INVESTIGACIONES ECONOMICAS, vol. XXI(2), 1997, 273-295

SEMIPARAMETRIC ESTIMATION OF SEASONAL LONG MEMORY MODELS: THEORY AND AN APPLICATION TO THE MODELING OF EXCHANGE RATES

IGNACIO N. LOBATO
University of Iowa

Several studies have reported evidence of long memory in the squared returns of financial series. This evidence has been usually interpreted as an indicator of persistence in the volatility process of the returns. In this study we examine how that evidence can be explained by a seasonal long memory component in the returns. Also, we review semiparametric estimates of long memory models in the frequency domain and extend Robinson's (1995a) pseudo maximum likelihood approach to the seasonal long memory case. Finally, using these tools we analyze the British pound-Deutsche mark exchange rate series.(*JEL* C14, C22, G15)

1. Introduction

The notion of long memory has been increasingly used in applied financial analysis during the last years. Although exchange rate and stock returns rarely exhibit long memory, their squares or absolute values frequently do (Baillie *et al.*, 1996; Bollerslev and Mikkelsen, 1996; Ding *et al.*, 1993).This fact has usually been associated with persistence in the volatility process of the returns. In this study we show that this does not need to be the case: in fact, a seasonal long memory component in a process may cause the presence of long memory in its squares.

In Section 2 we introduce the concept of long memory. In Section 3 we consider semiparametric estimates for the long memory parameters

I thank the editor and two referees for valuable comments.

and, in particular, extend Robinson's (1995a) estimation procedure to seasonal long memory. In Section 4 we provide a Lemma that establishes that evidence of long memory in the squares of a series may be induced by a seasonal long memory component in the level of that series. In Section 5 we analyze an exchange rate series that illustrates this phenomenon. Section 6 concludes.

2. Long memory processes

Some stationary economic processes exhibit a high degree of persistence. Intuitively, by this we mean that the long run component of the process is relatively more important than the short run component. If we consider covariance stationary processes it is reasonable to characterize persistence by looking at the behavior of second-order moments. Persistence would be reflected both in the spectral density function (with a relatively high concentration of mass around zero frequency) and in the autocovariance function (with relatively high concentration of mass at high lags). This is the idea behind the concept of long memory. Notice that long memory is based on second-order moments, i.e., on autocovariances, or alternatively, on the spectral density function (sdf, hereafter). Alternatively, persistence could be characterized using broader concepts of strong dependence which could be based on higher-order moments, higher-order spectra or in the cumulative joint distribution function.

Although autoregressive moving average (ARMA) models have been frequently employed in modelling economic series, they are not suitable to characterize long memory processes. This is due to the fact that ARMA models are only appropriate for processes whose autocovariances, eventually, decay exponentially (in an ARMA model the eventual behavior of the autocovariances is determined by a stable finite order difference equation, Yule-Walker, whose solution is a linear combination of exponential terms). Exponential decay is a very fast decay, so ARMA models cannot be used for processes that exhibit persistence as defined above. The models designed for this purpose are called long memory models.

Long memory processes are characterized by a slow decay of the autocovariances, which even though they may be rather small, their aggregate effect is not-negligible. Formally, let x_t be a covariance stationary process and let γ_j denote the autocovariance at lag j, of x_t, $\gamma_j = E[(x_t - \mu)(x_{t+j} - \mu)]$, where μ denotes the mean of x_t. A strictly long memory process is characterized by

$$\sum_{j=-\infty}^{\infty} |\gamma_j| = \infty . \qquad [1]$$

This definition, although appealing from a theoretical point of view, is too general to work with. This is why the literature has considered two alternative definitions.

Long memory can be defined in the frequency or in the time domains. In the frequency domain long memory is defined as follows. Assume that the sdf of x_t is defined and denote it by $f(\lambda)$, then x_t exhibits long memory if

$$f(\lambda) \sim C\lambda^{-2d_0} \text{ as } \lambda \rightarrow 0^+ , \qquad [2]$$

where $\sim$ means that the ratio of both sides tends to 1 as $\lambda \rightarrow 0^+$, C is a positive constant and $d_0 < 1/2$, $d_0 \neq 0$. First, note that the sdf is defined only in a neighborhood of zero frequency. This means that this definition is concerned with the long term properties and nothing is assumed about short run dynamics. Also, note that [2] includes two different cases. First, when $d_0 \in (0, 1/2)$, the sdf tends to infinity as it is evaluated at frequencies that tend to zero. Strictly speaking, this is the long memory case. Second, when $d_0 < 0$, the sdf tends to zero as it is evaluated at frequencies that tend to zero. In this case the time series process has been called antipersistent (Mandelbrot, 1969).

