Finally, g_u is a known function coming from the spectral density of u_t:

$$f_u(\lambda) = \frac{\sigma^2}{2\pi} g_u(\lambda;\tau), \qquad -\pi < \lambda \leq \pi.$$

Note that these tests are parametric and, therefore, they require specific modelling assumptions about the short-memory specification of u_t. In particular, if u_t is a white noise, $g_u \equiv 1$, whilst if it is an AR process of the form $\phi(L)u_t = \varepsilon_t$, $g_u = |\phi(e^{i\lambda})|^{-2}$, with $\sigma^2 = V(\varepsilon_t)$, with the AR coefficients being a function of τ.

In this context of h = 1, Robinson (1994) showed that:

$$\hat{R} \to_d \chi_1^2, \qquad as \quad T \to \infty, \tag{13}$$

where "$\to_d$" stands for convergence in distribution. Therefore, unlike in the case of other (unit root / fractional) procedures, this is a classical large-sample testing situation. If the test is carried out in the context of (10), the null H_o (9) will be rejected against the alternative H_a: $d \neq d_o$ if $\hat{R} > \chi_{1,\alpha}^2$, with Prob $(\chi_1^2 > \chi_{1,\alpha}^2) = \alpha$.

[77] For this particular version of the tests of Robinson (1994), the spectrum has the singularity at the zero frequency, so, j runs from 1 to T-1.

2.3 The extension to the non-linear case

In this section, we substitute the first equation in (10) by a non-linear deterministic form. In particular, we try (1) with the FFF approach of equation (3) (or equivalently its fractional formulation in (4)). Thus, the examined model under the null hypothesis (9) is now:

$$y_t = \alpha + \beta t + \sum_{k=1}^{n} \lambda_k \sin(2\pi k t / T) + \sum_{k=1}^{n} \lambda_k \cos(2\pi k t / T) + x_t, \qquad (1-L)^{d_o} x_t = u_t, \tag{14}$$

which becomes,

$$y_t^* = \alpha 1_t^* + \beta t_t^* + \sum_{k=1}^{n} \lambda_k \sin_{k,t}^* + \sum_{k=1}^{n} \lambda_k \cos_{k,t}^* + x_t, \qquad t = 1, 2, ..., \tag{15}$$

under the appropriate transformation, with $\sin_{k,t}^* = (1-L)^{d_o} \sin(2\pi k t / T)$ and $\cos_{k,t}^* = (1-L)^{d_o} \cos(2\pi k t / T)$. Thus, though the deterministic function in (4) is non-linear, it becomes linear in parameters in (15) and the coefficients can be estimated by standard methods. Thus, the only change with respect to (12) is the specification of the estimated residuals which are now:

$$\hat{u}_t = y_t - \hat{\alpha} 1_t^* + \hat{\beta} t_t^* + \sum_{k=1}^{n} \hat{\lambda}_k \sin_{k,t}^* + \sum_{k=1}^{n} \hat{\gamma}_k \cos_{k,t}^*, \qquad t = 1, 2, ..., \tag{16}$$

where $\hat{\theta}^T = \left(\hat{\alpha}, \hat{\beta}, \hat{\lambda}_1, ..., \hat{\lambda}_n, \hat{\gamma}_1, ..., \hat{\gamma}_n\right)^T$ is estimated by standard OLS methods, i.e.,

$$\hat{\theta} = \frac{\sum_{t=1}^{T} w_t^* y_t^*}{\sum_{t=1}^{T} w_t^{*2}},$$

with $w_t^* = (1_t^*, t_t^*, \sin_{1t}^*, \ldots, \sin_{nt}^*, \cos_{1t}^*, \ldots, \cos_{nt}^*)^T$.

We finish this section by presenting some simple realizations of the processes given by (14). For simplicity, we suppose $\alpha = \beta = 0$, and start in Figure 1 with n = 1, and $\lambda_1 = \gamma_1 = 1$. We consider two sample sizes, T = 100 and T = 500 and supposed that d takes the following values: 0 (short memory), 0.25 (stationary long memory), 0.75 (nonstationary mean reverting process) and d = 1 (a unit root).

[Insert Figures 1 - 3 about here]

We observe in Figure 1 that the non-linear structure is apparent in all cases with a somewhat cyclical pattern. If d = 0 or 0.25, the series remain stationary but as d increases to 0.75 and 1 the series are clearly nonstationary. Across Figures 2 and 3, we still maintain n = 1 but we allow now for fractional FFFs. We try with j_1 = 0.5 (in Figure 2) and j_1 = 1.5 (in Figure 3) for the same T- and d-values as in Figure 1. By using further unit values for k in the FFFs, that is extending n to 2 and 3 (k = 2 and k = 3) in Figures 4 and 5, we found patterns similar to those obtained for j_1 = 1.5 in Figure 3.

[Insert Figures 4 - 5 about here]

3. Monte Carlo simulations

In this section, we examine different versions of the test described in Section 2 by means of Monte Carlo simulations. Initially, we assume that the Data Generating Process (DGP) is the following one,

$$y_t = 0.6 \sin(2\pi t / T) + 0.4 \cos(2\pi t / T) + x_t; \quad (1-L)^d x_t = u_t, \tag{17}$$

with d = 0, 0.25, 0.75 and 1. For this purpose, we generate Gaussian series using the routines GASDEV and RAN3 of Press, Flannery, Teukolsky and Vetterling (1986), using different sample sizes, T = 100, 300, 500 and 1000, and taking 10,000 replications for each case.[88] We present the results of the test statistic $\hat{r} = (\hat{R})^{1/2}$, with $\hat{R}$ given by (12) and with the residuals given by (16), for a nominal size of 5%, and consider one–sided alternatives of the form H_a: $d > d_o$ and H_a: $d < d_o$. Thus, a one-sided 100α%-level of (9) against the alternative $d > d_o$ will be given by the rule:

[88] The codes are available from the authors upon request.

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

where the probability that a standard normal variate exceeds z_α is α. In the same way, an approximate one-sided 100α%-level of (9) against the alternative $d < d_o$ will be given by the rule:

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

Across Table 1, we examine the size and the power properties of the test in the case of the model given by (17) with d = 0, 0.25, 0.75 and 1, and look at the rejection frequencies of $\hat{r}$ with d_o = d - 0.5, d – 0.25, d, d + 0.25 and d + 0.5. Thus, for example, if the Data Generating Process (DGP) is such that x_t in (17) follows an I(0) process, i.e., d = 0, we look at alternatives with d = -0.5, -0.25, 0, 0.25 and 0.50, and the values corresponding to d_o = 0 (for example in the first table panel) will indicate the size of the test. Similarly for the rest of the panels with d_o = 0.25, d_o = 0.75 and d_o = 1.00.

