Universidad de Navarra

Facultad de Ciencias Económicas y Empresariales

Working Paper nº 10/08

Fractional integration and data frequency

Luis A. Gil-Alana
Facultad de Ciencias Económicas y Empresariales
Universidad de Navarra

Guglielmo M. Caporale
Brunel University
London, England

Fractional integration and data frequency
Luis A. Gil-Alana and Guglielmo M. Caporale
Working Paper No.10/08
November 2008

ABSTRACT

This paper examines the robustness of fractional integration estimates to different data frequencies. We show by means of Monte Carlo experiments that if the number of differences is an integer value (e.g., 0 or 1) there is no distortion when data are collected at wider intervals; however, if it is a fractional value, the distortion increases as the number of periods between the observations increases, which results in lower orders of integration than those of the true DGP. An empirical application using the S&P500 index is also carried out.

Luis A. Gil-Alana
Universidad de Navarra
Depto. Métodos Cuantitativos
Campus Universitario
31080 Pamplona
alana@unav.es

Guglielmo M. Caporale
Brunel University
Centre for Empirical Finance
Uxbridge, Middlesex, UB8 3ph
England
Guglilemo-Maria.Caporale@brunel.uk.ac

1. Introduction

Time series techniques for modelling macroeconomic and financial data has developed considerably over the last twenty years. Till the 80s, the most common approach to describe nonstationarity in time series was to assume deterministic (usually, linear) functions that were regressed on time. Subsequently, unit root models became very popular (Box and Jenkins, 1970; Dickey and Fuller, 1979), especially after the influential paper of Nelson and Plosser (1982), who reported strong evidence of unit roots in US annual macroeconomic series.

Unit roots and linear trend models both have the advantage of conceptual and computational simplicity, but are both based on very specific assumptions about the nonstationary behaviour of a series. They are naturally thought of as rival models because a unit root without or with a drift implies a constant or linear trend function, the difference being in the disturbance term. The issue of stochastic (unit root) versus deterministic trend models in macroeconomics occupied a significant proportion of the literature in the 80s and the 90s, and it still remains an open question. Already in the 70s, Chan, Hayya and Ord (1977) studied both inappropriate detrending of integrated series and inappropriate differencing of trending series, and showed that the former produced spurious variation in the detrending series, while the latter produced spurious variation in the differenced series at high frequencies. These results were later extended by Nelson and Kang (1981, 1984) and Durlauf and Phillips (1988).

In the last thirty years there has been a growing body in the literature which studies the source of nonstationarity in macroeconomics and finance in terms of fractional differencing. The idea that is behind this specification is that the number of differences required to make a series I(0) might not necessarily be an integer value but could be any real one. For the purpose of the present paper, we define an I(0) process $\{u_t, t = 0, \pm 1, \dots\}$ as a covariance stationary process with a spectral density function that is positive and finite at the zero frequency. In this context, we say that x_t is integrated of order d, and denoted by $x_t \approx I(d)$ if :

$$(1 - L)^d \; x_t = u_t, \qquad t = 1, 2, \dots, \tag{1}$$

where d can be any real number, L is the lag operator ($Lx_t = x_{t-1}$) and u_t is I(0).[1] The polynomial on the left-hand-side in (1) can be expressed in terms of its Binomial expansion,

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

such that, for all real d,

$$(1 - L)^d x_t = x_t - d x_{t-1} + \frac{d(d-1)}{2} x_{t-2} - \dots .$$

Thus, if d is an integer value, x_t in (1) will be a function of a finite number of past observations, while if d is not an integer, x_t depends strongly upon values of the time series far away in the past, and the higher d is, the higher is the level of association between the observations. These processes (with $d > 0$) were initially introduced by Robinson (1978), Granger (1980, 1981) and Hosking (1981), and they have been widely employed in recent years to describe the dynamic behaviour of economic and financial data. (Diebold and Rudebusch, 1989; Baillie, 1996; Gil-Alana and Robinson, 1997; etc.).

[1] Throughout this paper we assume that $x_t = 0$ for $t \leq 0$, which is a standard assumption in applied work using fractional integration techniques (see the type I and II definitions of fractional integration, e.g. in Gil-Alana and Hualde, 2008).

On the other hand, it has been shown in recent years that in many empirical applications the fractional differencing parameter substantially changes depending on the data frequency used, especially so in the case of financial data (see, e.g., Cunado et al., 2005). In this paper we examine this issue by means of Monte Carlo experiments, and show that if the true DGP is I(d) with a non-integer value d, the estimated order of integration becomes smaller as the data are collected at wider intervals, and as the number of time periods between the collection points increases.

The outline of the paper is as follows. In Section 2 we briefly review the literature on the estimation and testing of I(d) models. Section 3 presents the Monte Carlo results. Section 4 discusses the empirical application, while Section 5 contains some concluding comments and suggested extensions.

2. A brief review of the literature on fractional integration

There exist many procedures for estimating and testing the fractional differencing parameter. These methods can be divided in three classes: the heuristic approaches (see, for example, Hurst, 1951; Higuchi, 1988, Lo, 1991 and Giraitis et al., 2003 among others); the semiparametric approaches, where the only requirement on u_t in (1) is to be I(0) but no functional form is assumed (see, for example, Geweke and Porter-Hudak, 1983; Robinson, 1995a,b; Velasco, 1999 and Phillips and Shimotsu, 2004, 2005 among others); and the maximum-likelihood methods (see, for example, Fox and Taqqu, 1986; Dahlhaus, 1989; Sowell, 1992; Robinson, 1994 and Tanaka, 1999 among other references).

