Statistica Sinica **15**(2005), 505-525

A POWERFUL TEST FOR CONDITIONAL HETEROSCEDASTICITY FOR FINANCIAL TIME SERIES WITH HIGHLY PERSISTENT VOLATILITIES

Julio Rodríguez and Esther Ruiz

Universidad Politécnica de Madrid and Universidad Carlos III de Madrid

Abstract: Traditional tests for conditional heteroscedasticity are based on testing for significant autocorrelations of squared or absolute observations. In the context of high frequency time series of financial returns, these autocorrelations are often positive and very persistent, although their magnitude is usually very small. Moreover, the sample autocorrelations are severely biased towards zero, especially if the volatility is highly persistent. Consequently, the power of the traditional tests is often very low. In this paper, we propose a new test that takes into account not only the magnitude of the sample autocorrelations but also possible patterns among them. This additional information makes the test more powerful in situations of empirical interest. The asymptotic distribution of the new statistic is derived and its finite sample properties are analyzed by means of Monte Carlo experiments. The performance of the new test is compared with various alternative tests. Finally, we illustrate the results analyzing several time series of financial returns.

Key words and phrases: Autocorrelations of non-linear transformations, GARCH, long-memory, McLeod-Li statistic, stochastic volatility.

1. Introduction

It is well known that high frequency time series of returns are characterized by evolving conditional variances. As a consequence, some non-linear transformations of returns, such as squares or absolute values, are autocorrelated. The corresponding autocorrelations are often small, positive and decay very slowly towards zero. This last characteristic has usually been related to long-memory in volatility; see, for example, Ding, Granger and Engle (1993). Two of the most popular models to represent the dynamic evolution of volatilities are the Generalized Autoregressive Conditional Heteroscedasticity (GARCH) model of Engle (1982) and Bollerslev (1986) and the Autoregressive Stochastic Volatility (ARSV) model proposed by Taylor (1984). Both models generate series with autocorrelated squares. However, Carnero, Peña and Ruiz (2004a) show that, unless the kurtosis is heavily restricted, GARCH models are not able to represent autocorrelations as small as those often observed in practice. Furthermore, little is known about the autocorrelations of squared or absolute returns generated by the most popular long-memory GARCH model, the Fractionally Integrated GARCH (FIGARCH) model proposed by Baillie, Bollerslev and Mikkelsen (1996); see Karanasos, Psaradakis and Sola (2004) and Palma and Zevallos (2004) for the autocorrelations of squares of long memory GARCH models closely related to FIGARCH. Moreover, Davidson (2004) shows that the dynamic properties of the FIGARCH model are somehow unexpected in the sense that the persistence of the volatility is larger the closer the long memory parameter is to zero. On the other hand, ARSV models are more flexible in representing the empirical characteristics often observed in high frequency returns and the statistical properties of the model are well defined. In particular, Harvey (1998) derives the autocorrelations of powers of absolute observations of Long Memory Stochastic Volatility (LMSV) models. Consequently, in this paper, we focus on testing for conditional homoscedasticity in the context of Stochastic Volatility (SV) models.

For the reasons previously given, the identification of conditional heteroscedasticity is often based on testing whether squared or absolute returns are autocorrelated; see, for example, Andersen and Bollerslev (1997) and Bollerslev and Mikkelsen (1999) among many others. Testing for lack of correlation of a particular transformation of returns, $f(y_t)$, can be carried out using the portmanteau statistic suggested by Ljung and Box (1978),

$$Q(M) = T \sum_{k=1}^{M} \tilde{r}^2(k), \tag{1.1}$$

where $\widetilde{r}(k) = \sqrt{(T+2)/(T-k)}\,r(k)$ is the standardized sample autocorrelation of order k, $r(k) = \sum_{t=k+1}^{T}(f(y_t)-\overline{f})(f(y_{t-k})-\overline{f})/\sum_{t=1}^{T}(f(y_t)-\overline{f})^2$, $\overline{f} = \sum_{t=1}^{T} f(y_t)/T$, and T is the sample size. In this paper, we consider two popular transformations, namely $f(y_t) = y_t^2$ and $f(y_t) = |y_t|$.

The $Q(M)$ statistic applied to squared observations was proposed by McLeod and Li (1983) who show that, if the eighth order moment of y_t exists, its asymptotic distribution can be approximated by a $\chi^2_{(M)}$ distribution. From now on, the statistic in (1) is named the McLeod-Li statistic even if it is applied to absolute returns. Notice that the $\chi^2_{(M)}$ asymptotic distribution of the Ljung-Box statistic requires that the series to be tested for uncorrelation is an independent sequence with finite fourth order moment; see Hannan (1970). Therefore, when the $Q(M)$ statistic is implemented with absolute values, only the fourth order moment of returns should be finite for the asymptotic distribution to hold.

Alternatively, Peña and Rodriguez (2002) proposed a portmanteau test based on the Mth root of the determinant of the autocorrelation matrix of order M,

$$D_M = T\left[1 - |\mathbf{R}_M|^{1/M}\right], \tag{1.2}$$

where

$$\mathbf{R}_M = \begin{bmatrix} 1 & \widetilde{r}(1) & \cdots & \widetilde{r}(M) \\ \widetilde{r}(1) & 1 & \cdots & \widetilde{r}(M-1) \\ \vdots & \vdots & \vdots & \vdots \\ \widetilde{r}(M) & \widetilde{r}(M-1) & \cdots & 1 \end{bmatrix}.$$

Peña and Rodriguez (2005a) have proposed a modified version of the D_M statistic based on the logarithm of the determinant. This new statistic has better size properties for large M and better power properties. If the eighth order moment of y_t exists, the asymptotic distribution of the D_M statistic applied to squared observations can be approximated by a Gamma distribution, $\mathcal{G}(\theta, \tau)$, with $\theta = 3M(M+1)/4(2M+1)$ and $\tau = 3M/2(2M+1)$. The same result holds for absolute values if the fourth order moment of returns is finite.

Although Peña and Rodriguez (2002) show that in general, for squared returns, the D_M test is more powerful than the McLeod-Li test, both tests have rather low power, especially when the volatility is very persistent; see, Pérez and Ruiz (2003) for exhaustive Monte Carlo experiments in the context of LMSV models. The low power could be attributed to substantial finite sample negative biases of the sample autocorrelations and to the small magnitude of the population autocorrelations. Therefore, these tests may fail to reject homoscedasticity when the returns are conditionally heteroscedastic.

However, notice that the sample autocorrelations of independent series with finite fourth order moment are asymptotically, not only identically distributed normal variables with zero mean and variance $1/T$, but also mutually independent; see Hannan (1970). Therefore, the estimated autocorrelations are not expected to have any distinct pattern in large samples. However, the McLeod-Li and Peña-Rodriguez tests only focus on the first implication of the null hypothesis, namely that the sample autocorrelations should have zero mean. These tests ignore the information on the patterns of successive estimated autocorrelations and, consequently, cannot distinguish between the correlogram of an uncorrelated variable whose autocorrelation coefficients are small and randomly distributed around zero, and the correlogram of a variable that has relatively small autocorrelations with a distinct pattern for very long lags. In this paper, we propose using a new statistic that considers the information about possible patterns in successive correlations to test for uncorrelatedness in non-linear transformations of returns . This test is based on ideas developed by Koch and Yang (1986) in the context of testing for zero cross-correlations between series of multivariate dynamic systems.

Finally, given that we are analyzing the performance of tests for conditional homoscedasticity in the context of SV models, we also consider the test proposed by Harvey and Streibel (1998) who focus on the ARSV(1) model given by

$$\begin{aligned} y_t &= \sigma_* \varepsilon_t \sigma_t,\ t = 1, \ldots, T, \\ \log(\sigma_t^2) &= \phi \log(\sigma_{t-1}^2) + \eta_t, \end{aligned} \tag{1.3}$$

where σ_* is a scale parameter, σ_t is the volatility, and ε_t and η_t are mutually independent Gaussian white noise processes with zero mean and variances one and σ_η^2, respectively. The model is stationary if $|\phi| < 1$. The same condition guarantees the existence of the fourth order moment; see Ghysels, Harvey and Renault (1996) for a detailed description of the statistical properties of SV models. The variance of the log-volatility process is given by $\sigma_h^2 = \sigma_\eta^2/(1-\phi^2)$, and it is assumed to be finite and fixed. Therefore, the variance of η_t can be written as a function of the persistence parameter as $\sigma_\eta^2 = (1-\phi^2)\sigma_h^2$. Observe that if, as is often observed in real time series of high frequency returns, the persistence parameter ϕ is close to one, then σ_η^2 should be close to zero for a given value of the variance of $\log(\sigma_t^2)$. In this case, the volatility evolves very smoothly through time. In the limit, if $\phi = 1$ then $\sigma_\eta^2 = 0$ and y_t is conditionally homoscedastic. Harvey and Streibel (1998) propose testing $\sigma_\eta^2 = 0$ using

$$NM = -T^{-1} \sum_{k=1}^{T-1} r(k)k. \tag{1.4}$$

They show that if the second order moment of $f(y_t)$ is finite, the NM statistic has asymptotically the Crámer-von Mises distribution, for which the 5% critical value is 0.461. Furthermore, the corresponding test is the Locally Best Invariant (LBI) test for the presence of a random walk; see Ferguson (1967) for the definition of the Locally Best test. They implement the test on squared and absolute observations, and show that the finite sample power is higher when the latter transformation is used.

