

Board of Governors of the Federal Reserve System

International Finance Discussion Papers

Number 685

November 2000

LOG-PERIODOGRAM ESTIMATION OF LONG MEMORY VOLATILITY DEPENDENCIES WITH
CONDITIONALLY HEAVY TAILED RETURNS

Jonathan H. Wright

NOTE: International Finance Discussion Papers are preliminary materials circulated to stimulate discussion and critical comment. References in publications to International Finance Discussion Papers (other than an acknowledgment that the writer has had access to unpublished material) should be cleared with the author or authors. Recent IFDPs are available on the Web at www.bog.frb.fed.us.

LOG-PERIODOGRAM ESTIMATION OF LONG MEMORY VOLATILITY DEPENDENCIES WITH CONDITIONALLY HEAVY TAILED RETURNS

Jonathan H. Wright*

Abstract: Many recent papers have used semiparametric methods, especially the log-periodogram regression, to detect and estimate long memory in the volatility of asset returns. In these papers, the volatility is proxied by measures such as squared, log-squared and absolute returns. While the evidence for the existence of long memory is strong using any of these measures, the actual long memory parameter estimates can be sensitive to which measure is used. In Monte-Carlo simulations, I find that the choice of volatility measure makes little difference to the log-periodogram regression estimator if the data is Gaussian conditional on the volatility process. But, if the data is conditionally leptokurtic, the log-periodogram regression estimator using squared returns has a large downward bias, which is avoided by using other volatility measures. In U.S. stock return data, I find that squared returns give much lower estimates of the long memory parameter than the alternative volatility measures, which is consistent with the simulation results. I conclude that researchers should avoid using the squared returns in the semiparametric estimation of long memory volatility dependencies.

Keywords: Semiparametric Methods, Fractional Integration, Stochastic Volatility, Stock Returns, Heavy Tails.

* Wright is an economist in the International Finance Division of the Federal Reserve Board. I am grateful to two anonymous referees for extremely helpful comments on an earlier draft. The views in this paper are solely the responsibility of the author and should not be interpreted as reflecting the views of the Board of Governors of the Federal Reserve System or of any person associated with the Federal Reserve System.

1. Introduction.

While asset returns appear to be close to being a martingale difference sequence, they are nevertheless characterized by a high degree of volatility clustering and so are not even close to being independently and identically distributed (iid). Persistence in the volatility of asset returns has been documented in papers going back to Fama (1965), in the enormous ARCH/GARCH literature pioneered by Engle (1982) and surveyed by Bollerslev, Engle and Nelson (1994), and in the closely related stochastic volatility literature (see, for example, Taylor (1994)). Recently, a number of papers have noted that the autocorrelation functions of squared, log-squared and absolute returns are best characterized by a slowly mean-reverting hyperbolic rate of decay. These papers include Andersen and Bollerslev (1997), Bollerslev and Wright (2000), Dacorogna et al. (1993), Ding, Granger and Engle (1993), Granger and Ding (1996), Lobato and Savin (1998) and Robinson (1991). This property of the autocorrelation function of squared, log-squared and absolute returns has been found in many exchange rate and stock returns, but is not consistent with the standard ARCH/GARCH or stochastic volatility models.

Long memory models are designed to represent time series with autocorrelations functions exhibiting a slow hyperbolic rate of decay. The leading long memory time series model is the fractionally integrated ARMA (ARFIMA) model (Granger and Joyeux (1980) and Hosking (1981)). Models with long memory in the volatil-

ity process have been proposed and found to match the autocorrelation functions of squared, log-squared and absolute asset returns. These include the long memory ARCH model in Robinson (1991), the fractionally integrated GARCH, or FIGARCH, model in Bollerslev and Mikkelsen (1996) and Baillie, Bollerslev and Mikkelsen (1996) and the fractionally integrated stochastic volatility model in Breidt, Crato and de Lima (1998). These models can imply that the autocorrelation functions of squared, log-squared and/or absolute returns have the same hyperbolic rate of decay as the volatility process. Many papers have sought to estimate the long memory parameter in the volatility process without specifying a particular parametric model by applying a semiparametric estimator to one or more of these volatility measures. These papers include Bollerslev and Wright (2000), Breidt, Crato and de Lima (1998), Granger and Ding (1996), Lobato and Robinson (1998) and Lobato and Savin (1998).

