

Memory in Returns and Volatilities of Futures' Contracts

NUNO CRATO

BONNIE K. RAY*

Various authors claim to have found evidence of stochastic long memory behavior in futures' contract returns using the Hurst statistic. This paper reexamines futures' returns for evidence of persistent behavior using a biased-corrected version of the Hurst statistic, a nonparametric spectral test, and a spectral regression estimate of the long-memory parameter. Results based on these new methods provide no evidence for persistent behavior in futures' returns. However, they provide overwhelming evidence of long memory behavior for the volatility of futures' returns. This finding adds to the emerging literature on persistent volatility in financial markets and suggests the use of new methods of forecasting volatility, assessing risk, and optimizing portfolios in futures' markets.

Keywords: ARFIMA, efficient market hypothesis, futures, long-memory, R/S analysis, volatility.

* Correspondence author, Bonnie K. Ray, Department of Mathematical Sciences, New Jersey Institute of Technology, Newark, NJ 07102.

Nuno Crato, Department of Mathematics, Instituto Superior de Economia e Gesto, UTL, Lisbon and Department of Mathematical Sciences, New Jersey Institute of Technology.

Bonnie K. Ray, Department of Mathematical Sciences, New Jersey Institute of Technology.

INTRODUCTION

An important issue in the study of financial markets is the evaluation of the stochastic memory of market returns. If returns have a stochastic memory, it is possible to obtain increased profits on the basis of price change predictions. Such predictability would contradict the efficient market hypothesis (see, *e.g.*, Ball 1989), which states that increased profits cannot be consistently made by trading on the basis of information given by past prices.

Over twenty years ago, Mandelbrot's work (1965, 1972) on the use of the Rescaled Adjusted Range Statistic (R/S) stimulated the use of this statistic for analyzing the fractal behavior and stochastic memory of financial time series. Using R/S analysis, Greene and Fielitz (1977), Booth, Kaen, and Koveos (1982a, 1982b) and other researchers arrived at the surprising conclusion that some financial time series have long-memory behavior. In particular, Helms *et al.* (1984) analyzed price changes in futures' contracts and claimed that the returns of the series display long-memory characteristics. The claim of long-memory on commodity futures' price changes goes further than the simple claim of some 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, so that the actual movements in the market are stochastically influenced by the recent to the most remote past.

More recent work on R/S analysis (Lo, 1991) has demonstrated that this statistic is biased when short-range correlation is present in a series, calling into question earlier claims of persistence in futures' returns. Using new statistical tools, many authors (Fung and Lo, 1993; Cheung and Lai, 1993; Crato, 1994) have found that financial returns on the major markets have no significant memory. However, the original claims of Helms *et al.* (1984) concerning the memory of futures' returns were never reevaluated in light of the new statistical tools. Doukas (1990) found unit roots on most of the returns series, although the existence of a unit root does not preclude long-memory behavior. For a long time, the issue was not investigated further. Recently, Barkoulas, Labys, and Onochie (1997) used the classical R/S analysis to reevaluate the memory of futures' returns. Studying a large new data set, they claim to have found persistent long memory in a significant group of futures' contracts.

This paper reexamines the memory of futures' returns using three different methods: (i) a modified version of the R/S statistic developed by Lo (1991), (ii) the nonparametric

spectral test of Lobato and Robinson (1998), and (iii) a test based on the estimator of the long memory parameter due to Geweke and Porter-Hudak (1983), henceforth GPH. Our results indicate no long memory behavior in futures' returns. However, a similar analysis applied to the volatility of the returns finds overwhelming evidence of persistence in volatility. This finding is consistent with recent work by Ding, Granger and Engle (1993), de Lima and Crato (1993), Bollerslev and Mikkelsen (1996), Baillie, Bollerslev and Mikkelsen (1996) and Breidt, Crato, and de Lima (1998), among others, which finds compelling evidence that the volatility of financial markets displays a long-memory structure. As Engle (1982), Bollerslev (1986) and others have shown, the volatility of financial returns may display a strong autocorrelation structure while the level of the returns display no memory and a random-walk type behavior. This finding in futures' series behavior adds to the emerging literature on persistent volatility in financial markets and suggests the use of new methods of forecasting volatility, assessing risk, and optimizing portfolios in futures' markets.

The plan of the rest of the paper is as follows. Section 2 defines the concepts of short- and long-memory of a random process and describes tests for long-memory behavior. Section 3 describes the data sets and presents the new empirical findings for both futures' returns and volatilities. Section 4 introduces the long-memory stochastic volatility (LMSV) model and illustrates its use for a particular series of futures' contract returns. Section 5 concludes.

THE MEMORY OF A TIME SERIES

The expected relationship between the value of a process at time t and its value at time $t - k$ is a measure of the correlation present in the series. A stationary time series has correlation which depends only on the time lag k between the two observations and decays to zero as k increases, reflecting the fact that the influence of the past values decreases with the lags under consideration. The speed of this decay is a measure of the stochastic "memory" of the random process.

A process in which all observations are uncorrelated is called a white noise, and the random process is said to have *no memory*. One commonly assumed implication of the Efficient Market Hypothesis is that futures' returns follow such a process. Processes with *short memory* are those that have autocorrelations decaying to zero at a geometric rate. AutoRegressive Moving Average (ARMA) models are examples of such processes.

In contrast, a long-memory process has autocorrelations that decay much more slowly, asymptotically following a hyperbolic decay. More precisely, a stationary process $\{X_t\}$ is said to have *long-memory* if, for some $C > 0$ and $d < 0.5$, its autocovariance function, $\gamma(k) = E[(X_t - \mu)(X_{t+k} - \mu)]$, has asymptotic behavior

$$|\gamma(k)| \sim C|k|^{2d-1} \quad \text{as } k \rightarrow \infty, \quad (1)$$

for $d \neq 0$. If $d > 0$, so that $\sum |\gamma(k)| = \infty$, the process is said to be persistent.