In this paper we first employ a parametric testing approach developed by Robinson (1994) that enables us to test any real value d, thus including stationary and nonstationary processes. This method is based on the Lagrange Multiplier principle and uses the Whittle function in the frequency domain. For simplicity, we consider the case where the d-differenced process is uncorrelated, and test the null hypothesis:

$$H_o : d = d_o, \qquad (2)$$

in (1) for any real value d_o. The functional form of the test statistic is then given by:

$$\hat{r} = \frac{T^{1/2}}{\hat{\sigma}^2} \hat{A}^{-1/2} \hat{a}, \qquad (3)$$

where T is the sample size and

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

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

where $I(\lambda_j)$ is the periodogram of u_t evaluated under the null, i.e.,

$$\hat{u}_t = (1 - L)^{d_o} y_t .$$

Based on the null hypothesis H_o (2), Robinson (1994) established that under very mild regularity conditions:

$$\hat{r} \to_d N(0,1) \quad as \quad T \to \infty, \qquad (4)$$

and also the Pitman efficiency of the tests against local departures from the null.[2]

Additionally, we also employ a semiparametric method (Robinson, 1995), which is basically a local "Whittle estimate" in the frequency domain, based on a band of frequencies that degenerates to zero. The estimate is implicitly defined by:

---

[2] That means that if the test is defined against local alternatives of the form: H_a: $d = d_o + \delta T^{-1/2}$, with $\delta \neq 0$, the limit distribution is normal, with unit variance and mean that cannot be exceeded in absolute value by any rival regular statistic under Gaussianity of u_t.

$$\hat{d} = \arg\min_d \left(\log \overline{C(d)} - 2d \frac{1}{m} \sum_{j=1}^{m} \log \lambda_j \right), \qquad (5)$$

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

where m is a bandwidth parameter number, 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:

$$\sqrt{m}\,(\hat{d} - d_o) \to_d N(0, 1/4) \quad \text{as } T \to \infty,$$

where d_o is the true value of d and with the only additional requirement that m $\to \infty$ slower than T.[3]

3. The Monte Carlo results

Throughout this section we assume that the data generating process (DGP) is a purely I(d) process with T = 3,600 observations.[4] In Table 1 we assume that d = 0 (Table 1(i)); 0.20 (Table 1(ii)); 0.40 (1(iii)); 0.60 (1(iv)); 0.80 (1(v)) and 1 (in Table 1(vi)), and examine the rejection frequencies of Robinson's (1994) tests for d_o-values equal to d ± 0.1, 0.2, 0.3, 0.4 and 0.5, first for the whole sample period (x_t), and then collecting the data every two periods (x_{2t}); four periods (x_{4t}); six (x_{6t}); eight (x_{8t}); ten (x_{10t}) and twelve periods (x_{12t}).

[3] Other recent developments of fractional integration in semiparametric contexts and based on the Whittle function have been proposed in Velasco (1999) and Phillips and Shimotsu (2004, 2005) but these methods require additional user-chosen parameters, which may be very sensitive to the estimation of d.

[4] We generate Gaussian series using the routines GASDEV and RAN3 of Press, Flannery, Teukolsky and Vetterling (1986).

Thus, the sample sizes are clearly different for each sub-sample, containing 1,800 observations (x_{2t}); 900 (x_{4t}); 600 (x_{6t}); 450 (x_{8t}); 360 (x_{10t}) and 300 observations in x_{12t}.

[INSERT TABLE 1 ABOUT HERE]

The first result emerging from Table 1 (i) – (vi) is that if d is an integer value (i.e., d = 0, in Table 1(i), and d = 1, in Table 1(vi)), the procedure correctly determines the true order of integration in all subsamples, and the empirical size varies slightly below and above the nominal size of 5%. However, a different picture emerges if the order of integration lies between 0 and 1. Thus, if d = 0.2 (Table 1(ii)) the lowest rejection probabilities take place at d = 0.1 for x_{2t}, x_{4t}, x_{6t}, x_{8t} and x_{10t}, and at d = 0 for x_{12t}. Increasing the order of integration of the series, the bias is even more pronounced, and if d = 0.4 (Table 1(iii)), the lowest probabilities occur at d = 0.2 for x_{8t}, x_{10t} and x_{12t}. Then, as we approach 1, the bias tends to disappear, and, if d = 1, the lowest values ocurr at d = 1 in all cases.

The results presented so far, however, might be biased because of the different sample sizes used for each series examined. Thus, in what follows, we employ a constant sample size for each case. We start with x_t, with t = 1, 2, …, 300. Then, for x_{2t}, we take T = 600 and since the data are collected every two periods we also employ a sample of 300 observations. Similarly, for the remaining cases, x_{it}, i = 4, 6, 8, 10 and 12 we generate initial samples of 1200, 1800, 2400, 3000 and 3600 observations respectively, implying that in all cases the sample size examined is 300.

[INSERT TABLE 2 ABOUT HERE]

The results in Table 2 are very similar to those reported in Table 1. Thus, if d = 0 (in Table 2(i)) or d = 1 (in Table 2 (vi)), the lowest rejection probabilities occur in all cases at the true values of d. However, as we depart from these cases the lowest values are obtained at smaller orders of integration when data are collected at more than 1 period interval. Thus, for example, if d = 0.4 (Table 2 (iii)) and we collect the observations every 10 or 12 periods (x_{10t}, x_{12t}), the lowest frequencies take place in both cases at d = 0.2. The same happens if d = 0.6 (Table 2 (iv)) and the bias tends to disappear as we approach d = 1.