Long memory can be defined in the time domain as follows. The process x_t exhibits long memory if

$$\gamma_j \sim Kj^{2d_0 - 1} \text{ as } j \rightarrow \infty , \qquad [3]$$

where K is a constant and d_0 takes the same values as above. Note that only long lag autocovariances are considered. Again, this means that this definition is concerned with the long term and nothing is assumed about the short term properties of the process. As it happens with [2], [3] includes two different cases. First, when $d_0 \in (0, 1/2)$ [1] holds; second, when $d_0 < 0$

$$\sum_{j=-\infty}^{\infty} |\gamma_j| < \infty . \qquad [4]$$

Fractional autoregressive integrated moving average (ARFIMA) and fractional Gaussian noise are the two most employed and analyzed examples in the literature (Hosking, 1981; Granger and Joyeux, 1980; Mandelbrot and Van Ness, 1968). For these models both conditions [2] and [3] hold.

In this paper we will focus on the long memory case in the strict sense, that is, $d_0 \in (0, 1/2)$. For this case, consider generalizing definitions [2] and [3] to:

$$f(\lambda) \sim C|\lambda_* - \lambda|^{-2d_s} \text{ as } \lambda \to \lambda_* \; , \qquad [5]$$

$$\gamma_j \sim K j^{2d_s-1} \cos(j\lambda_*) \text{ as } j \to \infty \; , \qquad [6]$$

where $\lambda_* \in (0,\pi)$ and $d_s \in (0, 1/2)$. Notice that now we denote the true value by d_s instead of d_0. When $\lambda_* = 0$, [2] and [3] are particular cases of [5] and [6] respectively. Expression [5] establishes that the sdf has the singularity at frequency λ_* rather than at the zero frequency. Equation [3] implies that all long lag autocovariances are positive (if $K > 0$) or negative (if $K < 0$). Equation [6] allows their sign to change. Expressions [5] and [6] characterize seasonal long memory. Notice that while condition [3] rules out seasonal long memory ([6] cannot hold for $\lambda_* \neq 0$ if [3] is true), conditions [2] and [5] for $\lambda_* \neq 0$ are compatible, that is, a process could exhibit at the same time long memory and seasonal long memory. In the same way as ARFIMA processes follow definitions [2] and [3], generalized fractional processes (Gray *et al.,* 1989) follow definitions [5] and [6].

3. Semiparametric estimates

The definitions of long memory in the frequency domain only specify the behavior of the sdf close to the zero frequency in the case of [2], or close to λ_* in the case of [5]. Similarly, in the time domain, the definitions only specify the behavior of the autocovariances at long lags in [3] and [6]. Hence, the relevant estimation procedures are those that employ sample information only in a neighborhood of the zero frequency, or λ_*, or at long lags. All these are called semiparametric procedures (Robinson, 1994a; Robinson, 1994c). The main feature of the semiparametric approach is to employ a bandwidth number m. In the frequency domain, in the case of [2], this number reflects the highest frequency $\lambda_m = 2\pi m/n$, where n is the sample size, at which the statistics used to estimate the sdf are evaluated. In the time domain, the bandwidth number reflects the lowest sample autocovariance employed (see Robinson, 1994c). In order to develop the asymptotic theory, m has to tend to infinity as n tends to infinity, but in such a way that their ratio tends to zero,

$$\tfrac{1}{m} + \tfrac{m}{n} \to 0 \text{ as } n \to \infty. \qquad [7]$$

The first part of [7] is necessary in order to have an accumulation of information as the sample size grows. The second part of [7] is needed so that the considered statistics refer only to the long term properties of the process.

These semiparametric procedures provide robust estimates of d_0 and d_s. In case of d_0 this means that consistency can be ensured under any short term behavior of the process. Notice that this property has an interesting implication when high frequency financial data are used. This type of data is affected by the existence of microstructure effects (infrequent trading effect or bid/ask price effect) that are short-run effects and have been handled in the Finance literature by filtering the data with a low order stationary ARMA process (Stoll and Whaley, 1990). If the main interest is inference about d_0 this filtering does not need to be performed because it would not affect in any extent the asymptotic properties of the semiparametric estimates (although it can have an important effect in small samples).

Semiparametric estimates achieve robustness at the expense of a loss in efficiency. That is, they are typically root-m instead of root-n consistent and therefore they are asymptotically inefficient compared with parametric estimates of properly specified models. Therefore, the importance of semiparametric estimation is based on the fragility of parametric estimates: a slight misspecification of the model would lead to inconsistency of the parametric estimates of all the coefficients and, in particular, of the long memory parameters. Giving more importance to robustness than to efficiency is the typical feature of nonparametric and semiparametric procedures and it means that these procedures are at their best for relatively large sample sizes.

