E C O N O M I C S   B U L L E T I N

On the Robustness of Ljung−Box and McLeod−Li Q Tests: A Simulation Study

Yi−Ting Chen
Sun Yat−Sen Institute for Social Sciences and Philosphy, Academia Sinica

Abstract

In financial time series analysis, serial correlations and the volatility clustering effect of asset returns are commonly checked by Ljung−Box and McLeod−Li Q tests and filtered by ARMA−GARCH models. However, this simulation study shows that both the size and power performance of these two tests are not robust to heavily tailed data. Further, these Q tests may reject processes without ARMA−GARCH structures simply because of nonlinearity and conditionally heteroskedastic higher−order moments. These results imply that, to avoid misleading interpretations on time series data, these two tests should be used with care in practical applications.

Citation: Chen, Yi−Ting, (2002) "On the Robustness of Ljung−Box and McLeod−Li Q Tests: A Simulation Study." *Economics Bulletin*, Vol. 3, No. 17 pp. 1−10
Submitted: July 23, 2002. **Accepted:** September 5, 2002.
URL: http://www.economicsbulletin.com/2002/volume3/EB−02C40014A.pdf

1 Introduction

Autocorrelation is a leading measure of serial dependence in time series analysis. To understand the dependence structure of time series data, researchers routinely calculate sample autocorrelations and check the joint significance of these statistics by the Q test of Box and Pierce (1970, the Q_{BP} test) or Ljung and Box (1978, the Q_{LB} test). In financial time series analysis, it is particularly important to check serial correlations of squared series; such dependence is referred to as the volatility clustering effect. The Q test of McLeod and Li (1983, the Q_{ML} test) is used for this purpose. The Q_{ML} test is indeed a particular Q_{LB} test based on sample autocorrelations of squared series, and it is asymptotically equivalent to the LM test of Engle (1982). To explain the serial correlations and volatility clustering effect detected by these Q tests, empirical researchers often filter time series data by certain autoregressive moving averages and generalized autoregressive conditional heteroskedasticity (ARMA-GARCH) models. These Q tests are also applied on the residuals of estimated models as diagnostic checks. Although this modelling strategy is quite popular in empirical studies, it may encounter some problems in practice.

Note that the above Q tests require that the process being tested be an independently and identically distributed (i.i.d.) sequence. This condition is much more stringent than the hypothesis of white noise, so that the Q_{BP} and Q_{LB} tests (the Q_{ML} test) may reject the null hypothesis not only for serial correlations (the volatility clustering effect) but also for other types of serial dependence. Further, the Q_{BP} and Q_{LB} tests (the Q_{ML} test) need data to possess finite fourth (eighth) moment. Therefore, the performance of these tests is likely to be influenced by the fatness of distributional tails. In fact, most risky assets' returns have heavily tailed distributions. Several studies even reported that such distributions may not have finite fourth moment; see Jansen and de Vries (1991), Loretan (1994) and Phillips, and de Lima (1997). Consequently, it is important to investigate the robustness of these Q tests in these situations.

For this purpose, we conduct a Monte Carlo simulation to assess finite sample performance of the Q_{LB} and Q_{ML} tests. The Q_{BP} test will not be discussed in this study because the Q_{LB} test is already a finite-sample-corrected Q_{BP} test. The simulation results show that distributional heavy-tails may distort the asymptotic null distributions of the Q_{LB} and Q_{ML} tests and reduce the power of these tests. Further, these Q tests may reject their null hypotheses because of nonlinearity or conditionally heteroskedastic skewness, even if the process being tested does not have any ARMA-GARCH structures. In consequence, these Q tests should be used with care to avoid yielding misleading interpretations on time series data, and it is important to consider the use of tests for moment conditions, nonlinearity, and independence to complement these Q tests before accepting ARMA-GARCH models.

The rest of this paper is organized as follows. In Section 2, we write the notation and the Q_{LB} and Q_{ML} test statistics. In Section 3, we introduce the simulation experiment design. The simulation results will be shown and discussed in Section 4. Finally, our conclusion is in Section 5.