The precise estimates of the long memory parameter can however be sensitive to the choice of volatility measure. In this paper I show that for daily U.S. stock return data, there is considerable sensitivity to the choice of volatility measure and that using squared returns gives much lower estimates of the long memory parameter than the other alternatives. I report the results of a Monte-Carlo simulation in which the data is generated from various long memory volatility models, and different volatility measures are used in the log-periodogram regression to estimate the degree of fractional integration. If the returns are Gaussian, conditional on the volatility

process, then the results are insensitive to the choice of volatility measure. But it is widely agreed that asset return data is not even conditionally Gaussian, so I also simulate data which is conditionally heavy tailed (t distributed on v degrees of freedom). In this case, the simulated bias is sensitive to the choice of volatility measure. In particular, if squared data is used, there is a large downward bias in the estimated degree of fractional integration. This ties in with the empirical findings.

The plan for the remainder of this paper is as follows. Section 2 describes some long memory volatility models and semiparametric estimators. The empirical work is in Section 3. Section 4 contains the Monte-Carlo results. Section 5 concludes.

2. Long Memory in Volatility

2.1 *Long Memory and Fractional Integration*

The ideas of fractional integration and long memory were introduced to econometrics by papers including Granger and Joyeux (1980) and Mandelbrot (1971). Recent and comprehensive reviews of this topic are provided by Baillie (1996) and Robinson (1994). Long memory time series are characterized by a slow rate of decay of the autocovariance function. A time series $\{h_t\}_{t=1}^T$ is said to exhibit long memory if it has a covariance function $\gamma(j)$ that is of the same order as j^{2d-1} (as $j \rightarrow \infty$), and a spectrum $f(\lambda)$ that is of the same order as λ^{-2d} (as $\lambda \rightarrow 0$), for $0 < d < 0.5$. The

definition of long memory can be extended to the case $d > 0.5$ or $d < 0$ (negative long memory). The ARFIMA model is the leading parametric long memory time series model. Define $\{h_t\}_{t=1}^T$ to be an ARFIMA(p, d, q) process if

$$a(L)(1 - L)^d h_t = b(L)u_t \quad (1)$$

where u_t is i.i.d. with mean zero and variance σ^2 , L is the lag operator, the lag polynomials $a(L)$ and $b(L)$ (of orders p and q , respectively) are assumed to have all of their roots outside the unit circle, d is any number on the real line and $(1 - L)^d$ is defined by the usual binomial expansion. For $d < 0.5$, $\{h_t\}$ is stationary, and its covariance function and spectrum have the required rates of decay.

2.2 *Fractionally Integrated Stochastic Volatility*

The slow rate of decay of the autocorrelations of log-squared, squared and absolute returns motivates the construction of models with long memory in the volatility process. The fractionally integrated stochastic volatility (FISV) model is one such specification, proposed by Breidt, Crato and de Lima (1998). This specifies that $\{y_t\}_{t=1}^T$ is a time series of asset returns such that

$$y_t = \exp(h_t/2)\sigma_\varepsilon\varepsilon_t, \quad (2)$$

where h_t is an ARFIMA process, as defined in equation (1) and ε_t is an i.i.d. sequence with unit variance that is independent of u_t . It is further assumed that u_t is Gaussian

and that $E \log(\varepsilon_t^2) < \infty$ (but ε_t is not generally assumed to be Gaussian in this paper).

The log-squared returns have the same persistence properties as the latent volatility process, h_t , since $\log(y_t^2) = \log(\sigma^2) + E \log(\varepsilon_t^2) + h_t + \xi_t$, where $\xi_t = \log(\varepsilon_t^2) - E \log(\varepsilon_t^2)$, and so, for $d < 0.5$,

$$\text{Cov}(\log(y_t^2), \log(y_{t-j}^2)) \sim j^{2d-1}, \quad (3)$$

as $j \rightarrow \infty$ ¹. If, furthermore, ε_t is Gaussian then Andersen and Bollerslev (1997) and Robinson (1999) show that

$$\text{Cov}(y_t^2, y_{t-j}^2) \sim j^{2d-1}, \quad (4)$$

and

$$\text{Cov}(|y_t|, |y_{t-j}|) \sim j^{2d-1}, \quad (5)$$

as $j \rightarrow \infty$, for $d < 0.5$. Equations 3-5 form the basis for applying semiparametric estimators of the long memory parameter to log-squared, squared or absolute returns so as to estimate the degree of long memory in the latent volatility process². I conjecture that equations (4) and (5) continue to hold with some nonnormal specifications for ε_t .

¹Throughout this paper, the \sim notation means that the limit of the ratio of the quantities on the left and right hand sides of the symbol is a finite positive constant.

