# PERSISTENCE AND CYCLES
# IN THE US FEDERAL FUNDS RATE

**Guglielmo Maria Caporale**
**Brunel University, London, UK**

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

**January 2016**

## Abstract

This paper uses long-range dependence techniques to analyse two important features of the US Federal Funds effective rate, namely its persistence and its cyclical behaviour. It examines annual, monthly, bi-weekly and weekly data, from 1954 until 2010. Two models are considered. One is based on an I(d) specification with AR(2) disturbances and the other one on two fractional differencing structures, one at the zero and the other at a cyclical frequency. Thus, the two approaches differ in the way the cyclical component of the process is modelled. In both cases we obtain evidence of long memory and fractional integration with cycles repeating approximately every 8 years. The in-sample goodness-of-fit analysis supports the second specification in the majority of cases. An out-of-sample forecasting experiment also suggests that the long-memory model with two fractional differencing parameters is the most adequate one, especially over long horizons.

**Keywords:** Federal Funds rate, persistence, cyclical behaviour, fractional integration

**JEL classification:** C32, E43

**Corresponding author:** Professor Guglielmo Maria Caporale, Department of Economics and Finance, Brunel University London, UB8 3PH, UK. Tel.: +44 (0)1895 266713. Fax: +44 (0)1895 269770. Email: Guglielmo-Maria.Caporale@brunel.ac.uk

# 1.    Introduction

The Federal Funds rate is the interest rate at which depository institutions in the US lend each other overnight (normally without a collateral) balances held at the Federal Reserve System (the Fed), which are known as Federal Funds. Such deposits are held in order to satisfy the reserve requirements of the Fed. The rate is negotiated between banks, and its weighted average across all transactions is known as the Federal Funds effective rate. It tends to be more volatile at the end of the reserve maintenance period, the so-called settlement Wednesday, when the requirements have to be met.[1] The Federal Funds target rate is instead set by the Chairman of the Fed according to the directives of the Federal Open Market Committee (FOMC), which holds regular meetings (as well as additional ones when appropriate) to decide on this target. It is therefore a policy rate, used to influence the money supply, and to make the effective rate (which by contrast is determined by the interaction of demand and supply) follow it. Specifically, the Trading Desk of the Federal Reserve Bank of New York conducts open market operations on the basis of the agreed target. This is considered one of the most important indicators for financial markets, whose expectations can be inferred from the prices of option contracts on Federal Funds futures traded on the Chicago Board of Trade.

Given the fact that the Fed implements monetary policy by setting a target for the effective Federal Funds rate which also affects other linked interest rates and the real economy through various transmission channels, it is not surprising that both the theoretical and the empirical literature on this topic are extensive. Theoretical contributions include a well-known paper by Bernanke and Blinder (1988), who propose a model of monetary policy transmission which they then test in a follow-up study (Bernanke and Blinder, 1992) showing that the Federal Funds rate is very useful to forecast real

---

[1] In empirical studies, therefore, the series is often adjusted to eliminate this effect (see, e.g., Sarno and Thornton, 2003).

macroeconomic variables, being a good indicator of monetary policy actions. Bartolini et al. (2002) instead develop a model of the interbank money market with an explicit role for central bank intervention and periodic reserve requirements that is consistent with the observed volatility pattern of the US Federal Funds rate.

On the empirical side, some papers examine the extent to which variables targeted by the Fed such as the output gap and inflation can explain the effective rate (see, e.g., Taylor, 1993 and Clarida et al., 2000); others analyse the daily market for Federal Funds (e.g., Hamilton, 1996, and Taylor, 2001). An influential study by Hamilton and Jorda (2002) introduced the autoregressive conditional hazard model for forecasting a discrete-valued time series such as the target; this specification is shown to outperform standard VAR models that are unable to differentiate between the effects of an increase in the target and those of an anticipated target decrease that did not take place. Other studies examine the predictive power of the effective rate of the target (Taylor, 2001) or other interest rates (e.g., Clarida et al., 2006). Sarno et al. (2005) provide the most extensive study of the forecasting performance of a variety of models of the Federal Funds rate proposed in the literature. They consider both univariate (randow walk, ARMA, EGARCH, Markov-switching etc.) and multivariate (M-TAR, BTAR, MS-VECM) specifications, and find that the best forecasting model is a univariate one using the current difference between the effective and the target rate to forecast the future effective rate (also, combination forecasts only yield marginal improvements). These findings are interpreted as suggesting that the Fed in fact follows a forward-looking interest rate rule.

Most of the models found in the literature to describe the behaviour of the Federal Funds rate (and of interest rates in general) assume nonstationarity and are based on first-differenced series. This is true, for instance of all the univariate specifications considered in Sarno et al. (2005), which imply that the series are I(1), without mean reversion and

3

with permanent effects of shocks. This is a rather strong assumption that is not justified on theoretical grounds. The classic alternative is to assume that the Federal Funds rate and interest rates in general are stationary I(0) variables, and to model them as autoregressive processes with roots close to the unit circle to allow for a high degree of persistence, with the additional problem of the well-known low power of standard unit root tests in such a case. In this study we overcome this dichotomy by estimating fractional integration models allowing for both nonstationary and mean-reverting behaviour.[2]

Using recent techniques based on the concept of long-range dependence, in this paper we explicitly model two well-known features of interest rates in general which also appear to characterise the Federal Funds rate, namely their persistence and cyclical behaviour, mostly overlooked in previous studies.[3] In particular, we use fractional integration methods with multiple poles or singularities in the spectrum not constrained at the zero frequency as in the usual case, but allowing instead for poles at zero and non-zero (cyclical) frequencies. In this way we are able to capture the two aforementioned features of interest rates: their high degree of persistence (described by the pole in the spectrum at the zero frequency) and their cyclical pattern (described by the pole at the non-zero frequency).[4] Overall, our results confirm that both these stylised facts are important features of the stochastic behaviour of the Federal Funds rates. Sensitivity to data frequency is then analysed by using annual, monthly, bi-weekly and weekly data, from 1955 until 2010 (see Caporale and Gil-Alana, 2010, showing that lower integration order might correspond to lower frequencies).

---

[2] Gil-Alana (2002, 2004) also modelled the US Federal Funds rate using fractional integration,. However, in those two papers the analysis focused exclusively on the long run or zero frequency, without taking into account the cyclical component of the series.

[3] These two features are expected in the Federal Funds rate which is commonly modelled using a Taylor rule and therefore is a linear combination of inflation, which is highly persistent (Piveta and Reis, 2007) and the output gap, which is cyclical.

[4] Other sources of persistence/cyclical patterns are described by the short-run (ARMA) dynamics of the process.

The remainder of the paper is structured as follows. Section 2 describes the econometric approach. Section 3 presents the empirical results. Section 4 provides some concluding remarks.

