QATAR UNIVERSITY

COLLEGE OF ARTS AND SCIENCES

PARAMETRIC AND NONPARAMETRIC PORTMANTEAU TESTS FOR LACK OF FIT IN TIME SERIES MODELS: A COMPARATIVE STUDY

BY

MUHAMMAD ARSHAD SHAIKH

A Thesis Submitted to

the College of Arts and Sciences

in Partial Fulfillment of the Requirements for the Degree of

Master of Arts in Applied Statistics

June 2022

COMMITTEE PAGE

The members of the Committee approve the Thesis of
Muhammad Arshad Shaikh defended on 10/05/2022

____

Dr. Esam Mahdi
Thesis/Dissertation Supervisor

____

Dr. Faiz Elfaki
Committee Member

____

Dr. Ayman Bakleezi
Committee Member

Approved:

____

Ahmed Elzatahry, Dean, College of Arts and Sciences

ABSTRACT

Shaikh, A., Muhammad, Masters, March: 2022, Applied Statistics

Title: Parametric and Nonparametric Portmanteau Tests for Lack of Fit in Time Series Models: A Comparative Study

Supervisor of Thesis: Esam Bashir Mahdi

Several diagnostic tests for the lack of fit time series models have been introduced using parametric and nonparametric portmanteau tests. Some tests have been proposed based on the asymptotic distributions. Others are based on the Bootstrapping and Monte-Carlo significance techniques. It has been shown that the Bootstrapping and Monte-Carlo tests are robust as they provide the correct size and tend to be more powerful than those based on the asymptotic distributions.

In this thesis, I conducted a comparison study of the size and power of some portmanteau tests commonly used in linear and nonlinear time series models. In particular, I considered the cases where the residuals follow Gaussian and non-Gaussian distribution under Autoregressive Moving Average (ARMA) and Generalized Autoregressive Heteroskedasticity (GARCH) models; where some parametric and nonparametric tests were applied based on the limiting distributions, Bootstrapping , and Monte-Carlo significance tests. The results show that the nonparametric Bootstrapping and Monte-Carlo significance tests provide the best performance comparing with tests based on the parametric asymptotic distribution. I applied the tests on a real application using the Qatar National Bank returns.

Keywords: ARMA models; Autocorrelation; Bootstrapping ; Cross-correlation; GARCH models; Monte-Carlo tests; Nonlinearity dependency; Portmanteau tests; Returns.

DEDICATION

I would like to dedicate my thesis to my wife, Sheema A. Shaikh, for her constant support, personally, morally and financially throughout my studies at Qatar University.

ACKNOWLEDGMENTS

I would like to thank My supervisor, Dr. Esam Mahdi for monitoring my progress very minutely yet very critically and patiently at every stage. I really appreciate his guiding and advising as and when required by me. I also would like to thank Professor Dr. M. Mazharul Islam, Professor Dr. Ayman Bakleezi and Professor Dr. Faiz Elfaki for their valuable feedbacks and expert suggestions on improving the thesis. I also would like to thank all the faculty members who taught me in the program. I have learned a lot from them, and they have broadened the horizons of my knowledge and cleared many concepts that were not clear enough before.

Table of Contents

LIST OF TABLES

LIST OF FIGURES

CHAPTER 1: INTRODUCTION

1.1 Background

Testing for the lack of fit in time-series models is essential to get an accurate forecasting. If the model is accurate, then the residuals should show no dependence structure. Usually this can be done by using portmanteau tests. Several parametric and non-parametric portmanteau tests have been developed to test for linear and nonlinear dependency in time series models.

Under the assumptions of ARMA models, Box and Pierce (1970) proposed the first portmanteau test to check the adequacy of ARMA model utilizing the square autocorrelations of the residuals obtained from a fitted model.

Ljung and Box (1978) improved the Box and Pierce (1970) by proposing a modified portmanteau test and they showed that the modified test has the same limiting distribution of the Box and Pierce (1970) test but it estimates the type I error more successfully with a higher power.

Monti (1994) proposed a portmanteau test based on the partial autocorrelations of the residuals. All the above three tests are asymptotically approximated by chi-square distribution.

Peňa and Rodríguez (2002) introduced a test based on the mth root of the determinant of the mth residual autocorrelations matrix. Their test statistic can be seen as a linear combination of independent chi-squared distributions. They approximated the distribution of their test by a gamma distribution. Peňa and Rodríguez (2002) showed that their test statistic can improve the estimate of the significant levels if the autocorrelation coefficients in the auto-correlation matrix are replaced with their standardized values. They also showed that their test statistic is more powerful than the competitors Box-pierece, Ljung-Box and Monti in detecting the inadequacy of ARMA models in many situations.

Peňa and Rodríguez (2006) modified the Peňa and Rodríguez (2002) test statistic by proposing another portmanteau test based on the log of the determinant of the same previously mentioned autocorrelation matrix. Their test statistic can be written as a weighted average based on partial autocorrelations. The larger weights were assigned to low order coefficients of their statistic test and smaller weights to high-order coefficients. Under the ARMA assumptions, Peňa and Rodríguez (2006) showed that the asymptotic distribution of the their test statistic can, also, be seen as a linear combination of independent chi-squared distributions, and they approximated it with gamma distribution as well as with normal distribution.

One problem pointed out by Lin and McLeod (2006) is that the sequence of the standardized residual autocorrelation is not always non-negative definite; hence, the test statistics proposed by Peňa and Rodríguez (2002, 2006) may not exist in many cases. To overcome this problem, Lin and McLeod (2006) proposed by using a Monte-Carlo significance test. They showed that the Monte-Carlo significance test provides a portmanteau test with the correct estimate of size and is almost always more powerful than the competitors statistics appearing in the literature.

Fisher and Gallagher (2012) introduced two weighted tests based on the trace of the autocorrelation matrix defined in Peňa and Rodríguez (2002, 2006). The first one can be considered as a weighted version of the Ljung-Box test, whereas, the second can be seen as a weighted version of the Monti test. They approximated the distribution of their statistics by Gamma. Their simulation study showed that the weighted test statistics have more powers than the competitors statistics appearing in the literature in many cases.

Anderson (1993) and Hong (1996[a], 1996[b]) showed that the normalized spectral density of the stationary process of residuals can be used to obtain a kernal portmanteau test without knowing the distribution of the innovation series. The only assumption is the iid with mean zero and finite fourth moment. Hong (1996[a], 1996[b]) showed that the distribution of their statistics can be asymptotically approximated by Gaussian and the power of the proposed tests are usually better than Ljung and Box (1978).

Gallagher and Fisher (2015) proposed three tests based on the kernel test idea of Hong (1996[a], 1996[b]) by considering three different weighting schemes.

Mahdi (2017) used the idea of Hong (1996[a], 1996[b]) to propose a kernel-based portmanteau test based on the autocorrelation matrix defined from Peňa and Rodríguez (2002). He showed that his test is asymptotically the same as the test statistic given by Peňa and

Rodríguez (2006).

All the above tests can be used to test for linearity in ARMA model and not designed to test for nonlinearity in other time series models including ARCH structure.