²The results in Robinson (1999) also apply to the autocovariance function of $|y_t|^\lambda$, for noninteger λ , a transformation considered by Ding, Granger and Engle (1993).

2.3 The FIGARCH and FIEGARCH Models

Fractional integration can also be introduced into models in the ARCH/GARCH family. These models specify that $\{y_t\}_{t=1}^T$ is a time series of asset returns such that

$$y_t = \sigma_t \varepsilon_t, \quad (6)$$

where ε_t is an i.i.d. sequence with unit variance. The FIGARCH(1,d,1) model, proposed by Bollerslev and Mikkelsen (1996) specifies that

$$\sigma_t^2 = \omega + \beta \sigma_{t-1}^2 + [1 - \beta L - (1 - \phi L)(1 - L)^d] y_t^2, \quad (7)$$

where the parameters satisfy restrictions which ensure that σ_t^2 is positive almost surely for all t (Bollerslev and Mikkelsen provide explicit sufficient conditions). For $d = 0$, this reduces to the standard GARCH(1,1) specification, as $\sigma_t^2 = \omega + \beta \sigma_{t-1}^2 + (\phi - \beta) y_{t-1}^2$. Higher order FIGARCH models can, of course, be constructed in the same way.

Stock return volatility appears to be negatively correlated with lagged returns. This so-called leverage effect was first noted by Black (1976) and is not generated by the GARCH or FIGARCH models. This motivated Nelson (1991) to propose the exponential GARCH, or EGARCH model. This can also be generalized to incorporate fractional integration in the FIEGARCH(1,d,1) model, proposed by Bollerslev and Mikkelsen (1996), which specifies that

$$\ln(\sigma_t^2) = \omega(1 - \phi) + \phi \ln(\sigma_{t-1}^2) + (1 + \psi L)(1 - L)^{-d} (\theta \varepsilon_{t-1} + \gamma [|\varepsilon_{t-1}| - E|\varepsilon_{t-1}|]). \quad (8)$$

These FIGARCH and FIEGARCH models may easily be estimated by a quasi-maximum-likelihood estimator. For a number of stock and exchange rate returns Bollerslev and Mikkelsen (1996) and Baillie, Bollerslev and Mikkelsen (1996) find that d is significantly different both from 0 and from 1. These models can also be used to motivate semiparametric estimation of the long memory volatility dependencies, since $Cov(y_t^2, y_{t-j}^2) \sim j^{2d-1}$ and $Cov(\log(y_t^2), \log(y_{t-j}^2)) \sim j^{2d-1}$, as $j \rightarrow \infty$, if $d < 0.5$, in the FIGARCH and FIEGARCH models, respectively.

2.4 *Semiparametric Estimators of the Long Memory Parameter*

The fully efficient estimation of any long memory time series model naturally requires that the model be completely specified. But if the researcher is interested mainly in the long memory parameter d , this may be estimated by semiparametric methods. Of these, the log-periodogram regression estimator is the most widely used. It was first proposed by Geweke and Porter-Hudak (1983), but the formal proofs of its consistency and asymptotic normality were elusive until these were provided by Robinson (1995a), in the Gaussian case.

Let $I(\lambda_j)$ denote the sample periodogram of an arbitrary long memory time series, z_t , where $\lambda_j = 2\pi j/T$ denotes the j th Fourier frequency. The log-periodogram regression (GPH) estimator of d , or \hat{d} , is then given by minus the estimate of β_1 in the regression equation

$$\log(I(\lambda_j)) = \beta_0 + \beta_1 \log[4\sin^2(\lambda_j/2)] + \zeta_j, \quad j = 1, \dots, m. \quad (9)$$

Robinson (1995a) showed that if z_t is Gaussian, $-0.5 < d < 0.5$, the bandwidth parameter m converges to infinity but is $o(T)$, and mild additional conditions are satisfied³, then $\sqrt{m}(\hat{d} - d) \rightarrow N(0, \frac{\pi^2}{24})$. The log-periodogram estimate is therefore consistent and asymptotically normal, but converges at a rate slower than the usual \sqrt{T} . Alternatively, the variance of \hat{d} may be consistently estimated by the conventional OLS variance estimate from the regression in equation (9). Some recent literature has extended Robinson's results. Deo and Hurvich (1999) proved the consistency and asymptotic normality of \hat{d} for certain nonnormal time series. Velasco (1999) proved the consistency of \hat{d} in the case $0.5 \leq d < 1$ and its asymptotic normality in the case $0.5 \leq d < 0.75$.

