PERSISTENCE AND CYCLES IN HISTORICAL OIL PRICES DATA

Luis A. Gil-Alana
Navarra Center for International Development and Faculty of Economics, University of Navarra, Edificio Amigos, E-31080 Pamplona, Spain
Email: alana@unav.es

and

Rangan Gupta
Department of Economics, University of Pretoria, Pretoria, 0002, South Africa
Email: rangan.gupta@up.ac.za.

ABSTRACT

This paper deals with the analysis of two observed features in historical oil prices data. In particular, persistence and cyclicity. Using monthly data from September 1859 to October 2013, we observe that the series presents two peaks in the spectrum, one occurring at the long run or zero frequency and the other at a cyclical frequency. These features can be well described in terms of a long memory model that incorporates both peaks in the spectrum. It is found that the order of integration at the zero frequency is about 0.6, and the one at the cyclical frequency is substantially smaller (of about 0.3) with the length of the cycles being approximately of about 74 periods (months), which is consistent with the length suggested by the business cycles theory.

JEL Classification: C22

Keywords: Oil prices; Cycles; Persistence; Long memory

Corresponding author: Luis A. Gil-Alana
University of Navarra
Faculty of Economics, Edificio Amigos
E-31080 Pamplona
SPAIN

Phone: +34 948 425 625
Fax: +34 948 425 626
Email: alana@unav.es

* The first named author acknowledges financial support from the Ministry of Education of Spain ECO2011-2014 ECON Y FINANZAS, Spain). We thank the editor and two anonymous referees for all their interesting comments and suggestions.

1. Introduction

This paper deals with the analysis of historical oil prices data. Two features that are commonly observed in this type of data are its persistence and its cyclical behaviour. Persistence is a measure of the degree of dependence between the observations across time, while cyclicity deals with the repetitive pattern observed over the years.

Persistence in time series has been determined for many years by means of a simple AR(1) process, the associated coefficient being a measure of persistence. In a more general AR(k) process, the sum of the AR coefficients has also been considered as an alternative measure. However, these two measures of persistence are based on the assumption that the underlying series is I(0) stationary, and it is a well known stylized fact that many macroeconomic series are nonstationary (Nelson and Plosser, 1982). In this context, unit root tests have become a standard practice to determine a priori if a series is stationary I(0) or nonstationary I(1), in the latter case it being necessary for the series to be first differenced to render the I(0) behaviour. If a series is I(1) we could say that it is highly persistent with shocks persisting forever. In this paper we consider a more flexible approach based on fractional integration which includes the two above mentioned cases (I(0) and I(1)) as particular cases of interest. If d is fractional and constrained between 0 and 1, shocks will be transitory, disappearing in the long run, however, the process of convergence will be smaller the longer the value of d is. Thus, the differencing parameter d can be considered as a plausible alternative measure of persistence. Moreover, an I(d) process with fractional value d can also display AR structures in the differenced process, and the AR coefficients can also contribute to the degree of persistence of the series. The I(d) processes with $d > 0$ are characterized because the spectral density is unbounded at the zero frequency.

Cyclicity is another feature that is present in many time series. Modelling cycles stochastically, the AR(2) process is a viable method, noting that if the roots of the AR polynomial are complex, a cyclical pattern can be observed in the data. Similarly, unit root cycles have been also proposed in recent years (Bierens, 2001), and in the same way as with the issue of persistence at the zero frequency, the spectral density function can be unbounded at a frequency away from zero. In such a case a process can be cyclically fractionally integrated (Gray et al., 1989, 1994), and the magnitude of the differencing parameter in this context will indicate the degree of cyclical dependence.

In this paper, we combine these two approaches (persistence and cyclicity) in a single framework, proposing a model that allows the spectrum to have two poles or singularities, one at zero and the other at a non-zero (cyclical) frequency. The model was initially proposed in a very general framework in Robinson (1994) though most of the applications of his model focussed exclusively on the long run or zero frequency. There are also some applications using cyclical (fractional) models (Arteche and Robinson, 2000; Gil-Alana, 2001; Ferrara and Guegan, 2001; Sadek and Khotanzad, 2004; etc.), but very little research has been conducted using the two approaches simultaneously (exceptions being: Caporale and Gil-Alana (2011, 2012, 2014) where persistence and cyclicity is applied simultaneously to the Euribor rate, hours worked in the US and the Federal funds rate, respectively), as is the case in this work. To the best of our knowledge, this is the first paper to apply these two approaches simultaneously to the oil price model. However, as recently pointed out by Huntington et al., (2013) and Baumeister (2014) large number of time series approaches exists in modelling (and forecasting) oil prices. Huntington et al., (2013) indicates that oil price models can be categorized into structural (accommodates fundamental microeconomic theories about the objectives, constraints, and behaviours of market actors in a mathematical framework), computational (typically general equilibrium models accounting for other energy or other sectors of the economy) and reduced-form (structural or atheoretical vector autoregressive (VAR)) models). Huntington et al., (2013) points out that oil price modelling has become more reduced-form in nature than structural in recent years, with focus on modelling nonlinearity, including machine learning methods, in both univariate and multivariate frameworks. Baumeister (2014) covers multivariate models such as the one based on the oil futures curve, the global oil market (which includes variables such as world oil production, world economic activity and changes in above-ground inventories of crude oil), spot price of raw industrial materials, and refined product spreads. However, Baumeister (2014) emphasizes that given the range of models available for oil prices, the best approach a policy maker can undertake, especially when it comes to forecasting oil prices, is to undertake a forecast combination approach, since forecast combination allows one to pool the information content in all the models, thus allowing one to retain the best features of each of the models. Interested readers are referred to these two papers and references cited therein for further details.

The structure of the paper is as follows: Section 2 presents a brief review of the empirical literature on long-run oil prices, with Section 3 describing the model. Section 4 presents the data and the main empirical results, while Section 5 concludes the manuscript.

2. A brief review of the empirical literature on long-run oil prices

To realize the contribution of our work, we present a brief summary of the existing literature that exists in modelling oil price covering long samples. Using data from 1870 to 1978, Slade (1982) found statistically significant evidence for a U-shaped curve, represented by a quadratic time trend in the prices of crude oil. Pindyck (1999) showed that the prices of oil, besides gas and coal, are mean-reverting to a quadratic trend line but the rate of reversion is slow, to the extent that the random walk assumption of oil price is not a bad approximation of its movement patterns. Extending Slade's (1982) data to 1990, Berk and Roberts (1996) found the price of oil is non-stationary based on standard unit root tests, like the Dickey-Fuller test. However, Ahrens and Sharma (1997), using more advanced unit root tests concluded that the price of oil was stationary around a determinist trend. Lee et al., (2006) obtained similar results by employing a Lagrange Multiplier test allowing up to two endogenously determined structural breaks, when they showed that the unit root hypothesis can be rejected for the price of oil.

