Estimation of Long Memory in Volatility

Rohit S. Deo and C. M. Hurvich

New York University

Abstract

We discuss some of the issues pertaining to modelling and estimating long memory in volatility. The main focus is on semi parametric estimation of the memory parameter in the long memory stochastic volatility model. We present the asymptotic properties of the log periodogram regression estimator of the memory parameter in this model. A modest simulation study of the estimator is also presented to study its behaviour when the volatility possesses only short memory. We conclude with a discussion of the appropriate choice of transformation of returns to measure persistence in volatility.

1. Introduction:

There has been great interest recently in modelling the temporal dependence in the volatility of financial time series such as stock and exchange rate returns. See, for example, Robinson (1991), Ding, Granger and Engle (1993), de Lima and Crato (1993), Breidt, Crato and de Lima (1996), Andersen and Bollerslev (1997 a and b), Andersen, Bollerslev, Diebold and Labys (1999), Baillie, Bollerslev and Mikkelsen (1996) and Henry and Payne (1998). Using measures of volatility such as powers or logarithms of squared returns, these authors have found that the sample autocorrelation function of volatility decays slowly to zero at high lags.

A series $\{X_t\}$ is said to have long memory if its correlation function $\rho(\cdot)$ is of the form

$$\rho(j) \sim C_1 j^{2d-1} \qquad j \to \infty, \tag{1}$$

where $C_1 \neq 0$ and $0 < d < 0.5$. The parameter d is called the long memory parameter and controls the rate of decay of the correlations. Note that the correlations in a long memory series decay at a hyperbolic rate and are not absolutely summable as opposed to the exponential rate which is obtained in short memory series such as ARMA models. An alternative definition of a long memory series is through its spectral density $f(\cdot)$. Under this definition, a series $\{X_t\}$ is said to possess long memory if its spectral density is of the form

$$f(\lambda) \sim C_2 \lambda^{-2d} \qquad \lambda \to 0, \tag{2}$$

where $C_2 > 0$ and $0 < d < 0.5$. Under certain conditions, the two definitions (1) and (2) are equivalent. In general, neither condition implies the other. See Robinson (1995a).

In this paper, we will discuss some of the issues pertaining to modelling and estimating long memory in volatility. In section 2, we present some models of long memory in volatility and discuss their properties. In section 3, we discuss estimation of the long memory parameter in these models, focusing mainly on semiparametric estimation in the particular class of long memory stochastic volatility models.

2. Models for Long Memory in volatility:

Models for volatility of returns must possess two attributes based upon economic theory and empirical observation: Returns are martingale differences and hence uncorrelated but powers of absolute values of returns are correlated. There are two popular classes of models which have been proposed to account for these phenomena: The observation driven models and the latent variable models.

The observation driven models for the return series $\{r_t\}$ are of the form

$$r_t = \sigma_t v_t, \tag{3}$$

where v_t is an independent identically distributed series with mean zero and finite variance and $\sigma_t \in \psi_{t-1}$, the sigma algebra of $\{r_{t-1}, r_{t-2}, ...\}$. The model (3) was first introduced by Engle (1982), who specified $\sigma_t^2 = \alpha_0 + \sum_{i=1}^{q} \alpha_i r_{t-i}^2$, where $\alpha_0, \alpha_1, ..., \alpha_q$ are unknown parameters such that $\alpha_0 > 0$, $0 \leq \alpha_i < 1$ for $i = 1, ..., q$ and $\sum_{i=1}^{q} \alpha_i < 1$. Engle (1982) called this model the ARCH(q) (AutoRegressive Conditional Heteroscedasticity) model. It is easy to verify that $E(r_t | \psi_{t-1}) = 0$ and that $\{r_t^2\}$ is correlated in the ARCH model. However, the ARCH (q) model is such that the correlation function of $\{r_t^2\}$ decays exponentially.