The paper is organized as follows. Section 2 describes the new statistic and derives its asymptotic distribution under the null hypothesis. In Section 3, we carry out Monte Carlo experiments to assess the finite sample size and power of the new statistic in short memory ARSV and LMSV models. These finite sample properties are compared with the properties of the McLeod-Li, Peña-Rodriguez and Harvey-Streibel tests. In Section 4, the test is used to test for uncorrelatedness in squared and absolute daily returns of several financial prices.

2. A New Test for Conditional Homoscedasticity

As mentioned earlier, conditionally heteroscedastic processes generate time series with autocorrelated squares and absolute observations. Consequently, we propose testing for conditional homoscedasticity using the information contained in the sample autocorrelations of non-linear transformations, f of the underlying process y_t. The new test for zero correlations of $f(y_t)$ takes into account that, under the null hypothesis, if the fourth order moment of $f(y_t)$ exists, the sample autocorrelations of $f(y_t)$ are asymptotically independent and identically distributed normal variables with zero mean and variance $1/T$. Therefore, this statistic not only tests whether the sample autocorrelations are significantly different from zero but also incorporates information about possible patterns among successive autocorrelation coefficients, $r(k)$. We propose the statistic

$$Q_i^*(M) = T \sum_{k=1}^{M-i} \Big[\sum_{l=0}^{i} \tilde{r}(k+l) \Big]^2, \; i = 0, \ldots, M-1, \tag{2.1}$$

where $\tilde{r}(k+l)$ is the standardized sample autocorrelation of order $k+l$. Notice that for each value of the number of autocorrelations considered, M, we have a collection of statistics, choosing different values of i. Each of these statistics has different information on the possible pattern of the sample autocorrelations. For example, when $i = 0$, the McLeod-Li statistic in (1) is obtained as a particular case. In this case, the statistic is obtained adding up the squared estimated autocorrelations. If all of these autocorrelations are small, the statistic will be small and the null hypothesis is not rejected. However, when $i = 1$, the statistic incorporates information about the correlation between sample autocorrelations one lag apart. In this case, if they are strongly correlated, the null hypothesis can be rejected even if the coefficients $r(j)$ are very small. When $i = 2$, the correlations between coefficients two lags apart are also considered, and so on.

The statistic $Q_i^*(M)$ is a quadratic form in $T^{1/2}\tilde{\boldsymbol{r}}_{(M)}$, $\tilde{\boldsymbol{r}}_{(M)} = (\tilde{r}(1), \ldots, \tilde{r}(M))$, given by

$$Q_i^*(M) = T \tilde{\boldsymbol{r}}'_{(M)} \mathbf{A}_{\boldsymbol{i}} \tilde{\boldsymbol{r}}_{(M)},$$

where $\mathbf{A}_i = \mathbf{C}_i' \mathbf{C}_i$ is a symmetric matrix of dimension M. In general $\mathbf{C}_i'$ is a matrix of dimension $M \times (M-i)$, where the jth column has $j-1$ zeroes followed by $i+1$ ones, followed by zeroes. Given that, under the null hypothesis, the asymptotic distribution of $T^{1/2}\tilde{\boldsymbol{r}}_{(M)}$ is $N(\mathbf{0}, \mathbf{I}_M)$, $Q_i^*(M)$ behaves asymptotically as

$$Q(W) \sim \sum_{i=1}^{M} \lambda_i \chi^2_{(1),i},$$

where $\chi^2_{(1),i}$ are independent chi-squared variables with one degree of freedom, and the λ_j are the eigenvalues of $\mathbf{A}_j$; see Box (1954). Therefore, the asymptotic distribution of $Q_i^*(M)$ depends on the eigenvalues of $\mathbf{A}_i$, and consequently on M and i. Although the percentiles of the asymptotic distribution can be directly obtained, it is easier to compute them using one of the following two approximations. The first is from Satterthwaite (1941, 1946) and Box (1954).

In particular, the distribution of $Q_i^*(M)$ can be approximated by a gamma distribution, $\mathcal{G}(\theta,\tau)$, with parameters $\theta = a^2/2b$ and $\tau = a/2b$, where $a = (i+1)(M-i)$ and $b = (M-i)(i+1)^2 + 2\sum_{j=1}^{i+1}(M-i-j)(i+1-j)^2$. Notice that a is the trace of the matrix $\mathbf{A}_i$, $tr(\mathbf{A}_i) = \sum_{j=1}^{M}\lambda_j$, and b the trace of $\mathbf{A}_i\mathbf{A}_i$, $tr(\mathbf{A}_i\mathbf{A}_i) = \sum_{j=1}^{M}\lambda_j^2$. For example, if $i = 0$, then $a = M$, $b = M$, and therefore the usual $\chi^2_{(M)}$ asymptotic distribution is obtained. On the other hand, if for instance $i = 1$, then the parameters of the Gamma distribution are given by $\theta = (M-1)^2/(3M-4)$ and $\tau = (M-1)/(3M-4)$. For reasons that will be made clearer later, another interesting case is $i = [M/3]-1$. In this case, the corresponding parameters are $\theta = (54M+72M^2+24M^3)/2(45M+12M^2+7M^3+54))$ and $\tau = (108M+162)/(45M+12M^2+7M^3+54)$. Finally, consider the case $i = M-1$, with $\theta = -6M/2(M^3-4M^2-M-2)$ and $\tau = -6/(M^3-4M^2-M-2)$. Here the asymptotic distribution of the statistic $Q_{M-1}^*(M)/M$ can be approximated by $\chi^2_{(1)}$.

The asymptotic distribution of $Q_i^*(M)$ can alternatively be approximated using the generalization of the Wilson-Hilferty cube root transformation for χ^2 random variables. This generalization has been proposed by Chen and Deo (2004) to improve the normality of test statistics in finite samples. In particular, they provide an expression for the appropriate value of a power transformation, β, that makes the approximate skewness zero. The asymptotic distribution of the transformed statistic $Q_i^*(M)^\beta$ is then given by

$$\frac{Q_i^*(M)^\beta - \{a^\beta + \beta(\beta-1)a^{\beta-2}b\}}{\beta a^{\beta-1}\sqrt{2b}} \sim N(0,1),$$

where $\beta = 1 - 2ac/3b^2$, $c = tr(\mathbf{A}_i\mathbf{A}_i\mathbf{A}_i)$, and the parameters a and b are defined as in the previous approximation. Finally, notice that once M and i have been chosen, all the parameters needed to obtain the normal transformed statistic are given.

3. Size and Power in Finite Samples

In this section, we analyze the finite sample performance of the $Q_i^*(M)$ statistic by means of Monte Carlo experiments. The main objectives are to investigate whether the asymptotic distribution is an adequate approximation to the finite sample distribution under the null hypothesis, and to compare the size and power of the new statistic with the McLeod-Li, Peña-Rodriguez and Harvey-Streibel statistics. Furthermore, we give some guidelines as to which values of M and i work well from an empirical point of view.

3.1 Size

To analyze the size of the test in finite samples, we have generated series of white noise processes with two different distributions, Normal and Student-t with

$\nu = 5$ degrees of freedom. The Student-t distributions have been chosen because it has been often observed in empirical applications that the marginal distribution of financial returns is leptokurtic, and we want to analyze the performance of the tests in the presence of leptokurtic, homoscedastic time series. Moreover, the degrees of freedom are selected in such a way that the eighth moment of returns exists when $\nu = 9$, and does not exit when $\nu = 5$. All results are based on 20,000 replicates.