Although not focusing on trend and cycle patterns, Dvir and Rogoff (2010) examine changes in persistence and volatility of crude oil prices across three periods from 1861 to 2009 and documented striking similarities between the period of 1861-1878 and 1973-2009. Erten and Ocampo (2012) use an asymmetric band-pass filter to study the super-cycles (periods lasting from 20 to 70 years) in the price of West Texas Intermediate crude oil. By filtering out the long-term trend, they argue that the super-cycles in the price of crude oil were rather modest in the early twentieth century, but became more pronounced after the 1970s. Zellou and Cuddington (2012) apply similar band-pass filter methods to the same crude oil price data as used in the paper by Mu and Ye (2012), and found, similar cyclical patterns, namely a short cycle of 6 years and a long cycle of 29 years.

Given the low power of standard unit root tests to differentiate between near unit root processes from stationary processes (as well as fractionally integrated processes), we model the persistence property by specifically looking at long-memory (fractional integration) properties of the oil price series. Further, given the wide-spread evidence of cyclical movements in oil prices, as discussed above, we extend the standard model of fractional integration to also allow for cycles. Finally, unlike the literature which has used annual data, as well as oil price measured from different sources, we use monthly data on the Western Texas Intermediate (WTI) Crude oil prices over the period of September, 1859 till October, 2013, thus covering the entire modern era in the petroleum industry, which started with the first oil well drilled in the US on August 27, 1859 in Titusville, Pennsylvania. Since we use an unique source of oil price data, we avoid measurement issues unlike the previous studies, which primarily used annual averages of oil prices in the US for 1861-1944, Arabian Light crude oil for 1945-1983 and Brent crude oil for 1984-current. Besides, the fact that we also look at the highest possible frequency of the data over this entire long-sample, we avoid issues of aggregation (to annual data), which is known to affect the persistence properties of the data. Our main contribution is the use of a novel econometric technique in the analysis of persistence and cyclicity in oil prices data. Also note that, unlike the previous studies in the literature, we use nominal rather than real oil price, since as suggested by Hamilton (2011), it is the nominal oil price that should be used for economic analysis, since deflating the nominal oil price by a measure of price-level tends to induce measurement errors.

3. Methodology

We suppose that $\{y_t, t = 1, 2, ..., T\}$ is the time series we observe that is generated by a function of a deterministic process $\{z_t, t = 1, 2, ...\}$ and a stochastic term $\{x_t, t = 1, 2, ...\}$ that we suppose is fractionally integrated at both the zero and the cyclical frequencies. For the deterministic part, a linear function is initially assumed though a non-linear trend will also be considered. In other words, the model under study is given by:

$$y_t = \beta^T z_t + x_t; \qquad (1 - B)^{d_1} (1 - 2\cos w_r B + B^2)^{d_2} x_t = u_t, \quad t = 1, 2, ..., \qquad (1)$$

where β is a (kx1) vector of unknown coefficients and z_t is a (kx1) vector of deterministic terms that may include an intercept (in the case of $z_t = 1$), a linear time trend ($z_t = (1, t)^T$) or even dummy variables; B is the lag operator ($Bx_t = x_{t-1}$); w_r is a real number, and d_1

and d_2 are also real numbers referring to the degrees of integration at the zero and the cyclical frequencies respectively as will be explained below; finally, u_t is supposed to be an I(0) process, defined for the purpose of the present work as a covariance stationary process with a spectral density function that is positive and finite at all frequencies in the spectrum. Thus, it may be a white noise process, but also a weakly autocorrelated process of the ARMA type. Next we examine in detail the two polynomials appearing in the second equation in (1).

We start with $(1 - B)^{d1}$, i.e., we can assume that $d_2 = 0$ in (1). This is the standard case of fractional integration, where d_1 is the fractional differencing parameter. This specification attracted much interest in the late 90s and early 2000s in both the theoretical and the empirical literature (Gil-Alana and Robinson, 1997; Hauser and Kunst, 1998; Bollerslev and Jubinski, 1999; Ray and Tsay, 2000; Michelacci and Zaffaroni, 2000; etc.) and nowadays many economic time series are modelled using this approach (Mayoral, 2006; Baillie et al., 2007; Christensen et al., 2010; etc.). It can also be shown that the polynomial $(1 - B)^{d1}$ can be expressed for all real $d_1 > 0$ in terms of its Binomial expansion such that

$$(1 - B)^{d_1} = \sum_{k=0}^{\infty} \pi_k(d_1) B^k, \text{ with } \pi_k(d_1) = \frac{\Gamma(k - d_1)}{\Gamma(1 - d_1)\,\Gamma(k + 1)} \quad ,$$

where Γ represents the Gamma function. Also, Bronshtein and Semendyayev (1998) show that

$$\pi_k \approx \frac{k^{-d_1 - 1}}{\Gamma(-d_1)},$$

implying an hyperbolic decay in the autocorrelations, and that the higher the value of d_1 is, the higher the degree of association between observations distant in time is.

Next we focus on the second polynomial in (1). For practical purposes we define $w_r = 2\pi r/T$, with $r = T/s$, 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 x_t. Note that, assuming $d_1 = 0$, if $r = 0$ (or $s = 1$), the fractional cyclical polynomial in (1) becomes $(1-B)^{2d_2}$, which is the polynomial associated with the previous case of fractional integration at the long-run or zero frequency. This type of process was introduced by Andel (1986) and subsequently analysed by Gray, Zhang and Woodward (1989, 1994), Giraitis and Leipus (1995), Gil-Alana (2001), Giraitis et al. (2001), Sadek and Khotanzad (2004), Hildalgo (2005) and Dalla and Hidalgo (2005) among many others.

Gray et al. (1989, 1994) showed that if $d_1 = 0$ in (1), the second polynomial can be expressed in terms of the Gegenbauer polynomial, such that, denoting $\mu = cosw_r$, for all $d_2 \neq 0$,

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

where $C_{j,d_2}(\mu)$ are orthogonal Gegenbauer polynomial coefficients recursively defined as:

$$C_{0,d_2}(\mu) = 1, \qquad C_{1,d_2}(\mu) = 2\mu d_2,$$

$$C_{j,d_2}(\mu) = 2\mu\left(\frac{d_2-1}{j}+1\right)C_{j-1,d_2}(\mu) - \left(2\frac{d_2-1}{j}+1\right)C_{j-2,d_2}(\mu), \quad j = 2, 3, \ldots,$$

(see, for instance, Magnus et al., 1966, Rainville, 1960, etc. for further details on Gegenbauer polynomials). Gray et al. (1989) showed that x_t in (2) is (covariance) stationary if $d_2 < 0.5$ for $|\mu = cosw_r| < 1$ and if $d_2 < 0.25$ for $|\mu| = 1$.