To detect nonlinearity dependency in time series models (including an ARCH structure), McLeod and Li (1983) introduced a portmanteau test based on the squared-residuals autocorrelations under the assumptions of the ARMA model. They showed that their test is asymptotically distributed as chi-squared.

Li and Mak (1994) proposed a portmanteau statistic based on the Autoregressive Heteroskedasticity (ARCH) assumptions. Their test statistic was proposed based on the standardized squared-residual autocorrelation obtained from a fitted ARCH model.

Peňa and Rodríguez (2002, 2006) replaced the residual autocorrelations in the Toplitz matrix by the squared-residual autocorrelations and used it to derive new portmanteau tests that can be used to detect nonlinearity in several time series.

Rodriguez and Ruiz (2005) proposed a portmanteau test for ARCH models using the information contained in the sample autocorrelations of nonlinear transformations of the underlying process. Their test statistic can be used to test for whether the autocorrelations of the residuals differ from zero and at the same time it can be used to test for possible relationship among successive autocorrelation coefficients.

Fisher and Gallagher (2012) showed that the proposed weighted test can be seen as a modified version of Li and Mak (1994) statistic that can be used to detect the nonlinearity presence in GARCH time series models.

Recently, Psaradakis and Vávra (2019) used the Lawrance and Lewis (1985, 1987) idea, which was based on using the generalized correlation of the residuals to detect nonlinear dependency in time series, and proposed four different statistics from stationary linear models to test for linearity. The generalized correlation of the residuals is used to measure the correlation between the residuals at different powers.

More recently, Mahdi and Fisher (2021) proposed a portmanteau test using the block matrix of autocorrelations and cross-correlations of residuals and squared-residual. Under the assumptions of ARMA(p, q) model, Mahdi and Fisher (2021) approximated the asymptotic distribution of C_m by Gamma. They showed that their C_m test can be seen as a linear combination of four weighted tests. Essentially, their test statistic generalized the test statistics proposed by Mahdi (2020b). The first and the second components of Mahdi and Fisher (2021) tests elaborate the partial autocorrelation of the residuals and the squared-residuals, respectively. The third and the fourth components elaborate the cross-correlation between the residuals and their squares and vice-versa, respectively. Hence, the C_m test can be used to detect, simultaneously, the linear and nonlinear dependency in stationary time series data. Their simulation study demonstrated that the C_m statistic tends to have higher power than the competitors' statistics appearing in the literature, particularly in detecting ARMA with GARCH errors and other nonlinear models. Mahdi and Fisher (2021) utilized the Randomly Weighted Bootstrap (RWB) approach which was proposed by Zhu (2016) to improve the size and the power of their statistic.

1.2 Thesis Layout

The thesis is divided into seven chapters.

The first chapter is Introduction which discusses a brief literature review about portmanteau tests commonly used in linear and nonlinear time series based on parametric and non parametric tests.

The second chapter discusses the test for adequacy of linear models. In this chapter, the Autoregressive Moving Average (ARMA) model is discussed followed by classical portmanteau tests, Kernel-based normalized spectral density portmanteau tests, Randomly Weighted Bootstrap (RWB), and Monte-Carlo portmanteau tests including the procedural details.

Chapter three focusses on test for nonlinearity. In this chapter the nonlinear models including Autoregressive Conditional Heteroskedasticity (ARCH) Generalized Autoregressive Conditional Heteroskedasticity (GARCH) models are discussed.

In Chapter four, I conducted a simulation study to evaluate emperical sizes and power of portmanteau tests under the Gaussian and non-Gaussian erros distributions.

An application is given in Chapter five and Chapter six I drew some conclusions and suggestion for the future work.

1.3 Research Objectives

The main objectives of this thesis are to:

- Review the most common portmanteau tests that are used in linear and nonlinear time series,
- perform a Monte-Carlo simulation to evaluate the portmanteau tests based on the asymptotic distribution and the bootstrap technique,
- follow up with the recent developments in the area of diagnostic checking of time series,
- implement the portmanteau tests on some real financial data.

CHAPTER 2: TEST FOR ADEQUACY OF ARMA MODEL

In this chapter I will study the common portmanteau tests that are used to test for linearity in ARMA models.

Definition 1 The Autoregressive Moving Average (ARMA) model of order (p, q) for n observations $z_1, z_2, \cdots, z_n$ of a stationary mean μ time series can be written

$$\phi_p(B)(z_t - \mu) = \theta_q(B)\varepsilon_t, \tag{1}$$

where

$$\begin{aligned} \phi_p(B) \quad &= 1 - \phi_1 B - \phi_2 B^2 - \cdots - \phi_p B^p, \\ \theta_p(B) \quad &= 1 + \theta_1 B + \theta_2 B^2 + \cdots + \theta_q B^q, \end{aligned}$$

where B is the backshift operator in t, so that $B^j(z_t) = z_{t-j}$. The polynomials $\phi_p(B)$ and $\theta_p(B)$ are assumed to have no common roots and all roots outside the unit circle. The innovation series $\{\varepsilon_t\}$ are assumed to be independent and identically distributed (i.i.d.) where $\mathrm{E}(\varepsilon_t) < \infty$ and $\mathrm{Var}(\varepsilon_t) = \sigma^2$.

Under the i.i.d. innovation assumption, the following null hypothesis should not be rejected:

$$\mathcal{H}_0: \rho_{11}(\pm 1) = \rho_{11}(\pm 2) = \cdots = \rho_{11}(\pm m) = 0,$$

where m is the maximum lag considered for significant autocorrelation and $\rho_{11}(k)$ is the correlation coefficient at lag k. Therefore, after fitting an ARMA (p, q) model to a series, we can estimate the residuals, $\hat{\varepsilon}_t$, by calculating the difference between the true value of z_t and the predicted one $\hat{z}_t$, for $t = 1, 2, \cdots, n$. If the model in (1) is correctly identified, then these residuals should, approximately, behave as the innovations behave i.e., the sample autocorrelation coefficients of the residuals, for all $k \neq 0$, should approximately equal to zero.

Definition 2 *The population generalized correlation coefficient at lag time k between the error term raised to the power r (ε_t^r) and the error term raised to the power s (ε_{t+k}^s), where r and s are natural numbers is given by*

$$\rho_{rs}(k) = \frac{\gamma_{rs}(k)}{\sqrt{\gamma_{rr}(0)}\sqrt{\gamma_{ss}(0)}}, \quad k = 0, \pm1, \pm2, \cdots, \pm m, \tag{2}$$

where $\gamma_{rs}(k)$ is the generalized covariance between ε_t^r and ε_{t+k}^s that is given by

$$\gamma_{rs}(k) = n^{-1}\sum_{t=1}^{n-k} f_r(\varepsilon_t)f_s(\varepsilon_{t+k}), \tag{3}$$

and $f_h(x_t) = x_t^h - n^{-1}\sum_{t=1}^{n} x_t^h$, for $h = 1, 2$.

