Applied Financial Economics, 1994, **4**, 33–39

Some international evidence regarding the stochastic memory of stock returns

NUNO CRATO

Department of Pure and Applied Mathematics, Stevens Institute of Technology, Hoboken, NJ 07030, USA and *Institute of Economics of The Technical University of Lisbon, Portugal*

The present paper studies international stock indexes of the G-7 countries in the last 40 years. Evidence about the statistical memory of the returns is presented, and only in one country could the existence of long memory be sustained. These results contradict various previous studies that were based on the *R/S* analysis and consistently claimed the existence of long memory in financial returns. A general ARFIMA model capable of reproducing long- and short-memory properties is directly fitted to the data. The conclusion is then based on the estimated parameters of the model.

I. INTRODUCTION

The evaluation of the statistical memory of stock returns has always been an attractive topic in theoretical and practical economics. The question about the existence of that statistical memory is directly related to the one off stock price predictability, and the common answer that is taught in finance courses is the one from the efficient market hypothesis theory. There are many formulations of this hypothesis (see, e.g. Ball, 1989), all leading to the conclusion that consistent increased profits cannot be made by trading on the market on the basis of information given by past prices.

This classic answer from finance theory has recently been challenged by various empirical studies using novel and more sophisticated statistical tools. A basic reference is Guimarães *et al.* (1989), and a recent survey from the forecaster point of view is Granger (1992).

In this paper a particular issue related to the efficient market hypothesis is discussed: the claim made by some researchers that stock markets have a specific type of memory behaviour called persistence, or long memory, or long-range dependence.

The claim of long memory on stock prices is a surprising one. In a certain way, it goes further than the simple claim of some sort of statistical dependence that could improve the predictability of the prices: it states that correlations between price changes die out very slowly, in a sense made precise below, being the actual movements in the market stochastically influenced by the recent to the most remote past.

This claim was suggested by the seminal work of Mandelbrot (1965, 1972) on the fractional Brownian motion and on the celebrated *R/S* analysis. Several researchers used the tools developed by Mandelbrot and, by means of *R/S* analysis, arrived at the surprising conclusion that some financial time series had a long-memory behaviour.

Recently, Lo (1991) contested that analysis, showing that the *R/S* statistic was biased and unable to distinguish, in finite samples, short-memory characteristics from long-memory characteristics of a time series. Lo developed a modified *R/S* statistic that overcomes these problems and applied it to various US indexes of stock returns. His conclusions contradicted previous research and imputed the results of preceding *R/S* analysis to the sole existence of short- memory in security prices.

Here a different approach is presented where a class of discrete-time long-memory models is fitted to the data, and the compatibility of the fitted parameters with a short-memory assumption is tested. The class of models used is the autoregressive fractionally integrated moving average (ARFIMA), a generalization, introduced by Granger and Joyeux (1980) and by Hosking (1981), of the well-known ARIMA models of Box–Jenkins. The main estimation procedure used here is the new exact maximum-likelihood approach of Sowell (1991). This way, the problem is readdressed from a new and more pragmatic angle: the question is whether or not if the large class of linear long-memory

0960–3107

models that generalizes the ARIMA processes needs a 'long-memory parameter' to better fit the data.

Also, the data is perhaps more representative of stock market price behaviour in modern developed economies. A long span of data of almost 40 years is used for the G-7 countries: Canada, France, Italy, Japan, UK, US and what was West Germany.

The plan of the paper is as follows. Section II defines the concepts of short and long memory of a random process. Section III critically reviews previous work claiming the existence of long memory in financial time series. Section IV presents a fractional model with long-memory behaviour, the ARFIMA model. Section V describes the data, the available estimation procedures, and discusses the results of the estimations. Section VI concludes.

II. THE MEMORY OF A TIME SERIES

Let the time series under consideration be represented by (X_t). In what follows, it is assumed that the time series are of second order, i.e. that their variance is finite for every t. It is also assumed that the time series, directly or after some transformation, are stationary, i.e. their first and second moments are independent of t: EX_t is constant and $\mathrm{Cov}\,[X_t, X_{t+h}]=:\gamma(h)$, is a function of the lag h alone. Also, ε_t will denote a white-noise process with mean zero and variance $\sigma^2<\infty$, in symbols $\varepsilon_t \sim WN(0, \sigma^2)$.