Robinson (1995b) proposed an alternative semiparametric estimator of d , which is asymptotically more efficient and the properties of which can be established under milder conditions (Robinson (1995b), Robinson and Henry (1999)). This estimator, called the Gaussian semiparametric (GSP) estimator, is $\tilde{d} = \arg \min_d R(d)$, where $R(d) = \log(\frac{1}{m} \sum_{j=1}^m \lambda_j^{2d} I(\lambda_j)) - \frac{2d}{m} \sum_{j=1}^m \log \lambda_j$. Robinson (1995b) proved that $\sqrt{m}(\hat{d}$

³Strictly, the proofs in Robinson (1995a) required some of the very lowest frequencies to be omitted from the regression in equation (9). Subsequently, Hurvich, Deo and Brodsky (1998) provided a version of this proof which does not require any such trimming.

$-d) \rightarrow N(0, \frac{1}{4})$, under conditions given in that paper.

The models of subsections 2.2 and 2.3 can imply that the autocovariance functions of squared, log-squared and absolute returns have the same long memory property as the underlying volatility process. This motivates researchers to apply semi-parametric methods such as \hat{d} and \tilde{d} to the squared, log-squared or absolute returns in a bid to estimate the long memory in the volatility dependencies without having to specify a complete parametric model. There is little theoretical reason to prefer one volatility measure over any of the others. Lobato and Robinson (1998) and Lobato and Savin (1998) use squared returns, Granger and Ding (1996) use absolute returns, Breidt, Crato and de Lima (1998) use log-squared returns and Bollerslev and Wright (2000) use all three of these volatility measures. The focus of the present paper is on examining which is the appropriate volatility measure to use in practice, when estimating long memory in the volatility of asset returns.

3. Empirical Results with U.S. Stock Market Data

In this section, I apply the log-periodogram regression to volatility measures constructed from the demeaned continuously compounded total returns on the value-weighted and equal-weighted daily stock indices, obtained from the Center for Research in Securities Prices (CRSP). The data covers the period July 3 1962 to December 31 1998, for a total of 9190 observations.

3.1 *The Modified Log-Squared Transformation*

An inlier problem often arises when dealing with log-squared returns; if the asset return is very close to zero, then the log-squared transformation yields a large negative number. Such an observation can then greatly affect the results of subsequent data analysis. In the extreme case, if the asset return is equal to zero, then the log-squared transformation is not even defined (though this does not arise in the present context because the returns are demeaned). Fuller (1996) proposed a slight modification of the log-squared transformation, which does not converge to minus infinity as the argument converges to zero. This specifies that the transformed series of asset returns is

$$y_t^* = \log(y_t^2 + \tau s^2) - \frac{\tau s^2}{y_t^2 + \tau s^2}, \quad (10)$$

where s^2 is the sample variance of y_t and τ is a small constant. In this paper, I use four volatility measures: squared, log-squared and absolute returns, and the modified log-squared returns, as defined in equation (10) (with τ set to 0.02, following Fuller).

3.2 *Results*

The sample autocorrelations of each of the four volatility measures applied to the equal-weighted returns are shown in Figure 1. They all exhibit the slow rate of decay that is the defining characteristic of a long memory model. Table 1 reports the log-periodogram estimate, \hat{d} , using all four volatility measures, for both equal-weighted

and value-weighted returns. Results using the GSP estimator, \tilde{d} , instead are given in Table 2. In both tables, results are reported for four choices of the bandwidth, $m=50, 100, 150$ and 200 . Two key things can be noticed in Tables 1 and 2. Firstly, there is strong and consistent evidence for long memory volatility dependencies ($d>0$). The results are not too sensitive to the bandwidth (which can often be a problem (Geweke (1998))), nor are they too sensitive to the choice of estimator. But secondly, much lower estimates of d are obtained with squared returns than with the other volatility measures. This is true across all bandwidths, and with both \hat{d} and \tilde{d} . In some cases with squared returns, the estimate of d is not significantly different from zero (at the 5% level). Meanwhile d is significantly positive at all conventional significance levels for all the other volatility measures.