[INSERT FIGURE 1 ABOUT HERE]

Figure 1 displays the estimates of d based on the semiparametric Whittle method of Robinson (1995). Here, we consider the same DGPs as in Tables 1 and 2. In the case of the nonstationary models (i.e., d = 0.60, 0.80 and 1.00) the estimates are based on the first-differenced data, then adding 1 to the estimated values to get the proper orders of differencing. In all cases, the sample size is fixed, at T = 300. The values reported in the table correspond to the arithmetic means of the 1,000 replications of the experiment. For each case, we report the estimated d for a range of values of m (the bandwidth number) from 1 to T/2.[5] We also report in Figure 1 the 95% confidence interval corresponding to the true value of d. We observe that, similarly to the parametric case, if d is an integer

[5] Some methods to calculate the optimal bandwidth numbers in semiparametric contexts have been examined in Delgado and Robinson (1996) and Robinson and Henry (1996). However, in the case of the Whittle estimator of Robinson (1995), the use of optimal values has not been theoretically justified. Other authors, such as Lobato and Savin (1998) use an interval of values for m, but we have preferred to report the results for the whole range of values of m.

value, practically all the estimates are within the 95% interval. However, if d is not an integer, the values are below the intervals for practically all bandwidth numbers, especially in the cases of x_{10t} and x_{12t}.

4. The empirical application

The dataset analysed in this section is the Standard & Poor's 500 stock market index, and, initially, we consider monthly values from January 1950 until May 2008. Then, we collect data every two months (starting at February, 1950), three months (March, 1950), four months (April, 1950), six months (July 1950) and twelve months (December, 1950). Thus, we consider initially 701 observations, and subsequently analyse time series of length 350, 232, 175, 116 and 58 observations respectively. Though not reported, the time series plots were similar in the six cases.

[INSERT TABLE 3 ABOUT HERE]

Table 3 displays the estimates of d for each series based on the Whittle function for the two cases of white noise (Table 4(i)) and AR(1) errors (Table 4 (ii)). It also reports the 95% interval of those values of d_o where H_o cannot be rejected at the 5% level using Robinson's (1994) parametric approach. Here, we test H_o (2) in (1) assuming that x_t in (1) are the errors in a regression model of the form:

$$y_t = \alpha + \beta t + x_t; \qquad t = 0, 1, 2, ..., \qquad (6)$$

and we consider separately the three cases of no regressors in (6) (i.e., $\alpha = \beta = 0$ a priori), an intercept (α unknown and $\beta = 0$), and an intercept with a linear time trend (α and β unknown). Table 3 shows that the results are very similar in all three cases.

Starting with the results based on white noise disturbances (seeTable 3 (i)), we find that the estimated value of d is equal to 1.01 when using all the observations (monthly frequency) and the unit root hypothesis (i.e. d = 1) cannot be rejected at the 5% level. However, when collecting the data at other lower frequencies, the estimated d's are much higher, ranging now between 1.04 (when collecting four observations per year) and 1.49 (two observations per year). We see that the unit root null is now rejected in the majority of the cases. When allowing autocorrelation in the form of an AR(1) specification (see Table 3 (ii)) the estimated d is equal to 1.02 using the whole sample size (x_t), and the unit root hypothesis again cannot be rejected. However, when using frequencies the results dramatically change. For instance, when using x_{2t}, x_{3t} and x_{4t} (i.e., collecting six, four and three observations per year) the order of integration is found to be above 1 in all cases. Further, when using x_{10t} and x_{12t} (i.e. two and one observations per year) the estimated value of d is strictly smaller than 1 in all cases. Therefore, it seems that the order of differencing is very sensitive to the data frequency employed, especially in the autocorrelated case.

[INSERT TABLE 4 ABOUT HERE]

In Table 4 we follow the same procedure as in Table 3 but for the squared return series. Starting with the original data (x_t) we see that the estimated d is found to be 0.10 in case of white noise disturbances and slightly higher (0.13) with autocorrelated disturbances. In both cases the null hypothesis of d = 0 is rejected in favour of long memory (d > 0). If we now collect the data every two periods (x_{6t}), d is slightly smaller (about 0.07) and the I(0) hypothesis cannot be rejected if u_t is AR(1), and the same happens in all the remaining cases even with uncorrelated disturbances. This second application seems to be more relevant than the previous one for the purposes of the present paper. It shows that the order of integration is smaller for lower data frequencies, which is in line with the Monte Carlo results of Section 3. Also, the conclusions are very different depending on the data frequency employed, evidence of long memory being found with monthly data, whilst short memory seems to characterise lower frequencies.

5. Conclusions

This paper analyses the robustness of fractional integration estimates to different data frequencies. We show by means of Monte Carlo experiments that if the true DGP is I(d) with d lying between 0 and 1, when the data are collected at lower frequencies, the methods to detect the order of differencing tend to produce lower estimates for d. This bias tends to be higher as we further depart from 0 or 1. Thus, one must be cautious when estimating the degree of dependence through the fractional differencing parameter if the data are available at different frequencies. We carried out an empirical application using monthly data for the S&P 500 index. The order of integration was estimated and it was found to be close to 1 for the two cases of white noise and autocorrelated disturbances. However, when collecting the data at lower frequencies, d was found to be above 1 in case of uncorrelated errors, and strictly below 1 for AR(1) disturbances for data collected at 1 and 2 observations per year. The same happened for the volatility processes when conducting the same type of analysis for the squared return series, as d was found to be strictly above 0 in the original daily series but the I(0) hypothesis could not be rejected at other frequencies.

References

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

Box, G.E.P. and G.M. Jenkins, 1970, Time Series Analysis: Forecasting and Control, San Francisco, Holden-Day.

Chan, K.H., J.C. Hayya and J.K. Ord, 1977, A note on trend removal methods: the case of polynomial regression versus variate differencing, Econometrica 45, 737-744.