In an attempt to incorporate long memory in squared returns, Robinson (1991) gave two general specifications of σ_t in (3). The first specification was of the form

$$\sigma_t^2 = \alpha_0 + \sum_{i=1}^{\infty} \alpha_i r_{t-i}^2, \tag{4}$$

which can be considered to be an ARCH(∞) model in the spirit of Engle's ARCH(q) model. A particular version of this model where the $\{\alpha_j\}$ coefficients were taken to be those from an ARFIMA(p, d, q) model was discussed by Baillie et al (1996), who named it the FIGARCH process. Robinson (1991) also proposed another specification of the form

$$\sigma_t^2 = \left(\sigma + \sum_{j=1}^{\infty} \alpha_j r_{t-j}\right)^2. \tag{5}$$

Giraitis, Robinson and Surgailis (1999) have shown that there exist weakly stationary solutions to equations (3) and (5), which exhibit long memory in $\{r_t^2\}$. More specifically, they have proved that if the coefficients $\{\alpha_j\}$ in (5) satisfy

$$\alpha_j \sim C_3 j^{d-1} \qquad 0 < d < 0.5,$$

then under additional conditions on $\sum \alpha_j^2$, the process $\{r_t\}$ defined by (3) and (5) is weakly stationary and satisfies

$$Corr\left(r_t^2, r_{t-j}^2\right) \sim C_4 j^{2d-1},$$

where C_4 is a positive constant. Giraitis, Kokoskza and Leipus (1998) have obtained sufficient conditions for weakly stationary solutions to equations (3) and (4), which do not cover long memory in the correlation function $\rho_{r_t^2}(\cdot)$ of the squared returns. Thus, as Robinson and Henry (1999) point out, the character of solutions of (3) and (4) remains open to further study.

Latent variable models for the return series $\{r_t\}$ are of the form

$$r_t = \sigma_t v_t, \tag{6}$$

where v_t is an independent identically distributed series with mean zero and finite variance and σ_t^2 is taken to be a positive function of an unobservable latent process $\{h_t\}$ which is assumed to be independent of $\{v_t\}$. Latent variable models were initially considered by Clark (1973) in an attempt to incorporate the effect of flow of information on the returns. Clark (1973) assumed however that the $\{h_t\}$ process was an independent series and thus did not allow for dependence in the squared returns. Taylor (1986) incorporated this dependence by modelling $\{\sigma_t\}$ as

$$\sigma_t = \exp\left(h_t/2\right), \tag{7}$$

where $\{h_t\}$ is a stationary short memory Gaussian process independent of $\{v_t\}$. In simultaneous work, Breidt et al (1998) and Harvey (1993) extended the model by allowing $\{h_t\}$ to be a long memory Gaussian process with memory parameter $d \in (0, 0.5)$. In the special case where $\{v_t\}$ is normally distributed, this model can be thought of as a natural discrete time analogue of the continuous time model developed by Comte and Renault (1996). In this paper, we will refer to the model given by (6) and (7), where $\{h_t\}$ is a Gaussian long memory series independent of $\{v_t\}$, as the Long Memory

Stochastic Volatility (LMSV) model.

Using the moment generating function of a Gaussian distribution, it can be shown for the LMSV model that

$$\rho_{r_t^2}(j) \sim Cj^{2d-1} \qquad j \to \infty,$$

and hence the squared returns possess long memory. As a matter of fact, it can be shown that for any positive value s,

$$\rho_s(j) \sim C_s j^{2d-1} \qquad j \to \infty,$$

where $\rho_s(j)$ denotes the correlation of $\{|r_t|^s\}$ at lag j. Thus, the correlations of the absolute returns raised to any power s have hyperbolic decay and always decay at the same rate which is governed by d.

Another latent variable volatility model has been suggested recently by Robinson and Zaffaroni (1998), who model the scale parameter σ_t in (6) as

$$\sigma_t = \mu + \sum_{i=1}^{\infty} \alpha_i \varepsilon_{t-i}. \tag{8}$$

They consider two specifications for $\{\varepsilon_t\}$. In one specification, they set $\varepsilon_t = v_t$, and thus the model for the returns $\{r_t\}$ depends on only one shock $\{v_t\}$. In the other specification, $\{\varepsilon_t\}$ is a zero mean independent series independent of $\{v_t\}$. This second model thus incorporates two shocks. Under suitable conditions on the asymptotic behaviour of $\{\alpha_j\}$ in (8), Robinson and Zaffaroni (1998) show that $\{r_t^2\}$ has long memory.

In the next section, we discuss results on parametric and semiparametric estimation of long memory in volatility for the LMSV model.