Table 1. Empirical sizes of the Q(M), D_M, NM and $Q_i^*(M)$ tests, $i = 1, [M/3]-1$ and $M-1$ for squared and absolute observations of homoscedastic Gaussian noise series of size $T = 50$, 100, 500 and $2,000$.

		y_t^2				$\|y_t\|$			
$M\backslash T$		50	100	500	2,000	50	100	500	2,000
12	D_M	0.029	0.037	0.048	0.050	0.042	0.043	0.051	0.050
	$Q(M)$	0.058	0.057	0.048	0.050	0.068	0.055	0.053	0.048
	$Q_1^*(M)$	0.055	0.055	0.048	0.048	0.065	0.057	0.052	0.053
	$Q_3^*(M)$	0.049	0.050	0.045	0.049	0.054	0.052	0.049	0.051
	$Q_{11}^*(M)$	0.013	0.025	0.040	0.046	0.014	0.026	0.044	0.047
24	D_M	0.021	0.027	0.045	0.048	0.029	0.032	0.048	0.049
	$Q(M)$	0.074	0.069	0.054	0.051	0.089	0.070	0.058	0.050
	$Q_1^*(M)$	0.077	0.073	0.056	0.052	0.090	0.075	0.059	0.054
	$Q_7^*(M)$	0.053	0.053	0.046	0.049	0.055	0.053	0.051	0.049
	$Q_{23}^*(M)$	0.002	0.013	0.034	0.046	0.002	0.015	0.039	0.045
36	D_M	0.009	0.019	0.041	0.048	0.013	0.022	0.050	0.049
	$Q(M)$	0.071	0.081	0.059	0.051	0.087	0.083	0.062	0.052
	$Q_1^*(M)$	0.083	0.087	0.059	0.056	0.101	0.092	0.064	0.054
	$Q_{11}^*(M)$	0.054	0.052	0.049	0.050	0.057	0.056	0.052	0.051
	$Q_{35}^*(M)$	0.001	0.007	0.028	0.043	0.001	0.006	0.033	0.045
	NM	0.047	0.049	0.050	0.047	0.050	0.050	0.053	0.051

Table 1 reports the empirical sizes of the D_M, $Q(M)$, NM and $Q_i^*(M)$ tests when the series are generated by homoscedastic Gaussian white noise processes. We consider $M = 12, 24$ and 36 and $i = 1$, $[M/3]-1$ and $M-1$. The nominal size is 5% and the sample sizes are $T = 50$, 100, 500 and $2,000$. The critical values have been obtained using the Gamma and Normal approximations described in Section 2, with similar results. Therefore, Table 2 only reports the empirical sizes obtained using the Gamma distribution. All the statistics are implemented for both squared and absolute observations. The results reported in Table 1 show that, with the exception of $Q_{M-1}^*(M)$, the empirical size of all tests considered is very close to the nominal one for moderate sample size. However, when the sample size is small, $T = 50$, 100, only the $Q_{[M/3]-1}^*(M)$ and NM tests have sizes reasonably close to the nominal.

Table 2. Empirical sizes of Q(M), D_M, NM and $Q_i^*(M)$ tests, $i = 1, [M/3]-1$ and $M-1$ for squared and absolute observations of homoscedastic Student-5 noise series of size $T = 50$, 100, 500 and $2,000$.

		y_t^2				$\|y_t\|$			
$M\backslash T$		50	100	500	2,000	50	100	500	2,000
12	D_M	0.024	0.036	0.061	0.069	0.032	0.036	0.047	0.051
	$Q(M)$	0.034	0.040	0.065	0.074	0.044	0.047	0.051	0.054
	$Q_1^*(M)$	0.030	0.035	0.052	0.061	0.040	0.045	0.049	0.053
	$Q_3^*(M)$	0.025	0.029	0.039	0.048	0.032	0.038	0.047	0.051
	$Q_{11}^*(M)$	0.008	0.016	0.028	0.037	0.009	0.021	0.041	0.048
24	D_M	0.014	0.026	0.064	0.078	0.024	0.027	0.044	0.049
	$Q(M)$	0.033	0.040	0.064	0.086	0.050	0.055	0.054	0.055
	$Q_1^*(M)$	0.039	0.041	0.055	0.073	0.056	0.054	0.054	0.054
	$Q_7^*(M)$	0.027	0.031	0.032	0.045	0.037	0.040	0.048	0.049
	$Q_{23}^*(M)$	0.001	0.009	0.023	0.036	0.001	0.010	0.036	0.046
36	D_M	0.006	0.016	0.062	0.082	0.010	0.019	0.039	0.048
	$Q(M)$	0.027	0.043	0.063	0.089	0.041	0.059	0.056	0.053
	$Q_1^*(M)$	0.039	0.045	0.053	0.079	0.055	0.064	0.059	0.056
	$Q_{11}^*(M)$	0.032	0.035	0.030	0.044	0.040	0.042	0.050	0.048
	$Q_{35}^*(M)$	0.000	0.004	0.020	0.034	0.001	0.005	0.031	0.043
	NM	0.039	0.041	0.044	0.046	0.048	0.049	0.048	0.049

Figure 1 plots the differences between the empirical and nominal sizes of the $Q_i^*(M)$ test as a function of i for $M = 12$, 24 and 36, and $T = 500$, $1,000$, $2,000$, and $4,000$. This figure shows that whatever the sample size is, the empirical size of the test is closer to the nominal size of 5% when i is chosen to be approximately equal to $[M/3] - 1$. Similar results have been obtained for other nominal sizes.

Table 2 reports the same quantities as in Table 1 when the series are generated by homoscedastic Student-5 white noises. In this case, it is possible to observe that all the tests considered can suffer from important size distortions when they are applied to squared observations. The empirical size is larger than the nominal for the D_M, $Q(M)$ and $Q_i^*(M)$ tests with small values of i. Consequently, the tests would reject the hypothesis of homoscedasticity more often than expected. Furthermore, the size distortions are not reduced when the sample size increases. For example, the sizes of D_{24}, $Q(24)$ and $Q_1^*(24)$ are 6.4%, 6.4% and 5.5%, respectively, when $T = 500$, and 7.8%, 8.6% and 7.3% when $T = 2,000$. Therefore it is evident that, as expected when the eighth moment is not defined, the asymptotic distribution is a bad approximation to the finite sample distribution of the statistics considered. The size distortions of $Q_i^*(M)$ are smaller than those for the $Q(M)$ and D_M tests, but they are still big enough to matter. Figure 2, which plots the differences between empirical and nominal sizes of $Q_i^*(M)$, $i = 0, \ldots, M-1$ and $M = 12$, 24 and 36, illustrates these size distortions. On the other hand, notice that the asymptotic distribution of the NM test only requires the second moment of y_t to be finite and consequently, as reflected in Table 2, its size is close to the nominal. Finally, Table 2 shows that when the tests are applied to absolute observations, the nominal and empirical sizes of all the tests considered are very close; see also Figure 2, which shows that the size of $Q_i^*(M)$ is rather close to the nominal for all M and i, even for moderate sample sizes.