Remarks:

- When $r = s = 1$, we have $\gamma_{11}(k) = \gamma_{11}(-k)$; hence, $\rho_{11}(k) = \rho_{11}(-k)$ where $\rho_{11}(k)$ is the traditional linear correlation coefficient between the error terms.
- When $r = s = 2$, we also have $\gamma_{22}(k) = \gamma_{22}(-k)$; hence, $\rho_{22}(k) = \rho_{22}(-k)$ where $\rho_{11}(k)$ is the correlation coefficient between the square values of the error terms.
- When $r > 1$ and $s \in \mathbb{N}$ or $s > 1$ and $r \in \mathbb{N}$, where $\mathbb{N}$ is the set of natural numbers, we obtain the generalized correlation coefficient used by Lawrance and Lewis (1985,1987). In this case, we have $\gamma_{rs}(k) \neq \gamma_{rs}(-k)$ but $\gamma_{rs}(k) = \gamma_{sr}(-k)$; hence, $\rho_{rs}(k) = \rho_{sr}(-k)$.
- $\hat{r}_{rs}(k)$ denotes the generalized sample correlation coefficient.

2.1 Classical portmanteau tests

Box and Pierce (1970) proposed to literature the first portmanteau statistic to test the adequacy of ARMA (p, q) model. Their test statistic utilized the square autocorrelations of the residuals obtained from a fitted mode which is given by

$$Q_{BP} = n\sum_{k=1}^{m} \hat{r}_{11}^2(k), \tag{4}$$

where m is the maximum lag considered for significant autocorrelation. They showed that the limiting distribution of the Q_{BP} statistic is chi-square with $m - p - q$ degrees of freedom. It is worth noting that there is no specific rule to select m, but it is commonly used select it as a value between zero and $n/2$ using $ad\ hoc$ or using $m = \lfloor\sqrt{n}\rfloor$, where $\lfloor x\rfloor$

denotes the largest integer not exceeding x or $m = \log(n)$.

Ljung and Box (1978) improved Q_{BP} by proposing another portmanteau test

$$Q_{11} = n(n+2)\sum_{k=1}^{m} (n-k)^{-1}\hat{r}_{11}^{2}(k), \tag{5}$$

and they showed that the modified test has the same limiting distribution of the Box and Pierce (1970) test but it estimates the type I error more successfully than Q_{BP} with a higher power.

Under the assumptions of ARMA models, Monti (1994) proposed another portmanteau test based on the partial autocorrelations of the residuals and showed that his limiting distribution can also be approximated by $\chi^2_{m-(p+q)}$:

$$M_{11} = n(n+2)\sum_{k=1}^{m} (n-k)^{-1}\hat{\pi}_{1,k}^{2} \tag{6}$$

where $\hat{\pi}_{1,k}$ is the partial autocorrelations of the residuals that is given by

$$\hat{\pi}_{1,k} = \frac{\hat{r}_{11}(k) - \mathbf{r}'_{(k-1)}\widehat{\mathbf{R}}_{11}^{-1}(k-1)\mathbf{r}^{\star}_{(k-1)}}{1 - \mathbf{r}'_{(k-1)}\widehat{\mathbf{R}}_{11}^{-1}(k-1)\mathbf{r}_{(k-1)}}, k = 1, \cdots, m, \tag{7}$$

where $\widehat{\mathbf{R}}_{11}(m)$ is the mth residual autocorrelations matrix defined by

$$\widehat{\mathbf{R}}_{11}(m) = \begin{pmatrix} 1 & \hat{r}_{11}(1) & \dots & \hat{r}_{11}(m) \\ \hat{r}_{11}(1) & 1 & \dots & \hat{r}_{11}(m-1) \\ \vdots & \dots & \vdots & \vdots \\ \hat{r}_{11}(m) & \hat{r}_{11}(m-1) & \dots & 1 \end{pmatrix}, \tag{8}$$

and $\mathbf{r}_{(m)} = (\hat{r}_{11}(1), \hat{r}_{11}(2), \cdots, \hat{r}_{11}(m))'$ and $\mathbf{r}^{\star}_{(k)} = (\hat{r}_{11}(k), \hat{r}_{11}(k-1), \cdots, \hat{r}_{11}(1))'$.

Peňa and Rodríguez (2002) introduced a powerful portmanteau test based on the mth root of the determinant of the mth residual autocorrelations matrix given by (8). Their test statistic is given by

$$D_{11} = n[1 - |\hat{\mathbf{R}}_{11}(m)|^{1/m}],$$

where $|\cdot|$ denotes the determinant of a matrix. Peňa and Rodríguez (2002) showed that asymptotic distribution of the D_{11} statistic can seen as a linear combination of independent chi-squared distributions $a\chi_b^2$. They approximated the distribution of their test by a gamma distribution $\Gamma(\alpha = b/2, \beta = 1/2a)$ with mean $\alpha/\beta = (m+1)/2 - (p+q)$ and variance $\alpha/\beta^2 = (m+1)(2m+1)/3m - 2(p+q)$, where the shape and scale parameters are given by

$$\alpha = \frac{3m[(m+1) - 2(p+q)]^2}{2[2(m+1)(2m+1) - 12m(p+q)]},$$

and

$$\beta = \frac{3m[(m+1) - 2(p+q)]}{2(m+1)(2m+1) - 12m(p+q)}. \quad (9)$$

Peňa and Rodríguez (2002) showed that the D_{11} test statistic can be improved in estimating the significant levels if the autocorrelation coefficients in (8) are replaced with their standardized values

$$\tilde{r}_{11}(k) = \sqrt{\frac{n+2}{n-k}}\hat{r}_{11}(k), k = 1, \cdots, m. \quad (10)$$

Peňa and Rodríguez (2002) also showed that their test statistic is more powerful than its competitors Q_{BP}, Q_{11} and M_{11} in detecting the inadequacy of twenty four different ARMA models.

Peňa and Rodríguez (2006) modified the D_{11} test statistic by proposing another portmanteau test based on the log of $|\widetilde{\mathbf{R}}_{11}(m)|$, where $\widetilde{\mathbf{R}}_{11}(m)$ is defined by (8), replacing $\hat{r}_{11}(k)$ by $\tilde{r}_{11}(k)$ defined by (12). Their test statistic is given by

$$\widetilde{D}_{11} = -\frac{n}{m+1}\log|\widetilde{\mathbf{R}}_{11}(m)|, \quad (11)$$

which can be written in a form that is proportional to a weighted average of the squared partial autocorrelation coefficients

$$\widetilde{D}_{11} = -n \sum_{k=1}^{m} \frac{(m+1-k)}{(m+1)} \log\left(1 - \hat{\pi}_{1,k}^2\right). \tag{12}$$

As clearly seen, the larger weights in (14) will be given to low order coefficients of $\widetilde{D}_{11}$ test and smaller weights will be given to high-order coefficients.

