Department of Economics and Finance

Economics and Finance Working Paper Series

Working Paper No. 2401

Guglielmo Maria Caporale and Luis Alberiko Gil-Alana

A Long-Memory Model for Multiple Cycles with an Application to the S&P500

February 2024

http://www.brunel.ac.uk/economics

A LONG-MEMORY MODEL FOR MULTIPLE CYCLES WITH AN APPLICATION TO THE S&P500

Guglielmo Maria Caporale, Brunel University London, UK

Luis Alberiko Gil-Alana, University of Navarra, Pamplona, Spain, and Universidad Francisco de Vitoria, Madrid, Spain

February 2024

Abstract

This paper proposes a long-memory model including multiple cycles in addition to the long-run component. Specifically, instead of a single pole or singularity in the spectrum, it allows for multiple poles and thus different cycles with different degrees of persistence. It also incorporates non-linear deterministic structures in the form of Chebyshev polynomials in time. Simulations are carried out to analyse the finite sample properties of the proposed test, which is shown to perform well in the case of a relatively large sample with at least 1000 observations. The model is then applied to weekly data on the S&P500 from 1 January 1970 to 26 October 2023 as an illustration. The estimation results based on the first differenced logged values (i.e., the returns) point to the existence of three cyclical structures in the series with a length of approximately one month, one year and four years respectively, and to orders of integration in the range (0, 0.20), which implies stationary long memory in all cases.

Keywords: Fractional integration; multiple cycles; stock market indices; S&P500

JEL Classification: C22; C15

Corresponding author: Professor Guglielmo Maria Caporale, Department of Economics and Finance, Brunel University London, Uxbridge, UB8 3PH, UK. Email: Guglielmo-Maria.Caporale@brunel.ac.uk; https://orcid.org/0000-0002-0144-4135

Prof. Luis A. Gil-Alana gratefully acknowledges financial support from the MINEIC-AEI-FEDER ECO2017-85503-R project from ‘Ministerio de Economía, Industria y Competitividad’ (MINEIC), `Agencia Estatal de Investigación' (AEI) Spain and `Fondo Europeo de Desarrollo Regional' (FEDER), and also from internal projects of the Universidad Francisco de Vitoria.

1. Introduction

It is common in economics and finance, as well as in other disciplines, to decompose a time series into a trend, a seasonal and a cyclical component. For this purpose, various statistical tools have been developed over the years, and, most recently, machine learning and other big data techniques have also been used. This paper proposes a new time series model incorporating all these components which is a special case of the very general and flexible testing framework developed in Robinson (1994). More specifically, the deterministic part of the model includes a constant, a linear time trend, and a cyclical structure or seasonal dummy variables; it also allows for non-linearities in the form of neural networks (Yaya et al., 2021) or Chebychev polynomials in time as in Cuestas and Gil-Alana (2016). In the stochastic part the spectral density function is allowed to have multiple poles or singularities, and thus multiple integer or fractional roots of arbitrary order anywhere in the unit circle in the complex plane. These are related to the previously mentioned components. In particular, the trend component is associated with the long-run or zero frequency, and the others (i.e., seasonal and cyclical) to other non-zero frequencies. Simulations are carried out to analyse the finite sample properties of this model, which is then applied to weekly data on the S&P500 from 1 January 1970 to 26 October 2023 as an illustration.

The remainder of the paper is structured as follows: Section 2 describes the model; Section 3 presents the test statistic; Section 4 reports on Monte Carlo evidence concerning the finite sample performance of the proposed test; Section 5 provides information about the data and discusses the empirical application; Section 6 offers some concluding remarks.

2. The Econometric Model

Let y(t) be the observed time series from t = 1, 2, …, T. We consider the following model,

$$y(t) \;=\; f(z(t);\,\psi) \;+\; x(t), \qquad t \;=\; 1, 2, \ldots, \tag{1}$$

where f can be a linear or a non-linear function of z(t), which is a (kx1) vector of observable deterministic variables, and γ is a (kx1) vector of unknown parameters to be estimated. Thus, if f is linear, it can include, for instance, an intercept and a linear time trend of the form advocated by Bhargava (1986), Schmidt and Phillips (1992) and others in the context of unit roots, i.e.,

$$f(z(t);\,\psi) \;=\; \alpha \;+\; b\,t \tag{2}$$

and, if it is non-linear, it can include, for example, Chebyshev polynomials in time of the form:

$$f(z(t);\,\psi) \;=\; \textstyle\sum_{i=0}^{m} \theta_i P_{iT}\,(t), \tag{3}$$

where m indicates the number of coefficients of the Chebyshev polynomial in time $P_{i,T}(t)$ defined as:

$$P_{0,T}(t) = 1, \text{ and } P_{i,T}(t) \;=\; \sqrt{2}\,cos(i\pi(t-0.5)/T),$$

and described in Hamming (1973) and Smyth (1998). Bierens and Martins (2010) propose the use of such polynomials in the case of time-varying cointegrating parameters. There are several advantages to using them. First, their orthogonality avoids the problem of near collinearity in the regressors matrix which arises with standard time polynomials. Second, according to Bierens (1997) and Tomasevic et al. (2009), they can approximate highly non-linear trends with rather low degree polynomials. Finally, they can approximate structural breaks in a much smoother way than the classical structural change models.