Semiparametric procedures in the time domain have been considered in Delgado and Robinson (1994). In this study we focus on the frequency domain. There are several semiparametric procedures in the frequency domain, but we just consider the three procedures whose properties are better known. The most applied one in the literature has been the log periodogram estimate (LPE) proposed by Geweke and Porter-Hudak (1983) and modified by Robinson (1995b). This estimate is easily computed (it just involves ordinary least squares), but it has two main drawbacks. In order to derive a Gaussian asymptotic distribution, Robinson needed to assume Gaussianity (which is very restrictive and probably not valid for most of financial series) and, furthermore, he needed to introduce an additional user-chosen number to trim out frequencies very close to zero.

The second semiparametric estimate in the frequency domain (and the first one whose consistency was rigorously proved) is the averaged periodogram estimate (APE) (see Robinson, 1994a). This estimate is also very easy to obtain but, as the LPE, it has two disadvantages. First, it requires an additional user-chosen number. Second, the asymptotic distribution is quite complicated. Since it is normal if $d_0 \in (0, 1/4)$, but it is nonnormal if $d_0 \in (1/4, 1/2)$, statistical inference with this estimate is quite problematic (see Lobato and Robinson, 1996).

The third procedure is a pseudo maximum likelihood estimate (PMLE) analyzed by Robinson (1995a). It is more difficult to obtain than the

APE or the LPE, but it has important advantages. First, it only depends on the bandwidth m, and not on additional user-chosen numbers. Second, in order to derive an asymptotic normal distribution for this estimate, Gaussianity does not need to be assumed. These two features make this estimate a valuable statistical inference tool. Furthermore, this estimate appears to be the most efficient semiparametric procedure developed so far.

PMLE is based on estimating d_0 by minimizing the criterion function (d denotes any admissible value)

$$R(d) = \log \hat{C}(d) - 2d \sum_{j=1}^{m} \log \lambda_j \quad , \quad \hat{C}(d) = \tfrac{1}{m} \sum_{j=1}^{m} \lambda_j^{2d} I(\lambda_j), \qquad [8]$$

over the compact set $d \in [\Delta_1, \Delta_2]$ with $0 < \Delta_1 < \Delta_2 < 1/2$; $\lambda_j = 2\pi j/n$, $j = 1, 2 \ldots$, are the Fourier frequencies and $I(\lambda)$ the periodogram at frequency λ

$$I(\lambda) = \frac{1}{2\pi n} \left| \sum_{t=1}^{n} x_t e^{-it\lambda} \right|^2 .$$

Denote this estimate by $\hat{d}_0$. Under a condition which is slightly stronger than [7] and additional (but rather mild) assumptions which do not include Gaussianity, Robinson established the following asymptotic distribution

$$\sqrt{m}(\hat{d}_0 - d_0) \to_d N(0, 1/4). \qquad [9]$$

Robinson considered the estimation of [2], but the same approach can be extended to [5]. In the rest of the section we develop this procedure. Instead of focusing on a neighborhood of zero frequency, we concentrate on a neighborhood of λ_*. For simplicity we consider m to be even. Based on the Whittle approximation to the Gaussian likelihood and assuming that the sdf behaves as [5] in a neighborhood of λ_*, we consider the following criterion function

$$Q_*(C, d) = \frac{1}{m} \sum_{S'} \left\{ \log \left[C|\lambda_l - \lambda_j|^{-2d} \right] + \frac{I(\lambda_j)}{C|\lambda_l - \lambda_j|^{-2d}} \right\},$$

where $\lambda_l = 2\pi l/n$ and l denotes the integer part of $n\lambda_*/2\pi$. The set $S' = \{j : j \in [l - m/2, l + m/2], j \neq 0\}$ if λ_* is a Fourier frequency and $S' = \{j : j \in (l - m/2, l + m/2]\}$ if not. Then

$$Q_*(C, d) = \frac{1}{m} \sum_S \left\{ \log \left[C|\lambda_j|^{-2d} \right] + \frac{I(\lambda_l + \lambda_j)}{C|\lambda_j|^{-2d}} \right\}$$

where $S = \{j : j \in [-m/2, m/2], j \neq 0\}$ if λ_* is a Fourier frequency and $S = \{j : j \in (-m/2, m/2]\}$ if not. Setting its partial derivative with respect to C equal to zero

$$\frac{\partial Q_*(C, d)}{\partial C} = \frac{1}{m} \sum_S \left\{ \frac{1}{C} - \frac{I(\lambda_l + \lambda_j)}{C^2|\lambda_j|^{-2d}} \right\} = 0$$

$$\Rightarrow \quad \hat{C} = \hat{C}_*(d) = \frac{1}{m} \sum_S \frac{I(\lambda_l + \lambda_j)}{|\lambda_j|^{-2d}}.$$

Concentrating out C, we can define the objective function to be

$$R_*(d) = Q_*(\hat{C}, d) = \log \hat{C}_*(d) - 2d \textstyle\sum_S \log |\lambda_j|. \qquad [10]$$

Notice that equation [8] can be considered as a particular case of [10] with $\lambda_l = 0$. Denote by $\hat{d}_s$ the estimate of d_s based on minimizing [10] in the set $d \in (\Delta_1, \Delta_2)$ with $0 < \Delta_1 < \Delta_2 < 1/2$. In the Appendix we provide a Theorem that establishes that $\hat{d}_s$ has the same asymptotic properties as $\hat{d}_0$, particularly, [9] still holds for $\hat{d}_s$.

