Spectral Properties of Temporally Aggregated Long Memory Processes

Leonardo Rocha Souza*

Fundação Getúlio Vargas, Graduate School of Economics.

March 2003

Abstract

This paper derives the spectral density function of aggregated long memory processes in light of the aliasing effect. The results are different from previous analyses in the literature and a small simulation exercise provides evidence in our favour. The main result point to that flow aggregates from long memory processes shall be less biased than stock ones, although both retain the degree of long memory. This result is illustrated with the daily US Dollar/ French Franc exchange rate series.

JEL classification: C14, C22, C43

Keywords: Temporal Aggregation, Long Memory, Aliasing, Fejer Kernel

* Corresponding author: Leonardo Rocha Souza
E-mail: leors@fgv.br
Phone: +55 21 93330145
Fax: +55 21 25524898
Address: EPGE – Fundação Getúlio Vargas.
Praia de Botafogo, 190, 11 andar, Rio de Janeiro
CEP 22253-900 Brazil

1 - Introduction

Temporally aggregated time series data appear frequently in Economics, either because the collecting process implicitly aggregates data from different time periods or because the econometrician wants to work on a time frequency different (lower) than that of the collected data. Temporal aggregation has distinct interpretations depending on the underlying variable be a stock or a flow variable. A stock variable is aggregated through skip-sampling while a flow variable is aggregated by summing subsequent observations letting no overlap in the sums.

The question as to use aggregated or disaggregated data does not have a clear answer. Using aggregated data has the disadvantage of reducing the sample size, while on the other hand often simplifies the model to be fitted to the data as it reduces the influence of some short run components. Moreover, long memory time series models require longer series data to be reasonably estimated than ARMA models, for example. If the time unit is small, high frequency data are collected and the sample size usually remains big after aggregation. However, if the original data are, e.g., monthly and aggregated to quarterly series, a small aggregated series is usually obtained. An example of high frequency series whose memory parameter (or alternatively integration order) is estimated across different levels of aggregation is the French Franc/German Mark exchange rate that appears in Bisaglia and Guégan (1998). As to low frequency data, Diebold and Rudebusch (1989) use annual and quarterly data in their (long memory) study of real US GNP and Chambers (1998) uses quarterly and annual flow data to estimate the memory parameter of a number of UK macroeconomic series. Tschernig (1995), in turn, uses daily, weekly, monthly and quarterly data to investigate the long memory in many exchange rates.

Specification (and identification) of aggregated long memory processes is important since it can shed a light on the question of which data (aggregated or disaggregated) is preferable to use. Tschernig (1995) and Teles, Wei and Crato (1999) studied the effects of temporal aggregation (the latter only for the flow type) on AutoRegressive Fractionally Integrated Moving Average (ARFIMA) models. They conclude that an ARFIMA(p,d,q) turns out to be an ARFIMA(p,d,∞) when aggregated. They do not provide, however, the MA polynomial form for non-overlapping aggregates[1]. Crato and Ray (2002) provide the spectral density function of aggregated (flow) ARFIMA processes, but it depends on the MA(∞) polynomial that is not derived completely in Tschernig (1995) and Teles, Wei and Crato (1999). Chambers (1998) investigates the spectral density function of stock and flow aggregated long memory processes, as well as continuous-time long memory processes observed at discrete-time intervals and cross-sectionally aggregated long memory processes. For the temporal aggregation of discrete-time processes, however, he does not take into account the aliasing effect as he does in the case of the continuous-time processes. All these papers agree that neither type of aggregation changes the memory parameter.

The related issue of estimating the fractional integration parameter of aggregated series is investigated in Souza and Smith (2002) for stock series, and in Souza and Smith (2003) for flow series. The former concludes that temporal aggregation of stock processes induces a non-negligible bias in the memory parameter and relates it to the aliasing effect. In contrast, the latter finds that temporal aggregation of flow processes generally does not induce any significant bias.

[1] They provide the MA(∞) polynomial (or an insight to it) for the overlapping type of aggregation, although they refer to non-overlapping aggregated processes.

The issue of forecasting aggregated versus disaggregated long memory series is investigated in Souza and Smith (2003). They compare the forecasting performance of aggregated (flow) ARFIMA series against the cumulative forecasts from the original series. They conclude that in general disaggregated models forecast better for $d>0$ and worse for $d<0$, comparing with aggregated models. This latter finding contrasts with the results from ARIMA models, where forecasts from the aggregated series are, at best, not much worse than aggregated forecasts from the underlying series (see Amemiya and Wu, 1972, Lütkepohl, 1986, Wei, 1989, and Hotta and Neto, 1993, to name a few).

The present paper is inserted among those on the issues of identification and estimation. It derives the spectral density function (alternatively speaking, spectral function, spectral density or spectrum) of aggregated stock and flow long memory processes in light of the Aliasing Theorem, unlike the previous works of Chambers (1998) and Crato and Ray (2002). The result is a completely different formula which a simulation study brings evidence in favour, as it matches with the periodogram averaged across 100 realizations of each process. The Aliasing Theorem is adapted to the case a discrete-time process is observed at a slower sampling rate. One may infer from the formulas that flow aggregates from long memory processes shall be less biased than stock ones. This result is illustrated with the daily US Dollar/ French Franc exchange rate from October 20, 1977 to October 23, 2002. In a long memory stochastic volatility model framework, the logarithm of the squared returns are analysed and the absence of long memory is rejected by the Lo's (1991) modified R/S test. The series is aggregated as flow and stock variables and flow aggregates yield estimates of the same order as the original series, while stock aggregates yields estimates somewhat lower. This result is also consistent with the aforementioned works of Souza and Smith (2002a, b).

A secondary result is related with two conditions usually taken as equivalent; the time-domain condition:

$$\rho_k \sim c_\rho(k) k^{2d-1} \text{ as k} \to \infty \qquad (1)$$

and the frequency-domain condition:

$$f(\lambda) \sim c_f(\lambda)|\lambda|^{-2d} \text{ as } \lambda \to 0 \qquad (2)$$

where ρ_k and $f(\lambda)$ are respectively the k-th order autocorrelation and the spectral density function of the process; $c_\rho(k)$ and $c_f(\lambda)$ are functions slowly varying as k tends to infinity and λ to zero, $c_\rho(k)$ having the same sign as d and $c_f(\lambda)$ being always positive; $\lambda \in (-\pi, \pi]$; and d is a real number. Stationarity holds for d < 0.5 so that it must apply in (1) and (2) in order to ρ_k and $f(\lambda)$ have the usual meaning. The result is that a negative memory parameter d, according to its frequency-domain condition, changes with the aggregation of stock variables while remains unchanged if taken as to the time-domain condition. This proves the non-equivalence for $d < 0$ of the time- and frequency-domain conditions usually considered in the literature to define the memory parameter, given respectively by equations (1) and (2). On the other hand, aggregating a flow variable does not change the memory parameter according to either definition.

The plan of the paper is as follows. Section 2 defines long memory and ARFIMA models. Section 3 derives the autocovariances and the spectral density function of aggregated long memory processes, while Section 4 provides a small Monte Carlo simulation. An example with the daily US Dollar/ French Franc exchange rate series illustrates the results in Section 5. Section 6 offers a final consideration. All technical details and proofs are relegated to the Appendix.

2 – Long Memory Models

Long memory in stationary processes has traditionally two alternative definitions, one in the frequency-domain and the other in the time-domain[2]. Although there are some works in the literature that state they are equivalent, Robinson (1995, p.1632) argues that this is true only if the autocovariances are quasimonotonically convergent to zero. The time- and frequency-domain definitions of stationary long memory processes are respectively as follows:

Definition 1: Let X_t be a stationary process such that (1) holds for a real number $d \in (0, 0.5)$. Then X_t displays long memory or equivalently long range dependence.

Definition 2: Let X_t be a stationary process such that (2) holds for a real number $d \in (0, 0.5)$. Then X_t is a stationary process with long memory and d is the memory parameter.

In a broader context, allowing d outside (0, 0.5), X_t is said to be antipersistent if $d \in (-0.5, 0)$; while if $d > 0.5$ the process is non-stationary (and still long memory); and if $d < 0.5$ the process is non-invertible. Short memory processes are those for which $d = 0$. The first long memory model to appear in the literature is the Fractional Gaussian Noise (Mandelbrot, 1965, Mandelbrot and Van Ness, 1968), while the most popular class of models is the ARFIMA (Hosking, 1981, Granger and Joyeux, 1980),

[2] Other definitions are available, e.g. the one in McLeod and Hipel (1978) that defines stationary processes with long memory those whose autocorrelations have infinite sum.

which is a generalization of the ARIMA models allowing a non-integer integration parameter. X_t follows an ARFIMA(p,d,q) model if $\Phi(\mathrm{B})(1-\mathrm{B})^{\mathrm{d}}\mathrm{X}_{\mathrm{t}} = \Theta(\mathrm{B})\varepsilon_{\mathrm{t}}$, where ε_t is a mean-zero, constant variance (σ_ε^2) white noise process, B is the backward shift operator such that $BX_t = X_{t-1}$, and $\Phi(B) = 1-\phi_1 B-\ldots-\phi_p B^p$ and $\Theta(B)=1+\theta_1 B+\ldots+\theta_q B^q$ are the short-run autoregressive and moving-average polynomials, respectively. If the roots of $\Phi(B)$ are outside the unit circle, the process is stationary and if the roots of $\Theta(B)$ are outside the unity circle the process is invertible. A non-integer difference can be expanded into an infinite autoregressive or moving average polynomial using the binomial theorem:

$$(1-B)^d = \sum_{k=0}^{\infty} \binom{d}{k} (-B)^k , \qquad (3)$$

where $\binom{d}{k} = \frac{\Gamma(d+1)}{\Gamma(k+1)\Gamma(d-k+1)}$ and $\Gamma(.)$ is the gamma function.