In this paper we use a joint model of the form as in (1) estimating the two fractional differencing parameters by means of the Whittle function in the frequency domain (Fox and Taqqu, 1086; Dahlhaus, 1989). Additionally, we use a version of a testing procedure developed by Robinson (1994) that allows us to obtain a region of (d_1,

d_2) – values where the null hypothesis cannot be rejected at a given statistical significant level. This method tests the null hypothesis:

$$H_o : (d_1, d_2) \; = \; (d_{1o}, d_{2o}), \qquad (2)$$

in the model given by equation (1) for any given real numbers d_{1o} and d_{2o}. The tests statistic follows a χ^2_2 – distribution, and it holds independently of the deterministic terms employed for z_t in equation (1) and the way of modeling the I(0) disturbances u_t in (1).

4. Data and empirical results

The oil price used here corresponds to the West Texas Intermediate (Cushing, Oklahama) crude oil price, and is obtained from the Global Financial database, covering the monthly period of 1859:09 till 2013:10.

Figure 1 displays the time series plot of the logged oil prices data. It can be seen that the values increase across the sample period. Figure 2 shows the first differenced data and the series may now have a stationary appearance. Figure 3 shows the correlogram of the first differenced, while Figures 4 and 5 display the periodogram at different frequencies. The correlogram presents several significant values at some lags