3. Estimation of LM in volatility:

Fully parametric estimation of long memory may be done by either using the time domain maximum likelihood estimator (MLE) or the frequency domain approximate maximum likelihood estimator (Whittle 1962) referred to as the quasi maximum likelihood estimator (QMLE). Asymptotic results on such estimators of linear long memory series have been established by Fox and Taqqu (1986) and Dahlhaus (1989) among others. There are a few results on fully parametric estimation of long memory in volatility in the literature. Zaffaroni (1998) has shown the asymptotic normality of the QMLE based on $\{r_t^2\}$ in the stochastic volatility model (6) and (8), when $\varepsilon_t = v_t$.

Hosoya (1997) has obtained results on fully parametric maximum likelihood estimators for multivariate linear long memory series. His results are also valid when the observed series has a state space representation. As we show below, the LMSV model, upon appropriate transformation, possesses a state space representation. Hence, Hosoya's results may be applied to obtain the limiting distribution of the QMLE of d in the LMSV model. Let $y_t = \log r_t^2$,where $\{r_t\}$ follows the LMSV model. It then follows that

$$y_t = \mu + h_t + u_t, \tag{9}$$

where $\mu = E\left(\log v_t^2\right)$ and $u_t = \log v_t^2 - E\left(\log v_t^2\right)$ is a zero mean series independent of $\{h_t\}$. Since $\{u_t\}$ will also be independently distributed, the series $\{y_t\}$ will possess long memory with the same memory parameter d as possessed by $\{h_t\}$. From equation (9), we see that $\{y_t\}$ has the required state space representation.

Several semiparametric estimators of long memory have been proposed and studied in detail for linear long memory time series. Two leading semiparametric estimators are the Geweke Porter-Hudak (GPH) estimator (Geweke and Porter-Hudak (1983)) and the Gaussian Semiparametric (GSE) Estimator (Künsch (1987), Robinson (1995b)). Robinson (1995a) and Hurvich Deo and Brodsky (1998) have established the asymptotic normality of the GPH estimator when the observed series is Gaussian, while Robinson (1995b) has shown that the GSE is asymptotically normal when the observed series is a linear series in martingale differences with constant conditional variance. It is very tempting to appeal to this second result in justifying the use of the GSE in estimating d in volatility for either squares or absolute values of returns, which are clearly non-Gaussian. However, it should be noted that there is no known model which guarantees a martingale difference structure for returns, admits a linear representation in martingale differences for either powers or logarithms of squared returns and also allows for long memory in volatility. Currently, the only results available on semiparametric estimation of long memory in volatility are due to Deo and Hurvich (1999). We now describe these results.

Let $\{r_t\}$ follow the LMSV model described in Section 2. Deo and Hurvich (1999) assumed that $\{h_t\}$ has spectral density $f_h(\cdot)$ which is of the form

$$f_h(\lambda) = \left|2\sin\left(\frac{\lambda}{2}\right)\right|^{-2d} g^*(\lambda), \tag{10}$$

where $d \in (0, 0.5)$ and $g^*(\cdot)$ is a spectral density continuous on $[-\pi, \pi]$ bounded above and away from zero with first derivative $g^{*\prime}(0) = 0$ and second and third derivatives bounded in a neighbourhood of zero. An example of a process with spectral density satisfying (10) is a stationary invertible ARFIMA(p, d, q). Deo and Hurvich (1999) also assumed that $\{u_t\}$ as defined in (9) has a finite eighth moment. A sufficient condition for this is that $\{v_t\}$ have a probability density that be bounded at the origin and obey a power law decay in the tails as would occur, for example, in all t or stable distributions.

Let $\{y_t\}$ be defined as in (9) and define the periodogram of the observations $y_0, y_1, ..., y_{n-1}$ at the j^{th} Fourier frequency $\omega_j = 2\pi j/n$ as

$$I_j = \frac{1}{2\pi n}\left|\sum_{t=0}^{n-1} y_t \exp(-it\omega_j)\right|^2.$$

The GPH estimator of d using the first m Fourier frequencies is then

$$\hat{d} = -\frac{1}{2S_{xx}} \sum_{j=1}^{m} a_j \log I_j,$$

where $a_j = X_j - \bar{X}$, $X_j = \log|2\sin(\omega_j/2)|$, $\bar{X} = m^{-1}\sum_{j=1}^{m} X_j$ and $S_{xx} = \sum_{j=1}^{m} a_j^2$. Note that the GPH estimator is invariant to the mean of $\{y_t\}$ since it is based upon the periodogram at non-zero Fourier frequencies. Deo and Hurvich (1999) obtained the following two theorems about $\hat{d}$.