[Insert Table 1 about here]

The first thing we observe in Table 1 is that the size of the test is clearly biased in small samples. Thus, for example, if T = 100 and d = 0, if we direct the test against d > 0, the size of the test is 0.014; however, when directed against d < 0, the value is much higher, 0.111. In both cases it is far from the theoretical size of 0.050. Nevertheless, as we increase the sample size, both numbers tend to approximate the theoretical 5% level, and when T = 1000, the values are, respectively, 0.043 and 0.058. Very similar values were obtained when d changes from 0 to the alternative values (0.25, 0.75 and 1) in the three remaining table panels. If we focus now on the rejection frequencies we notice that the lower sizes associated to the (d > d_o) alternatives produce lower rejection frequencies, though even for the smallest sample size (T = 100), they are close to 1 if d = d_o ± 0.5.

[Insert Table 2 about here]

Table 2 is similar to Table 1 but including a linear time trend. Thus, the DGP that we consider now is:

$$y_t = 1 + 0.5t + 0.6 \sin\left(2\pi t/T\right) + 0.4 \cos\left(2\pi t/T\right) + x_t; \quad (1-L)^d x_t = u_t, \tag{18}$$

and the results are fairly similar to those presented in Table 1, though with a larger bias in the size in finite samples, probably due to the over-parameterization of the model, but considerably improving as we increase the sample size. The same experiment was also conducted in non-Gaussian environments. In particular, we consider the case where u_t in (17) and (18) is distributed with a Student-t distribution with 3 degrees of freedom. This distribution is of interest because it just satisfies the second moment condition required in the test, while the third moments do not exist. The results, though not reported, produced essentially the same type of conclusions, with a slight bias in the sizes with the small sample size (T = 100) but improving considerably as we increase the sample size. Similar evidence was obtained when multiple frequencies were employed.

4. Empirical applications

For the empirical analysis, we consider both cyclical and trend stationary cases. First, for cyclical series, we consider the US unemployment rate. Also, as applied in Omay et al. (2016), we considered US GNP per capita for the case of trend stationarity. Both time series were sourced from the FRED database of the Federal Reserve Bank of St Louis (https://www.stlouisfed.org/). The choice of unemployment of US is a result of its cyclical pattern, its hysteresis and its non-linearity has been already analysed in the literature (Furuoka, 2014; 2016). There have also been arguments on shocks to US real GDP whether the shocks have temporary or permanent effects, thus checking whether output of the US economy as measured by real GDP or GNP is trend stationary or difference stationary (Omay et al., 2016). All these reasons are behind our choice of these two initial datasets. Our results, however, did not indicate any evidence of non-linearities of the form advocated in this work, and based on Fourier approximations. Because of that, in addition to these two datasets, we also employed the Purchasing Power Parity (PPP) of some G7 countries (as measured by GDP per Capita). These time series were sourced from the International Monetary Fund (IMF) website (www.imf.org).

4a. US unemployment

The time series refers to the monthly US civilian unemployment rate from January 1948 to August 2018 covering a total of 848 data points.[99] The plot of the series is displayed in Figure 6 and we clearly observe a cyclical pattern in its behaviour.

We consider the model given by equation (14) with n = 1, assuming that x_t is an I(d) process of the form as in (2), under the assumption that the I(0) disturbances u_t are uncorrelated (white noise) and autocorrelated, in the latter case following the non-parametric approach of Bloomfield (1973).[1010] Thus, under the null hypothesis (9), the model becomes:

$$y_t = \alpha + \beta t + \lambda_1 \sin(2\pi t/T) + \gamma_1 \cos(2\pi t/T) + x_t; \quad (1-L)^d x_t = u_t, \tag{18}$$

and the estimated parameters, along with their corresponding t-values are reported in Table 3. In case of the differencing parameter d, we report the confidence band of the non-rejection values of d at the 5% level using the test described in Section 2.3.

[Insert Figure 6 and Table 3 about here]

We observe in Table 3 that under both specifications (white noise and Bloomfield-autocorrelation), the estimated value of d is significantly higher than 1 but the estimated coefficients for the deterministic terms are all insignificantly different from zero. Thus, we do

[99] A quarterly frequency version of the US unemployment rate was examined in Caporale and Gil-Alana (2007) and Gil-Alana and Caporale (2008).

[1010] This is a non-parametric specification in the sense that the model is only implicitly determined by its spectral density function. This model of Bloomfield (1973) produces autocorrelations decaying exponentially fast as in the AR(MA) case. However, unlike that case, it is stationary across all its values and it accommodates extremely well in the context of fractional integration. (See, e.g., Gil-Alana, 2004).

not find any evidence supporting the FFF model with this data. Surprisingly, if we impose a priori that d = 0, the four coefficients become significant. However, that assumption is very unrealistic based on the estimated values of d. If we impose d = 1 (also rejected with our test) only the intercept becomes statistically significant.

4b. US real GNP per capita

This series is quarterly and spans from Q1 of 1947 to Q1 of 2018 covering a total of 285 (see Figure 7).

[Insert Figure 7 and Table 4 about here]

Table 4 reports the estimates for the real GNP per capita data. We observe now that the estimated values of d are 1.26 under white noise and 1.11 with autocorrelation, and in the latter case, the I(1) hypothesis cannot be rejected. The intercept and time trend coefficients are now significant in both cases, but once again, we do not find any evidence of non-linearities of the FFF form since the coefficients are both insignificant. As with the previous case, if we impose d = 0, the two non-linear terms become significant but we know this hypothesis has been rejected with the non-linear test. Estimating the model with d = 1, we obtain that the intercept and the time trend are again significant but not the non-linear terms.

4.c A group of GDP per capita series

In this third application, we use GDP per capita (PPP), annually, from 1960 to 2017, for the U.S., Canada, Japan and the U.K., and the results are reported in Table 5.

As in the previous cases, we report the estimated coefficients for the two cases of uncorrelated and autocorrelated (Bloomfield) errors. Starting with the case of white noise errors, the first thing we observe is that the unit root null hypothesis (i.e., d = 1) cannot be rejected for the cases of the U.S., Japan and the U.K., being this hypothesis rejected in favour of d > 1 for Canada. We also observe that the two non-linear terms are statistically significant for the U.S.; one of the coefficients is significant in the case of Japan, and both coefficients are insignificant in the cases of Canada and the U.K.

[Insert Table 5 about here]

Imposing autocorrelated disturbances, the confidence intervals are extremely wide in case of the U.S. and Canada, observing also a substantial reduction in the degree of integration; for the U.S. neither the I(0) nor the I(1) hypotheses can be rejected and for Canada it is the I(1) null which cannot be rejected. For the other two countries the estimated values of d are very close to those based on white noise disturbances. Focussing now on the non-linear terms, this hypothesis can neither be rejected for the U.S. nor for either of the coefficients in the cases of Canada and Japan, while it is decisively rejected in the case of the U.K.