Figure 1: Log of the oil prices

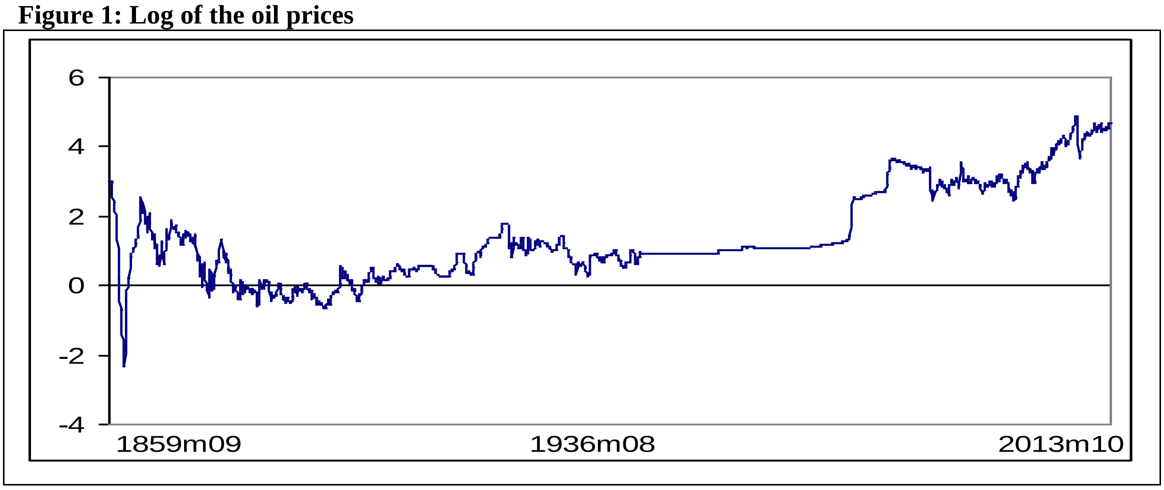


Figure 2: First differences of the logged prices

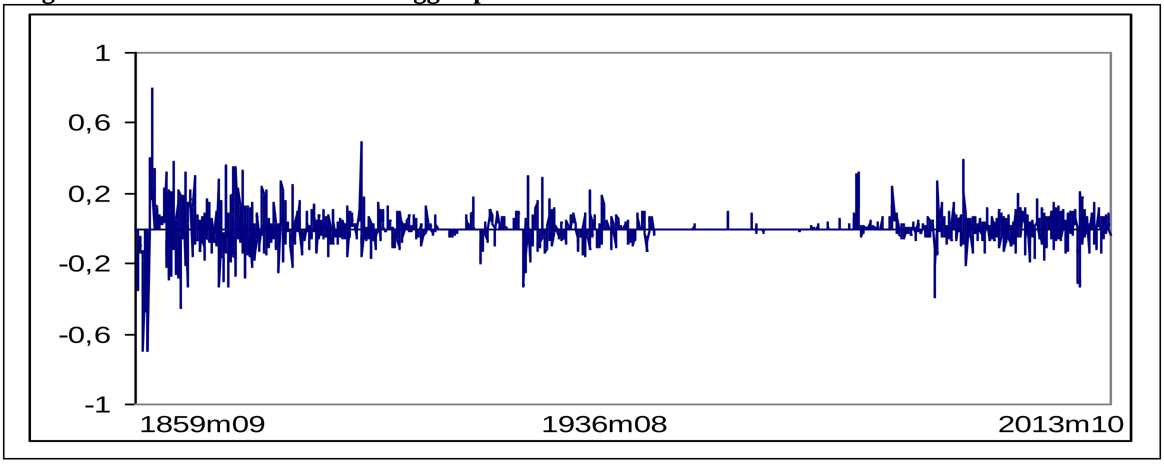


Figure 3: Correlogram of the differenced (log-)series

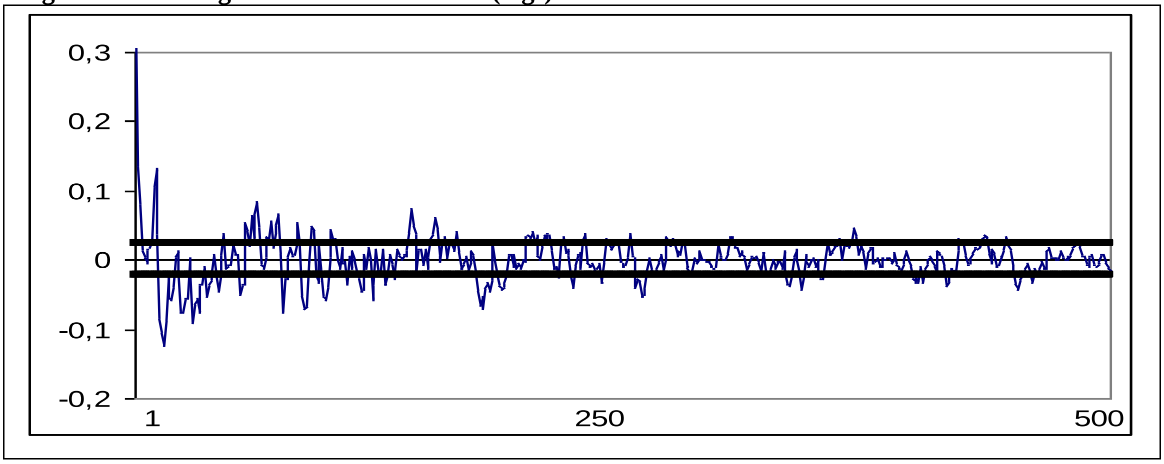


Figure 4: Periologram of the differenced (log-)series

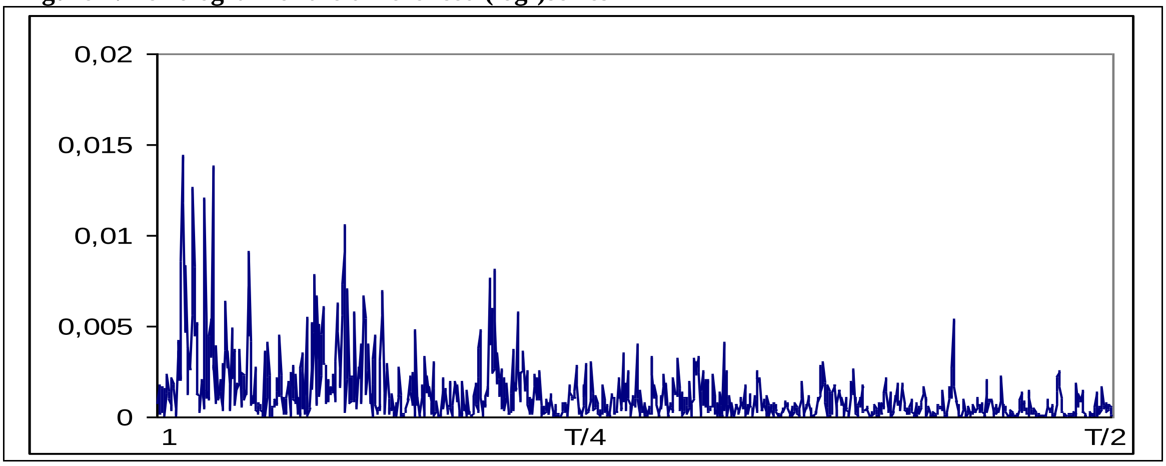
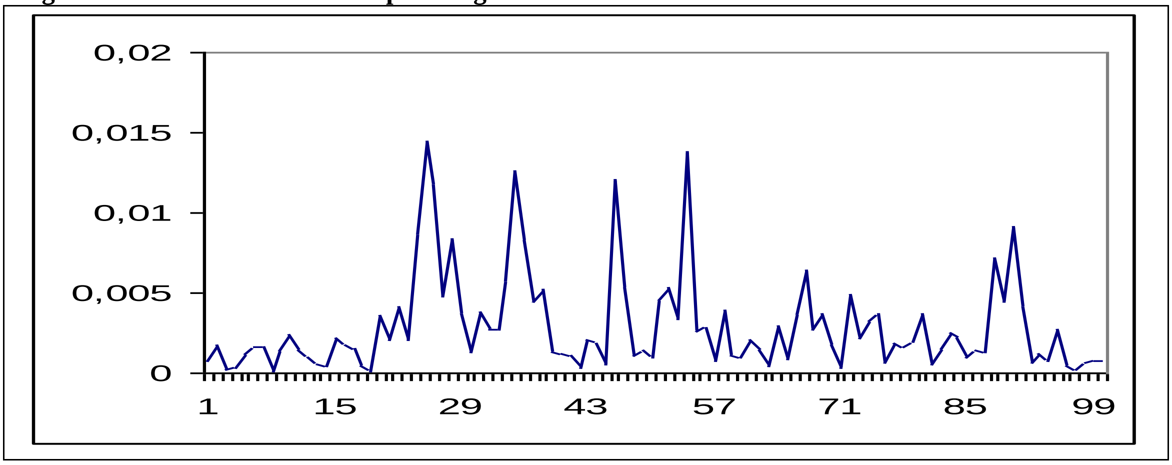


Figure 5: First 100 values in the periodogram



which may be an indication that first differencing is not sufficient to remove the dependence across the data. Moreover, we observe an underlying cyclical pattern in the values. This is also observed in the periodogram showing the largest value at a non-zero frequency. In fact, Figure 5 shows that the largest value correspond to frequency 25, which correspond to T/25 = 74 periods per cycle.

As can be seen based on the unit root tests reported in Table A1 in the Appendix, the decision on whether the log-level of the series is stationary or not is ambiguous, and depends strongly not only on the test conducted, but also on the specification of the test – a result which serves as an indication of fractional integration, given the low power of standard unit root tests to differentiate between near unit root processes from stationary processes or I(d) processes.[1] Furthermore, based on the standard lag-length tests (setting a maximum lag order of 8), as reported in Table A2 in the Appendix, the natural log of the oil price series is shown to have an optimal lag-length of 2, suggesting potential cycles.

Due to the ambiguity in the unit root test results, we then consider the possibility of fractional integration, i.e., we consider a model with a single pole in the spectrum at

[1]Authors such as Diebold and Rudebusch (1991), Hassler and Wolters (1994) and Lee and Schmidt (1996) showed that most commonly used unit root tests have very low power if the alternative are of a fractional form.

the zero frequency, using the classical concept of fractional integration. The model examined is the following:

$$y_t = \beta_1 + \beta_2 t + x_t; \qquad (1 - B)^d x_t = u_t, \quad t = 1, 2, ..., \qquad (3)$$

assuming that the error term is a white noise process, but also an AR(1), an AR(2) and a non-parametric approach developed by Bloomfield (1973) that approximates ARMA structure with a few number of parameters.[2] We present the estimates of d (and their corresponding 95% confidence intervals) for the three standard cases of i) no deterministic terms in the regression model (3)); ii) an intercept (β_1 unknown, and $\beta_2 = 0$); and iii) an intercept with a linear time trend (both β_1 and β_2 unknown). For this purpose, we employ the Whittle function in the frequency domain as suggested in Fox and Taqqu (1986) and Dahlhaus (1989). Additionally, we employ a simple version of the LM tests of Robinson (1994) along with other estimates of d based on the time domain. The results were very similar in all cases and those based on the Whittle function are displayed in Table 1.