Cunado, J., L.A. Gil-Alana and F. Perez de Gracia, 2005, A test for rational bubbles in the Nasdaq stock index. A fractionally integrated approach, Journal of Banking and Finance 29, 2633-2654.

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

Delgado, M.A. and P.M. Robinson, 1996, Optimal spectral bandwidth for long memory. Statistica Seneca, 6, 97-112.

Dickey, D. A. and W.A. Fuller, 1979, Distribution of the estimators for autoregressive time series with a unit root. Journal of the American Statistical Association, 74, 427-431.

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

Durlauf, S. and P.C.B. Phillips, 1988, Trends versus random walks in time series analysis, Econometrica 56, 1333-1354.

Fox, R. and M.S. Taqqu, 1986, Large sample properties of parameter estimates for strongly dependent stationary Gaussian time series, Annals of Statistics 14, 517-532.

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

Gil-Alana, L.A. and J. Hualde, 2008, Fractional integration and cointegration. A review and an empirical application, Mills and Patterson, Palgrave Handbook in Econometrics, Vol. 2, forthcoming.

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.

Giraitis, L., P. Kokoszka, R. Leipus and G. Teyssiere, 2003. Rescaled variance and related tests for long memory in volatiluty and levels. *Journal of Econometrics* 112, 265-294.

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

Granger, C.W.J., 1981, Some properties of time series data and their use in econometric model specification, Journal of Econometrics 16, 121-130.

Higuchi, T., 1988, Approach to an irregular time series on the basis of the fractal theory, Physica D 31, 277-283.

Hosking, J.R.M., 1981, Fractional differencing. Biometrika 68, 168-176.

Hurst, H.E., 1951, Long term storage capacity of reservoirs, Transactions of the American Society of Civil Engineers 116, 770-799.

Lo, A., 1991, Long term memory in stock prices, Econometrica 59, 1279-1313.

Lobato, I.N. and N.E. Savin, 1998, Real and spurious long memory properties of stock market data. Journal of Business and Economic Statistics, 16, 261-283.

Nelson, C.R. and H. Kang, 1981, Spurious periodicity in inappropriately detrended time series, Econometrica 49, 741-751.

Nelson, C.R. and H. Kang, 1984, Pitfalls in the use of time as an explanatory variable in regression, Journal of Business and Economic Statistics 2, 73-82.

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

Phillips, P.C.B., Shimotsu, K., 2004, Local Whittle estimation in nonstationary and unit root cases, Annals of Statistics 32, 656-692.

Phillips, P.C.B., Shimotsu, K., 2005, Exact local Whittle estimation of fractional integration, Annals of Statistics 33, 1890-1933.

Press, W.H., B.P. Flannery, S.A. Teukolsky and W.T. Vetterling, 1986, Numerical recipes: The Art of Scientific Computing, Cambridge University Press, Cambridge.

Robinson, P.M., 1978, Statistical inference for a random coefficient autoregressive model, Scandinavian Journal of Statistics 5, 163-168.

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

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.

Robinson, P.M. and M. Henry, 1996, Bandwidth choice in Gaussian semiparametric estimation of long-range dependence, P.M. Robinson and M. Rosenblatt eds. Athens Conference on Applied Probability in Time Series Analysis, Vol.II, New York, 220-232.

Sowell, F., 1992, Maximum likelihood estimation of stationary univariate fractionally integrated time series models, Journal of Econometrics 53, 165-188.

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

Velasco, C., 1999, Gaussian semiparametric estimation of nonstationary time series, Journal of Time Series Analysis 20, 87-127.

Table 1(i): Rejection probabilities in Robinson (1994) with d = 0

d_o	-0.5	-0.4	-0.3	-0.2	-0.1	**0**	0.1	0.2	0.3	0.4	0.5
x_t	1.000	1.000	1.000	1.000	0.986	**0.054**	0.993	1.000	1.000	1.000	1.000
x_{2t}	1.000	1.000	1.000	1.000	0.839	**0.046**	0.857	1.000	1.000	1.000	1.000
x_{4t}	1.000	1.000	1.000	0.972	0.529	**0.055**	0.560	0.978	1.000	1.000	1.000
x_{6t}	1.000	1.000	0.996	0.910	0.369	**0.046**	0.380	0.915	0.999	1.000	1.000
x_{8t}	1.000	0.999	0.969	0.789	0.276	**0.056**	0.287	0.797	0.983	1.000	1.000
X_{10t}	0.998	0.989	0.924	0.649	0.227	**0.076**	0.295	0.699	0.951	0.998	1.000
X_{12t}	0.996	0.980	0.892	0.547	0.181	**0.046**	0.227	0.621	0.903	0.983	1.000

In bold, the lowest probabilities across d_o.

Table 1(ii): Rejection probabilities in Robinson (1994) with d = 0.20

d_o	-0.3	-0.2	-0.1	0	0.1	**0.2**	0.3	0.4	0.5	0.6	0.7
x_t	1.000	1.000	1.000	1.000	0.986	**0.054**	0.993	1.000	1.000	1.000	1.000
x_{2t}	1.000	1.000	1.000	0.986	**0.292**	0.405	0.997	1.000	1.000	1.000	1.000
x_{4t}	1.000	1.000	0.980	0.561	**0.066**	0.576	0.987	1.000	1.000	1.000	1.000
x_{6t}	1.000	0.987	0.823	0.277	**0.085**	0.541	0.959	1.000	1.000	1.000	1.000
x_{8t}	0.995	0.936	0.650	0.157	**0.109**	0.483	0.907	0.995	1.000	1.000	1.000
x_{10t}	0.977	0.856	0.502	0.118	**0.117**	0.477	0.837	0.981	0.999	1.000	1.000
x_{12t}	0.936	0.763	0.368	**0.091**	0.094	0.403	0.795	0.954	0.996	1.000	1.000