5. Concluding comments

In this article, we have presented a new testing procedure for the potentially fractional order of integration in time series which follow non-linear deterministic terms which are approximated by Fourier functions. The method is a slight modification of the one proposed in Robinson (1994) for the non-linear case and follows the same line of research as in Cuestas and Gil-Alana (2016) for the case of the Chebyshev polynomials in time. As in Robinson (1994) and Cuestas and Gil-Alana (2016), the proposed test statistic has a standard N(0, 1) distribution and several experiments conducted via Monte Carlo show that it performs relatively well even with small samples. Several empirical applications conducted in the paper show evidence of non-linear structures in the case of the U.S. GDP per capita but not in case of the U.K.

This paper can be extended in several directions. For example, other non-linear structures can be examined, such as discrete and continuous threshold models in time functions, that is, the Markov Switching (MS), Threshold Autoregressive (TAR) and Smooth Transition Autoregressive (STAR) models. In addition, asymmetry and symmetry of STAR models could be examined, still in the context of the Robinson's (1994) set-up, though we need, in such cases, some a priori knowledge about certain parameters to maintain the linear structure in the parameters once the fractional integration polynomial is incorporated in the non-linear structure (see Caporale and Gil-Alana, 2007). Finally, the extension of the tests to the case where the singularity in the spectrum takes place at frequencies away from zero, such as in the seasonal and cyclical models, will also be examined in future papers.

References

Abbritti, M., L.A. Gil-Alana, Y. Lovcha and A. Moreno (2016). Term Structure Persistence, *Journal of Financial Econometrics* 14, 2, 331-352.

Bai, J. and Perron P. (2003). Computation and analysis of multiple structural change models. Journal of Applied Econometrics, 18(1), 1–22.

Baillie, R. T. (1996). Long memory processes and fractional integration in econometrics. Journal of Econometrics, 73 (1), 5-59.

Becker, R., Enders, W. and Hurn, S. (2004). A general test for time dependence in parameters. Journal of Applied Econometrics, Vol. 19, pp. 899–906.

Becker, R., Enders, W. and Lee, J. (2006). A stationary test with an unknown number of smooth breaks. Journal of Time Series Analysis, Vol. 27, pp. 381–409.

Bierens, H.J. (1997). Testing for a unit root with drift hypothesis against nonlinear trend stationarity, with an application to the US price level and interest rate. Journal of Econometrics, Vol. 81, pp. 29–64.

Bierens, H.J. (2001). Complex unit roots and business cycles, Are they real? Econometric Theory 17, 962-983.

Bisognin, C. and S.R.C. Lopes (2009), Properties of seasonal long memory processes, Mathematical and Computer Modelling 49, 9-10, 1837-1851.

Bloomfield, P., (1973). An exponential model in the spectrum of a scalar time series. Biometrika, 60(2), 217-226.

Campbell, J.Y. and P. Perron, (1996). Pitfalls and opportunities. What macroeconomists should know about unit roots. NBER Macroeconomic Annual 1141-1201.

Caporale, G.M. and L.A. Gil-Alana (2007). Nonlinearities and fractional integration in the US unemployment rate, Oxford Bulletin of Economics and Statistics 69, 4, 521-544.

Cuestas, J.C. and L.A. Gil-Alana, (2016). Testing for long memory in the presence of non-linear deterministic trends with Chebyshev polynomials. Studies in Nonlinear Dynamics and Econometrics 20(1), 57-74.

Davies, R. B. (1987). Hypothesis testing when a nuisance parameter is only identified under the Alternative. Biometrika, Vol. 47, pp. 33–43.

DeJong, D., J. Nankervis, N.E. Savin and C.H. Whiteman, (1992). Integration versus trend stationarity in time series, Econometrica 60, 423-433.

Dickey, D.A and Fuller, W. A. (1979). Distributions of the Estimators for Autoregressive Time Series with a Unit Root, Journal of American Statistical Association, 74, 366, 427-481.

Dickey, D.A., D.P. Hasza and W.A. Fuller, (1984). Testing for unit roots in seasonal time series, Journal of the American Statistical Association 79, 355-367.

Diebold, F.X., Rudebusch, G.D. (1991). On the power of Dickey-Fuller test against fractional alternatives. Econ. Lett. 35, 155–160.

Dolado, J.J., Gonzalo, J. and Mayoral, L. (2002). A fractional Dickey-Fuller test for unit roots. Econometrica, 70(5): 1963-2006.

Elliot, G., Rothenberg, T.J and Stock, J.H. (1996). Efficient Tests for an Autoregressive Unit Root. Econometrica, 64, 813-836.

Enders, W. and Lee, J. (2012a). A unit root test using a Fourier series to approximate smooth breaks. Oxford Bulleting of Economics and Statistics, 74: 574-599.

Enders, W. and Lee, J. (2012b). The flexible Fourier form and Dickey-Fuller-type unit root tests. Economic Letters, 117: 196-199.

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

Furuoka, F. (2014). Are unemployment rates stationary in Asia-Pacific countries? New findings from Fourier ADF test. Economic Research, 27(1): 34-45.

Furuoka, F. (2016). A new approach to testing unemployment hysteresis. Empirical Economics, DOI: 10.1007/s00181-016-1164-7.

Gallant, A. R. (1981). On the bias in flexible functional forms and an essentially unbiased form: the flexible Fourier form. Journal of Econometrics, 15: 211–245.

Gallant, R. and Souza, G. (1991). On the asymptotic normality of Fourier flexible form estimates, Journal of Econometrics, 50: 329–353.

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

Gil-Alana, L.A. (2002). Seasonal long memory in aggregate output, Economics Letters 74, 333-337.

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., (2005). Statistical modelling of the temperatures in the Northern hemisphere using fractional integration techniques, Journal of Climate 18, 24, 5357-5369.

Gil-Alana, L.A., (2007). Testing the existence of multiple cycles in financial and economic time series. Annals of Economics and Finance 1, 1-20.

Gil-Alana, L.A. and Caporale, G.M. (2008). Modelling the US, the UK and Japanese unemployment rates. Fractional integration and structural breaks. Faculty Working Papers 11/08, School of Economics and Business Administration, University of Navarra.

Gil-Alana, L.A. and E.H. Huijbens (2018). Tourism in Iceland. Persistence and seasonality, Annals of Tourism Research 68, 20-29.

Gil-Alana, L.A. and A. Moreno (2012). Uncovering the term premium. An alternative route, Journal of Banking and Finance 36, 4, 1181-1193.

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 seasonal fractional integration in the UK and Japanese consumption and income, Journal of Applied Econometrics 16, 95-114.

Granger, C.W.J. and Teräsvirta, T. (1993). Modelling Nonlinear Economic Relationships. Oxford: Oxford University Press. Advanced Texts in Econometrics, Chinese edition 2006: Shanghai University of Finance and Economics Press.

Hasslers, U., Wolters, J., (1994). On the power of unit root tests against fractional alternatives. Economic Letters, 45: 1–5.