The Wold representation of an ARFIMA model is obtained through inverting the non-integer difference and the AR polynomials as follows:

$$X_t = (1-B)^{-d}\Phi^{-1}(B)\Theta(B)\varepsilon_t \qquad (4)$$

The term $(1\text{-B})^{-d}$ is computed using equation (3), replacing *d* by *–d*. The AR polynomial can be inverted more easily if it is first factored in smaller terms $(1-c_i B)$, i = 1, …, p, inverted and then convoluted back. The ARFIMA spectral density function is given by:

$$f(\lambda) = \frac{\sigma_\varepsilon^2}{2\pi} \left|1-e^{-j\lambda}\right|^{-2d} \frac{\left|\Theta(e^{-j\lambda})\right|^2}{\left|\Phi(e^{-j\lambda})\right|^2}, \quad -\pi < \lambda \le \pi , \qquad (5)$$

where $j^2 = -1$ and λ is the frequency (Hosking, 1981). The spectral density function of ARFIMA processes has a pole in the frequency zero if d > 0, whereas if d < 0 it is bound to zero in that frequency. An ARFIMA model satisfies Definitions 1 and 2 if 0

< d < 0.5, having thus long memory while being stationary. An extensive overview of the long memory literature up to is given in Beran (1994).

3 – Aggregation of Long Memory Models

By temporal aggregation we mean that a process is observed at a frequency slower than that it is generated at. Let *n* be level of aggregation. If the variable is a stock variable it is observed every *n*-th period while if it is a flow variable a sum of the *n*-th and the *n*-1 preceding periods is observed every *n*-th period.

Definition 3: The aggregated variable Y_t is observed as follows:

3a) If X_t is a stock variable, then $Y_t = X_{nt}$, t = 1, …, T.

3b) If X_t is a flow variable, then $Y_t = \sum_{i=0}^{n-1} X_{nt-i} = \sum_{i=0}^{n-1} B^i X_{nt}$, t = 1, …, T.

The difference between both is that a moving average filter $\sum_{i=0}^{n-1} B^i$ = (1 + B + … + B^{n-1}) is applied to X_t in the case of a flow variable before skip-sampling while the stock variable is simply "skip-sampled". The term skip-sampling is used in Otero and Smith (2000) and Souza and Smith (2002), while Brewer (1973) and Weiss (1984) use the term systematic sampling for the action of producing Y_t as in Definition 3a. In some of these works these terms contrast with (temporal) aggregation, used for the action of producing Y_t as in Definition 3b. Teles and Wei (2002) as well as Teles, Wei and Crato (1999) use the term temporal aggregation solely to describe Definition 3b. We follow the nomenclature found in Tschernig (1995) and Chambers (1998), which address temporal aggregation for both actions, discerning between them by the series type, stock or flow.

Temporal aggregation as defined includes at some part the act of skip-sampling. This causes the aliasing phenomenon, well known in the signal processing literature for continuous-time processes observed at discrete-time intervals. The literature seems to give little heed, however, to the fact that the same causes for the aliasing to appear when observing a continuous process at discrete-time are also present in the act of skip-sampling. In fact, a number of signal processing and time series books (e.g. Priestley, 1981, Oppenheim and Schafer, 1989, Hamilton, 1994) explain the aliasing effect only as a phenomenon which arises when observing continuous-time processes at discrete intervals. Koopmans (1974), although hinting that aliasing would appear when sampling discrete-time processes at a lower sampling rate (p. 70, figure), states (p. 71): "The aliasing problem arises when the spectrum of interest is that of the original, continuous-time series.". However, the explanation of this phenomenon and the derivation of its effects when observing a discrete process at a lower sampling rate are almost identical. The difference lies in that the spectral density function of discrete-time processes is defined only over the range $(-\pi, \pi]$ while that of continuous-time processes is defined over the real line $\Re$.

An intuitive explanation of the phenomenon occurring in discrete processes observed at a lower sampling frequency is the following. When the sampling frequency is lower than that of the underlying process by a factor n, a component with frequency ω in the original process will have frequency $n\omega$ in the newly sampled series, possibly falling outside $(-\pi, \pi]$. Alternatively, the frequency interval $(-\pi, \pi]$ in the spectrum of the aggregated process is equivalent to $(-\pi/n, \pi/n]$ in the original process. Clearly some frequencies of the original process will not be directly observed in the aggregated process (and therefore will not appear in its spectrum), for they will complete more than an entire cycle between two subsequent observations, since their respective periods are smaller than the sampling period. Instead, components with these frequencies will have an apparent (lower) frequency in the aggregated process, different from the "real" frequency. All frequencies under the same apparent frequency will be observed together. This is, loosely speaking, the aliasing effect and is equivalent to folding the spectrum *n* times into the interval (-π/*n*, π/*n*].

The aliasing effect arising from aggregating discrete-time processes is given in the following theorem:

Theorem 1: Let X_t be a covariance stationary discrete-time process with spectral density function $f_x(\omega)$ and $Y_t = X_{nt}$. The spectral density function of Y_t, $f_y(\lambda)$, is given by:

1a) If *n* is an odd number:

$$f_y(\lambda) = \frac{1}{n} \sum_{i=-\frac{n-1}{2}}^{\frac{n-1}{2}} f_x\left(\frac{\lambda}{n} + \frac{2i\pi}{n}\right), \qquad -\pi < \lambda \le \pi \tag{6}$$

1b) If *n* is an even number:

$$f_y(\lambda) = \begin{cases} \dfrac{1}{n} \displaystyle\sum_{i=1-\frac{n}{2}}^{\frac{n}{2}} f_x\left(\frac{\lambda}{n} + \frac{2i\pi}{n}\right), & -\pi < \lambda \le 0 \\ \dfrac{1}{n} \displaystyle\sum_{i=-\frac{n}{2}}^{\frac{n}{2}-1} f_x\left(\frac{\lambda}{n} + \frac{2i\pi}{n}\right), & 0 < \lambda \le \pi \end{cases} \tag{7}$$

Theorem 1 gives the relationship between the spectra of the original and the aggregated stock process. To have the general formula for the spectrum of aggregated flow processes given further in Corollary 1, however, one needs an additional result. This result is given in Theorem 2 as follows.

Theorem 2: Let X_t be a covariance stationary discrete-time process with spectral density function $f_x(\omega)$ and $Z_t = \sum_{i=0}^{n-1} X_{t-i} = \sum_{i=0}^{n-1} B^i X_t$ its overlapping aggregated process. The spectral density function of Z_t, $f_z(\lambda)$, is given by:

$$f_z(\lambda) = 2\pi n.F_n(\lambda).f_x(\lambda), \qquad -\pi < \lambda \le \pi \tag{8}$$

where $F_n(\lambda) = \frac{1}{2\pi n} \lim_{\theta \to \lambda} \frac{sin^2(n\theta/2)}{sin^2(\theta/2)}$ is the Fejer Kernel.

The Fejer Kernel is periodic with period equal to 2π. For λ restricted to the interval (-π, π], it has the highest peak at the frequency zero (far higher than the subsidiary peaks) and zeros at frequencies that are nonzero multiples of $2\pi/n$ as shown in Figure 1 for n = 6. The frequency $2\pi/n$ is called the Nyquist frequency if the process will be further “skip-sampled” as in a flow aggregation. This mean that after applying a moving average filter $(1+B+\ldots+B^{n-1})$, the low frequencies predominate. Furthermore, the Fejer kernel first derivative is zero at all (zero and nonzero) multiples of the Nyquist frequency and the nonzero multiples will be folded into the frequency zero after a further skip-sampling. These properties will be explored ahead in Sections 3.2 and 3.3 in respect to flow temporal aggregation. The relationship between the spectra of the original and the aggregated flow process follows directly from Theorems 1 and 2 and is given in Corollary 1 below.

Corollary 1: Let X_t be a covariance stationary discrete-time process with spectral density function $f_x(\omega)$ and $Y_t = \sum_{i=0}^{n-1} X_{nt-i} = \sum_{i=0}^{n-1} B^i X_{nt}$. The spectral density function of Y_t, $f_y(\lambda)$, is given by:

2a) If n is an odd number:

$$f_y(\lambda) = 2\pi \sum_{i=-\frac{n-1}{2}}^{\frac{n-1}{2}} \left[F_n\left(\frac{\lambda}{n} + \frac{2i\pi}{n}\right) . f_x\left(\frac{\lambda}{n} + \frac{2i\pi}{n}\right)\right], \qquad -\pi < \lambda \leq \pi \tag{9}$$

2b) If n is an even number:

$$f_y(\lambda) = \begin{cases} 2\pi \sum_{i=1-\frac{n}{2}}^{\frac{n}{2}} \left[F_n\left(\frac{\lambda}{n} + \frac{2i\pi}{n}\right) . f_x\left(\frac{\lambda}{n} + \frac{2i\pi}{n}\right)\right], & -\pi < \lambda \leq 0 \\ 2\pi \sum_{i=-\frac{n}{2}}^{\frac{n}{2}-1} \left[F_n\left(\frac{\lambda}{n} + \frac{2i\pi}{n}\right) . f_x\left(\frac{\lambda}{n} + \frac{2i\pi}{n}\right)\right], & 0 < \lambda \leq \pi \end{cases} \tag{10}$$

The proof of Theorem 1 (see Appendix) is adapted from the proof of the Aliasing Theorem for continuous-time processes observed at discrete-time intervals, easily found in the Spectral Analysis books, e.g. Priestley (1981) and Oppenheim and Schafer (1989). The term aliasing is due to John W. Tukey (Priestley, 1981 p. 505) and its motivation is that the energy in the frequencies λ such that $|\lambda| > \pi$ are not observed directly in the spectrum but as aliases of lower frequencies. The proof of Theorem 2 (see Appendix) follows from the moving average filter representation of (1 + B + … + B^{n-1}) after some trigonometric manipulation.