Equivalently, a long-memory process can be characterized in the frequency domain by the following form of its spectral density function, $f(\lambda)$, at frequency λ :

$$f(\lambda) = |1 - e^{-i\lambda}|^{-2d} f_U(\lambda), \quad (2)$$

where $d \neq 0$ and f_U is finite and bounded above from zero at the origin.

There are many tests and statistics to detect the existence of long memory in a time series. One of the first tools to be developed was the R/S method.

R/S analysis. Let X_1, X_2, \dots, X_n represent the observations in n successive periods, and let \bar{X} represent the empirical average. The *adjusted range*, R , is defined as

$$R(d) := \max_{0 \leq k \leq n} \left\{ \sum_{i=1}^k X_i - k\bar{X} \right\} - \min_{0 \leq \ell \leq n} \left\{ \sum_{i=1}^{\ell} X_i - \ell\bar{X} \right\}.$$

A normalization factor S is an estimate of the standard deviation of the process. The estimator can be written in the general form

$$S(n, q) := \sqrt{\sum_{j=-q}^q w_q(j) \hat{\gamma}(j)},$$

where $\hat{\gamma}(j)$ is an estimate of the process autocovariance at lag j and $w_q(j)$ are weights. The R/S statistic is then defined as

$$Q(n, q) := \frac{R(n)}{S(n, q)}. \quad (3)$$

In the classical R/S analysis applied by Helms *et al.* (1984) and by Barkoulas, Labys, and Onochie (1997), $q = 0$, $w_0(0) = 1$ and so $S(n, q)$ is simply the usual sample standard deviation estimate. The distribution of this classical R/S statistic ($q = 0$) is not known,

and short-memory components can affect the R/S values in a non-negligible way. A modified R/S statistic, having $q \neq 0$, was constructed by Lo (1991) to handle these problems. Lo used the weights proposed by Newey and West (1987), $w_q(j) := 1 - \frac{j}{q+1}$ with $q < n$. For choosing the truncation parameter q , Lo followed Andrews (1987) suggestion of setting q as the greatest integer less than or equal to the data dependent quantity

$$\left(\frac{3n}{2}\right)^{1/3} \left(\frac{2\hat{\rho}(1)}{1 - \hat{\rho}^2(1)}\right)^{2/3}, \quad (4)$$

where $\hat{\rho}(1)$ is an estimate of the first order autocorrelation of the process. The modified estimator is invariant over a general class of short-memory processes but is sensitive to the presence of long-memory. If only short memory is present, then $Q(n, q)/\sqrt{n}$ converges weakly to the range of the Brownian bridge on the unit interval. The distribution function for this range is given by Kennedy (1976) and Siddiqui (1976). It can be expressed as $F(v) = \sum_{k=-\infty}^{\infty} (1 - 4v^2k^2)e^{-2v^2k^2}$, from which critical values can be computed to test the null hypothesis of no long-range dependence.

For ARMA and other short memory processes, the values of $Q(n, q)$ converge to n^J with $J = 1/2$. The parameter J is called the *Hurst exponent* and is related to the long-memory parameter d discussed above by $J = d + 1/2$. Mandelbrot and co-workers proved in a series of papers (see, for instance, Mandelbrot and Taqqu 1979) that when $J > 1/2$, a process has long-memory. Various methods can be used in R/S analysis to estimate the J . A natural estimate for a series of length n is simply

$$\hat{j} = \frac{\log Q(n, q)}{\log n}. \quad (5)$$

No distributional results are currently available for this estimate of J .

GPH spectral regression. The second method we have used for determining the existence of long-range dependence in a time series is based on the spectral form of a long-range dependent process, as given in (2). Geweke and Porter-Hudak (1983) suggested regressing the log of the estimated spectrum of the series on the log of the frequency values themselves. They also suggested that the regression be performed using a set of Fourier frequencies close to zero, where the slope of the log spectrum relative to the frequency is directly dependent on the long-memory parameter d . They argued that their regression estimator could capture the long-memory characteristic of the process

without being “contaminated” in the estimation by short-memory correlation in the time series evident at higher frequencies.

Use of this method requires choice of a truncation parameter m to determine the number of Fourier frequencies to be considered. Based on simulations, Geweke and Porter-Hudak suggest the use of $m = [n^u]$ with $u = .5$, where n is the number of observations and $[\cdot]$ is the greatest integer function. Robinson (1995) argued that a consistent estimator for d could be obtained if an additional low-order truncation $l > 1$ was introduced, thus using the Fourier frequencies $j = l, 2, \dots, m$. More recent results by Hurvich, Deo, and Brodsky (1998) find that this lower truncation is not necessary. They conclude that setting $l = 1$, *i.e.*, no lower truncation and $[n^{.6}]$ for the upper truncation is a more optimal choice. We have tried a range of possible values for both upper and lower truncations in order to have more robust results.

The distribution of the estimated d parameter can be shown to be Gaussian under certain conditions and thus a test for $H_0 : d = 0$ (no long-memory) can be performed as a usual t -test, using the standard deviation given by the regression.

Nonparametric LM spectral test. A third tool we have used is the nonparametric spectral test constructed by Lobato and Robinson (1998). This is a test for stationary short-memory, *i.e.*, integration of order zero of the time series, $X_t \sim I(0)$ against fractional alternatives, $X_t \sim I(d)$ with $d \neq 0$.

This test uses the periodogram ordinates $I(\omega_j)$ at the Fourier frequencies $\omega_j = 2\pi j/n$, where $j = 1, 2, \dots, m < n/2$ and n is the length of the series. The truncation parameter m is chosen in order to use a sensible number of low-order frequencies, whose values are essentially determined by the long-memory component of the process. Without the previous knowledge of the memory parameter, there is no clear rule for determining the value of the truncation m . We have thus used a set of values consistent with the simulation results available in the original paper.

The testing statistic is based on the Lagrange Multiplier (LM) principle and takes the form

$$t = -\sqrt{n} \sum_{j=1}^m \nu_j I(\omega_j) / \sum_{j=1}^m I(\omega_j), \quad (6)$$

where $\nu_j = \ln j - m^{-1} \left(\sum_{j=1}^m \ln(j) \right)$. This statistic is asymptotically standard normal under the short-memory assumption. Rejection of the null hypothesis in the direction

of persistent long memory ($d > 0$) occurs at the upper tail of the distribution.