Hylleberg, S., R.F. Engle, C.W.J. Granger and B.S. Yoo, (1990). Seasonal integration and cointegration. Journal of Econometrics 44, 215-238.

Johansen, S., Juselius, K., Frydman, R. and Goldberg, M. (2010). Testing hypotheses in an I(2) model with piecewise linear trends. An analysis of the persistent long swings in the Dmk/$ rate. Journal of Econometrics, 158(1), 117-129.

Kapetanios, G., Y. Shin, and A. Snell. (2003). Testing for a Unit Root in the Nonlinear STAR Framework. Journal of Econometrics 112: 359–379.

Kunst, R. (2001), Testing for cyclical nonstationarity in autoregressive processes, Journal of Time Series Analysis 18, 2, 123-135.

Lee, D. and P. Schmidt (1996). On the power of the KPSS test of stationary against fractionally integrated alternatives, Journal of Econometrics 73, 285-302.

Leybourne, S. J., Newbold, P. and Vougas, D. (1998). Unit roots and smooth transitions. Journal of Time series Analysis, 19: 83-97.

Nelson, C.R., J. Piger and E. Zivot, (2001). Markov regime switching and unit root tests, Journal of Business and Economics Statistics 19, 404-415.

Omay, T. (2015). Fractional frequency flexible fourier form to approximate smooth breaks in unit root testing. Economic Letters, 134: 123-126.

Omay, T., Gupta, R. and Bonaccolto, G. (2017). The US real GNP is trend stationary after all. Applied Economics Letters, 24, 8, 510-514.

Phillips, P.C.B. and Perron, P. (1988). Testing for a unit root in time series regression', Biometrika 75, 2, 335–346.

Press, W.H., B.P. Flannery, S.A. Teukosky and W.T. Wetterling, (1986). Numerical recipes. The art of scientific computing, Carmbridge University Press, Cambridge.

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

Rodrigues, P.M.M., A.M.R. Taylor (2012), "The flexible Fourier form and local generalised least squares de-trended unit root tests", Oxford Bulletin of Economics and Statistics, 74(5), 736-759.

Sadek, N. and A. Khotanzad (2004). K-factor Gegenbauer ARMA process for network traffic simulation. Computers and Communications 2, 963-968.

Schmidt, P. and Phillips, P.C.B. (1992). LM tests for a unit root in the presence of deterministic trends. Oxford Bulletin of Economics and Statistics, 54: 257-287.

Tanaka, K., (1999). The nonstationary fractional unit root, Econometric Theory 15, 549-582.

Figure 1: Simple realization of the model in (14) with $\alpha = \beta = 0$, $n = 1$, $\lambda_1 = \gamma_1 = 1$.

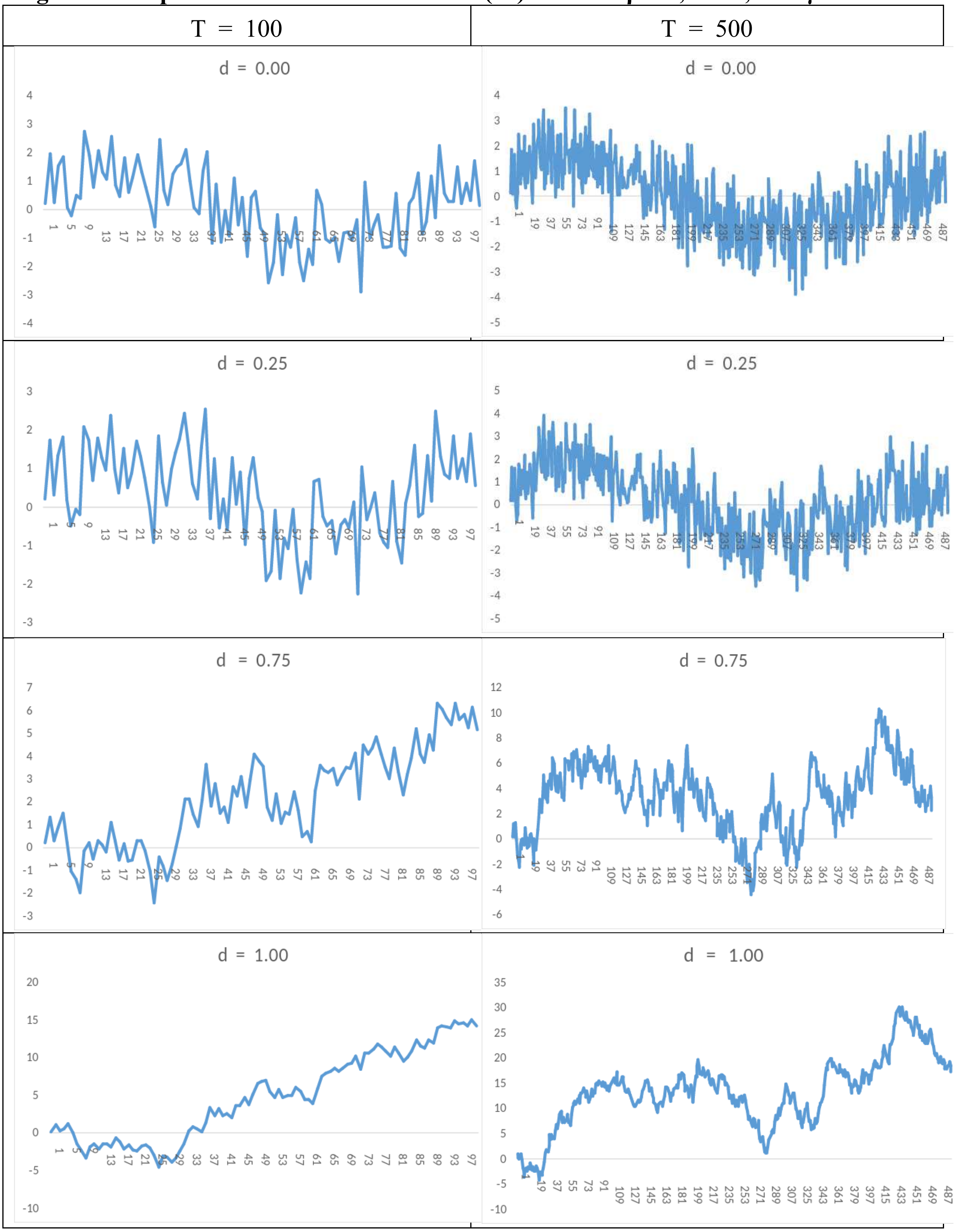

Figure 2: Simple realization of the model in (14) with $\alpha = \beta = 0$, $n = 1$, $j_1 = 0.5$, $\lambda_1 = \gamma_1 = 1$

