Facultad de Ciencias Económicas y Empresariales
Universidad de Navarra

Working Paper nº 02/11

LONG MEMORY AND FRACTIONAL INTEGRATION IN HIGH FREQUENCY BRITISH POUND / US DOLLAR SPOT EXCHANGE RATES

Guglielmo Maria Caporale
Brunel University London
London, United Kingdom

and

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

Long memory and fractional integration in high frequency British pound / US dollar spot exchange rates
Guglielmo M. Caporale and Luis A. Gil-Alana
Working Paper No.02/11
January 2011

ABSTRACT

This paper analyses the long-memory properties of high frequency financial time series. It focuses on temporal aggregation and the influence that this might have on the degree of dependence of the series. Fractional integration or I(d) models are estimated with a variety of specifications for the error term. In brief, we find evidence that a lower degree of integration is associated with lower data frequencies. In particular, when the data are collected every 10 minutes there are several cases with values of d strictly smaller than 1, implying mean-reverting behaviour. This holds for all four series examined, namely Open, High, Low and Last observations for the British pound/US dollar spot exchange rate.

Luis A. Gil-Alana
University of Navarra
Faculty of Economics
Edificio Biblioteca, E. Este
E-31080 Pamplona, Spain
alana@unav.es

Guglielmo Maria Caporale
Centre for Empirical Finance
Brunel University
West London, UB8 3PH
UK
Guglielmo-Maria.Caporale@brunel.ac.uk

1. Introduction

The Efficient Market Hypothesis (EMH) in its weak form rules out the possibility of abnormal systematic profits over and above transaction costs and risk premia, as prices should fully reflect available information (see Fama, 1970). The implication is that stock prices should follow a random walk process, which implies unpredictable returns (see Summers, 1986). Therefore, a finding of mean reversion in stock prices is seen as inconsistent with equilibrium asset pricing models (see, e.g., Poterba and Summers, 1988 and Fama and French, 1988). A large number of studies have been carried out to establish whether prices are indeed I(1) and, consequently, stock market returns I(0) series, although business cycle variation and short-range dependence might also lead to a rejection of long memory in stock prices (see Lo, 1991). However, as we argued in Caporale and Gil-Alana (2002), the assumptions imposed by standard unit root tests might be too restrictive, and the possibility of fractional orders of integration with a slow rate of decay should be considered. Therefore that study performed tests allowing for fractional alternatives that incorporates the I(0) and the I(1) models as particular cases of interest, and found that US real stock returns are close to being I(0) (note that if shocks are weakly autocorrelated, markets will not be efficient). Fractional integration models (at the long run or zero frequency) have also been used for inflation and interest rates (see, e.g., Shea, 1991; Backus and Zhin, 1993; Hassler and Wolters, 1995; Baillie et al., 1996, etc.).

A subsequent contribution (see Caporale and Gil-Alana, 2007) decomposed the stochastic process followed by US stock prices into a long-run component described by the fractional differencing parameter (d) and a short-run (ARMA) structure. Finally, in Caporale and Gil-Alana (2008) we introduced a more general model which, instead of

considering exclusively the component affecting the long-run or zero frequency, also takes into account the cyclical structure. Specifically, a procedure was applied which allows to test simultaneously for unit roots with possibly fractional orders of integration at both the zero and the cyclical frequencies. Modelling simultaneously the zero and the cyclical frequencies can solve at least to some extent the problem of misspecification that might arise with respect to these two frequencies.

However, the fractional differencing parameter may be very sensitive to the data frequency. In fact, it has often been claimed that aggregation is behind fractional integration: Robinson (1978) and Granger (1980) showed that the aggregation of heterogeneous individual AR process may produce fractional integration. In a more recent paper, Souza (2005) gives a formula for the spectral density of the aggregated process and shows that long memory is very likely under aggregation. On the other hand, it is well known that temporal aggregation leads to finite sample biases in the estimates of the fractional differencing parameter (see, e.g. Souza and Smith, 2002). This is the issue that will be investigated in the present study by using high frequency data on the British pound-US dollar spot exchange rate collected every 1, 2, 3, 5, and 10 minutes. As in Caporale and Gil-Alana (2008), we start the analysis using a long memory model for both the zero and cyclical frequencies; however, since the evidence clearly suggested orders of integration close to zero for the cyclical frequencies, we then focus exclusively on the long run or zero frequency. Excellent surveys on the use of high frequency financial data in econometrics are among others the papers of Campbell, Lo and MacKinlay (1997), Goodhart and O`Hara (1997), Wood (2000), Andersen (2000), Ghysels (2000), Bawens and Giot (2001), Gourieroux and Jasiak (2001), Lyons (2001), Dacarogna et al. (2001), Tsay (2001) and Engle and Russel (2004).

Fractional integration in exchange rates markets has been examined in a number of papers. A large number of papers deal with the theory of Purchasing Power Parity (PPP), which occupies a central place in international economics. In fact, it is a key building block in monetary models of exchange rate determination. In a flexible-price monetary model, PPP is assumed to hold continuously. In a sticky-price model, PPP does not hold, but is a maintained assumption for the long run. The question of interest is then to determine if deviations from PPP are transitory or permanent. Applying R/S techniques to daily rates for the British pound, French franc and Deutsche mark, Booth, Kaen and Koveos (1982) found positive memory during the flexible exchange rate period (1973-1979) but negative one (i.e., anti-persistence) during the fixed exchange rate period (1965-1971). Later, Cheung (1993) also found evidence of long memory behaviour in foreign exchange markets during the managed floating regime. On the other hand, Baum, Barkoulas and Caglayan (1999) estimated ARFIMA models for real exchange rates in the post-Bretton Woods era and found almost no evidence to support long run PPP. Additional papers on exchange rate dynamics using fractional integration are Fang, Lai and Lai (1994), Crato and Ray (2000) and Wang (2004).

The layout of the paper is as follows. Section 2 describes the econometric methodology used. Section 3 provides details of the data and discusses the empirical results. Section 4 summarises the main findings and offers some concluding remarks.

2. Methodology

There are two definitions of long memory, one in the frequency domain and the other in the time domain. Let us consider a zero-mean covariance stationary process $\{x_t,$

$t = 0, \pm 1, \dots \}$ with autocovariance function $\gamma_u = E(x_t x_{t+u})$. The time domain definition of long memory states that:

$$\sum_{u=-\infty}^{\infty} |\gamma_u| = \infty.$$

Assuming that x_t has an absolutely continuous spectral distribution, so that it has a spectral density function of the following type:

$$f(\lambda) = \frac{1}{2\pi} \left(\gamma_0 + 2 \sum_{u=1}^{\infty} \gamma_u \cos(\lambda u) \right),$$

according to the frequency domain definition of long memory the spectral density function is unbounded at some frequency λ in the interval $[0, \pi)$.

Most of the existing empirical literature considers the case when the singularity or pole in the spectrum occurs at the zero frequency. This is the standard case of $I(d)$ models of the form:

$$(1 - L)^d x_t = u_t, \quad t = 0, \pm 1, \dots, \qquad (1)$$

$$x_t = 0, \quad t \leq 0,$$

with d > 0, where L is the lag-operator ($Lx_t = x_{t-1}$) and u_t is $I(0)$, being defined as a covariance stationary process with a spectral density function that is positive and finite at any frequency. This includes a wide range of model specifications such as the white noise, the stationary autoregression (AR), moving average (MA), stationary ARMA etc.[1]

The I(d) models of the form given by equation (1) were introduced by Granger (1980, 1981), Granger and Joyeux (1980) and Hosking (1981) and since then have been widely employed to describe the behaviour of many economic time series (see, e.g.,

Diebold and Rudebusch, 1989; Sowell, 1992; Gil-Alana and Robinson, 1997; etc.). Note that the parameter d plays a crucial role in describing the degree of dependence of the series. Specifically, if d = 0 in (1), $x_t = u_t$, and the series is I(0), potentially including ARMA structures, with the autocorrelations decaying then exponentially rapid. If d belongs to the interval (0, 0.5), the series is still covariance stationary but the autocorrelations take longer to disappear than in the I(0) case. If d is in the interval [0.5, 1), the series is no longer covariance stationary; however, it is still mean-reverting with shocks affecting it disappearing in the long run. Finally, if d ≥ 1 the series is nonstationary and non-mean-reverting.

u   In this paper we analyse the long memory (fractional integration) property of high frequency financial data, noting that the fractional differencing parameter can change substantially depending on the data frequency employed. The methodology employed here to estimate the fractional differencing parameter is based on the Whittle function in the frequency domain (Dahlhaus, 1989). We also employ a testing procedure developed by Robinson (1994) allowing for any real value of d in I(d) models. The latter is a Lagrange Multiplier (LM) procedure which is the most efficient one in the context of fractional integration. It tests the null hypothesis H_o: d = d_o for any real value d_o in (1), and given the fact that the test statistic follows a standard (normal) limit distribution it is possible to construct confidence bands for the non-rejection values.[2]

---

[1] In fact, if u_t in (1) is ARMA(p,q), y_t is said to follow a fractionally integrated ARMA, ARFIMA(p, d, q) model.