Table 1(iii): Rejection probabilities in Robinson (1994) with d = 0.40

d_o	-0.1	0	0.1	0.2	0.3	**0.4**	0.5	0.6	0.7	0.8	0.9
x_t	1.000	1.000	1.000	1.000	0.986	**0.054**	0.993	1.000	1.000	1.000	1.000
x_{2t}	1.000	1.000	1.000	0.986	**0.277**	0.468	0.997	1.000	1.000	1.000	1.000
x_{4t}	1.000	1.000	0.969	0.529	**0.090**	0.683	0.993	1.000	1.000	1.000	1.000
x_{6t}	1.000	0.980	0.788	0.244	**0.150**	0.671	0.980	1.000	1.000	1.000	1.000
x_{8t}	0.994	0.923	0.569	**0.138**	0.168	0.632	0.953	0.999	1.000	1.000	1.000
x_{10t}	0.968	0.812	0.411	**0.107**	0.186	0.605	0.929	0.993	0.999	1.000	1.000
x_{12t}	0.912	0.690	0.310	**0.091**	0.201	0.566	0.888	0.984	0.997	1.000	1.000

Table 1(iv): Rejection probabilities in Robinson (1994) with d = 0.60

d_o	0.1	0.2	0.3	0.4	0.5	**0.6**	0.7	0.8	0.9	1	1.1
x_t	1.000	1.000	1.000	1.000	0.986	**0.054**	0.993	1.000	1.000	1.000	1.000
x_{2t}	1.000	1.000	1.000	0.995	0.439	**0.295**	0.995	1.000	1.000	1.000	1.000
x_{4t}	1.000	1.000	0.990	0.735	**0.097**	0.442	0.972	1.000	1.000	1.000	1.000
x_{6t}	1.000	0.996	0.899	0.462	**0.088**	0.434	0.921	0.999	1.000	1.000	1.000
x_{8t}	0.998	0.977	0.782	0.286	**0.091**	0.418	0.887	0.992	0.999	1.000	1.000
x_{10t}	0.993	0.911	0.644	0.224	**0.092**	0.403	0.805	0.981	0.997	1.000	1.000
x_{12t}	0.964	0.837	0.526	0.163	**0.115**	0.353	0.754	0.955	0.996	0.999	1.000

Table 1(v): Rejection probabilities in Robinson (1994) with d = 0.80

d_o	0.3	0.4	0.5	0.6	0.7	**0.8**	0.9	1	1.1	1.2	1.3
x_t	1.000	1.000	1.000	1.000	0.986	**0.054**	0.993	1.000	1.000	1.000	1.000
x_{2t}	1.000	1.000	1.000	0.999	0.667	**0.125**	0.966	1.000	1.000	1.000	1.000
x_{4t}	1.00	1.000	1.000	0.898	0.265	**0.163**	0.844	0.997	1.000	1.000	1.000
x_{6t}	1.000	0.999	0.978	0.722	0.180	**0.166**	0.719	0.978	1.000	1.000	1.000
x_{8t}	1.000	0.994	0.918	0.565	**0.128**	0.155	0.648	0.948	0.998	1.000	1.000
x_{10t}	0.998	0.971	0.838	0.444	**0.107**	0.151	0.552	0.898	0.989	0.999	1.000
x_{12t}	0.989	0.941	0.744	0.372	**0.100**	0.152	0.485	0.832	0.977	0.999	0.999

Table 1(vi): Rejection probabilities in Robinson (1994) with d = 1.00

d_o	0.5	0.6	0.7	0.8	0.9	**1**	1.1	1.2	1.3	1.4	1.5
x_t	1.000	1.000	1.000	1.000	0.986	**0.054**	0.993	1.000	1.000	1.000	1.000
x_{2t}	1.000	1.000	1.000	1.000	0.851	**0.045**	0.850	1.000	1.000	1.000	1.000
x_{4t}	1.000	1.000	1.000	0.975	0.541	**0.056**	0.557	0.980	1.000	1.000	1.000
x_{6t}	1.000	1.000	0.996	0.885	0.375	**0.062**	0.418	0.896	0.997	1.000	1.000
x_{8t}	1.000	1.000	0.973	0.775	0.259	**0.057**	0.314	0.830	0.980	1.000	1.000
x_{10t}	0.999	0.994	0.931	0.672	0.228	**0.064**	0.261	0.683	0.950	0.993	0.999
x_{12t}	0.998	0.979	0.863	0.573	0.212	**0.064**	0.228	0.603	0.899	0.987	0.998

Table 2(i): Rejection probabilities in Robinson (1994) with d = 0 (T = 300)

	-0.5	-0.4	-0.3	-0.2	-0.1	**0**	0.1	0.2	0.3	0.4	0.5
x_t	1.000	1.000	1.000	0.972	0.554	**0.055**	0.548	0.988	1.000	1.000	1.000
x_{2t}	1.000	1.000	1.000	0.975	0.546	**0.061**	0.524	0.979	1.000	1.000	1.000
x_{4t}	1.000	1.000	1.000	0.972	0.529	**0.055**	0.560	0.978	1.000	1.000	1.000
x_{6t}	1.000	1.000	1.000	0.982	0.547	**0.051**	0.552	0.988	1.000	1.000	1.000
x_{8t}	1.000	1.000	1.000	0.971	0.538	**0.065**	0.534	0.976	1.000	1.000	1.000
x_{10t}	1.000	1.000	1.000	0.968	0.534	**0.054**	0.551	0.989	1.000	1.000	1.000
x_{12t}	1.000	1.000	1.000	0.970	0.534	**0.060**	0.544	0.982	1.000	1.000	1.000

In bold, the lowest probabilities across d_o.