4. Monte-Carlo Results

The main goal of this paper is to compare the properties of the semiparametric estimators \hat{d} and \tilde{d} , using the different volatility measures to determine which are good/bad choices and to account for some of the results reported in Table 1. Theory gives us little guidance as to which measure should be used, and so, in this section I report the results of a Monte-Carlo simulation using three data generating processes, calibrated to the CRSP equal-weighted returns:

(1) The FISV model given by equations (1) and (2), where h_t is an ARFIMA(1,d,0) time series, so that $(1 - aL)(1 - L)^d h_t = u_t$ and u_t is Gaussian. I set $a = 0.52$, $d = 0.475$ and $\sigma^2 = 0.118$. These were the estimated values from fitting the FISV model to the CRSP equal-weighted returns (as in section 3 above) by the frequency-domain maximum-likelihood procedure of Breidt, Crato and de Lima (1998). I set $\sigma_\varepsilon^2 = 1$, without loss of generality⁴.

(2) The FIGARCH model given by equations (6) and (7). I set $\omega = 0.022$, $\beta = 0.436$, $\phi = 0.309$ and $d = 0.473$. These were the estimated values from fitting the FIGARCH model to the CRSP equal-weighted returns by maximum-likelihood (as in Bollerslev and Mikkelsen (1996)).

(3) The FIEGARCH model given by equations (6) and (8). I set $\omega = -0.65$, $\theta = 0.463$, $\gamma = -0.14$, $\psi = -0.661$, $\phi = 0.833$ and $d = 0.382$. These were the estimated values from fitting the FIEGARCH model to the CRSP equal-weighted returns by maximum-likelihood.

In each model, it is further specified that ε_t is t-distributed on v degrees of freedom, $v = 2, 3, 5, 7, \infty$, so that ε_t has $v - 1$ finite moments. The case $v = \infty$ corresponds to normality. The sample size T is $2^{12} = 4096$ or $2^{13} = 8192$; the latter is a typical sample size for daily U.S. stock return data. In the FISV model, the Gaussian ARFIMA

⁴Changing σ_ε^2 amounts to multiplying y_t^2 or $|y_t|$ by a constant, or adding a constant to $\log(y_t^2)$; the log-periodogram regression is numerically invariant to these transformations.

series were generated by the algorithm of Davies and Harte (1987); the FIGARCH and FIEGARCH series were simulated exactly as described in Bollerslev and Mikkelsen (1996). For each simulation, 1000 replications were conducted. Results in which the Monte-Carlo experiment is instead calibrated to the CRSP value-weighted returns are similar and are available from the author on request, but are not shown, so as to conserve space.

The simulated biases of the GPH and GSP estimators using the four different volatility measures, in the alternative models, are shown in Figures 2-5. The simulated standard deviations are in Figures 6-9. In all cases, the bias and standard deviation is plotted against the bandwidth m , for values of m between 30 and 300. The key points to note are:

- (1) The estimators generally exhibit some downward bias. This is consistent with findings in a number of other papers, including Bollerslev and Wright (2000), where solutions to the bias using high frequency data were discussed. In the present paper, it is assumed that no such data is available. Accordingly, temporal aggregation is not an issue.
- (2) The choice of volatility measure has little to do with the magnitude of the bias, provided that the data is conditionally Gaussian ($v = \infty$).
- (3) However, using squared returns gives much more downward bias than any of the

other three volatility measures, if the data is conditionally leptokurtic. This is true in the FISV, FIGARCH and FIEGARCH models. Squared returns give the worst bias even for $v = 7$ (not very heavy tails). The bias using squared returns continues to worsen as v falls, and is enormous for $v = 2$ or $v = 3$. The other volatility measures yield comparable bias and show little tendency to become more biased as the tail thickness increases, except that using absolute returns gives more bias than log-squared or modified log-squared returns if $v = 2$ (a rather extreme specification). Even with $v = 2$, although the bias using absolute returns is large, it is substantially smaller than the bias using squared returns.

(4) The choice of estimator (\hat{d} or \tilde{d}) and the sample size both have little to do with the magnitude of the bias. In the FISV and FIEGARCH models, increasing the bandwidth makes the bias somewhat worse.

(5) The standard deviation of the estimators is not greatly affected by the choice of volatility measure, but is typically somewhat higher using squared returns, except in the case $v = 2$.