In the next section, these three methods are applied to two sets of futures' returns data .

TESTING FOR LONG-MEMORY IN FUTURES' RETURNS

The first data set to be considered is the set of six daily returns series previously analyzed by Helms *et al.* (1984). They correspond to soybean oil (BO), soybean (SO) and soybean meal (SM) contracts for March 1976 (M76) and for January 1977 (J77), comprising approximately 230 observations each. The second data set to be considered is an extension of the data set used by Barkoulas, Labys, and Onochie (1997). It consists of 17 commodities, five major currencies, and one U.S. stock index. The price data are daily settlement prices of the nearest-to-expiration contract over approximately the last 20 years. At the beginning of the month in which the near contract matures, the price of the successive contract is used as the next price value. The earliest start date is January 3, 1977 and the latest end date is November 28, 1997. The appendix lists the data sets, along with the exchange in which the trade takes place and the start and end dates of the series. The data were obtained from Market Research Inc., associated with the Chicago Board of Trade. For each contract, the time series of returns is analyzed. For a given series of futures' prices, p_t , the returns, r_t^* , are computed as the first differences of the log prices: $r_t^* = \log p_t - \log p_{t-1}$.

Table I gives the results for the three methods analysis of the Helms data. Table II presents corresponding results for the extended data set. The first three columns of the tables show the estimated Hurst coefficients, J , computed from the R/S statistic as in (5). In the first column, the truncation parameter q for the computations is zero, *i.e.*, the estimates are based on the classical R/S statistic. The second and third columns show results for $q \neq 0$ with estimates based on the corrected statistic of Lo (1991). The first corrected estimates, in the second column, use a reasonably high truncation parameter, yielding conservative tests for long memory. The second corrected estimates, in the third column, use the value q^* chosen by Andrews' data-dependent formula presented in (4). An extensive set of truncation values q was also tried; these results are available upon request. The estimates are essentially similar for reasonable non-zero values of q .

The values given within square brackets below the estimated J are p -values corresponding to the probability of obtaining a value at least as high for the Hurst exponent

as the one obtained, under the null hypothesis of short-memory. For the corrected R/S of Lo ($q \neq 0$), the p -values of the estimates for J are computed as the tail probabilities of the corresponding values of the R/S statistic under the short-memory null. The p -values shown for the classical R/S statistic ($q = 0$) are appropriate only if the series are uncorrelated.

The next three columns of the tables show the GPH estimates of d for different truncation values l and m . The values given within parentheses below the estimated d are the estimated standard deviations of the estimates for d , computed as the standard error given by the regression.

The last three columns show the results of the nonparametric LM test. The table shows the values of the statistic t from equation (6) for different truncations. The values given within brackets below the statistics are the p -values of the one-sided tests for persistent long memory ($d > 0$).

— — *INSERT TABLES I and II AROUND HERE* — —

A simple inspection of Table I reveals the contrast between the results based on classical R/S estimates ($q = 0$) and those based on the modified R/S statistics. With the possible exception of the soybean oil 76 contract (BO M76), all the classical R/S estimates clearly indicate values of $J > 0.5$, *i.e.*, they suggest the presence of long memory. However, the p -values for the corrected estimates are all high, well above the conventional significance values for statistical testing. At 1%, 5%, and even 10%, one cannot reject the null of short memory, $J = 0$, for any of the time series. The GPH method gives estimates of d oscillating about zero, with standard errors of the same order as the estimates. None of the t -tests rejects the hypothesis $d = 0$, confirming the results of the modified R/S analysis. The third approach confirms these results. In all cases, the values of the statistic are close to zero and the corresponding p -values are high, well above the conventional thresholds for significance.

The three methods provide a consistent indication of no long memory for the returns.

— — *INSERT FIGURES 1 and 2 AROUND HERE* — —

Findings for the extended data set, given in Table II, are based on much longer series and corroborate these findings. An example of returns of such a data series is shown in Figure 1. The corresponding sample autocorrelations are shown in Figure 2. A simple graphical analysis of these two figures suggests that there is no significant memory on

the series of returns. It is not surprising that the corrected R/S ($q \neq 0$) statistic, the GPH tests, and the LM tests reject unequivocally the existence of long-memory. Using the classical R/S statistic ($q = 0$), short memory is rejected for several series, although the estimated values of the Hurst exponent are only slightly larger than 0.5. Using the GPH and the LM tests, only in a few cases are there some apparently significant values. However, since 23 series are analyzed and six tests are conducted for each case, one or two apparently significant values should be taken as natural result of the random error due to the repetition of the same tests. Nevertheless, there is not a single case in which the different methods and different truncations simultaneously point to long memory.

The results are very clear: the indication of long memory given by the classical R/S is a statistical artifact of the short-range correlations present in the data.

TESTING FOR LONG-MEMORY IN VOLATILITY OF FUTURES' RETURNS

Although the lack of long-memory in the movement of financial markets is currently well-accepted, more recent research has focused on the existence of persistent behavior in the volatility of financial markets. Evidence of long-range dependence in volatilities has been found by many authors, including Ding, Granger and Engle (1993), de Lima and Crato (1993), Bollerslev and Mikkelsen (1996), Baillie, Bollerslev and Mikkelsen (1996), Breidt, Crato, and de Lima (1998) and Lobato and Savin (1998). However, a study of persistence in *futures'* volatility has not yet been presented in the literature. This analysis is important for forecasting, risk assessment, and portfolio optimization.

Various measures of volatility can be constructed, in particular the absolute value of returns, the square of returns, and the logarithm of the squared returns. Empirical findings indicate that evidence concerning long-range dependence is essentially the same for all these measures (see, *e.g.*, Breidt, Crato, and de Lima, 1998). This study analyzes the series of logarithms of the mean-corrected returns, $r_t = \log(r_t^* - \bar{r}^*)^2$, using the same three tests applied in Section 3. Breidt, Crato, and de Lima (1998) and Deo and Hurvich (1998) study the behavior of these tests for the detection of long memory in volatility data.