Table 2(ii): Rejection probabilities in Robinson (1994) with d = 0.20 (T = 300)

	-0.3	-0.2	-0.1	0	0.1	**0.2**	0.3	0.4	0.5	0.6	0.7
x_t	1.000	1.000	1.000	0.972	0.554	**0.055**	0.548	0.988	1.000	1.000	1.000
x_{2t}	1.000	1.000	0.997	0.836	**0.173**	0.228	0.902	0.999	1.000	1.000	1.000
x_{4t}	1.000	1.000	0.980	0.561	**0.066**	0.576	0.987	1.000	1.000	1.000	1.000
x_{6t}	1.000	1.000	0.943	0.418	**0.086**	0.715	0.997	1.000	1.000	1.000	1.000
x_{8t}	1.000	0.999	0.930	0.327	**0.125**	0.769	0.997	1.000	1.000	1.000	1.000
x_{10t}	1.000	1.000	0.903	0.248	**0.149**	0.825	1.000	1.000	1.000	1.000	1.000
x_{12t}	1.000	0.999	0.875	0.217	**0.171**	0.857	0.999	1.000	1.000	1.000	1.000

Table 2(iii): Rejection probabilities in Robinson (1994) with d = 0.40 (T = 300)

	-0.1	0	0.1	0.2	0.3	**0.4**	0.5	0.6	0.7	0.8	0.9
x_t	1.000	1.000	1.000	0.972	0.554	**0.055**	0.548	0.988	1.000	1.000	1.000
x_{2t}	1.000	1.000	0.997	0.818	**0.177**	0.256	0.912	0.999	1.000	1.000	1.000
x_{4t}	1.000	1.000	0.969	0.529	**0.090**	0.683	0.993	1.000	1.000	1.000	1.000
x_{6t}	1.000	0.999	0.928	0.370	**0.152**	0.815	0.999	1.000	1.000	1.000	1.000
x_{8t}	1.000	0.999	0.905	0.256	**0.225**	0.899	1.000	1.000	1.000	1.000	1.000
x_{10t}	1.000	0.998	0.854	**0.187**	0.289	0.941	1.000	1.000	1.000	1.000	1.000
x_{12t}	1.000	0.998	0.802	**0.151**	0.346	0.969	1.000	1.000	1.000	1.000	1.000

Table 2(iv): Rejection probabilities in Robinson (1994) with d = 0.60 (T = 300)

	0.1	0.2	0.3	0.4	0.5	**0.6**	0.7	0.8	0.9	1	1.1
x_t	1.000	1.000	1.000	0.972	0.554	**0.055**	0.548	0.988	1.000	1.000	1.000
x_{2t}	1.000	1.000	0.996	0.895	0.254	**0.172**	0.861	0.999	1.000	1.000	1.000
x_{4t}	1.000	1.000	0.996	0.735	**0.097**	0.442	0.972	1.000	1.000	1.000	1.000
x_{6t}	1.000	1.000	0.985	0.964	**0.096**	0.551	0.990	1.000	1.000	1.000	1.000
x_{8t}	1.000	1.000	0.981	0.584	**0.088**	0.656	0.995	1.000	1.000	1.000	1.000
x_{10t}	1.000	0.999	0.971	0.534	**0.088**	0.702	0.997	1.000	1.000	1.000	1.000
x_{12t}	1.000	1.000	0.965	0.522	**0.103**	0.724	1.000	1.000	1.000	1.000	1.000

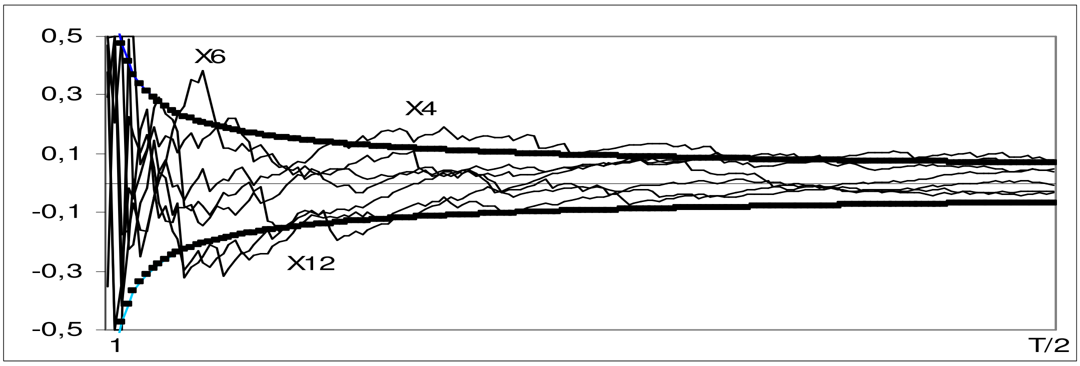
Table 2(v): Rejection probabilities in Robinson (1994) with d = 0.80 (T = 300)

	0.3	0.4	0.5	0.6	0.7	**0.8**	0.9	1	1.1	1.2	1.3
x_t	1.000	1.000	1.000	0.972	0.554	**0.055**	0.548	0.988	1.000	1.000	1.000
x_{2t}	1.000	1.000	0.999	0.932	0.407	**0.092**	0.717	0.996	1.000	1.000	1.000
x_{4t}	1.000	1.000	1.000	0.898	0.265	**0.163**	0.843	0.997	1.000	1.000	1.000
x_{6t}	1.000	1.000	0.998	0.882	0.272	**0.197**	0.870	0.996	1.000	1.000	1.000
x_{8t}	1.000	1.000	0.998	0.885	0.222	**0.217**	0.899	0.999	1.000	1.000	1.000
x_{10t}	1.000	1.000	0.995	0.873	**0.209**	0.232	0.909	1.000	1.000	1.000	1.000
x_{12t}	1.000	1.000	0.998	0.854	**0.201**	0.247	0.902	1.000	1.000	1.000	1.000