## 2. Methodology

We consider the following model:

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

$$(1 - L)^{d_1}(1 - 2\cos w_r L + L^2)^{d_2} x_t = u_t, \quad t = 1, 2, ..., \tag{2}$$

with $x_t = 0$, $t \leq 0$, and where $y_t$ is the observed time series; $\beta$ is a (kx1) vector of unknown parameters, and $z_t$ is a (kx1) vector of deterministic terms, that might include, for example, an intercept (i.e. $z_t = 1$) or an intercept with a linear trend ($z_t = (1,t)^T$); L is the lag operator (i.e., $L^s x_t = x_{t-s}$); $d_1$ is the order of integration corresponding to the long-run or zero frequency; $w_r = 2\pi/r$, $r = T/j$, with r representing the number of periods per cycle and j indicating the frequency with a pole at the spectrum; $d_2$ is the order of integration with respect to the non-zero (cyclical) frequency, and $u_t$ is assumed to be an I(0) process, defined for the purposes of the present study as a covariance-stationary process, with a spectral density function that is positive and finite at any frequency on the spectrum. Note that $d_1$ and $d_2$ are allowed to be any real values and thus are not restricted to be integers.

The set-up described in (1) and (2) is fairly general, including the standard ARMA model (with or without trends), if $d_1 = d_2 = 0$ and $u_t$ is weakly autocorrelated; the I(1) model if $d_1 = 1$ or, more generally, the ARIMA case if $d_1$ is an integer and $d_2 = 0$; the standard ARFIMA specification, if $d_1$ has a fractional value and $d_2 = 0$, along with other more complex representations.

We now focus on equation (2), and first assume that $d_2 = 0$. Then, for any $d_1 > 0$, the spectral density function of $x_t$ is given by

$$f(\lambda) = \frac{\sigma^2}{2\pi} |g_u(\lambda)|^2 |1 - e^{i\lambda}|^{-2d_1},$$

where $g_u(\lambda)$ corresponds to the potential ARMA structure in $u_t$. It can be easily shown that this function $f(\lambda)$ contains a pole or singularity at the long-run or zero frequency, i.e.,

$$f(\lambda) \rightarrow \infty, \quad as \quad \lambda \rightarrow 0^+.$$

Further, note that the polynomial $(1 - L)^{d_1}$ can be expressed in terms of its Binomial expansion, such that, for all real $d_1$,

$$(1 - L)^{d_1} = \sum_{j=0}^{\infty} \binom{d_1}{j} (-1)^j L^j = 1 - d_1 L + \frac{d_1 (d_1 - 1)}{2} L^2 - \dots, \quad (3)$$

implying that the higher the value of $d_1$ is, the higher the degree of dependence between observations distant in time will be. Thus, the parameter $d_1$ plays a crucial role in determining the degree of long-run persistence of the series. Estimation and testing procedures in this context include Sowell (1992a), Agiakloglou and Newbold (1994) and Beran (1995) in the time domain, and Geweke and Porter-Hudak (1983), Robinson (1994, 1995a,b) among others in the frequency domain. Examples of applications using this model can be found in Diebold and Rudebusch (1989), Sowell (1992b), Baillie (1996), and Gil-Alana and Robinson (1997) among many others.[5]

On the other hand, if $d_1 = 0$ in (2), then for any $d_2 \neq 0$, the process $x_t$ has a spectral density function

$$f(\lambda) = \frac{\sigma^2}{2\pi} |g_u(\lambda)|^2 |2(\cos(\lambda) - \cos(w_r))|^{-2d_2},$$

which is characterised by a pole at a non-zero frequency, i.e.,

$$f(\lambda) \rightarrow \infty, \quad as \quad \lambda \rightarrow \lambda^*, \quad \lambda^* \in (0, \pi).$$

---

[5] Empirical studies estimating I(d) models of this form for interest rates include Lai (1997), Tsay (2000), Meade and Maier (2003) and Couchman, Gounder and Su (2006).

Moreover, the polynomial $(1 - 2\cos w_r L + L^2)^{d_2}$ can be expressed as a Gegenbauer polynomial, such that, defining $\mu = \cos w_r$, for all $d_2 \neq 0$,

$$(1 - 2\mu L + L^2)^{-d_2} = \sum_{j=0}^{\infty} C_{j,d_2}(\mu) L^j, \qquad (4)$$

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

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

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

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

(see, inter alia, Magnus et al., 1966, or Rainville, 1960, for further details). Gray et al. (1989, 1994) showed that this process is stationary if $d_2 < 0.5$ for $|\mu = \cos w_r| < 1$ and if $d_2 < 0.25$ for $|\mu| = 1$. Estimation methods include the parametric approach of Giraitis et al. (2001, and the semiparametric method of Hidalgo (2005) and Wald (W) and Lagrange Multiplier (LM) tests of short memory against long memory as in (2) (with d1 = 0) can be found in Dalla and Hidalgo (2005). If $d_2 = 1$, the process is said to contain a unit root cycle (Ahtola and Tiao, 1987; Bierens, 2001); and other empirical applications using fractional values of d2 can be found in Gil-Alana (2001), Anh, Knopova and Leonenko (2004) and Soares and Souza (2006).

In the empirical analysis carried out in the following section we use a very general procedure to test the model given by equations (1) and (2). It was initially developed by Robinson (1994) on the basis of the Lagrange Multiplier (LM) principle that uses the Whittle function in the frequency domain. It can be applied to test the null hypothesis:

$$H_o: \ d \equiv (d_1, d_2)^T = (d_{1o}, d_{2o})^T \equiv d_o, \qquad (5)$$

7

in (1) and (2) where $d_{10}$ and $d_{20}$ can be any real values, thus encompassing stationary and nonstationary hypotheses. The specific form of the test statistic (denoted by $\hat{R}$) is presented in the Appendix. Under very general regularity conditions, Robinson (1994) showed that for this particular version of his tests,

$$\hat{R} \rightarrow_d \chi_2^2, \quad as \quad T \rightarrow \infty, \tag{6}$$

where T indicates the sample size, and "$\rightarrow_d$" stands for convergence in distribution. Thus, unlike in other procedures, we are in a classical large-sample testing situation. A test of (5) will reject $H_o$ against the alternative $H_a$: $d \neq d_o$ if $\hat{R} > \chi_{2,\alpha}^2$, where Prob ($\chi_2^2 > \chi_{2,\alpha}^2$) = $\alpha$. Note that, as mentioned before, when testing the null $H_o$: $d_2 = 0$, the model reduces to the standard case of fractional integration at the zero frequency as in Gil-Alana and Robinson (1997), and testing $H_o$ (5) with $d_1 = 0$ becomes a test for the order of integration at a cyclical frequency as in Gil-Alana (2001).