Time series studied in economics and finance have a characteristic economic behaviour in that their autocovariances decay to zero, reflecting the fact that the influence of the past values decreases with the lags under consideration. The speed of that decay is a measure of the internal memory of the random event.

If a white noise is the appropriate model then the random event is said to have no memory. It can be said that such an extreme case seldom occurs in practice, but one commonly assumed implication of the efficient market hypothesis is that stock returns are such a phenomenon.

Models with short memory are, for instance, the so-called auto regressive moving average (ARMA) ones. The rationale is that the decay of the autocovariances of an ARMA is geometrically bounded: there exist constants $C>0$ and r in the open interval $]0, 1[$ such that $|\gamma(h)| \leqslant Cr^{|h|}$. Hence, $\gamma(h)$ is absolutely summable.

In contrast, long-memory models have autocovariances that decay much more slowly, following asymptotically a hyperbolic decay. More precisely, a second-order stationary process (X_t) would be called of long-memory type if, for some $C>0$ and $\alpha<0$, its autocovariance function has the following asymptotic behaviour:[1]

$$|\gamma(h)| \sim C|h|^{\alpha} \quad \text{as } h \to \infty \tag{1}$$

If, in addition, $\alpha>-1$ and so $\sum|\gamma(h)|=\infty$, it will be said that X_t is persistent.

There are many examples of data series that reveal long-memory characteristics. Historically, one of the most important examples is the series of yearly water levels of the Nile river, that puzzled probabilists for more than a decade.

The English physicist Harold Edwin Hurst, working in Cairo, studied extensively the historic records of the Nile river levels (Hurst, 1951) and developed a new statistic tool appropriated to his purposes: the R/S analysis.

Let $X_1, X_2, \ldots, X_d$ represent the inputs (e.g. river levels, rainfalls, etc.) in d successive years, and let $\bar{X}$ represent the empirical average. The adjusted range R is defined as

$$R(d) := \max_{0 \leqslant l \leqslant d} \left\{ \sum_{i=1}^{l} X_i - l\bar{X} \right\} - \min_{0 \leqslant l \leqslant d} \left\{ \sum_{i=1}^{l} X_i - l\bar{X} \right\}$$

and the normalization factor S is the following version of the standard deviation:

$$S(d) := d^{-1/2} \left(\sum_{i=1}^{d} X_i^2 - d\bar{X}^2 \right)^{1/2}$$

The statistic that Hurst used was the rescaled adjusted range Q, also called R/S statistic:

$$Q(d) := \frac{R(d)}{S(d)}$$

In the studies of the Nile and other rivers, Hurst averaged the values $R(d)/S(d)$ for different starting points t_0. He found that this average fluctuates around d^J, with $J \simeq 0.74$. The parameter J is called the Hurst exponent and the fact that $J>1/2$ for such records and other natural and economic data (Mandelbrot, 1983) became known as the Hurst phenomenon, a disturbing fact for traditional time-series modelling. The point is that, for ARMA models and for virtually all stationary models conceived and used in traditional time-series analysis, we have $J=1/2$. The fact that $J>1/2$ can be interpreted as a sign of long memory, as Mandelbrot proved in a series of papers (see, for instance, Mandelbrot and Taqqu, 1979).

Various methods can be used in R/S analysis in order to estimate the Hurst exponent J. A natural estimate, using a sample of size d, is simply

$$\hat{J} = \frac{\log[R(d)/S(d)]}{\log d} \tag{2}$$

III. R/S ANALYSIS IN ECONOMICS AND FINANCE

The first economic studies on long-range dependence used R/S analysis and tried to detect the existence of long memory in stock returns. Greene and Fielitz (1977) studied

[1] The symbol $\sim$ is used in the standard way: $a_n \sim cb_n$, where $c \neq 0$ iff $a_n/b_n \to c$.

200 stock series with 1220 daily return observations each, and compared their estimates of J with biases obtained by Wallis and Matalas (1970) from simulations of series that presented either no memory or short memory due to a first-order autoregressive term. They found a significant number of cases in which the J estimates were out of the bias limits of the simulations.

Their findings were perhaps the first to report long memory in economic time series, but a number of objections can be raised. First, the distribution of the R/S statistic is not known, and its capability of discriminating between long-range dependence and short-range autoregressive models is not known in finite samples. Second, with a sample of 200 series, it is expected, even without long-memory, that $200 \times \alpha$ series fall out of the empirical α-confidence interval.