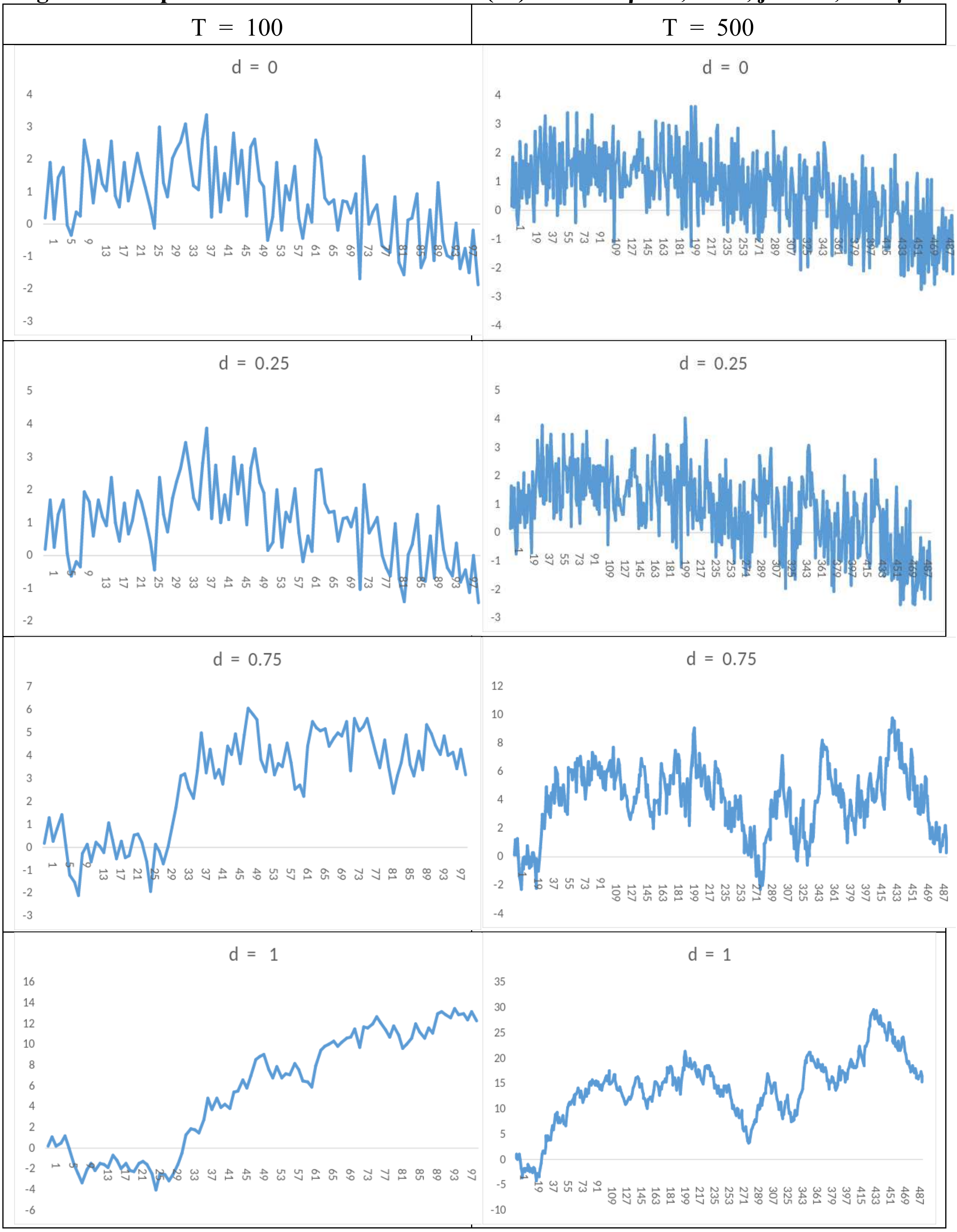

Figure 3: Simple realization of the model in (14) with $\alpha = \beta = 0$, $n = 1$, $j_1 = 1.5$, $\lambda_1 = \gamma_1 = 1$

T = 100	T = 500

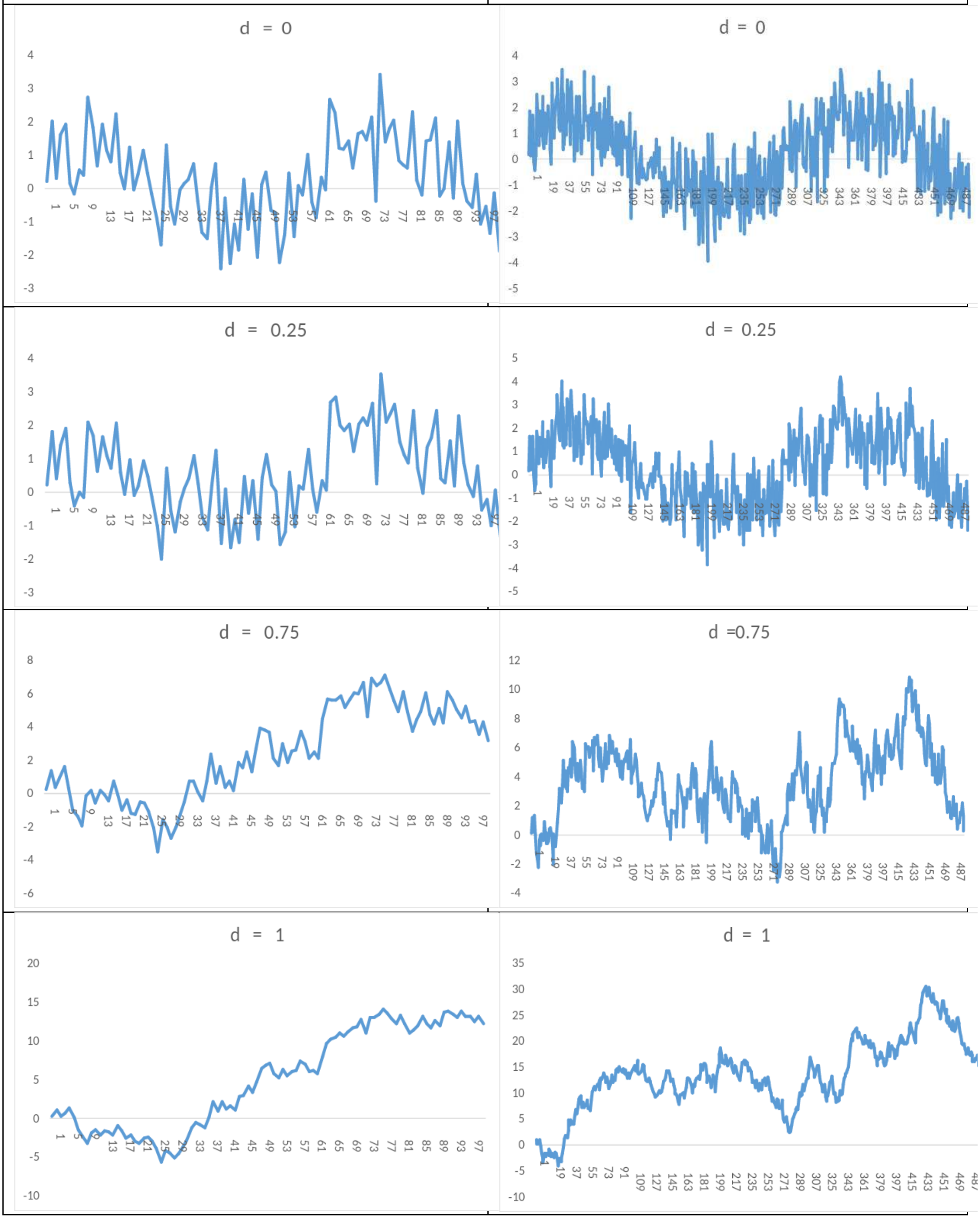

Figure 4: Simple realization of the model in (14) with α = β = 0, n = 2, $\lambda_k = \gamma_k = 1$