There are several reasons for using this approach. First, this test is the most efficient in the Pitman sense against local departures from the null, that is, if it is implemented against local departures of the form: $H_a$: $d = d_o + \delta T^{-1/2}$, for $\delta \neq 0$, the limit distribution is a $\chi_2^2(v)$, with a non-centrality parameter v that is optimal under Gaussianity of $u_t$. Moreover, Gaussianity is not necessary for the implementation of this procedure, a moment condition of only order 2 being required. As in other standard large-sample testing situations, Wald and LR test statistics against fractional alternatives have the same null and limit theory as the LM test of Robinson (1994). Lobato and Velasco (2007) essentially employed such a Wald testing procedure, even though it requires a consistent estimate of d and concerns exclusively the long run or zero frequency; therefore the LM test of Robinson (1994) used here seems computationally more attractive.

## 3.    Empirical results

The series examined is the US Federal Funds effective rate, from 1954 till 2010, at annual,
monthly, bi-weekly and weekly frequencies. The source of the data is the Federal Reserve
Bank of St. Louis database.

**[Insert Figures 1 – 3 about here]**

Figure 1 displays plots of the series at the four frequencies considered, the pattern
being similar in all four cases.[6] Note that during the years 2009 and 2010 there was a
downward movement in the Federal Funds rate, which hit the zero lower bound (ZLB) and
stayed between 0% and 0.25%. Removing these two years of observations, however, did
not affect the overall picture. Figure 2 displays the correlograms; the two features
mentioned above can clearly be seen: there is a slow decay in the sample autocorrelation
values possibly due to persistence, and a cyclical pattern. The same two features are
exhibited by the periodograms in Figures 3a and 3b, with the highest peaks occurring at the
smallest frequency (long-run persistence) and at frequency 7 corresponding to T/7 periods
per cycle, namely to approximately 8 years in all cases.

First, we examine the degree of persistence considering only the long-run or zero
frequency, that is, we specify a model such as (1) and (2) with $d_2 = 0$ a priori and $z_t = (1,t)^T$, $t \geq 1$; $(0, 0)^T$, otherwise, i.e.,

$$y_t = \beta_0 + \beta_1 t + x_t, \qquad t = 1, 2,..., \tag{7}$$

$$(1 - L)^{d_1} x_t = u_t, \qquad t = 1, 2, ..., \tag{8}$$

with $x_t = 0$ for $t \leq 0$, under the assumption that the disturbance term $u_t$ is white noise,
AR(1) and AR(2) respectively.[7] The latter specification is used to describe the cyclical
component throughout the potential presence of complex roots in the AR polynomial.

---

[6] The monthly series are aggregated daily data. In the case of the weekly and bi-weekly series we use the
Wednesday observations.

[7] Although a time trend appears not to be required (see Figure 1), we have included it to consider a more
general case. Note also that, when $d_1 = 1$, the time trend disappears for $t > 1$.

Higher AR orders and other MA (ARMA) structures were also considered, with very similar results. We employ here a simple version of Robinson's (1994) procedure, testing $H_o$: $d_1 = d_{1o}$, for $d_{1o}$-values from 0 to 2 with 0.001 increments, (i.e., $d_{1o} = 0, 0.001, 0.002,$ …, 1.999 and 2), and reporting the estimates of $d_1$ along with the 95% confidence intervals of the non-rejection values of $d_1$ based on the testing procedure.[8]

**[Insert Table 1 about here]**

We obtain estimates for the three standard cases examined in the literature, i.e., with no regressors in the undifferenced regression (7) ($\beta_0 = \beta_1 = 0$); with an intercept ($\beta_0$ unknown and $\beta_1 = 0$); and with an intercept and a linear time trend ($\beta_0$ and $\beta_1$ unknown). The results for the time trend were found to be statistically insignificant in all cases, while the intercept was always significant. Thus, in what follows, we only consider the case of an intercept.[9] Table 1 displays the estimates of $d_1$ based on the Whittle function in the frequency domain (Dahlhaus, 1989) along with the 95% confidence interval of the non-rejection values of $d_1$ using Robinson's (1994) method.[10]

When $u_t$ is assumed to be a white noise process, the results change substantially depending on the data frequency. In particular, for annual data the estimated value of $d_1$ is 0.937 and the I(1) null hypothesis cannot be rejected. It is rejected instead for monthly and bi-weekly data in favour of values of $d_1$ above 1. Finally, for weekly data, the estimated $d_1$ is significantly smaller than 1, implying mean reversion. When allowing for autocorrelated errors, if $u_t$ is assumed to be AR(1) values of $d_1$ below 1 supporting mean reversion are obtained in the annual and monthly cases; for bi-weekly and weekly data, $d_1$ is instead slightly above 1 and the unit root null is rejected in favour of $d_1 > 1$ in the weekly case. Finally, if $u_t$ is assumed to be AR(2) the unit root cannot be rejected in any single case and

---

[8] See Gil-Alana and Robinson (1997) for details of the testing procedure in this context.
[9] Note that if $d_1 = 1$ and $u_t$ is white noise, for $t > 1$ the model becomes the simple driftless random walk model.
[10] The main results did not change when considering 90% and 99% intervals.

the estimated values of $d_1$ range between 0.722 (with annual data) and 1.045 (weekly data). As mentioned before, the case of AR(2) disturbances is interesting because it allows to capture the cyclical pattern of the series through a short-memory I(0) process for $u_t$.[11]

Likelihood Ratio (LR) tests and other likelihood criteria (not reported) suggest that the model with AR(2) disturbances outperforms the others. These results, however, might be biased owing to the long memory in the cyclical structure of the series having been overlooked. Thus, in what follows we examine the possibility of long memory but focusing on the case where the spectrum has a pole or singularity at a non-zero (cyclical) frequency. For this purpose and as a preliminary step, we employ the LM parametric test of Dalla and Hidalgo (2005), and given the results displayed in Table 1 we use the first differenced data. We consider different specifications for the disturbance term and in all cases the null hypothesis of no cycles is rejected at the 5% level.