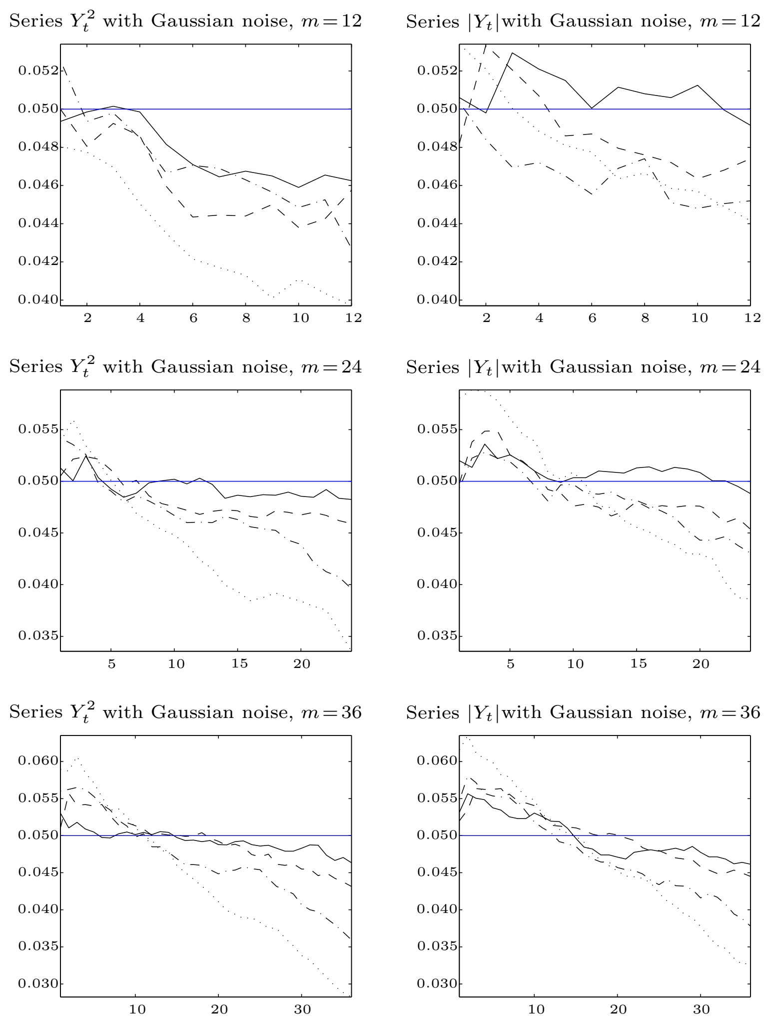

Figure 1. Differences between empirical and nominal rejection probabilities, $\alpha = 0.05$, of the $Q_i^*(M)$ test, $i = 0, \ldots, M-1$, for non-linear transformations of Gaussian noises with T=500 $(\cdots)$, $T = 1,000$ $(-\cdot)$, $T = 2,000$ $(--)$ and $T = 4,000$ (—).

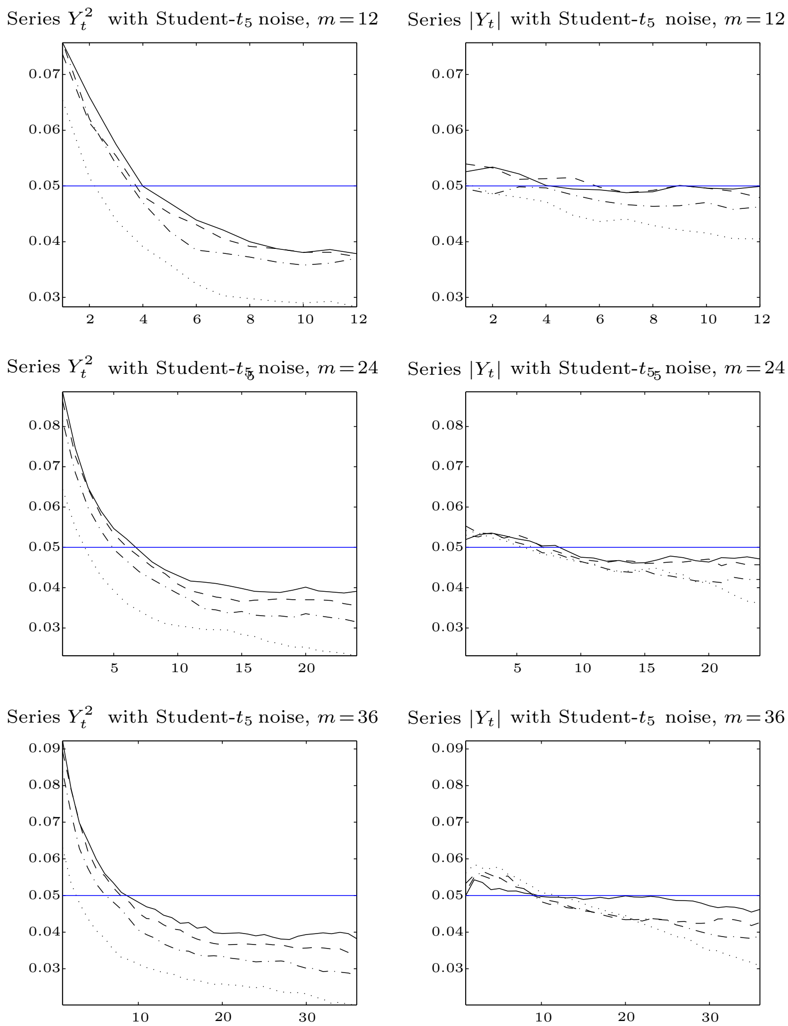

Figure 2. Differences between empirical and nominal rejection probabilities, $\alpha = 0.05$, of the $Q_i^*(M)$ test, $i = 0, \ldots, M-1$, for non-linear transformations of Student-t$_5$ noises with $T = 500$ $(\cdots)$, $T = 1,000$ $(-\cdot)$, $T = 2,000$ $(--)$ and $T = 4,000$ (—).

Focusing now on the results for absolute observations, Figures 1 and 2, the empirical sizes of $Q_i^*(M)$ decrease with i. For $M = 12, 24$ and 36, the smallest differences are obtained when i is approximately 3, 5 and 11 respectively. Therefore, it seems that the size is close to the nominal if $i = [M/3] - 1$. In any case, it is important to point out that for all values of M and i considered, the size of $Q_i^*(M)$ is remarkably close to the nominal, especially for large sample sizes.

Summarizing, the Monte Carlo experiments reported in this section show that, if $i = [M/3] - 1$ and the fourth order moment of y_t exists, the asymptotic distribution provides an adequate approximation to the sample distribution of $Q_i^*(M)$ when it is applied to absolute returns. If this is implemented with squared observations, the eighth moment should be finite. Consequently, we recommend testing for conditional heteroscedasticity using absolute returns; see also Harvey and Streibel (1998) and Pérez and Ruiz (2003). From now on, in this paper, all results are based on implementing the various statistics considered with absolute observations.

3.2. Power for short memory models

To analyze the power of $Q_i^*(M)$ in finite samples, we have generated artificial series by the ARSV(1) in (1.3) for squared $C.V. = \exp(\sigma_h^2) - 1 = 0.22$, 0.82 and 1.72. These values have been chosen to resemble the parameter values often estimated when the ARSV model is fitted to time series of financial returns; see Jacquier, Polson and Rossi (1994). The sample sizes considered are $T = 100$, 512 and 1,024. The scale parameter σ_* has been fixed to one without loss of generality.

Figure 3 plots the percentage of rejections of the $Q(M)$, D_M, NM and $Q_i^*(M)$ tests as a function of the persistence parameter, ϕ, for $M = 24$ and $T = 100$ and 512. Remember that when $\phi = 1$, $\sigma_\eta^2 = 0$ and, consequently, given that the series are homoscedastic, the percentage of rejections is 5%. First of all, this figure shows that the power of the NM test is highest when ϕ is close to the boundary and, consequently, the series is close to the null hypothesis of homoscedasticity, and when the sample size is very small. When the persistence of volatility decreases or the sample size is moderately large, the NM has important losses of power relative to its competitors. Notice that for the sample sizes usually encountered in the empirical analysis of financial time series, the powers of the $Q(M)$, D_M and $Q_i^*(M)$ tests are larger than the power of NM. On the other hand, comparing the powers of $Q(M)$, D_M and $Q_i^*(M)$, we can observe that if the sample size is $T = 100$, the power of $Q_i^*(M)$ is the largest for all the values of the parameters considered in Figure 3, i.e., $\phi \geq 0.8$. Finally, if $T = 512$, the power of $Q_i^*(M)$ is larger when $C.V. = 0.22$, and very similar for the other two $C.V.$ considered. The powers of the three tests are similar and close to one for large sample sizes.

To illustrate how the proposed test has higher power than its competitors, even in the more persistent cases, when the McLeod-Li and Peña-Rodriguez tests are well known to have difficulties in identifying the presence of conditional heteroscedasticity, Table 3 reports the powers of these tests using absolute observations for $M = 12$, 24 and 36, $T = 50$, 100, 512 and 1,024, $\phi = 0.98$, and $\sigma_\eta^2 = 0.1$. The power of the $Q_{[M/3]-1}^*(M)$ test is the highest for all M and T. For example, if $T = 512$ and $M = 12$, the power is 67.6% if the D_M test is implemented for absolute returns, 74% if the McLeod-Li is used, and 83.3% if the new test with $i = [M/3] - 1$ is implemented. Therefore, the new test has higher power and better size properties without a significant increase in the computational burden.