3. Data and empirical results

The data used for the analysis are taken from Reuters, and are intraday data for several days at the 1, 2, 3, 5, 10-minute frequency. Specifically, the series whose properties are being investigated is the nominal exchange rate of the British pound vis-à-vis the US dollar.

[Insert Figures 1 – 4 about here]

Figure 1 shows plots of the four series, i.e., Open, High, Low and Last values of the exchange rate collected every minute, where High (Low) stands for the highest (lowest) price and Open (Last) for the initial (last) price observed in that time interval respectively. Their corresponding returns, obtained as the first differences of the log-prices, are shown in Figure 2. Figures 3 and 4 display the correlograms and the periodograms of the return series. The values of the former seem to indicate that the original series may be I(1), suggesting the possibility of random walk behaviour; however, the presence of some significant values even at lags far away from zero might indicate weak autocorrelation and/or fractional integration.[3]

First, we estimate the value of d for the four series at the highest frequency (i.e., with data collected each minute). For this purpose we consider the following model:

$$y_t \;=\; \alpha \;+\; \beta t \;+\; x_t; \qquad (1 - L)^d x_t \;=\; u_t; \qquad t = 1, 2, ..., \quad (2)$$

where y_t is the time series observed, α and β are the deterministic terms (an intercept and a linear time trend respectively), and x_t is assumed to be I(d), where d can be any real number. Different assumptions will be made about the error term u_t.[4]

[Insert Table 1 about here]

Table 1 displays the results of the Whittle estimates of d along with the 95% confidence interval of the non-rejection values according to Robinson's (1994) parametric approach. The error term u_t is assumed to be a white noise in Table 1a, an AR(1) process in Table 1b, whilst it is specified using the exponential spectral model of Bloomfield (1973) in Table 1c. This is a non-parametric approach to modelling I(0) terms that produces autocorrelations decaying exponentially as in the AR(MA) case.

Table 1 shows the results of the estimated values of d, for the three standard cases of no regressors (i.e., $\alpha = \beta = 0$ in (2)), an intercept (α unknown and $\beta = 0$), and an intercept with a linear time trend (α and β unknown). Starting with the case of white noise errors (Table 1a), it can be seen that for "Open" and "Last" the estimates are slightly below 1, though the unit root null cannot be rejected in any case. However, for "High" and "Low" the unit root hypothesis is rejected in favour of higher degrees of integration in the cases of an intercept and an intercept with a linear time trend. When allowing autocorrelation in the form of an AR(1) process (in Table 1b), the results vary depending on the inclusion or not of deterministic terms. Specifically, if no regressors are included in the regression model, d is found to be strictly higher than 1 for all four series; however, when including deterministic terms, the unit root null is almost never rejected. The only exception is "High" with a linear trend, when the estimated value of d is found to be 0.936, and the interval excludes the unit root in favour of mean reversion. When adopting the more general Bloomfield specification (Table 1c), the unit root null hypothesis is never rejected.

[Insert Figures 5 and 6 about here]

---

[3] Some authors may argue that it does not make sense to model each series separately and treat them in the same way. However, a simple inspection at Figure 1 shows that the four series move roughly in a very similar way.

Next we focus on the variance of the return series and examine the squared and absolute returns, which are used as proxies for volatility. These two measures have been widely employed in the financial literature to measure volatility.[5] Plots of the absolute return series are displayed in Figure 5, while Figure 6 shows the squared returns.

[Insert Tables 2 and 3 about here]

Tables 2 and 3 report the estimates of d for the absolute and squared returns respectively under the assumption that the error term is white noise. Very similar results were obtained imposing weakly autocorrelated errors. The estimates are significantly positive in all cases, the values ranging between 0.142 ("Last" with an intercept) and 0.162 ("High" with no regressors) in case of the absolute returns, and between 0.096 ("Low" with a linear trend) and 0.109 ("Last" with an intercept) for the squared returns.

The results presented so far are consistent with those reported in the literature for lower frequency data, that is, the exchange rates appear to be I(1) implying that returns are I(0), and the associated volatility is I(d) with a positive and small value of d.

In the context of high frequency data, it is interesting to investigate if the same result holds as the distance between observations increases. For this purpose we examine again the long memory property of the same variables but now using series which are collected every 2, 3, 5 and 10 minutes respectively.

[Insert Table 4 about here]

Table 4 displays the results using these lower frequencies. Starting with data collected every 2 minutes (see Table 4a), it can be seen that the unit root null is almost

---

[5] Absolute returns were employed among others by Ding et al. (1993), Granger and Ding (1996), Bollerslev and Wright (2000), Gil-Alana (2005), Cavalcante and Assaf (2004), Sibbertsen (2004) and Cotter (2005), whereas squared returns were used in Lobato and Savin (1998), Gil-Alana (2003), Cavalcante and Assaf (2004) and Cotter (2005).

never rejected. The only two exceptions are "Low" with an intercept, and with an intercept and a linear trend, where d is strictly above 1. Focusing now on the data collected every 3 minutes (Table 4b), it can be seen that the estimated values of d are slightly smaller, and the unit root null hypothesis is never rejected. In general the estimates of d are smaller by about 0.020 compared with those reported in Table 4a. Table 4c concerns the data collected every 5 minutes. Once more the values are smaller than in previous tables, and the same happens in Table 4d which concerns data collected every 10 minutes. In this case, even values which are strictly smaller than 1 are found, implying a small degree of mean-reverting behaviour.

[Insert Tables 5 and 6 about here]

Tables 5 and 6 display the estimates of d for the absolute and squared returns series respectively, again assuming white noise errors. The results here are slightly more ambiguous as there is no monotonic decrease in the value of d as the time distance between the observations increases. For example, in the two cases of absolute and squared returns the highest values for "Last" occur for data collected every 3 minutes, and for "High" and "Low" for data collected every 5 minutes. This lack of a relationship between data frequency and the order of integration in the volatility processes is also found in the case of autocorrelated errors.

Finally, we employ a semiparametric method to estimate the values of d in our series of interest. Therefore, no functional form is assumed for the error term. We employ here a procedure developed by Robinson (1995). This method is essentially a local 'Whittle estimator' in the frequency domain, which uses a band of frequencies that degenerates to zero. The estimator is implicitly defined by:

$$\hat{d} = \arg\min_{d} \left(\log \overline{C(d)} - 2d \frac{1}{m} \sum_{s=1}^{m} \log \lambda_s \right), \qquad (3)$$

$$\overline{C(d)} = \frac{1}{m} \sum_{s=1}^{m} I(\lambda_s) \lambda_s^{2d}, \qquad \lambda_s = \frac{2\pi s}{T}, \qquad \frac{m}{T} \to 0,$$

where $I(\lambda_s)$ is the periodogram of the raw time series, x_t, given by:

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

and $d \in (-0.5, 0.5)$. Under finiteness of the fourth moment and other mild conditions, Robinson (1995) proved that:

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

where d_o is the true value of d.[6]

[Insert Figures 7 – 10 about here]

The results based on the above approach are displayed in Figures 7 – 10. It can be seen that the values are similar for the four series. Along with the estimates we also present the 95% confidence band corresponding to the I(1) hypothesis. We display the estimates for a range of values of the bandwidth parameter m. The highest estimates correspond to the highest frequency, while the lowest ones correspond to the series with data collected every 10 minutes.