Once we have determined the presence of a pole or singulatiry at a non-zero (cyclical) frequency, we next implement the methods proposed in Giraitis et al. (2001) and Hidalgo (2005) to estimate the differencing parameter and the pole in a model of the form given by (2) with $d_1 = 0$. The first of these methods is parametric and we impose white noise disturbances for the error term $u_t$, while the second is semiparametric and no functional form is assumed for $u_t$. The analysis was conducted using both the original data and the first differences and the results were consistent in the two cases. Specifically, for the original data the estimated pole was found at the zero frequency, although with the Giraitis et al.'s (2001) method the order of integration changes substantially depending on the frequency of the data. For the first differenced data we obtained estimated values for the poles at frequencies 8, 98, 210 and 426 respectively for the annual, monthly, bi-weekly

---

[11] The roots of the AR(2) polynomials (not reported) were in all cases in the complex plane, which is consistent with the cyclical pattern observed in the data. When using other approaches such as Sowell's (1992a) maximum likelihood method, the results based on the AIC also suggest that the ARIMA(2, d, 0) is the most adequate specification.

and weekly data, and once again the order of integration changes substantially across frequencies. This lack of robustness could be partly due to the fact that we do not allow for any degree of flexibility with respect to the order of integration at the zero frequency, which is imposed to be 0 or 1 respectively for the original and first differenced data. Thus, in what follows we consider a more general model that allows for fractional degrees of integration at both the zero and the cyclical frequencies. In particular, we use a model such as (1) and (2), again with $z_t = (1,t)^T$, $t \geq 1$; $(0, 0)^T$, otherwise, i.e., the model now becomes:

$$y_t = \beta_0 + \beta_1 t + x_t, \tag{9}$$

$$(1 - L)^{d_1}(1 - 2\cos w_r L + L^2)^{d_2} x_t = u_t, \tag{10}$$

again with I(0) (potentially ARMA) $u_t$. Note that this model includes the previous one of long memory at the zero frequency if $d_2 = 0$. The results, for the case of an intercept, which is the most realistic one on the basis of the t-values (not reported), are displayed in Table 2.

**[Insert Table 2 about here]**

The estimated values of j and thus r = T/j (the number of periods per cycle) for the four series is now close to 8 years. Specifically, j is found to be 8 in the case of the annual data; 97 (and thus 97/12 = 8.089 years) for the monthly data; and 8.15 years (212/26 and 424/52) for bi-weekly and weekly data. This is consistent with the plots of the periodograms displayed in Figure 3 and with the results obtained applying the Giraitis et al. (2001) and Hidalgo (2005) methods to the first differenced data. Focusing now on the fractional differencing parameters, it can be seen that $d_1$ is close to (although below) 1 and $d_2$ is slightly above 0 for the four series. For $d_1$ the unit root null is rejected in favour of mean reversion in the case of annual, bi-weekly and weekly data; however, for monthly data, even though $d_1$ is still below 1, the unit root null cannot be rejected at conventional significance levels. As for the cyclical fractional differencing parameter, $d_2$, is estimated to be 0.094 in the annual case and the I(0) null hypothesis cannot be rejected. In the

12

remaining three cases, $d_2$ is significantly above 0 (thus displaying cyclical long memory), ranging from 0.145 (weekly data) to 0.234 (monthly data). Very similar values for $d_1$ and $d_2$ are obtained in the case of autocorrelated disturbances; LR and no-autocorrelation tests strongly support the white noise specification for $u_t$ for each of the four series.[12]

Finally, we investigate which of the two specifications (the I(d) one with AR(2) disturbances or the one with the two fractional differencing structures) has a better in-sample performance, and also better forecasting properties. For the first of these two purposes we employ several goodness-of-fit measures based on the likelihood function. For the forecasting experiment, we use instead various statistics including the modified Diebold and Mariano (1995) (M-DM) statistic. Remember that the two models considered are:

$$y_t = \beta_0 + x_t; \qquad (1-L)^{d_1} x_t = u_t; \qquad u_t = \phi_1 u_{t-1} + \phi_2 u_{t-2} + \varepsilon_t, \qquad \text{(M1)}$$

and

$$y_t = \beta_0 + x_t; \quad (1-L)^{d_1}(1 - 2\cos w_r L + L^2)^{d_2} x_t = \varepsilon_t, \qquad \text{(M2)}$$

and therefore they differ in the way the cyclical component is modelled, model (M1) and (M2) adopting respectively an AR(2) process and a Gegenbauer (fractional) specification for the $d_1$-differenced (demeaned) series.

For the in-sample goodness of fit analysis we carry out first a Likelihood Ratio (LR) test, comparing (M1) with (M2) with AR(2) errors, which is clearly nested in (M1). Note that using in (M2) the equations given by (9) and (10) with $\beta_1 = 0$, $d_2 = 0$ and AR(2) $u_t$ we obtain (M1). The results support the modified (M2) specification for three of the four series examined - the AR(2) coefficients were found to be statistically insignificant suggesting model (M2). Only for the annual data (M1) seems to be preferable at the 5% level. This is consistent with the results displayed in Tables 1 and 2, noting that the only

---

[12] We use here the Box-Pierce and Ljung-Box-Pierce statistics (Box and Pierce, 1970; Ljung and Box, 1978).

confidence interval in Table 2 where $d_2 = 0$ is not excluded is precisely that for the annual series. Other likelihood criteria (AIC and SIC) lead essentially to the same conclusions.[13]

Next we focus on the forecasting performance of the two models. For this purpose we calculate one- to twenty-step ahead forecasts for each of the four series at four different data frequencies. The forecasts were constructed according to a recursive procedure conditionally upon information available up to the forecast date which changes recursively.

We computed the Root Mean Squared Errors (RMSE) and the Mean Absolute Deviation (MAD) for the two specifications of each series. The results (not reported here for reasons of space, but available from the authors upon request) indicate that the fractional structure outperforms the AR(2) model in practically all cases.[14]

However, the above two criteria and other methods such as the Mean Absolute Prediction Error (MAPE), Mean Squared Error (MSE), etc., are purely descriptive devices.[15] Several statistical tests for comparing different forecasting models are now available. One of them, widely employed in the time series literature, is the asymptotic test for a zero expected loss differential due to Diebold and Mariano (1995).[16] Harvey, Leybourne and Newbold (1997) note that the Diebold-Mariano test statistic could be seriously over-sized as the prediction horizon increases, and therefore provide a modified Diebold-Mariano test statistic (M-DM) given by:

$$M - DM = DM \sqrt{\frac{n + 1 - 2h + h(h-1)/n}{n}},$$

---

[13] Note, however, that these criteria might not necessarily be the best criteria in applications involving fractional differences, as they focus on the short-term forecasting ability of the fitted model and may not give sufficient attention to the long-run properties of the fractional models (see, e.g. Hosking, 1981, 1984 and more recently, Beran et al., 1998).

[14] These two models also outperform the random walk model in all cases, which is not surprising since this hypothesis was decisively rejected with the tests of Robinson (1994) in practically all cases.

[15] The accuracy of different forecasting methods is a topic of continuing interest and research (see, e.g., Makridakis et al., 1998 and Makridakis and Hibon, 2000, for a review of the forecasting accuracy of competing forecasting models).