— — INSERT FIGURE 3 AROUND HERE — —

Figure 3 shows an example of such a volatility series. A simple graphical analysis reveals an interesting pattern, completely different from the one displayed by the series

of returns. Underneath the natural variability, there are apparent slow movements on the volatility levels.

Tables III and IV present the formal testing results for all the futures' contract volatility series. The contrast between the analysis of the returns and of the volatilities is striking. While returns do not display persistence, the volatilities show a strong and significant persistence for almost all studied series. For the Helms data, the only exceptions are two of the shorter series, soybeans for March 77 (BO M76) and soybean meal for the same date (SM M76), which have conflicting evidence of long-memory. Results for the extended data set, shown in Table IV, consistently indicate significant long-memory behavior.

— — *INSERT TABLES III AND IV AROUND HERE* — —

It is interesting to note from Table IV that the persistence showed by the commodity futures' volatility is substantially higher than the persistence shown by the currency futures' volatility. Higher estimates for the memory parameters J and d and higher values of the LM statistic t indicate more persistent memory in the time series. With the exception of the British Pound (BP), the estimates for the commodity futures are substantially higher than the estimates for the currency futures. This means that shocks driving volatility persist longer, on average, on the commodity futures' transactions. Markets seem to absorb more quickly the instability in currency futures than in commodity futures.

A VOLATILITY MODEL FOR FUTURES DATA

Breidt, Crato, and de Lima (1998) introduced the long-memory stochastic volatility (LMSV) model to describe the type of persistent dependence structure observed for the futures' volatility series. The model is an extension of the short-range dependent stochastic volatility model of Melino and Turnbull (1990) and Harvey, Ruiz, and Shephard (1994), which has been used to describe the behavior of log squared stock returns. A LMSV process is defined by

$$r_t = \sigma_t \psi_t, \quad \sigma_t = \sigma \exp(v_t/2), \quad (7)$$

where $\sigma > 0$, $\{v_t\}$ is independent of $\{\psi_t\}$, $\{\psi_t\}$ is a sequence of independent and identically distributed (*i.i.d.*) random variables with mean zero and variance one, and $\{v_t\}$ is a long-memory process having parameter d , with $0 < d < 0.5$. The fractionally integrated

noise model of Hosking (1981) and Granger and Joyeux (1980), *i.e.*, $(1 - B)^d v_t = \eta_t$, where B denotes the backward shift operator ($B^j x_t = x_{t-j}$) and η_t is a short-memory process, can be used to parameterize the long memory process. Typically, v_t is assumed to be Gaussian. The conditional variance σ_t^2 measures the volatility of r_t . By taking the logarithm of the squares of the mean corrected returns, $\{r_t\}$, we have

$$\begin{aligned} y_t = \log(r_t^2) &= [\log(\sigma^2) + \mathbf{E}(\log \psi_t^2)] + v_t + [\log(\psi_t^2) - \mathbf{E}(\log \psi_t^2)] \\ &\equiv \mu + v_t + \epsilon_t, \end{aligned} \tag{8}$$

i.e. $\{y_t\}$ is a Gaussian long-memory signal plus a non-Gaussian white noise. If ψ_t is assumed to be Gaussian, ϵ_t follows a $\log \chi^2(1)$ distribution with variance $\pi^2/2$.

Using this formulation, a fully parametric model can be fit to y_t by quasi-maximum likelihood estimation (for details, see Section 4 of Breidt, Crato, and de Lima 1998). As an illustration of this technique, an estimated LMSV model for the volatility of coffee futures (KC) is presented below. The volatility is modeled as

$$\log(r_t^2) = -0.000086 + v_t + \epsilon_t,$$

with ϵ_t independent and identically distributed with mean zero and variance 2.721, and v_t an ARFIMA(1, d ,1)

$$(1 - 0.308B)(1 - B)^{0.445} = (1 - 0.612B)\eta_t,$$

where η_t is a zero-mean white noise with variance 3.017.

Note the high estimated value of d , relatively close to the stationarity boundary 0.5, implying strong persistence in volatility. As seen in Figure 4, the autocorrelation structure of the data is closely matched by this model, providing further evidence for the long memory behavior of the volatility series.

— — INSERT FIGURE 4 AROUND HERE — —

CONCLUSIONS

This paper reexamines returns on futures' contracts using updated statistical tools. The results provide *no evidence of long memory on the series of returns*. This contradicts the findings of previous researchers, who have used results from classical R/S analysis

as evidence of the existence of long memory in futures' returns. On the other hand, the analysis of the *volatility of futures' returns* indicates a strongly persistent behavior. These are the first findings of this type using futures data. By showing that changes in the variance of financial time series are persistent, *i.e.*, that they remain important for forecasts in large horizons, these results indicate the need for new models for risk assessment of futures' portfolios. In order to adequately evaluate and forecast the risk of a portfolio, analysts should take into consideration the persistent volatility of these financial series.

An additional interesting conclusion deserving further investigation is the fact that currency futures' volatility seem to have a less persistent memory than commodity futures' volatility. In the latter case, shocks driving instability in the markets seem to be absorbed faster.

Acknowledgements

The work of Bonnie K. Ray was partially funded under NSF grant DMS-9623884. The authors gratefully acknowledge the computational help of students Chernok Chernilos, Steven Davidson, Jean-Charles Gresset, and Roxanne Sosa during the course of a statistics project concerning long memory time series.

References

- Andrews, D. W. Q. (1991): "Heteroskedasticity and autocorrelation consistent covariance matrix estimation," *Econometrica*, 59, 817-858.
- Baillie, R. T., Bollerslev, T., and Mikkelsen, H. O. (1996): "Fractionally integrated generalized autoregressive conditional heteroskedasticity," *Journal of Econometrics*, 74, 3-30.
- Ball, R. (1989): "What do we know about stock market 'efficiency'?", in Guimarães, Rui M. C., Kingsman, Brian and Taylor, Stephen, *A Reappraisal of the Efficiency of Financial Markets*, Springer-Verlag.
- Barkoulas, J. T., Labys, W. C., and Onochie, J. I. (1997): "Fractional dynamics in international commodity prices", *Journal of Futures Markets* 2, 737-745.