Under the ARMA (p, q) assumptions, Peňa and Rodríguez (2006) showed that asymptotic distribution of the $\widetilde{D}_{11}$ statistic can also be seen as a linear combination of independent chi-squared distributions $a\chi_b^2$. Therefore, they approximated the distribution of their test by a Gamma distribution $\Gamma(\alpha = b/2, \beta = 1/2a)$ with mean $\alpha/\beta = m/2 - (p+q)$ and variance $\alpha/\beta^2 = m(2m+1)/(3(m+1)) - 2(p+q)$, where the parameters are given by

$$\alpha = \frac{3(m+1)[m-2(p+q)]^2}{2[2m(2m+1) - 12(m+1)(p+q)]}, \tag{13}$$

and

$$\beta = \frac{3(m+1)[m-2(p+q)]}{2m(2m+1) - 12(m+1)(p+q)}. \tag{14}$$

They also approximated $\widetilde{D}_{11}$ by a Normal distribution:

$$D_{11}^{\star} = (\alpha/\beta)^{-1/\lambda}(\lambda/\sqrt{\alpha})[(\widetilde{D}_{11})^{1/\lambda} - (\alpha/\beta)^{1/\lambda}\{1 - \frac{1}{\alpha}(\frac{\lambda-1}{\lambda^2})\}], \tag{15}$$

where

$$\lambda = \{1 - \frac{2(m/2-(p+q))(m^2/(4(m+1)) - (p+q))}{3(m(2m+1)/(6(m+1)) - (p+q))^2}\}^{-1} \tag{16}$$

They showed $D_{11}^{\star}$ is asymptotically distributed as standard Normal when m is moderately large $\lambda \approx 4$ and the values of α and β are obtained in (15) and (16), respectively.

Fisher and Gallagher (2012) introduced two weighted portmanteau tests based on the trace of the square of the autocorrelation matrices $\widehat{\mathbf{R}}_{11}^2(m)$ defined in (8). The first one can be considered as a weighted version of the Ljung-Box Q_{11} and it is given by

$$Q_{11}^{w} = n(n+2)\sum_{k=1}^{m} \frac{(m-k+1)}{m}\frac{\hat{r}_{11}^{2}(k)}{n-k}, \tag{17}$$

whereas the second can be seen as a weighted version of the Monti M_{11} and it is given by

$$M_{11}^{w} = n(n+2)\sum_{k=1}^{m} \frac{(m-k+1)}{m}\frac{\hat{\pi}_{1,k}^{2}}{n-k}. \tag{18}$$

Fisher and Gallagher (2012) showed that the distribution of both statistics can asymptotically distributed as $\sum_{k=1}^{m} \lambda_k \chi_k^2$, where $\{\chi_k^2\}$ are independent chi-squared random variables with one degree of freedom and $\{\lambda_k\}$ are the eigenvalues of the matrix $(\mathbf{I}-\mathbf{Q})\mathbf{W}$, where $\mathbf{W}$ is a weighted diagonal matrix with elements $w_{kk} = (m+1-k)/m, k = 1,2,\cdots,m$ (Box, 1954). They approximated the distribution of their statistics by gamma distribution, $\Gamma(\alpha = K_1^2/K_2, \beta = K_2/K_1)$, where

$$K_1 = (m+1)/2, \tag{19}$$

and

$$K_2 = \frac{(m+1)(2m+1)}{3m} - 2(p+q). \tag{20}$$

Their simulation study showed that the weighted test statistics have more powers than the competitors statistics appearing in the literature in many cases.

2.2 Kernel-based normalized spectral density portmanteau test

Anderson (1993); Hong (1996[a, b]) showed that the normalized spectral density of the stationary process $\{\varepsilon_t\}$ can be written in the following form

$$f(\omega) = (2\pi)^{-1}\sum_{\ell\in\mathbb{Z}} \rho_\ell \cos(\ell\omega), \;\; where \;\; \omega \in [-\pi,\pi]. \tag{21}$$

In this case, the null hypothesis, $\mathcal{H}_0$, equals to the null normalized spectral density $f(\omega) = f_0(\omega) = 1/2\pi$. Based on this, Hong (1996[a, b]) used the kernel-based normalized spectral density estimators of $f(\omega)$ to propose three classes of portmanteau tests without knowing the distribution of the innovation series. The only assumption is the iid with mean zero and finite fourth moment. His tests measure the distance, based on Hellinger metric, quadratic norm, and Kullback-Leibler information criterion, between a kernel-based normalized spectral density estimator and null normalized spectral density. The tests are based on

$$\hat{f}_n(\omega) = \frac{1}{2\pi}\sum_{\ell=-n+1}^{n-1} k\left(\frac{\ell}{m_n}\right)\hat{r}_\ell \cos(\ell\omega), \omega \in [-\pi, \pi], \tag{22}$$

where $k(.) > 0$ is a symmetric kernel function which satisfies the following conditions:

1. The kernel function $k: \mathbb{R} \to [-1,1]$ is symmetric differentiable except at a finite number of points, with $k(0) = 1$ and $\int_{-\infty}^{\infty} k^2(u)du < \infty$.
2. $\int_{-\pi}^{\pi} |k(u)|du < \infty$ and for $\omega \in (-\infty, \infty)$, the Fourier transform $K(\omega)$ for $k(u)$ exist and defined as

$$K(\omega) = (2\pi)^{-1}\int_{-\infty}^{\infty} k(u)e^{-iu\omega}du \geq 0.$$

3. $\{\varepsilon_t\}$ is a mean zero fourth order stationary process with $\sum_{\ell=-\infty}^{\infty} \gamma_\ell^2 < \infty$ and

$$\sum_i \sum_j \sum_l |k_4(i,j,l)| < \infty,$$

where $k_4(i,j,l)$ is the fourth joint cumulant of the distribution of $\{\varepsilon_t, \varepsilon_{t+i}, \varepsilon_{t+j}, \varepsilon_{t+l}\}$ and defined as

$$k_4(i,j,l) = e(\varepsilon_t\varepsilon_{t+i}\varepsilon_{t+j}\varepsilon_{t+l}) - e(\tilde{\varepsilon}_t\tilde{\varepsilon}_{t+i}\tilde{\varepsilon}_{t+j}\tilde{\varepsilon}_{t+l}),$$

where $\{\tilde{\varepsilon}_t\}$ is a Gaussian sequence with the same mean and covariance as $\{\varepsilon_t\}$.

Note that $m_n \to \infty$ where $m_n/n \to 0$ is the bandwidth which is depending on the sample size. Hong (1996[a,b]) showed that the distribution of their statistics can be asymptotically approximated by Gaussian and and the power of the proposed tests are usually better than Ljung and Box (1978).

Recently, Gallagher and Fisher (2015) proposed three tests modifying the Fisher and Gallagher (2012) test in (20) by considering three different weighting schemes: the squared Daniell kernel-based weights proposed by Hong (1996[a,b])

$$w_\ell = (n+2)(n-\ell)^{-1}K^2\left(\frac{\ell}{m}\right),$$

the geometrically decaying weights,

$$w_\ell = (p+q)a^{\ell-1},\ for\ some\ \ 0 < a < 1,$$

and the data-adaptive weights which gives the data-adaptive weights test

$$Q_D = n(n+2)\sum_{\ell=1}^{m_0}(n-\ell)^{-1}\hat{r}_\ell^2 + n\sum_{\ell=m_0+1}^{m} w_\ell \hat{r}_\ell^2, \tag{23}$$

where the first m_0 terms obtain the standardizing weight $(n+2)/(n-\ell)$ from the Ljung-Box statistic, and the remaining weights (from m_0+1 to m) selected from the data to be summable $w_\ell = -\log(1-|\hat{\pi}_\ell|)$. It is worth noting that $m_0 \geq m$ in (26) will give the Ljung-Box test.