Whether the underlying process is discrete-time or continuous-time determines if the sums in the RHS of (6-7) and (9-10) are respectively finite or infinite, for the spectrum of a continuous process is defined over the real line and the spectrum of a discrete process is defined only over the range (-π, π]. Besides, the multiplicative term $1/n$ is the Jacobian of the frequency transformation $\lambda = n\omega$.

The aliasing phenomenon will then be present in the spectral function of aggregated processes and fractionally integrated processes are not an exception. To formalize our results in respect to long memory processes, let assume that the underlying variable X_t is fractionally integrated, having the following Wold representation.

Assumption 1. X_t is covariance stationary and has the Wold representation given by:

$$(1-B)^d X_t = W(B)\varepsilon_t = \sum_{i=0}^{\infty} w_i \varepsilon_{t-i} \,, \qquad (11)$$

where d < 0.5; ε_t is a white noise with variance σ_ε^2; $w_0 = 1$, and $\sum_{i=0}^{\infty} w_i^2 < \infty$.

For a fractionally integrated process Z_t with d > 0.5, one can always difference Z_t δ times to obtain X_t following Assumption 1 such that $X_t = (1-B)^\delta Z_t$, where δ is an integer. Covariance stationarity is required for the autocovariances to be defined as a function of the lag in time and the spectral function to take its usual definition (given by equation (A1) in the Appendix).

3.1 – Autocovariances

In this section the general form for the autocovariances of aggregated fractionally integrated processes is derived. Although this paper emphasizes frequency-domain behaviour, its time-domain counterpart is also important as the reader can note by the two alternative definitions for long memory. In fact there are a number of both time- and frequency-domain estimators for the integration order d.

Note that if X_t is a stock variable then the autocovariances of the aggregated process are $\gamma_k^y = \gamma_{nk}^x$ and if X_t is a flow variable then (for k≥0) $\gamma_{-k}^y = \gamma_k^y = \sum_{i=1-n}^{n-1} (n-|i|)\gamma_{nk+i}^x$. The variance of the aggregated process Y_t is hence obtained by making k = 0: $\gamma_0^y = \gamma_0^x$ for stock variables and, for flow variables,

$\gamma_0^y = \sum_{i=1-n}^{n-1}(n-|i|)\gamma_i^x = n\gamma_0^x + 2\sum_{i=1}^{n-1}(n-i)\gamma_i^x$. It is easier to use these relationships to obtain the autocovariances of the aggregated process if those of the original process are known. However, if only the model specification is available, the autocovariances of the aggregated processes can be obtained from the Wold representation of the original process (given by Equation (4) for ARFIMA processes) and are expressed in terms of infinite sums. Whether they converge, these infinite sums can (almost) always be approximated using numerical methods. The result for fractionally integrated processes is formulated in the following proposition.

Proposition 1: Let Assumption 1 hold so that X_t has the Wold representation $X_t = \sum_{i=0}^{\infty} a_i \varepsilon_{t-i} = (1-B)^{-d}\sum_{i=0}^{\infty} w_i \varepsilon_{t-i}$, and k-th order autocovariances γ_k^X. Then:

1a) If X_t is a stock variable, the autocovariances of the aggregated process Y_t are given by: $\gamma_k^y = \gamma_{nk}^x = \sigma_\varepsilon^2 \sum_{i=0}^{\infty} a_i a_{i+|nk|}$. (12)

The k-th order autocorrelation of Y_t is given by:

$$\rho_k^y = \rho_{nk}^x = \frac{\sum_{i=0}^{\infty} a_i a_{i+|nk|}}{\sum_{i=0}^{\infty} a_i^2} \qquad (13)$$

1b) If X_t is a flow variable, the autocovariances of the aggregated process Y_t are given by: $\gamma_k^y = \sigma_\varepsilon^2 \sum_{i=0}^{\infty} b_i b_{i+|nk|}$. (14)

The k-th order autocorrelation of Y_t is given by:

$$\rho_k^y = \frac{\sum_{i=0}^{\infty} b_i b_{i+|nk|}}{\sum_{i=0}^{\infty} b_i^2}, \qquad (15)$$

where $b_i = \sum_{k=i-n+1}^{i} a_k$ and $a_k = 0$ for $k < 0$.

The question of identifying the aggregated process from the disaggregated one, however, is more difficult to address. Tschernig (1995) and Teles, Wei and Crato (1999) found that a temporally aggregated ARFIMA(p,d,q) process turns out to be an ARFIMA(p,d,∞) when aggregated. The result is not complete, however, as they leave the MA(∞) polynomial in the original process time unit while the AR(p) and the I(d) polynomials are in the aggregated time unit. Beran and Ocker (2000), in turn, show that an aggregated ARFIMA(p,d,q) process tends to an I(d) process as the level of aggregation n tends to infinity. The exact ARFIMA polynomial form of a general aggregated ARFIMA process, if any, remains undiscovered to the author knowledge.

3.2 – Spectral Density Function

Having stated Theorems 1 and 2, as well as Corollary 1, it is straightforward to calculate the spectrum of temporal aggregates of fractionally integrated processes based on the spectrum of the underlying process given by equation (5). Amongst the many existing memory parameter estimators, the most popular[3] are in the frequency-domain, as they do not suffer from the need to estimate the mean of the process since it influences only the spectrum at very low frequencies, which are commonly discarded. It is well known that the low rate of convergence of the mean of long memory processes turns its estimation imprecise for small to medium samples (a comprehensive discussion on the matter is found in Beran, 1994). The importance of

[3] For example, the Geweke and Porter-Hudak (1983) estimator, the Gaussian semi-parametric estimator of Robinson (1995) and the one based on maximizing the approximate frequency-domain Gaussian log-likelihood, following from Whittle (1951).

the spectral density in the long memory estimation is thus central. The spectral density function of aggregated long memory models is given in the following lemma:

Lemma 1: Let Assumption 1 hold so that X_t has the Wold representation $X_t = (1-B)^{-d}\sum_{i=0}^{\infty} w_i \varepsilon_{t-i}$. Then, for $-\pi < \lambda \le \pi$:

2a) If X_t is a stock variable, the spectral density function of the aggregated variable Y_t is given by:

$$f_y(\lambda) = \frac{\sigma_\varepsilon^2}{2n\pi}\sum_{i=0}^{n-1}\left[\left|1-e^{-j(\lambda+i2\pi)/n}\right|^{-2d} W(e^{-j(\lambda+i2\pi)/n})W(e^{j(\lambda+i2\pi)/n})\right] \quad (16)$$

where W(.) is defined in equation (11).

2b) If X_t is a flow variable, the spectral density function of the aggregated variable Y_t is given by:

$$f_y(\lambda) = \sigma_\varepsilon^2\sum_{i=0}^{n-1}\left[\left|1-e^{-j(\lambda+i2\pi)/n}\right|^{-2d} W(e^{-j(\lambda+i2\pi)/n})W(e^{j(\lambda+i2\pi)/n})F_n((\lambda+i2\pi)/n)\right] \quad (17)$$

Lemma 1 follows directly from Theorems 1 and 2 and Corollary 1 and gives the relationship between the spectral densities of the aggregated and the underlying processes. Note that the difference between aggregated stock and flow processes, apart from a multiplicative constant, is the Fejer kernel multiplying the spectral density of flow processes. As commented before in Section 3, it has a major peak at the frequency zero and zeros at all nonzero multiples of the Nyquist frequency ($2\pi/n$). Furthermore, it has the first derivative equal to zero at all (zero and nonzero) multiples of the Nyquist frequency. These frequencies will "alias" right onto the frequency zero. This ensures that the aliasing effect in the neighbourhood of the frequency zero is offset if aggregating a flow variable. On the other hand, if the process is a stock variable, the aliasing effect is felt in its full. If $d > 0$, however, the peak at the zero frequency still dominates the energy from other frequencies due to aliasing if there is no peak at the neighbourhood of them (which is not a restrictive assumption), keeping the same parameter d.

Even though the aggregation of stock variables retain the spectrum behaviour in a small neighbourhood of zero, the aggregation of flow variables do it in a far wider neighbourhood. These different frequency-domain behaviours of stock and flow variables will affect the (semiparametric) estimation of aggregated fractionally differenced processes based on the low-frequency periodogram ordinates. If the process is a flow variable, less bias is likely to be induced by aggregation, while if it is a stock variable, it is likely that aggregation will incur some bias. In particular, if d is negative, the aggregation of stock fractionally differenced processes, with its consequent aliasing effect, will destroy condition (2). However, for positive d, when the sample size increases the bias tend to disappear in both cases as a smaller band of frequencies is used for estimation.

These findings are largely consistent with Souza and Smith (2002) and Souza and Smith (2003). The former shows that aggregation as in Definition 3a does induce a non-negligible bias towards zero in the estimation of the integration order for $-0.5 < d < 0.5$. For $d < 0$ this bias has almost the same absolute value as d, turning it (almost) impossible to detect $d < 0$ after temporal aggregation of stock variables. They conjecture that this bias is due to the aliasing effect and provide a heuristic formula for it. The latter shows that aggregation as in Definition 3b does not induce any considerable bias in the long memory estimation.

The next section formalizes some implications of Lemma 1, namely the invariance (or not) of the memory parameter with respect to aggregation.