Table 1: Estimates of d and 95% confidence interval

	No regressors	An intercept	A linear time trend
White noise	1.175 (1.131, 1.212)	**1.302** **(1.251, 1.356)**	1.175 (1.251, 1.356)
AR(1)	xxx	**0.899** **(0.722, 1.043)**	0.895 (0.774, 1.044)
AR(2)	xxx	**0.896** **(0.771, 1.093)**	0.895 (0.774, 1.044)
Bloomfield-type	1.010 (0.941, 1.107)	**1.053** **(0.982, 1.104)**	0.890 (0.783, 1.089)

In bold, the significant models according to the deterministic terms. xxx means that convergence was not achieved.

It can be seen in Table 1 that, for the case of an intercept (which seems to be the most realistic one according to the t-values, unreported), the estimated value of d is significantly higher than 1 if u_t is white noise; it is below 1 for the cases of AR(1) and

[2]This non-parametric approach accommodates extremely well in the context of fractional integration. For an application of this model in the context of Robinson's (1994) tests, see Gil-Alana (2004).

AR(2) disturbances, though the unit root null (i.e., d = 1) cannot be rejected; and it is slightly above 1 with the model of Bloomfield (1973) and the unit root cannot be rejected either in this case.[3]

Due again to the disparity of the results now depending on the specification of the error term, we also employed a semiparametric approach (Robinson, 1995), which is a Whittle estimate of d based on a band of frequencies that degenerates to zero. The estimate of d is implicitly defined by:

$$\hat{d} = \arg\min_d \left(\log \overline{C(d)} - 2d \frac{1}{b} \sum_{j=1}^{b} \log \lambda_j \right), \qquad (4)$$

$$\text{for } d \in (-1/2, 1/2); \quad \overline{C(d)} = \frac{1}{b} \sum_{j=1}^{b} I(\lambda_j)\, \lambda_j^{2d}, \quad \lambda_j = \frac{2\pi j}{T}, \quad \frac{1}{b} + \frac{b}{T} \rightarrow 0,$$

where T is the sample size, b is the bandwidth parameter, and $I(\lambda_j)$ is the periodogram of the time series, x_t, given by:

$$I(\lambda_j) = \frac{1}{2\pi T} \left| \sum_{t=1}^{T} x_t e^{i \lambda_j t} \right|^2 .$$

Under finiteness of the fourth moment and other mild conditions, Robinson (1995) proved that:

$$\sqrt{b}\,(\hat{d} - d_o) \rightarrow_d N(0, 1/4) \quad \text{as } T \rightarrow \infty,$$

where d_o is the true value of d and with the only additional requirement that $m \rightarrow \infty$ slower than T.

[3] Table A3 in the Appendix reproduces the results with the data starting in 1882: 11, the time a break was found with Bai and Perron's (2003) methodology. The estimates are fairly similar to those reported in Table 1.

Table 2: Estimates of d based on a Whittle semiparametric approach

Bandwidth number	d	Lower I(1) bound	Upper I(1) bound
10	0.872	0.739	1.260
20	0.901	0.816	1.183
30	0.650*	0.849	1.150
40	0.655*	0.869	1.130
43	**0.683***	**0.874**	**1.254**
50	0.712*	0.883	1.115
60	0.737*	0.893	1.106
70	0.785*	0.901	1.098
80	0.830*	0.908	1.091
90	0.852*	0.913	1.086
100	0.913*	0.917	1.082

In bold the estimate corresponding to m = $(T)^{0.5}$; *: Evidence of mean reversion (d < 1) at the 5% level.

The results for a group of values of b (the bandwidth parameter) are displayed in Table 2.[4] Given the potential nonstationary nature of the series examined, the values are estimated using first differences, then adding 1 to obtain the proper orders of integration of the series. It can be seen that the estimated values of d are all below 1 and, for b > 20, the unit root null hypothesis is rejected in favour of mean reversion (d < 1), which contradicts the results based on the parametric approach above.[5]

This lack of consistency in all the above results may be a consequence of model misspecification since the I(d) model presented does not take into consideration the cyclical nature of the series. In what follows, we consider a more general approach taking account both, persistence at the long run or zero frequency and cyclicity.

The model considered here is the one given by equation (1) with z_t =1, i.e.,

$$y_t = \beta + x_t; \qquad (1-B)^{d_1}(1-2\cos w_r B + B^2)^{d_2} x_t = u_t, \quad t = 1, 2, ..., \qquad (5)$$

[4] When choosing the bandwidth there is a trade-off between bias and variance: the asymptotic variance is decreasing whilst the bias is increasing with b. An approach to this problem can be found in Henry (2007). In empirical applications, b = $(T)^{0.5}$ is commonly employed.

[5] Results with data starting in 1882m11 are presented in the Appendix in Table A4.

assuming that the error term is white noise, AR(1) and Bloomfield. The parameters to be estimated are the intercept (β), the two differencing parameters (d_1 and d_2), along with s (s = T/r) referring to the number of time periods per cycle, and those affecting u_t in case of autocorrelated errors. The results are displayed in Table 3.

Table 3: Estimated coefficient in the model given by equation (3)

	s	d_1	d_2	β
White noise	74	0.701 (0.523, 1,327)	0.289 (-0,190, 0,401)	2.994 (5.235)
AR(1)	74	0.579 (0.433, 1386)	0.370 (0.026, 0.460)	2.452 (5.128)
Bloomfield	76	0.688 (0.509, 1,221)	0.223 (-0,211, 0,348)	2,657 (4.332)

In parenthesis in columns 3 and 4, the 95% confidence bands of the differencing parameters. In column 5, the values in parenthesis are the t-values.

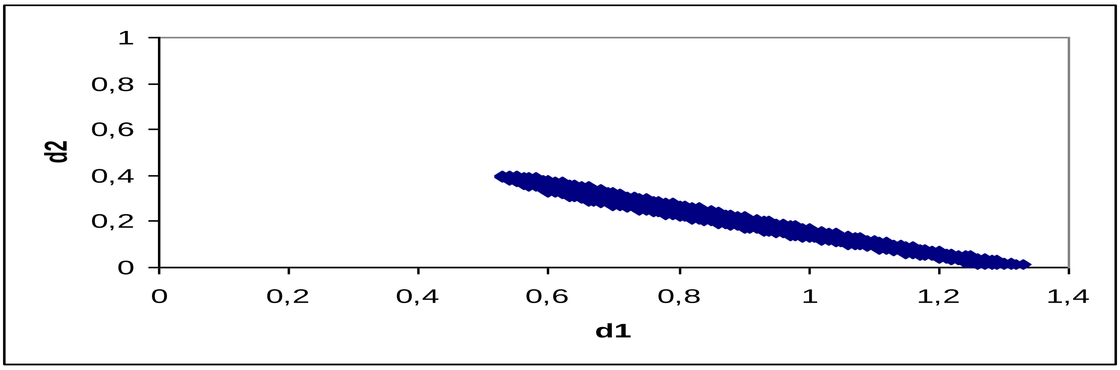
The first thing we observe in this table is that the estimated value of s is 74 for the cases of white noise and AR(1) disturbances, and 76 with the model of Bloomfield (1973), which is completely in line with the largest value detected in the periodogram of the series. This value was 25 which corresponded precisely to 74 (T=1850/25) periods per cycle. If we focus now on the estimated differencing parameters, we see that d_1 ranges between 0.579 (with AR(1) disturbances) and 0.801 (white noise), while d_2 ranges between 0.223 (Bloomfield) and 0.570 (AR(1)). Thus, both degrees of differentiation are fractional being higher at the long run or zero frequency.[6]