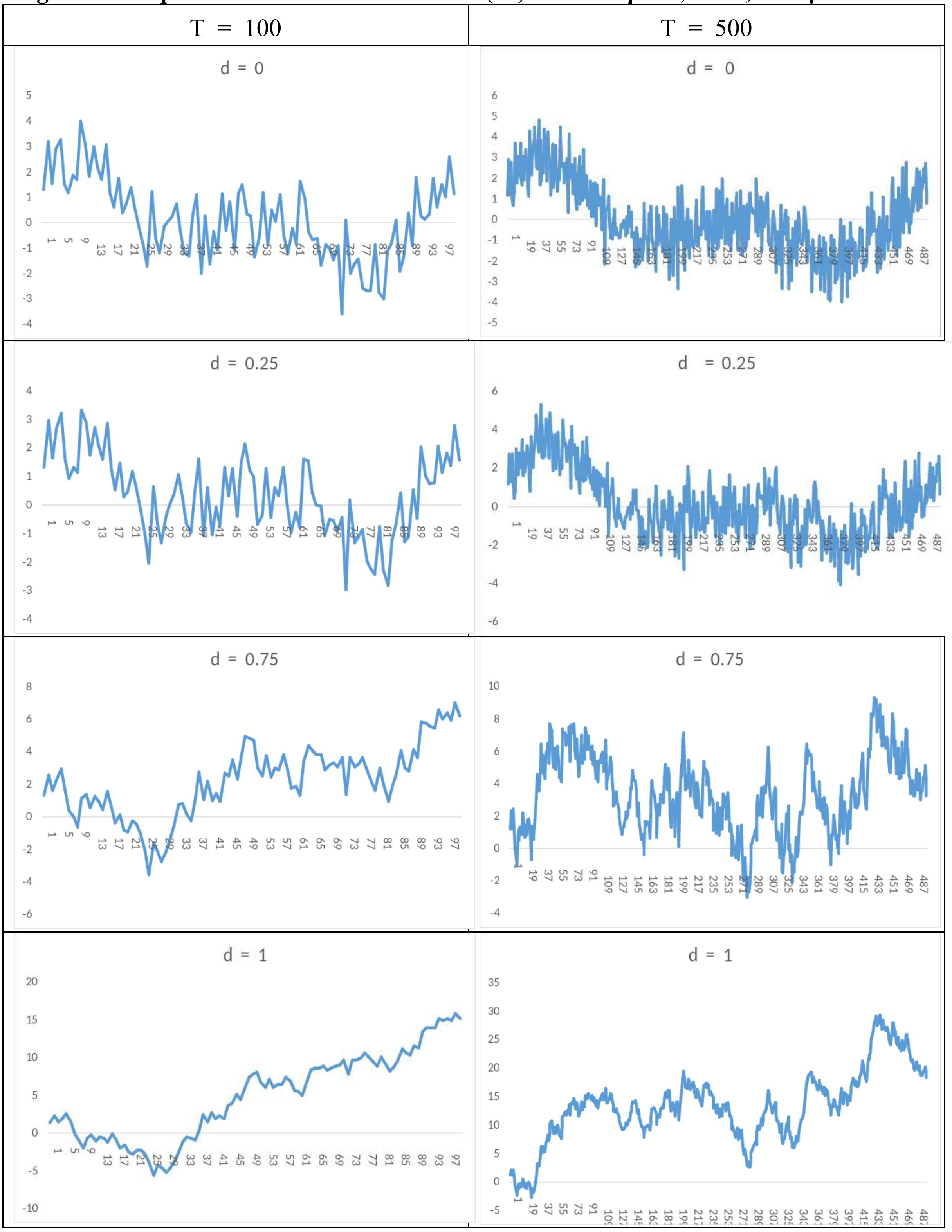

Figure 5: Simple realization of the model in (14) with α = β = 0, n = 3, $\lambda_k = \gamma_k = 1$