More recently, Mahdi (2017) proposed a kernel-based portmanteau test based on the Toeplitz autocorrelation matrix defined from Peňa and Rodríguez (2002). He showed that his test is asymptotically the same as the test statistic given by Peňa and Rodríguez (2006) and might be seen as a Kullback-Leibler discrimination information test proposed by Hong (1996[a]). His test statistic is

$$K_m^\star = \frac{-n(m+1)^{-1}\log|\hat{\mathbf{R}}_{11}(m)| - C_n(k)}{\sqrt{2D_n(k)}} \xrightarrow{d} \mathcal{N}(0,1), \tag{24}$$

where $C_n(k) = \sum_{\ell=1}^{m-1}(1-\ell/n)k^2(\ell/m)$, $D_n(k) = \sum_{\ell=1}^{m-2}(1-\ell/n)(1-(\ell+1)/n)k^4(\ell/m)$, and $k(.)$ is the estimated kernel which can be computed from the Daniell kernel

$$k(u) = \sin(\pi u)/\pi u\ \ for\ u \in (-\infty,\infty). \tag{25}$$

2.3 Monte-Carlo (MC) portmanteau tests

One problem pointed by Lin and McLeod (2006) is that the sequence of the standardized residual autocorrelation defined in (12) is not always non-negative definite; hence, the test statistics proposed by Peňa and Rodríguez (2002, 2006) may not exist in many cases. To overcome this problem, Lin and McLeod (2006) proposed the using a Monte-Carlo significance test. They showed that the Monte-Carlo significance test provides a portmanteau test with the correct estimate of size and is almost always more powerful than the competitors statistics appearing in the literature. The p-value for the portmanteau test statistics using the Monte-Carlo test can be computed by the algorithm outlined below (see Lin and McLeod (2006), Mahdi (2011), and Mahdi and McLeod (2012)). Alok (2020) also worked on crosscorrelation of square of residualsused using the Monte-Carlo approach unlike Mahdi who used the asymptotic distribution approach.

The following are the steps taken to calculate the p-value using Monte-Carlo technique.

Step 1: Simulate a time series from a given model (say ARMA (p,q)).

Step 2: Fit a suitable model to this simulated data and obtain the residuals (say AR (p)).

Step 3: Use the residuals and compute the observed value of the portmanteau test statistic for a set of lags m (say $\mathfrak{D}_m^{(o)}$).

Step 4: Use the residuals from Step 3 and simulate a time series data using estimated parameters obtained in Step 2. This step is replicated B times (say $B = 1000$).

- In each replicate, fit the model AR(p) to the simulated data and obtain the residuals.
- Use the residuals and calculate the corresponding test value of the portmanteau test statistic (say $\mathfrak{D}_m^{(i)}, i = 1,2,\cdots,B$).
- Count the average number of times that the test values of the portmanteau test statistic greater than or equal to the observed statistic.
- The estimated p-value is given by,

$$\hat{p}_j = \frac{\#\{\mathfrak{D}_m^{(i)} \geq \mathfrak{D}_m^{(o)}, i = 1,2,\ldots,B\} + 1}{B + 1}. \tag{26}$$

The larger the B number, the more accurate estimate of the p-value. The approximate 95% margin of error for the p-value computed using the normal approximation for the binomial is $\pm 1.96\sqrt{\hat{\mathrm{p}}_j(1-\hat{\mathrm{p}}_j)/B}$.

Step 5: Repeat Steps 1-4 for N times, wh**E**re N is the number of simulations (Usually, $N \geq 1000$) and each time, calculate $\hat{\mathrm{p}}_j, j = 1,2,\cdots,N$.

Step 6: Calculate the average of the p-values obtained in the previous step. This will gives the estimated p-value based Monte-Carlo test using N simulations with B replications:

$$p - value = \sum_{j=1}^{N} \hat{\mathrm{p}}_j/N.$$

2.4 Random Weighted Bootstrap (RWB) tests

Recently, Lee (2016) proposed using the Wild bootstrap of the Ljung-Box portmanteau test in time series when ARCH is presented. Zhu (2016) proposed a Randomly Weighted Bootstrap (RWB) which can be seen as a variant of the Wild bootstrap approach. A bootstrapping method is a nonparatmetric way shown to be robust when distributional assumptions are violated. The Randomly Weighted Bootstrap algorithm is as follows:

Step 1: Estimate the model from (1) and calculate the correlation coefficients based on the fitted residuals.

Step 2: Generate a sequence of iid random variables, say $\boldsymbol{w}^* = \{w_1, w_2, \dots, w_n\}$ independent of the data from a common distribution satisfies $p(w_i \geq 0) = 1$ with mean and variance both equal to 1.

Step 3: Calculate $\delta = \boldsymbol{w}^*\{\sqrt{n}(\hat{\boldsymbol{r}}_m^* - \hat{\boldsymbol{r}}_m)\}$.

Step 4: Repeat steps 2 and 3 B times (say $B = 1000$), to obtain $\{\delta_1, \delta_2, \dots, \delta_B\}$ and compute its covariance matrix and its associated eigenvalues.

Step 5: Generate N (say 1000) iid random data $\{z_1^{(j)}, z_2^{(j)}, \dots, z_m^{(j)}\}_{j=1}^{N}$ from multivariate normal distribution with identity matrix and compute the sequence $\{G^{(j)}\}_{j=1}^{N}$ by

$$G^{(j)} = \sum \lambda_i^* \, (z_i^{(j)})^2.$$

Step 6: The sequence $\{G^{(j)}\}_{j=1}^{N}$ is the bootstrap sampling of the portmanteau distribution and the critical p-value can be calculated in the same maner of Equation (28).

CHAPTER 3: TEST FOR NONLINEARITY

One problem pointed by Granger and Anderson (1978) and Tong and Lim (1980) was that the squared residuals of many of Box and Jenkins (1970) ARMA models are significantly autocorrelated even though the residual autocorrelations are not. This suggests that these models are not adequate and the innovation of these models might be uncorrelated but not independent. Granger and Anderson (1978) and Tong and Lim (1980) suggested of using the autocorrelation function of the squared values of the residuals to detect the nonlinear dependency.

Engle (1982) showed that the Box and Pierce (1970) and Ljung and Box (1978) test, based on the autocorrelation function of the residuals, might fail to detect the presence of the ARCH in many financial time series. In this respect, he proposed a Lagrange multiplier test based on the autocorrelations of the squared-residuals and used it to test for ARCH structure.