3.3 – Invariance of the memory parameter to aggregation

Chambers (1998) argues that the main implication of his Theorem 1, given in his Proposition 1, is that aggregation of fractionally integrated processes does not change the integration order. Even though I argue here that Chambers's (1998) Theorem 1 is wrong, its implications are still valid under certain conditions. If d is taken from the frequency-domain definition, for a negative d the invariance applies only for the flow type of aggregation, whereas for a positive d this applies for both types. On the other hand, if d is taken from the time-domain definition, the memory parameter remains unchanged after aggregation of either type. These results prove that the time- and frequency-domain conditions are not equivalent if $d < 0$, more specifically (1) does not imply (2) for $d < 0$, otherwise as stated by Robinson (1995, p.1632) for $-0.5 < d < 0$.

Although our results are valid only for covariance stationary processes, it is likely that the integration order remains untouched after aggregation of non-stationary ones, as it does in the case of unit root processes ($d = 1$). The results are given in the following propositions.

Proposition 2:

2a) If X_t satisfies equation (1) with $d < 0.5$ then its aggregated process Y_t also satisfies equation (1) with the same integration order d.

2b) If X_t is covariance stationary and satisfies equation (2) with $d > 0$ and has its spectral function finite and with finite first derivative in the neighbourhood of nonzero multiples of the Nyquist frequency ($2\pi/n$), then Y_t also satisfies equation (2) with the same integration order d.

2c) If X_t satisfies equation (2) with $d < 0$ and has its spectral function positive and finite and with finite first derivative in the neighbourhood of nonzero multiples of the Nyquist frequency, then Y_t satisfies equation (2) with the same integration order d if X_t is a flow variable but not if X_t is a stock variable.

Propositions 2a and 2c imply a secondary result for the study of long memory since it refers to negative values of d. It is given in Proposition 3 below.

Proposition 3:

Condition (1) does not imply (2) for negative values of d.

Proposition 2 says that covariance stationary fractionally integrated processes, when aggregated, should retain the integration order if d is positive or X_t is a flow variable. If d is negative and X_t a stock variable, the invariance in the memory parameter depends whether d is taken from condition (1) or (2). This property gives rise to Proposition 3, stating the non-equivalence of (1) and (2) for negative values of d. This result is of lesser practical importance because a negative d is rarely observed empirically, but arises frequently from overdifferencing a process. As practitioners usually aggregate series before differencing them and not otherwise, it is unlikely that a (stock) process with negative d will be aggregated.

4 – Simulation

In this Section a small simulation is carried out to compare the spectral density function derived here in Lemma 1 for aggregated long memory processes with that given by Theorem 1 of Chambers (1998). I argue here that in his paper Chambers (1998) did not take into account the aliasing effect for discrete processes[4] and that is why the formula derived here is different. A formula for the spectral function of aggregated (flow) ARFIMA processes is also given in Crato and Ray (2002). It is based on Tschernig (1995) and Teles, Wei and Crato (1999) results on aggregated ARFIMA models, basically that a flow ARFIMA(p,d,q) turns out to be an ARFIMA(p,d,∞) when aggregated. However, the MA(∞) polynomial is derived incompletely (the AR and I(d) polynomials are derived for non-overlapping aggregated processes while the MA polynomial given is that of an overlapping aggregated process), making the result of little use.

Figure 2 shows the spectral function derived in this paper for aggregated stock ARFIMA processes and the one derived in Chambers (1998), together with the periodogram ordinates averaged across 100 realizations of the process. The X axis shows the indices $i = 1, 2, \ldots, \lfloor T/2 \rfloor$ representing the Fourier frequencies $i2\pi/T$. Figure 3 does the same as Figure 2, but for flow processes. The processes are aggregated from ARFIMA(0,0.3,0), with $n = 3$; ARFIMA(1,0.3,0) with $\phi = 0.8$ and $n = 4$; ARFIMA(0,0.3,1) with $\theta = -0.8$ and $n = 4$; ARFIMA(1,0.3,1) with $\phi = -0.4$, $\theta = -0.8$ and $n = 3$. The aggregated series length is 512 observations and the error variance is taken as $\sigma_\varepsilon^2 = 1$.

As we can see for both stock and flow aggregated ARFIMA processes the averaged periodogram ordinates (dots) are scattered around the solid line, which represents the formula derived in this paper. The formula derived in Chambers (1998),

[4] He did take it into account to derive the spectral function of a continuous ARFIMA process observed at discrete-time intervals.

represented by a dashed line, yields values somewhat different from the observed in the simulation experiment.

5 – Real example

Chambers (1998) studies a number of UK macroeconomic flow series. He finds that the estimation using semiparametric methods is consistent with that the memory parameter remains unchanged after aggregation. The development of the theoretical results derived here suggests that his empirical results might be different if the variables were of the stock type. To verify the findings of the present paper, the daily US Dollar/ French Franc (US$/FF) exchange rate series is considered from October 20, 1977 to October 23, 2002 (25 years). More specifically, the natural logarithm of the squared returns is analysed. There are 68 (approximately 1.09%) zero returns existent in the 6264 workdays which were simply skipped, as well as the holidays. The series, its autocorrelation function (ACF) up to lag 1000 and its periodogram are shown in Figures 4-6, where the reader can notice the apparent long memory features such as persistently positive ACF (up to lag 250), and the periodogram scattered around a frequency power near the frequency zero.

The study of long-memory in the US Dollar/ French Franc exchange rate is not new in the literature. A number of authors have studied this property for different frequencies, time spans and series transformations. Booth, Kaen and Koveos (1982) and Cheung (1993) have investigated the presence of long memory respectively in the daily and weekly changes of the US Dollar/ French Franc exchange rate for different time spans as the one considered in this paper. Both find evidence of long memory. Vilasuso (2002) studies the squared returns of the daily US Dollar/ French Franc exchange rate, among other exchange rates, for a time span (1979-1999) similar to the one used in this paper. He finds long memory evidence, confirmed by better out-of-sample forecasts from a FIGARCH compared with GARCH and IGARCH modelling. Tschernig (1995) concludes that the long memory in changes of exchange rates is a phenomenon intimately linked to the US Dollar and possibly the American economy. Unlike these studies, which analyse long memory in the change (return) or squared return of exchange rates, the present analysis is consistent with the Long Memory Stochastic Volatility (LMSV) model of Breidt, Crato and Lima (1998), in which the returns are uncorrelated[5] and are given by the following relation:

$$R_t = \sigma \exp(Y_t / 2)\varepsilon_t , \tag{18}$$

where Y_t is a stationary Gaussian long memory process independent of ε_t, mean zero iid white noise. The analysed series is then:

$$Z_t \equiv \log(R_t^2) = \mu + Y_t + v_t , \tag{19}$$

where $\mu = (\log \sigma^2 + E[\log \varepsilon_t^2])$ and $v_t = (\log \varepsilon^2 - E[\log \varepsilon_t^2])$ is iid mean zero. Z_t is then a sum of a Gaussian long memory process and a white noise. The kurtosis of the series in study is approximately 3.68 and the skewness –0.79, so that the Jarque-Bera test rejects the hypothesis of Gaussianity at 1% confidence level. This does not mean that the Gaussianity of Y_t is rejected since it is contaminated by the noise v_t in the observed Z_t.

For the long memory study of the series, we shall utilize some semiparametric estimation methods. These are more adequate for the study than parametric methods, since it aims at the spectrum behaviour in the vicinity of the frequency zero as the

---

[5] If one wants to allow the returns to exhibit serial correlation, one may relax the assumption that ε_t is iid white noise in the model, but in this case the noise v_t is not iid. For example, one may imagine that (a long memory) exogeneity influences both the returns and the volatility of the exchange rate series and thus long memory are present in both as the literature seem to find for the US Dollar/ French Franc exchange rate and others. See Robinson and Zaffaroni (1998) for an alternative stochastic volatility model that allows for autocorrelated raw series.

arguments developed in Section 3.2 refer only to low frequency behaviour. The first method (GPH) was proposed by Geweke and Porter-Hudak (1983) and estimates *d* using property (2). Taking the log of (2) on both sides,

$$\log I(\lambda_j) = \log c_f(\lambda) - 2d[\log \lambda_j] + \varepsilon_j \qquad (20)$$

for frequencies near zero, where $I(\lambda_j)$ is the periodogram. Defining j = 1, …, g(T), where $g(T) = T^{\alpha}$ and $\alpha \in (0, 1)$, ordinary least squares regression on the set of Fourier frequencies λ_j gives the estimate of *d*. Hurvich, Deo and Brodsky (1998) provide the conditions which ensure the consistency of this estimator and Deo and Hurvich (2001) prove that the asymptotic distribution of this estimator for Z_t is identical as for Y_t provided a restriction in the bandwidth choice is observed.

The second method, denoted here by SMGPH, was proposed by Hassler (1993) and is similar to the GPH but uses a smoothed estimator for the spectrum (smoothed periodogram) instead of the raw periodogram. The third and last method, denoted by GSPR, is referred to as the Gaussian semi-parametric approach of Robinson (1995). This method is based on maximising the approximate form of the frequency domain Gaussian likelihood, where discrete averaging is carried out over a neighbourhood of zero frequency:

$$R(d) = \log\left(\frac{1}{m}\sum_{j=1}^{m}\lambda_j^{2d} I_j\right) - \frac{2d}{m}\sum_{j=1}^{m}\log(\lambda_j) \qquad (21)$$

where $m=T^{\alpha}$ and $\alpha \in (0, 1)$. Robinson (1995) outlines the conditions under which the estimator is consistent.

The series Z_t displays long memory, as the modified R/S test of Lo (1991) rejects the hypothesis of short memory at a 0.5% level[6]. Moreover, breaking up the

---

[6] Teverovsky et al (1999) has shown that this test is conservative in that it is biased towards accepting the null of no long memory.

series in five smaller sub-series of 1240 days[7] yields estimates of the memory parameter consistent with the presence of long memory, as shown in Table 1.