4. Spurious or real long memory in financial series

Long memory properties of financial data (exchange rate and stock returns) have been analysed in several papers. The dominant view has been that long memory is not present at the levels of the returns (Lo, 1991), but it is usually present in some transformations such as the squares (Baillie *et al.*, 1996; Bollerslev and Mikkelsen, 1996; Ding *et al.*, 1993). An interesting question is whether this evidence is induced by an intrinsic characteristic of the behavior of the returns (e.g. persistence of the volatility process) or if it is spurious. Spurious evidence may arise for several reasons such as nonstationarity of the mean or aggregation effects. (An analysis of some of these causes with stock market data is Lobato and Savin, 1996). Nonstationarity in the mean of a process is the most usual reason to cause spurious evidence of long memory. So, careful checking for stationarity when analysing empirical data should always be performed. This cause has been analysed in several papers (see e.g., Klemes, 1974; Bhattacharya *et al.*, 1983). Aggregation may be a relevant cause when analysing indices.

In this section we consider another cause: the presence of a seasonal long memory component in the return process may imply (spurious) evidence of long memory in the squared return process. It is not clear how relevant this argument is for financial data. On the one hand, our experience with stock market and exchange rate daily suggests that the presence of seasonal long memory is unusual (however, in the next section we analyse an exchange rate series which may be indicative of this phenomenon). On the other hand, strong seasonality patterns are usually present with intradaily data (see Goodhart and O'Hara, 1995; or Guillaume *et al.*, 1995) suggesting that this cause could have some relevance with this type of data.

In the following lemma we show that if x_t is a covariance stationary linear process that exhibits seasonal long memory then, under certain conditions, the x_t^2 process will exhibit long memory.

LEMMA. *Let x_t be a covariance stationary process that admits the following representation as a zero mean linear process*

$$x_t = \sum_{k=0}^{\infty} a_k e_{t-k}$$

with e_t a martingale difference sequence, $E(e_t|F_{t-l}) = 0$, where F_t is the σ-field of events generated by $\{e_s, s \leq t\}$, with $E|e_t| < \infty$, $E(e_t^2|F_{t-1}) = 1$ for normalization and $\sum_{k=0}^{\infty} a_k^2 < \infty$. Assume for simplicity that the cumulant of x_t, x_t, x_{t+j} and x_{t+j}, denoted by $cum(x_t, x_t, x_{t+j}, x_{t+j})$, is equal to zero for $t, j = 1, \ldots$ Let x_t exhibit seasonal long memory so its autocovariances are given by [6] with $d_s \in [1/4, 1/2)$, then x_t^2 will exhibit long memory in the sense that [1] will be satisfied for the autocovariances of x_t^2.

PROOF. Define

$$\gamma_0 = Ex_t^2 = \sum_{k=0}^{\infty} a_k^2.$$

The lag-j autocovariance of x_t^2 will be

$$\begin{aligned}\nu_j &= E[(x_t^2 - \gamma_0)(x_{t+j}^2 - \gamma_0)] = Ex_t^2 x_{t+j}^2 - \gamma_0^2 \\ &= 2\gamma_j^2 + cum(x_t, x_t, x_{t+j}, x_{t+j}) = 2\gamma_j^2,\end{aligned}$$

where we have used that for zero mean random variables u, v, w and z

$$\begin{aligned}E(uvwz) &= E(uv)E(wz) + E(uw)E(vz) \\ &+E(uz)E(vw) + cum(u, v, w, z).\end{aligned}$$

Then

$$\nu_j = 2\gamma_j^2 \sim 2K^2 j^{4d_s - 2} \cos^2(j\lambda_*).$$

Showing [1] for ν_j is immediate because $\cos^2(j\lambda_*)$ is nonnegative and periodic in j with period π/λ_* and $\sum_{j=1}^{\infty} j^{4d_s-2} = \infty$ for $d_s \in [1/4, 1/2)$. □

This lemma is interesting because it provides a potential theoretical reason to explain the empirical features of financial processes commented above. Furthermore, it shows that we do not necessarily need a nonlinear model to characterize a series which does not exhibit long memory but its squares do.

5. Empirical illustration

In this section we examine the long memory properties of a spot exchange rate series, the British pound against the Deutsche mark. We analyze weekly data (the closing rate every Friday) from January 1989 to July 1994 so that the sample size is 292. These data are examined in Lobato (1995) where four spot exchange rates series are considered, the British pound against the US dollar, the Japanese yen, the Swiss franc and the Deutsche mark. The dollar and the yen series do not present evidence of seasonal long memory, whereas the other two do (the behavior of the Swiss franc is very similar to the mark). We focus on the mark because our intention is to provide an example that could serve as an illustration of the argument developed in the previous section.