2 The Ljung-Box and McLeod-Li Q tests

Let $\{y_t\}$ be a time series with the lag-k autocorrelation:

$$\rho_1(k) = \frac{\text{cov}(y_t, y_{t-k})}{\text{var}(y_t)}, \qquad k = 1, 2, \ldots.$$

The lag-k sample autocorrelation of $\{y_t\}$ is

$$\hat{\rho}_1(k) = \frac{\sum_{t=k+1}^{T}(y_t - \bar{y}_T)(y_{t-k} - \bar{y}_T)}{\sum_{t=1}^{T}(y_t - \bar{y}_T)^2},$$

where T denotes the sample size; $\bar{y}_T$ is the sample mean of $\{y_t\}$. To test if $\{y_t\}$ is a white noise sequence, it is common to check the hypotheses consisted of the first m autocorrelations:

$$H_o \ : \ \rho_1(1) = \rho_1(2) = \ldots = \rho_1(m) = 0,$$

$$H_1 \ : \ \rho_1(k) \neq 0, \quad \text{for some } k = 1, 2, \ldots, m.$$

For these hypotheses, the Q_{LB} test statistic is of the form:

$$Q_{LB}(m) = T(T+2)\sum_{k=1}^{m}\frac{\hat{\rho}_1^2(k)}{T-k}.$$

This statistic has the asymptotic null distribution $\chi^2(m)$ provided that $\{y_t\}$ is an i.i.d sequence and that y_t has a finite fourth moment.

The Q_{ML} test is a particular Q_{LB} test based on sample autocorrelations of the squared series $\{y_t^2\}$. Its test statistic is

$$Q_{ML}(m) = T(T+2)\sum_{k=1}^{m}\frac{\hat{\rho}_2^2(k)}{T-k},$$

where

$$\hat{\rho}_2(k) = \frac{\sum_{t=k+1}^{T}(y_t^2 - \hat{\sigma}_T^2)(y_{t-k}^2 - \hat{\sigma}_T^2)}{\sum_{t=1}^{T}(y_t^2 - \hat{\sigma}_T^2)^2}, \qquad \hat{\sigma}_T^2 = \frac{1}{T}\sum_{t=1}^{T} y_t^2.$$

This statistic also has the asymptotic null distribution $\chi^2(m)$ provided that $\{y_t\}$ is an i.i.d. sequence and that y_t possesses a finite eighth moment. This test is often used to check the whiteness of $\{y_t^2\}$:

$$\rho_2(k) := \frac{\text{cov}(y_t^2, y_{t-k}^2)}{\text{var}(y_t^2)} = 0, \qquad \text{for all } k = 1, 2, \ldots, m,$$

against the volatility clustering effect. As we have noted previously, the Q_{LB} and Q_{ML} tests also serve as diagnostic checks for estimated ARMA-GARCH models. However, in this case the degrees of freedom of the asymptotic null distribution of $Q_{LB}(m)$ and $Q_{ML}(m)$ have to be reduced by the number of parameters estimated.

3 The data generation processes

In this simulation experiment, we consider five types of data generation processes (DGPs): i.i.d. sequences, a simple nonlinear process, a conditionally heteroskedastic skewness process, an autoregressive (AR) process, and a GARCH process with different innovation distributions. The first three DGPs are designed to represent processes without ARMA-GARCH structures. Supposedly, the Q_{LB} and Q_{ML} tests should be powerful against the last two DGPs but not the first three DGPs. The details of these DGPs are described as follows.

IID sequences

In this case,

$$y_t = \varepsilon_t, \tag{1}$$

where $\{\varepsilon_t\}$ is an i.i.d. sequence. To study the influence of the tail behavior of ε_t on the Q tests, we consider three distributions for ε_t: the standard normal distribution $N(0,1)$, Student t distribution with the degrees of freedom ν, denoted as $t(\nu)$, and the standardized log-normal distribution:

$$\varepsilon_t = \frac{\exp(\lambda u_t - \lambda^2/2) - 1}{(\exp(\lambda^2) - 1)^{1/2}}, \qquad u_t \sim N(0,1),$$

with the skewness parameter λ, denoted by $L(\lambda)$. The latter two distributions can yield different tail properties by adjusting their parameters. It is well known that $t(\nu)$ is a symmetric distribution without finite ν-th moment. As compared to $N(0,1)$, $t(\nu)$ has heavier tails and higher peakedness when ν is small. As $\nu \to \infty$, this distribution converges to $N(0,1)$. On the other hand, $L(\lambda)$ is a right-skewed distribution with zero mean and unit variance. The fatness of its right tail, and hence skewness and kurtosis, increase with the value of $|\lambda|$. This distribution also encompasses $N(0,1)$ as $\lambda \to 0$; see Johnson et al. (1995).

