# Bootstrapping the Li-Mak and McLeod-Li Portmanteau Tests for GARCH Models

Gul Nisa [1*] • Farhat Iqbal [2]

[1] Balochistan University of Information Technology, Engineering and Management Sciences, Quetta, Pakistan
[2] Department of Statistics, University of Balochistan, Quetta, Pakistan

farhatiqb@gmail.com

**Abstract:** In this paper, blocks-of-blocks (BOB) bootstrap method is employed for the commonly used diagnostic tests for generalized autoregressive conditional heteroscedastic (GARCH) models. More specifically, the single block-of-blocks and double blocks-of-blocks bootstrap techniques, using three different block lengths of size 4, 10, and 20, are implemented for bootstrapping the Li-Mak and Mcleod-Li portmanteau tests. Using Monte Carlo simulations, the size and power of both tests under the standard normal and Student-*t* errors are investigated. It was found that the discrepancy between the true and nominal probability of rejection was reduced for both the tests using single block-of-blocks and double blocks-of-blocks bootstrap methods. The power of the Li-Mak test for the GARCH (1, 1) model was found slightly better than the Mcleod-Li test. An empirical example using the monthly data of currency exchange rate (US $ per Pak Rupees) is also reported.

**To cite this article**
[Nisa, G., & Iqbal, F. (2018). Bootstrapping the Li-Mak and McLeod-Li Portmanteau Tests for GARCH Models. *The Journal of Middle East and North Africa Sciences*, 4(01), 32-38]. (P-ISSN 2412- 9763) - (e-ISSN 2412-8937). www.jomenas.org. **4**

**Keywords**: Blocks-of-blocks bootstrap; GARCH; Portmanteau tests; bootstrapped *p*-values

## 1. Introduction:

In the classical linear regression model, one of the common assumptions is that the residuals of the estimated regression line are stochastically independent from each other. To determine the adequacy of the fitted model, various diagnostic tests are used based on the autocorrelation function (ACF) of the residuals. Box and Pierce (1970) developed one of the most commonly used portmanteau tests. Ljung and Box (1978) proposed a modification of the Box-Pierce test. The Ljung-Box test gives good approximations and is adequate for many practical purposes.

Granger and Andersen (1978) suggested that for autoregressive moving average (ARMA) models if the statistical dependence in the residuals was found to be non-linear than squared residual autocorrelations may be useful. It was observed from some of the time series models that squared values of residuals are highly correlated as compared to the residuals itself. McLeod and Li (1983) analyzed autocorrelation of squared residuals of ARMA models. They developed a test based on the square residual autocorrelations. For large sample size, the autocorrelation of squared residuals is asymptotically normally distributed with a mean of zero and unit covariance matrix.

One of the assumptions of an econometric model is that the model has constant forecast variance. Engle (1982) developed autoregressive conditional heteroscedastic (ARCH) a process which allows the conditional variance varying with time.

The conditional variance, also known as volatility, in ARCH models is a function of the past squared errors. Bollerslev (1986) introduced the generalized ARCH (GARCH) model in which the current conditional variance equation also includes the past conditional variance as explanatory variables. Li and Mak (1994) developed the portmanteau statistic which depends on the squared standardized residuals autocorrelations. This test is considered useful for the diagnostic testing of nonlinear time series with conditional heteroscedasticity.

Bootstrap is a technique that can be used for determining the distribution of test statistic or an estimator by resampling the available data. The bootstrap work as an alternative method in that situation in which the asymptotic distributions are difficult to obtain. For finite samples, the first-order asymptotic theory often does not gives an accurate approximation to the distributions of test statistic as the bootstrap method. As a consequence, the test depends on the asymptotic critical values can lead to the nominal levels very different from the true levels (Horowitz; 1997).

A simulation study conducted by Chen (2002) indicated that for heavily tailed data the size and power performance of the Ljung-Box and Mcleod-Li tests are not robust. It was concluded that for serial correlations and volatility clustering effects other portmanteau tests which are robust to nonlinearity, conditional heteroscedastic higher order moments and distributional heavily tailed are required to be developed. Tsui (2004) analyzed the empirical size and power of Wooldridge (1991), Li-Mak (1994) and Tse (2002) diagnostic tests for the univariate conditional heteroscedastic model. It was observed that Tse and Li-Mak tests are powerful diagnostic tests for univariate conditional heteroscedastic models.

Horowitz et al. (2006) used bootstrapping methods to test the hypothesis that in the existence of statistical dependence the first *K* autocorrelations are zero of a covariance stationary time series. The *p*-values were obtained for Box-Pierce test from blocks-of-blocks bootstrap methods with pre-whitening. It was observed that the double blocks-of-blocks bootstrap reduced the difference much more between the true and nominal probability as compared to single block-of-blocks bootstrap.

Iqbal et al. (2013) analyzed the size and power of Ljung-Box and Li-Mak diagnostic tests using univariate autoregressive conditional heteroscedastic models for symmetric and asymmetric errors. When the distribution of the errors was asymmetrical the performance of the Li-Mak test was observed better than the Ljung-Box test. For asymmetrical heavily-tailed data, the empirical power of the Li-Mak test was found better.