In Figure 1 we present the plot of the logarithm of the exchange rate series. Since it seems to be nonstationary, following a standard practice in the Finance literature, we will concentrate on analyzing the weekly nominal returns. These are calculated by first differencing the logarithm of the exchange rates series and are plotted in Figure 2. These returns constitute our basic series and are denoted by BP/DM. Their squares are in Figure 3.

FIGURE 1
British Pound / Deutsche Mark

BP/DM appears to be stationary but the plots may also indicate the presence of a structural break in the mean of the squared returns around September 1992. As commented in the previous section, a

Figure 2
BP / DM

shift in its mean may be interpreted as (spurious) evidence of long memory in a process. So, it is interesting to test for constancy of the mean of squared BP/DM. The testing procedure is not straightforward, though. Traditional structural break tests have been designed for a short memory context. The only reference we know to test for structural break in the presence of long memory is Hidalgo and Robinson (1996). But even this procedure has an important shortcoming: in order to derive a χ^2 asymptotic null distribution they assume that the process is Gaussian. In order to check this in Table 1 we report the summary statistics for BP/DM and its squares. Since Gaussianity is overwhelmingly rejected there is no guarantee that the null asymptotic distribution of the Hidalgo and Robinson test applied to squared BP/DM is a χ^2. Fortunately, as they point out, the results of Yajima (1991) indicate that the same asymptotic results can be obtained when considering linear processes whose moments are all finite. Again, this may not be the case for squared BP/DM and, in general, for most financial series (see Loretan and Phillips, 1993). But, at any rate, it is not clear how to test robustly

FIGURE 3
Squared BP/DM

for finiteness of moments in a time series context. (Robust procedures such as the one analyzed by Hall (1982) are valid for the random sample case). Therefore, despite its limitations, we have applied Hidalgo and Robinson procedure to test for a shift in the mean of the squares of BP/DM choosing as break point September 18, 1992 (other choices do not change the results). This implies two subsamples made up of 193 and 98 observations. Implementation of this test requires to estimate the long memory parameter, d_0, using a better-than-log(n) estimate. Any of the three semiparametric estimates considered in Section 3 are valid. Hidalgo and Robinson use the APE; we have tried both the APE and the PMLE for different values of m. The test results shown in Table 2 indicate that the null hypothesis of no shift in the mean of the squared-return process is never rejected (the 5% critical value of a χ^2 with one degree of freedom is 3.84).

In Figure 4 we present an estimate of the sdf for BP/DM. We employ

TABLE 1
Summary Statistics

	BP/DM	Squared BP/DM
Mean	-0.000938	0.000115
Variance	0.000114	$0.104\ 10^{-6}$
Skewness	-0.858	8.88
Excess Kurtosis	5.61	102.8
Jarque-Bera test	417.6	131980

TABLE 2
Stationarity test (Hidalgo and Robinson)

	m			
	20	30	40	50
Test (PMLE)	2.86	2.44	2.47	1.86
Test (APE)	3.60	3.22	2.81	2.50

a smoothed periodogram estimate for the normalized sdf

$$\hat{f}(\lambda) = \frac{1}{2\pi} \sum_{u=1-n}^{n-1} k\left(\frac{u}{M}\right) \hat{\rho}_u e^{-iu\lambda},$$

using as lag window the modified Bartlett window

$$k(x) = \begin{cases} 1 - |x|, & if \quad |x| \leq 1 \\ 0, & otherwise \end{cases}$$

where M is a bandwidth parameter. The autocorrelations ($\rho_u = \gamma_u/\gamma_0$) are estimated using

$$\hat{\rho}_u = \left(\sum_{t=1}^{n-|u|} (x_t - \bar{x})(x_{t+|u|} - \bar{x})\right) / \left(\sum_{t=1}^{n} (x_t - \bar{x})^2\right), \text{with } \bar{x} = \frac{1}{n}\sum_{t=1}^{n} x_t.$$

The bandwidth M has been set equal to 35. Lower M produces smoother plots whereas higher M produces the opposite, but the main features remain the same. Notice that, in order to facilitate interpretation, on the X-axis we state the periods instead of the associated frequencies. There are two points worth noticing in Figure 4. First, there is no pronounced peak close to the origin what suggests no evidence of long memory. Second, there is a very marked peak at a frequency close to π; this suggests a periodic behavior with a period a bit longer than two weeks. Figure 5 presents the smoothed periodogram estimate for the sdf of the squares of BP/DM. The peak at zero frequency is the main feature of this plot.