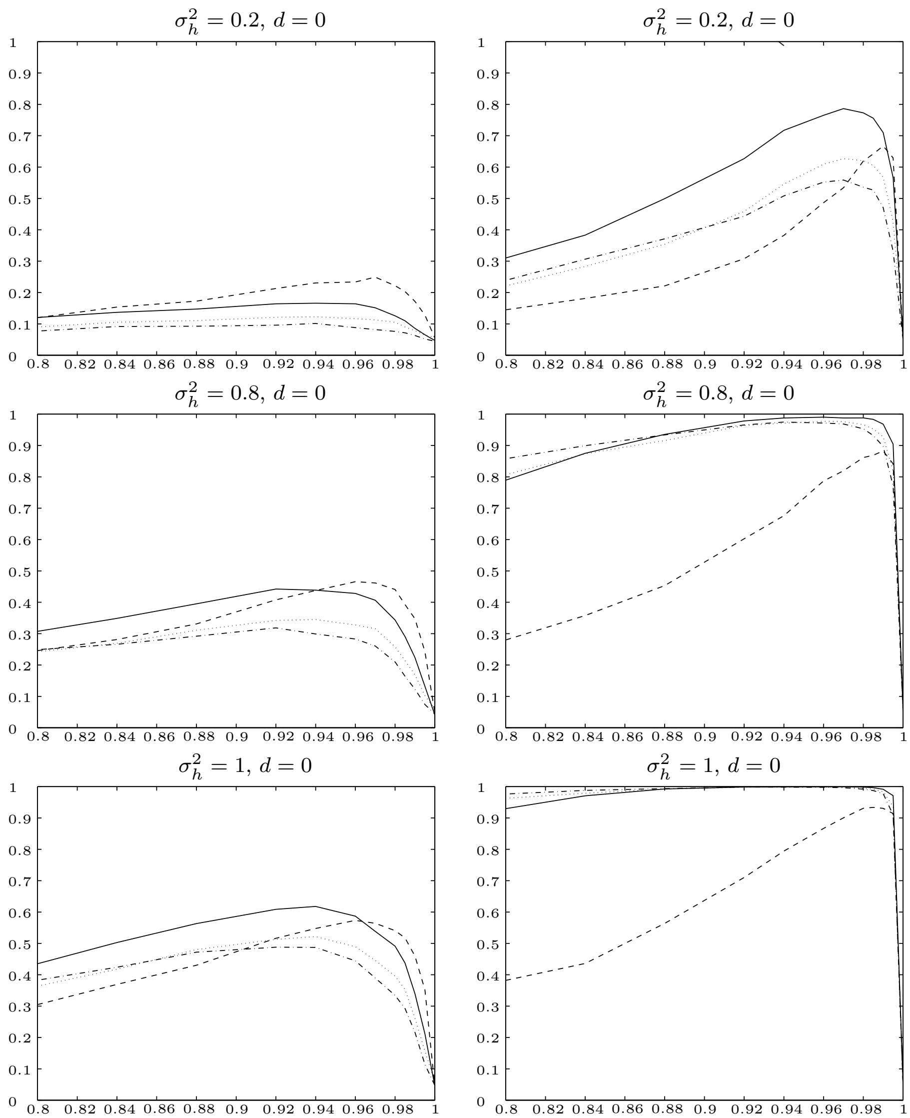

Figure 3. Empirical powers of the Q(24) ($\cdots$), D_{24} ($-\cdot$), NM ($--$) and $Q_7^*(24)$ (—) tests for absolute observations of $ARSV(1)$ processes with $T = 100$ (left panels), $T = 512$ (right panels), and C.V.=0.22 (first row), 0.82 (second row) and 1.72 (third row).

Notice that in Figure 3, the power of $Q_i^*(M)$ has been considered as a function of ϕ for fixed $C.V.$ Therefore, in this figure, ϕ and σ_η^2 are moving together. However, it is of interest to analyze how the power depends on both parameters separately. To analyze this point, Figure 4 plots the powers of $Q_i^*(24)$ for $T = 500$ as a function of ϕ and σ_η^2. This figure illustrates clearly that, as expected, power is an increasing function of both σ_η^2 and ϕ. It also shows that the power depends more heavily on the persistence parameter ϕ than on the variance σ_η^2. When ϕ is relatively small, power is low even if σ_η^2 is large. However, if ϕ is large, power is large even if σ_η^2 is small.

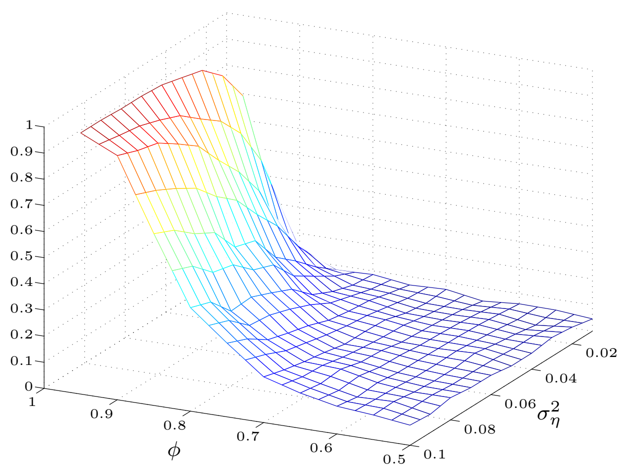

Figure 4. Powers of $Q_i^*(24)$ test, in ARSV(1), as a function of the persistence parameter, ϕ, and the variance of volatility, σ_η^2, when $T = 500$.

3.3. Power for long memory models

Another result often observed in the sample autocorrelations of squared and absolute returns is their slow decay towards zero, suggesting that volatility may have long memory. In this section, we analyze the power of $Q_i^*(M)$ in the presence of long memory.

In the context of SV models, Breidt, Crato and de Lima (1998) and Harvey (1998) have independently proposed the LMSV model where the log-volatility follows an ARFIMA(p, d, q) process. The corresponding LMSV(1,d,0) model is given by

$$\begin{aligned} y_t &= \sigma_* \varepsilon_t \sigma_t, \ t = 1, \ldots, T, \\ (1-L)^d (1-\phi L) \log(\sigma_t^2) &= \eta_t, \end{aligned} \tag{3.1}$$

where $0 \leq d < 1$ is the long memory parameter and L is the lag operator such that $L^j x_t = x_{t-j}$. All parameters and noises are defined as in the short memory ARSV(1) model in (1.3) except the variance of η_t that, in model (3.1), is given by $\sigma_\eta^2 = \{[[\Gamma(1-d)]^2(1+\phi)]/[\Gamma(1-2d)F(1;1+d;1-d;\phi)]\}\sigma_h^2$, where $\Gamma(\cdot)$ and $F(\cdot;\cdot;\cdot;\cdot)$ are the Gamma and Hypergeometric functions, respectively. It is important to notice that although the noises ε_t and η_t in (3.1) are assumed to be independent Gaussian processes, the noise corresponding to the reduced form representation of $|y_t|$ is neither independent nor Gaussian; see Breidt and Davis (1992). The same result applies to y_t^2. Hosking (1996) shows that in this case, the sample autocorrelations have the standard behavior of asymptotic normality and asymptotic variance of order T^{-1} if $d < 0.25$; see also Wright (1999) for this result applied to the logarithm of squares of observations of LMSV models. At the moment, as far as we know, there are no results on the asymptotic properties of the sample autocorrelations of non-linear transformations of observations generated by LMSV processes when $0.25 \leq d \leq 0.50$.

To analyze whether the $Q_i^*(M)$ statistic is also more powerful than its competitors in the presence of long memory, 5,000 time series were generated by model (3.1) with the same $C.V.$ as in Section 3, and long memory parameter $d = \{0.2, 0.4\}$. As in the short memory case, the parameter σ_* was set to one because the statistics are invariant to its value. The values of d have been chosen because, as mentioned above, when $d = 0.2$, the sample autocorrelation of square and absolute observations are expected to have the usual asymptotic properties. Therefore, because of arguments as in Hong (1996), we expect that our proposed test is consistent in this case. On the other hand, we have also generated series with $d = 0.4$ because asymptotic results on the sample autocorrelations have not yet been derived and we want to investigate the finite sample behavior of the statistic. Figures 5 and 6 plot the powers of the $Q(M)$, D_M, NM and $Q_i^*(M)$ tests for $d = 0.2$ and $d = 0.4$, respectively. These figures show that the conclusions about the relative performance in terms of the power of the alternative tests are the same as in the short memory case. The NM is more powerful only when the series are very close to homoscedastic and the sample size is small. On the other hand, the $Q_i^*(M)$ test clearly overperforms the $Q(M)$ and D_M tests when the sample size or the $C.V.$ are small. Finally, notice that for large sample sizes and $C.V.$, the power of the three tests is very similar and close to one. To illustrate the dependence of power on the parameter d, Figure 7 plots the powers of the $Q_0^*(12)$ and $Q_4^*(12)$ as a function of d for $\phi = 0.9$, $\sigma_\eta^2 = 0.01$, and $\phi = 0$, $\sigma_\eta^2 = 0.1$. The power of both statistics seem to depend heavily on the parameter ϕ. When $\phi = 0$ and $T = 1,024$, large values of d (over 0.35) are needed for the power to be over 20%. However, when $\phi = 0.9$, the power is bigger than 20% for all values of d. Furthermore, Figure 7 illustrates the gains for power of the $Q_i^*(M)$ test with respect to the McLeod-Li test.