A simple nonlinear process

This nonlinear process is of the form:

$$y_t = \begin{cases} \mu_1 + \sigma_1 \varepsilon_t, & y_{t-1} > 0, \\ \mu_2 + \sigma_2 \varepsilon_t, & y_{t-1} \le 0, \end{cases} \tag{2}$$

where $\{\varepsilon_t\}$ is an i.i.d. $N(0,1)$ sequence. The magnitudes of $|\mu_1 - \mu_2|$ and $|\sigma_1 - \sigma_2|$ control the significance of the location and scale switching behavior of $\{y_t\}$, respectively. Obviously, this process has no ARMA-GARCH structures. Granger and Teräsvirta (1999) used a particular case of this process, $(\mu_1, \mu_2; \sigma_1, \sigma_2) = (1, -1; 1, 1)$, to illustrate that autocorrelations may yield misleading linear properties.

A conditionally heteroskedastic skewness process

Some studies recently aimed at investigating conditionally heteroskedastic skewness of asset returns; see e.g., Hansen (1994) and Harvey and Siddique (1999). To represent this DGP, we consider that y_t is distributed as $L(\lambda_t)$ with the skewness parameter λ_t, which follows the first-order AR process:

$$\lambda_t = \delta_o + \delta_1 \lambda_{t-1} + \varepsilon_t, \tag{3}$$

where $\{\varepsilon_t\}$ is an i.i.d. $N(0,1)$ sequence. The process $\{y_t\}$ has zero conditional mean and unit conditional variance, so that it does not have serial correlations and the volatility clustering effect. However, it is still a serially dependent process governed by the conditionally heteroskedastic skewness parameters $\{\lambda_t\}$.

An AR process

This case considers the first-order AR process:

$$y_t = \alpha y_{t-1} + \varepsilon_t, \tag{4}$$

where $\{\varepsilon_t\}$ is an i.i.d. sequence with $N(0,1)$, $t(\nu)$, or $L(\lambda)$. In this DGP, $\{y_t\}$ has serial correlations but no volatility clustering effect.

A GARCH process

This GARCH process is of the form:

$$y_t = \varepsilon_t h_t^{1/2}, \quad h_t = \beta_o + \beta_1 y_{t-1}^2 + \beta_2 h_{t-1}, \tag{5}$$

where $\{\varepsilon_t\}$ is an i.i.d. sequence with $N(0,1)$, $t(\nu)$, or $L(\lambda)$. This process has volatility clustering effect but no serial correlations.

4 The simulation results

To implement this experiment, we consider $T = 100, 500$, $m/T = 5\%, 10\%, 30\%$, $\nu = 1, 3, 5$, $\lambda = 0.5, 1, 2$, $\alpha = \pm 0.3, \pm 0.7$, $(\beta_o, \beta_1, \beta_2) = (1, 0.3, 0.6), (1, 0.05, 0.9)$, $(\mu_1, \mu_2; \sigma_1, \sigma_2) = (1, -1; 1, 1)$, (1,1;1,4), (1,-1;1,4), (δ_o, δ_1)= (1,0.3),(1,0.5),(1,0.7),(1,0.9), and many other parameter settings. However, due to the reason of space limitations, we only report the results of $T = 500$ with selected parameter settings. Other simulation results are available upon request. For comparing $\chi^2(m)$ and the empirical distributions of $Q_{LB}(m)$ and $Q_{ML}(m)$, the nominal level θ considered includes θ =50%, 25%, 10%, 5%, and 1%. In Table 1, we show the empirical rejection frequencies of the Q_{LB} and Q_{ML} tests for DGPs (1), (2), and (3). The results of DGPs (4) and (5) are shown in Table 2. These rejection frequencies are calculated from 5000 replications.

[Tables 1 and 2 about here.]