Table 2(vi): Rejection probabilities in Robinson (1994) with d = 1.00 (T = 300)

	0.5	0.6	0.7	0.8	0.9	**1**	1.1	1.2	1.3	1.4	1.5
x_t	1.000	1.000	1.000	0.972	0.554	**0.055**	0.548	0.988	1.000	1.000	1.000
x_{2t}	1.000	1.000	1.000	0.972	0.569	**0.061**	0.522	0.974	1.000	1.000	1.000
x_{4t}	1.000	1.000	1.000	0.975	0.541	**0.056**	0.556	0.980	1.000	1.000	1.000
x_{6t}	1.000	1.000	1.000	0.969	0.554	**0.058**	0.538	0.981	1.000	1.000	1.000
x_{8t}	1.000	1.000	1.000	0.971	0.518	**0.057**	0.578	0.987	1.000	1.000	1.000
x_{10t}	1.000	1.000	0.999	0.975	0.545	**0.043**	0.544	0.985	1.000	1.000	1.000
x_{12t}	1.000	1.000	1.000	0.976	0.548	**0.054**	0.531	0.988	1.000	1.000	1.000

Figure 1(i): Whittle semiparametric estimates (Robinson, 1995) with d = 0.00



The horizontal axis refers to the bandwidth while the vertical one displays the values of d. We also show the 95% confidence interval for the I(d) hypothesis.