[16] An alternative approach is the bootstrap-based test of Ashley (1998), though his method is computationally more intensive.

where DM is the original Diebold-Mariano statistic, h is the prediction horizon and n is the time span for the predictions. Harvey et al. (1997) and Clark and McCracken (2001) show that this modified test statistic performs better than the DM test statistic, and also that the power of the test is improved when p-values are computed with a Student t-distribution.

We further evaluate the relative forecast performance of the different models by making pairwise comparisons based on the M-DM test statistic. We consider here 5, 10, 15, 20 and 25-period ahead forecasts. The results are displayed in Table 3.

**[Insert Table 3 about here]**

They show that for the 5-step and 10-step ahead predictions it cannot be inferred that one model is statistically superior to the other. By contrast, over longer horizons there are several cases where the fractional model (M2) outperforms (M1). As a final remark, we should note that the M-DM test may have very low power under some circumstances, especially in the case of non-linear models (see, e.g., Costantini and Künst, 2011). Thus, these results here should be taken with caution.[17]

### 4. Conclusions

This paper uses long-range dependence techniques to analyse two important features of the US Federal Funds effective rate, namely its persistence and its cyclical behaviour. In particular, it examines annual, monthly, bi-weekly and weekly data, from 1954 until 2010. The main results are the following. When estimating a simple I(d) model, the estimates suggest that d is close to 1, in some cases below 1 indicating mean reversion, and in others above 1 implying a rejection of the I(1) hypothesis depending on the data frequency and the type of disturbances considered (white noise or AR(1)). If these are modelled as AR(2), which is an appropriate specification according to the likelihood criteria and also highly

---

[17] Using non-linear deterministic trends of the Chebyshev form (Bierens, 1997; Cuestas and Gil-Alana, 2012) produced insignificant coefficients in virtually all cases.

plausible in view of the cyclical pattern of the series under examination, the results indicate that the I(1) null cannot be rejected at any of the four frequencies. The LM test of Dalla and Hidalgo (2005) on the first differenced data suggests the existence of strong cycles at a non-zero frequency, which is estimated to be at around 8 years in the four cases considered. In order to allow for a more flexible specification, we consider a model that incorporates the zero and the non-zero poles into the analysis, estimating simultaneously the two differencing parameters, one corresponding to the long-run or zero frequency ($d_1$), and the other one to the cyclical structure ($d_2$). When using this specification the results indicate that the order of integration at the zero frequency ranges between 0.802 (bi-weekly frequency) and 0.966 (monthly), whilst that of the cyclical component ranges between 0.094 (annual) and 0.234 (bi-weekly), with the cycles repeating themselves every 8 years. Both the in-sample and out-of-sample evidence suggest that the long memory model with two fractional structures (one at zero and the other at the cyclical frequency) outperforms the other models.

Our results are not directly comparable to those of Sarno et al. (2005), who examine different specifications for the difference between the effective and the target rate (finding that a variation of Taylor's (2001) model is the best among them), whilst we focus on the effective rate itself. Nevertheless, our analysis produces valuable evidence for interest rate modelling, since it shows that an I(d) specification including a cyclical component outperforms both classical I(0) and I(1) models and simple I(d) representations. This confirms the importance of adopting an econometric framework such as the one chosen here, which explicitly takes into account both persistence and cyclical patterns, to model the behaviour of the US Federal Funds effective rate and interest rates in general. Finally, as mentioned before, non-linearities might also be present in the data, in particular in the cyclical structure. Figure 1 shows a big spike in the middle of the sample, which could be

an indication of time variation in the parameters. Therefore we also conducted the analysis for two subsamples, but the results did not substantially differ in terms of the number of periods per cycle. Moreover, our findings appear to be robust across data frequencies. Since fractional integration and potential breaks are issues which are intimately related, (Diebold and Inoue, 2001; Granger and Hyung, 2004), it would also be interesting to apply the method of Ohanissian et al. (2008) that uses different frequencies. All these issues will be examined in future papers.

## Appendix

The test statistic proposed by Robinson (1994) for testing $H_o$ (5) in the model given by equations (1) and (2) is given by:

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

where T is the sample size, and

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

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

$$\psi(\lambda_j)' = [\psi_1(\lambda_j), \psi_2(\lambda_j)]; \quad \hat{\varepsilon}(\lambda_j) = \frac{\partial}{\partial\tau}\log g_u(\lambda_j;\hat{\tau}); \quad \psi_1(\lambda_j) = \log\left|2\sin\frac{\lambda_j}{2}\right|;$$

$\psi_2(\lambda_j) = \log\left|2(\cos\lambda_j - \cos w_r)\right|$, with $\lambda_j = 2\pi j/T$, and the summation in * is over all frequencies which are bounded in the spectrum. $I(\lambda_j)$ is the periodogram of

$$\hat{u}_t = (1-L)^{d_{1o}}(1 - 2\cos w_r L + L^2)^{d_{2o}}y_t - \hat{\beta}'\bar{z}_t, \text{ with}$$

$$\hat{\beta} = \left(\sum_{t=1}^T\bar{z}_t\bar{z}_t'\right)^{-1}\sum_{t=1}^T\bar{z}_t(1-L)^{d_{1o}}(1 - 2\cos w_r L + L^2)^{d_{2o}}y_t;$$

$\bar{z}_t = (1-L)^{d_{1o}}(1 - 2\cos w_r L + L^2)^{d_{2o}}z_t$, evaluated at $\lambda_j = 2\pi j/T$ and $\hat{\tau} = \arg\min_{\tau \in T^*}\sigma^2(\tau)$, with $T^*$ as a suitable subset of the $R^q$ Euclidean space. Finally, the function $g_u$ above is a known function coming from the spectral density of $u_t$:

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

Note that this test is purely parametric and, therefore, it requires specific modelling assumptions about the short-memory specification of $u_t$. Thus, if $u_t$ is white noise, $g_u \equiv 1$,

and if $u_t$ is an AR process of the form $\phi(L)u_t = \varepsilon_t$, $g_u = |\phi(e^{i\lambda})|^{-2}$, with $\sigma^2 = V(\varepsilon_t)$, so that the AR coefficients are a function of $\tau$.

The point estimates were obtained by choosing the values that minimise Robinson's (1994) test statistic over a grid of values for $d_1$, $d_2$ and r. These parameter estimates were practically identical to those obtained by maximising the Whittle function in the frequency domain (Dahlhaus, 1989). The confidence intervals were obtained by choosing the values of the differencing parameters were the null hypotheses tested could not be rejected at the 5% level.

**References**

Agiakloglou, C. and P. Newbold (1994) Lagrange multiplier tests for fractional difference. Journal of Time Series Analysis 14, 253—262.

Ahtola, J.. and Tiao, G.C., 1987, Distributions of least squares estimators of autoregressive parameters for a process with complex roots on the unit circle, Journal of Time Series Analysis 8, 1-14.