Figure 6 displays the (d_1, d_2)-region of non-rejection values using Robinson's (1994) parametric approach for the two parametric cases of white noise and AR(1) disturbances for u_t in equation (5). In other words, we test H_o (2) in the model given by (5), with d_{1o} and d_{2o} values from 0 to 2 with 0.001 increments. Then, we compute the set of (d_{1o},d_{2o}) where the null cannot be rejected. We can see in this figure that there is a

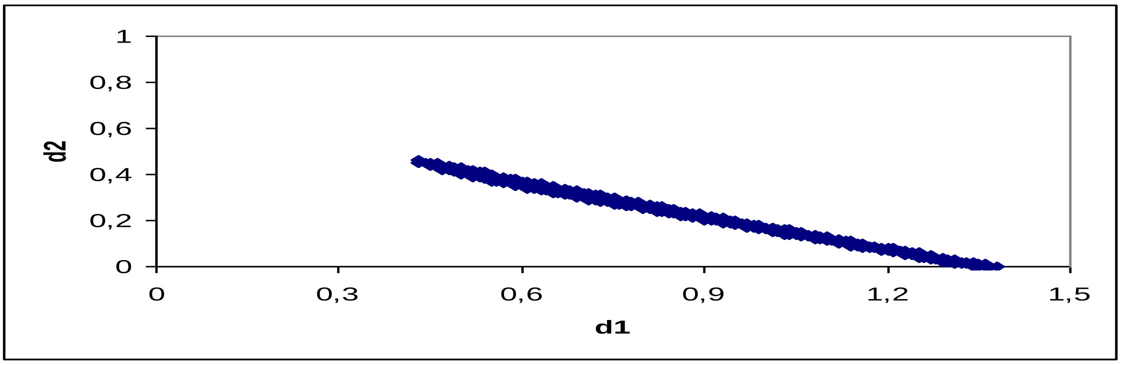
[6] The estimated values of the differencing parameters with data starting in 1882:11 are presented in Table A5 (Appendix) and are once more similar to those based on the whole sample.

Figure 6: (d_1, d_2) region of non-rejection values using Robinson's (1994) approach

i) White noise disturbances



ii) AR(1) disturbances



d_1 refers to the order of integration at the long run or zero frequency, while d_2 refers to the non-zero (cyclical) frequency.

clear negative correlation between the two fractional differencing parameters, the higher the value of d_1 is, smaller the value of d_2 is, implying a kind of competition between the two parameters in describing the dependency in the long run and the cyclical behaviour.

Finally, we compare three different model specifications (an AR(2) process (Table A2), the I(d) one with AR(1) disturbances (d = 0.899, Table 1), and the one given by model (1) with AR(1) disturbances (d_1 = 0.579 and d_2 = 0.370, Table 3) in terms of diagnostic testing and likelihood criteria, and in all cases the latter model was preferred. Note that the AR(2) model is nested in the I(d) one with AR(2) disturbances presented in Table 1, and the null hypothesis of d = 0 is decisively rejected in that table. In the same way, the I(d) model with AR(2) disturbances is nested in the one given by model (5) with AR(2) u_t, and the null of d_2 = 0 is also rejected in such a case. Thus, all the evidence suggests that the long memory model with two differencing parameters may be an adequate specification for the oil data examined in this work.

With respect to the deterministic terms, different models were examined, including segmented trends determined according to different oil prices shocks. In all cases, the estimated coefficients were found to be statistically insignificant. Moreover, using non-linear deterministic trends of the Chebyshev form (Bierens, 1997; Cuestas and Gil-Alana, 2012) produced also insignificant coefficients in virtually all cases.[7]

5. Concluding comments

We have examined in this paper historical oil prices data, namely the West Texas Intermediate (Cushing, Oklahama) crude oil (spot) price, for the time period September 1859 - October 2013, and we focus mainly on two important features of the data; in particular its degree of persistence and its cyclicity. For the former, we focus on standard I(d) models that are characterized by a singularity or pole in the spectrum at the long run or zero frequency. However, a visual inspection of the periodogram of the series shows us that there exists a peak as well as a non-zero frequency, which may be related with the cyclical nature of the data. Thus, we consider a more general model, that contains two differencing parameters, one at the long run or zero frequency, and another one at a non-zero (cyclical) frequency. Using this approach, the estimated parameter at the zero frequency is about 0.58 and the one referring to the cyclical frequency is about 0.37 with the cycles repeating every 6 years approximately.

Our results have two important (policy) implications: (1) We show that the preferred model for oil price is a framework that incorporates both persistence and cyclicity; however, ignoring cyclicity tends to lead to an overestimation of the degree of

[7] The values are available from the authors upon request.

persistence in oil price shocks. Given that oil price is believed to be a leading indicator for growth and inflation of the US economy (Stock and Watson, 2003), an overestimated degree of persistence is likely to send wrong signals to the monetary policy authority, who might feel the need to affect interest rates to mitigate the impact of oil prices on the economy, thinking that the effect of oil prices are going to last longer than they would actually in reality; (2) Also, given that long-term data of oil prices is characterized by cycles of specific length, once the policy maker is aware at which point in the cycle the oil price is, one can predict the future behaviour of oil prices, and the economy in general, given the leading indicator role of oil prices. Thus, the cyclicity component, along with the persistence aspect of oil price, is likely to provide the policy maker information about the impending effect on growth and inflation, and hence, assist in the appropriate policy design.