Theorem 1 *Let* $n \to \infty$, $m \to \infty$ *and* $n^{-2d}m^{2d}\log^2 m \to 0$. *Then*

$$E\left(\hat{d} - d\right) = -(2\pi)^{2d}\frac{\sigma_u^2}{2\pi g^*(0)}\frac{d}{(2d+1)^2}\left(\frac{m^{2d}}{n^{2d}}\right) + O\left(\frac{\log^3 m}{m}\right) + o\left(\frac{m^{2d}}{n^{2d}}\right)$$

and

$$Var\left(\hat{d}\right) = \frac{\pi^2}{24m} + o\left(m^{-1}\right) + O\left(\frac{m^{4d}}{n^{4d}}\log^2 m\right).$$

Theorem 1 implies that $\hat{d}$ is consistent for d if $m = Kn^{\delta}$ for any $0 < \delta < 1$. The first term in the bias, which is due to the noise $\{u_t\}$, is dominant if and only if $\delta > (2d+1)^{-1}2d$. Hence, $\hat{d}$ will tend to have an increasingly negative bias as m becomes sufficiently large. The quantity $\sigma_u^2/(2\pi g^*(0))$ is a measure of the relative importance of the noise term $\{u_t\}$ compared to the short memory component of $\{h_t\}$. As σ_u^2 increases, the bias in $\hat{d}$ increases. Under stronger conditions on m, asymptotic normality is obtained for $\hat{d}$.

Theorem 2 *Let* $n \to \infty$, $m \to \infty$ *and* $n^{-4d}m^{4d+1}\log^2 m \to 0$ *and* $\log^2 n = o(m)$. *Then*

$$m^{1/2}\left(\hat{d} - d\right) \xrightarrow{D} N\left(0, \frac{\pi^2}{24}\right).$$

Theorem 2 shows that the limiting distribution of $\hat{d}$ remains unchanged compared to the Gaussian case considered by Robinson (1995a) and Hurvich Deo and Brodsky (1998). However, the conditions on m here are much stronger. The limiting distribution will hold if and only if $\delta < (4d+1)^{-1}4d$, which can be arbitrarily small if d is sufficiently close to zero. Hence, it is crucial for the finite sample performance that the lowest frequencies not be dropped when computing the GPH estimator in the LMSV framework. Deo and Hurvich (1999) found in their simulation study that, due to leverage effects, even the deletion of just the first two frequencies caused a substantial inflation in the MSE of the estimator for samples of size $n = 6144$.

Robinson (1995a) obtained the limiting distribution of a modified version of $\hat{d}$ for a Gaussian long memory process $\{Z_t\}$ under the assumption that its spectral density was of the form

$$f_Z(\lambda) = C\lambda^{-2d}\left(1 + O\left(\lambda^{\beta}\right)\right) \qquad \lambda \to 0,$$

where $C > 0$ and $0 < \beta \leq 2$. He showed that the feasible range of values for m required to obtain asymptotic normality for $\hat{d}$ depended on β. More specifically, Robinson showed that the condition on m was of the form $m^{2\beta+1}/n^{2\beta} \to 0$. In the LMSV context, it can be easily shown that for the $\{y_t\}$ process,

$$f_y(\lambda) = \lambda^{-2d} g^*(0)\left[1 + O\left(\lambda^{2d}\right)\right].$$

This is similar to Robinson's formulation with $\beta = 2d$ and hence it is not surprising that the conditions on m in Theorem 2 depend on d.

4. Monte Carlo:

An issue not addressed in the work of Deo and Hurvich (1999) is the asymptotic distribution of $\hat{d}$ when $d = 0$ in the LMSV model. This distribution would be required for construction of tests of the null hypothesis that $d = 0$, i.e. the absence of long memory in volatility. When $d > 0$, the periodogram of the logarithm of the squared returns near the origin behaves like the periodogram of the Gaussian process $\{h_t\}$. This fact was exploited in Deo and Hurvich (1999) in obtaining the limiting distribution of $\hat{d}$ for $d > 0$. When $d = 0$, however, the contribution from the periodograms of both $\{h_t\}$ and $\{u_t\}$ are of the same order. Hence, it is hard to see how the periodogram of $\{\log r_t^2\}$ could be approximated by the periodogram of a linear series, much less a Gaussian one. Here we study the performance of $\hat{d}$ when $d = 0$ in the LMSV model via simulations.