Overall, in these simulations, there is little difference between the performance of the semiparametric estimators of d using alternative volatility measures, except that less reliable results are obtained with the squared returns (point (4)). This observation is the main focus of this paper and is consistent with the empirical results of the

previous section. It is clear that asset returns have heavy tails, and perhaps even have infinite kurtosis (Loretan and Phillips (1994)). Volatility clustering with conditional Gaussianity can account for some of this leptokurtosis, but it is widely agreed that ε_t must have heavier tails to approximate the actual tail behavior of stock market and exchange rate returns (see e.g. Bollerslev (1987), Bollerslev, Engle and Nelson (1994), Fridman and Harris (1998), Gallant, Hsieh and Tauchen (1997), Geweke (1998) and Liesenfeld and Jung (1997)⁵). While ε_t is unlikely to be t-distributed on 2 degrees of freedom, it seems very reasonable to think of it as being t-distributed on 7 or 5 (or even fewer) degrees of freedom. Accordingly, the downward bias using squared returns found in these simulations can be expected to be an empirically relevant phenomenon. So, squared returns should not be used as the volatility measure in the log-periodogram regression estimation of long memory volatility dependencies.

In the FISV, FIGARCH and FIEGARCH models, if the m th moment of ε_t is infinite, then the m th unconditional moment of y_t must also be infinite. Davis and Mikosch (1998) show that the sample autocorrelation function of a general stationary infinite variance process has a random limit. They also show that the sample autocorrelation function of a general stationary process with finite variance but infinite

⁵For example, Fridman and Harris (1998) and Liesenfeld and Jung (1997) both consider the parametric estimation of a short memory autoregressive stochastic volatility model, by simulated maximum likelihood, allowing for conditional nonnormality. Both papers find evidence of significant conditional nonnormality in stock returns.

fourth moment converges to its population counterpart, but at a rate slower than the usual $T^{1/2}$ rate. The nonstandard asymptotics of sample autocorrelation functions of heavy-tailed processes are relevant to the results I find in these Monte-Carlo simulations. If the fourth moment of ε_t is finite, but its eighth moment is infinite, then y_t^2 must have an infinite fourth moment, but $|y_t|$ may have a finite fourth moment. Similarly, if ε_t has finite variance but an infinite fourth moment, then y_t^2 must have infinite variance, but $|y_t|$ may have finite variance. The squared returns are most likely to be more vulnerable to the difficulties that arise with heavy-tailed variables.

A possible intuition for the downward bias when using squared returns with conditionally leptokurtic data is as follows⁶. The denominator of the sample autocorrelation of $|y_t|^p$ will involve extremely large terms if y_t does not have $2p$ finite moments. Meanwhile, the numerator of the sample autocorrelation at lag j will involve large terms only if $|y_t|^p$ and $|y_{t-j}|^p$ are both large. Hence the sample autocorrelations of $|y_t|^p$ shrink towards zero and the associated semiparametric estimator of d is biased downwards. This problem is worse for squared returns than for absolute returns, because absolute returns will always have more finite moments than squared returns.

⁶I am grateful to an anonymous referee for suggesting this intuition.

5. Conclusion

Researchers use a variety of volatility measures to estimate the long memory parameter in the volatility of asset returns. For daily U.S. stock returns, using squared returns gives the lowest estimates. In Monte-Carlo simulations, I have found that using squared returns causes a large downward bias, if the returns are sufficiently leptokurtic, conditional on the volatility process. Since conditional leptokurtosis is a widely accepted property of asset returns, I conclude that researchers should avoid using squared returns as the volatility measure in semiparametric estimation of long memory volatility dependencies.

References

- Andersen, T.G. and T. Bollerslev (1997): Heterogeneous Information Arrivals and Return Volatility Dynamics: Uncovering the Long Run in High Frequency Returns, *Journal of Finance*, 52, pp.975-1005.
- Baillie, R.T. (1996): Long Memory Processes and Fractional Integration in Econometrics, *Journal of Econometrics*, 73, pp.5-59.
- Baillie, R.T., T. Bollerslev and H.O. Mikkelsen (1996): Fractionally Integrated Generalized Autoregressive Conditional Heteroskedasticity, *Journal of Econometrics*, 74, pp.3-30.
- Black, F. (1976): Studies of Stock Price Volatility Changes, *Proceedings from the American Statistical Association, Business and Economic Statistics Section*, pp.177-181.
- Bollerslev, T. (1987): A Conditional Heteroskedastic Time Series Model for Speculative Prices and Rates of Return, *Review of Economics and Statistics*, 69, pp.542-547.
- Bollerslev, T., R.F. Engle and D.B. Nelson (1994): ARCH Models in *Handbook of Econometrics*, Vol 4, (R.F. Engle and D.L. McFadden, eds.), Elsevier, Amsterdam.
- Bollerslev, T. and H.O. Mikkelsen (1996): Modeling and Pricing Long Memory in Stock Market Volatility, *Journal of Econometrics*, 73, pp.151-184.