Figure 1(ii): Whittle semiparametric estimates (Robinson, 1995) with d = 0.20

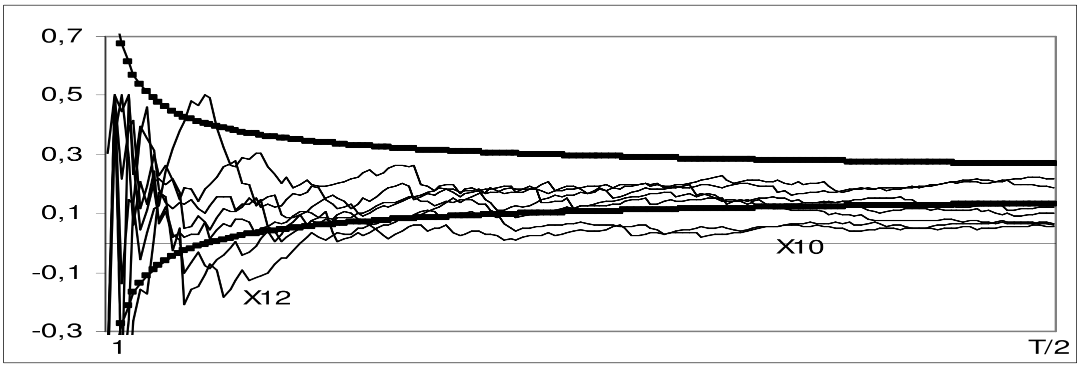


Figure 1(iii): Whittle semiparametric estimates (Robinson, 1995) with d = 0.40

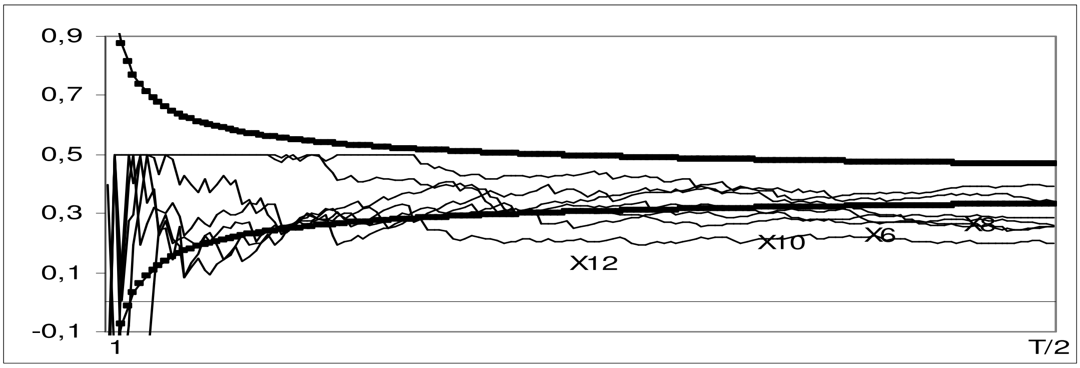


Figure 1(iv): Whittle semiparametric estimates (Robinson, 1995) with d = 0.60

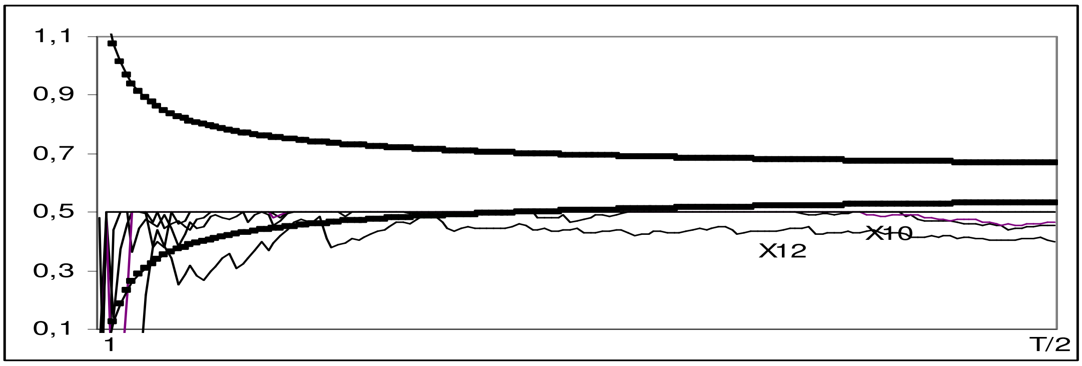


Figure 1(v): Whittle semiparametric estimates (Robinson, 1995) with d = 0.80

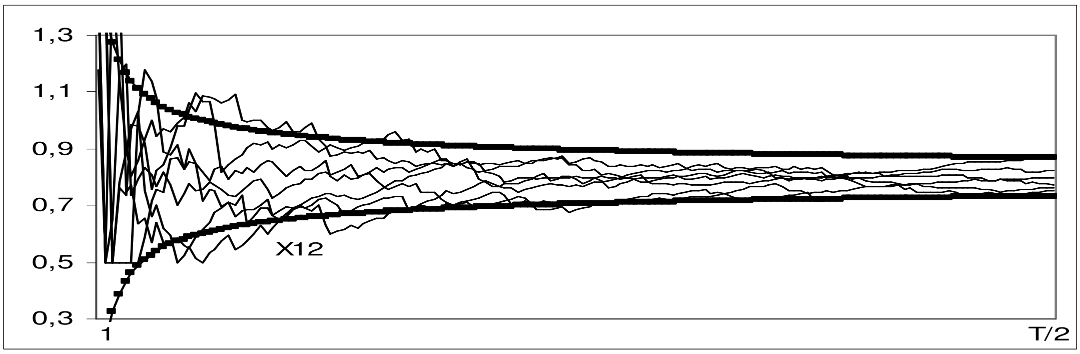


Figure 1(vi): Whittle semiparametric estimates (Robinson, 1995) with d = 1.00

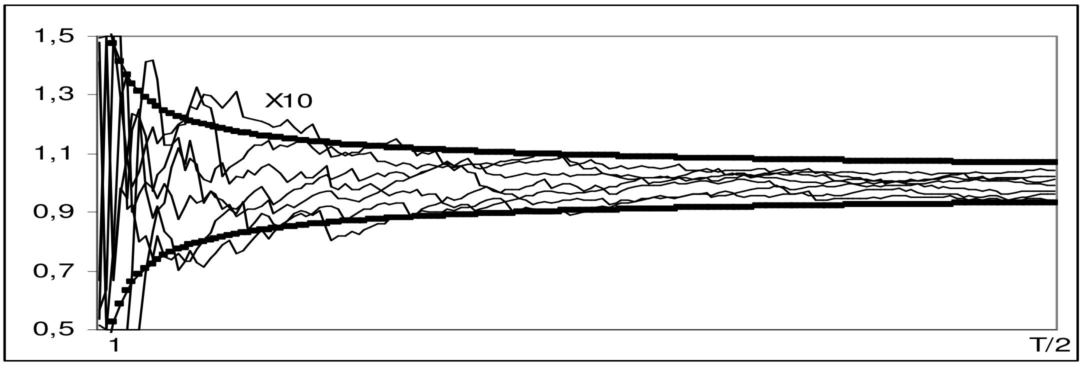


Table 3(i): Estimates of d in the S&P500 under the assumption of white noise u_t

Observations per year		Sample size	No regressors	An intercept	A linear trend
Twelve (12)	(x_t)	701	1.01	1.01	1.01
Six (6)	(x_{2t})	350	1.07	1.07	1.07
Four (4)	(x_{3t})	233	1.04	1.05	1.05
Three (3)	(x_{4t})	175	1.10	1.10	1.10
Two (2)	(x_{6t})	116	1.49	1.49	1.49
One (1)	(x_{12t})	58	1.21	1.21	1.21

The values are the estimates of d. In brackets the 95% confidence bands.

Table 3(ii): Estimates of d in the S&P500 under the assumption of AR(1) u_t

Observations per year		Sample size	No regressors	An intercept	A linear trend
Twelve (12)	(x_t)	701	1.02	1.02	1.02
Six (6)	(x_{2t})	350	1.10	1.10	1.10
Four (4)	(x_{3t})	233	1.20	1.21	1.21
Three (3)	(x_{4t})	175	1.32	1.33	1.33
Two (2)	(x_{6t})	116	0.80	0.78	0.76
One (1)	(x_{12t})	58	0.62	0.73	0.74

Table 4(i): Estimates of d in the squared returns under white noise u_t

Observations per year		Sample size	No regressors	An intercept	A linear trend
Twelve (12)	(x_t)	701	0.10	0.10	0.10
Six (6)	(x_{2t})	350	0.07	0.07	0.07
Four (4)	(x_{3t})	233	0.06	0.06	0.06
Three (3)	(x_{4t})	175	0.06	0.06	0.06
Two (2)	(x_{6t})	116	0.14	0.14	0.11
One (1)	(x_{12t})	58	-0.08	-0.08	-0.09

The values are the estimates of d. In brackets the 95% confidence bands.

Table 4(ii): Estimates of d in the squared returns under the assumption of AR(1) u_t

Observations per year		Sample size	No regressors	An intercept	A linear trend
Twelve (12)	(x_t)	701	0.13	0.13	0.13
Six (6)	(x_{2t})	350	0.08	0.08	0.07
Four (4)	(x_{3t})	233	0.15	0.15	0.15
Three (3)	(x_{4t})	175	0.11	0.11	0.12
Two (2)	(x_{6t})	116	-0.22	-0.21	-0.21
One (1)	(x_{12t})	58	-0.21	-0.21	-0.40