The series Z_t is of stock type and is related to the instant volatility of the returns, so that its aggregation would give a picture of this volatility measure once in a pre-determined period, say, a week. Note that this measure of volatility is not the usual squared returns[8] but the log of it. Also, as proposed by Crato and Ray (2002), one might aggregate the series Z_t as a flow variable so as to decrease the signal-to-noise ratio with estimation purposes. Remember that Z_t is a long memory process with added noise. Table 2 compares the estimates for the original and the aggregated series considering $n = 5$, with the aggregated time unit roughly corresponding to one week. There are n possible (different) aggregated series for each case (stock or flow), corresponding to the original series commencing at time $t = 1, 2, \ldots, n$. All five are shown for each case and the same bandwidth choices as in Table 1, $\alpha = 0.5, 0.6$. It is also shown their average and standard deviation across different commencing times in the original series.

It is apparent in Table 2 that the long memory estimates of the flow aggregates are of the same order as the original series, whereas the stock aggregates have estimates somewhat lower. The GPH yields smaller estimates for the flow aggregates, but higher on average than their stock counterparts (although in the first column, corresponding to $\alpha = 0.5$, this is due only to the fifth stock aggregate estimate). One should note that, by the nature of the aggregation type, the standard deviation of the flow aggregates estimates is much lower than that of stock aggregates. This is consistent with the model, which says that Z_t is a long memory process with added

[7] The five series correspond respectively to observations 1-1240; 1241-2480; 2481-3720; 3721-4960; and 4961-6196

[8] Andersen and Bollerslev (1998) propose the use of the "realized volatility" (the sum of squared returns at a higher frequency to get an estimate of the volatility for a lower frequency) as a much less noisier proxy for the volatility than the squared returns.

noise, together with the findings of Crato and Ray (2002), who shows that flow aggregation reduces the signal-to-noise ratio, improving the long memory estimation. The example is very illustrative of a result proposed in this paper, namely that stock aggregates should be more biased in respect to long memory than flow aggregates, although the stock and flow aggregation of the same series may seem arbitrary. The clear picture of the result was made possible mainly because of the large amount of observations in the original series.

6 – Concluding Remarks

This paper derives the spectral density function of temporally aggregated long memory processes in light of the aliasing effect. The aggregation process is considered here the fact of observing the process at a slower sampling rate, giving rise to distinct interpretations: if the process is a stock variable, every n-th observation is considered, while if it is a flow variable, a sum of every n-th and their n-1 preceding observations is considered. Temporal aggregation is shown to maintain the integration order (the memory parameter d), with the exception of stock aggregates when $d < 0$ in the original series if the frequency-domain definition of d is considered (the integration order is always retained in the time-domain definition). A small Monte Carlo experiment is run, showing that the formulas for the spectral density function derived here match with the periodogram of synthetic series averaged across series, whereas some previous papers that addressed the subject do not provide accurate formulas.

From the spectral function of aggregated process derived here, I infer that flow aggregates from long memory processes shall have the estimates of d less biased than stock aggregates. This result is consistent with Souza and Smith (2002a, b) Monte

Carlo results and is illustrated in Section 5 by an example with the daily US Dollar/ French Franc exchange rate series.

Acknowledgements

The author would like to thank FAPERJ for the financial support, EPGE for its kind hospitality and Marcelo Fernandes for helpful comments on previous versions of this work.

Appendix

Proof of Theorem 1

The spectral density function of a covariance stationary discrete-time process X_t is defined by:

$$f_x(\omega) = \frac{1}{2\pi}\sum_{k=-\infty}^{\infty}\gamma_k^x e^{-jk\omega}\,, \qquad -\pi < \omega \leq \pi \tag{A1}$$

where γ_k^x is the k-th order autocovariance of X_t. As the autocovariance function of real valued processes is an even function, (A1) reduces to:

$$f_x(\omega) = \frac{1}{2\pi}\sum_{k=-\infty}^{\infty}\gamma_k^x \cos k\omega\,, \qquad -\pi < \omega \leq \pi \tag{A2}$$

$f_x(\omega)$ is then defined as a Fourier cosine series whose coefficients are the autocovariances of X_t. As $\cos k\omega$, k = 0, 1, 2, ..., is a complete orthogonal set over the interval (–π, π] for even functions (and the spectral density is an even function) the relation given by (A2) is equivalent to:

$$\gamma_k^x = \int_{-\pi}^{\pi} f_x(\omega)\cos k\omega\, d\omega = \int_{-\pi}^{\pi}\cos k\omega\, dF_x(\omega), \qquad k = 0, \pm 1, \pm 2, ... \tag{A3}$$

where $F_x(\omega)$ is the spectral distribution function of X_t. The autocovariances of $Y_t = X_{nt}$ are given thus by:

$$\gamma_k^y = \gamma_{nk}^x = \int_{-\pi}^{\pi} f_x(\omega)\cos nk\omega\, d\omega = \int_{-\pi}^{\pi}\cos nk\omega\, dF_x(\omega), \qquad k = 0, \pm 1, \pm 2, ... \tag{A4}$$

First take the simpler case where *n* is an odd number. The integral in (A4) can be split into:

$$\gamma_k^y = \sum_{i=-\frac{n-1}{2}}^{\frac{n-1}{2}} \int_{(2i-1)\pi/n}^{(2i+1)\pi/n} \cos nk\omega \, dF_x(\omega) =$$

$$= \sum_{i=-\frac{n-1}{2}}^{\frac{n-1}{2}} \int_{-\pi/n}^{\pi/n} \cos(nk\omega + 2ik\pi)\, dF_x(\omega + 2i\pi/n) \qquad (A5)$$

Since cos(a + 2iπ) = cos(a), where i is an integer number, *(A5)* rewrites to:

$$\gamma_k^y = \sum_{i=-\frac{n-1}{2}}^{\frac{n-1}{2}} \int_{-\pi/n}^{\pi/n} \cos(nk\omega)\, dF_x(\omega + 2i\pi/n) \qquad (A6)$$

Making $\lambda = n\omega$ where λ is the frequency measured in the same time unit as Y_t, we can write:

$$\gamma_k^y = \sum_{i=-\frac{n-1}{2}}^{\frac{n-1}{2}} \int_{-\pi}^{\pi} \cos(k\lambda)\frac{1}{n}\, dF_x(\lambda/n + 2i\pi/n) \qquad (A7)$$

However, by (A3) we can write the k-th order autovariance of Y_t as:

$$\gamma_k^y = \int_{-\pi}^{\pi} f_y(\lambda)\cos k\lambda \, d\lambda = \int_{-\pi}^{\pi} \cos k\lambda \, dF_y(\lambda), \qquad k = 0, \pm1, \pm2, \ldots \qquad (A8)$$

The fact that cos $k\lambda$, k = 0, 1, 2, ..., is a complete orthogonal set over the interval (–π, π] for even functions, together with (A7) and (A8) imply (6).

Now if n is an even number (A5) rewrites to:

$$\gamma_k^y = \sum_{i=1-\frac{n}{2}}^{\frac{n}{2}} \int_{-\pi/n}^{0} \cos(nk\omega + 2ik\pi)\, dF_x(\omega + 2i\pi/n) +$$

$$+ \sum_{i=-\frac{n}{2}}^{\frac{n}{2}-1} \int_{0}^{\pi/n} \cos(nk\omega + 2ik\pi)\, dF_x(\omega + 2i\pi/n) \qquad (A9)$$

and the rest of the proof follows as in the case n is odd.

Proof of Theorem 2

Let $Z_t = \sum_{i=0}^{n-1} B^i X_t$ be the overlapping aggregated process of X_t. The moving average representation of $(1 + B + \ldots + B^{n-1})$ straightforwardly gives the following relationship between the spectra of Z_t and X_t:

$$f_z(\lambda) = f_x(\lambda)\left|\sum_{k=0}^{n-1} e^{-jk\lambda}\right|^2, \qquad -\pi < \lambda \le \pi \qquad (A10)$$

The quantity $\left|\sum_{k=0}^{n-1} e^{-jk\lambda}\right|^2$ rewrites as:

$$\left|\sum_{k=0}^{n-1} e^{-jk\lambda}\right|^2 = \left|\sum_{k=0}^{n-1} \cos k\lambda - j\sin k\lambda\right|^2 = \left|\sum_{k=0}^{n-1} \cos k\lambda\right|^2 + \left|\sum_{k=0}^{n-1} \sin k\lambda\right|^2 =$$

$$= \sum_{k=0}^{n-1}\left(\cos^2 k\lambda + \sin^2 k\lambda\right) + 2\sum_{k=0}^{n-1}\sum_{i=k+1}^{n-1}\left(\cos k\lambda \cos i\lambda + \sin k\lambda \sin i\lambda\right) \qquad (A11)$$

Trigonometric manipulation on the RHS of (A11) gives

$$\left|\sum_{k=0}^{n-1} e^{-jk\lambda}\right|^2 = n + 2\sum_{k=0}^{n-1}\sum_{i=k+1}^{n-1}\cos[(i-k)\lambda] = \sum_{k=1-n}^{n-1}(n-|k|)\cos k\lambda \qquad (A12)$$

Using equation (6.1.43) of Priestley (1981, p. 400), which is based on further trigonometric manipulation, in (A12) and correcting for the case $\lambda = 0$, gives

$$\left|\sum_{k=0}^{n-1} e^{-jk\lambda}\right|^2 = \lim_{\theta\to\lambda}\frac{\sin^2(n\theta/2)}{\sin^2(\theta/2)} = 2\pi n.F_n(\lambda), \qquad (A13)$$

where $F_n(\lambda)$ is the Fejer kernel. Equations (A10) and (A13) imply equation (8) and the proof is complete.