4.1 DGPs without ARMA-GARCH structures

We first discuss DGP (1), the case of i.i.d. sequences. For the distribution $N(0,1)$, Table 1 shows that the empirical rejection frequencies of $Q_{LB}(m)$ and $Q_{ML}(m)$ are quite close to their respective nominal levels when $m/T = 5\%, 10\%$. This result supports that $\chi^2(m)$ is a proper approximation to the null distributions of these two statistics in this standard case. However, when $m/T = 30\%$, the empirical sizes of $Q_{LB}(m)$ and $Q_{ML}(m)$ are 9.6% and 8.3% (4.0% and 3.2%) for $\theta = 5\%$ (1%), respectively; that is, these Q tests are somewhat over-sized for large m/T. This evidence is likely caused by the fact that the variance of $Q_{LB}(m)$ is larger than which of $\chi^2(m)$ when m/T is large; see Davies et al. (1977).

de Lima (1997) reported that the Q_{ML} test is sensitive to the heavy tails of Pareto distributions. Table 1 further shows that both the Q_{LB} and Q_{ML} tests are not robust to the tail behavior of $t(\nu)$ and $L(\lambda)$. For example, given the distribution $L(2)$ and $m/T = 5\%$, the empirical rejection frequencies of the Q_{LB} (Q_{ML}) test are 20.6%, 14.5%, 10.5%, 8.6%, 6.6% (10.8%, 9.1%, 7.6%, 6.9%, 5.6%) for θ =50%, 25%, 10%, 5%, and 1%, respectively. This illustrates that $\chi^2(m)$ is inappropriate to approximate the null distributions of $Q_{LB}(m)$ and $Q_{ML}(m)$ in this case. Similar examples can also be found for other distributions listed in Table 1. From this table, we also observe that the Q_{ML} test is much more sensitive to heavy tails of data. This is because the Q_{ML} test requires a stricter moment condition than the Q_{LB} test.

Table 1 also shows that the Q_{LB} and Q_{ML} tests are not only sensitive to the tail behavior of data but also not robust to nonlinearity and conditionally heteroskedastic higher-order moments of the process being tested. Recall that DGP (2) with $(\mu_1, \mu_2; \sigma_1, \sigma_2) = (1, -1; 1, 1)$ is a nonlinear process with switching locations and a fixed scale parameter. In this case, the Q_{LB} test rejects the null with the frequency 100% for all values of θ and m/T considered. On the other hand, the Q_{ML} test has proper size performance when $m/T = 5\%, 10\%$. This result reminds us that this nonlinear process might be fitted by a misleading ARMA model, if the significance of $Q_{LB}(m)$ is misleadingly interpreted as a signal of the need for certain ARMA models to filter the "detected serial correlations" of data. In fact, the Q_{ML} test may also over-reject the null hypothesis when this nonlinear process has regime switching scale parameters. For example, given $(\mu_1, \mu_2; \sigma_1, \sigma_2) = (1, -1; 1, 4)$, $\theta = 5\%$, and $m/T = 5\%$, the empirical rejection frequency of the Q_{ML} test is 29.4%; even though, this process has no GARCH structures.

For the conditionally heteroskedastic skewness process with $(\delta_o, \delta_1) = (1, 0.5)$, given $m/T =$ 5%, the Q_{LB} (Q_{ML}) test has the empirical rejection frequencies 36.1% and 3.4% (17.9% and 6.6%) for θ =50% and 1%, respectively. Hence, $\chi^2(m)$ is not a proper approximation to the null distributions of $Q_{LB}(m)$ and $Q_{ML}(m)$ in this case. For certain parameter settings, the Q_{LB} and Q_{ML} tests may reject their null hypotheses with high probabilities, even if there is no serial correlations and the volatility clustering effect. For instance, given $(\delta_o, \delta_1) = (1, 0.9)$ and $m/T =$ 5%, these two tests have the empirical rejection frequencies 52.9% and 32.0%, respectively.