Anh, V.V., V.P. Knopova and N.N. Leonenko, 2004, Continuous-time stochastic processes with cyclical long range dependence, Australian and New Zealand Journal of Statistics 46, 275-296.

Ashley, R., 1998, A new technique for postsample model selection and validation, Journal of Economics Dynamics and Control 22, 647-665.

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

Bartolini, L., Bertola, G. and A. Prati (2002), "Day-to-day monetary policy and the volatility of the Federal Funds interest rate", Journal of Money, Credit and Banking, 34, 1, 137-159.

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

Beran, J., Bhansali, R. J. and D. Ocker (1998). On unified model selection for stationary and nonstationary short- and long-memory autoregressive processes. *Biometrika*, 85, 921-934.

Bernanke, B.S. and A.S. Blinder (1988), "Credit, money and aggregate demand", American Economic Review Papers and Proceedings, 78, 435-439.

Bernanke, B.S. and A.S. Blinder (1992), "The Federal Funds rate and the channels of monetary transmission", American Economic Review, 82, 4, 901-921.

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.

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

Box, G.E.P. and D.A. Pierce, 1970, Distribution of residual autocorrelations in autoregressive integrated moving average time series models, Journal of the American Statistical Association 65, 1509-1526.

Caporale, G.M. and L.A. Gil-Alana, 2010, Fractional integration and data frequency, Journal of Statistical Computation and Simulation 30, 121-132.

Clarida, R.H., Gali, J. and M. Gertler (2000), "Monetary policy rules and macroeconomic stability: evidence and some theory", Quarterly Journal of Economics, 115, 147-180.

Clarida, R.H., Sarno, L., Taylor, M. and L. Valente (2006), "The role of asymmetries and regime shifts in the term structure of interest rates", Journal of Business, 79, 3, 1193-1224.

Clark, T.E. and M.W. McCracken, 2001, Tests of forecast accuracy and encompassing for nested models, Journal of Econometrics 105, 85-110.

Costantini, M. and R.M. Künst, 2011, On the usefulness of the Diebold and Mariano test in the selection of prediction models, Econometric Series 276, Institute for Advanced Studies.

Couchman, J., R. Gounder and J.J. Su, 2006, Long memory properties of real interest rates for 16 countries, Applied Financial Economics Letters 2, 25-30.

Cuestas, J.C. and L.A. Gil-Alana, 2015, A Non-linear Approach with Long Range Dependence based on Chebyshev Polynomials, Studies in Nonlinear Dynamics and Econometrics, forthcoming.

Dahlhaus, R. (1989), "Efficient Parameter Estimation for Self-similar Process", *Annals of Statistics*, 17, 1749-1766.

Dalla, V. and J. Hidalgo, (2005), A parametric bootstrap test for cycles, Journal of Econometrics, 129 (1-2). pp. 219-261.

Diebold, F.X. and A. Inoue, 2001, Long memory and regime switching. Journal of Econometrics, 105, 131-159.

Diebold, F.X. and R.S. Mariano, 1995, Comparing predictive accuracy, Journal of Business, Economics and Statistics 13, 253-263.

Diebold, F.X. and G.D. Rudebusch, 1989, Long memory and persistence in the aggregate output. Journal of Monetary Economics 24, 189-209.

Geweke, J., and S. Porter-Hudak (1983). The estimation and application of long memory time series models. *Journal of Time Series Analysis* 4(4): 221-238.

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

Gil-Alana, L.A., 2002, A mean shift break in the US interest rate, Economics Letter 77, 357-363.

Gil-Alana, L.A., 2004, Modelling the US interest rate in terms of I(d) statistical models, The Quarterly Review of Economics and Finance 44, 476-486.

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.

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

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

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.

Hamilton, J.D. (1996), "The daily market for Federal Funds", Journal of Political Economy, 5, 1135-1167.

Hamilton, J.D. and O. Jorda (2002), "A model for the Federal Funds rate target", Journal of Political Economy, 110, 1135-1167.

Harvey, D.I., S.J. Leybourne and P. Newbold, 1997, Testing the equality of prediction mean squared errors, International Journal of Forecasting 13, 281-291.

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

Hosking, J.R.M., 1981, Fractional differencing, Biometrika 68, 165-176.

Hosking, J.R.M., 1984, Modelling persistence in hydrological time series using fractional differencing, Water Resources Research 20, 1898-1908.

Lai, K.S., 1997, Long term persistence in the real interest rate: Some evidence of a fractional unit root, International Journal of Finance and Economics 2, 225-235.

Ljung, G.M. and G.E.P. Box, 1978, On a measure of lack of fit in time series models, Biometrika 65, 297-303.

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

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

Makridakis, S. and M. Hibon, 2000, The M-3 competition: results, conclusions and implications, International Journal of Forecasting, 16, 451-476.

Makridakis, S., S. Wheelwright and R. Hyndman, 1998, Forecasting methods and applications, 3rd Edition, John Wiley & Sons.

Meade, N. and M.R. Maier, 2003, Evidence of long memory in short term interest rates, Journal of Forecasting 22, 553-568.

Ohanissian, A, J.R. Russel and R.S. Tsay, 2008, True or spurious long memory. A new test, Journal of Business and Economic Statistics 26, 2, 161-175.

Pivetta, F and R. Reis, 2007, The persistence of inflation in the United States, Journal of Economics, Dynamics and Control 31, 4, 1326-1358.

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

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

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

Robinson, P. M. (1995b), 'Log-periodogram regression of time series with long range dependence', Annals of Statistics 23, 1048_1072.

Sarno, L. and D.L. Thornton (2003), "The dynamic relationship between the Federal Funds rate and the Treasury bill rate: an empirical investigation", Journal of Banking and Finance, 27, 1079-1110.

Sarno, L., Thornton, D.L. and G. Valente (2005), "Federal Funds rate prediction", Journal of Money, Credit and Banking, 37, 3, 449-471.

Soares, L.J. and L.R. Souza, 2006, Forecasting electricity demand using generalized long memory, International Journal of Forecasting 22, 17-28.

Sowell, F.B., 1992a. Maximum Likelihood Estimation of Stationary Univariate Fractionally Integrated Time Series Models. Journal of Econometrics 53, 165–188.

Sowell, F., 1992b, Modelling long run behaviour with the fractional ARIMA model, Journal of Monetary Economics 29, 2, 277-302.

Taylor, J.B. (1993), "Discretion versus policy rules in practice", Carnegie Rochester Conference Series on Public Policy, 39, 195-214.

Taylor, J.B. (2001), "Expectations, open market operations, and changes in the Federal Funds rate", Federal Reserve Bank of St. Louis Review, 83, 33-47.