- Bollerslev, T. and J.H. Wright (2000): Semiparametric Estimation of Long Memory Volatility Dependencies: The Role of High Frequency Data, *Journal of Econometrics*, 98, pp.81-106.
- Breidt, F.J., N. Crato and P. de Lima (1998): The Detection and Estimation of Long Memory in Stochastic Volatility, *Journal of Econometrics*, 83, pp.325-348.
- Dacorogna, M.M., U.A. Müller, R.J. Nagler, R.B. Olsen and O.V. Pictet (1993): A Geographical Model for the Daily and Weekly Seasonal Volatility in the Foreign Exchange Market, *Journal of International Money and Finance*, 12, pp.413-438.
- Davies, R.B. and D.S. Harte (1987): Tests for Hurst Effect, *Biometrika*, 74, pp. 95-102.
- Davis, R.A. and T. Mikosch (1998): The Sample Autocorrelations of Heavy-Tailed Processes with Applications to ARCH, *Annals of Statistics*, 26, pp.2049-2080.
- Deo, R.S. and C.M. Hurvich (1999): On the Log Periodogram Regression Estimator of the Memory Parameter in Long Memory Stochastic Volatility Models, unpublished manuscript, New York University.
- Ding, Z., C.W.J. Granger and R.F. Engle (1993): A Long Memory Property of Stock Returns and a New Model, *Journal of Empirical Finance*, 1, pp.83-106.
- Engle, R.F. (1982): Autoregressive Conditional Heteroskedasticity with Estimates of the Variance of UK Inflation, *Econometrica*, 50, pp.987-1007.
- Fama, E.F. (1965): The Behavior of Stock Market Prices, *Journal of Business*, 38, pp.34-105.
- Fridman, M. and L. Harris (1998): A Maximum Likelihood Approach for Non-Gaussian Stochastic Volatility Models, *Journal of Business and Economic Statistics*, 16, pp.284-291.
- Fuller, W. A. (1996): *Introduction to Statistical Time Series*, Wiley, New York.
- Gallant, A.R., D. Hsieh and G. Tauchen (1997): Estimation of Stochastic Volatility Models with Diagnostics, *Journal of Econometrics*, 81, pp.159-192.
- Geweke, J. (1998): Comment on Real and Spurious Long-Memory Properties of Stock-Market Data (by I. Lobato and N.E. Savin), *Journal of Business and Economic Statistics*, 16, pp.269-271.
- Geweke, J. and S. Porter-Hudak (1983): The Estimation and Application of Long Memory Time Series Models, *Journal of Time Series Analysis*, 4, pp.221-237.
- Granger, C.W.J. and Z. Ding (1996): Varieties of Long Memory Models, *Journal of Econometrics*, 73, pp.61-77.
- Granger, C.W.J. and R. Joyeux (1980): An Introduction to Long Memory Time Series Models and Fractional Differencing, *Journal of Time Series Analysis*, 1, pp.15-29.
- Hosking, J.R.M. (1981): Fractional Differencing, *Biometrika*, 68, pp.165-176.
- Hurvich, C.M., R.S. 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, pp.19-46.

- Liesenfeld, R. and R.C. Jung (1997): Stochastic Volatility Models: Conditional Normality versus Heavy-Tailed Distributions, unpublished manuscript, University of Tübingen.
- Lobato, I. and P.M. Robinson (1998): A Nonparametric Test for $I(0)$, *Review of Economic Studies*, 65, pp.475-495.
- Lobato, I and N.E. Savin (1998): Real and Spurious Long Memory Properties of Stock-Market Data, *Journal of Business and Economic Statistics*, 16, pp.261-268.
- Loretan, M. and P.C.B. Phillips (1994): Testing the Covariance Stationarity of Heavy-Tailed Time Series: An Overview of the Theory with Applications to Several Financial Datasets, *Journal of Empirical Finance*, 1, pp.211-248.
- Mandelbrot, B. (1971): When Can Price Be Arbitraged Efficiently? A Limit to the Validity of the Random Walk and Martingale Models, *Review of Economics and Statistics*, 53, pp.225-236.
- Nelson, D.B. (1991): Conditional Heteroskedasticity in Asset Returns: A New Approach, *Econometrica*, 59, pp.347-370.
- Robinson, P.M. (1991): Testing for Strong Serial Correlation and Dynamic Conditional Heteroskedasticity in Multiple Regression, *Journal of Econometrics*, 47, pp.67-84.
- Robinson, P.M. (1994): Time Series with Strong Dependence, in *Advances in Econometrics, Sixth World Congress* (C.A. Sims, ed.), Cambridge University Press, Cambridge.
- Robinson, P.M. (1995a): Log-Periodogram Regression for Time Series with Long Range Dependence, *Annals of Statistics*, 23, pp.1048-1072.
- Robinson, P.M. (1995b): Gaussian Semiparametric Estimation of Long Range Dependence, *Annals of Statistics*, 23, pp.1630-1661.
- Robinson, P.M. (1999): The Memory of Stochastic Volatility Models, unpublished manuscript, London School of Economics.
- Robinson, P.M. and M. Henry (1999): Long and Short Memory Conditional Heteroskedasticity in Estimating the Memory Parameter of Levels, *Econometric Theory*, 15, pp.299-336.
- Taylor, S.J. (1994): Modeling Stochastic Volatility, *Mathematical Finance*, 4, pp.183-204.
- Velasco, C. (1999): Non-Stationary Log-Periodogram Regression, *Journal of Econometrics*, 91, pp.325-371.