For weak ARMA models, Zhu (2016) employed the random weighting procedure to bootstrap the critical values of Ljung-Box, Monti (1994), weighted Ljung-Box and weighted Monti portmanteau tests. These four tests were also implemented to investigate the adequacy of power-GARCH models. The outcomes from simulation conducted revealed that critical values of weighted Ljung-Box and weighted Monti test have higher power than un-weighted portmanteau tests.

The aim of this article is to develop a bootstrap method for obtaining the rejection probabilities for Li-Mak and Mcleod-Li diagnostic tests that are commonly used for the diagnosis of GARCH models. The single block-of-blocks (SBOB) bootstrap and double blocks-of-blocks (DBOB) bootstrap techniques are used with three block lengths of size 4, 10 and 20. Monte Carlo simulations are used to examine the true rejection probabilities of both the tests with the *p*-values of BOB bootstrap. Data were generated using the standard normal and Student-*t* distributions. Our results indicate that the *p*-values obtained, from the bootstrapping of both the tests, provide the close approximation to the true level of rejection than the asymptotic *p*-values. The DBOB bootstrap reduced the distinction between the true level of rejection and nominal probabilities more than SBOB bootstrap method.

Therefore, it is recommended that a close approximation to the true level of rejection is possible with high accuracy by using the bootstrap method.

This paper is arranged as follows: The GARCH model and diagnostic tests are briefly introduced in Section 2. In Section 3, the SBOB and DBOB bootstrap methods for Li-Mak and Mcleod-Li diagnostic tests are described. Results of simulations are presented and discussed in Section 4. Section 5 provides an empirical example and finally, Section 6 concludes the article.

## 2. GARCH Models and Portmanteau Tests:

For the modeling of asset returns, it is known that the residuals of the estimated models are no longer homoscedastic and their variances vary over time. The ARCH process is generally used to reflect this type of effect. The GARCH model developed by Bollerslev (1986) is the generalized form of the ARCH model. The process $\{Y_t; t \in Z\}$ is considered. Observe $\{Y_t; 1 \leq t \leq n\}$ such that

$$Y_t = \sigma_t e_t$$

$$\sigma_t^2 = \alpha_o + \sum_{i=1}^{p} \alpha_i Y_{t-i}^2 + \sum_{j=1}^{q} \beta_j \sigma_{t-j}^2$$

Where $\alpha_0 > 0$, $\alpha_i \geq 0$, $i$=1,2,…,$p$, $\beta_j \geq 0$, $j$=1,2,…,$q$and $e_t$ are i.i.d white noise error terms. The conditions $\alpha_0 > 0$, $\alpha_i \geq 0$ and $\beta_j \geq 0$ are necessary for the conditional variance to be positive. The GARCH process is said to stationary if $\sum \alpha_i + \sum \beta_j < 1$.The GARCH (1, 1) model is commonly found adequate in financial applications.

To test the assumption that no autocorrelation is present between the white noise residuals of the fitted models. The portmanteau tests are used to test the null hypothesis of no autocorrelation. The portmanteau tests are distributed asymptotically as chi-square with a degree of freedom equal to the number of autocorrelation coefficient minus the number of parameters to be estimated in the equation.

Conducting different diagnostic tests is an essential step in time series model building. When a model is fitted, diagnostic tests are used to evaluate the model which then serves as a test of model adequacy. To test the null hypothesi $H_K: \rho(1) = \cdots = \rho(K) = 0, K = 1, 5, 10$ where $\rho$ is autocorrelation. The Mcleod-Li (1983) and Li-Mak (1994) tests are commonly used to test this type of hypothesis in GARCH-type models. The Mcleod-Li test is written as

$$Q_{ML}(K) = n(n+2) \sum_{k=1}^{K} \frac{\hat{r}_{1k}^2}{n-k}$$

Where the lag-*K* squared residual autocorrelation is given by

$$\hat{r}_{1k} = \frac{\sum_{t=k+1}^{n} (\hat{e}_t^2 - \bar{e})(\hat{e}_{t-k}^2 - \bar{e})}{\sum_{t=1}^{n} (\hat{e}_t^2 - \bar{e})^2} \text{for } k = 1,2,\dots,K$$

Where $\bar{e} = \frac{1}{n} \sum \hat{e}_t^2$ and $n$ is the sample size. For large $n$ the autocorrelations of squared residuals are asymptotically normally distributed with zero mean and unit covariance matrix. The asymptotic distribution of Mcleod-Li portmanteau test is chi-square with degrees of freedom $K$ when the eight-order moment exists. The Li-Mak test for GARCH models is defined as

$$Q_{LM}(K) = n\sum_{k=1}^{K}\hat{r}_{2k}^2$$

The standardized square residuals autocorrelation at lag-$K$ is defined as

$$\hat{r}_{2k} = \frac{\sum_{t=k+1}^{n}(\hat{e}_t^2-1)(\hat{e}_{t-k}^2-1)}{\sum_{t=1}^{n}(\hat{e}_t^2-1)^2} \text{ for } k = 1,2,\dots,K$$

Where $\hat{e}_t$ is standardized residuals obtained from GARCH models. The Li-Mak test is asymptotically distributed as chi-square distribution with $K$-$p$-$q$ degree of freedom.

## 3. Bootstrapping Portmanteau Tests:

Bootstrap is an important technique developed by Efron (1979). The bootstrap method worked as Monte Carlo method to simulate from the available data without making an assumption. The basic purpose of bootstrap testing is that it is useful to characterize the distribution of test statistics whose distribution is unknown under the null hypothesis. Through simulation, many artificial data sets are generated which are called bootstrap samples and for each sample, the statistic is calculated. The simulation-based estimate for the unknown distribution is then obtained by using the empirical distribution function (EDF) of these bootstrap statistics.