The errors $\{\varepsilon_t\}$ in (1) follows an Autoregressive Conditional Heteroskedasticity (ARCH) model with order b model if it can be written as follows:

$$\varepsilon_t = \xi_t \sigma_t, \qquad \sigma_t^2 = \omega + \sum_{i=1}^{b} \alpha_i \varepsilon_{t-i}^2, \tag{27}$$

where $\{\xi_t\}$ is a sequence of i.i.d. random variables with a mean value of 0 and variance value of 1, $\omega > 0$, $\alpha_i \geq 0$.

The ARCH model has been generalized by Bollerslev (1986) to the Generalized Autoregressive Conditional Heteroskedasticity (GARCH) model. The innovation $\{\varepsilon_t\}$ in (29) follows a GARCH (b, a) model given by

$$\varepsilon_t = \xi_t \sigma_t, \qquad \sigma_t^2 = \omega + \sum_{i=1}^{b} \alpha_i \varepsilon_{t-i}^2 + \sum_{j=1}^{a} \beta_j \sigma_{t-j}^2, \tag{28}$$

where $\{\xi_t\}$ is a sequence of iid random variables with a mean value of 0 and variance value of 1, $\omega > 0$, $\alpha_i \geq 0, \beta_j \geq 0$, and $\sum_{i=1}^{max\,(b,a)} (\alpha_i + \beta_i) < 1$.

Since then, the ARCH and GARCH models become essential statistical tools for modeling financial time series and modern option pricing theory and practice. In literature, there are several types of GARCH models that can be used to model the dynamic behavior of conditional ARCH in time series applications (more details are provided in Tsay (2005) and Carmona (2014)). The most commonly used models are:

• The Exponential Generalized Autoregressive Conditional Heteroskedastic (EGARCH (p,q)) model of Nelson (1991) assumes the form

$$\log\sigma_t^2 = \omega + \sum_{i=1}^{b} \alpha_i f(Z_{t-i}) + \sum_{j=1}^{a} \beta_j \log\sigma_{t-j}^2, \tag{29}$$

where $f(Z_t) = \theta Z_t + \lambda(|Z_t| - E|Z_t|)$ is a function allows the sign and the magnitude of the $Z_t \sim \mathcal{N}(0,1)$ (or Z_t ~a generalized error distribution) to have separate effects on the volatility, θ and λ are coefficients. The EGARCH(b,a) can be rewritten in the as

$$\varepsilon_t = \xi_t\sigma_t, \qquad \log\sigma_t^2 = \omega + \frac{1 + \sum_{i=1}^{a-1} \beta_i \mathbf{B}^i}{1 - \sum_{j=1}^{b-1} \alpha_j \mathbf{B}^j} f(\varepsilon_{t-1}), \tag{30}$$

where $\mathbf{B}$ is the back-shift (or lag) operator such that $\mathbf{B}f(\varepsilon_t) = f(\varepsilon_{t-1})$, $\sum_{i=1}^{a-1} \beta_i \mathbf{B}^i$ $(i = 1,2,\cdots,a-1)$ and $\sum_{j=1}^{b-1} \alpha_j \mathbf{B}^j$ $(j = 1,2,\cdots,b-1)$ are polynomials with zeros outside the unit circle and have no common factors.

• The GARCH-in-mean (GARCH-M(b,a)) model proposed by Engle et al. (1987), where "M" denotes the GARCH in the mean. The GARCH-M adds a Heteroskedasticity term into the mean equation that can be interpreted as a risk premium.

$$z_t = \mu + \lambda\sigma_t^2 + \varepsilon_t, \quad \varepsilon_t = \xi_t\sigma_t, \tag{31}$$

$$\sigma_t^2 = \omega + \sum_{i=1}^{b} \alpha_i \varepsilon_{t-i}^2 + \sum_{j=1}^{a} \beta_j \sigma_{t-j}^2, \tag{32}$$

where μ and λ are constants. The parameter λ denotes the risk premium parameter. When $\lambda > 0$ then z_t is positively related to its volatility. There are many other formulation for GARCH -M(p,q) including $z_t = \mu + \lambda\sigma_t + \varepsilon_t$ and $z_t = \mu + \lambda\log\sigma_t^2 + \varepsilon_t$.

• The Threshold GARCH (TGARCH(b, a)) model by Glosten et al. (1993) and Zakoian (1994) is commonly used to handle leverage effects is the threshold GARCH defined by

$$\sigma_t^2 = \omega + \sum_{i=1}^{b} (\alpha_i + \gamma_i I_{t-i})\varepsilon_{t-i}^2 + \sum_{j=1}^{a} \beta_j \sigma_{t-j}^2, \tag{33}$$

where I_{t-i} is an indicator for negative ε_{t-i}; that is

$$I_{t-i} = \begin{cases} 1 & \text{if } \varepsilon_{t-i} \leq s, \\ 0 & \text{if } \varepsilon_{t-i} \geq s, \end{cases}$$

where s is the threshold is used to separate the impacts of past shocks. When $s = 0$, then it can clearly be seen that $\varepsilon_{t-i} > 0$ contributes $\alpha_i \varepsilon_{t-i}^2$ to σ_{t-j}^2, whereas $\varepsilon_{t-i} < 0$ has a larger impact $(\alpha_i + \gamma_i)\varepsilon_{t-i}^2$ with $\gamma_i > 0$.

McLeod and Li (1983) introduced a portmanteau test to detect the presence of ARCH based on the squared-residuals autocorrelations. Their test statistic is given by

$$Q_{22} = n(n+2) \sum_{k=1}^{m} (n-k)^{-1} \hat{r}_{22}^2(k). \tag{34}$$

It is worth noting that the McLeod and Li (1983) test was derived under the assumptions of the ARMA model but it asymptotically distributed as χ_m^2 which does not depend on the order of the fitted ARMA model, (p, q). Note also that the Q_{22} is widely used to detect nonlinearity dependency in several time series including GARCH models.

Li and Mak (1994) proposed a portmanteau test statistic based on the ARCH assumptions. Their test statistic is given by

$$L_b = n \sum_{k=1}^{m} \hat{r}_{22}^{\star 2}(k),$$

where $\hat{r}_{22}^{\star}(k)$is the standardized squared-residual autocorrelation obtained from a fitted GARCH model that is defined by

$$\hat{r}_{22}^{\star}(k) = \frac{\sum_{t=k+1}^{n} (\hat{\varepsilon}_t^2/\hat{\sigma}_t - \bar{\varepsilon})(\hat{\varepsilon}_{t-k}^2/\hat{\sigma}_{t-k} - \bar{\varepsilon})}{\sum_{t=1}^{n} (\hat{\varepsilon}_t^2/\hat{\sigma}_t - \bar{\varepsilon})^2},$$

where $\bar{\varepsilon} = n^{-1} \sum \hat{\varepsilon}_t^2 / \hat{\sigma}_t$ and $\hat{\sigma}_t$ are the estimated sample conditional variances of the GARCH (b, a) model defined in (31).