4.   Conclusions

Despite the existence of a very extensive literature, there is still lack of consensus on what is the most appropriate model specification for many financial series. For instance,

---

[6] This method has been further examined and refined by Velasco (1999), Velasco and Robinson (2000), Phillips and Shimotsu (2004, 2005) and others. However, such refined methods require additional user-

whether asset returns of asset prices are predictable or not is still controversial: the efficiency market hypothesis suggests that they should follow a random walk (see Fama, 1970), but mean reversion is often found (see, e.g., Poterba and Summers, 1988). More recently, it has become clear that it is essential to consider the possibility of fractional integration in order to analyse the long-memory properties and to allow for a much richer dynamic specification. Various models have been suggested, increasingly general (see, e.g., Caporale and Gil-Alana, 2002, 2007, 2008). However, a potentially crucial issue which has been overlooked is the extent to which the fractional differencing parameter might be sensitive to the data frequency. This has been analysed in the present paper by using high frequency data on the British pound-US dollar spot exchange rate. In particular, we examined intra-day data (collected every 1, 2, 3, 5 and 10 minutes) of the open, close, high and low values of the exchange rate. In brief, we find evidence that a lower degree of integration is associated with lower data frequencies. In particular, when the data are collected every 10 minutes there are several cases with values of d strictly smaller than 1, implying a certain degree of mean-reverting behaviour. This holds for all the four series examined.

It might be asked whether the lower degrees of dependence estimated for the lower frequencies is the result of small sample bias. However, it should be noted that even at the lowest data frequencies the sample size is large enough to justify the estimation of a fractional integration model. Other approaches could be applied to these and other high frequency data such as the one suggested by Ohanissian et al. (2008) in their study on fractional integration, structural breaks and data frequency. Moreover, the presence of other issues in high frequency financial data such as non-linear structures (Dunis and

---

chosen parameters, and the estimates of d may be very sensitive to the choice of these parameters. In this

Zhou, 1998), turbulent cascades (Ghashghaie et al., 1996) and various other market microstructure issues will be examined in future papers.

respect, the method of Robinson (1995) seems computationally simpler.

References

Andersen, T.G. (2000) Some reflections on analysis of high frequency data, Journal of Business and Economic Statistics 18, 146-153.

Backus, D. and S. Zin, 1993, Long memory inflation uncertainty: Evidence of term structure of interest rate, Journal of Money, Credit and Banking 25, 687-700.

Baillie, R.T., C-F. Chung and M.A. Tieslau, 1996, Analysing inflation by the fractionally integrated ARFIMA-GARCH model, Journal of Applied Econometrics 11, 23-40.

Baum, C.F., J. Barkoulas and M. Caglayan (1999) Persistence in the international inflation rates. Southern Economic Journal 65, 900-913.

Bauwens, L. and P. Giot (2001), Econometric Modelling of Stock Market Intraday Activity, Kluwer.

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

Bollerslev, T., Wright, J.H., 2000, High frequency data, frequency domain inference and volatility forecasting, Review of Economics and Statistics 83, 596-602.

Booth, G.G., F.R. Kaen and P.E. Koveos (1982) R/S analysis of foreign exchange markets under two international monetary regimes. Journal of Monetary Economics 10, 407-415.

Campbell, J.Y., A.W. Lo and A.C. MacKinlay (1997), The Econometrics of Financial Markets, Princeton University Press, New Jersey.

Caporale, G.M. and L.A. Gil-Alana, 2002, Fractional integration and mean reversion in stock prices, Quarterly Review of Economics and Finance 42, 599-609.

Caporale, G.M. and L.A. Gil-Alana, 2007, Long run and cyclical dynamics in the US stock market, CESifo Working Paper no. 2046.

Caporale, G.M. and L.A. Gil-Alana (2008), "Modelling long-run trends and cycles in financial time series data", CESifo Working Paper no. 2330.

Cavalcante, J., Assaf, A., 2004, Long range dependence in the returns and volatility of the Brazilian stock market, European Review of Economics and Finance 3, 5-22.

Cheung, Y.W. (1993) Long memory in foreign exchange rates. Journal of Business and Economic Statistics 11, 93-101.

Cotter, J., 2005, Uncovering long memory in high frequency UK futures, European Journal of Finance 11, 325-337.

Crato, N. and B.K. Ray (2000) Memory in returns and volatilities of future's contracts. Journal of Futures Markets 20, 525-543.

Dacarogna, M.M., R. Gencay, U.A. Muller, R.B. Olsen and O.V. Pictet (2001) An Introduction to High Frequency Finance, Academic Press.

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

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

Ding, Z., Granger, C.W.J., Engle, R.F., 1993, A long memory property of stock markets and a new model, Journal of Empirical Finance 1, 83-106.

Dunis, C. and B. Zhou (1998) Non-linear modelling of high frequency financial time series, John Wiley & Sons.

Engle, R.F. and J. Russell (2004) Analysis of High Frequency Financial Data, Handbook of Financial Econometrics.

Fama, E.F. (1970), "Efficient capital markets: a review of theory and empirical work", Journal of Finance, 25, 383-417.

Fama, E.F. and K.R. French, 1988, Permanent and transitory components of stock prices, Journal of Political Economy 96, 246-273.

Fang, H., K.S. Lai and M. Lai (1994) Fractal structure in currency futures price dynamics. Journal of Futures Markets 14, 169-181.

Ghashghaie, S., W. Breymann, J. Peinke, P. Talkner and Y. Dodge (1996) Turbulent cascades in foreign exchange markets, Nature 381, 767-770.

Ghysels, E. (2000), Some econometric recipes for high frequency data cooking, Journal of Business and Economic Statistics 18, 154-163.

Gil-Alana, Luis A. (2000), "Mean reversion in the real exchange rates", Economics Letters 69, 285-288.

Gil-Alana, L.A., 2003, Fractional integration in the volatility of asset returns, European Review of Economics and Finance 2, 41-52.

Gil-Alana, L.A., 2005, Long memory in daily absolute and squared returns in the Spanish stock market, Advances in Investment Analysis and Portfolio Management 1, 198-217.

Gil-Alana, L.A. and J. Hualde (2009) Fractional integration and cointegration. An overview with an empirical application. The Palgrave Handbook of Applied Econometrics, Volume 2. Edited by Terence C. Mills and Kerry Patterson, MacMillan Publishers, pp. 434-472.

Gil-Alana, L.A. and Robinson, P.M., 1997, Testing of unit roots and other nonstationary hypotheses in macroeconomic time series, Journal of Econometrics 80, 241-268.

Goodhart, C.A.E. and M. O`Hara (1997) High frequency data in financial markets. Issues and applications, Journal of Empirical Finance 4, 73-114.

Gourieroux, C. and J. Jasiak (2001) Financial Econometrics. Problems, models and methods, Princeton University Press, New Jersey.

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.

Granger, C.W.J., Ding, Z., 1996, Varieties of long memory models, Journal of Econometrics 73, 61-78.

Granger, C.W.J. and R. Joyeux, 1980, An introduction to long memory time series and fractionally differencing, Journal of Time Series Analysis 1, 15-29.

Hassler, U. and J. Wolters, 1995, Long memory in inflation rates. International evidence, Journal of Business and Economic Statistics 13, 37-45.

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

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

Lobato, I.N., Savin, N.E., 1998, Real and spurious long memory properties of stock market data, Journal of Business and Economic Statistics 16, 261-268.

Lyons, R. (2001), The Microstructure Approach to Exchange Rates, MIT Press.

Ohanissian, A. J.R. Russell and R.S. Tsay (2008), True or spurious long memory? A new test. *Journal of Business and Economic Statistics*, 26, 161-175.

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

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

Poterba, J. and L. Summers, 1988, Mean reversion in stock returns. Evidence and implications, Journal of Financial Economics 22, 27-60.

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., 1995, Gaussian semi-parametric estimation of long range dependence, Annals of Statistics 23, 1630-1661.

Shea, G., 1991, Uncertainty and implied variance bounds of long memory models of the interest rate term structure, Empirical Economics 16, 287-312.

Sibbertsen, P., 2004, Long memory in volatilities of German stock returns, Empirical Economics 29, 477-488.

Souza, L.R., (2005) A note on Chamber's long memory and aggregation in macroeconomic time series, International Economic Review 46, 3, 1059-1062.

Souza, L.R. and J. Smith (2002) Bias in the memory parameter for different sampling rates, International Journal of Forecasting 18, 489-522.

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

Summers, L.H., 1986, Does the stock market rationally reflect fundamental values?, Journal of Finance 41, 591-601.

Tsay, R.S. (2001) Analysis of Financial Time Series, John Wiley & Sons, Inc.

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

Velasco, C. and P.M. Robinson, 2000, Whitle pseudo maximum likelihood estimation for nonstationary time series, Journal of the American Statistical Association 95, 1229-1243.