Efron (1979) bootstrap method is applicable to observations which are independent identically distributed. In time series data the observations are often dependent and the usual methods of bootstrap fail to provide consistent estimates because all the information related to dependence nature of observations is lost (Singh, 1981). Hall (1985) Introduced the block bootstrap technique, the overlapping block bootstrap developed by Kunsch (1989) and the no-overlapping block bootstrap proposed by Carlstein (1986) are mostly used as bootstrapping technique for time series data to maintain the dependence structure of the observations within a block.

The block bootstrap is a technique used to produce the bootstrap samples in those situations when a parametric model is not available. The blocking is a procedure in which data is divided into blocks and sampling is done randomly from these blocks with replacement. In the presence of statistical dependence for the testing of individual autocorrelation coefficients, Romano and Thombs (1996) produced robust inferences using the block bootstrap method. The block bootstrap method provides the closest approximation to the test statistic distribution under analysis. For asymptotically pivotal test statistic the block bootstrap procedure gives more precise approximation than the first order asymptotic theory (Hall et al., 1995; Hall and Horowitz, 1996; Andrews, 2002).

### 3.1 Single bootstrap:

To apply the single block-of-blocks (BOB) bootstrap a new matrix $(\varepsilon_1,\varepsilon_2,\dots,\varepsilon_{n-K})$ having order $(K+1)\times(n-K)$ is defined where $\varepsilon_i = (\hat{e}_i,\hat{e}_{i+1},\dots,\hat{e}_{i+k})' = (\hat{e}_i^1,\hat{e}_i^2,\dots,\hat{e}_i^{K+1})'$, $K$ is lag length and $e_i$ are residuals. Without the use of pre-whitening, the bootstrap sample of size $n$ from the blocks is produced by resampling blocks from the $K$+1 dimensional series. Suppose $b$ represents the block size and let $h=n/b$. let $B_i$ be a matrix of order $(K+1)\times b$ written as $B_i = \varepsilon_i,\dots,\varepsilon_{i+b-1}$, where $i = 1,2,\dots,z$ and $z = n-b-K+1$. Using the following three steps the SBOB bootstrap test is derived.

1. The number of blocks is randomly selected $h$ times with replacement from the collection $\{B_1,\dots,B_z\}$. These make a set of blocks $B_1^*,\dots,B_h^*$. The blocks selected are arranged end-to-end to make a matrix of order $(K+1)\times n$ known as bootstrap sample and is represented by $\varepsilon^* = (\varepsilon_1^*,\dots,\varepsilon_n^*)$ and the bootstrap replicate of $\varepsilon_i$ is $\varepsilon_i^* = (\hat{e}_i^{1*},\hat{e}_i^{2*},\dots,\hat{e}_i^{(K+1)*})'$.
2. Compute the tests $Q_{ML}^S(K) = n(n+2)\left[\sum_{k=1}^{K}\frac{(\hat{r}_{1k}^*-\hat{r}_{1bk})^2}{n-k}\right]$ and $Q_{LM}^S(K) = n[\sum_{k=1}^{K}(\hat{r}_{2k}^*-\hat{r}_{2bk})^2]$ for Mcleod-Li and Li-Mak tests, respectively, from the bootstrap samples. Since overlapping blocks are used, in the collection $\{B_i,\dots,B_z\}$, some measurements received more weight than others. Therefore, both the statistics $Q_{ML}^S(K)$ and $Q_{LM}^S(K)$ are centered using the estimator $\hat{r}_{1bk}$ and $\hat{r}_{2bk}$ respectively.

$$\hat{r}_{1k}^* = \frac{\sum_{t=1}^{n}\{(\hat{e}_t^2)^{1*}-(\overline{e})^{1*}\}\{(\hat{e}_t^2)^{(k+1)*}-(\overline{e})^{(k+1)*}\}}{[\sum_{t=1}^{n}\{(\hat{e}_t^2)^{1*}-(\overline{e})^{1*}\}^2\sum_{t=1}^{n}\{(\hat{e}_t^2)^{(k+1)*}-(\overline{e})^{(k+1)*}\}^2]^{1/2}}$$

Where $\overline{e}^{j*} = 1/n\sum_{t=1}^{n}(\hat{e}_t^2)^{j*}$

$$\hat{r}_{1bk} = \frac{\sum_{t=1}^{n-K}w_t\{(\hat{e}_t^2)^1-(\overline{e})^1\}\{(\hat{e}_t^2)^{(k+1)}-(\overline{e})^{(k+1)}\}}{\left[\sum_{t=1}^{n-k}w_t\{(\hat{e}_t^2)^1-(\overline{e})^1\}^2\sum_{t=1}^{n-k}w_t\{(\hat{e}_t^2)^{(k+1)}-(\overline{e})^{(k+1)}\}^2\right]^{1/2}}$$