FIGURE 4
Smoothed Periodogram of BP/DM

In Table 3 we provide the pseudo maximum likelihood estimates for the level of the time series and its squares. We report estimates for m equal to 20 and 30 and the 95% asymptotic confidence interval based on [9]. Table 3 shows the absence of evidence of long memory at zero frequency in the series, but the presence of long memory in the squares. Notice that the width of this asymptotic confidence

Figure 5
Smoothed Periodogram of squared BP/DM

interval is $1.96/\sqrt{m}$ so that for m=20 it is 0.44 and for m=30 it is 0.36. Clearly, the intervals are very broad and this is the main disadvantage of the semiparametric approach compared to the parametric one. As noted in Section 3, robustness is achieved at the expense of efficiency (Robinson, 1994a; Robinson, 1994b). Note, however, that the efficiency of the semiparametric estimate can be improved in a multivariate framework. In Lobato (1995) BP/DM and the three exchange rates returns previously mentioned are analysed jointly, and for instance, for m=20, instead of 0.18(-0.04, 0.40) we obtain 0.24(0.12, 0.36), which shows that the asymptotic confidence interval is substantially reduced.

We now show that the evidence of long memory in the squared returns can be explained by seasonal long-memory in the returns. Figure 4 shows a very strong peak at a frequency close to π, therefore we fit a model like [5] to BP/DM. Notice that the estimation procedure outlined in Section 3 is only valid for a known fixed λ_* and does not cover the case of data dependent λ_*. In practice λ_* is unknown and

TABLE 3
Estimates of d_0

	m	
	20	30
BP/DM	-0.07 (-0.28, 0.15)	0.17 (-0.02, 0.33)
Squared BP/DM	0.18 (-0.04, 0.40)	0.23 (0.05, 0.41)

has to be estimated. In general, this may not be an easy task, but for our series, based on Figure 4 we can establish λ_l to be about $136\pi/146$ or $137\pi/146$. In Table 4 we present the estimates for d_s for a set of λ_*. From this table we see that in this case the selection of a specific λ_* is not crucial. In Table 4 we report the estimates of d_s based on minimizing [10] and the asymptotic confidence intervals based on [9] for d_s. Notice that the estimates of d_s are almost always in the region (1/4, 1/2) and the asymptotic confidence intervals usually over the interval. This supports the hypothesis that the evidence of long memory in the squares of BP/DM could be due to a strong seasonal long memory behavior of BP/DM.

6. Conclusion

In this paper we have analysed semiparametric procedures for long memory series and applied them to an exchange rate return series. We have showed that there is strong evidence of seasonal long memory in that series and that this could be causing the appearance of long memory in their squared returns. It is difficult to ponder the importance of the former result. Taken by its own value, this result suggest that it is possible to forecast to some extent these returns. Obviously this does not mean that the efficient market hypothesis is violated. The improvement in the forecasting performance may not be enough to compensate for transaction costs. Furthermore, that result can be due to some unknown institutional factor that could have affected that exchange rate market during the period considered. The second result would imply that the appearance of long memory in the squared returns is spurious. This would question the

Table 4
Estimates of d_s for BP/DM

λ_*	m: 20	m: 30
134.5/146	0.30 (0.09, 0.52)	0.30 (0.12, 0.47)
135 /146	0.40 (0.18, 0.62)	0.31 (0.13, 0.48)
135.5/146	0.41 (0.19, 0.62)	0.30 (0.12, 0.48)
136 /146	0.34 (0.12, 0.56)	0.20 (0.02, 0.38)
136.5/146	0.45 (0.23, 0.66)	0.26 (0.09, 0.44)
137 /146	0.42 (0.20, 0.63)	0.19 (0.01, 0.37)
137.5/146	0.44 (0.22, 0.65)	0.22 (0.04, 0.40)
138 /146	0.42 (0.20, 0.64)	0.20 (0.02, 0.38)

approach of analyzing the returns using ARCH or stochastic volatility models.

Notice that, as was pointed out at the beginning of Section 5, from our original data set neither the British pound against the US dollar nor the British pound against the Japanese yen present evidence of seasonal long memory. It would be interesting to analyse other exchange rate series and check if the presence of seasonal long memory is more general.

Appendix

In this appendix we extend Robinson (1995a) semiparametric estimation procedure to the seasonal long memory case. We just state the Assumptions and the Theorem and provide a heuristic proof. Consider x_t a covariance stationary process with sdf f(λ). Assume:

Assumption 1. *For some* $\beta \in (0, 2]$

$$f(\lambda) \sim G_0|\lambda_* - \lambda|^{-2d_s}(1 + O(\lambda^{\beta})) \quad as \lambda \to \ \lambda_*,$$

with $G_0 \in (0, \infty)$, $d_s \in (0, 1/2)$.