A similar investigation using R/S analysis was later performed by Booth *et al.* (1982a, b) over gold prices and exchange rates. In their first paper, they obtained the estimates of $\hat{J}=0.642$ and $\hat{J}=0.660$. They claimed that persistence could be the explanation for the existence of cycles of unequal duration reported by simple inspection, cycles that have been the basis for the success in the application of filter rules to the gold market. In a second paper, Booth *et al.* analysed the daily spot prices for the British pound, French franc and German mark in terms of the US dollar. Data corresponding to a fixed exchange regime (1965–71) and to a flexible regime (1973–79) were used. An anti-persistent long-term dependence ($J<1/2$) was estimated for the first regime, and a persistent form of long memory ($J>1/2$) was estimated for the data corresponding to the second regime.

Other authors (Helms *et al.*, 1984) analysed the memory in commodity futures contracts. Their findings were also in the direction of persistent long memory for the six contracts studied (soybean, soybean oil and soybean meal contracts for January 1977 and for March 1976, comprising approximately 230 observations of intra-day price changes for each contract).

All these last three studies can be subjected to the same remarks as presented in regards to the paper of Greene and Fielitz (1977). In all cases, the authors were cautious and argued that the deviations of the estimated coefficients from 1/2 were greater than it would be expected in the absence of long memory. They compared their results with the existing simulations of Wallis and Matalas for the case of first-order autocorrelations similar to those presented in the empirical series under consideration. Nevertheless, the results were near the border line, and, what is more important, simulations for autocorrelations of higher order are unknown, and so no valid reference existed. In the case of the study of Helms *et al.* (1984), another major remark should be made: they analysed relative price changes in a period when the rate of change was increasing. The hypothesis of non-stationarity can then be introduced in parallel with the one of long memory.

Given the difficulty in making reliable statistical inferences using the R/S statistic, efforts were moved in the direction of constructing formal statistic tests, but the tests revealed a very low power and the necessity of very long samples.

More recently, Lo (1991) constructed a modified R/S statistic to handle those problems. The basic idea is the construction of a statistic that is invariant over a general class of short-memory processes but that is sensitive to the presence of long memory.

The modified R/S statistic, which will be represented by $Q^{L}(n, q)$, differs from the usual $Q(n)$ only in the denominator. The new estimator for the variance is

$$\begin{aligned} S^2(n,q) &:= \frac{1}{n}\sum_{j=1}^{n}(X_j-\bar{X}_n)^2 \\ &\quad + \frac{2}{n} w_q(j) \sum_{j=1}^{q} \sum_{i=j+1}^{n} (X_i-\bar{X}_n)(X_{i-j}-\bar{X}_n) \\ &= \hat{\sigma}_n^2 + 2\sum_{j=1}^{q} w_q(j)\hat{\gamma}(j) \end{aligned} \tag{3}$$

where $\hat{\sigma}_n^2$ and $\hat{\gamma}(j)$ are the usual estimators for the variance and the autocovariance, and the weights $w_q(j)$ can be defined as

$$w_q(j) := 1 - \frac{j}{q+1}, \quad q<n \tag{4}$$

$S^2(n, q)$ is a consistent estimator of the partial sums' variance, which, in presence of autocorrelations, is not simply the sum of the variances of the individual terms, as in $S(n)$, includes the autocovariances. The weights above are those proposed by Newey and West (1987) and yield a consistent estimator under very general conditions. Regarding the rule for choosing the truncation parameter q, Lo provides some Monte Carlo evidence and suggests the formula

$$q := [k_n], \quad k_n := \left(\frac{3n}{2}\right)^{1/3}\left(\frac{2\hat{\rho}(1)}{1-\hat{\rho}^2(1)}\right)^{2/3} \tag{5}$$

where $[\cdot]$ denotes the greatest integer function and $\hat{\rho}(1)$ is the usual estimator of the first-order autocorrelation coefficient.