With $\overline{e}^k = \sum_{t=1}^{n-K}w_t\hat{e}_t^k$

$$\hat{r}_{2k}^* = \frac{\sum_{t=1}^{n}\{(\hat{e}_t^2)^{1*}-1\}\{(\hat{e}_t^2)^{(k+1)*}-1\}}{[\sum_{t=1}^{n}\{(\hat{e}_t^2)^{1*}-1\}^2\sum_{t=1}^{n}\{(\hat{e}_t^2)^{(k+1)*}-1\}^2]^{1/2}}$$

$$\hat{r}_{2bk} = \frac{\sum_{t=1}^{n-K}w_t\{(\hat{e}_t^2)^1-1\}\{(\hat{e}_t^2)^{(k+1)}-1\}}{[\sum_{t=1}^{n-k}w_t\{(\hat{e}_t^2)^1-1\}^2\sum_{t=1}^{n-k}w_t\{(\hat{e}_t^2)^{(k+1)}-1\}^2]^{1/2}}$$

Where,

$$w_t = t/b(n-K-b+1),\quad t = 1,\dots,b-1$$
$$= 1/(n-K-b+1),\quad t = b,\dots,n-K-b+1$$
$$= (n-K+1-t)/b(n-K-b+1),\quad t = n-K-b+2,\dots,n-K.$$

3. Step 1 and 2 are performed $M_1$ times.

From single bootstrap, the distribution of $Q_{LM}$(K) and $Q_{ML}(K)$ is estimated by the empirical distribution of the $M_1$ values of $Q_{LM}^S(K)$ and $Q_{ML}^S(K)$ respectively. The single BOB bootstrap estimate of *p*-value is $p_k^*$ where $p_k^* = \#\left(Q_{LM}^S(K) > Q_{LM}(K)\right)/M_1$ the number of $Q_{LM}^S(K)$ greater than $Q_{LM}(K)$ divided by $M_1$ and the *p*-values for Mcleod-Li are computed using the same formula. The SBOB bootstrap test rejects $H_K$ if $p_k^* < \alpha$ under the given nominal level of $\alpha$.

The test based on bootstrap *p*-values $p_k^*$ has probability of rejection $\alpha$ if $P\ (p_k^* < \alpha|H_K) = \alpha$, when the $p_k^*$ distribution is uniform on [0, 1]. If the distribution is found to be not uniform then this will indicates the presence of some $\beta$ such that $P\ (p_k^* < \beta|H_K) = F_{p^*}(\beta) = \alpha$. The unknown $\beta$ is the inverse of empirical distribution $F_{p^*}$ obtained at $\alpha$, $\beta = {F_{p^*}}^{-1}(\alpha)$. When the estimate of $F_{p^*}$ is available both $\beta$ and the error in the *p*-values can be estimated. The $F_{p^*}$ and $\beta$ can be estimated by using the double bootstrap method.

### 3.2 Double bootstrap:

The size *n* double bootstrap sample from the blocks is produced by resampling blocks from a bootstrap sample$\varepsilon_1^*, \ldots, \varepsilon_n^*$. Here also, the block size is *b*, where *n=hb*. Suppose $B_i^*$ represents the block of *b* successive observations starting with $\varepsilon_i^*$ that is, $B_i^* = \varepsilon_i^*, \ldots, \varepsilon_{i+b-1}^*$, where $i = 1,2,\ldots,z$ and $z = n - b - K + 1$. The following steps described the DBOB bootstrap test.

Perform the above step (1) and (2).

$1'$. For each single bootstrap sample, the number of blocks is randomly selected *h* times with replacement from the collection $\{\ B_1^*, \ldots, B_z^*\ \}$. These make a collection of blocks$\{B_1^{**}, \ldots, B_h^{**}\}$. The blocks selected are then put end-to-end to make a time series of length *n*, which is known as double bootstrap sample $\varepsilon^{**} = (\varepsilon_1^{**}, \ldots, \varepsilon_n^{**})$ where $\varepsilon_i^{**} = (e_i^{1**}, e_i^{2**}, \ldots, e_i^{(k+1)**})'$.

$2'$. Compute the statistic $Q_{ML}^D(K)\ and\ Q_{LM}^D(K)$ from double bootstrap sample

$$Q_{ML}^D(K) = n(n+2)\left[\sum_{k=1}^{K} \frac{(\hat{r}_{1k}^{**} - \hat{r}_{1bk}^{*})^2}{n-k}\right]$$

Where,

$$\hat{r}_{1k}^{**} = \frac{\sum_{t=1}^{n}\{(\hat{e}_t^2)^{1**} - (\overline{e})^{1**}\}\{(\hat{e}_t^2)^{(k+1)**} - (\overline{e})^{(k+1)**}\}}{[\sum_{t=1}^{n}\{(\hat{e}_t^2)^{1**} - (\overline{e})^{1**}\}^2 \sum_{t=1}^{n}\{(\hat{e}_t^2)^{(k+1)**} - (\overline{e})^{(k+1)**}\}^2]^{1/2}}$$

and $\overline{e}^{j**} = 1/n\sum_{t=1}^{n}(\hat{e}_t^2)^{j**}$

$$Q_{LM}^D(K) = n\sum_{k=1}^{K}(\hat{r}_{2k}^{**} - \hat{r}_{2bk}^{*})^2$$

Where,

$$\hat{r}_{2k}^{**} = \frac{\sum_{t=1}^{n}\{(\hat{e}_t^2)^{1**} - 1\}\{(\hat{e}_t^2)^{(k+1)**} - 1\}}{[\sum_{t=1}^{n}\{(\hat{e}_t^2)^{1**} - 1\}^2 \sum_{t=1}^{n}\{(\hat{e}_t^2)^{(k+1)**} - 1\}^2]^{1/2}}$$

The procedure for calculating $\hat{r}_{1bk}^*$ and $\hat{r}_{2bk}^*$ is similar to $\hat{r}_{1bk}$ and $\hat{r}_{2bk}$ of the SBOB bootstrap sample respectively.