Assumption 2. x_t *has a linear representation*

$$x_t = Ex_0 + \sum_{j=0}^{\infty} \alpha_j e_{t-j} \quad , \quad \sum_{0}^{\infty} \alpha_j^2 < \infty,$$

with

$$E(e_t|F_{t-1}) = 0, \qquad E(e_t^2|F_{t-1}) = 1,$$
$$E(e_t^3|F_{t-1}) = \mu_3, \qquad E(e_t^4) = \mu_4,$$

where μ_3 and μ_4 are finite constants.

Now, define $\alpha(\lambda) = \sum_{j=0}^{\infty} \alpha_j e^{-ij\lambda}$ and introduce:

ASSUMPTION 3. *In a neighbourhood $(\lambda_* - \delta,\ \lambda_* + \delta)$ $\alpha(\lambda)$ is differentiable and*

$$\frac{d}{d\lambda}\alpha(\lambda) = O\left(\frac{|\alpha(\lambda)|}{\lambda}\right) \quad as\ \lambda \to \lambda_*\ .$$

ASSUMPTION 4. *As $n \to \infty$*

$$\frac{1}{m} + \frac{m^{1+2\beta}(\log m)^2}{n^{2\beta}} \to 0.$$

Assumptions 1 and 3 are trivial extensions of A1' and A2' in Robinson (1995a) while Assumptions 2 and 4 are just the same as A3' and A4'.

THEOREM: *Under Assumptions 1 to 4*

$$\sqrt{m}(\hat{d}_s - d_s) \rightarrow_d N\left(0, \frac{1}{4}\right),$$

where $\hat{d}_s$ is the argument that minimizes [10] over the set $[\Delta_1, \Delta_2]$ with $0 < \Delta_1 < \Delta_2 < 1/2$.

PROOF: A formal proof would basically consist in rewriting the proof in Robinson (1995a). This basically entails to prove:

(1) $\hat{d}_s \rightarrow_p d_s$,

(2) $\frac{dR_*(d_s)}{d(d)} \rightarrow_d N\left(0, \frac{1}{4}\right)$, and

(3) $for\ any\ \tilde{d}_s \rightarrow_p d_s$: $\frac{d^2 R_*(\tilde{d}_s)}{d(d)^2} \rightarrow_p 4$.

The proofs of (1) and (3) follow similarly as in Theorem 1 and Theorem 2 in Robinson (1995a). With respect to (2), the main difference is that now instead of expression (4.11) in Robinson (1995a), we have:

$$\sqrt{m}\frac{dR_*(d_s)}{d(d)} = \frac{2}{\sqrt{m}}\sum_{S_1'}\nu_j\left(\frac{I(\lambda_l+\lambda_j)}{g_j}-1\right)$$

$$+\tfrac{2}{\sqrt{m}}\sum_{S_2'}\nu_j\left(\tfrac{I(\lambda_l+\lambda_j)}{g_j}-1\right)+o_p(1) \qquad \text{[A1]}$$

where $S_1' = [1, 2, \ldots, m/2]$ while $S_2' = [-m/2, \ldots, -2, -1]$ if λ_* is a Fourier frequency and $S_2' = [-m/2-1, \ldots, -1, 0]$ if not, while

$$\nu_j = \log j - \frac{2}{m}\sum_{j=1}^{m/2}\log j \ \ , \ g_j = C{\lambda_j}^{-2d_s}.$$

Notice that calling $m/2 = m^*$, we can rewrite [A1] as:

$$\frac{1}{\sqrt{2}}\left(\frac{2}{\sqrt{m^*}}\sum_{S_1}\nu_j\left(\frac{I^*(\lambda_j)}{g_j}-1\right)\right.$$

$$\left.+\frac{2}{\sqrt{m^*}}\sum_{S_2}\nu_j\left(\frac{I^*(\lambda_j)}{g_j}-1\right)\right)+o_p(1),$$

where $I^*(\lambda_j) = I(\lambda_l+\lambda_j)$ and $S_1 = [1, 2, \ldots, m^*]$ while
$S_2 = [-m^*, \ldots, -2, -1]$ if λ_* is a Fourier frequency and
$S_2 = [-m^*-1, \ldots, -1, 0]$ if not. Following a similar procedure as in Robinson (1995a), we get that both terms in the main bracket are asymptotically N(0,4), so, in order to establish that [A1] converges in distribution to a random variable distributed as a N(0,4), we need to prove that both terms are asymptotically independent. This proof is tedious, but basically similar to the proof of expression (14) in Theorem 1 in Lobato and Robinson (1996) that relies heavily on Robinson (1994b). The heuristic idea is that the intervals in which we evaluate the periodograms do not overlap.

References

Baillie, R.T., T. Bollerslev and H.O.A. Mikkelsen (1966): "Fractionally integrated generalized autoregressive conditional heteroskedasticity", *Journal of Econometrics,* forthcoming.