4.2 DGPs with an AR or GARCH structure

For the AR (GARCH) process, as expected, Table 2 shows that the Q_{LB} (Q_{ML}) test can powerfully reject the null hypotheses when the innovations are distributed as $N(0,1)$. For example, given $m/T = 5\%$ and $\theta = 5\%$, the Q_{LB} test has the power 99.4% against the AR process with $\alpha = 0.3$; the Q_{ML} test rejects the GARCH process with $(\beta_o, \beta_1, \beta_2) = (1, 0.3, 0.6)$ and $(1, 0.05, 0.9)$ with the frequencies 99.3% and 87.0%, respectively. However, the power performance of these two tests is still influenced by the tail behavior of innovation distribution. Given $m/T = 10\%$ and $\theta = 5\%$, the power of the Q_{LB} test against the AR process with $\alpha = 0.3$ drops from 96.8% to 70.8% when the innovation distribution $N(0,1)$ is replaced by $L(2)$; on the other hand, the Q_{ML} test has the power 98.3% against the GARCH process with $(\beta_o, \beta_1, \beta_2) = (1, 0.3, 0.6)$ when the innovations are distributed as $N(0,1)$, but the power is reduced to 36.2% for $L(1)$ distributed innovations.

From this study, we also see that the Q_{LB} (Q_{ML}) test may reject the null hypothesis with certain probabilities because of the volatility clustering effect (serial correlations). Table 2 shows that, given $m/T = 5\%$ and $\theta = 5\%$, the Q_{LB} test has the rejection frequency 85.9% for the GARCH process with $(\beta_o, \beta_1, \beta_2) = (1, 0.3, 0.6)$ and $N(0,1)$, even though this GARCH process is a white noise sequence. On the other hand, given the same m/T and θ, the Q_{ML} test has the rejection frequency 16.5% against the AR process with $\alpha = 0.3$ and $N(0,1)$ innovations. This process contains no GARCH structures, but the above rejection frequency is much larger than the 5% nominal level. This evidence demonstrates the difficulty of using these two Q tests to distinguish between serial correlations and the volatility clustering effect.

5 Conclusions

This simulation study shows that both the size and power performance of the Q_{LB} and Q_{ML} tests are not robust to heavily tailed data. Furthermore, these two tests may reject their null hypotheses because of certain types of nonlinearity and conditionally heteroskedastic higher-order moments, even if the process being tested does not have ARMA-GARCH structures. For the processes with AR or GARCH structures, these two tests may not be able to distinguish between serial correlations and the volatility clustering effect. In consequence, we may not completely rely on these Q tests to construct and check ARMA-GARCH models. The results of this study also support the importance of developing new portmanteau tests for serial correlations which can be robust to the volatility clustering effect; see e.g., Lobato et al. (2001) and the references therein. However, we still need other portmanteau tests for serial correlations and the volatility clustering effect that are robust to nonlinearity, conditional heteroskedastic higher-order moments, and distributional heavy tails.

References

Box, G. E. D. and D. Pierce (1970). Distribution of autocorrelations in autoregressive moving average time series models, *Journal of the American Statistical Association*, **65**, 1509–1526.

Davies, N., C. M. Triggs, and P. Newbold (1977). Significance of the Box-Pierce portmanteau statistics in finite samples, *Biometrika*, **64**, 517–522.

de Lima, P. J. F. (1997). On the robustness of nonlinearity tests to moment condition failure, *Journal of Econometrics*, **76**, 251–280.

Engle, R. F. (1982). Autoregressive conditional heteroskedasticity with estimates of the variance of United Kingdom inflation, *Econometrica*, **50**, 987–1006.

Granger, C. W. and T. Teräsvirta (1999). A simple nonlinear time series model with misleading linear properties, *Economics Letters*, **62**, 161–165.

Hansen, B. E. (1994). Autoregressive conditional density estimation, *International Economic Review*, **35**, 705–730.

Harvey, C. R. and A. Siddique (1999). Autoregressive conditional skewness, *Journal of Finance and Quantitative Analysis*, **34**, 465–487.

Jansen, D. W. and C. G. de Vries (1991). On the frequency of large stock returns: Putting booms and busts into perspective, *Review of Economics and Statistics*, **73**, 18–24.

Johnson, N. L., S. Kotz, and N. Balakrishnan (1995), *Continuous Univariate Distributions- Volume 2*, New York: Wiley.

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

Lobato, I., J. Nankervis, and N. E. Savin (2001). Testing for Autocorrelation Using a Modified Box-Pierce Q Test, *International Economic Review*, **42**, 187–205.

Loretan, M. and P. C. B. Phillips (1994). Testing the covariance stationarity of heavy-tailed time series, *Journal of Empirical Finance*, **1**, 211–248.