The modified rescaled range is thus defined as

$$\begin{aligned} Q^{L}(n,q) &:= \frac{1}{S(n,q)}\left(\max_{0\leq k\leq d}\left\{\sum_{i=1}^{k} X_i - k\bar{X}\right\}\right. \\ &\quad \left. - \min_{0\leq l\leq d}\left\{\sum_{i=1}^{l} X_i - l\bar{X}\right\}\right) \end{aligned} \tag{6}$$

In the absence of long memory $Q^{L}(n, q)/\sqrt{n}$ converges weakly to the well-known range of the Brownian bridge $[W^{0}(t)]$ on the unit interval, and so it is easy to compare $Q^{L}(n, q)$ with the critical values for the null hypothesis.

Lo (1991) studied the monthly and daily stock returns given by the value- and equal-weighted indexes of the

Center for Research in Security Prices (CRSP). The daily series had 6409 observations, ranging from July 1962 to December 1987; the same period yield 1330 observations of the weekly returns. Monthly indexes had 744 observations each, ranging from 1926 to 1987. Annual data was used from Standard and Poor's composite index – 115 observations from 1872 to 1986.

The results contradicted previous research. In most of the cases the values for $Q(n)/\sqrt{n}$ given by the classic R/S statistic were in the critical region, at 5%. But the values for $Q^{L}(n, q)/\sqrt{n}$ given by the modified R/S statistic were never significant, for any of the series analysed and for any of the truncation values for q that were used. Moreover, the differences between $Q(n)$ and $Q^{L}(n, q)$ were substantial: the former statistic gives 'biases' (relatively to the latter) in the order of magnitude of 50% and 100%.

Of course, both statistics rely solely on the asymptotic theory, and the only advantage of Lo's proposal is that he offers the exact asymptotic distribution of $Q^{L}(n, q)$ under very general conditions, being possible to tabulate critical values regardless of the particular form of short-range dependence that could be present.

Lo also provided some simulations in order to access the power of his test. He showed that the power was sensitive to the choice of q and that a reasonable power against fractional noise could only be obtained in large samples: 500, 1000 observations. Nevertheless, these results are for particular models, and we are still left in a fuzzy terrain when using the R/S statistic with real data.

A parametric analysis with long-memory models is then an appealing approach.

IV. FRACTIONAL LONG-MEMORY MODELS FOR FINANCIAL TIME SERIES

A very general class of long-memory and persistent models is based on the fractionally differenced noise, introduced independently by Granger and Joyeux (1980) and by Hosking (1981). These models have been proved as valuable tools in various areas of economic theory (see e.g. Diebold and Rudebush, 1989), and in economic forecasting (see e.g. Geweke and Porter-Hudak, 1983; Crato, 1991).

In order to introduce the models, the notation will follow closely Brockwell and Davis (1991), where a more detailed development can be found. The symbols B and ∇ will represent the backwards and the differencing operators, respectively, i.e. $BX_t := X_{t-1}$ and $\nabla X_t := (1-B)X_t = X_t - X_{t-1}$. Integer powers of the these operators are defined in the usual way, i.e. $B^n = B \cdot B^{n-1}$ and $\nabla^n = \nabla \cdot \nabla^{n-1}$.

If d is any real number, the fractional difference operator ∇^d is defined as

$$\nabla^d = (1-B)^d := \sum_{k=0}^{\infty} \binom{d}{k} (-B)^k \tag{7}$$

where

$$\binom{d}{k} = \frac{d}{k}\frac{d-1}{k-1} \cdots \frac{d-k+1}{1} \tag{8}$$

When d is a positive integer, the sum is reduced to the terms $k = 1, \ldots, d$, since then, for $k > d$, $\binom{d}{k} = 0$. When the operator is applied to a random process, the sum of the series is to be understood as a mean-square limit.

A process (X_t) is called fractionally differenced noise or fractional noise, with d in the open interval $]-0.5, 0.5[$ if

$$\nabla^d X_t = \varepsilon_t, \quad \text{with } \varepsilon_t \sim WN(0, \sigma^2) \tag{9}$$

Equation 9 provides an autoregressive representation of the process

$$\sum_{k=0}^{\infty} \pi_k X_{t-k} = \varepsilon_t \tag{10}$$

with $\pi_k = \binom{d}{k}(-1)^k$. Explicitly,

$$X_t - dX_{t-1} + \frac{d(d-1)}{2} X_{t-2} - \frac{d(d-1)(d-2)}{6} X_{t-3} + \cdots = \varepsilon_t$$