Bhattacharya, R.N., V.K. Gupta and E. Waymire (1983): "The Hurst effect under trends", *Journal of Applied Probability* 20, pp. 649-662.

Bollerslev, T. and H.O.A. Mikkelsen (1966): "Modeling and pricing long-memory in stock market-volatility", *Journal of Econometrics* 73, pp. 151-184.

Delgado, M. and P.M. Robinson (1994): "New methods for the analysis of long memory time series: application to Spanish inflation", *Journal of Forecasting* 13, pp. 97-107.

Ding, Z., C.W.J. Granger and R.F. Engle (1993): "A long memory property of stock market returns and a new model", *Journal of Empirical Finance* 1, pp. 83-106.

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

Goodhart, C.A.E. and M. O'Hara (1995): "High frequency data in financial markets: issues and applications", *Journal of Empirical Finance,* forthcoming.

Granger, C.W.J. and R. Joyeux (1980): "An introduction to long-memory time series and fractional differencing", *Journal of Time Series Analysis* 1, pp. 15-30.

Gray, H.L., N. Zhang and W.A. Woodward (1989): "On generalized fractional processes", *Journal of Time Series Analysis* 10, pp. 233-257.

Guillaume, D., M. Dacorogna, R. Dave, U. Muller, R. Olsen and O. Pictet (1944): "From the bird's eye to the microscope: a survey of new stylized facts of the intra-daily foreign exchange markets", O&A Research Group, preprint.

Hall, P. (1982): "On some simple estimates of an exponent of regular variation ", *Journal of the Royal Statistical Society* B, 44, pp. 37-42.

Hidalgo, J. and P.M. Robinson (1966): "Testing for structural change in a long-memory environment", *Journal of Econometrics* 70, pp. 159-174.

Hosking, J.R.M. (1981): "Fractional differencing", *Biometrika* 68, pp. 165-176.

Klemes, V. (1974): "The Hurst phenomenon: a puzzle?", *Water Resources Research* 10, pp. 675-688.

Lo, A.W. (1991): "Long term memory in stock market prices", *Econometrica* 59, pp. 1279-1313.

Lobato, I. (1995): "Multivariate quasi maximum likelihood analysis of long memory series", University of Iowa, preprint.

Lobato. I. and P.M. Robinson (1996): "Averaged periodogram estimation of long memory", *Journal of Econometrics* 73, pp. 303-324.

Lobato, I. and N.E. Savin (1996): "Real and spurious long memory properties of daily stock market data", *Journal of Business and Economic Statistics,* forthcoming.

Loretan, M. and P.C.B. Phillips (1993): "Testing the covariance stationarity of heavy-tailed time series: an overview of the theory with applications to several financial datasets", *Journal of Empirical Finance* 1, pp. 211-248.

Mandelbrot, B.B. (1969): "Long-run linearity, locally Gaussian processes, H-spectra and infinite variances", *International Economic Review* 10, pp. 82-111.

Mandelbrot, B.B. and J.W. Van Ness (1968): "Fractional Brownian motions, fractional noises and applications", *SIAM Review* 10, pp. 422-437.

Robinson, P.M. (1994a): "Semiparametric analysis of long-memory time series", *Annals of Statistics* 22, pp. 515-539.

Robinson, P.M. (1994b): "Rates of convergence and optimal spectral bandwidth for long range dependence", *Probability Theory and Related Fields* 99, pp. 443-473.

Robinson, P.M. (1994c): "Time series with strong dependence", in C.A. Sims (ed.), *Advances in Econometrics: Sixth World Congress*, Vol. 1, Cambridge University Press, Cambridge, pp. 47-95.

Robinson, P.M. (1995a): "Gaussian semiparametric estimation of long range dependence", *Annals of Statistics* 23, pp. 1630-1661.

Robinson, P.M. (1995b): "Log-periodogram regression of time series with long range dependence", *Annals of Statistics* 23, pp. 1048-1072.

Stoll, H.R. and R.E. Whaley (1990): "The dynamics of stock index and stock index futures returns", *Journal of Financial and Quantitative Analysis* 25, pp. 441-468.

Yajima, Y. (1991): "Asymptotic properties of the LSE in a regression model with long-memory stationary errors", *Annals of Statistics* 19, pp. 158-177.

Resumen

Diversos artículos han encontrado evidencia de memoria larga en los cuadrados de los rendimientos de varias series financieras. Esto ha sido interpretado como indicativo de que el proceso de volatilidad de estos rendimientos es persistente. En este artículo analizamos cómo esa evidencia puede ser explicada por la presencia de un componente estacional de memoria larga en los rendimientos. Además analizamos estimadores semiparamétricos de modelos de memoria larga en el dominio de la frecuencia y aplicamos el estimador de pseudo máxima verosimilitud de Robinson (1995a) al caso de memoria larga estacional. Por último, estas técnicas son empleadas en la serie de tipos de cambio libra-marco.