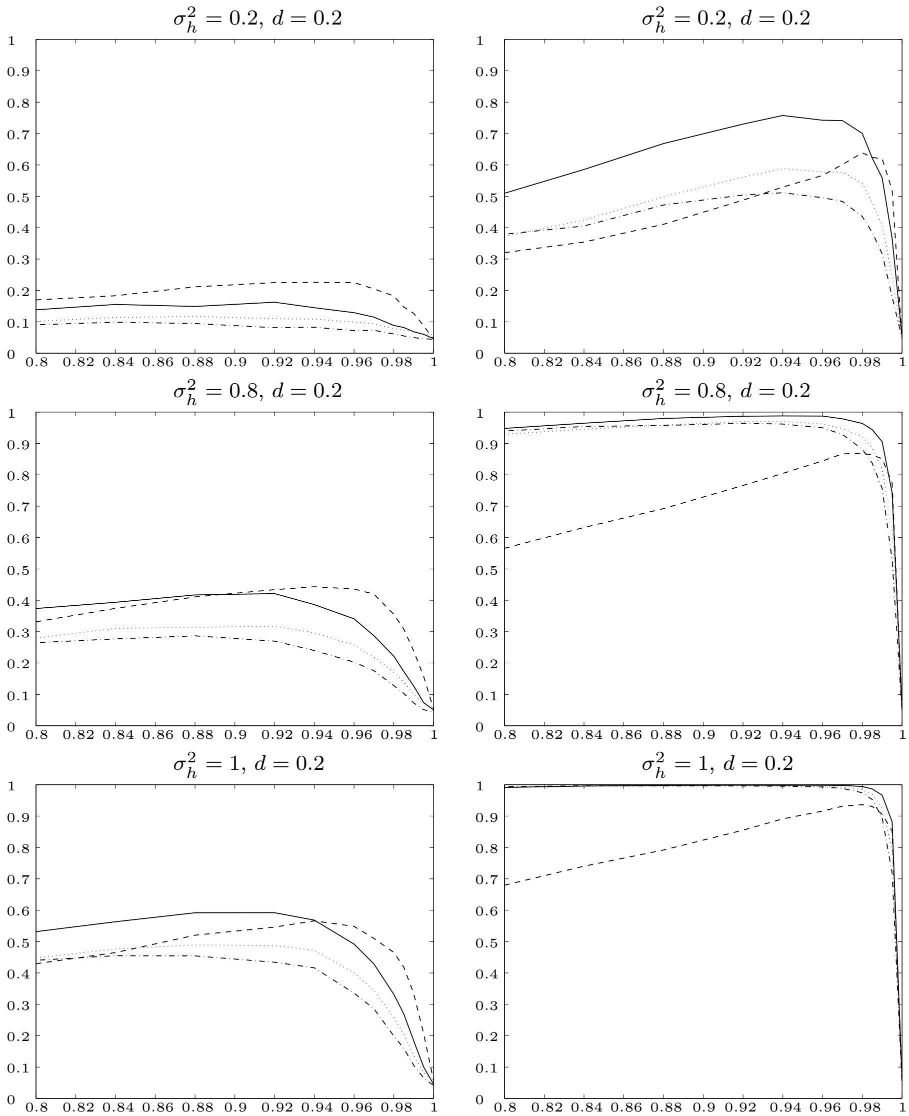

Figure 5. Empirical powers of the Q(24) $(\cdots)$, D_{24} $(-\cdot)$, NM $(--)$ and $Q_7^*(24)$ (—) tests for absolute observations of $LMSV$ processes with long memory parameter $d = 0.2$, sample size $T = 100$ (left panels), $T = 512$ (right panels), and C.V.=0.22 (first row), 0.82 (second row) and 1.72 (third row).

Finally, to illustrate the power gains of the $Q_i^*(M)$ test in the context of LMSV models, Table 3 reports the results of the Monte Carlo experiments for some selected designs which are characterized by generating series where it is hard to detect conditional heteroscedasticity. In particular, we consider $\phi = 0.9$, $d = 0.2$, $\sigma_\eta^2 = 0.01$ and $\phi = 0$, $d = 0.4$, $\sigma_\eta^2 = 0.1$. In this table it is possible to observe that, in the presence of long-memory, the gains in power of $Q_i^*(M)$

with respect to the $Q(M)$, D_M and NM statistics can be very important. For example, when $\phi = 0.9$, $d = 0.2$, $\sigma_\eta^2 = 0.01$ and $T = 512$, the powers of the $Q(12)$, D_{12} and NM tests are 45%, 39.2% and 37.6% respectively, while the power of $Q_3^*(12)$ is 59.3%. This is a substantial increase in power. Even for relatively large sample sizes such as $T = 1,024$, the powers of the $Q(12)$, D_{12} and NM are 71.5%, 69% and 46.6% respectively, while the power of $Q_3^*(M)$ is 85.7%. Consequently,

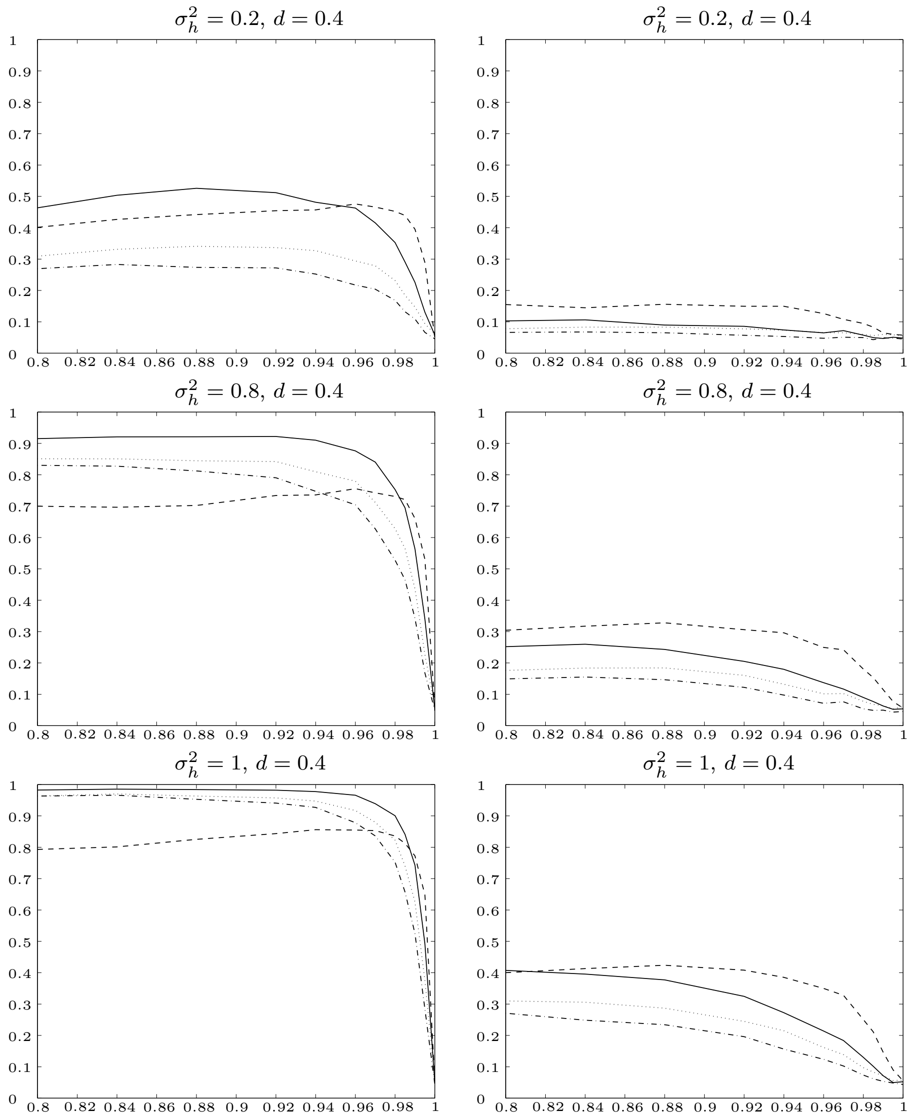

Figure 6. Empirical powers of the Q(24) $(\cdots)$, D_{24} $(-\cdot)$, NM $(--)$ and $Q_7^*(24)$ (—) tests for absolute observations of $LMSV$ processes with long memory parameter $d = 0.4$, sample size $T = 100$ (left panels), $T = 512$ (right panels), and C.V.=0.22 (first row), 0.82 (second row) and 1.72 (third row).