It can be argued that the analysis presented in this work is simplistic, in the sense that we do not take into account possible alternative features of our long time series data. In particular, we did not carry out a thorough check for the possibility of structural breaks or non-linearities. A preliminary analysis of structural breaks, based on the Bai and Perron (2003) sequential and repartition methodologies, applied to the AR(2) model however, revealed only one break in 1882:10, details of which are available upon request from the authors. Nevertheless, we also conduct the analysis of fractional integration at zero and at zero and the cyclical frequencies with data starting in 1882:11, and the results, reported in the Appendix did not show significant differences with respect to those reported across the paper. Admittedly, these are relevant issues and the connections between fractional integration, non-linearities and structural breaks have recently attracted the attention of many researchers (Smith, 2005; Haldrup et al., 2010; Choi et al., 2010; Haldrup and Kruse, 2014; etc.). Undoubtedly, these important factors will be investigated in future papers. The forecasting properties of the model developed in this paper, will also be examined in the future.

References

Ahrens, W. A. and V.R. Sharma, 1997, Trends in natural resource commodity prices: Deterministic or stochastic?, Journal of Environmental Economics and Management, 33,1, 59-74.

Andel, J., 1986, Long memory time series models, Kybernetika, 22, 105-123.

Arteche, J. and P.M. Robinson, 2000, Semiparametric inference in seasonal and cyclical long memory processes, Journal of Time Series Analysis 21, 1-25.

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

Baillie, R.T., Y.W. Han, R.J. Myers and J. Song, 2007, Long memory models for daily and high frequency commodity future returns, Journal of Future Returns 27, 643-668.

Baumeister, C, 2014, The Art and Science of Forecasting the Real Price of Oil," Bank of Canada Review Spring, 21-31.

Berk, P. and M. Roberts, 1996, Natural resource prices: will they ever turn up?, Journal of Environmental Economics and Management, 31, 65-78.

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., 2001, Complex unit roots and business cycles: Are they real?, Econometric Theory, 17, 962-983.

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

Bollerslev, T. and D. Jubinski, 1999, Equity trading volume and volatility. Latent information arrivals and common long run dependencies, Journal of Business Economic Statistics 17, 1, 9-21.

Bronshtein, I.N. and K.A. Semendyayev, 1998, Handbook of Mathematics, Springer. Springer-Verlag Berlin and Heidelberg GmbH & Co. K.

Caporale, M.G. and Gil-Alana, L.A. 2011, Persistence and cyclical dependence in the monthly Euribor rate, CESifo Working Paper Series, No. 3653.

Caporale, M.G. and Gil-Alana, L.A. 2012b, Persistence and cycles in the US Federal funds rate, CESifo Working Paper Series, No. 4035.

Caporale, M.G. and Gil-Alana, L.A. 2014, Persistence and cycles in US hours worked, Economic Modelling, forthcoming.

Christensen, B.J., M.Ø. Nielsen, and J. Zhu, 2010, Long memory in stock market volatility and the volatility-in-mean effect: the FIEGARCH-M, Journal of Empirical Finance 17, 460-470.

Choi, K., Yu, W.-C. and E. Zivot, 2010, Long memory versus structural breaks in modeling and forecasting realized volatility, Journal of International Money and Finance 29, 857-875.

Cuestas, J.C. and L.A. Gil-Alana, 2012, A Non-linear Approach with Long Range Dependence based on Chebyshev Polynomials, Working Paper 2012-03, Working Paper, University of Sheffield.

Dahlhaus, R., 1989, Efficient parameter estimation for self-similar processes, Annals of Statistics 17, 1749-1766.

Dalla, V. and Hidalgo, J. 2005, A parametric bootstrap test for cycles, Journal of Econometrics, 129, 219-261.

Dickey, D. and Fuller, W. 1981, Likelihood ratio statistics for autoregressive time series with a unit root,Econometrica, 49,1057-1072.Diebold, F.X. and Inoue, A. 2001,Long memory and regime switching. Journal of Econometrics, 105, 131-159.

Diebold, F.X. and Rudebusch, G.D. 1991, On the power of Dickey-Fuller test against fractional alternatives, Economics Letters, 35, 155-160.

Dvir, E. and K.S. Rogoff, 2010, The Three Epochs of Oil, NBER working paper 14927.

Erten, B. and J.A. Ocampo, 2012, Super-cycles of commodity prices since the mid-nineteenth century, UN-DESA working paper No. 110.

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, 1986, Large-sample properties of parameter estimates for strongly dependent stationary Gaussian time series, Annals of Statistics 14, 17-532.

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 the 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 (2), 241-268.

Giriatis, L., J. Hidalgo and P.M. Robinson (2001) Gaussian estimation of parametric spectral density with unknown pole. Annals of Statistics 29, 987-1023.

Giraitis, L. and Leipus, R. 1995, A generalized fractionally differencing approach in long memory modeling, Lithuanian Mathematical Journal 35, 65-81.

Gray, H.L., Yhang, N. and Woodward, W.A. 1989, On generalized fractional processes, Journal of Time Series Analysis 10, 233-257.

Gray, H.L., Yhang, N. and Woodward, W.A. 1994, On generalized fractional processes. A correction, Journal of Time Series Analysis 15, 561-562.

Haldrup, N. and R. Kruse (2014), Discriminating between fractional integration and spurious long memory, CREATES Research Paper 2014-19.

Haldrup, N., F.S. Nielsen and M.Ø. Nielsen, 2010, A vector autoregressive model for electricity prices subject to long memory and regime switching, Energy Economics 32, 1044-1058.

Hamilton, J.D., 2011, Nonlinearities and the macroeconomic effects of oil prices. Macroeconomic Dynamics 15, 472–497.

Hasslers, U. and Wolters J. 1994, On the power of unit root tests against fractional alternatives,Economics Letters, 45, 1-5.

Hauser, M. and R. Kunst, 1998, Fractionally integrated models with ARCH errors. With an application to the Swiss 1-month euromarket interest rate, Review of Quantitative Finance and Accounting 10, 95-113.

Henry, M., 2007, Bandwidth Choice, Optimal Rates and Adaptivity in Semiparametric Estimation of Long Memory, Long Memory in Economics, Part I, 157-172. Eds. G. Teyssière and A. P. Kirman.

Hidalgo, J., 2005, Semiparametric estimation for stationary processes whose spectra have an unknown pole. Annals of Statistics 33(4), 1843–1889.

Huntington, H. Al-Fattah, S.M., Zhou, H., Gucwa, M., and A Nouri, 2013, Oil markets and price movements: A survey of models, 13-129.

Kwiatkowski, D., Phillips, P., Schmidt, P. and Shin, J. 1992, Testing the null hypothesis of stationarity against the alternative of a unit root, Journal of Econometrics, 54:159-178.

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

Lee, J., J. A. List and M.C. Strazicich, 2006, Non-renewable resource prices: Deterministic or stochastic trends?, Journal of Environmental Economics and Management, 51, 3, 354-370.