As for the stochastic part of the model, x(t) is specified as:

$$\prod_{j=1}^{m}\left(1-2\cos w_j^r\, L+L^2\right)^{d_j} x(t) \;=\; u(t), \quad t \;=\; 1,2,\ldots, \qquad (4)$$

where w_j^r = 2πr/T, r = T/j is a real scalar value, L is the lag operator, i.e., Lx(t) = x(t-1), d_j is another real value corresponding to the order of integration of the cycle that explodes (i.e., it goes to infinity) in the spectrum at λ = j; m stands for the number of cyclical structures, and u(t) is a short-memory process integrated of order 0 or I(0). Such a process is defined as a covariance stationary one with a spectral density function that is positive and finite at all frequency in the spectrum. Thus, it could be a white noise process but also display weak autocorrelation as in a stationary and invertible AutoRegressive Moving Average (ARMA) model. In the present study, in order to avoid overparameterisation, we follow the exponential spectral approach of Bloomfield (1973) to model u(t). This is a non-parametric framework that is implicitly defined in terms of its spectral density function:

$$f(\lambda;\,\tau) \;=\; \left[\frac{\sigma 2}{2\pi}\right]\exp\left[\,2\sum_{i=0}^{n}\tau_i\cos(\lambda\, i)\,\right]\,, \qquad (5)$$

where σ^2 is the variance of the error term and n denotes the number of short-run dynamic terms, and whose logged form approximates fairly well autoregressive processes. Bloomfield (1973) showed that for a stationary and invertible ARMA (p, q) process of the form:

$$u(t) \;=\; \sum_{r=1}^{p}\varphi_r\, u(t-r) \;+\; \varepsilon_t \;+\; \sum_{s=1}^{q}\theta_s\varepsilon(t-s)\,,$$

where ε_t is a white noise process, the spectral density function is given by:

$$f(\lambda;\tau) \;=\; \frac{\sigma^2}{2\pi}\left|\frac{1\;+\;\sum_{s=1}^{q}\theta_s e^{i\lambda s}}{1\;-\;\sum_{r=1}^{p}\varphi_r e^{i\lambda r}}\right|^2.$$

According to Bloomfield (1973), the log of the above expression can be well approximated by Equation (5) when p and q are small values, and thus it does not require the estimation of as many parameters as in the case of ARMA models. In addition, Bloomfield's (1973) model is stationary across all its values (see Gil-Alana, 2004).

In the empirical application carried out below we assume that p = 4, and that w(r;1) = 0, so the first cyclical component corresponds to the long run or zero frequency. In such a case, the summand $(1 - 2\ \cos w(r;j)\, L + L^2)^{d_j}$ becomes $(1 - 2\ L + L^2)^{d_1}$, which can be expressed as $(1 - L)^{2d_1}$, with the pole or singularity in the spectrum going to infinity at the zero frequency (Granger, 1980, Granger and Joyeux, 1980 and Hosking, 1981). For the other two cyclical structures, we choose the frequencies on the basis of the values of the periodogram, which is an estimator of the spectral density function.

3. The Test Statistic

The test statistic can be easily derived from Robinson (1994) extending the function f in Equation (1) to the non-linear case, and specifying the errors to follow Bloomfield's (1973) model.

Specifically, we test the null hypothesis:

$$H_o\!: d = d_o, \qquad (6)$$

where d_o is a vector of real numbers and dimension m, each element corresponding to the order of integration at a given frequency. Given this null hypothesis, the residuals in (1) and (4) are

$$r(t) = \prod_{j=1}^{m} \left(1 - 2\cos w_j^r\, L + L^2\right)^{d_{jo}} xx(t)$$

where xx(t) are the residuals of the linear or non-linear model in (1), and the periodogram of r(t) is computed as

$$P(\lambda j) = \ \left| \frac{1}{(2\,\pi\,T)^{\frac{1}{2}}} \sum_{t=1}^{T} r(t) e^{i\,\lambda j\, t} \right|^2 .$$