$3'$. Steps $1'$ and $2'$are performed $M_2$ times.

$4'$. Steps 1, 2 and $3'$ are performed $M_1$ times.

There are $M_2$ values of $Q_{LM}^D(K)$ and $Q_{ML}^D$(K) test statistics for every $M_1$ single bootstrap samples. Therefore the *p*-values of double bootstrap will be $M_1$ represented by $p_k^{**}$ where $p_k^{**} = \#\left(Q_{LM}^D(K) > Q_{LM}^S(K)\right)/M_2$ and similar formula of *p*-values is used for Mcleod-Li test. $F_{p^{**}}$ is used to represent the empirical distribution function of $M_1$ *p*-values and it provides an estimate of $F_{p^*}$. $\beta^* = {F_{p^{**}}}^{-1}(\alpha)$ is an estimate of $\beta$. If $p_k^* < \beta^*$, then double BOB test rejects the hypothesis that is the DBOB bootstrap test rejects if $p_{ka}^* = F_{p^{**}}(p_k^*) < \alpha$ where $p_{ka}^*$ is the adjusted *p*-values (Davison and Hinkley, 1997). The formula to estimate the adjusted *p*-values is $p_{ka}^* = \#\,(p_k^{**} \le p_k^*)/M_1$ given by Hinkley (1989).

## 4. Results and Discussion:

In this section, we derived the probabilities of rejection for the Li-Mak and Mcleod-Li tests used to test the hypothesis $H_K\colon \rho(1) = \cdots = \rho(K) = 0, K = 1, 5, 10$ for the GARCH (1, 1) model adequacy. Single BOB and double BOB bootstrap techniques are used with sample size of *n* = 300 with block lengths *b* = 4, 10, and 20. We use 1000 independent replications with $M_1$ = 299 and $M_2$ = 200. For the asymptotic *p*-values we use 25000 replications. All the computations are performed using R software.

The data are generated from the GARCH (1, 1) model $\sigma_t^2 = 0.001 + 0.05Y_{t-1}^2 + 0.90\ \sigma_{t-1}^2$. For these parameters values the condition that the fourth moment of $Y_t$exist is satisfied (He and Teräsvirta, 1999). The errors are generated from the standard normal and Student-*t* distributions with 3 degree of freedom.

Table 1 and Table 2 show the rejection probabilities in percentage for both Mcleod-Li and Li-Mak tests, respectively. The asymptotic *p*-values of both the tests over reject all three hypotheses at lag 1, 5 and 10. At lag 10 the largest over rejection occurred of both the tests for the GARCH (1, 1) model with normal errors. The asymptotic *p*-values of Li-Mak test under reject the null hypothesis at the lag length 1, 5 and 10 with student-*t* errors.

The difference between the nominal and the empirical probabilities of rejection was reduced by DBOB bootstrap method for both the tests at lag length 1, 5 and 10 under both types of error distributions. Horowitz et al. (2006) also found that using the blocks-of-blocks bootstrap method for bootstrapping of the Box-Pierce (1970) test reduced the difference between the nominal and true probabilities of rejection more than the asymptotic *p*-values. Our study is different from Horowitz et al. (2006)

in the sense that we used the Li-Mak (1994) and Mcleod-Li (1983) tests. The results observed are independent to the selection of the block lengths. Davison and Hinkley (1997) also found that the blocks-of-blocks bootstrap is not influenced by the selection of block length.

Next, we investigate the power of both the tests. For the power of the tests, the following two data generating processes are used.The GARCH (1, 1) model denoted by A1 and the ARCH (2) model denoted by A2.

A1: $Y_t = \sigma_t e_t,\quad \sigma_t^2 = 0.5 + 0.3Y_{t-1}^2 + 0.2\sigma_{t-1}^2$
A2: $Y_t = \sigma_t e_t,\quad \sigma_t^2 = 0.1 + 0.8Y_{t-1}^2 + 0.1Y_{t-1}^2$

The power is calculated by fitting the GARCH (1, 1) model to A2 and fitting the ARCH (2) model to A1.Table 3 shows the empirical powers of DBOB $Q_{LM}$ and DBOB $Q_{ML}$ tests under the normal and Student-*t* errors. The results show that both the tests have low powers at lag 1 and the power of tests increase with lag lengths. The power of both the tests with *t* (3) errors was found to be low as compared to the tests with N (0, 1) errors. For testing that the fitted GARCH (1, 1) model is adequate the Li-Mak test is observed more powerful than the Mcleod-Li test using all the error distributions examined in this study. For univariate conditional heteroscedastic model Tsui (2004) also analyzed the empirical size and power of Wooldridge (1991), Li-Mak (1994) and Tse (2002) diagnostic tests and it was observed that Tse and Li-Mak test are more powerful diagnostic tests for univariate conditional heteroscedastic models.