Proof of Proposition 1

The proof of 1a) is straightforward since Assumption 1 assures that X_t is covariance stationary and hence $\sum_{i=0}^{\infty} a_i^2 < \infty$. Requiring the covariance stationarity of X_t in Assumption 1 is necessary because if, for example, $(1-B)^{d^*} X_t = \varepsilon_t$, where $1 > d^* > 0.5$, the representation $(1-B)^{d^*/2} X_t = \sum_{i=0}^{\infty} w_i \varepsilon_{t-i} = (1-B)^{-d^*/2} \varepsilon_t$ would satisfy the Wold representation in (11) but $\sum_{i=0}^{\infty} a_i^2$ would diverge.

Proof of 1b): By the Definition 3b) $Y_t = \sum_{i=1}^{n} B^i X_{nt}$. Let $Z_t = \sum_{i=1}^{n} B^i X_t$ so that $Y_t = Z_{nt}$.

Since $X_t = \sum_{i=0}^{\infty} a_i \varepsilon_{t-i}$ then $Z_t = \sum_{k=0}^{n-1} B^k \sum_{i=0}^{\infty} a_i \varepsilon_{t-i} = \sum_{i=0}^{\infty} b_i \varepsilon_{t-i}$, where $b_i = \sum_{k=i-n+1}^{i} a_k$ and $a_k = 0$ for $k < 0$. It remains to prove that $\sum_{i=0}^{\infty} b_i^2 < \infty$ and then the proof of 1b) is similar the proof of 1a). This condition is equivalent to say that the variance of Z_t is finite. Since X_t has finite variance, $VAR(Z_t) \le n^2\ VAR(X_t)$ and is finite. The proof is complete.

Proof of Lemma 1

2a) The proof of 2a follows directly from Theorem 1. The change in the summation indices is to simplify the formula. Each of the indices "i" in the summations of (6) and (7) has one and only one equivalent index in {0, 1, …, n-1} for $e^{-j(\lambda+i2\pi)/n}$ and all summations have *n* different terms. Suppose *n* is even. For example, $e^{-j(\lambda-(n/2)2\pi)/n}$ is indistinguishable from $e^{-j(\lambda+(n/2)2\pi)/n}$, that is, the results for i = -n/2 and i = n/2 are the same.

2b) The proof of 2b follows directly from Corollary 1, using the same change in the summation indices as in 2a.

Proof of Proposition 2

2a) If X_t is a stock variable, the autocorrelations of the aggregated process behave like

$\rho_k^y = \rho_{nk}^x \sim c_{\rho x}(nk)(nk)^{2d-1} = c_{\rho y}(k)k^{2d-1}$ as k tends to infinity, which satisfies (1).

If X_t is a flow variable, we only need to prove that the overlapping aggregated

variable $Z_t = \sum_{i=1}^{n} B^i X_t$ satisfies (1). Hence, using the fact that $Y_t = Z_{nt}$, the remaining

of the proof is identical to the proof for a stock variable. By equation (1), $\gamma_k^x \sim \sigma_x^2$

$c_\rho(k)k^{2d-1}$ as $k \to \infty$. For the sake of simplicity assume $E[X_t] = 0$. So, $\gamma_k^x = E[X_t X_{t+k}]$.

The autocovariances of Z_t ($k \geq 0$) are given then by

$$\gamma_{-k}^z = \gamma_k^z = E[Z_t Z_{t+k}] = E\left[\sum_{i=0}^{n-1} X_{t-i} \sum_{i=0}^{n-1} X_{t+k-i}\right] = \sum_{i=-n+1}^{n-1} (n - |i|)\gamma_{k+i}^x . \qquad (A14)$$

Take the Taylor series expansion of $(k+i)^{2d-1}$ around k^{2d-1}:

$$(k+i)^{2d-1} = k^{2d-1} + i.(2d-1)\, k^{2d-2} + (i^2/2)(2d-1)(2d-2)\, k^{2d-3} + O(k^{2d-4}). \qquad (A15)$$

If the right hand side of equation (A14), substituting (1) for γ_{k+i}^x, is expanded in Taylor series of this type (A15), the series second terms cancel out in the summation and hence $\gamma_k^z \sim n^2\sigma_x^2\, c_\rho(k)k^{2d-1} + O(k^{2d-3})$ as $k \to \infty$, and consequently $\rho_k^z \sim c_{\rho z}(k)k^{2d-1}$. The proof of 2a is complete.

2b) If X_t is a covariance stationary stock variable and $\lambda \geq 0$, Theorem 1 says that

$f_y(-\lambda) = f_y(\lambda) = \frac{1}{n} \sum_{i=-\lfloor n/2 \rfloor}^{\lfloor (n-1)/2 \rfloor} f_x([\lambda + i2\pi]/n)$, where $\lfloor . \rfloor$ is the integer operator. As

$f_x(i2\pi/n)$, $i \neq 0$, is finite and has finite neighbourhood, and X_t satisfies Definition (2), the term in the summation corresponding to $i = 0$ dominates the others as $\lambda \to 0$. The spectral function of Y_t thus satisfies:

$f_y(\lambda) \sim (1/n).c_{fx}(\lambda/n)\, |\lambda/n|^{-2d} = n^{2d-1}.c_{fx}(\lambda/n)\, |\lambda|^{-2d}$ as $\lambda \to 0$.

Now if X_t is a flow variable, then $f_y(\lambda) = 2\pi \sum_{i=-\lfloor n/2 \rfloor}^{\lfloor (n-1)/2 \rfloor} \left[f_x([\lambda + i2\pi]/n).F_n([\lambda + i2\pi]/n) \right]$

as stated in Corollary 1. The Fejer kernel $F_n(.)$ and its first derivative are zero valued in the nonzero multiples of the Nyquist frequency $2\pi/n$ and hence the term corresponding to i = 0 in the summation dominates the others however value d takes. Furthermore $F_n(0) = n/2\pi$ so that the spectrum of Y_t behaves like: $f_y(\lambda) \sim n.f_x(\lambda/n) \sim n^{2d+1}.c_{fx}(\lambda/n)\,|\lambda|^{-2d}$ as $\lambda \rightarrow 0$, satisfying (2) with the same integration order d. The proof of 2b is complete.

2c) The proof that condition (2) with a negative memory parameter d still holds after flow aggregation is given in the proof of 2b. It does not apply for stock aggregation because if d < 0, the term corresponding to i = 0 in the summation in (16) is dominated by the other positive terms, as $\lim_{\lambda\rightarrow 0} |\lambda|^{-2d} = 0$ and condition (2) is thus violated. The proof of 2c is complete.

Proof of Proposition 3

Suppose (1) and (2) hold for a stock variable X_t with d < 0, for example X_t might be an ARFIMA(0,d,0), d < 0. Take the aggregated variable Y_t. By Proposition 2a condition (1) still holds for Y_t while by 2c condition (2) does not hold for Y_t. Hence condition (1) does not imply condition (2) for d < 0. The spectrum finiteness at the nonzero multiples of the Nyquist frequency is a dispensable (maybe redundant) assumption for Proposition 3.

References

Amemiya, T. and R. Y. Wu, 1972, The effect of aggregation on prediction in the autoregressive model, Journal of the American Statistical Association 67, 628-632.

Andersen, T.G. and T. Bollerlev, 1998, Answering the skeptics: yes, standard volatility models do provide accurate forecasts, International Economic Review 39, 885-905.

Beran, J., 1994, Statistics for Long Memory Processes (Chapman & Hall, London).

Beran, J. and D. Ocker, 2000, Temporal aggregation of stationary and nonstationary FARIMA(p,d,0) models, working paper.

Bisaglia, L. and D. Guégan, 1998, A comparison of techniques of estimation in long-memory processes, Computational Statistics & Data Analysis 27, 61-81.

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

Breidt, F.J., N.Crato and P. Lima, 1998, The detection and estimation of long memory in stochastic volatility, Journal of Econometrics 83, 325-348.

Brewer, K.R.W., 1973, Some consequences of temporal aggregation and systematic sampling for ARMA and ARMAX models, Journal of Econometrics 1, 133-154.

Chambers, M. J., 1998, Long memory and aggregation in macroeconomic time series, International Economic Review 39, 1053-1072.

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

Crato, N. and B. K. Ray, 2002, Semi-parametric smoothing estimators for long-memory processes with noise, Journal of Statistical Planning and Inference 105, 2, 283-297.

Deo, R. S. and C. M. Hurvich, 2001, On the log periodogram regression estimator of the memory parameter in long memory stochastic volatility models, Econometric Theory 17, 4, 686-710.

Diebold, F. X. and G. D. Rudebusch, 1989, Long Memory and Persistence in Aggregate Output, Journal of Monetary Economics 24, 189-209.

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

Geweke, J. and S. Porter-Hudak, 1983, The estimation and application of long memory time series models, Journal of Time Series Analysis 4, 221-237.

Hamilton, J. D., 1994, Time Series Analysis (Princeton University Press, New Jersey).

Hassler, U., 1993, Regression of spectral estimators with fractionally integrated time series, Journal of Time Series Analysis 14, 369-380.

Hosking, J., 1981, Fractional differencing, Biometrika 68, 1, 165-176.

Hotta, L. K. and J. C. Neto, 1993, The effect of aggregation on prediction in autoregressive integrated moving-average models, Journal of Time Series Analysis 14, 261-269.

Hurvich, C. M., R. Deo and J. Brodsky, 1998, The mean square error of Geweke and Porter-Hudak's estimator of the memory parameter of a long-memory time series, Journal of Time Series Analysis 19, 19-46.