The test statistic takes the form:

$$\text{NLROB} = \frac{T}{\hat{\sigma}^4} \hat{a}' \hat{A}^{-1} \hat{a}, \qquad (7)$$

where *T* is the sample size, and

$$\hat{a} = \frac{-2\pi}{T} \sum_f^* \psi(\lambda_j) g_u(\lambda_j; tau)^{-1} P(\lambda_j),$$

$$\hat{\sigma}^2 = \sigma^2(tau) = \frac{2\pi}{T} \sum_{f=1}^{T-1} g_u(\lambda_j\ tau)^{-1} P(\lambda_j),$$

$$\hat{A} = \frac{2}{T} \left\{ \sum_f^* \psi(\lambda_j)\psi(\lambda_j)' - \sum_f^* \psi(\lambda_j)\hat{\xi}(\lambda_j)' \left[\sum_f^* \hat{\xi}(\lambda_j)\hat{\xi}(\lambda_j)' \right]^{-1} \sum_f^* \hat{\xi}(\lambda_j)\psi(\lambda_j)' \right\};$$

$$\psi(\lambda_f) = log \left| 2\, sin \frac{\lambda_f}{2} \right|; \quad \hat{\xi}(\lambda_j) = \frac{\partial}{\partial tau} \log\ g_{xx}(\lambda_j; tau),$$

where $\lambda_j = 2\pi j/T$, and * indicates that the sums are taken over all frequencies bounded in the spectrum, namely removing those that explode or go to infinity. Also, tau is definted as the $arg \quad min_{\tau \in T^*} \quad \sigma^2(\tau)$, where T^* is a subset of the R^q Euclidean space.

It can be easily proven that extending the conditions in Robinson (1994) to the non-linear structure in (1), which is satisfied by Condition (*) in his paper,

$$NLROB \underset{d}{\rightarrow} \chi^2{}_m \quad \text{as } T \rightarrow \infty, \qquad (8)$$

where T indicates the sample size, and "→d" stands for convergence in distribution. Thus, unlike in the case of other procedures, the present one is a classical large-sample testing situation. Moreover, this test is the most efficient in the Pitman sense against local departures from the null, i.e., if it is implemented against local departures, the limit distribution is $\chi^2{}_m$ (v) with a non-centrality parameter v which is optimal under Gaussianity of the error term. The latter is not necessary for the implementation of this procedure, a moment condition of only order 2 being required.

4. Finite Sample Properties

This section reports on the finite sample performance of the test described above. Specifically, we carry out Monte Carlo simulations to analyse the size and the power of this test against various alternatives. A similar experiment was conducted in Gil-Alana (2001), though in that case m = 1 and thus only a single cyclical structure was allowed. By contrast, in the present study we assume that m = 3 and that the Data Generating Process (DGP) is characterised by the following orders of integration: d1= 0.75, d2 = 0.50 and d3 = 0.25; this implies that the first cyclical structure is highly persistent and non-stationary, the second one is on the borderline between the stationary and nonstationary case, and the third one is stationary. For the length of the cycles we impose j = 10, 100 and 250, with different sample sizes. These values are arbitrary, though the results were found to be robust to choosing other values. For the alternative hypotheses we consider values for the three orders of integration of 0.25, 0.50 and 0.75. For j we choose the same values as in the true model, therefore the size of the test is reported in the tables for d = $(0.75, 0.50, 0.25)^T$.

INSERT TABLES 1, 2 AND 3 ABOUT HERE

Table 1 reports the results based on T = 1,000. It can be seen that the size of the test is too large, the rejection frequency being 0.119 for a nominal size of 0.050 and very large (above 0.700) in all cases. Table 2 shows that with a bigger sample size, i.e. T=2,000, the power becomes closer to the nominal size, 0.094, and the rejection probabilities are now higher than 0.800 in all cases, being equal to 1 in 7 out of the 27 cases. Finally, Table 3 displays the results for T=3,000; in this case the nominal size is 0.066, and there are 12 cases when the rejection probabilities are equal to 1, these being higher than 0.900 in all cases.

5. An Empirical Application

We use weekly data on the S&500 closing prices from 1 January 1970 to 26 October 2023. The source is Yahoo finance. Figure 1 displays plots of the original data and their log transformation in the upper panel, and their corresponding periodograms in the lower one. Both exhibit a very large value at the smallest (zero) frequency, which might be an indication of the need for (fractional or non-fractional) differentiation of the data.

INSERT FIGURES 1 AND 2 ABOUT HERE

Figure 2 displays the plots of the first differenced data. Much higher volatility is observed in the second half of the sample when using the original data, but not in the case of the logged ones. Therefore, we use the latter series for the empirical application below.