- Bollerslev, T. (1986): "Generalized autoregressive conditional heteroskedasticity," *Journal of Econometrics*, 31, 307-327.
- Bollerslev, T. and Mikkelsen H. O. A. (1996): "Modeling and pricing long-memory in stock market volatility," *Journal of Econometrics*, 73, 151-184.
- Booth, G., Kaen, F. R. and Koveos, P. E. (1982a): "Persistent dependence in gold prices", *The Journal of Financial Research* 5, 1, 85-93.
- Booth, G., Kaen, F. R. and Koveos, P. E. (1982b): "R/S analysis of foreign exchange rates under two international monetary regimes", *Journal of Monetary Economics* 10, 407-415.
- Breidt, F. J., Crato, N., and de Lima, P. J. F. (1998): "The detection and estimation of long memory in stochastic volatility," *Journal of Econometrics*, 83, 325-348.
- Brockwell, P. J. and Davis, R. A. (1991): *Time Series: Theory and Methods*, second edition, Springer-Verlag.
- Cheung, Y.-W. and Lai, K. S. (1993): "Do gold market returns have long memory?", *The Financial Review* 28, 181-202.
- Crato, N. (1994): "Some international evidence regarding the stochastic memory of stock returns," *Applied Financial Economics* 4, 1, 33-39.
- De Lima, P. and Crato, N. (1993): "Long-memory in stock returns and volatilities," American Statistical Association, Proceedings of the Business and Economic Statistics Section.
- Deo, R. S. and Hurvich C. M. (1998): "On the log periodogram regression estimator of the memory parameter in long memory stochastic volatility models," Stern School of Business, New York University, mimeo.
- Ding, Z., Granger, C., and Engle, R. F. (1993): "A long memory property of stock market returns and a new model," *Journal of Empirical Finance*, 1, 83-106.
- Doukas, J. (1990): "On unit roots and the persistent dependence of futures prices," *Advances in Econometrics* 8, JAI Press, 295-306.
- Engle, R. F. (1982): "Autoregressive conditional heteroskedasticity with estimates of the variance of United Kingdom inflation," *Econometrica*, 50, 987-1007.
- Fung, H. G. and Lo, W. C. (1993): "Memory in interest rate futures", *The Journal of Futures Markets* 13, 865-872.
- Geweke, J. and Porter-Hudak, S. (1983): "The estimation and application of long memory time series models", *Journal of Time Series Analysis* 4, 4, 221-238.

- Granger, C. W. J. and Joyeux, Roselyne (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**, 339-349.
- Harvey, A. C., Ruiz, E., and Shephard, N. (1994): "Multivariate stochastic variance models," *Review of Economic Studies*, 61: 247-264.
- Helms, B. P., Kaen, F. R. and Rosenman, R. E. (1984): "Memory in commodity futures contracts", *Journal of Futures Markets* **4**, 4, 559-567.
- Hosking, J. R. M. (1981): "Fractional differencing", *Biometrika* **68**, 1, 165-176.
- Hurvich, C. M., Deo, R. S., and Brodsky, J. (1998): "The mean squared 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.
- Hurst, H. E. (1951): "Long-term storage capacity of reservoirs", *Transactions of the American Society of Civil Engineers* **116**, 770-808.
- Kennedy, D. (1976): The distribution of the maximum Brownian excursion. *Journal of Applied Probability* **13**, 371-376.
- Lo, A. W. (1991): "Long-term memory in stock market prices", *Econometrica* **59**, 1279-1313.
- Lobato, I. N. and Robinson, Peter M. (1998): "A nonparametric test for $I(0)$ ", *Review of Economic Studies*, 65 : 475-495.
- Lobato, I. N. and Savin, N. E. (1998): "Real and spurious long memory properties of stock market data", *Journal of Business and Economic Statistics*, 16: 261-267.
- Mandelbrot, 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**, 3274-3277.
- Mandelbrot, B. (1972): "Statistical methodology for nonperiodic cycles: from the covariance to R/S analysis", *Annals of Economic and Social Measurement* **1**, 259-290.
- Mandelbrot, 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.*, vol. 48, 2, 69-104.
- Newey, W. K. and West, K. D. (1987): "A simple, positive semi-definite, heteroskedasticity and autocorrelation consistent covariance matrix", *Econometrica* **55**, 3, 703-708.
- Robinson, Peter M. (1995): "Log-periodogram regression of time series with long range dependence", *Annals of Statistics* **23**, 1048-1072.

Siddiqui, M. (1976): The asymptotic distribution of the range and other functions of partial sums of stationary processes. *Water Resources Research* **12**, 1271-1276.

APPENDIX

Description of the Extended Dataset

Commodity	Symbol	Exchange	Start date	End date	Contract Months
Soybean Oil	BO	CBOT	01/03/77	11/28/97	1, 3, 5, 7, 8, 9, 10, 12
British Pound	BP	CME	01/03/77	11/28/97	1, 3, 4, 6, 7, 9, 10, 12
Corn	C	CBOT	01/03/77	11/28/97	3, 5, 7, 9, 12
Canadian Dollar	CD	CME	01/03/77	11/28/97	1, 3, 4, 6, 7, 9, 10, 12
Cotton	CT	NYCE	01/03/77	11/26/97	3, 5, 7, 10, 12
Deutsche Mark	DM	CME	01/03/77	11/28/97	1, 3, 4, 6, 7, 9, 10, 12
Feeder Cattle	FC	CME	01/03/77	11/28/97	1, 3, 4, 5, 8, 9, 10, 11
Gold	GC	COMEX	01/03/77	11/26/97	2, 4, 6, 8, 10, 12
Copper, HighGrade	HG	COMEX	01/03/77	11/26/97	ALL
Heating Oil	HO	NYMEX	11/15/78	11/26/97	ALL
Japanese Yen	JY	CME	01/03/77	11/28/97	1, 3, 4, 6, 7, 9, 10, 12
Coffee "C"	KC	CSCE	01/03/77	11/26/97	4, 5, 7, 9, 12
Live Cattle	LC	CME	01/03/77	11/28/97	2, 4, 6, 8, 10, 12
Live Hogs	LH	CME	01/03/77	11/26/97	2, 4, 6, 7, 8, 10, 12
Oats	O	CBOT	01/03/77	11/28/97	3, 5, 7, 9, 12
Pork Bellies, Fresh	PB	CME	01/03/77	11/28/97	1, 3, 5, 7, 8, 9, 11
Soybean	S	CBOT	01/03/77	11/28/97	1, 3, 5, 7, 8, 9, 11
Sugar #11	SB	CSCE	01/03/77	11/26/97	3, 5, 7, 10
Swiss Franc	SF	CME	01/03/77	11/28/97	1, 3, 4, 6, 7, 9, 10, 12
Silver	SI	COMEX	01/03/77	11/26/97	3, 5, 7, 9, 12
Soybean Meal	SM	CBOT	01/03/77	11/28/97	1, 3, 5, 7, 8, 9, 10, 12
S&P 500 Index	SP	CME	04/21/82	11/28/97	3, 6, 9, 12