Peňa and Rodríguez (2002) and Peňa and Rodríguez (2006) replaced the residual autocorrelations in the matrix defined in (8) by the squared-residual autocorrelations and used it to derive new portmanteau tests that can be used to detect nonlinearity in several time series. The Peňa and Rodríguez (2002) test for nonlinearity is given by

$$D_{22} = n[1 - |\widehat{\mathbf{R}}_{22}(m)|^{1/m}], \tag{35}$$

and the Peňa and Rodríguez (2006) test is

$$\widetilde{D}_{22} = -\frac{n}{m+1} \log|\widetilde{\mathbf{R}}_{22}(m)|, \tag{36}$$

where $\widehat{\mathbf{R}}_{22}(m)$ is the matrix of mth order of squared-residual autocorrelation:

$$\widehat{\mathbf{R}}_{22}(m) = \begin{pmatrix} 1 & \hat{r}_{22}(1) & \dots & \hat{r}_{22}(m) \\ \hat{r}_{22}(1) & 1 & \dots & \hat{r}_{22}(m-1) \\ \vdots & \dots & \vdots & \vdots \\ \hat{r}_{22}(m) & \hat{r}_{22}(m-1) & \dots & 1 \end{pmatrix}. \tag{37}$$

The asymptotic distribution of D_{22} is gamma $\Gamma(\alpha, \beta)$, where α and β are the same as those defined in (10) and (11), respectively, where $p + q = 0$. Also, the asymptotic distribution of $\widetilde{D}_{22}$ has two approximations by using the Gamma and the Normal distributions. For the gamma, $\Gamma(\alpha, \beta)$ distribution, the shape and the scale α and β are the same as those defined in (15) and (16) , respectively, where $p + q = 0$. For the normal distribution, Peňa and Rodríguez (2006) showed that $\widetilde{D}_{22}$ has the same normal distribution of $\widetilde{D}_{11}$ defined in (17), where λ is given in (18) where $p + q = 0$.

Rodriguez and Ruiz (2005) proposed a portmanteau test for ARCH models using the information contained in the sample autocorrelations of non-linear transformations of the underlying process. Their test statistic can be used to test for whether the autocorrelations of the residuals differ from zero and at the same time it can be used to test for possible relationship among successive autocorrelation coefficients. Their test is

$$Q_i^{\star}(M) = n \sum_{k=1}^{M-i} [\sum_{\ell=0}^{i} \tilde{r}_{11}(k+\ell)]^2, i = 0,1,\cdots, M-1, \tag{38}$$

where $\tilde{r}_{11}(k+\ell)$ is the standardized sample autocorrelation of order $k+\ell$.

For the different value of i for M we can have different collection of statistics all these have different informations on the possible sample correlation pattern. If one wants to have McLeod-Li statistic in (36) then choose $i = 0$. For this scenario, the statistic is obtained by adding up the squared estimated autocorrelations. In case if all of these autocorrelations are small, the statistic will be small and hence we will reject the null hypothesis . But when $i = 1$ the statistics obtained will be a correlation between sample auto correlation one leg apart. If they are strongly correlated then the null hypothesis can be rejected even if the coefficient $r_{11}(j)$ are very small.

Rodriguez and Ruiz (2005) showed that $Q_i^{\star}(M)$ behaves asymptotically as a linear combination of independent chi-squared variables with one degree of freedom that can be approximated by a gamma distribution $\Gamma(\alpha = a^2/2b, \beta = a/2b)$, where

$$a = (i+1)(M-i), \tag{39}$$

and

$$b = (M-i)(i+1)^2 + 2\sum_{j=1}^{i+1} (M-i-j)(i+1-j)^2. \tag{40}$$

Lin and McLeod (2006) also showed that the Peňa and Rodríguez (2002, 2006) can distort the size of the test; hence, they suggested to use the Monte-Carlo significance test to test for nonlinearity. They also showed that the Monte-Carlo significance test provides a portmanteau test with the correct estimate of size and is almost always more powerful than the competitors statistics appearing in the literature.

Fisher and Gallagher (2012) modified their weighted test defined in (20) to detect the nonlinearity presence in time series models. Their modified test is given by

$$Q_{22}^{w} = n(n+2)\sum_{k=1}^{m} \frac{(m-k+1)}{m}\frac{\hat{r}_{22}^{2}(k)}{n-k}, \tag{41}$$

which asymptotically distributed as gamma $\Gamma(\alpha = K_1^2/K_2, \beta = K_2/K_1)$, where

$$K_1 = (m+1)/2 \tag{42}$$

and

$$K_2 = (m+1)(2m+1)/3m \tag{43}$$

In addition, Fisher and Gallagher (2012) proposed a weighted test that can be seen as a modified version of Li and Mak (1994) statistic. The modified Li and Mak (1994) weighted test is given by

$$L_b^w = n\sum_{k=1}^{m} \frac{m-k+(b+1)}{m}\hat{r}_{22}^{\star 2}(k), \tag{44}$$

where L_b^w statistic is asymptotically distributed as $\chi^2_{m-(b+a)}$, where (a, b) are the order of the fitted GARCH model.

The problem with the aforementioned statistics that they respond well to ARMA and GARCH models but they tend to have a lack of power compared to other types of time series models, especially when the residuals to different powers might be correlated.

Recently, Psaradakis and Vávra (2019) used the Lawrance and Lewis (1985, 1987) idea, which was based on using the generalized correlation of the residuals to detect nonlinear dependency in time series, and proposed four different statistics from stationary linear models to test for linearity. Their test statistics are given by

$$\tilde{Q}_{rs} = n\sum_{k=1}^{m} \hat{r}_{rs}^{2}(k), \qquad (r \neq s \in \mathbb{N}). \tag{45}$$

and

$$Q_{rs} = n(n+2)\sum_{k=1}^{m} (n-k)^{-1}\hat{r}_{rs}^{2}(k), \qquad (r \neq s \in \mathbb{N}). \tag{46}$$

where $\hat{r}_{rs}(k)$ is the generalized correlation coefficient between the residuals to the power

r and the residuals to the power s.

Psaradakis and Vávra (2019) preliminary analysis suggested that the Q_{rs} tests control the Type I error probability somewhat more successfully than the $\tilde{Q}_{rs}$ tests; hence, they restricted their simulation study focusing on the test statistics Q_{rs}. The authors reported that $\tilde{Q}_{rs}$ and Q_{rs} are asymptotically distributed as χ_m^2 and tend to have more power in detecting nonlinearity in time series models comparing to the McLeod and Li (1983) test statistic. Motivated by the ideas of Lawrance and Lewis (1985, 1987), and Psaradakis and Vávra (2019), I proposed in the next section new test statistics that can be used to detect nonlinearity in time series models.

The $\tilde{Q}_{rs}$ and $\tilde{Q}_{sr}$ statistics are similar in spirit to Box and Pierce (1970) defined in (4) but the authors replaced the autocorrelations of the residuals defined in (4) by the generalized correlations at different powers getting the two tests: $\tilde{Q}_{rs}$ which is based on the cross-correlations between ε_t^r and ε_{t+k}^s; and $\tilde{Q}_{sr}$ which is based on the cross-correlations between ε_t^s and ε_{t+k}^r. Similarly, Q_{rs} and Q_{sr} statistics can be seen as modify tests similar in spirit to McLeod and Li (1983) defined in (36) obtained by replacing the autocorrelations of the squared-residuals defined by (36) by the cross-correlations between ε_t^r and ε_{t+k}^s and the cross-correlations between ε_t^s and ε_{t+k}^r, respectively.