Initially we set m = 5, and obtained $j_1 = 1$, as one would have expected in view of the periodograms displayed in Figure 1. The estimated value of d_1 is then 1, which is consistent with the results of standard unit root tests (Dickey and Fuller, 1979; Phillips and Perron, 1988; Kwitakowski et al., 1992; Elliot et al., 1996; Ng and Perron, 2001; etc.). We also allowed for fractional integration by estimating the equation (1-L)x(t) = u(t) instead of imposing (4). The results support the unit root null hypothesis with values of d equal to 0.981 and 0.992 for the original and logged transformed data respectively, the corresponding 5% confidence intervals being (0.972, 1.009) and (0.983, 1.017). Next, we focus on the first differenced log-values, in this case setting m = 4, and finding an order of integration not significantly different from zero; finally we set m = 3.

Note that standard unit roots are a special case of (2) with m = 1 and w = 0 such that the model becomes:

$$(1 - 2L + L^2)^{d_j} x(t) = u(t), \quad t = 1, 2, \ldots, \qquad (9)$$

and, denoting $2d_j = d$, corresponds to the standard I(d) model at the long-run or zero frequency and a unit root if d =1:

$$(1 - L)^{d} x(t) = u(t), \quad t = 1, 2, \ldots \; . \qquad (10)$$

To allow for some degree of generality we assume that x(t) are the errors in a regression model with an intercept and a time trend, i.e.,

$$y(t) = \alpha + \beta t + x(t), \qquad (11)$$

and make two alternative assumptions for u(t), namely that it is a white noise process or that it follows the exponential spectral model of Bloomfield (1973) in turn.

We start with the linear case. Table 4 displays the estimates of d along with their associated 95% confidence bands, under the three standard specifications with: i) no deterministic terms ($\alpha = \beta = 0$ in (8)); ii) an intercept only ($\beta = 0$) and iii) a constant and a linear time trend. Our preferred model is selected on the basis of the statistical significance of the estimated coefficients (shown in bold in the table). We report the results for both the original and log-transformed data. The coefficient on the linear trend is significant in three out of the four cases (the exception is represented by the original data with Bloomfield disturbances), and, although d is maller than 1 in all four cases, the unit root null hypothesis cannot be rejected for any of them.

INSERT TABLES 4 AND 5 ABOUT HERE

Table 5 concerns the nonlinear model. One can see that the nonlinear terms are significant in some cases (especially with white noise errors), and again the estimates of d are within the unit root interval, which implies that first differencing is required for both the linear and the nonlinear terms.

INSERT TABLES 6 AND 7 ABOUT HERE

As mentioned before, we estimate the model given by equations (1) and (4) for the return series, and first set m = 4. In this case, the order of integration of one of the cyclical structures is found to be equal to zero. Next, we set m = 3. The results are very similar for the linear (Table 6) and non-linear (Table 7) models, in both cases the orders of integration of all three components being significant.

INSERT TABLES 8 AND 9 ABOUT HERE

6. Conclusions

This paper proposes a long-range dependence framework allowing for multiple cycles which is then applied to analyse the behaviour of the weekly S&P500. Specifically, instead of a single pole or singularity in the spectrum as in standard models, our model allows for multiple poles and thus different cycles with different degrees of persistence. It also incorporates non-linear deterministic structures in the form of Chebyshev polynomails in time. Montecarlo simulations show that in finite samples the proposed test behaves well if the sample size is relatively large, namely if the number of observations is at least 1,000.

The empirical application using weekly data on the S&P500 provides evidence of a large value in the periodogram at the zero frequency. Unit and fractional root tests also suggest the need to take first differences. The estimation results based on the first differenced logged values (i.e., the returns) point to the existence of three cyclical structures in the series with a length of approximately one month, one year and four years respectively, and to orders of integration in the range (0, 0.20), which implies stationary long memory in all cases.

Future research could develop the framework presented in this paper in several ways. For instance, the number of structures *m* could be endogenised; estimation methods such as those proposed by Giraitis and Leipus (1995), Woodward et al. (1998), Ferrara and Guegan (2001), and Sadek and Khotanzad (2004) could be extended to allow for non-linear trends, and breaks in the data could also be modelled. Moreover, these methods could also be used to examine the stochastic behaviour of a wide range of macro variables such as GDP, inflation or unemployment.

References

Bhargava, A. 1986. On the theory of testing for unit roots in observed time series. Review of Economic Studies 52, 369–84

Bierens, H. J. 1997. "Testing the Unit Root with Drift Hypothesis Against Trend Stationarity with an Application to the US Price Level and Interest Rate." *Journal of Econometrics* 81: 29–64.10.1016/S0304-4076(97)00033-X

Bierens, H. J., and L. F. Martins. 2010. "Time Varying Cointegration." *Econometric Theory* 26: 1453–1490.10.1017/S0266466609990648

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

Cuestas, J.C. and L.A. Gil-Alana (2016), A non-linear approach with long range dependence based on Chebyshev polynomials, Studies in Nonlinear Dyanmics and Econometrics 20, 1, 57-70.