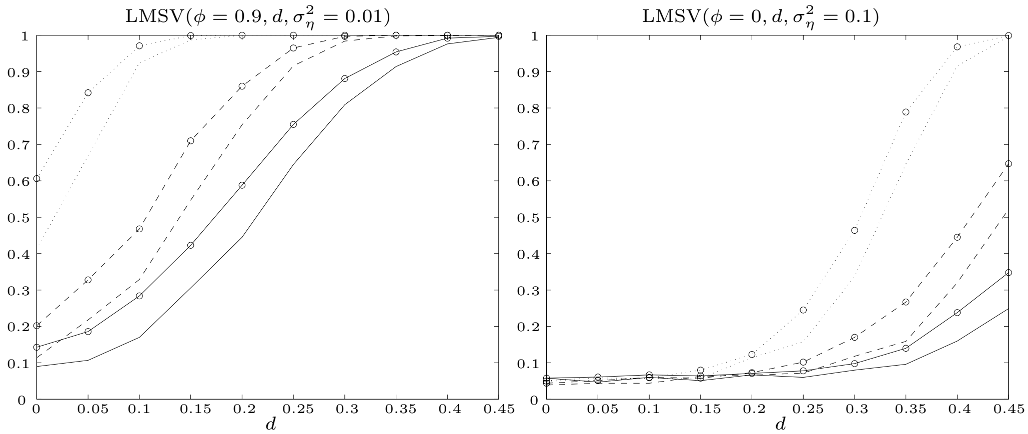

Figure 7. Empirical powers of the nominal 5% $Q_{ML}(12)$ (lines) and $Q_4^*(12)$ (lines with circles) tests for the absolute transformation of $LMSV(1,d,0)$ processes with $T = 512$ (—), T=1,024 (−−) and $T = 4,096$ (···).

Table 3. Empirical powers of the D_M, $Q(M)$, $Q^*_{[M/3-1]}$ and NM tests for absolute observations of $LMSV(1,d,0)$ processes.

	$T = 50$			$T = 100$			$T = 512$			$T = 1,024$		
M	12	24	36	12	24	36	12	24	36	12	24	36
$\{\phi, d, \sigma_\eta^2\} = \{0.98, 0, 0.01\}$												
D_M	0.042	0.028	0.012	0.083	0.057	0.032	0.676	0.674	0.647	0.939	0.946	0.938
$Q(M)$	0.070	0.082	0.090	0.125	0.118	0.113	0.727	0.725	0.713	0.965	0.968	0.964
$Q^*_{[M/3]-1}$	0.070	0.079	0.102	0.165	0.166	0.171	0.828	0.855	0.860	0.986	0.994	0.990
NM	0.132	0.132	0.132	0.266	0.266	0.266	0.665	0.665	0.665	0.775	0.775	0.775
$\{\phi, d, \sigma_\eta^2\} = \{0.9, 0.2, 0.01\}$												
D_M	0.046	0.029	0.015	0.080	0.054	0.036	0.392	0.370	0.326	0.706	0.690	0.646
Q^*_0	0.072	0.080	0.082	0.103	0.105	0.117	0.450	0.405	0.365	0.746	0.715	0.664
$Q^*_{[M/3]-1}$	0.081	0.093	0.118	0.134	0.148	0.160	0.593	0.573	0.563	0.857	0.850	0.830
NM	0.126	0.126	0.126	0.176	0.176	0.176	0.376	0.376	0.376	0.466	0.466	0.466
$\{\phi, d, \sigma_\eta^2\} = \{0, 0.45, 0.1\}$												
D_M	0.039	0.030	0.010	0.061	0.039	0.025	0.229	0.209	0.189	0.493	0.476	0.453
Q^*_0	0.068	0.080	0.084	0.077	0.085	0.096	0.253	0.251	0.220	0.504	0.516	0.514
$Q^*_{[M/3]-1}$	0.065	0.073	0.082	0.091	0.100	0.110	0.363	0.380	0.392	0.661	0.683	0.693
NM	0.083	0.083	0.083	0.133	0.133	0.133	0.382	0.382	0.382	0.564	0.564	0.564

the gains in power of the new test proposed in this paper, compared with the alternative tests, can be very important, especially in the presence of long-memory in the volatility process.

4. Empirical Application

In this section, we implement the $Q(M)$, D_M, NM and $Q_i^*(M)$ statistics to test for conditional homoscedasticity of returns of exchange rates of the Canadian Dollar, Euro and Swiss Franc against the US Dollar, observed daily from April 1, 2,000 until May 21, 2003, with $T = 848$. The data are freely available from the web page of Professor Werner Antweiler, University of British Columbia, Vancouver BC, Canada. The exchange rates have been transformed into returns as usual by taking first differences of logarithms and multiplying by 100, i.e., $y_t = 100(\log(p_t) - \log(p_{t-1}))$. To avoid the influence of large outliers on the properties of the homoscedasticity tests, returns larger than 5 sample standard deviations have been filtered out; see Carnero et al. (2004b) for the influence of outliers on tests for conditional homoscedasticity in the context of GARCH models. The three series of returns have been plotted in Figure 8, together with the corresponding autocorrelations of absolute returns. These are rather small, always under 0.1 in absolute value. However, observe that with the exception of the Swiss Franc, the autocorrelations are mainly positive, which is incompatible with independent observations.

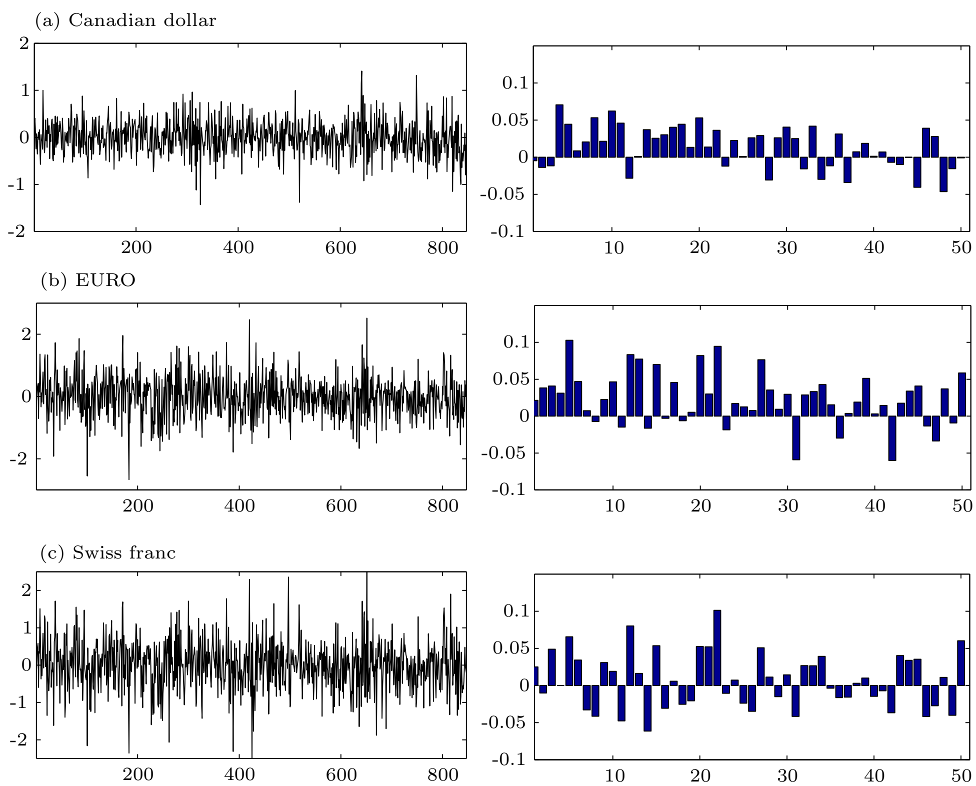

Figure 8. Series of returns with the corresponding autocorrelations of absolute returns.