Tsay, W.J., 2000, The long memory story of the real interest rate, Economics Letters 67, 325-330.
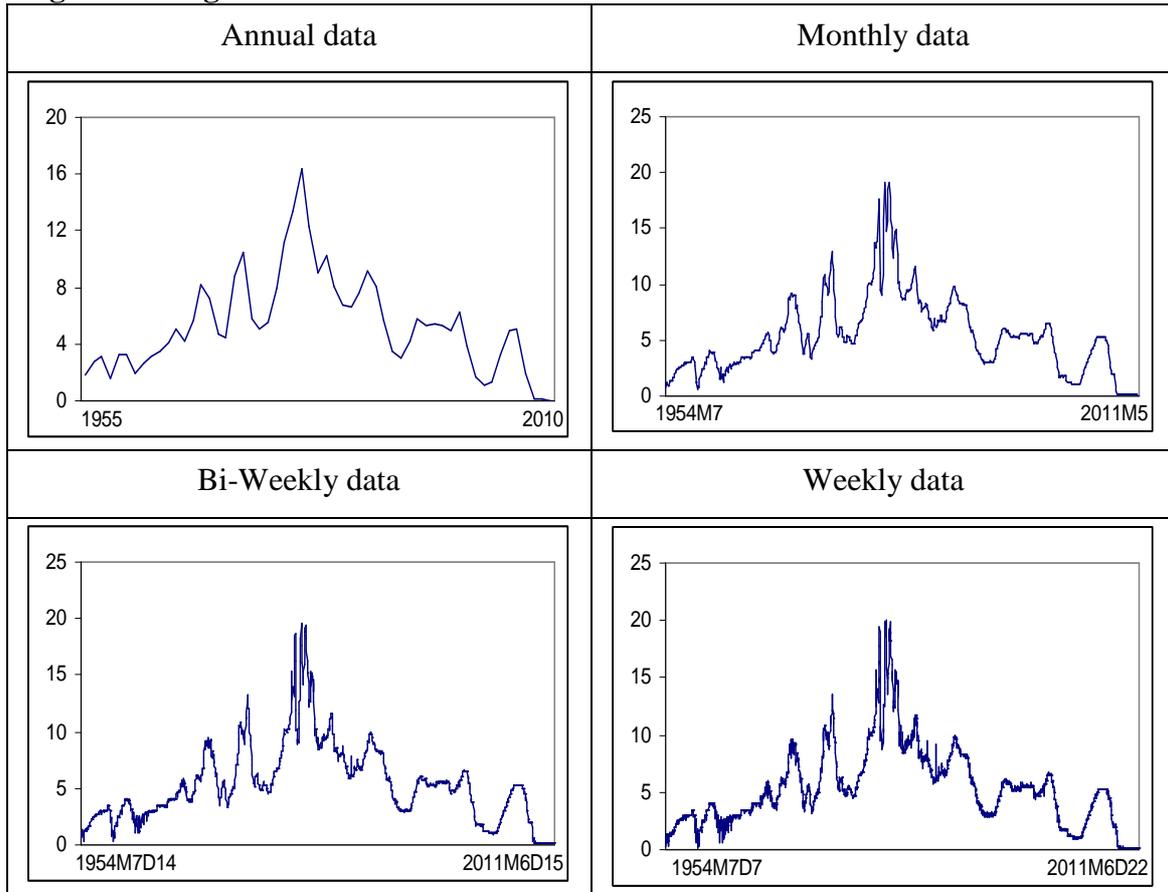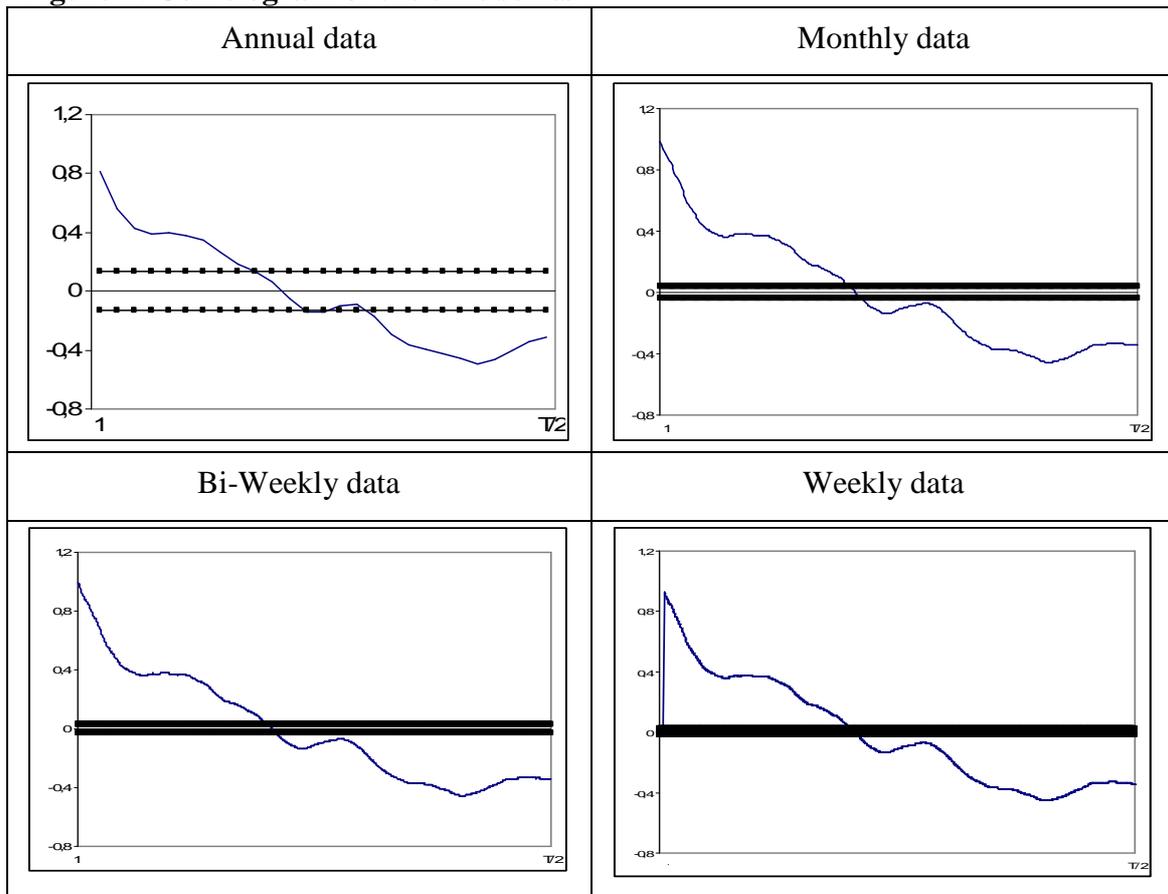
**Figure 1: Original time series data**

| Annual data | Monthly data |
|---|---|
|  |  |
| Bi-Weekly data | Weekly data |
|  |  |

**Figure 2: Correlogram of the time series**

| Annual data | Monthly data |
|:---:|:---:|
|  |  |
| **Bi-Weekly data** | **Weekly data** |
|  |  |

Note: The thick lines refer to the 95% confidence band for the null hypothesis of no autocorrelation.