By the laws of exponents,

$$X_t = \nabla^{-d} \varepsilon_t \tag{11}$$

Equation 11 can be rewritten as

$$X_t = \varepsilon_t + d\varepsilon_{t-1} + \frac{d(d+1)}{2} \varepsilon_{t-2} + \frac{d(d+1)(d+2)}{6} \varepsilon_{t-3} + \cdots$$

The natural generalization of the fractional noise is the fractionally integrated ARMA model, i.e. an ARIMA where the order of integration is the non-integer $d \in]-0.5, 0.5[$. In this way, the relatively small flexibility existent in fractional noise, controlled only by the two parameters σ^2 and d, is increased by the $p+q$ parameters of the autoregressive and moving-average polynomials.

An autoregressive fractionally integrated moving-average process, ARFIMA(p, d, q) with p, q non-negative integers and d in the open interval $]-0.5, 0.5[$, is a stationary ARMA driven by fractional noise, i.e. (X_t) is an ARFIMA if it is a stationary solution of the difference equation

$$\phi(B)X_t = \theta(B)\xi_t, \quad \text{with } \xi_t := \nabla^{-d}\varepsilon_t, \quad \varepsilon_t \sim WN(0, \sigma^2) \tag{12}$$

where $\phi(B)$ and $\theta(B)$ are lag polynomials of order p and q, respectively. Explicitly,

$$X_t - \phi_1 X_{t-1} - \ldots - \phi_p X_{t-p} = \xi_t + \theta_1 \xi_{t-1} + \ldots + \theta_q \xi_{t-q} \tag{13}$$

with

$$\xi_t = \varepsilon_t + d\varepsilon_{t-1} + \frac{d(d+1)}{2} \varepsilon_{t-2} + \frac{d(d+1)(d+2)}{6} \varepsilon_{t-3} + \cdots$$

It can be proved that if (X_t) is a fractional noise or, more generally, a stationary ARFIMA, then its autocorrelation function has the following asymptotic expression:

$$\gamma(h) \sim Ch^{2d-1} \quad \text{as } h \to \infty. \tag{14}$$

Hence, for $d \neq 0$ it is a long-memory process, for $-1/2 < d < 0$ it is antipersistent, and for $0 < d < 1/2$ it is persistent. For details see again Brockwell and Davis (1991, pp. 520–34).

V. FORTY YEARS IN SEVEN COUNTRIES

The data that will be considered represent the G-7 countries stock price monthly return indexes, as in Citibase, from the first week of 1950 to the first week of 1989.

The maximum-likelihood Gaussian estimation used the program GQSTFRAC of Sowell and estimated 16 alternative ARFIMA models, with $p = 0, 1, 2, 3$ and $q = 0, 1, 2, 3$. In order to select what could be reasonable ARFIMA models, appropriate to each country's stock returns, AICc and SIC criteria were used. These are perhaps the most common selection criteria used in today's time-series analysis. They measure the goodness of fit of the model to the data and subtract a penalization that increases with the number of parameters of the model. AICc stands for Akaike information criteria in its corrected version, as suggested by Hurvich and Tsai (1989), and SIC is the Bayesian criterion of Schwarz (1978). For details, the reader is referred to Brockwell and Davis (1991, pp. 301–6) and to de Gooijer *et al.* (1985).

Next step, all selected models were inspected and the estimated values for the differenced parameter d were tested. The standard t-test for the nullity of the parameter was performed using the estimated standard deviation of the estimate for d. The non-rejection of the null hypothesis $d = 0$ leads to the conclusion of short memory in the series and the rejection leads to the existence of long memory. The results are presented on Table 1. Estimations were also performed on a broader set of ARFIMA models, using the approximate maximum-likelihood estimation program LONGMEM, from the new version of the package ITSM of Brockwell and Davis. The results obtained confirmed the selected models on Table 1.