CME = Chicago Mercantile Exchange, CBOT = Chicago Board of Trade, COMEX = Commodity Exchange, NYCE = New York Cotton Exchange, NYMEX = New York Mercantile Exchange, CSCE = Coffee, Sugar, and Cocoa Exchange.

Table I: Memory Estimates for the Helms Data Returns

Series	Hurst estimates \hat{J}			GPH estimates \hat{d}			Nonparametric LM test		
	$q = 0$	$q = 5$	$q = q^*$	$m = [n^{0.50}]$ $l = 1$	$m = [n^{0.60}]$ $l = 1$	$m = [n^{0.60}]$ $l = 2$	$m = [n^{.50}]$	$m = [n^{.55}]$	$m = [n^{.60}]$
BO M76	0.536 [0.500]	0.540 [0.460]	0.534 [0.525]	0.211 (0.182)	0.189 (0.190)	-0.223 (0.110)	0.029 [0.488]	-0.199 [0.579]	-0.431 [0.667]
BO J77	0.506 [0.781]	0.513 [0.721]	0.515 [0.706]	0.136 (0.161)	0.176 (0.125)	0.164 (0.253)	-0.310 [0.622]	-0.135 [0.554]	-0.109 [0.543]
SO M76	0.525 [0.605]	0.529 [0.571]	0.528 [0.582]	0.021 (0.131)	-0.076 (0.083)	0.195 (0.272)	0.779 [0.218]	0.193 [0.424]	0.288 [0.387]
SO J77	0.541 [0.440]	0.536 [0.492]	0.541 [0.440]	0.265 (0.296)	0.057 (0.172)	0.230 (0.178)	-0.001 [0.506]	0.313 [0.377]	0.177 [0.430]
SM M76	0.532 [0.549]	0.557 [0.319]	0.540 [0.480]	0.118 (0.228)	-0.032 (0.152)	-0.239 (0.221)	-0.255 [0.601]	-0.081 [0.532]	-0.643 [0.740]
SM J77	0.546 [0.407]	0.555 [0.328]	0.553 [0.347]	0.015 (0.113)	0.183 (0.135)	0.264 (0.198)	-0.001 [0.506]	0.312 [0.377]	0.177 [0.430]

Values in brackets below the estimates for J (Hurst estimates) and below the statistics t (LM test) are the p -values for the one-sided test for persistent long memory. Values in parentheses below the estimates for d (GPH estimates) are the standard errors for the estimates.