We generated 500 replications of the time series $\{r_t\}$ of length $n = 6144$ given by the process (6) and (7), where $\{h_t\}$ was a Gaussian autoregressive process of order 1 with unit innovation variance and lag one autocorrelation of 0.9. The $\{v_t\}$ process was taken to be a Gaussian white noise process with unit variance. For every replication, $\hat{d}$ was computed using three values of m, viz. $m = [n^{0.3}]$, $m = [n^{0.4}]$ and $m = [n^{0.5}]$. In Table 1 we report the simulation means and standard deviations of the values of $\hat{d}$. For comparison, we also report the value $\pi/\sqrt{24 S_{xx}}$, which is an approximation to the standard deviation of $\hat{d}$ motivated by ordinary linear regression heuristics. Simulation results in Deo and Hurvich (1999) have shown that this quantity is a closer approximation to the true standard deviation than $\pi/\sqrt{(24m)}$ in the case $d > 0$. As can be seen from Table 1, $\hat{d}$ is slightly positively biased due to the large autoregressive coefficient but the standard deviation is quite close to the value $\pi/\sqrt{24 S_{xx}}$. In spite of the lack of rigorous theoretical results regarding the limiting distribution of $\hat{d}$ when $d = 0$, we conjecture that the limiting distribution of Theorem 2 remains valid and hence tests of $d = 0$ may be constructed as usual.

As noted in Section 2, any positive power of $|r_t|$ possesses long memory with the same value of d in the LMSV model. We thus carried out another simulation to study the extent to which this invariance holds in $\hat{d}$ based upon different such transformations. We generated 500 replications of the time series $\{r_t\}$ of length $n = 6144$ given by the process (6) and (7). The $\{h_t\}$ series was an ARFIMA$(0, d, 0)$ given by

$$(1 - B)^d h_t = \eta_t,$$

where $\{\eta_t\}$ is Gaussian white noise with variance $\sigma_\eta^2 = 0.8$. The $\{v_t\}$ process was Gaussian white noise with unit variance. For each realization, we computed $\hat{d}_{\log}$, $\hat{d}_{abs}$ and $\hat{d}_{sq}$, based on $\log r_t^2$, $|r_t|$ and r_t^2 respectively using the same three values of m as above. For each estimator of d, the bias becomes more negative as m increases, as predicted in Theorem 1 for the case $\log r_t^2$. Furthermore, for a given value of m, the squared returns show less persistence than either absolute returns or log squared returns. This is in keeping with the observation made by Ding, Granger and Engle (1993) in their empirical analysis of stock return volatility.

The lower persistence in squared returns can be attributed to the fact that sample correlations of squared returns at a given lag tend to be substantially smaller than sample correlations at the same lag for absolute or log squared returns. This effect can be seen in the left side of Figure 1, in which we have plotted the sample correlation function of r_t^2, $|r_t|$ and $\log r_t^2$ for a single realization of a $(0, d, 0)$ LMSV model with $d = 0.47$. We propose here a graphical explanation for this phenomenon. Note first that the sample correlation at lag j for a stationary series x_t is almost identical to the slope in the least squares regression of x_t on x_{t-j}. The right side of Figure 1 shows the scatter plot and corresponding least squares regression line at lag 20 for the three transformations of r_t. The slope of this line is largely determined by a relatively few outliers in the lagged variable r_{t-20}^2, which act as highly influential points due to their high leverage. Since the corresponding r_t^2 values for these points are small, the slope of the regression line is damped substantially. Though there are some extreme values of r_t^2 occurring at small values of r_{t-20}^2, these are not highly influential due to the presence of a large number of smaller values of r_t^2 in that region. This damping effect of the influential observations translates into small values of the sample correlations and thus lead to strongly negatively biased estimates of d based on r_t^2. By contrast, the outliers do not possess such a great damping effect in the case of $|r_t|$, and the outliers are completely nullified by the $\log r_t^2$ transformation.