Dickey, D. A and Fuller, W. A. (1979). Distributions of the estimators for autoregressive time series with a unit root, Journal of American Statistical Association, 74 (366), 427-481.

Elliot, G., T.J. Rothenberg and J.H. Stock (1996). Efficient tests for an autoregressive unit root, Econometrica 64, 813-836.

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

Gil-Alana, L.A. 2004. "The use of the Bloomfield model as an approximation to ARMA processes in the context of fractional integration." *Mathematical and Computer Modelling*. 39: 429-436. doi: 10.1016/S0895-7177(04)90515-8ç

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

Granger, C.W.J., (1980), Long memory relationships and the aggregation of dynamic models, JOURNAL OF ECONOMETRICS, 14, 227-238.

Granger, C.W.J., and R. Joyeux, (1980), An introduction to long memory time series models and fractional differencing, JOURNAL OF TIME SERIES ANALYSIS 1, 15-39.

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

Hamming, R. W. (1973). Numerical Methods for Scientists and Engineers. Dove

Kwiatkowski, D., Phillips, P. C., Schmidt, P. and Shin, Y. (1992). Testing the null hypothesis of stationarity against the alternative of a unit root. Journal of Econometrics, 54(1-3), 159-178.

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.

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

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

Phillips (1992), LM Tests for a Unit Root in the Presence of Deterministic Trends, Oxford Bulletin of Economics and Statistics 54, 257-287.

Smyth, G.K. (1998). Polynomial Approximation. John Wiley & Sons, Ltd, Chichester.

Tomasevic, N., M. Tomasevic, and T. Stanivuk. 2009. "Regression Analysis and Approximation by Means of Chebishev Polynomials. *Informatologia* 42: 166–172.

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

Yaya, O., A. Oghonna, F. Furuoka and L.A. Gil-Alana (2021), A new unit root analysis for testinghysteresis in unemployment, Oxford Bulletin of Economics and Statistics 83, 4, 960-981.

Table 1: Rejection frequencies of the test for a sample size T = 1,000

d_1	d_2	d_3	Rejection freq.
0.25	0.25	0.25	0.889
0.25	0.25	0.50	0.947
0.25	0.25	0.75	1.000
0.25	0.50	0.25	0.799
0.25	0.50	0.50	0.845
0.25	0.50	0.75	0.945
0.25	0.75	0.25	0.904
0.25	0.75	0.50	0.967
0.25	0.75	0.75	1.000
0.50	0.25	0.25	0.866
0.50	0.25	0.50	0.923
0.50	0.25	0.75	0.988
0.50	0.50	0.25	0.777
0.50	0.50	0.50	0.814
0.50	0.50	0.75	0.908
0.50	0.75	0.25	0.890
0.50	0.75	0.50	0.923
0.50	0.75	0.75	0.998
0.75	0.25	0.25	0.815
0.75	0.25	0.50	0.901
0.75	0.25	0.75	0.945
0.75	0.50	0.25	**0.119**
0.75	0.50	0.50	0.807
0.75	0.50	0.75	0.833
0.75	0.75	0.25	0.812
0.75	0.75	0.50	0.865
0.75	0.75	0.75	0.939

In bold, the value corresponding to the size of the test. Nominal size: 5%

Table 2: Rejection frequencies of the test for a sample size T = 2,000

d_1	d_2	d_3	Rejection freq.
0.25	0.25	0.25	0.911
0.25	0.25	0.50	1.000
0.25	0.25	0.75	1.000
0.25	0.50	0.25	0.801
0.25	0.50	0.50	0.906
0.25	0.50	0.75	1.000
0.25	0.75	0.25	0.998
0.25	0.75	0.50	1.000
0.25	0.75	0.75	1.000
0.50	0.25	0.25	0.899
0.50	0.25	0.50	0.978
0.50	0.25	0.75	1.000
0.50	0.50	0.25	0.839
0.50	0.50	0.50	0.848
0.50	0.50	0.75	0.922
0.50	0.75	0.25	0.955
0.50	0.75	0.50	0.988
0.50	0.75	0.75	1.000
0.75	0.25	0.25	0.890
0.75	0.25	0.50	0.977
0.75	0.25	0.75	0.994
0.75	0.50	0.25	**0.094**
0.75	0.50	0.50	0.883
0.75	0.50	0.75	0.847
0.75	0.75	0.25	0.890
0.75	0.75	0.50	0.914
0.75	0.75	0.75	0.978