The maximum-likelihood procedure was supplemented with the estimation technique of Geweke and Porter-Hudak (1983), that suggested a regression over the log of the periodogram at low frequencies, where its slope is directly dependent on the long-memory parameter d. The rationale lies in the form of the spectral density $f(\lambda)$ of an ARFIMA process near zero,

$$f(\lambda) \sim C|1 - e^{-i\lambda}|^{-2d} \quad \text{as } \lambda \to 0. \tag{15}$$

Geweke and Porter-Hudak argued that their regression estimator could capture the long-memory characteristic of the process, without being 'contaminated' in the estimation

Table 1. *Results of the maximum-likelihood estimation (models chosen by AICc and SIC criteria, estimated d and corresponding t-values)*

Country	AICc	SIC
Canada	(2, d, 2) $\hat{d} = -0.45$ $t = -2.24$	(0, d, 0) $\hat{d} = 0.06$ $t = 1.35$
France	(1, d, 1) $\hat{d} = -0.04$ $t = -0.99$	(0, d, 0) $\hat{d} = 0.01$ $t = 0.01$
Italy	(0, d, 2) $\hat{d} = 0.18$ $t = 1.78$	(0, d, 2) $\hat{d} = 0.18$ $t = 1.78$
Japan	(0, d, 1) $\hat{d} = 0.02$ $t = 0.42$	(0, d, 0) $\hat{d} = 0.12$ $t = 3.05$
UK	(2, d, 3) $\hat{d} = -0.07$ $t = -1.11$	(0, d, 1) $\hat{d} = -0.02$ $t = 0.43$
US	(0, d, 1) $\hat{d} = -0.02$ $t = -0.32$	(0, d, 1) $\hat{d} = 0.02$ $r = -0.32$
West Germany	(1, d, 0) $\hat{d} = 0.09$ $t = 1.26$	(0, d, 0) $\hat{d} = 0.23$ $t = 6.04$

by the other $p + q$ parameters of the ARFIMA model. Using the periodogram as an estimator of the spectral density, d can be estimated by regression. As a truncation parameter m that would choose the low Fourier frequencies to be considered, simulations by Geweke and Porter-Hudak suggest the use of m around $\sqrt{n}$, where n is the number of observations. The results are presented on Table 2. Again, the test on the nullity of the parameter d is performed as a usual t-test, using the standard deviation given by the regression.

An inspection of the two tables shows immediately that the null hypothesis of short memory is not rejected for France, UK and US. The t-values are very small. At a level of 5% of significance, none of the estimation procedures leads to a parameter d significantly different from zero. The results are very clear and do not support any claim for the existence of long memory in these stock return series.

Some doubts can arise in the cases of Canada, Italy and Japan.

In the case of Canada, the regression method, with results as summarized in Table 2, never leads to an estimated parameter significantly different from zero. But the maximum-likelihood procedure leads to two competitive models, an ARFIMA (2, d, 2) and an ARFIMA (0, d, 0), and in the first one a value of d significantly different from zero is found (at least for a significance level of 5%; the null hypothesis is not rejected for a level of 1%).

Table 2. *Results of the estimation by regression over the log-periodogram (estimated d and t-values for different numbers m of low-frequency ordinates used)*

Country	$m=20$	$m=30$	$m=40$
Canada	$\hat{d}=-0.22$ $t=-1.43$	$\hat{d}=-0.08$ $t=-0.70$	$\hat{d}=-0.08$ $t=0.84$
France	$\hat{d}=-0.02$ $t=-0.15$	$\hat{d}=0.13$ $t=1.04$	$\hat{d}=0.03$ $t=0.25$
Italy	$\hat{d}=0.09$ $t=0.51$	$\hat{d}=0.36$ $t=2.26$	$\hat{d}=0.22$ $t=1.70$
Japan	$\hat{d}=0.15$ $t=0.84$	$\hat{d}=0.20$ $t=1.42$	$\hat{d}=0.11$ $t=0.89$
UK	$\hat{d}=-0.14$ $t=-1.15$	$\hat{d}=-0.02$ $t=0.14$	$\hat{d}=-0.01$ $t=-0.11$
US	$\hat{d}=-0.28$ $t=-1.22$	$\hat{d}=-0.15$ $t=-0.94$	$\hat{d}=-0.13$ $t=1.00$
West Germany	$\hat{d}=0.20$ $t=1.26$	$\hat{d}=0.19$ $t=1.70$	$\hat{d}=0.21$ $t=2.22$

Table 3. *Results of the R/S analysis (Hurst exponents estimated with $Q(n)$ and with $Q^L(n, q)$; unilateral test p-values displayed for the second estimates)*