McLeod, A. I. and W. K. Li (1983). Diagnostic checking ARMA time series models using squared-residual autocorrelations, *Journal of Time Series Analysis*, **4**, 269–273.

Table 1: Rejection frequencies of $Q_{LB}(m)$ and $Q_{ML}(m)$ for DGPs (1), (2), and (3).

			Ljung-Box $Q(m)$					McLeod-Li $Q(m)$				
DGP	m/T	$\theta =$	50%	25%	10%	5%	1%	50%	25%	10%	5%	1%
(1) i.i.d. sequences												
	5%		50.7	24.8	10.4	5.6	1.3	44.9	22.2	9.7	5.4	1.5
$N(0,1)$	10%		48.2	25.1	11.4	6.6	1.6	44.9	23.4	10.2	5.8	1.7
	30%		48.0	28.5	14.7	9.6	4.0	41.8	24.3	12.3	8.3	3.2
	5%		44.2	22.4	9.1	5.0	1.4	20.1	13.6	9.3	7.3	4.8
$t(3)$	10%		43.0	21.4	9.7	5.7	1.7	20.9	14.2	9.4	7.4	4.3
	30%		39.1	22.9	12.1	7.4	2.9	16.9	10.2	5.9	4.2	2.3
	5%		16.2	12.5	9.5	7.8	5.8	7.7	6.5	5.5	5.0	4.2
$t(1)$	10%		16.4	12.4	9.5	8.3	5.9	8.9	7.5	6.3	5.9	4.9
	30%		10.7	6.9	4.1	2.9	1.2	4.5	2.6	1.4	0.8	0.4
	5%		38.6	19.3	8.8	5.5	2.1	18.3	13.8	10.9	9.3	6.4
$L(1)$	10%		36.4	19.1	9.3	5.9	2.2	19.1	13.9	10.4	8.4	6.0
	30%		34.0	19.6	10.8	7.2	2.9	13.9	8.5	4.9	3.5	1.7
	5%		20.6	14.5	10.5	8.6	6.0	10.8	9.1	7.6	6.9	5.6
$L(2)$	10%		21.9	14.7	10.5	8.0	4.9	12.3	10.4	8.5	7.6	6.0
	30%		17.2	10.9	6.6	4.4	1.9	7.7	4.6	2.5	1.8	0.7
(2) nonlinear process with the parameters $(\mu_1, \mu_2; \sigma_1, \sigma_2)$												
	5%		100	100	100	100	100	47.7	23.9	9.7	5.3	1.5
(1,-1;1,1)	10%		100	100	100	100	100	47.6	24.7	10.8	5.8	1.4
	30%		100	100	100	100	100	45.2	26.3	14.2	8.8	3.6
	5%		99.5	98.2	95.5	92.7	85.7	70.8	53.3	37.9	29.4	17.5
(1,-1;1,4)	10%		98.6	95.8	91.0	86.6	76.4	61.6	43.9	29.3	22.4	12.7
	30%		94.9	88.9	80.9	75.0	61.9	49.9	35.3	24.3	18.7	10.5
(3) conditionally heteroskedastic skewness process with the parameter (δ_o, δ_1)												
(1,0.5)	5%		36.1	20.8	11.1	7.5	3.4	17.9	13.9	10.5	9.0	6.6
	10%		33.1	18.6	10.3	6.8	3.2	17.0	13.4	10.4	8.9	6.2
	30%		28.7	16.7	9.7	6.3	2.8	11.2	7.1	4.5	3.2	1.6
(1,0.9)	5%		62.2	59.9	57.7	56.0	52.9	39.1	36.9	35.2	34.0	32.0
	10%		50.4	47.4	44.3	42.7	39.2	31.4	29.8	28.2	27.0	24.9
	30%		14.6	12.3	10.9	10.2	8.5	8.9	7.7	6.8	6.3	5.4

Note: The entries are rejection frequencies in percentages; θ denotes the nominal size; $T = 500$.

Table 2: Rejection frequencies of $Q_{LB}(m)$ and $Q_{ML}(m)$ for DGPs (4) and (5).