For each exchange rate, Table 4 reports the ratio between the value of the statistic and the corresponding 5% critical value for the NM, D_M, $Q(M)$ and $Q^*_{[M/3]-1}(M)$ statistics, for $M = 10$, 20, 30 and 50, when implemented on absolute returns. Looking at the results for the Canadian Dollar, the McLeod-Li and Peña-Rodriguez tests do not reject the null hypothesis of conditional homoscedasticity for any value of M. The Harvey-Streibel test rejects the null, although the statistic is relatively close to the critical value. Finally, the $Q^*_{[M/3]-1}(M)$ statistic clearly rejects the null hypothesis, especially for large values of M such as $M = 30$ or 50. Conclusions on whether the Canadian Dollar returns are homoscedastic are contradictory, depending on the particular statistic used to test the null.

Table 4. Ratio between the value of the statistic and the corresponding 5% critical value for the NM, D_M, $Q(M)$ and $Q^*_{[M/3-1]}(M)$ statistics for $M = 10, 20, 30$ and 50 when implemented on absolute returns.

	(a) Canadian dollar				(b) EUROS				(c) Swiss franc			
M	10	20	30	50	10	20	30	50	10	20	30	50
$Q(M)$	0.70	0.75	0.68	0.61	0.94	1.30	1.31	1.17	0.59	0.91	1.01	0.91
$Q_i^*(M)$	1.05	1.72	2.29	1.97	1.77	2.60	3.66	3.48	0.56	0.31	0.50	0.37
D_M	0.51	0.65	0.64	0.62	0.84	1.04	1.11	1.07	0.56	0.83	0.98	0.95
NM	1.28	1.28	1.28	1.28	3.34	3.34	3.34	3.34	0.69	0.69	0.69	0.69

A similar result is obtained for the returns of the exchange rates of the Euro against the Dollar. In this case, the McLeod-Li and Peña-Rodriguez tests are just on the boundary of the non-rejection region when the size is 5% and the number of correlations considered in the statistic is 20, 30 or 50. Furthermore, if $M = 10$, these tests suggest that this series is conditionally homoscedastic. However, the Harvey-Streibel and the $Q^*_{[M/3]-1}(M)$ statistics clearly reject the null hypothesis.

Finally, looking at the results for the Swiss-Franc exchange rates, the situation is somehow reversed. As in previous examples, the $Q(M)$ and $D(M)$ tests are close to the boundary of the rejection region when the size is 5%. However, the other two tests are more conclusive and they do not reject the null.

In the three examples considered in this section, it seems that taking into account not only the magnitude but also the pattern of the sample autocorrelations of absolute returns helps to obtain a clearer answer on whether the corresponding returns are homoscedastic or heteroscedastic. In the three cases, the statistics that only account for the magnitude of the autocorrelations are rather inconclusive while our proposed test gives a clearer answer. In these empirical examples, the result of the test proposed by Harvey and Streibel (1998) is in concordance with the $Q^*_{[M/3]-1}(M)$ statistic.

Acknowledgement

We are very grateful to Mike Wiper for his help with the final draft. This research has been supported by project BEC2002.03720 from the Spanish Goverment and the Fundación BBVA.

References

Andersen, T. G. and Bollerslev, T. (1997). Heterogenous information arrivals and return volatility dynamics: Uncovering the long-run in high frequency returns. *J. Finance* **52**, 975-1005.

Baillie R. T., Bollerslev, T. and Mikkelsen H. O. (1996). Fractionally integrated autoregressive conditional heteroskedasticity. *J. Econometrics* **74**, 3-30.

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

Bollerslev, T. and Mikkelsen, H. O. (1999). Long-term equity anticipation securities and stock market volatility dynamics. *J. Econometrics* **92**, 75-99.

Box, G. E. P. (1954). Some theorems on quadratic forms applied in the study of analysis of variance problems I: Effect of the inequality of variance in the one-way classification. *Ann. Math. Statist.* **25**, 290-302.

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

Breidt, F. J. and Davis, R. A. (1992). Time-Reversibility, identifiability and indendence of innovations for stationary time series. *J. Time Ser. Anal.* **25**, 265-281.

Carnero, M. A., Peña, D. and Ruiz, E. (2004a). Persistence and kurtosis in GARCH and stochastic volatility models. *J. Finan. Econom.* **13**, 377-390.

Carnero, M. A., Peña, D. and Ruiz, E. (2004b). Spurious and hidden volatility. WP 04-20 (07), Universidad Carlos III de Madrid.

Chen, W. W. and Deo, R. S. (2004). On power transformations to induce normality and their applications. *J. Roy. Statist. Soc. Ser. B* **66**, 117-130.

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

Engle, R. F. (1982). Autoregressive conditional heteroskedasticity with estimates of the variances. *Econometrica* **50**, 987-1007.

Ferguson, T. S. (1967). *Mathematical Statistics: A Decision Theoretic Approach.* Academic Press, New York.

Ghysels, E., Harvey, A. C. and Renault, E. (1996). Stochastic volatility. In *Statistical Methods in Finance* (Edited by C. R. Rao and G. S. Maddala), 119-191. North-Holland, Amsterdam.

Hannan, E. J. (1970). *Multiple Time Series.* Wiley, New York.

Harvey, A. C. (1998). Long memory in Stochastic Volatility. In *Forecasting Volatility in Financial Markets* (Edited by J. Knight and S. Satchell), 307-320. Butterworth-Haineman, Oxford.

Harvey, A. C. and Streibel, M. (1998). Testing for a slowly changing level with special reference to stochastic volatility. *J. Econometrics* **87**, 167-189. In *Recent Developments in Time Series*, 2003 (Reproduced by P. Newbold and S. J. Leybourne). Edward Elgar Publishing.

Hong, Y. (1996). Consistent testing for serial correlation of unknown form. *Econometrica* **64**, 837-864.

Jacquier, E., Polson, N. G. and Rossi, P. E., (1994). Bayesian analysis of stochastic volatility models (with discussion). *J. Bus. Econom. Statist.* **12**, 371-417.

Karanasos, M., Psaradakis, Z. and Sola, M. (2004). On the autocorrelation properties of long-memory GARCH processes. *J. Time Ser. Anal.* **25**, 265-281.

Koch, P. D. and Yang, S. (1986). A method for testing the independence of two time series that accounts for a potential pattern in the cross-correlation function. *J. Amer. Statist. Assoc.* **81**, 533-544.

Ljung, G. M. and Box, G. E. P. (1978). On a measure of lack of fit in time series models. *Biometrika* **65**, 297-303.

McLeod, A. I. and Li, W. K. (1983). Diagnostic checking ARMA time series models using squared-residual autocorrelations. *J. Time Ser. Anal.* **4**, 269-273.

Palma, W. and Zevallos, M. (2004). Analysis of the correlation structure of square time series. *J. Time Ser. Anal.* **25**, 529-550.

Peña, D. and Rodríguez, J. (2002). A powerful portmanteau test of lack of fit for time series. *J. Amer. Statist. Assoc.* **97**, 601-610.

Peña, D. and Rodríguez, J. (2005a). The log of the determinant of the autocorrelation matrix for testing goodness of fit in time series. *J. Statist. Plann. Inference.* Forthcoming.

Peña, D. and Rodríguez, J. (2005b). Detecting nonlinearity in time series by model selection criteria. *Internat. J. Forecasting.* Forthcoming.

Pérez, A. and Ruiz, E. (2003). Properties of the sample autocorrelations of non-linear transformations in long memory stochastic volatility models. *J. Financial Econometrics* **1**, 420-444.

Satterthwaite, F. E. (1941). Synthesis of Variance. *Psychometrica* **6**, 309-316.

Satterthwaite, F. E. (1946). An approximate distribution of estimates of variance components. *Biometrics Bull.* **2**, 110-114.

Taylor, J. H. (1986). *Modelling Financial Time Series.* Wiley, New York.

Tol, R. S. J. (1996). Autoregressive conditional heteroscedasticity in daily temperatures measurements. *Environmetrics* **7**, 67-75.

Departamento de Estadística, Universidad Carlos III de Madrid, C/ Madrid 126, 28903 Getafe, Madrid, Spain.

E-mail: ortega@est-econ.uc3m.es

Laboratorio de Estadística, Universidad Politécnica de Madrid, Avda. Ramiro de Maeztu, 7, E-28040 Madrid, Spain.

E-mail: puerta@etsii.upm.es

(Received November 2003; accepted December 2004)