Wang, C. (2004) Futures trading activity and predictable foreign exchange movements. Journal of Banking and Finance 28, 1023-1041.

Wood, R.A. (2000) Market microstructure research databases. History and projections, Journal of Business and Economic Statistics 18, 140-145.
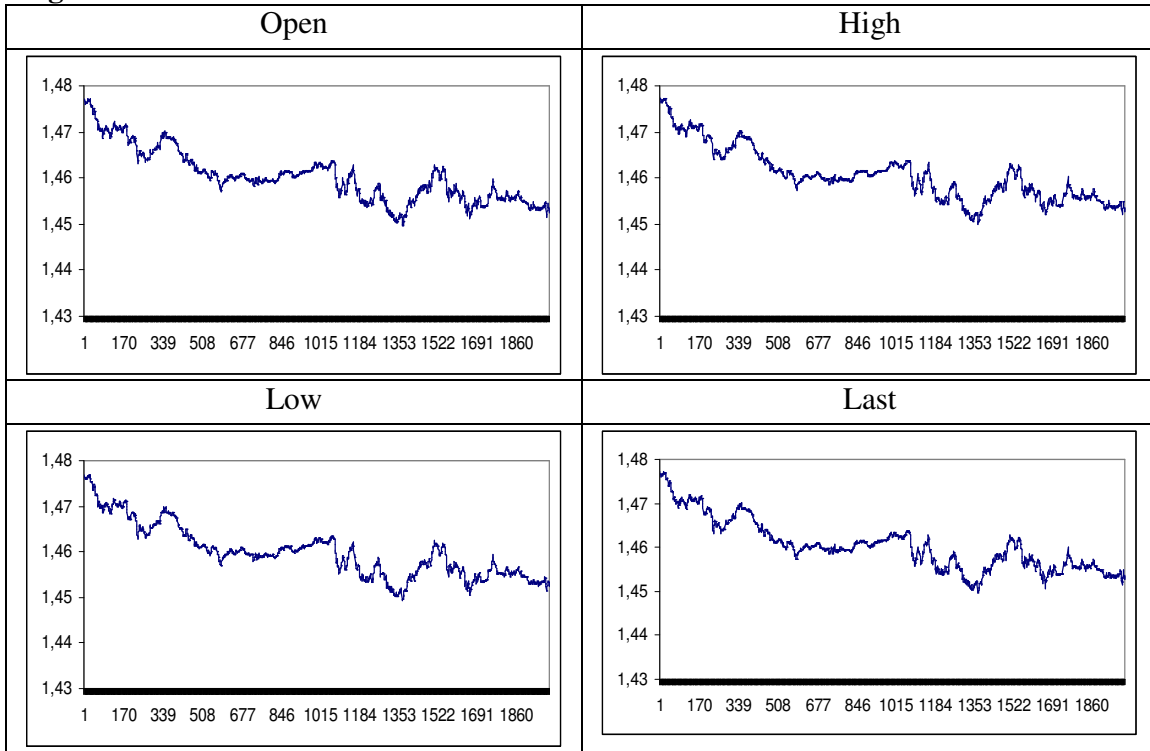
Figure 1: Series in levels

| Open | High |
| Low | Last |

Figure 2: Returns

| Open | High |
| Low | Last |
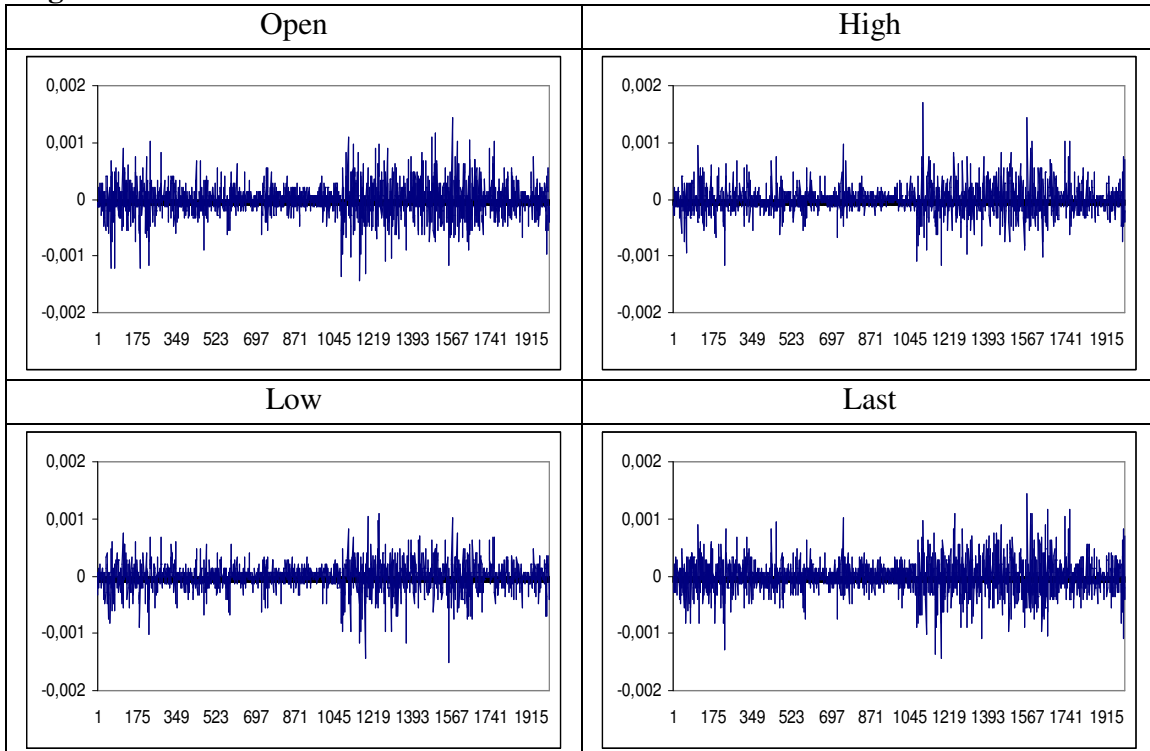
Figure 3: Correlograms of returns

Open	High

Low	Last
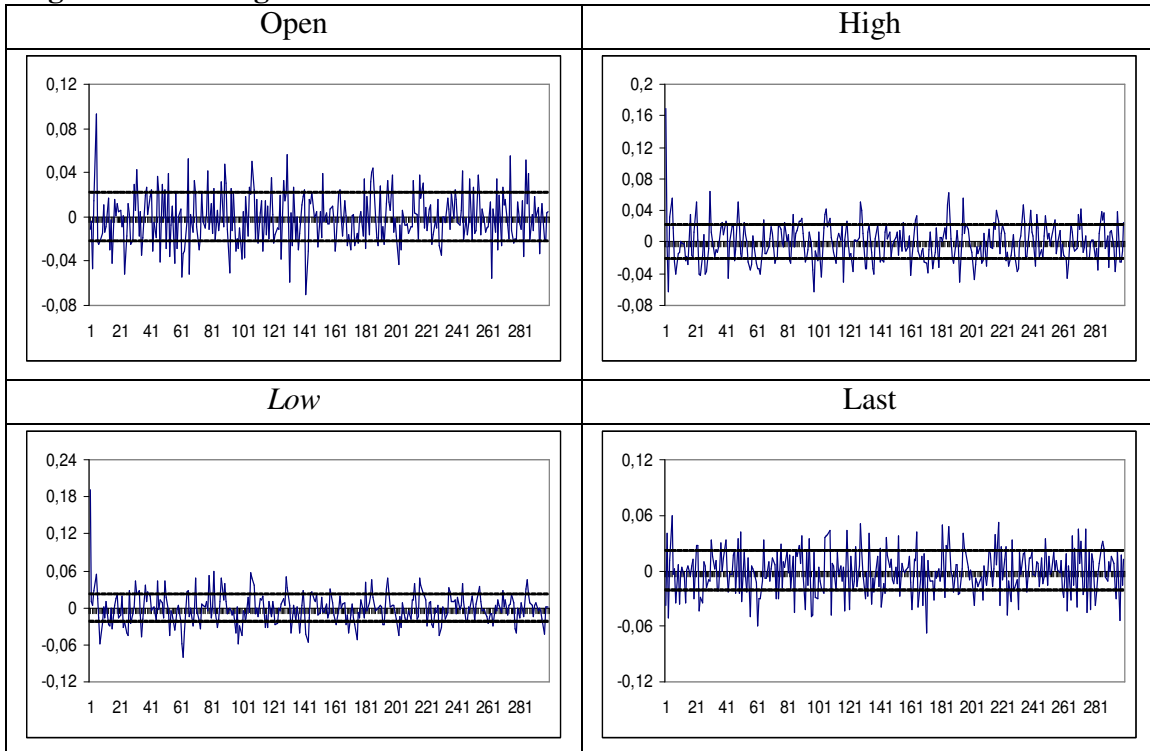
Figure 4: Periodograms of returns