Koopmans, L.H., 1974, The Spectral Analysis of Time Series (Academic Press, New York).

Lo, A.W., 1991, Long-term memory in stock market prices, Econometrica 59, 5, 1279-1313.

Lütkepohl, H., 1986, Comparison of predictors for temporally and contemporaneously aggregated time series, International Journal of Forecasting 2, 461-475.

Mandelbrot, B. B., 1965, Une classe de processus stochastiques homothétiques à soi; application à la loi climatologique de H. E. Hurst, Comptes Rendus Acad. Sci. Paris 260, 3274-3277.

Mandelbrot, B. B. and J. W. Van Ness, 1968, Fractional Brownian motion, fractional noises and applications, SIAM Review 10, 422-437.

McLeod, A.I. and K.W. Hipel, 1978, Preservation of the rescaled range I. A reassessment of the Hurst phenomenon, Water Resources Research 14, 491-508.

Oppenheim, A. V. and R. W. Schafer, 1989, Discrete-Time Signal Processing, (Prentice-Hall, New Jersey).

Otero, J. and J. Smith, 2000, Testing for cointegration: power versus frequency of observation - further Monte Carlo results, Economics Letters 67, 1, 5-9.

Priestley, M. B., 1981, Spectral Analysis and Time Series (Academic Press, London).

Robinson, P. M., 1995, Gaussian semi-parametric estimation of long range dependence, Annals of Statistics 23, 5, 1630-1661.

Robinson, P. M. and P. Zaffaroni, 1998, Nonlinear time series with long memory: a model for stochastic volatility, Journal of Statistical Planning and Inference 68, 359-371.

Souza, L. R. and J. Smith, 2002, Bias in the memory parameter for different sampling rates, International Journal of Forecasting 18, 299-313.

Souza, L. R. and J. Smith, 2003, Effects of temporal aggregation on estimates and forecasts of fractionally integrated processes: A Monte Carlo study, International Journal of Forecasting, forthcoming.

Teles, P. and W. W. S. Wei, 2002, The use of aggregate time series in testing for Gaussianity, Journal of Time Series Analysis 23, 1, 95-116.

Teles, P., W. W. S. Wei and N.Crato, 1999, The use of aggregate series in testing for long memory, Bulletin of the International Statistical Institute, 52nd Session Book 3, 341-342.

Teverovsky, V., M.S. Taqqu and W. Willinger, 1999, A critical look at Lo's modified R/S statistic, Journal of Statistical Planning and Inference 80, 211-227.

Tschernig, R., 1995, Long memory in foreign exchange rates revisited, Journal of International Financial Markets, Institutions and Money 5, 53-78.

Vilasuso, J., 2002, Forecasting exchange rate volatility, Economics Letters 76, 59-64.

Wei, W. W. S., 1989, Time series analysis: univariate and multivariate methods (Addison-Wesley, Boston, MA).

Weiss, A. A., 1984, Systematic sampling and temporal aggregation in time series models, Journal of Econometrics 26, 271-281.

Whittle, P., 1951, Hypothesis testing in time series analysis, (Hafner, New York).

Table 1: Estimates of the memory parameter with the three semiparametric estimators for the whole series Z_t and for its sub-series, considering the bandwidth parameter $\alpha = 0.5$ and 0.6.

	$\alpha = 0.5$			$\alpha = 0.6$		
	GPH	SMGPH	GSPR	GPH	SMGPH	GSPR
original series	0,309	0,338	0,333	0,309	0,298	0,299
1st sub-s.	0,274	0,366	0,372	0,318	0,421	0,393
2nd sub-s.	0,285	0,257	0,284	0,114	0,158	0,182
3rd sub-s.	0,468	0,393	0,401	0,407	0,355	0,322
4th sub-s.	0,318	0,285	0,288	0,265	0,293	0,289
5th sub-s.	0,357	0,245	0,280	0,139	0,083	0,118

Table 2: Estimates of the memory parameter with the three semiparametric estimators for the whole series Z_t and for the aggregated flow and stock series, considering the bandwidth parameter $\alpha = 0.5$ and 0.6.

	$\alpha = 0.5$			$\alpha = 0.6$		
	GPH	SMGPH	GSPR	GPH	SMGPH	GSPR
orig. series	0,309	0,338	0,333	0,309	0,298	0,299
1^{st} flow agg.	0,252	0,324	0,338	0,268	0,296	0,300
2^{nd} flow agg.	0,260	0,325	0,339	0,281	0,307	0,309
3^{rd} flow agg.	0,259	0,328	0,347	0,276	0,306	0,309
4^{th} flow agg.	0,256	0,326	0,347	0,259	0,301	0,304
5^{th} flow agg.	0,254	0,327	0,347	0,271	0,301	0,304
average	0,256	0,326	0,344	0,271	0,302	0,305
st. dev.	0,003	0,002	0,005	0,008	0,004	0,004
1^{st} stock agg.	0,255	0,250	0,225	0,253	0,217	0,223
2^{nd} stock agg.	0,280	0,201	0,250	0,320	0,216	0,216
3^{rd} stock agg.	0,240	0,146	0,217	0,181	0,159	0,164
4^{th} stock agg.	0,239	0,310	0,331	0,167	0,196	0,212
5^{th} stock agg.	0,000	0,093	0,124	0,112	0,151	0,143
average	0,203	0,200	0,229	0,206	0,188	0,192
st. dev.	0,115	0,085	0,074	0,081	0,031	0,036

Figure 1: The Fejer kernel for n = 6, restricted to $(-\pi, \pi]$

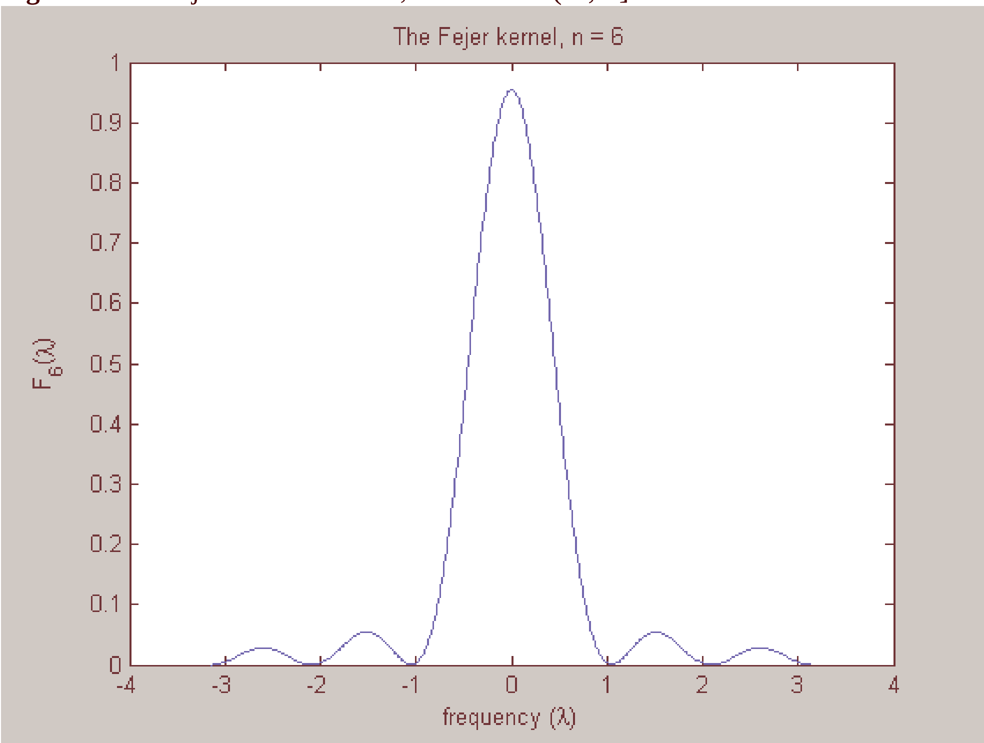

Figure 2: Comparison between Chambers's (1998) and this paper's theoretical spectral functions for aggregated stock ARFIMA processes, respectively the dashed and the continuous lines. The dots are the periodogram ordinates averaged across 100 realizations of the processes.

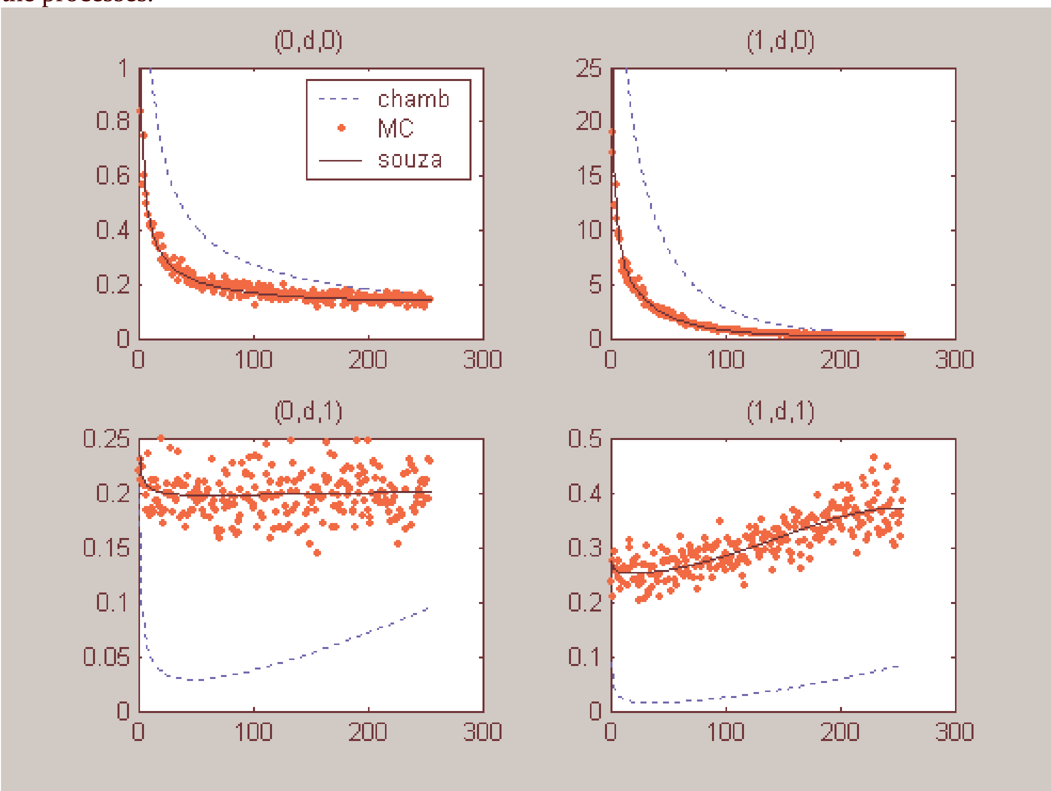

Figure 3: Comparison between Chambers's (1998) and this paper's theoretical spectral functions for aggregated flow ARFIMA processes, respectively the dashed and the continuous lines. The dots are the periodogram averaged across 100 realizations of the processes.