In bold, the value corresponding to the size of the test. Nominal size: 5%

Table 3: Rejection frequencies of the test for a sample size T = 3,000

d_1	d_2	d_3	Rejection freq.
0.25	0.25	0.25	0.989
0.25	0.25	0.50	1.000
0.25	0.25	0.75	1.000
0.25	0.50	0.25	0.991
0.25	0.50	0.50	0.911
0.25	0.50	0.75	1.000
0.25	0.75	0.25	1.000
0.25	0.75	0.50	1.000
0.25	0.75	0.75	1.000
0.50	0.25	0.25	0.939
0.50	0.25	0.50	1.000
0.50	0.25	0.75	1.000
0.50	0.50	0.25	0.965
0.50	0.50	0.50	0.934
0.50	0.50	0.75	0.980
0.50	0.75	0.25	0.999
0.50	0.75	0.50	1.000
0.50	0.75	0.75	1.000
0.75	0.25	0.25	0.956
0.75	0.25	0.50	1.000
0.75	0.25	0.75	0.999
0.75	0.50	0.25	**0.066**
0.75	0.50	0.50	0.909
0.75	0.50	0.75	0.917
0.75	0.75	0.25	0.943
0.75	0.75	0.50	0.993
0.75	0.75	0.75	1.000

In bold, the value corresponding to the size of the test. Nominal size: 5%

Table 4: Estimates of d in the linear model given by Equation (1)

i) Results based on white noise errors			
	No terms	With an intercept	With a time trend
Original	0.97 (0.94, 1.00)	0.97 (0.94, 1.00)	**0.97 (0.94, 1.00)**
Logged values	0.99 (0.96, 1.02)	0.98 (0.96, 1.01)	**0.98 (0.96, 1.01)**
ii) Results based on autocorrelated (Bloomfield) errors			
	No terms	With an intercept	With a time trend
Original	0.95 (0.92, 0.99)	**0.95 (0.92, 1.00)**	0.95 (0.92, 1.00)
Logged values	0.97 (0.93, 1.02)	0.99 (0.95, 1.04)	**0.99 (0.95, 1.04)**

The values in parenthesis are the 95% confidence intervals. Those in bold correspond to the models selected on the basis of the statistical significance of the deterministic terms.

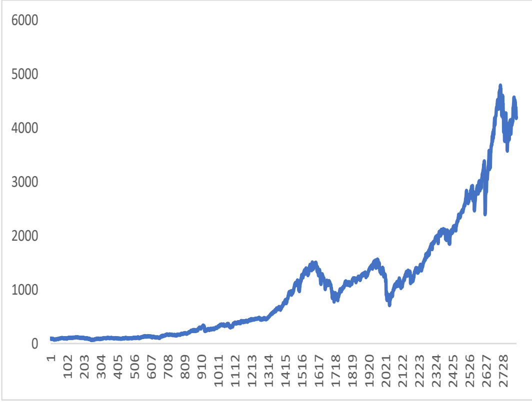
Table 5: Estimates of d in the non-linear model given by Equation (2)

i) Results based on white noise errors					
	d	θ_1	θ_2	θ_3	θ_4
Original	0.97 (0.93, 1.01)	**1811.38 (1.92)**	**-1381.76 (-2.44)**	**484.63 (1.67)**	**-335.25 (-1.72)**
Logged	1.01 (0.98, 1.04)	**5.400 (6.81)**	-0.058 (-0.12)	**-0.575 (-2.39)**	**1.167 (7.30)**
ii) Results based on autocorrelated (Bloomfield) errors					
	d	θ_1	θ_2	θ_3	θ_4
Original	0.96 (0.94, 1.02)	1043.07 (1.59)	**-702.52 (-1.93)**	307.28 (1.48)	-185.24 (-1.32)
Logged	1.00 (0.97, 1.04)	**0.915 (6.81)**	**-3.429 (-4.32)**	**1.096 (2.77)**	0.047 (0.17)

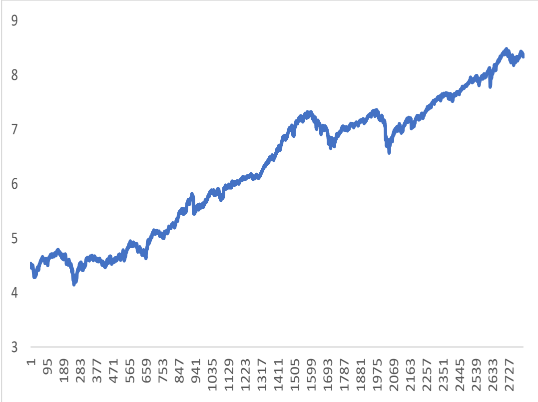
The values in parenthesis in column 2 are the 95% confidence intervals. Those in bold correspond to the significant Chebyshev polynomials.