Open	High

Low	Last
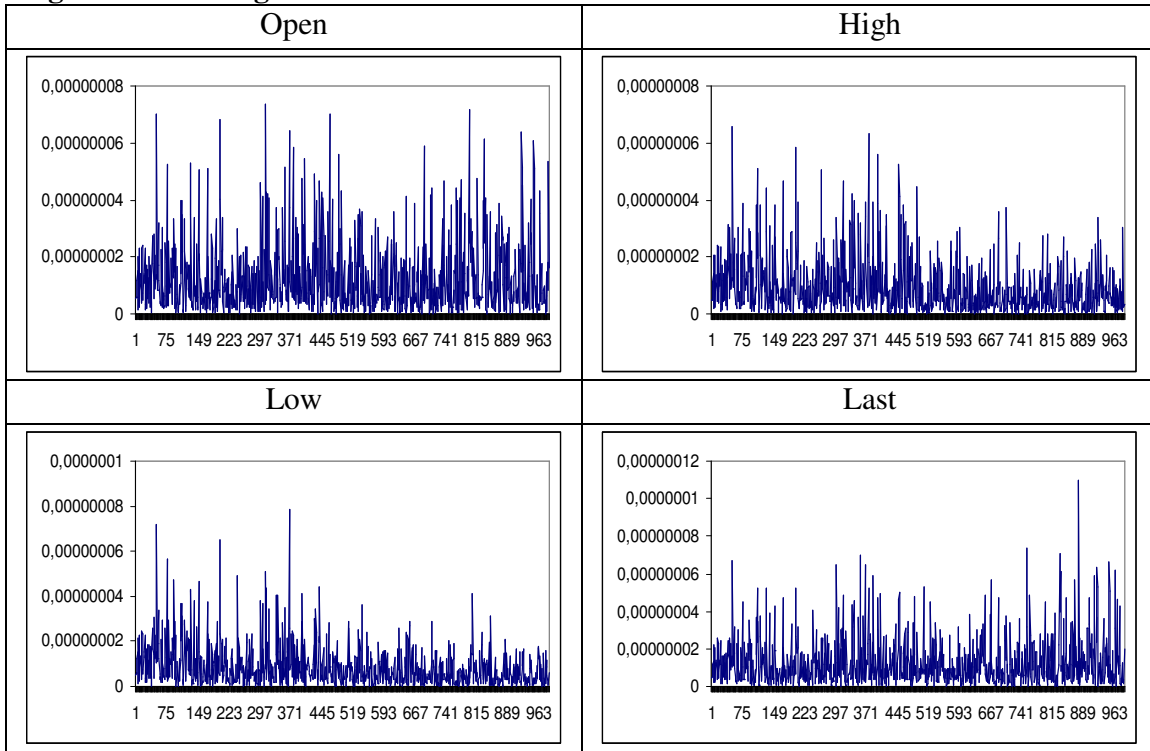
Table 1a: Estimates of d based on a model with white noise errors

	No regressors	An intercept	A linear time trend
Open	0.997 (0.970, 1.027)	0.983 (0.955, 1.015)	0.983 (0.956, 1.015)
High	0.998 (0.971, 1.028)	1.101 (1.066, 1.141)	1.101 (1.066, 1.141)
Low	0.997 (0.970, 1.027)	1.130 (1.095, 1.169)	1.130 (1.095, 1.169)
Last	0.998 (0.970, 1.028)	0.977 (0.950, 1.007)	0.977 (0.950, 1.007)

The values in parentheses give the 95% confidence band for the non-rejection values of d.

Table 1b: Estimates of d based on a model with AR(1) errors

	No regressors	An intercept	A linear time trend
Open	1.381 (1.328, 1.441)	0.973 (0.923, 1.031)	0.974 (0.924, 1.031)
High	1.382 (1.329, 1.442)	0.934 (0.879, 0.996)	0.936 (0.883, 0.996)
Low	1.381 (1.327, 1.440)	0.969 (0.907, 1.037)	0.970 (0.910, 1.037)
Last	1.382 (1.329, 1.442)	1.004 (0.954, 1.060)	1.004 (0.955, 1.060)

The values in parentheses give the 95% confidence band for the non-rejection values of d.

Table 1c: Estimates of d based on a model with Bloomfield errors

	No regressors	An intercept	A linear time trend
Open	0.997 (0.944, 1.041)	0.963 (0.922, 1.029)	0.970 (0.923, 1.029)
High	0.991 (0.950, 1.042)	0.962 (0.914, 1.006)	0.962 (0.915, 1.006)
Low	0.990 (0.951, 1.047)	0.988 (0.939, 1.047)	0.988 (0.940, 1.047)
Last	0.998 (0.950, 1.049)	1.010 (0.955, 1.057)	1.010 (0.955, 1.057)

The values in parentheses give the 95% confidence band for the non-rejection values of d.
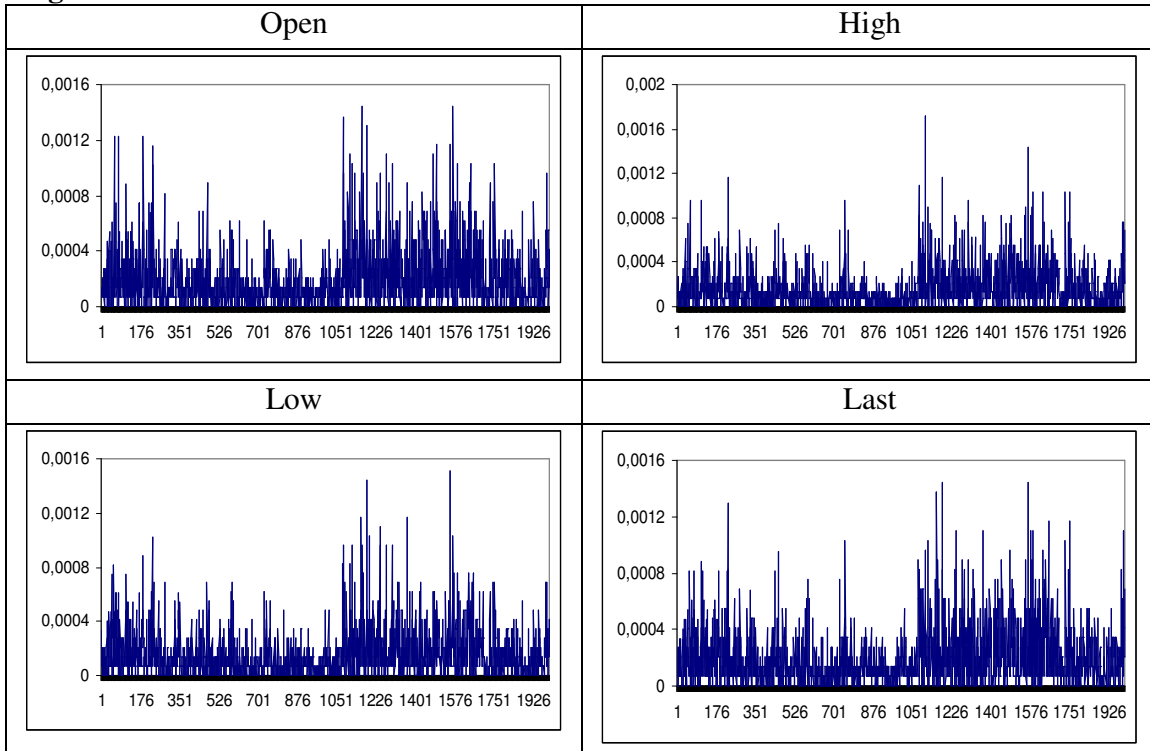
Figure 5: Absolute returns
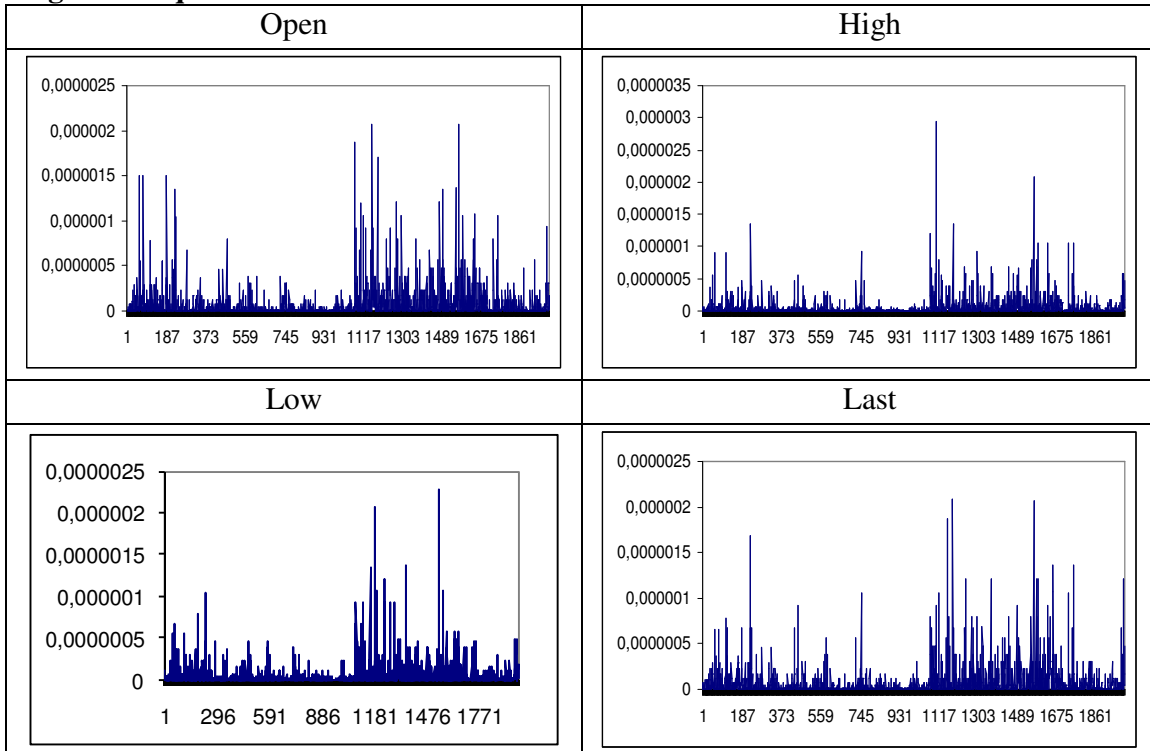