**Figure 3: Periodogram of the time series**

| Annual data | Monthly data |
|---|---|



| Bi-Weekly data | Weekly data |
|---|---|
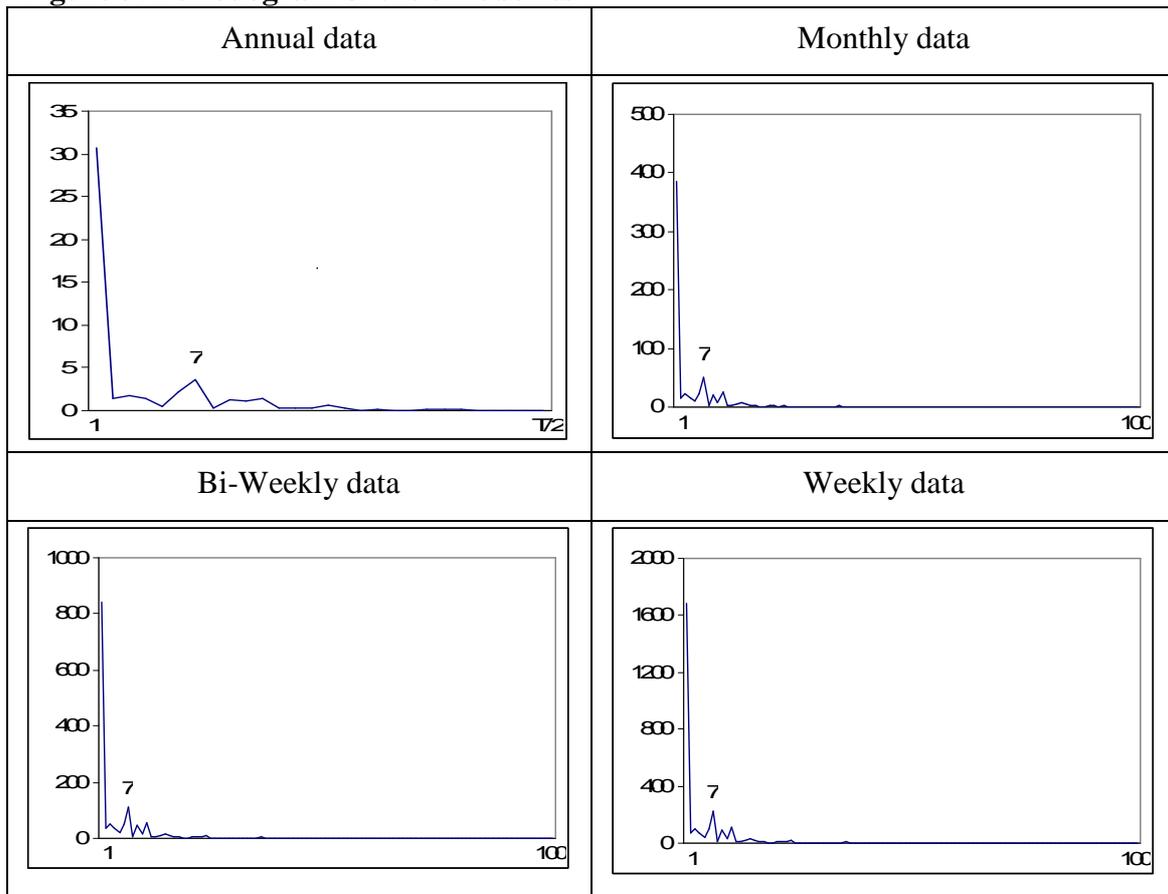


Note: The horizontal axis refers to the discrete Fourier frequencies $\lambda_j = 2\pi j/T$, $j = 1, \ldots, T/2$.

28

**Table 1: Estimates of d and 95% confidence interval in an I(d) model with an intercept**

|  | White noise | AR(1) disturbances | AR(2) disturbances |
|---|---|---|---|
| Annual | 0.937 (0.704, 1.450) | 0.544 (0.429, 0.700) | 0.722 (0.334, 1.495) |
| Monthly | 1.277 (1.189, 1.383) | 0.821 (0.742, 0.913) | 0.852 (0.679, 1.016) |
| Bi-Weekly | 1.168 (1.122, 1.213) | 1.025 (0.891, 1.146) | 0.824 (0.633, 1.008) |
| Weekly | 0.973 (0.954, 0.994) | 1.086 (1.044, 1.127) | 1.045 (0.984, 1.101) |

The values are Whittle estimates of d in the frequency domain (Dahlhaus, 1989). Those in parentheses are the 95% confidence interval of the non-rejection values of d using Robinson (1994).

**Table 2: Estimates of $d_1$ and $d_2$ in the model with two fractional structures**

| Frequency | j ( = T / r ) | $d_1$ | $d_2$ |
|---|---|---|---|
| Annual | j = 7  (r = 8) | 0.932 (0.561, 0.983) | 0.094 (-0.008, 0.233) |
| Monthly | j = 683  (r = 97) | 0.966 (0.895, 1.128) | 0.145 (0.109, 0.217) |
| Bi-Weekly | j = 1486  (r = 212) | 0.802 (0.661, 0.977) | 0.234 (0.158, 0.299) |
| Weekly | j = 2973  (r = 424) | 0.817 (0.722, 0.903) | 0.156 (0.114, 0.198) |

The values in parentheses in the third and fourth columns are the 95% confidence interval of the non-rejection values of d using Robinson (1994).

**Table 3: Modified DM statistic: 5, 10, 15, 20 and 25-step ahead forecasts**

| (M1) vs (M2) | 5 | 10 | 15 | 20 | 25 |
|---|---|---|---|---|---|
| Annual | 1.435 | 1.764 | 1.114 | -1.698 | **-4.311 (M2)** |
| Monthly | 1.872 | 1.554 | -1.050 | **-3.564 (M2)** | **-12.344 (M2)** |
| Bi-weekly | 1.115 | 1.355 | **-3.211 (M2)** | **-5.667 (M2)** | **-10.093 (M2)** |
| Weekly | 0.998 | -0.065 | -1.445 | **-4.443 (M2)** | **-8.005 (M2)** |

In bold the cases where model (M2) outperforms model (M1) in statistical terms.