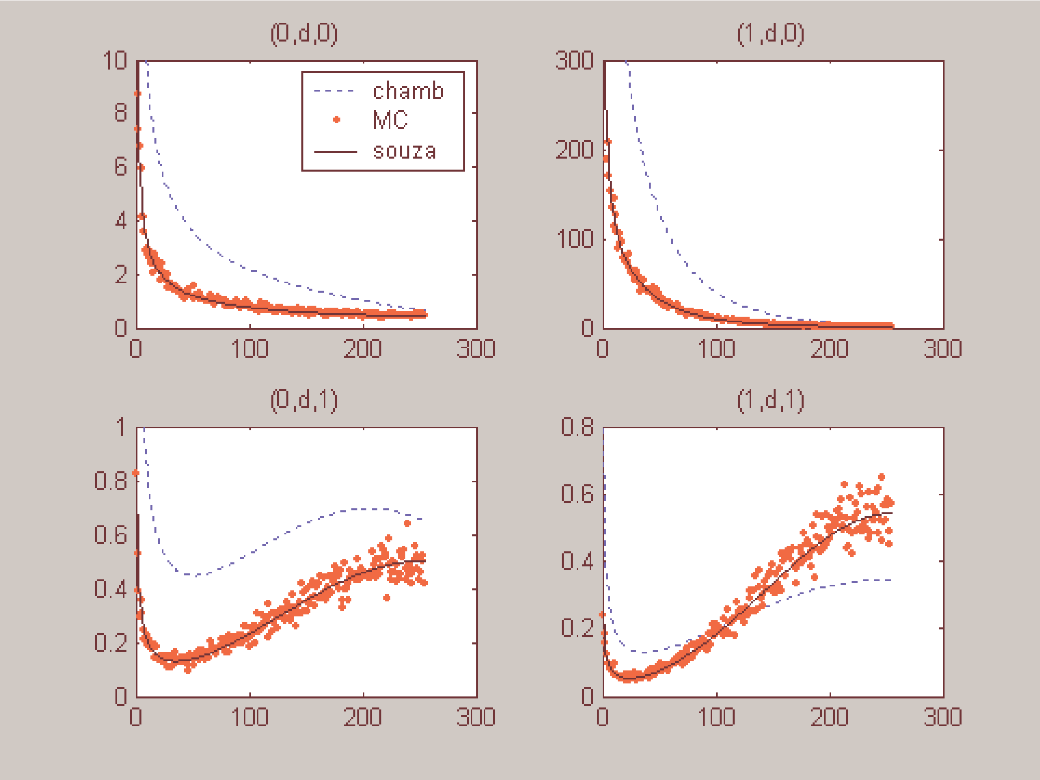

Figure 4: US$/FF exchange rate, logarithm of the squared returns from October 20, 1977 to October 23, 2002. The series.

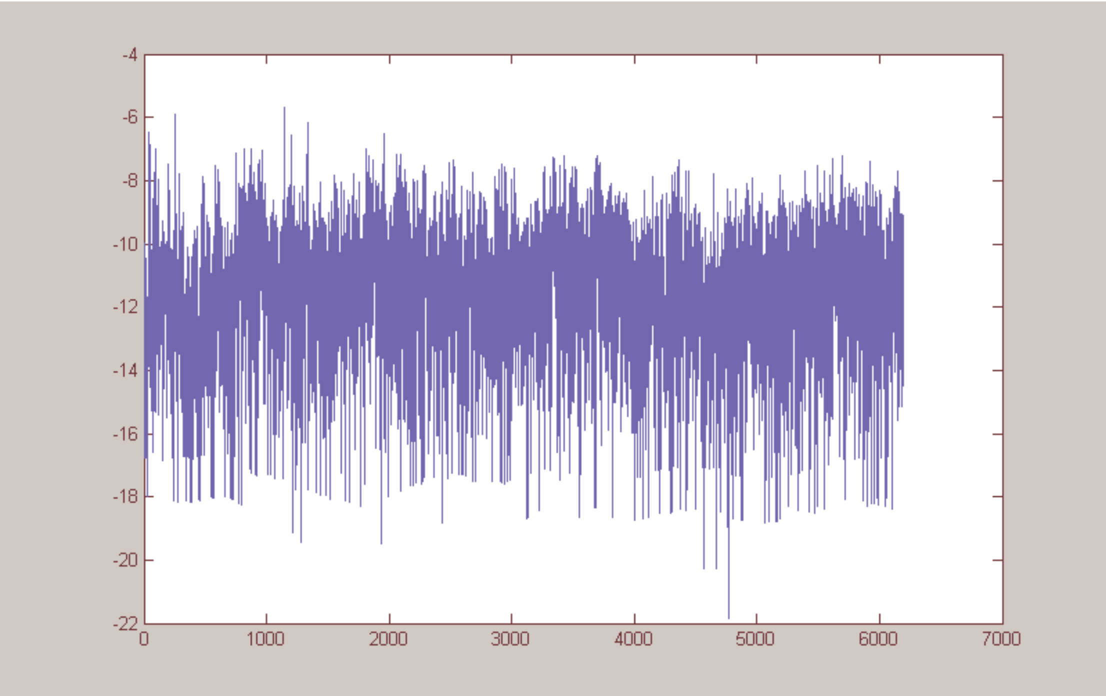

Figure 5: US$/FF exchange rate, logarithm of the squared returns from October 20, 1977 to October 23, 2002. ACF up to lag 1000.

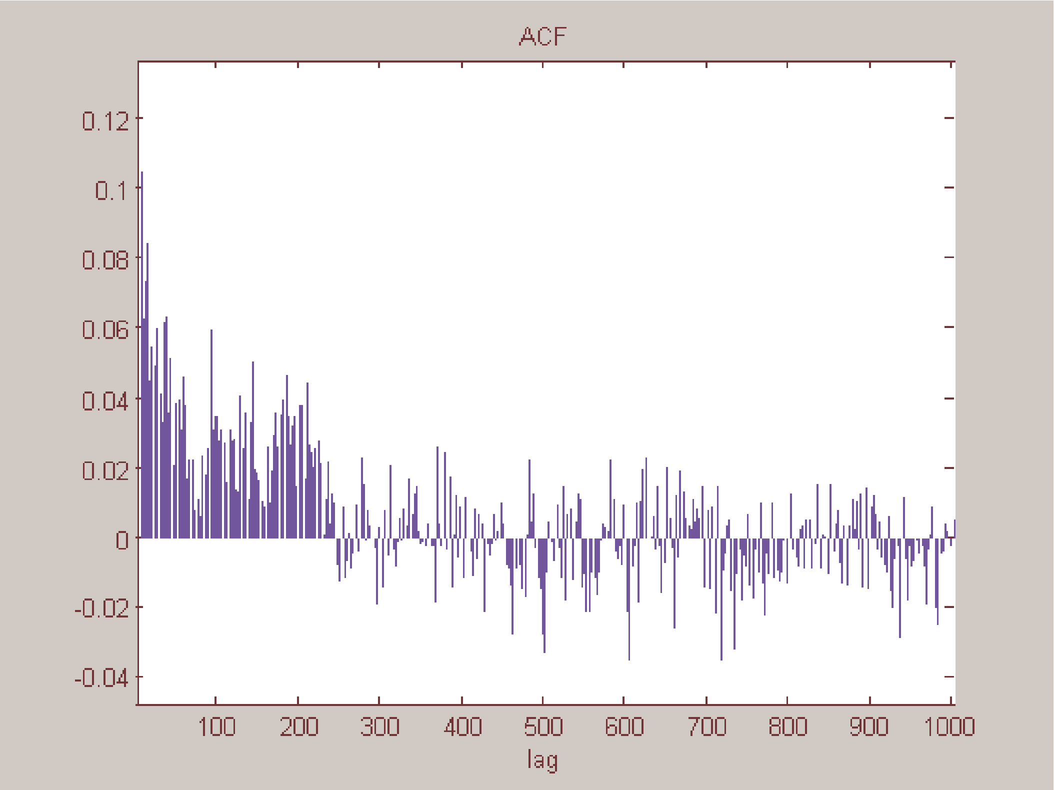

Figure 6: US$/FF exchange rate, logarithm of the squared returns from October 20, 1977 to October 23, 2002. Periodogram ordinates in log-log scale.