Figure 6: Squared returns

Table 2: Estimates of d for the absolute returns in a model with white noise errors

	No regressors	An intercept	A linear time trend
Open	0.149 (0.131, 0.171)	0.148 (0.130, 0.168)	0.144 (0.126, 0.165)
High	0.162 (0.142, 0.185)	0.159 (0.140, 0.181)	0.156 (0.136, 0.178)
Low	0.154 (0.134, 0.171)	0.151 (0.132, 0.172)	0.149 (0.129, 0.176)
Last	0.143 (0.123, 0.167)	0.142 (0.124, 0.163)	0.136 (0.117, 0.158)

Table 3: Estimates of d for the squared returns in a model with white noise errors

	No regressors	An intercept	A linear time trend
Open	0.106 (0.088, 0.126)	0.107 (0.089, 0.127)	0.103 (0.085, 0.124)
High	0.098 (0.078, 0.121)	0.099 (0.080, 0.122)	0.094 (0.074, 0.118)
Low	0.098 (0.079, 0.120)	0.099 (0.080, 0.121)	0.096 (0.077, 0.118)
Last	0.106 (0.088, 0.128)	0.109 (0.090, 0.130)	0.102 (0.082, 0.124)

The values in parentheses give the 95% confidence band for the non-rejection values of d.

Table 4a: Estimates of d based on a model with white noise errors (2 minutes)

	No regressors	An intercept	A linear time trend
Open	0.994 (0.956, 1.038)	0.980 (0.939, 1.028)	0.980 (0.940, 1.028)
High	0.994 (0.956, 1.038)	1.034 (0.989, 1.087)	1.034 (0.989, 1.087)
Low	0.995 (0.957, 1.039)	1.062 (1.017, 1.116)	1.062 (1.017, 1.115)
Last	0.994 (0.957, 1.039)	0.989 (0.948, 1.035)	0.989 (0.949, 1.035)

The values in parentheses give the 95% confidence band for the non-rejection values of d.

Table 4b: Estimates of d based on a model with white noise errors (3 minutes)

	No regressors	An intercept	A linear time trend
Open	0.992 (0.946, 1.047)	0.962 (0.912, 1.019)	0.963 (0.914, 1.019)
High	0.992 (0.946, 1.047)	1.003 (0.950, 1.066)	1.003 (0.951, 1.065)
Low	0.993 (0.947, 1.048)	1.041 (0.984, 1.108)	1.041 (0.985, 1.107)
Last	0.992 (0.946, 1.048)	0.958 (0.907, 1.016)	0.958 (0.910, 1.016)

The values in parentheses give the 95% confidence band for the non-rejection values of d.

Table 4c: Estimates of d based on a model with white noise errors (5 minutes)

	No regressors	An intercept	A linear time trend
Open	0.990 (0.930, 1.064)	0.941 (0.872, 1.024)	0.942 (0.877, 1.024)
High	0.990 (0.931, 1.064)	0.948 (0.880, 1.030)	0.949 (0.885, 1.030)
Low	0.990 (0.931, 1.064)	0.981 (0.910, 1.069)	0.982 (0.913, 1.068)
Last	0.989 (0.930, 1.063)	0.942 (0.874, 1.024)	0.944 (0.879, 1.023)

The values in parentheses give the 95% confidence band for the non-rejection values of d.

Table 4d: Estimates of d based on a model with white noise errors (10 minutes)

	No regressors	An intercept	A linear time trend
Open	0.977 (0.895, 1.088)	0.831 (0.719, 0.957)	0.848 (0.761, 0.961)
High	0.978 (0.895, 1.089)	0.869 (0.766, 0.990)	0.881 (0.794, 0.991)
Low	0.977 (0.895, 1.088)	0.860 (0.750, 0.987)	0.873 (0.784, 0.988)
Last	0.978 (0.895, 1.089)	0.861 (0.755, 0.983)	0.872 (0.785, 0.985)

The values in parentheses give the 95% confidence band for the non-rejection values of d.

Table 5a: Estimates of d for the absolute returns (2 minutes)

	No regressors	An intercept	A linear time trend
Open	0.188 (0.157, 0.225)	0.182 (0.153, 0.217)	0.179 (0.149, 0.215)
High	0.181 (0.152, 0.216)	0.179 (0.151, 0.211)	0.174 (0.146, 0.208)
Low	0.176 (0.148, 0.210)	0.171 (0.144, 0.202)	0.168 (0.141, 0.200)
Last	0.143 (0.116, 0.173)	0.140 (0.116, 0.169)	0.136 (0.111, 0.166)

The values in parentheses give the 95% confidence band for the non-rejection values of d.

Table 5b: Estimates of d for the absolute returns (3 minutes)

	No regressors	An intercept	A linear time trend
Open	0.159 (0.124, 0.202)	0.157 (0.124, 0.197)	0.151 (0.116, 0.192)
High	0.178 (0.143, 0.221)	0.176 (0.143, 0.216)	0.171 (0.136, 0.212)
Low	0.165 (0.127, 0.212)	0.159 (0.124, 0.202)	0.156 (0.120, 0.200)
Last	0.167 (0.131, 0.211)	0.168 (0.135, 0.210)	0.160 (0.124, 0.204)

The values in parentheses give the 95% confidence band for the non-rejection values of d.

Table 5c: Estimates of d for the absolute returns (5 minutes)

	No regressors	An intercept	A linear time trend
Open	0.189 (0.140, 0.247)	0.176 (0.133, 0.231)	0.175 (0.131, 0.230)
High	0.194 (0.144, 0.259)	0.190 (0.144, 0.249)	0.186 (0.138, 0.247)
Low	0.216 (0.165, 0.281)	0.203 (0.157, 0.262)	0.202 (0.155, 0.261)
Last	0.149 (0.106, 0.204)	0.150 (0.109, 0.202)	0.144 (0.102, 0.198)

The values in parentheses give the 95% confidence band for the non-rejection values of d.