References

Andersen, T. and Bollerslev, T. (1997a). Heterogenous information arrivals and return volatility dynamics: Uncovering the long-run in high frequency returns. *The Journal of Finance,* **LII,** 975-1005.

Andersen, T. and Bollerslev, T. (1997b). Intraday periodicity and volatility persistence in financial markets. *Journal of Empirical Finance,* **4,** 115-158.

Andersen, T., Bollerslev, T., Diebold, F. and Labys, P. (1999). The distribution of exchange rate volatility. Working Paper.

Baillie, R. Bollerslev, T. and Mikkelsen, H. (1996). Fractionally integrated generalized autoregressive conditional heteroscedasticity. *Journal of Econometrics.* **74,** 3-30.

Bollerslev, T. (1986). Generalized autoregressive conditional heteroscedasticity. *Journal of Econometrics*, **31,** 307-327.

Bollerslev, T. and Mikkelsen, H. (1996). Modelling and pricing long memory in stock market volatility. *Journal of Econometrics,* **73,** 151-184.

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

Clark, P.K. (1973). A subordinated stochastic process model with fixed variance for speculative prices. *Econometrica,* **41**, 135-156.

Comte F. and Renault, E. (1996). Long memory continuous time models. *Journal of Econometrics,* **73,** 101-149.

Dahlhaus, R. (1989). Efficient parameter estimation for self similar processes, *Annals of Statistics,* **17,** 1749-1766.

de Lima, P. and Crato, N. (1993). Long range dependence in the conditional variance of stock returns. Proceedings of the Business and Economics Statistics Section, Joint Statistical Meetings.

Deo, R.S. and Hurvich, C.M. (1999). On the log periodogram regression estimator of the memory parameter in long memory stochastic volatility models. NYU Working Paper. SOR 99-9.

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

Dunsmuir, W. (1979). A central limit theorem for parameter estimation in stationary vector time series and its application to models for a signal observed with noise. *Annals of Statistics,* **7,** 490-506.

Engle, R. F. (1982). Autoregressive conditional heteroscedasticity with estimates of the variance of United Kingdom inflation. *Econometrica,* **50**, 987-1007.

Fox, R. and Taqqu, M. S. (1986). Large sample properties of parameter estimates for strongly dependent stationary time series. *Annals of Statistics,* **14,** 517-532.

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

Giraitis L., Kokoskza P. and Leipus R. (1998). Stationary ARCH models: dependence structure and Central Limit Theorem. Working paper.

Giraitis, L., Robinson, P. and Surgailis, D. (1999). A model for long memory conditional heteroskedasticity. Working Paper.

Harvey, A. C. (1993). Long memory in stochastic volatility. Preprint.

Henry, M. and Payne, R. (1998). An investigation of long range dependence in intra-day foreign exchange rate volatility. Preprint.

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

Hosoya, Y. (1997). A limit theory for long range dependence and statistical inference on related models. *Annals of Statistics,* **25,** 105-137.

Künsch, H. (1987). Statistical aspects of self similar processes. *Proceedings of the First World Congress of the Bernoulli Society,* pp 67-74. VNU Science Press.

Lobato, I. N. and Savin, N. (1996) Real and spurious long memory properties of stock market data. Preprint.

Lobato, I. N. (1999) A semiparametric two step estimator in a multivariate long memory model. *Journal of Econometrics,* **90**, 129-153.

Robinson, P.M. (1991). Testing for strong serial correlation and dynamic conditional heteroskedasticity in multiple regression. *Journal of Econometrics,* **47,** 67-84.

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

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

Robinson, P.M. and Henry, M. (1999). Long and short memory conditional heterscedasticity in estimating the memory parameter of levels. *Econometric Theory*, **15,** 299-336.