Table II: Memory Estimates for the Extended Data Returns

Series	Hurst estimates \hat{J}			GPH estimates \hat{d}			Nonparametric LM test		
	$q = 0$	$q = 5$	$q = q^*$	$m = [n^{0.50}]$ $l = 1$	$m = [n^{0.60}]$ $l = 1$	$m = [n^{0.60}]$ $l = 3$	$m = [n^{.50}]$	$m = [n^{.55}]$	$m = [n^{.60}]$
BO	0.480 [0.961]	0.478 [0.969]	0.477 [0.970]	-0.165 (0.069)	0.010 (0.053)	0.095 (0.057)	-0.804 [0.789]	0.204 [0.419]	0.384 [0.351]
BP	0.562 [0.067]	0.562 [0.066]	0.561 [0.072]	0.033 (0.075)	0.059 (0.053)	0.035 (0.059)	0.757 [0.225]	1.452 [0.073]	0.928 [0.177]
JC	0.533 [0.351]	0.506 [0.756]	0.506 [0.756]	0.043 (0.085)	-0.003 (0.054)	0.012 (0.061)	-0.297 [0.617]	-0.732 [0.768]	-0.768 [0.779]
CD	0.569 [0.037]	0.568 [0.039]	0.568 [0.038]	-0.022 (0.085)	-0.038 (0.051)	-0.087 (0.056)	-0.152 [0.560]	0.010 [0.496]	-0.292 [0.615]
CT	0.559 [0.080]	0.531 [0.384]	0.531 [0.388]	0.030 (0.077)	-0.020 (0.046)	-0.051 (0.051)	0.492 [0.311]	-0.201 [0.580]	-0.198 [0.578]
DM	0.552 [0.132]	0.549 [0.160]	0.552 [0.132]	0.049 (0.065)	0.032 (0.045)	0.019 (0.050)	1.080 [0.140]	0.679 [0.249]	0.945 [0.172]
FC	0.570 [0.032]	0.560 [0.073]	0.562 [0.066]	0.092 (0.094)	-0.026 (0.057)	-0.047 (0.064)	0.540 [0.295]	0.477 [0.317]	-0.372 [0.645]
GC	0.575 [0.019]	0.576 [0.017]	0.577 [0.016]	-0.019 (0.072)	0.152 (0.052)	0.162 (0.058)	-0.773 [0.780]	1.193 [0.116]	2.518 [0.006]
HG	0.504 [0.778]	0.504 [0.772]	0.503 [0.786]	-0.025 (0.073)	0.016 (0.048)	0.023 (0.053)	-0.155 [0.562]	-0.787 [0.784]	0.109 [0.457]
HO	0.518 [0.595]	0.517 [0.607]	0.513 [0.655]	-0.034 (0.089)	0.048 (0.055)	0.041 (0.061)	-0.869 [0.808]	-0.215 [0.585]	0.505 [0.307]
JY	0.558 [0.085]	0.555 [0.113]	0.558 [0.085]	0.056 (0.076)	0.144 (0.053)	0.169 (0.059)	0.769 [0.221]	0.955 [0.170]	2.834 [0.002]
KC	0.499 [0.831]	0.502 [0.797]	0.500 [0.826]	0.024 (0.089)	-0.004 (0.056)	0.014 (0.062)	0.511 [0.305]	-0.343 [0.634]	-0.127 [0.551]
LC	0.510 [0.700]	0.509 [0.715]	0.508 [0.725]	-0.015 (0.091)	-0.078 (0.055)	-0.095 (0.061)	-0.664 [0.747]	-0.826 [0.796]	-1.707 [0.956]
LH	0.499 [0.831]	0.502 [0.797]	0.500 [0.826]	0.024 (0.089)	-0.004 (0.056)	0.014 (0.062)	0.511 [0.305]	-0.343 [0.634]	-0.127 [0.551]
JO	0.520 [0.550]	0.520 [0.560]	0.518 [0.584]	-0.091 (0.092)	-0.098 (0.056)	-0.049 (0.062)	-0.657 [0.744]	-1.568 [0.942]	-1.247 [0.894]
PB	0.501 [0.814]	0.488 [0.919]	0.490 [0.910]	-0.104 (0.080)	-0.069 (0.060)	-0.077 (0.067)	-0.751 [0.774]	-0.973 [0.835]	-1.235 [0.892]
JS	0.487 [0.926]	0.491 [0.905]	0.487 [0.930]	-0.040 (0.075)	0.074 (0.053)	0.122 (0.058)	-0.148 [0.559]	0.933 [0.175]	1.456 [0.073]
SB	0.513 [0.658]	0.519 [0.569]	0.511 [0.681]	-0.024 (0.078)	-0.013 (0.051)	0.016 (0.057)	-0.412 [0.660]	-0.012 [0.505]	-0.767 [0.779]
SF	0.544 [0.209]	0.542 [0.232]	0.544 [0.216]	0.007 (0.085)	0.016 (0.048)	0.024 (0.054)	0.905 [0.183]	0.863 [0.194]	1.542 [0.062]
SI	0.556 [0.102]	0.550 [0.149]	0.551 [0.143]	0.006 (0.067)	0.101 (0.045)	0.107 (0.050)	0.217 [0.414]	0.931 [0.176]	1.628 [0.052]
SM	0.505 [0.769]	0.505 [0.765]	0.503 [0.790]	-0.022 (0.084)	0.035 (0.054)	0.084 (0.060)	-0.205 [0.581]	0.349 [0.364]	0.694 [0.244]
SP	0.497 [0.848]	0.505 [0.762]	0.497 [0.848]	-0.001 (0.091)	-0.080 (0.051)	-0.107 (0.058)	-0.525 [0.700]	-0.446 [0.672]	-1.650 [0.951]
JW	0.532 [0.370]	0.532 [0.366]	0.531 [0.384]	-0.025 (0.079)	-0.024 (0.048)	0.010 (0.053)	-0.551 [0.709]	-0.375 [0.646]	-0.880 [0.811]

Values in brackets below the estimates for J (Hurst estimates) and below the statistics t (LM test) are the p -values for the one-sided test for persistent long memory. Values in parentheses below the estimates for d (GPH estimates) are the standard errors for the estimates.

Table III: Memory Estimates for the Helms Data Volatility

Series	Hurst estimates \hat{J}			GPH estimates \hat{d}			Nonparametric LM test		
	$q = 0$	$q = 5$	$q = q^*$	$m = [n^{0.50}]$ $l = 1$	$m = [n^{0.60}]$ $l = 1$	$m = [n^{0.60}]$ $l = 2$	$m = [n^{.50}]$	$m = [n^{.55}]$	$m = [n^{.60}]$
BO M76	0.600 [0.050]	0.585 [0.109]	0.604 [0.040]	0.162 (0.149)	0.158 (0.149)	0.382 (0.151)	1.737 [0.04]	1.400 [0.081]	2.328 [0.010]
BO J77	0.706 [0.000]	0.658 [0.001]	0.700 [0.000]	0.459 (0.168)	0.193 (0.110)	0.172 (0.137)	3.870 [0.000]	4.650 [0.000]	4.853 [0.000]
SO M76	0.615 [0.018]	0.584 [0.108]	0.602 [0.041]	0.454 (0.165)	0.347 (0.123)	0.412 (0.123)	2.029 [0.020]	2.604 [0.005]	3.157 [0.001]
SO J77	0.711 [0.000]	0.673 [0.000]	0.692 [0.000]	0.454 (0.165)	0.347 (0.123)	-0.001 (0.116)	4.894 [0.000]	5.413 [0.000]	5.767 [0.000]
SM M76	0.597 [0.073]	0.581 [0.142]	0.590 [0.098]	0.162 (0.149)	0.158 (0.149)	0.077 (0.184)	0.572 [0.280]	0.158 [0.437]	0.592 [0.277]
SM J77	0.728 [0.000]	0.659 [0.000]	0.659 [0.000]	0.459 (0.168)	0.193 (0.110)	0.179 (0.191)	4.277 [0.000]	5.295 [0.000]	5.953 [0.000]

Returns volatility is computed as the series of logs of mean-corrected squared returns. Values in brackets below the estimates for J (Hurst estimates) and below the statistics t (LM test) are the p -values for the one-sided test for persistent long memory. Values in parentheses below the estimates for d (GPH estimates) are the standard errors for the estimates.