Table 5d: Estimates of d for the absolute returns (10 minutes)

	No regressors	An intercept	A linear time trend
Open	0.143 (0.077, 0.233)	0.136 (0.076, 0.213)	0.133 (0.071, 0.212)
High	0.088 (0.028, 0.174)	0.086 (0.028, 0.161)	0.084 (0.025, 0.160)
Low	0.133 (0.058, 0.233)	0.121 (0.055, 0.207)	0.121 (0.054, 0.207)
Last	0.088 (0.024, 0.176)	0.083 (0.024, 0.156)	0.083 (0.024, 0.160)

The values in parentheses give the 95% confidence band for the non-rejection values of d.

Table 6a: Estimates of d for the squared returns (2 minutes)

	No regressors	An intercept	A linear time trend
Open	0.129 (0.096, 0.164)	0.130 (0.101, 0.164)	0.126 (0.100, 0.162)
High	0.115 (0.087, 0.149)	0.118 (0.090, 0.151)	0.110 (0.080, 0.145)
Low	0.113 (0.086, 0.145)	0.115 (0.088, 0.146)	0.111 (0.083, 0.143)
Last	0.097 (0.072, 0.127)	0.100 (0.075, 0.129)	0.093 (0.067, 0.124)

The values in parentheses give the 95% confidence band for the non-rejection values of d.

Table 6b: Estimates of d for the squared returns (3 minutes)

|	No regressors	An intercept	A linear time trend
Open	0.106 (0.073, 0.148)	0.108 (0.075, 0.149)	0.103 (0.068, 0.145)
High	0.123 (0.090, 0.165)	0.125 (0.092, 0.166)	0.119 (0.084, 0.161)
Low	0.128 (0.089, 0.176)	0.129 (0.091, 0.176)	0.126 (0.087, 0.174)
Last	0.126 (0.090, 0.171)	0.130 (0.094, 0.174)	0.123 (0.084, 0.169)

The values in parentheses give the 95% confidence band for the non-rejection values of d.

Table 6c: Estimates of d for the squared returns (5 minutes)

	No regressors	An intercept	A linear time trend
Open	0.136 (0.088, 0.198)	0.135 (0.089, 0.195)	0.133 (0.085, 0.194)
High	0.144 (0.094, 0.210)	0.147 (0.097, 0.211)	0.141 (0.089, 0.208)
Low	0.160 (0.111, 0.224)	0.159 (0.111, 0.221)	0.157 (0.108, 0.220)
Last	0.110 (0.067, 0.166)	0.113 (0.070, 0.169)	0.107 (0.061, 0.164)

The values in parentheses give the 95% confidence band for the non-rejection values of d.

Table 6d: Estimates of d for the squared returns (10 minutes)

|	No regressors	An intercept	A linear time trend
Open	0.088 (0.027, 0.171)	0.090 (0.028, 0.171)	0.085 (0.021, 0.168)
High	0.044 (-0.015, 0.125)	0.045 (-0.016, 0.127)	0.041 (-0.023, 0.125)
Low	0.083 (0.019, 0.177)	0.084 (0.020, 0.169)	0.082 (0.016, 0.168)
Last	0.049 (-0.011, 0.132)	0.051 (-0.011, 0.133)	0.049 (-0.015, 0.132)

The values in parentheses give the 95% confidence band for the non-rejection values of d.

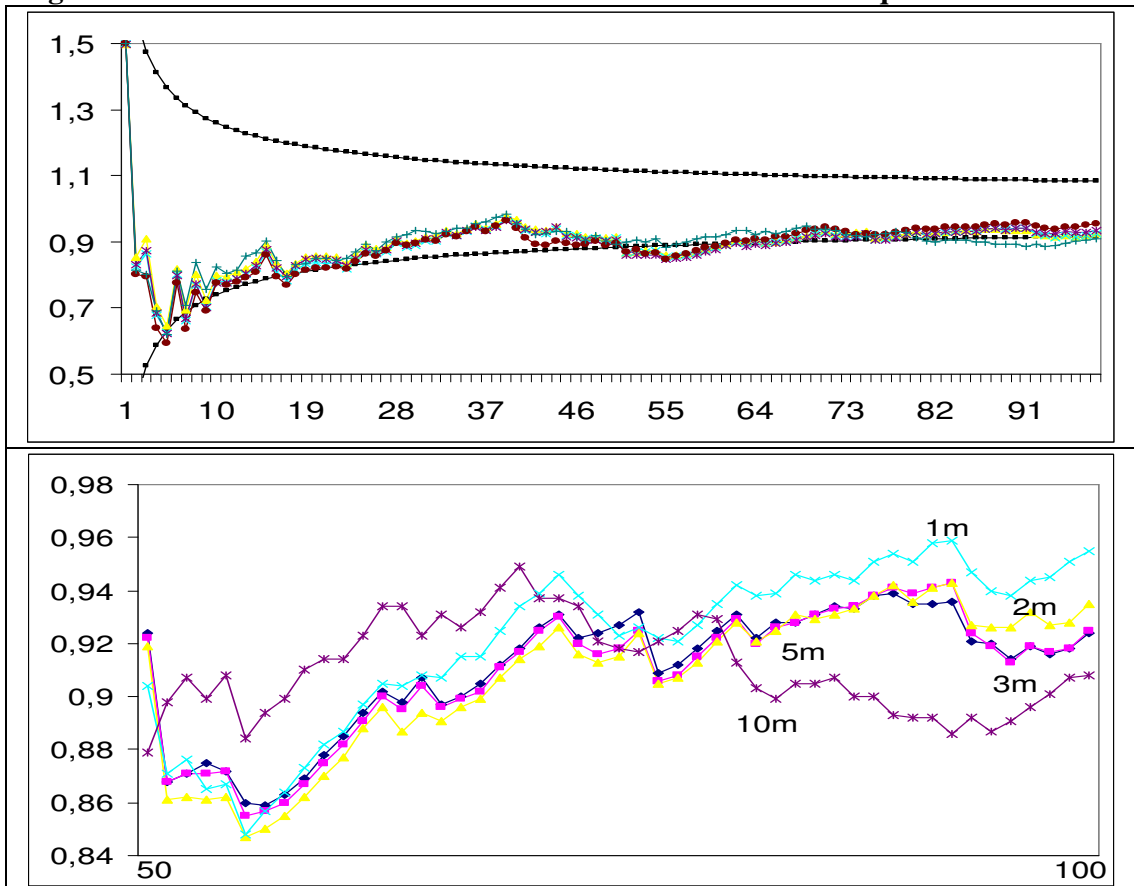**Figure 7: Estimates of d for the OPEN series at different data frequencies**
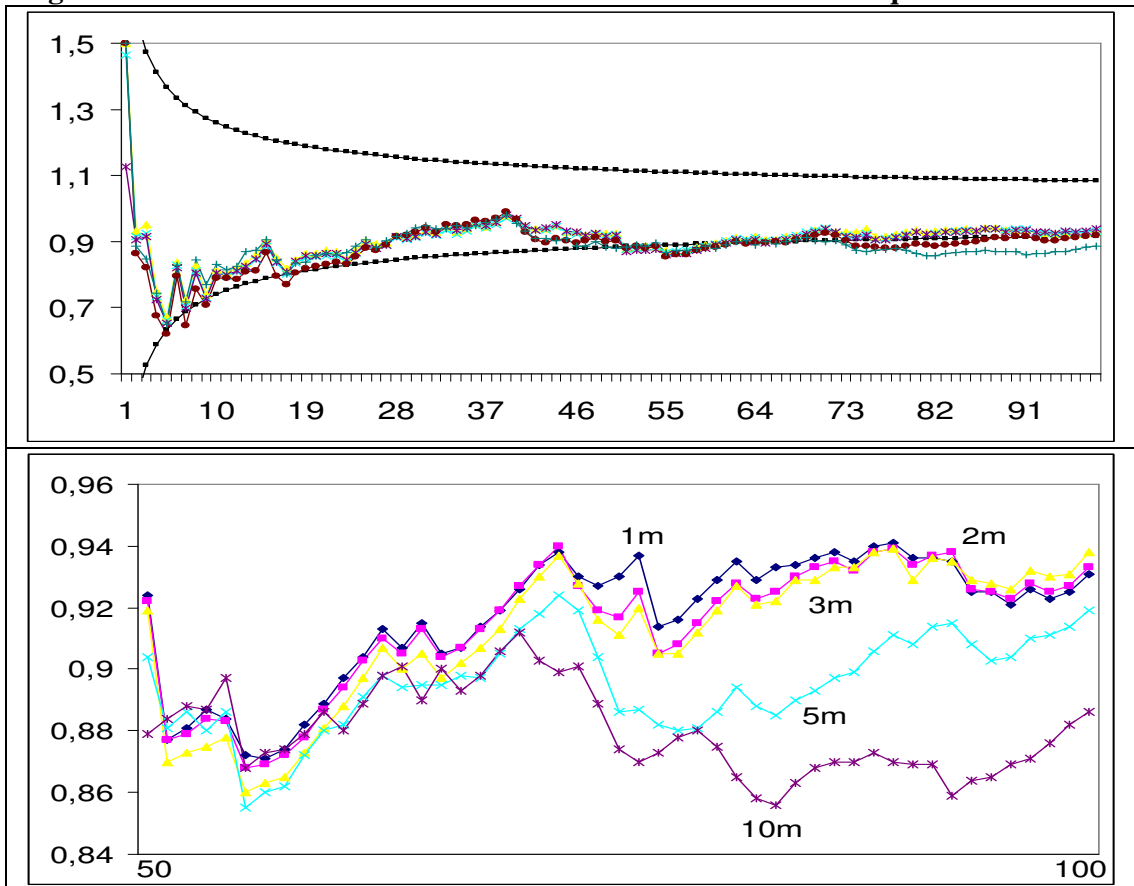
The horizontal axis concerns the bandwidth parameter while the vertical one refers to the estimated value of d.