Table 1

**Log-Periodogram Estimates of Long Memory in the Volatility of U.S. Stock Returns
(Standard Errors in Parentheses)**

Volatility Measure	$m = 50$	$m = 100$	$m = 150$	$m = 200$
Squared VW returns	0.138 (0.054)	0.161 (0.034)	0.158 (0.026)	0.132 (0.021)
Log-Squared VW returns	0.578 (0.096)	0.425 (0.066)	0.407 (0.048)	0.444 (0.042)
MLSQ VW returns	0.582 (0.104)	0.456 (0.068)	0.437 (0.050)	0.467 (0.045)
Absolute VW returns	0.389 (0.082)	0.402 (0.065)	0.427 (0.050)	0.411 (0.042)
Squared EW returns	0.066 (0.069)	0.092 (0.050)	0.114 (0.038)	0.111 (0.031)
Log-Squared EW returns	0.409 (0.093)	0.353 (0.063)	0.394 (0.054)	0.388 (0.047)
MLSQ EW returns	0.424 (0.096)	0.333 (0.064)	0.382 (0.052)	0.379 (0.046)
Absolute EW returns	0.311 (0.093)	0.234 (0.058)	0.319 (0.050)	0.340 (0.047)

Notes: This table reports the log-periodogram estimates of the long memory parameter in four measures of the volatility of value weighted (VW) and equal weighted (EW) stock returns. These returns are the demeaned, total continuously compounded index returns, from CRSP. The standard errors are the conventional asymptotic standard errors (Robinson (1995a)). The data is daily, covering the period July 3 1962 to December 31 1998, for a total of 9190 observations. The returns labelled MLSQ have been transformed by the modified log-squared transformation, as in equation (10). The bandwidth parameter is m .

Table 2

Gaussian Semiparametric Estimates of Long Memory in the Volatility of U.S. Stock Returns
(Standard Errors in Parentheses)

Volatility Measure	$m = 50$	$m = 100$	$m = 150$	$m = 200$
Squared VW returns	0.129 (0.071)	0.167 (0.050)	0.171 (0.041)	0.149 (0.035)
Log-Squared VW returns	0.493 (0.071)	0.423 (0.050)	0.436 (0.041)	0.457 (0.035)
MLSQ VW returns	0.491 (0.071)	0.442 (0.050)	0.462 (0.041)	0.474 (0.035)
Absolute VW returns	0.369 (0.071)	0.400 (0.050)	0.438 (0.041)	0.426 (0.035)
Squared EW returns	0.077 (0.071)	0.089 (0.050)	0.123 (0.041)	0.127 (0.035)
Log-Squared EW returns	0.351 (0.071)	0.334 (0.050)	0.378 (0.041)	0.366 (0.035)
MLSQ EW returns	0.370 (0.071)	0.336 (0.050)	0.384 (0.041)	0.370 (0.035)
Absolute EW returns	0.290 (0.071)	0.250 (0.050)	0.313 (0.041)	0.304 (0.035)

Notes: This table reports the Gaussian semiparametric estimator (Robinson (1995b)) of the long memory parameter in four measures of the volatility of value weighted (VW) and equal weighted (EW) stock returns. These returns are the demeaned, total continuously compounded index returns, from CRSP. The data is daily, covering the period July 3 1962 to December 31 1998, for a total of 9190 observations. The returns labelled MLSQ have been transformed by the modified log-squared transformation, as in equation (10). The bandwidth parameter is m .