Table IV: Memory Estimates for the Extended Data Volatility

Series	Hurst estimates \hat{J}			GPH estimates \hat{d}			Nonparametric LM test		
	$q = 0$	$q = 5$	$q = q^*$	$m = [n^{0.50}]$ $l = 1$	$m = [n^{0.60}]$ $l = 1$	$m = [n^{0.60}]$ $l = 3$	$m = [n^{.50}]$	$m = [n^{.55}]$	$m = [n^{.60}]$
BO	0.697	0.663	0.678	0.367	0.313	0.283	8.635	12.405	15.570
	[0.000]	[0.000]	[0.000]	(0.078)	(0.047)	(0.053)	[0.000]	[0.000]	[0.000]
BP	0.732	0.686	0.693	0.627	0.384	0.345	13.123	17.258	21.944
	[0.000]	[0.000]	[0.000]	(0.073)	(0.047)	(0.052)	[0.000]	[0.000]	[0.000]
JC	0.642	0.594	0.601	0.339	0.359	0.414	4.600	7.908	12.156
	[0.000]	[0.002]	[0.001]	(0.090)	(0.056)	(0.061)	[0.000]	[0.000]	[0.000]
CD	0.643	0.608	0.619	0.424	0.323	0.337	6.251	8.694	11.219
	[0.000]	[0.000]	[0.000]	(0.071)	(0.049)	(0.055)	[0.000]	[0.000]	[0.000]
CT	0.663	0.627	0.634	0.363	0.269	0.293	7.201	9.213	11.230
	[0.000]	[0.000]	[0.000]	(0.071)	(0.045)	(0.050)	[0.000]	[0.000]	[0.000]
DM	0.699	0.666	0.681	0.444	0.395	0.381	9.156	11.050	15.189
	[0.000]	[0.000]	[0.000]	(0.066)	(0.046)	(0.052)	[0.000]	[0.000]	[0.000]
FC	0.772	0.718	0.721	0.342	0.349	0.311	15.576	21.068	27.603
	[0.000]	[0.000]	[0.000]	(0.082)	(0.048)	(0.052)	[0.000]	[0.000]	[0.000]
GC	0.787	0.721	0.717	0.607	0.385	0.312	19.331	25.999	33.522
	[0.000]	[0.000]	[0.000]	(0.078)	(0.050)	(0.054)	[0.000]	[0.000]	[0.000]
HG	0.691	0.651	0.660	0.436	0.335	0.315	11.155	14.857	19.366
	[0.000]	[0.000]	[0.000]	(0.079)	(0.048)	(0.054)	[0.000]	[0.000]	[0.000]
HO	0.734	0.677	0.677	0.506	0.490	0.486	9.584	13.485	19.143
	[0.000]	[0.000]	[0.000]	(0.082)	(0.051)	(0.058)	[0.000]	[0.000]	[0.000]
JY	0.661	0.629	0.638	0.259	0.258	0.279	5.855	7.328	10.661
	[0.000]	[0.000]	[0.000]	(0.084)	(0.048)	(0.052)	[0.000]	[0.000]	[0.000]
KC	0.714	0.663	0.666	0.424	0.393	0.378	9.510	13.264	17.352
	[0.000]	[0.000]	[0.000]	(0.096)	(0.059)	(0.066)	[0.000]	[0.000]	[0.000]
LC	0.739	0.701	0.713	0.498	0.339	0.294	14.911	19.177	23.732
	[0.000]	[0.000]	[0.000]	(0.080)	(0.052)	(0.058)	[0.000]	[0.000]	[0.000]
LH	0.714	0.663	0.666	0.424	0.393	0.378	9.510	13.264	17.352
	[0.000]	[0.000]	[0.000]	(0.096)	(0.059)	(0.066)	[0.000]	[0.000]	[0.000]
JO	0.637	0.597	0.602	0.428	0.263	0.285	5.894	9.056	12.055
	[0.000]	[0.000]	[0.000]	(0.105)	(0.056)	(0.061)	[0.000]	[0.000]	[0.000]
PB	0.657	0.630	0.643	0.303	0.242	0.202	5.741	8.042	11.105
	[0.000]	[0.000]	[0.000]	(0.088)	(0.052)	(0.057)	[0.000]	[0.000]	[0.000]
JS	0.717	0.670	0.680	0.397	0.408	0.386	9.156	13.183	18.259
	[0.000]	[0.000]	[0.000]	(0.080)	(0.052)	(0.058)	[0.000]	[0.000]	[0.000]
SB	0.774	0.722	0.722	0.468	0.292	0.229	17.011	20.767	26.254
	[0.000]	[0.000]	[0.000]	(0.092)	(0.053)	(0.058)	[0.000]	[0.000]	[0.000]
SF	0.672	0.646	0.661	0.441	0.389	0.382	8.216	11.005	15.066
	[0.000]	[0.000]	[0.000]	(0.079)	(0.054)	(0.060)	[0.000]	[0.000]	[0.000]
SI	0.751	0.700	0.703	0.488	0.368	0.320	14.114	19.245	24.743
	[0.000]	[0.000]	[0.000]	(0.088)	(0.052)	(0.057)	[0.000]	[0.000]	[0.000]
SM	0.697	0.652	0.658	0.300	0.309	0.277	7.944	11.597	15.288
	[0.000]	[0.000]	[0.000]	(0.085)	(0.046)	(0.051)	[0.000]	[0.000]	[0.000]
SP	0.708	0.673	0.693	0.625	0.378	0.335	11.430	14.381	18.323
	[0.000]	[0.000]	[0.000]	(0.089)	(0.055)	(0.062)	[0.000]	[0.000]	[0.000]
JW	0.632	0.609	0.622	0.230	0.208	0.206	5.130	6.521	8.704
	[0.000]	[0.000]	[0.000]	(0.082)	(0.048)	(0.054)	[0.000]	[0.000]	[0.000]

The volatility is computed as the series of logs of mean-corrected squared returns. Values in brackets below the estimates for J (Hurst estimates) and below the statistics t (LM test) are the p -values for the one-sided test for persistent long memory. Values in parentheses below the estimates for d (GPH estimates) are the standard errors for the estimates.