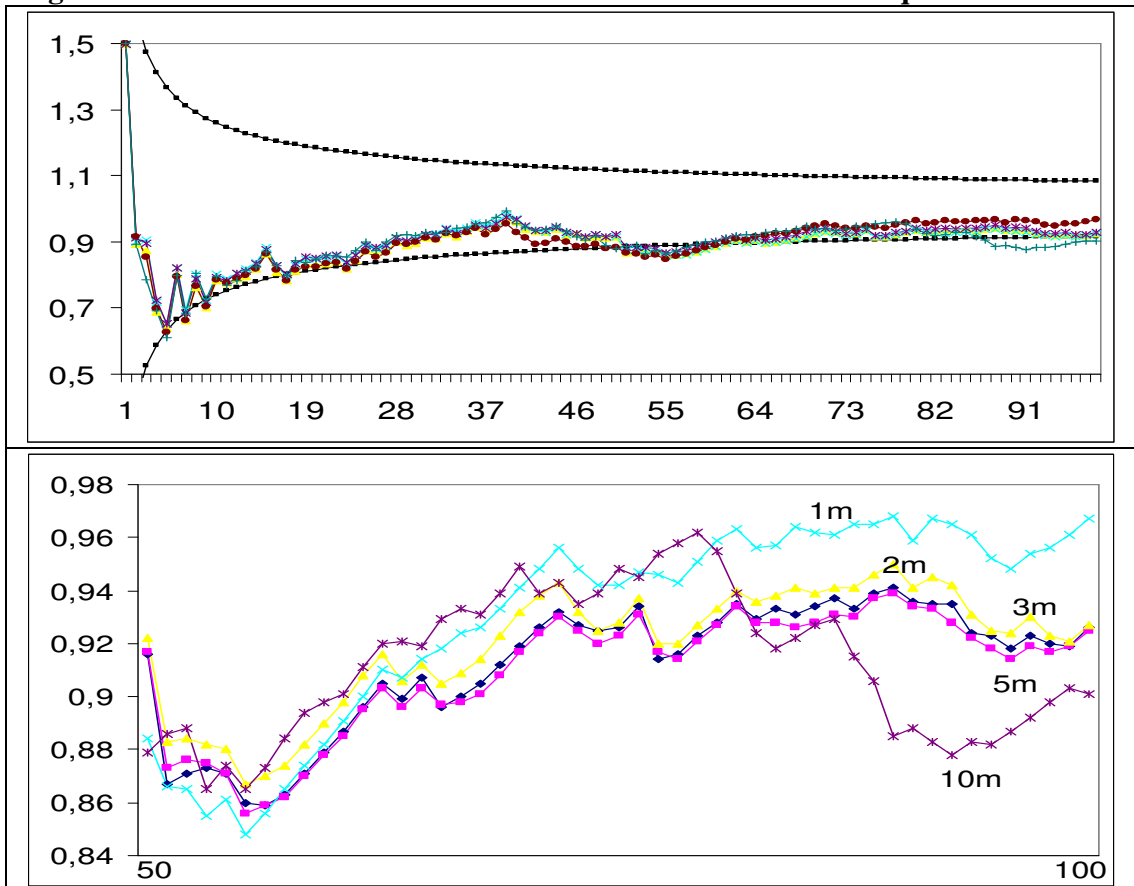**Figure 8: Estimates of d for the HIGH series at different data frequencies**
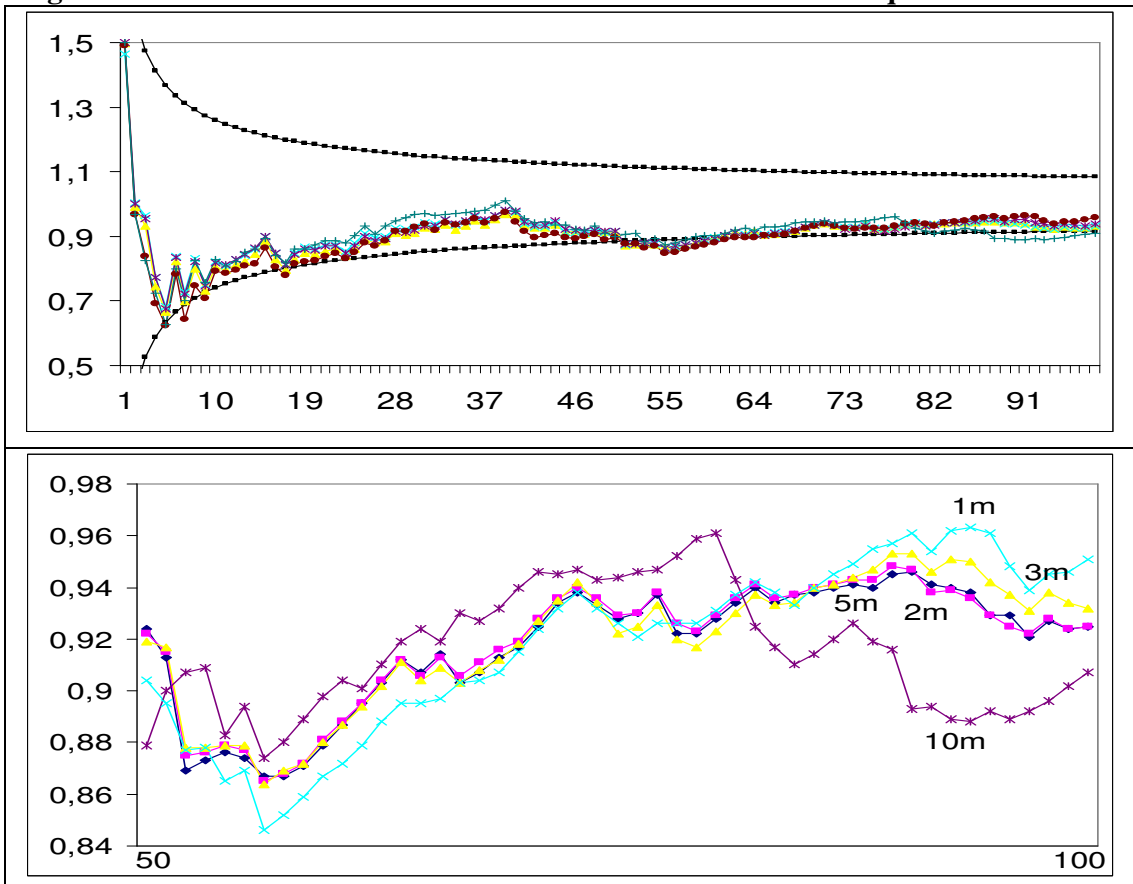
The horizontal axis concerns the bandwidth parameter while the vertical one refers to the estimated value of d.

Figure 9: Estimates of d for the LOW series at different data frequencies

The horizontal axis concerns the bandwidth parameter while the vertical one refers to the estimated value of d.

Figure 10: Estimates of d for the LAST series at different data frequencies

The horizontal axis concerns the bandwidth parameter while the vertical one refers to the estimated value of d.