## 5. Empirical example:

In this section, we derived the rejection probabilities for the Li-Mak and Mcleod-Li portmanteau tests using an empirical example. The data set for our analysis includes the monthly data of foreign exchange rate (US $ per Pak Rupees) from July 1981 through April 2013 (the total number of observations recorded is 361).

The dataset was taken from the website (http://www.sbp.org.pk/stats/survey/index.asp).

The asymptotic *p*-values of Li-Mak and Mcleod-Li tests for testing the hypothesis at lag 5 using the nominal rejection probability of 5% are 0.068 and 0.076 respectively both the tests over reject the hypothesis and high difference is observed between the nominal rejection probability and the true rejection probability. For testing the hypothesis at lag 5 using 5% nominal rejection probability the SBOB (DBOB) bootstrap *p*-values of Li-Mak test with the block lengths $b$ = 4, 10 and 20 are 0.049(0.051), 0.052(0.050), 0.057(0.051) and SBOB (DBOB) bootstrap *p*-values of Mcleod-Li test with the block lengths $b$ = 4, 10 and 20 are 0.052(0.048), 0.055(0.051), 0.060(0.051). The result reveals that the difference between the true and nominal significance level is reduced by both the tests using the blocks-of-blocks bootstrap technique as compared to the asymptotic *p*-values of the tests.

## 6. Conclusion:

In this study, we employed blocks-of-blocks bootstrap methods for two commonly used diagnostic tests for GARCH models. Both the single block-of-blocks and double blocks-of-blocks bootstrap method was implemented using the three different block lengths of size 4, 10 and 20 for the bootstrapping of Li-Mak and Mcleod-Li tests. Using Monte Carlo simulations the power of both the tests under the two types of error distributions standard normal and student-*t* was investigated. It was observed that the difference between the true and nominal probability of rejection was eliminated by the BOB bootstrap method than the asymptotic *p*-values of the tests. The power of the Li-Mak test for the GARCH (1, 1) model was observed moderately better than Mcleod-Li test. We conclude that the tests using the *p*-values obtained from the bootstrap method provided that the method of bootstrapping is accurately implemented give us the closest approximation to the true level of rejection as compared to the asymptotic *p*-value.

Table 1: *Percentage of rejection probabilities of Mcleod-Li test for the GARCH (1, 1) model with n = 300*

| Tests | H | | | $H_5$ | | | $H_{10}$ | | |
|---|---|---|---|---|---|---|---|---|---|
| | 1 | 5 | 10 | 1 | 5 | 10 | 1 | 5 | 10 |
| Normal distribution | | | | | | | | | |
| $Q_{ML}$ | 1.6 | 7.8 | 13.5 | 1.8 | 7.4 | 12.9 | 2.3 | 8.1 | 14.5 |
| SBOB $Q_{ML}$ | | | | | | | | | |
| $b$ = 4 | 1.0 | 5.6 | 10.7 | 0.7 | 5.1 | 10.4 | 1.1 | 4.8 | 9.9 |
| $b$ = 10 | 1.3 | 6.4 | 11.5 | 1.1 | 5.6 | 11.0 | 1.1 | 5.4 | 11.1 |
| $b$ = 20 | 1.9 | 6.9 | 12.7 | 1.3 | 6.1 | 12.8 | 1.2 | 5.8 | 13.6 |
| DBOB $Q_{ML}$ | | | | | | | | | |
| $b$ = 4 | 0.9 | 5.1 | 10.4 | 0.6 | 4.7 | 9.2 | 0.8 | 4.0 | 8.1 |
| $b$ = 10 | 1.1 | 5.3 | 10.4 | 0.9 | 5.0 | 10.1 | 0.7 | 4.6 | 9.7 |
| $b$ = 20 | 1.0 | 5.6 | 10.8 | 1.1 | 5.2 | 10.9 | 0.8 | 4.7 | 10.2 |
| Student-*t* (3) distribution | | | | | | | | | |
| $Q_{ML}$ | 2.3 | 7.6 | 13.8 | 4.4 | 13.2 | 16.9 | 5.0 | 13.4 | 21.5 |
| SBOB $Q_{ML}$ | | | | | | | | | |
| $b$ = 4 | 1.5 | 6.7 | 11.7 | 0.9 | 5.6 | 11.5 | 1.2 | 5.2 | 10.1 |
| $b$ = 10 | 1.6 | 6.9 | 12.6 | 1.1 | 6.6 | 12.9 | 1.1 | 5.4 | 11.6 |
| $b$ = 20 | 2.2 | 7.4 | 14.4 | 1.6 | 7.2 | 14.6 | 1.3 | 6.3 | 13.4 |
| DBOB $Q_{ML}$ | | | | | | | | | |
| $b$ = 4 | 1.0 | 5.6 | 10.2 | 0.6 | 4.5 | 9.6 | 0.8 | 4.6 | 8.3 |
| $b$ = 10 | 1.0 | 5.5 | 10.6 | 0.7 | 5.3 | 10.9 | 0.8 | 4.8 | 9.4 |
| $b$ = 20 | 1.3 | 5.6 | 11.7 | 0.8 | 5.7 | 11.5 | 0.9 | 6.5 | 10.1 |

Notes: $H_1$, $H_5$ and $H_{10}$ the numbers given below represents the nominal rejection probabilities. $Q_{ML}$ is the Mcleod-Li test. Implementing the *p*-values of BOB bootstrap method the total replications are 1000 for the $Q_{ML}$ test. The total replications for the asymptotic $Q_{ML}$ test are 25000.