Magnus, W., Oberhettinger, F. and Soni, R.P. 1966, Formulas and theorems for the special functions of mathematical physics, Springer, Berlin.

Mayoral, L., 2006, Further evidence on the statistical properties of real GNP, Oxford Bulletin of Economics and Statistics 68, 901-920.

Michelacci, C. and P. Zaffaroni, 2000, Fractional Beta Convergence, Journal of Monetary Economics 45, 129-153.

Mu, X. and H. Ye, 2012, Small Trends and Big Cycles in Crude Oil Prices, USAEE Working Paper No. 12-147.

Nelson, C.R. and Plosser, C.I. 1982, Trends and random walks in macroeconomic timeseries, Journal of Monetary Economics 10, 139-162.

Ng, S. and Perron, P. 2001, Lag length selection and the construction of unit root tests with good size and power,Econometrica, 69, 1519-1554.

Phillips, P. And Perron. P. 1988, Testing for a unit root in time series regression, Biometrika, 75, 335–346.

Pindyck, R.S., 1999, The long-run evolution of energy prices, The Energy Journal, 20, 2, 1-27.

Rainville, E.D. 1960, Special functions, MacMillan, New York.

Ray, B. and R. Tsay, 2000, Long range dependence in daily stock volatilities, Journal of Business Economic and Statistics 18, 2, 254-262.

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

Robinson, P.M. 1995, Gaussian semiparametric estimation of long range dependence. Annals of Statistics, 23, 1630-1661.

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

Slade, M.E., 1982, Trends in natural resource commodity prices: an analysis of the time domain, Journal of Environmental Economics and Management, 9, 122-137.

Smith, A., 2005, Level shifts and the illusion of long memory in economic time series, Journal of Business Economic and Statistics 23, 321-239.

Stock, J.H. and M.W. Watson, 2003, Forecasting output and inflation: the role of asset prices, Journal of Economic Literature 41, 788-829.

Zellou, A. and J.T. Cuddington, 2012, Is There Evidence of Super Cycles in Crude Oil Prices?, SPE Economics and Management (forthcoming)

APPENDIX:

Table A1: Unit root test results of the (log-)series

	Intercept			Intercept and Trend		
	Level	1st Difference	Conclusion	Level	1st Difference	Conclusion
ADF	-1.726	-12.486	I(1)	-4.554***	----------	I(0)
PP	-1.717	-28.625	I(1)	-4.496***	----------	I(0)
Ng-Perron	-3.731	-17699.9	I(1)	-3.961	-14160.3	I(1)
KPSS	3.849***	----------	I(0)	0.783***	----------	I(0)

Notes: *** indicates stationarity at 1% level of significance. The critical values are:

- ADF and PP with intercept (Intercept and Trend): -3.433, -2.863 and -2.567 (-3.963, -3.412 and -3.128) at the 1%, 5% and 10% level of significance, respectively;
- Ng-Perron with intercept (Intercept and Trend): -2.58, -1.98 and -1.62 (-3.42, -2.91and -2.62) at the 1%, 5% and 10% level of significance, respectively;
- KPSS with intercept (Intercept and Trend): 0.739, 0.463 and 0.347 (0.216, 0.146 and 0.119) at the 1%, 5% and 10% level of significance, respectively.
- All the tests have a null of unit root barring the KPSS test, which has a null of stationarity.

Table A2: AR Lag Order Selection

Lag	LogL	LR	FPE	AIC	SC	HQ
0	-3111.123	NA	1.721211	3.380905	3.383902	3.382010
1	1821.004	9853.538	0.008115	-1.976104	-1.970110	-1.973894
2	1968.619	294.7496*	0.006921*	-2.135382*	-2.126390*	-2.132067*
3	1968.772	0.304643	0.006927	-2.134462	-2.122472	-2.130041
4	1970.267	2.982487	0.006923	-2.135000	-2.120013	-2.129474
5	1971.084	1.627789	0.006925	-2.134800	-2.116817	-2.128169
6	1971.155	0.142716	0.006932	-2.133792	-2.112811	-2.126056
7	1971.205	0.097862	0.006939	-2.132759	-2.108780	-2.123918
8	1971.910	1.403376	0.006941	-2.132439	-2.105463	-2.122492

Notes: * indicates lag order selected by the criterion;
LR: sequential modified LR test statistic (each test at 5% level);
FPE: Final prediction error;
AIC: Akaike information criterion;
SC: Schwarz information criterion;
HQ: Hannan-Quinn information criterion.

Table A3: Estimates of d and 95% confidence interval with data starting at 1882m10

	No regressors	An intercept	A linear time trend
White noise	1.244 (1.192, 1.313)	**1.255** **(1.205, 1.313)**	1.255 (1.202, 1.313)
AR(1)	0.814 (0.744, 0.919)	**0.822** **(0.756, 0.913)**	0.807 (0.714, 0.919)
AR(2)	0.822 (0.736, 0.956)	**0.834** **(0.736, 0.970)**	0.821 (0.708, 0.972)
Bloomfield-type	0.941 (0.885, 1.027)	**0.939** **(0.897, 1.025)**	0.939 (0.895, 1.025)

In bold, the significant models according to the deterministic terms. xxx means that convergence was not achieved.

Table A4: Estimates of d based on a Whittle semiparametric approach with data starting at 1882m10

Bandwidth number	D	Lower I(1) bound	Upper I(1) bound
10	0.946	0.739	1.260
20	0.987	0.816	1.183
30	0.838*	0.849	1.150
40	0.815*	0.869	1.130
43	**0.832***	**0.874**	**1.254**
50	0.811*	0.883	1.115
60	0.830*	0.893	1.106
70	0.835*	0.901	1.098
80	0.837*	0.908	1.091
90	0.854*	0.913	1.086
100	0.875*	0.917	1.082

In bold the estimate corresponding to m = $(T)^{0.5}$; *: Evidence of mean reversion (d < 1) at the 5% level.

Table A5: Estimated coefficient in the model given by equation (3) with data starting at 1882m10

	s	d_1	d_2	β
White noise	24	0.688 (0.503, 1,311)	0.217 (-0,134, 0,345)	2.336 (5.115)
AR(1)	21	0.599 (0.411, 1.321)	0.334 (0.017, 0.433)	2.151 (5.009)
Bloomfield	22	0.612 (0.511, 1,112)	0.139 (-0,200, 0,312)	2,037 (4.111)

In parenthesis in columns 3 and 4, the 95% confidence bands of the differencing parameters. In column 5, the values in parenthesis are the t-values.