			Ljung-Box $Q(m)$					McLeod-Li $Q(m)$				
DGP	m/T	$\theta =$	50%	25%	10%	5%	1%	50%	25%	10%	5%	1%
(4) AR(1) process with the parameter $\alpha = 0.3$ and the distribution of ε_t												
	5%		100	100	99.7	99.4	97.5	65.1	42.3	24.1	16.5	7.3
$N(0,1)$	10%		99.9	99.6	98.4	96.8	92.0	59.5	37.4	20.8	14.1	5.9
	30%		99.3	97.2	93.1	89.3	78.7	53.3	34.5	21.0	15.0	7.4
	5%		99.9	99.7	99.4	98.9	96.4	32.2	22.6	15.9	12.5	8.7
$t(3)$	10%		99.9	99.3	98.1	96.4	89.6	26.8	19.2	13.4	11.0	7.1
	30%		97.1	94.0	88.9	84.8	73.7	21.1	14.2	09.2	6.8	3.6
	5%		99.5	99.3	98.9	98.5	95.7	8.2	7.2	6.2	5.7	4.8
$t(1)$	10%		88.4	67.2	56.2	50.1	39.9	9.7	8.3	7.2	6.6	5.6
	30%		44.3	38.7	33.9	30.4	24.2	6.1	4.0	2.5	1.8	0.8
	5%		100	100	99.9	99.7	97.4	27.4	21.4	16.7	14.7	10.9
$L(1)$	10%		99.9	99.4	98.2	96.0	87.4	24.7	18.3	14.1	11.5	8.2
	30%		95.8	91.9	85.0	79.9	66.9	16.6	11.3	7.0	5.2	2.6
	5%		99.9	99.9	99.8	99.8	97.4	13.4	11.2	9.5	8.6	7.1
$L(2)$	10%		97.2	87.2	77.7	70.8	58.5	14.5	12.5	10.6	9.6	7.8
	30%		64.2	57.0	49.1	44.0	34.7	8.7	5.9	3.5	2.7	1.3
(5) GARCH(1,1) process with the parameters $(\beta_o, \beta_1, \beta_2)$ and the distribution of ε_t												
(1,0.3,0.6)	5%		96.9	93.2	89.2	85.9	78.6	99.8	99.6	99.5	99.3	99.0
$N(0,1)$	10%		93.6	87.5	81.1	76.8	68.4	99.1	98.7	98.5	98.3	97.8
	30%		69.4	58.9	50.7	45.6	35.9	95.3	94.6	93.7	93.0	91.9
(1,0.05,0.9)	5%		83.4	67.2	51.0	41.3	26.5	95.6	92.7	89.3	87.0	81.8
$N(0,1)$	10%		81.4	65.6	50.1	41.4	28.4	93.0	89.1	84.5	81.7	76.6
	30%		68.5	52.6	38.8	31.3	19.5	83.4	77.4	71.7	67.7	61.2
(1,0.3,0.6)	5%		99.7	99.3	98.6	97.4	95.3	98.9	98.6	98.4	98.1	97.8
$t(3)$	10%		98.9	97.8	96.4	95.1	92.3	97.7	97.2	96.7	96.3	95.5
	30%		81.3	75.8	69.2	65.1	57.2	90.9	89.7	88.2	87.5	85.9
(1,0.05,0.9)	5%		95.9	90.6	83.1	77.2	66.6	97.8	96.9	95.7	95.2	93.3
$t(3)$	10%		96.4	91.6	85.6	81.6	71.8	96.5	95.1	94.0	93.3	91.6
	30%		85.3	77.8	69.9	64.3	54.0	89.4	87.5	85.4	83.7	80.8
(1,0.3,0.6)	5%		79.5	64.4	49.2	41.7	30.1	62.3	55.1	49.4	45.8	40.4
$L(1)$	10%		67.6	51.5	38.1	31.3	21.3	50.9	44.3	38.9	36.2	30.6
	30%		44.6	30.5	20.7	15.9	8.9	31.1	25.1	19.9	16.9	12.1
(1,0.05,0.9)	5%		74.7	56.9	40.5	32.4	21.0	46.3	39.2	34.3	31.4	26.4
$L(1)$	10%		68.0	50.0	34.8	27.7	17.8	42.1	35.1	29.5	25.9	21.0
	30%		47.0	33.1	21.5	16.1	8.9	22.7	16.6	11.3	8.8	5.8

Note: The entries are rejection frequencies in percentages; θ denotes the nominal size; $T = 500$.