Table 2: *Percentage of rejection probabilities of Li-Mak test for the GARCH(1, 1) model with n = 300*

| Tests | $H_1$ | | | $H_5$ | | | $H_{10}$ | | |
|---|---|---|---|---|---|---|---|---|---|
| | 1 | 5 | 10 | 1 | 5 | 10 | 1 | 5 | 10 |
| Normal distribution | | | | | | | | | |
| $Q_{LM}$ | 1.6 | 6.6 | 12.3 | 1.8 | 7.5 | 13.2 | 1.9 | 7.3 | 12.8 |
| SBOB $Q_{LM}$ | | | | | | | | | |
| $b$ = 4 | 1.0 | 5.5 | 10.5 | 0.9 | 5.0 | 10.1 | 1.0 | 5.1 | 9.9 |
| $b$ = 1 | 1.2 | 5.9 | 10.9 | 1.6 | 5.3 | 6.5 | 1.0 | 5.5 | 10.9 |
| $b$ = 2 | 1.7 | 6.2 | 11.6 | 1.1 | 6.0 | 11.8 | 1.1 | 5.9 | 11.2 |
| DBOB $Q_{LM}$ | | | | | | | | | |
| $b$ = 4 | 0.9 | 5.1 | 10.2 | 0.8 | 4.5 | 9.5 | 0.9 | 4.2 | 10.1 |
| $b$ = 1 | 1.1 | 5.2 | 10.3 | 0.8 | 5.0 | 10.1 | 0.8 | 4.7 | 10.4 |
| $b$ = 2 | 2.5 | 7.0 | 10.8 | 1.0 | 5.1 | 10.9 | 0.9 | 5.5 | 11.5 |
| Student-*t* (3) distribution | | | | | | | | | |
| $Q_{LM}$ | 0.8 | 2.2 | 7.3 | 0.8 | 3.4 | 7.8 | 0.8 | 4.3 | 8.2 |
| SBOB $Q_{LM}$ | | | | | | | | | |
| $b$ = 4 | 1.3 | 6.0 | 10.9 | 0.8 | 5.5 | 10.4 | 1.1 | 4.9 | 10.3 |
| $b$ = 1 | 1.5 | 6.2 | 11.1 | 0.9 | 5.7 | 10.9 | 1.2 | 4.9 | 10.9 |
| $b$ = 2 | 1.9 | 6.8 | 12.1 | 1.4 | 6.9 | 11.2 | 1.3 | 5.6 | 12.9 |
| DBOB $Q_{LM}$ | | | | | | | | | |
| $b$ = 4 | 1.0 | 5.3 | 10.3 | 0.5 | 4.5 | 9.8 | 0.7 | 4.5 | 9.9 |
| $b$ = 1 | 1.1 | 5.4 | 10.4 | 0.6 | 5.1 | 10.4 | 0.8 | 4.7 | 10.3 |
| $b$ = 2 | 1.3 | 5.5 | 10.9 | 0.7 | 6.0 | 11.1 | 0.9 | 6.3 | 11.1 |

Notes: $H_1$, $H_5$ and $H_{10}$ the numbers given below represents the nominal rejection probabilities. $Q_{LM}$ is the Li-Mak test. Implementing the *p*-values of BOB bootstrap method the total replications are 1000 for the $Q_{LM}$ test. The total replications for the asymptotic $Q_{LM}$ test are 25000.

Table 3: *Empirical power (percent) of portmanteau tests using 0.05 nominal level*

| Error Distribution | DGP | EM | block size | $H_1$ $Q_{LM}$ | $H_1$ $Q_{ML}$ | $H_5$ $Q_{LM}$ | $H_5$ $Q_{ML}$ | $H_{10}$ $Q_{LM}$ | $H_{10}$ $Q_{ML}$ |
|---|---|---|---|---|---|---|---|---|---|
| N(0, 1) | A2 | GARCH (1, 1) | $b$=4 | 7.8 | 6.7 | 8.8 | 6.7 | 10.5 | 6.5 |
| | | | $b$=10 | 8.5 | 7.5 | 10.2 | 7.1 | 10.6 | 6.7 |
| | | | $b$=20 | 8.9 | 8.5 | 10.9 | 7.1 | 11.2 | 6.6 |
| | A1 | ARCH (2) | $b$=4 | 5.5 | 6.5 | 8.0 | 6.3 | 15.6 | 7.3 |
| | | | $b$=10 | 3.8 | 6.7 | 9.3 | 6.5 | 16.1 | 7.6 |
| | | | $b$=20 | 5.9 | 7.0 | 10.2 | 6.8 | 17.3 | 7.8 |
| t(3) | A2 | GARCH (1, 1) | $b$=4 | 7.8 | 6.2 | 8.1 | 6.8 | 9.2 | 6.3 |
| | | | $b$=10 | 8.3 | 6.5 | 9.2 | 6.9 | 10.4 | 6.5 |
| | | | $b$=20 | 8.3 | 6.8 | 11.4 | 6.2 | 10.8 | 9.1 |
| | A1 | ARCH (2) | $b$=4 | 4.8 | 6.4 | 8.0 | 6.1 | 12.1 | 7.1 |
| | | | $b$=10 | 5.5 | 6.5 | 8.8 | 7.1 | 13.1 | 7.3 |
| | | | $b$=20 | 5.6 | 6.0 | 9.6 | 6.8 | 17.1 | 7.7 |

Note: DGP: data generating process, EM: estimated model, $Q_{LM}$ and $Q_{ML}$ are Li-Mak and Mcleod-Li test, respectively.