Figure 1: Data in levels and periodograms

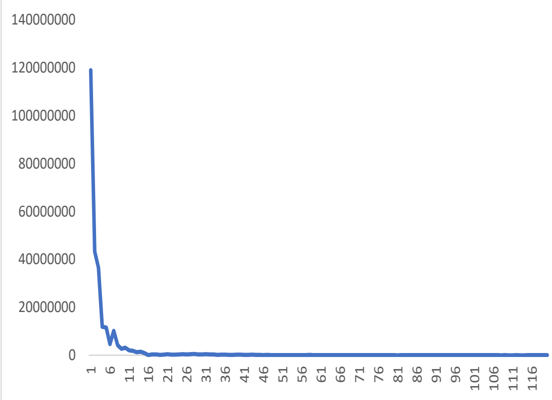
Original data



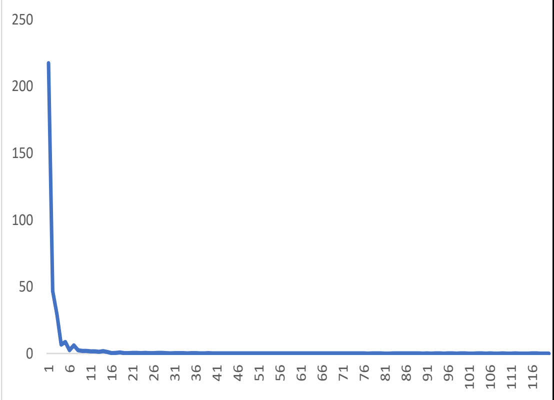
Logged transformed data



Periodogram original data

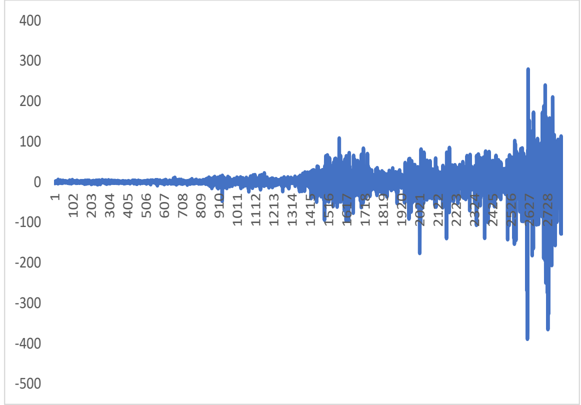
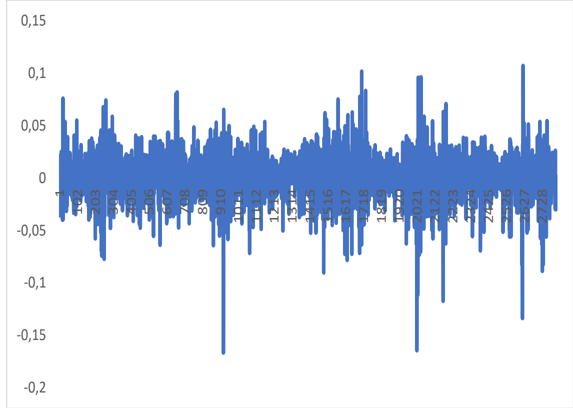
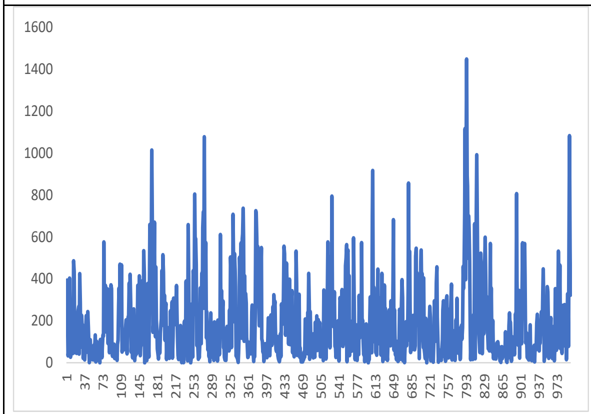
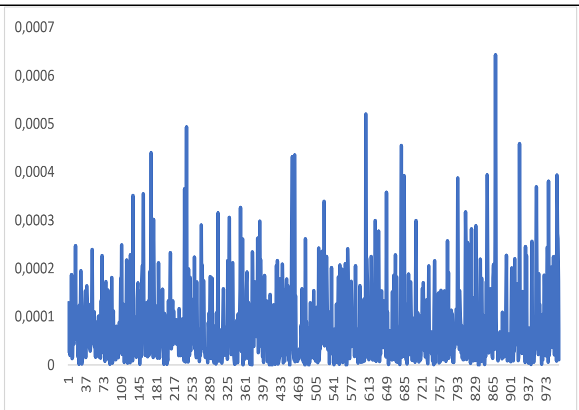
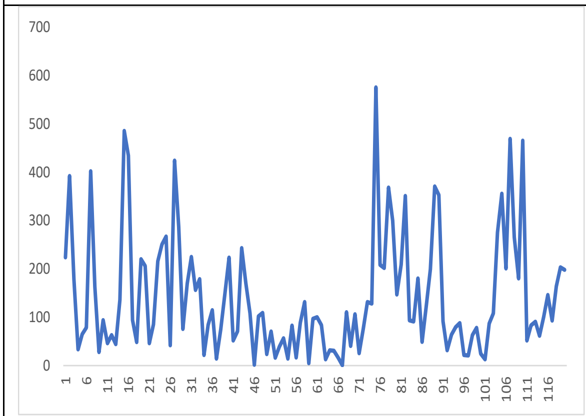
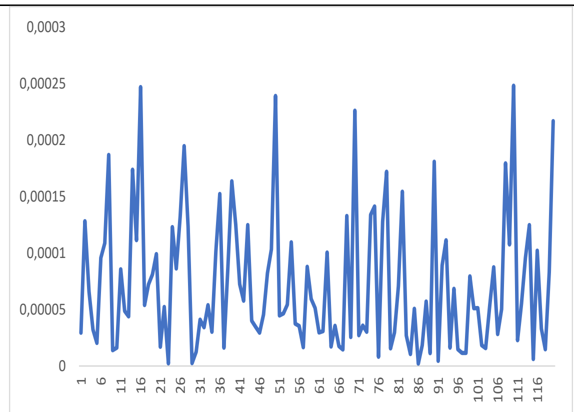