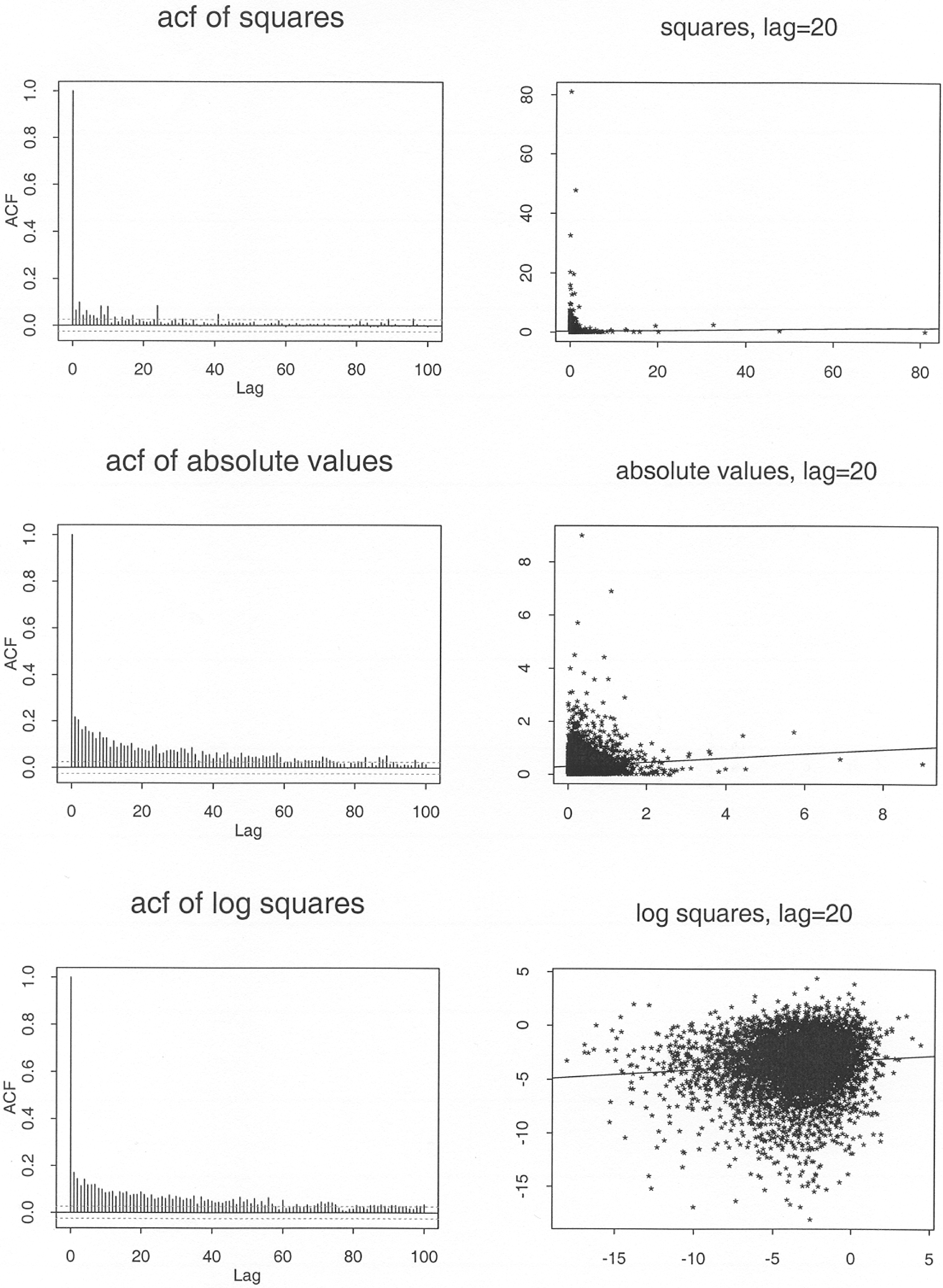
Robinson, P.M. and Zaffaroni, P. (1998). Nonlinear time series with long memory: a model for stochastic volatility. *Journal of Statistical Planning and Inference,* **68**, 359-371.

Taylor, S. (1986). Modelling Financial Time Series. Wiley, New York.

Whittle, P. (1962). Gaussian estimation in stationary time series. *Bulletin of the International Statistical Institute.* **39**, 105-129.

Zaffaroni, P. (1998). Gaussian inference of certain long-range dependent volatility models. Working paper.

Figure 1


acf of squares
squares, lag=20
acf of absolute values
absolute values, lag=20
acf of log squares
log squares, lag=20
ACF
Lag
0.0
0.2
0.4
0.6
0.8
1.0
0
20
40
60
80
100
2
4
6
8
-15
-10
-5
5