**Corresponding Author:**
Farhat Iqbal, Ph.D.
Department of Statistics, University of Balochistan, Quetta, Pakistan.
E-mail: farhatiqb@gmail.com

**References:**

1. Andrews, D. W. (2002). Higher-order improvements of a computationally attractive k-step bootstrap for extremum estimators. *Econometrica, 70*(1), 119-162.
2. Bollerslev, T. (1986). Generalized autoregressive conditional heteroskedasticity. *Journal of Econometrics, 31*(3), 307-327.
3. Box, G. E., & Pierce, D. A. (1970). Distribution of residual autocorrelations in autoregressive-integrated moving average time series models. *Journal of the American Statistical Association, 65*(332), 1509-1526.
4. Chen, Y. T. (2002). On the robustness of Ljung-Box and McLeod-Li Q tests: a simulation study. *Economics Bulletin, 3*(17), 1-10.
5. Carlstein, E. (1986). The use of subseries values for estimating the variance of a general statistic from a stationary sequence. *The Annals of Statistics*, 1171-1179.
6. Davison, A. C., & Hinkley, D. V. (1997). *Bootstrap methods and their application* (Vol. 1): Cambridge University Press.
7. Engle, R. F. (1982). Autoregressive conditional heteroscedasticity with estimates of the variance of United Kingdom inflation. *Econometrica: Journal of the Econometric Society*, 987-1007.
8. Efron, B. (1979). Bootstrap methods: another look at the jackknife. *Annals of Statistics*, 7: 1–26.
9. Granger, C., & Anderson, A. (1978). An Introduction to Bilinear Time Series Models. Vandenhoeck and Ruprecht: Gottingen.
10. He, C., & Teräsvirta, T. (1999). Properties of moments of a family of GARCH processes. *Journal of Econometrics, 92*(1), 173-192.
11. Hinkley, D. V. (1989). Bootstrap significance tests, Proceedings of the 47th Session of the International Statistical Institute. 65-74.
12. Horowitz, J. L. (1997). Bootstrap methods in econometrics: theory and numerical performance. *Econometric Society Monographs, 28*, 188-222.
13. Horowitz, J. L., Lobato, I., Nankervis, J. C., & Savin, N. (2006). Bootstrapping the Box-Pierce Q test: a robust test of uncorrelatedness. *Journal of Econometrics, 133*(2), 841-862.
14. Hall, P., & Horowitz, J. L. (1996). Bootstrap critical values for tests based on generalized-method-of-moments estimators. *Econometrica: Journal of the Econometric Society*, 891-916.
15. Hall, P. (1985). Resampling a coverage pattern. *Stochastic Process and Their Application Journal*, 20(2), 231-246.
16. Hall, P., Horowitz, J. L., & Jing, B.-Y. (1995). On blocking rules for the bootstrap with dependent data. *Biometrika, 82*(3), 561-574.
17. Iqbal, F., Jafri, Y., & Talpur, G. (2013). Diagnostics for GARCH-Type Models under Symmetric and

Asymmetric Errors. *Sindh University Research Journal-SURJ (Science Series), 45*(3), 529-633.

18. Kunsch, H. R. (1989). The jackknife and the bootstrap for general stationary observations. *The Annals of Statistics*, 1217-1241.
19. Li, W., & Mak, T. (1994). On the squared residual autocorrelations in non-linear time series with conditional heteroskedasticity. *Journal of Time Series Analysis, 15*(6), 627-636.
20. Ljung, G. M., & Box, G. E. (1978). On a measure of lack of fit in time series models. *Biometrika*, 297-303.
21. McLeod, A. I., & Li, W. K. (1983). Diagnostic checking ARMA time series models using squared-residual autocorrelations. *Journal of Time Series Analysis, 4*(4), 269-273.
22. Monti, A. C. (1994). A proposal for a residuals autocorrelation test in linear models. *Biometrika*, 81, 776-780.
23. Romano, J. P., & Thombs, L. A. (1996). Inference for autocorrelations under weak assumptions. *Journal of the American Statistical Association, 91*(434), 590-600.
24. Singh, K. (1981). On the asymptotic accuracy of Efron's bootstrap. *The Annals of Statistics*, 1187-1195.
25. Tse, Y. K. (2002). Residual-based diagnostics for conditional heteroscedasticity models. *The Econometrics Journal, 5*(2), 358-374.
26. Tsui, A. K. (2004). Diagnostics for conditional heteroscedasticity models: some simulation results. *Mathematics and Computers in Simulation, 64*(1), 113-119.
27. Wooldridge, J . M. (1991). On the application of robust, regression-based diagnostic to models of conditionals means and conditionals variances. *Journal of Econometrics*, 47, 5-46.
28. Zhu, K. (2016). Bootstrapping the portmanteau tests in weak auto-regressive moving average models. *Journal of the Royal Statistical Society: Series B (Statistical Methodology), 78*(2), 463-485.

Received December 04, 2017; revised December 18, 2017; accepted December 28, 2017; published online January 01, 2018