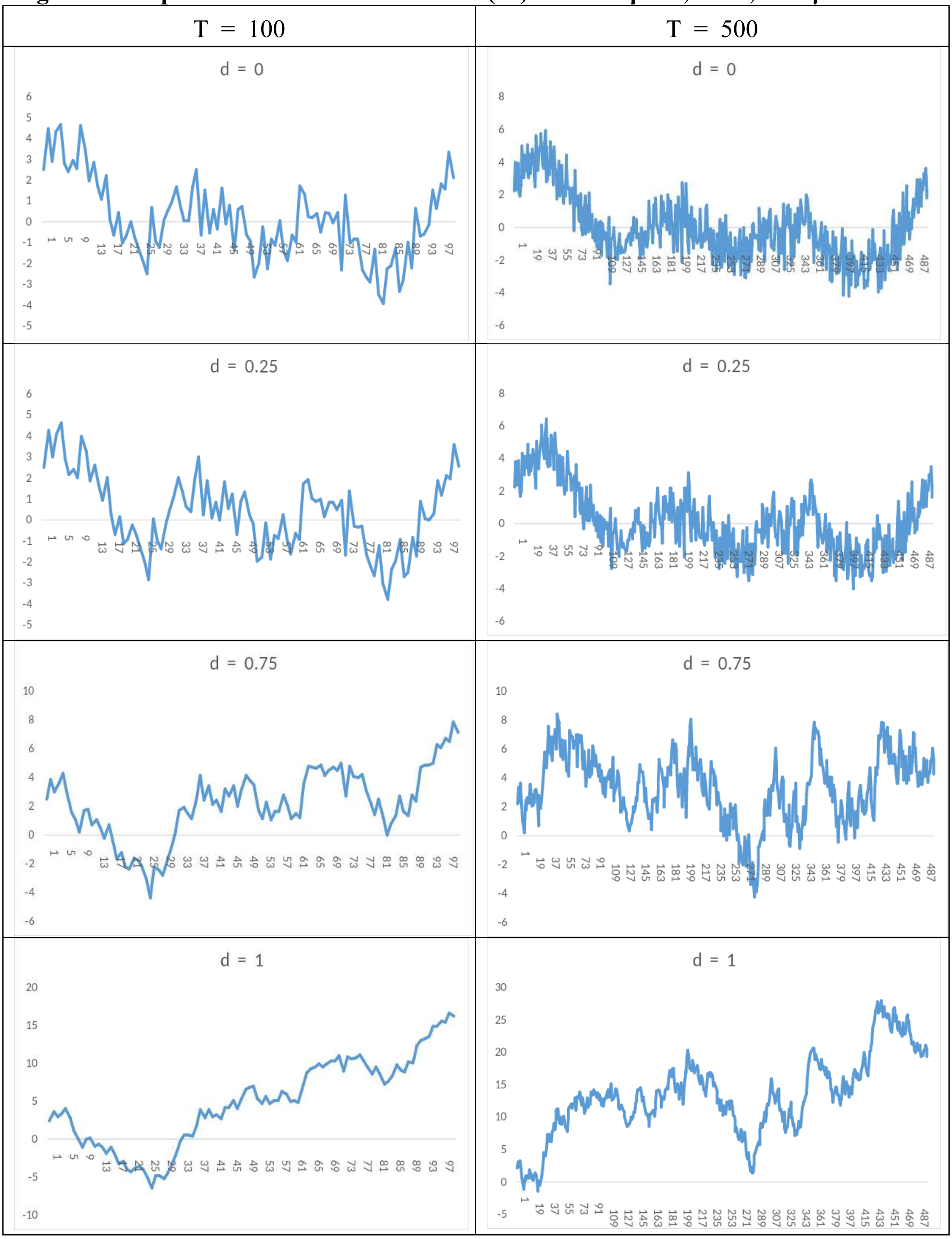

Figure 6: US monthly unemployment rate

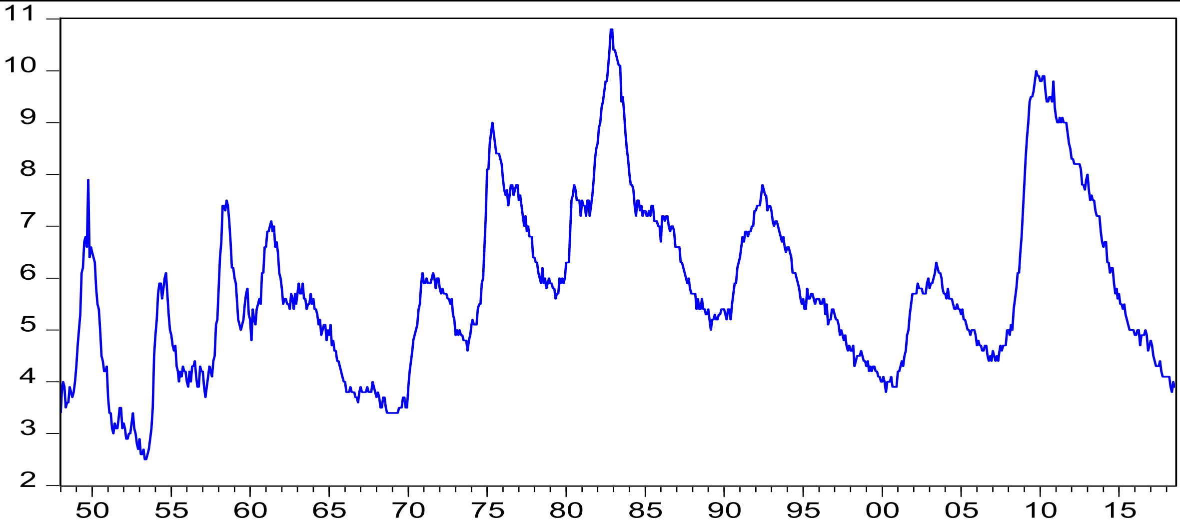

Figure 7: US real GNP per capita

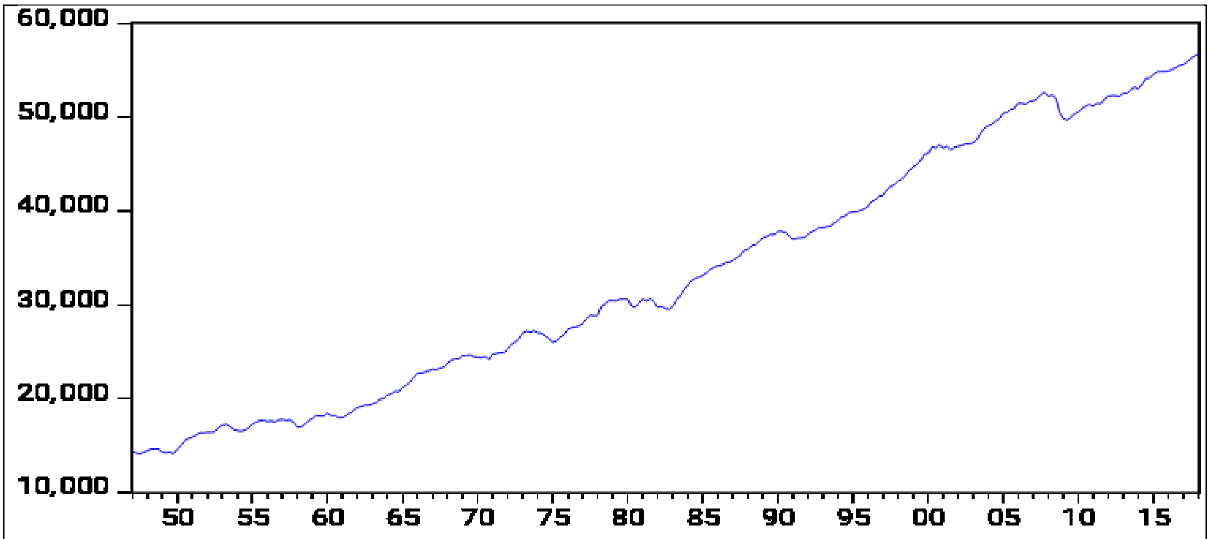
60,000
50,000
40,000
30,000
20,000
10,000
50
55
60
65
70
75
80
85
90
95
00
05
10
15