Periodogram logged transformed data



The values on the horizontal axis correspond to the discrete Fourier frequencies, $\lambda j = 2\pi j/T$, $j = 1, 2, \ldots T/2$.

Figure 2: Data in first differences and periodograms

Original data	Logged transformed data
	
Periodogram original data (j = 1, …, 1000)	Periodogram logged data (j = 1, …, 1000)
	
Periodogram original data, (j = 1, …120)	Periodogram logged data (j = 1, …, 120)
	

The values on the horizontal axis correspond to the discrete Fourier frequencies, λj = 2πj/T, j = 1, 2, … T/2.

Table 6: Frequencies with the highest values at periodograms, with j = 1, …, 1000

(1 – L) Data			(1 – L) Log data		
j	T / j	Value at Periodogram	J	T / j	Value at Periodogram
794	3.53	1448.09	871	3.22	0.000642
998	2.81	1082.75	607	4.62	0.000520
274	10.24	1076.58	242	11.60	0.000493
170	16.51	1013.94	920	3.05	0.000458
814	3.45	990.76	679	4.13	0.000454
608	4.61	916.17	170	16.51	0.000639

J indicates the discrete frequency in the periodogram; T/j indicates the number of periods per cycle, while the third and sixth columns indicate the corresponding value at the periodogram.

Table 7: Frequencies with the highest values at periodograms, with j = 1, …, 120

(1 – L) Data			(1 – L) Log data		
j	T / j	Value at Periodogram	J	T / j	Value at Periodogram
75	37.44	575.65	110	25.52	0.000248
15	187.20	485.48	16	175.50	0.000247
107	26.25	469.23	50	56.16	0.000239
110	25.52	465.70	70	40.11	0.000226

J indicates the discrete frequency in the periodogram, T/j indicates the number of periods per cycle, and the third and sixth columns the corresponding value of the periodogram.

Table 8: Frequencies and orders of integration in a linear model

	j_1	j_2	j_3	d_1	d_2	d_3
White noise	602 (4.66)	240 (11.70)	14 (200.57)	0.09 (0.02, 1.17)	0.06 (0.01, 0.09)	0.13 (0.11, 0.14)
Bloomfield	601 (4.67)	241 (11.65)	14 (200.57)	0.06 (-0.01, 1.17)	0.07 (0.02, 0.10)	0.05 (-0.02, 0.10)

The values in parenthesis in columns 2, 3 and 4 are the number of periods per cycle. Those in parenthesis in columns 5, 6 and 7 are the 95% confidence bands for the orders of integration at the respective frequencies.

Table 9: Frequencies and orders of integration in a non-linear model

	j_1	j_2	j_3	d_1	d_2	d_3
White noise	600 (4.69)	236 (11.89)	13 (216.00)	0.07 (-0.01, 0.14)	0.04 (-0.05, 0.08)	0.12 (0.05, 0.16)
Bloomfield	609 (4.61)	238 (11.78)	14 (200.57)	0.05 (-0.02, 0.19)	0.03 (-0.03, 0.07)	0.11 (0.04, 0.17)

The values in parenthesis in columns 2, 3 and 4 are the number of periods per cycle. Those in parenthesis in columns 5, 6 and 7 are the 95% confidence bands for the orders of integration at the respective frequencies.