Country	$Q(n)$	$Q^L(n, q)$
Canada	$\hat{J}=0.518$	$\hat{J}=0.509$ $0.10<p<0.20$
France	$\hat{J}=0.568$	$\hat{J}=0.568$ $0.10<p<0.20$
Italy	$\hat{J}=0.635$	$\hat{J}=0.586$ $0.05<p<0.10$
Japan	$\hat{J}=0.567$	$\hat{J}=0.544$ $0.30<p<0.40$
UK	$\hat{J}=0.589$	$\hat{J}=0.560$ $0.20<p<0.30$
US	$\hat{J}=0.563$	$\hat{J}=0.530$ $0.40<p<0.50$
West Germany	$\hat{J}=0.641$	$\hat{J}=0.590$ $0.01<p<0.05$

The case of Japan is very similar. The model chosen by one of the criteria leads to a non-null value for d ($t=3.05$), while with the other competitive model the hypothesis of short memory is not rejected ($t=0.42$). The regression method, however, always leads to small t-values.

The case of Italy is just the opposite: the maximum-likelihood procedure leads to a unique model in which the t-test accepts the null hypothesis $d=0$. But with the regression method and a truncation of the periodogram at $m=30$, the value of d is significantly different from zero. This result, nevertheless, is not consistent: with the other truncations also present in Table 2, the nullity is not rejected.

But the really interesting case is the one of West Germany. The maximum-likelihood method leads to two competitive models in which the more parsimonious, chosen by SIC, the ARFIMA (0, d, 0), rejects strongly the null hypothesis of short-memory and gives $\hat{d}=0.23$. In addition, the Geweke and Porter-Hudak method that estimates d independently of the possible short-memory parameters from the ARMA part of the process, consistently points to values of d that are significantly different from zero and also around 0.20.

For the sake of completeness, the traditional and the modified *R/S* analyses were also performed on this set of series. The estimates for the Hurst exponent J were obtained using Equation 2. The statistics $Q(n)$ and $Q^L(n, q)$ were computed following the steps described in Section III and using the complete series. The results are presented in Table 3.

In all the countries but Canada, the classical estimates for the Hurst exponent J [using $Q(n)$] point in the direction of persistent long memory. As discussed above, rigorous statistical tests are not possible. However, comparing the results with the simulations of Wallis and Matalas (1970) for similar first-order sample autocorrelations, three countries, Italy, UK and West Germany, are found to have significant positive values for the exponent J, while France, Japan and US are on the border line.

The corrected R/S estimates [using $Q^L(n, q)$] can be compared with the fractiles for the Brownian bridge reproduced in Lo (1991, p. 1288). At conventional levels, the alternative hypothesis of persistent long memory could be sustained only in Italy and West Germany.

VI. CONCLUSIONS

A broad class of time-series models capable of reproducing long-memory behaviour was fitted to a long international series of stock returns that could be found for the G-7 countries (Canada, France, Italy, Japan, UK, US and West Germany).

The time-series models that were used are the autoregressive fractionally differenced moving-average models. ARFIMA.

In these ARFIMA models, the hypothesis of nullity of the differencing parameter d corresponds to the hypothesis of non-existence of long memory. This hypothesis was tested in the seven series.

In three of the countries, France, UK and US, no test could reject the nullity of the parameter d. In three other countries, Canada, Italy and Japan, short- and long-memory models provided competitive tools to explain the behaviour of the stock returns. Nevertheless, consistent fractional estimates for d were not found, and with this parametric analysis the alternative hypothesis of long memory could not be based on solid grounds. In the final country, West Germany, the estimations and the tests performed strongly suggest the existence of long memory in the stock returns.

These results were complemented with the classical and the modified R/S analyses. Aditional evidence of long memory was found for Italy and West Germany.

On the whole, the results obtained with this parametric approach contradicted the findings of various previous researchers, who had used the non-parametric tool supplied by *R/S* analysis and had consistently claimed the existence of long memory in financial time series. Lo (1991) had argued that what had been detected was simply the existence of short memory in stock returns. For only one out of the seven countries studied, the former West Germany, could that critique of Lo be contested with the parametric approach developed here.

ACKNOWLEDGEMENTS

I am grateful both to Peter Brockwell and Fallaw Sowell for graciously supplying their ARFIMA estimation programs. Valuable suggestions from an anonymous referee led to improvements of the paper. This work was initiated at the Lisbon Institute of Economics (ISEG). A grant from JNICT, Portugal, is gratefully acknowledged.