Table 1: Rejection frequencies against one-sided alternatives with FFF model and k = 1

d_o = 0						
	d_o > 0			d_o < 0		
T	d = -0.50	d = -0.25	d = 0	d = 0	d = 0.25	d = 0.50
100	0.996	0.474	**0.014**	**0.111**	0.940	1.000
300	1.000	0.991	**0.024**	**0.099**	1.000	1.000
500	1.000	1.000	**0.039**	**0.078**	1.000	1.000
1000	1.000	1.000	**0.043**	**0.058**	1.000	1.000
d_o = 0.25						
	d_o > 0.25			d_o < 0.25		
T	d = -0.25	d = 0.	d = 025	d = 0.25	d = 0.50	d = 0.75
100	0.978	0.479	**0.014**	**0.113**	0.930	1.000
300	1.000	0.991	**0.023**	**0.099**	1.000	1.000
500	1.000	1.000	**0.037**	**0.076**	1.000	1.000
1000	1.000	1.000	**0.042**	**0.058**	1.000	1.000
d_o = 0.75						
	d_o > 0.50			d_o < 0.50		
T	d = 0.25	d = 0.50	d = 0.75	d = 0.75	d = 1.00	d = 1.25
100	1.000	0.507	**0.018**	**0.105**	0.917	1.000
300	1.000	0.993	**0.030**	**0.093**	1.000	1.000
500	1.000	1.000	**0.041**	**0.074**	1.000	1.000
1000	1.000	1.000	**0.046**	**0.055**	1.000	1.000
d_o = 1.00						
	d_o > 1.00			d_o < 1.00		
T	d = 0.50	d = 0.75	d = 1.00	d = 1.00	d = 1.25	d = 1.50
100	0.986	0.579	**0.020**	**0.102**	0.922	1.000
300	1.000	0.995	**0.033**	**0.091**	1.000	1.000
500	1.000	1.000	**0.042**	**0.070**	1.000	1.000
1000	1.000	1.000	**0.047**	**0.054**	1.000	1.000

In bold the size of the test

Table 2: Rejection frequencies against one-sided alternatives with an intercept, a linear trend, and FFF with k = 1

$d_o = 0$						
	$d_o < 0$			$d_o > 0$		
T	d = -0.50	d = -0.25	d = 0	d = 0	d = 0.25	d = 0.50
100	0.935	0.358	**0.007**	**0.132**	0.952	1.000
300	1.000	0.982	**0.020**	**0.109**	1.000	1.000
500	1.000	1.000	**0.031**	**0.090**	1.000	1.000
1000	1.000	1.000	**0.038**	**0.067**	1.000	1.000
$d_o = 0.25$						
	$d_o < 0.25$			$d_o > 0.25$		
T	d = -0.25	d = 0.	d = 025	d = 0.25	d = 0.50	d = 0.75
100	0.922	0.367	**0.009**	**0.130**	0.942	1.000
300	1.000	0.984	**0.020**	**0.109**	1.000	1.000
500	1.000	1.000	**0.032**	**0.090**	1.000	1.000
1000	1.000	1.000	**0.038**	**0.069**	1.000	1.000
$d_o = 0.75$						
	$d_o < 0.50$			$d_o > 0.50$		
T	d = 0.25	d = 0.50	d = 0.75	d = 0.75	d = 1.00	d = 1.25
100	0.948	0.435	**0.012**	**0.128**	0.936	1.000
300	1.000	0.988	**0.022**	**0.103**	1.000	1.000
500	1.000	1.000	**0.033**	**0.088**	1.000	1.000
1000	1.000	1.000	**0.040**	**0.059**	1.000	1.000
$d_o = 1.00$						
	$d_o < 1.00$			$d_o > 1.00$		
T	d = 0.50	d = 0.75	d = 1.00	d = 1.00	d = 1.25	d = 1.50
100	0.971	0.465	**0.014**	**0.128**	0.935	1.000
300	1.000	0.990	**0.023**	**0.103**	1.000	1.000
500	1.000	1.000	**0.033**	**0.087**	1.000	1.000
1000	1.000	1.000	**0.041**	**0.059**	1.000	1.000

In bold the size of the test

Table 3: Estimated parameters for the US unemployment rate

u_t	d	α	β	λ_1	γ_1
White noise	1.17 (1.13, 1.21)	3.775 (1.19)	0.002 (0.11)	-0.066 (-0.02)	-0.448 (-0.14)
Bloomfield	1.46 (1.34, 1.57)	2.002 (0.13)	0.046 (0.39)	1.859 (0.15)	1.233 (0.08)
Imposing d = 0					
-----	0.00	**4.696** **(31.52)**	**0.002** **(7.68)**	**0.248** **(2.17)**	**-0.720** **(-10.07)**
Imposing d = 1					
-----	1.00	**1.360** **(5.40)**	0.0001 (0.12)	-0.038 (-0.15)	-0.137 (-0.55)

In bold in column 3 – 7, significant coefficients. T-values in parenthesis.

Table 4: Estimated parameters for the US real GNP per capita

u_t	d	α	β	λ_1	γ_1
White noise	1.26 (1.16, 1.37)	**12830.11** **(4.64)**	**144.16** **(2.47)**	-1452.05 (-0.58)	1409.40 (0.51)
Bloomfield	1.11 (0.90, 1.38)	**12603.14** **(8.11)**	**148.95** **(5.45)**	-1364.34 (-0.93)	1604.15 (1.04)
Imposing d = 0					
-----	0.00	**11.765.14** **(71.01)**	**149.90** **(137.53)**	**-1310.02** **(-10.33)**	**1682.71** **(21.19)**
Imposing d = 1					
-----	1.00	**9.609** **(261.41)**	**0.004** **(8.74)**	-0.035 (-0.98)	0.046 (-1.30)

In bold in column 3 - 7, significant coefficients. T-values in parenthesis.

Table 5: GDP per capita series

i) White noise disturbances					
u_t	d	α	β	λ_1	γ_1
U.S.A.	1.14 (0.92, 1.43)	**8.207** **(175.70)**	**0.051** **(13.24)**	**-0.104** **(-2.53)**	**-0.238** **(-5.57)**
Canada	1.27 (1.06, 1.57)	**7.939** **(36.95)**	**0.047** **(2.20)**	-0.093 (-0.50)	-0.225 (-1.11)
Japan	1.04 (0.76, 1.42)	**6.841** **(33.75)**	**0.076** **(5.31)**	-0.088 (-0.49)	**-0.745** **(-4.18)**
U.K.	1.21 (0.75, 1.74)	**7.471** **(28.15)**	**0.058** **(2.35)**	-0.305 (-1.31)	-0.273 (-1.10)
ii) Autocorrelation (Bloomfield)					
u_t	d	α	β	λ_1	γ_1
U.S.A.	0.71 (-0.02, 1.26)	**8.196** **(346.04)**	**0.052** **(54.99)**	**-0.091** **(-4.70)**	**-0.240** **(-14.49)**
Canada	0.67 (0.29, 1.39)	**7.893** **(92.79)**	**0.055** **(16.97)**	-0.010 (-0.15)	**-0.247** **(-4.31)**
Japan	1.05 (0.88, 1.22)	**6.848** **(33.09)**	**0.076** **(5.13)**	-0.085 (-0.48)	**-0.744** **(-4.07)**
U.K.	1.20 (0.96, 1.49)	**7.473** **(28.77)**	**0.058** **(2.43)**	-0.304 (-1.34)	-0.275 (-1.13)

In bold in column 3 – 7, significant coefficients. T-values in parenthesis.