REFERENCES

Ball, R. (1989) What do we know about stock market 'efficiency'?, in *A Reappraisal of the Efficiency of Financial Markets*, Guimarães, R. M. C., Kingsman B. and Taylor, S. (eds), Springer, Berlin.

Booth, G. Geoffrey, K., Fred R. and Koveos, P. E. (1982a) Persistent dependence in gold prices, *The Journal of Financial Research*, **5**, 85–93.

Booth, G. Geoffrey, K., Fred R. and Koveos, P. E. (1982b) *R/S* analysis of foreign exchange rates under two international monetary regimes' *Journal of Monetary Economics*, **10**, 407–15.

Brockwell, P. J. and Davis, R. A. (1991) *Time Series: Theory and Methods*, 2nd edn, Springer, Berlin.

Crato, N. (1991) Forecasting with long-memory models, invited paper to the 11th International Symposium on Forecasting, New York, June 1991.

de Gooijer, J. G., Abraham, B., Gould, A. and Robinson, L. (1985) Methods of determining the order of an autoregressive–moving average process: a survey, *International Statistical Review*, **53**, 301–29.

Diebold, F. X. and Rudebush, G. D. (1989) Long memory and persistence in aggregate output, *Journal of Monetary Economics*, **24**, 189–209.

Geweke, J. and Porter-Hudak, S. (1983) The estimation and application of long memory time series models, *Journal of Time Series Analysis*, **4**, 221–38.

Granger, C. W. J. (1992) Forecasting stock market prices: lessons for forecasters, *International Journal of Forecasting*, **8**, 3–13.

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

Greene, M. T. and Fielitz, B. D. (1977) Long-term dependence in common stock returns, *Journal of Financial Economics*, **4**, 339–49.

Guimarães, R. M. C., Kingsman, B. G. and Taylor, S. (1989) *A Reappraisal of the Efficiency of Financial Markets*, Springer, Berlin.

Haubrich, J. G. and Lo, A. W. (1989) The sources and nature of long-term memory in the business cycle, *NBER Working Paper*, 2951.

Helms, B. P., Kaen, F. R. and Rosenman, R. E. (1984) Memory in commodity futures contracts, *Journal of Futures Markets*, **4**, 559–67.

Hosking, J. R. M. (1981) Fractional differencing, *Biometrika*, **68**, 165–76.

Hurvich, C. M. and Tsai, C. L. (1989) Regression and time series model selection in small samples, *Biometrika*, **76**, 297–307.

Hurst, H. E. (1951) Long-term storage capacity of reservoirs, *Transactions of the American Society of Civil Engineers*, **116**, 770–808.

Lo, A. W. (1991) Long-term memory in stock market prices, *Econometrica*, **59**, 1279–313.

Mandelbrot, B. B. (1965) Une classe de processus stochastiques homothétiques à soi; application à la loi climatologique de H. E. Hurst, *Comptes Rendues des Séances de l'Académie des Sciences, Paris*, **260**, 3274–7.

Mandelbrot, B. B. (1972) Statistical methodology for nonperiodic cycles: from the covariance to R/S analysis, *Annals of Economic and Social Measurement*, **1**, 259–90.

Mandelbrot, B. B. (1983) *The Fractal Geometry of Nature*, Freeman, New York.

Mandelbrot, B. B. and Taqqu, M. S. (1979) Robust *R/S* analysis of long-run serial correlation, *Proceedings of the 42nd Session of the International Statistical Institute, Manila, Bulletin of the I.S.I.*, **48**, 69–104.

Newey, W. K. and West, K. D. (1987) A simple, positive semi-definite, heteroskedasticity and autocorrelation consistent covariance matrix, *Econometrica*, **55**, 703–8.

Schwarz, G. (1978) Estimating the dimension of a model, *Annals of Statistics*, **6**, 461–4.

Sowell, F. (1992) Maximum likelihood estimation of stationary univariate fractionally integrated time series models, *Journal of Econometrics*, **53**, 165–88.

Wallis, J. R. and Matalas, N. C. (1970) Small sample properties of H. and K-estimators of the Hurst coefficient *h*, *Water Resources Research*, **6**, 1583–94.