More recently, Mahdi and Fisher (2021) proposed a portmanteau test using the block matrix of autocorrelations and cross-correlations of residuals and squared-residual, $\widehat{\mathbf{R}}(m)$, defined by

$$\widehat{\mathbf{R}}(m) = \begin{bmatrix} \widehat{\mathbf{R}}_{11}(m) & \widehat{\mathbf{R}}_{12}(m) \\ \widehat{\mathbf{R}}_{12}'(m) & \widehat{\mathbf{R}}_{22}(m) \end{bmatrix}_{2(m+1)\times 2(m+1)}, \tag{47}$$

where $\widehat{\mathbf{R}}_{11}(m)$ and $\widehat{\mathbf{R}}_{22}(m)$ are defined by (8) and (40), respectively, and $\widehat{\mathbf{R}}_{12}(m)$ is the matrix of cross-correlations between residuals and their squares which is given by

$$\widehat{\mathbf{R}}_{12}(m) = \begin{pmatrix} \hat{r}_{12}(0) & \hat{r}_{12}(1) & \dots & \hat{r}_{12}(m) \\ \hat{r}_{12}(-1) & \hat{r}_{12}(0) & \dots & \hat{r}_{12}(m-1) \\ \vdots & \dots & \vdots & \vdots \\ \hat{r}_{12}(-m) & \hat{r}_{12}(-m+1) & \dots & \hat{r}_{12}(0) \end{pmatrix}. \tag{48}$$

The test statistic proposed by Mahdi and Fisher (2021) is given by

$$C_m = -\frac{n}{m+1}\log|\widehat{\mathbf{R}}(m)|, \tag{49}$$

where $|\cdot|$ denotes the determinant of a matrix.

Under the assumptions of ARMA (p, q) model, Mahdi and Fisher (2021) proposed approximated the asymptotic distribution of C_m by gamma and showed that the distribution has a mean of $\alpha\beta = 2m + 5 - (p + q)$. They showed that the C_m test can be seen as a linear combination of four weighted tests. The first and the second components elaborate the partial autocorrelation of the residuals and the squared-residuals, respectively. The third and the fourth components elaborate the cross-correlation between the residuals and their squares and vice-versa, respectively. Hence, the C_m test can be used to detect, simultaneously, the linear and nonlinear dependency in stationary time series data. Their simulation study demonstrated that the C_m statistic tends to have higher power than the competitors statistics appearing in the literature, particularly in detecting ARMA with GARCH errors and other nonlinear models. They utilized the Randomly Weighted Bootstrap (RWB) approach which was proposed by Zhu (2016) to improve the size and the power of their statistic and showed that the RWB be robust and give the correct size and tend to be more powerful than the one based on the asymptotic distributions.

CHAPTER 4: SIMULATION STUDY

In this section, I presented the simulation results regarding the finite-sample size and power properties of the aforementioned tests. I invistigated the effects of Gaussian and non-Gaussian[1] noise on performance of the protmanteau tests. First, I calculated the significant level based on 5% nominal levels using the limiting distributions, random weighted Bootstrapping (RWB), and Monte-Carlo (MC) significance tests. After that, I studied the power of the portmanteau tests based on the three different technques. For brevity, I considered the test at lags $m = 5$, and 10. Monte-Carlo simulation was based on 1000 simulations, where each simulation has used 500 replications. In my numerical study, I used the **R** package ***portes*** in a parallel framework (Mahdi, 2020a).

4.1 Empirical sizes

In this section, I tested the adequacy of fitted ARMA and GARCH models. First, I generated the data from the following four linear models:

$\boldsymbol{A_1}$**:** AR(1): $z_t = -0.8z_{t-1} + \varepsilon_t$;

$\boldsymbol{A_2}$**:** AR(2): $z_t = 0.5z_{t-1} - 0.4z_{t-2} + \varepsilon_t$;

$\boldsymbol{A_3}$**:** MA(1): $z_t = 0.9\varepsilon_{t-1} + \varepsilon_t$;

$\boldsymbol{A_4}$**:** ARMA(1,1): $z_t = 0.7z_{t-1} + 0.3\varepsilon_{t-1} + \varepsilon_t$;

In each case, I fitted the true model and calculated the Type-I error at lags $m = 5$ and 10 with sample sizes $n = 100, 300,$ and 500 to cover small, moderate and large sample sizes based on Gaussian and non-Gaussian. For non-Gaussian case, I considered the t-distribution with degrees of freedoms of 3, 6, 9, and 12, a range of values which are sufficiently representative of mild asymmetry and heavy tailed leptokurtosis distributions in many finnancial time series. For brevity, I averaged the relative rejection frequncies across the non-Gaussian cases. In my simulation study, I considered the test statistics

---

[1] The results based on Skewed Normal and Skewed Students' t-distributions with skewness parameters skewness $\{-2, -1.5, -0.5, 0.5, 1, 1.5, 2\}$ and degress of freedom $\{3,6,9,12\}$ are given in the Appendix.

Ljung-Box (1978), Q_{11}, Fisher and Gallagher (2012), Q_{11}^{w}, and Mahdi and Fisher (2021), C_m.

Tables 1-3 shows the results. The results show that the asymptotic distribution of the test statistics can distort the size, especially for skewed data with small sizes. On the other hand, the MC and RWB approaches provided accuartae results for estimating the type I error for both Gaussian and non-Gaussian cases and regardless of the sample size.

Table 1: Empirical sizes, for 5% nominal tests, based on models $A_1 - A_4$, n=100

Model	Lags	Asymptotic distribution			MC			RWB		
		Q_{11}	Q_{11}^{w}	C_m	Q_{11}	Q_{11}^{w}	C_m	Q_{11}	Q_{11}^{w}	C_m
		Gaussian distribution								
A_1	5	5.5	5.8	5.7	3.4	6.7	3.1	6.0	6.5	6.0
	10	5.3	4.4	4.3	3.6	6.3	3.2	5.6	6.3	5.8
A_2	5	5.3	3.6	5.7	6.4	3.6	6.7	6.0	6.2	3.8
	10	3.5	2.2	3.9	6.1	3.8	6.4	5.6	6.0	3.9
A_3	5	9.4	9.0	6.5	3.6	3.2	6.1	6.1	5.9	3.9
	10	6.5	6.6	5.2	3.9	3.4	5.9	5.9	5.7	4.0
A_4	5	6.5	5.3	4.6	3.3	3.6	3.4	3.4	3.8	6.4
	10	4.8	3.2	3.7	3.5	3.7	3.5	3.6	3.9	6.0
		Student-t distribution								
A_1	5	3.9	3.4	8.1	2.8	7.3	2.8	3.4	6.4	3.5
	10	5.2	3.1	7.4	2.9	6.8	2.9	3.6	6.1	3.7
A_2	5	3.5	2.6	9.0	8.2	8.1	2.9	3.2	3.3	6.3
	10	4.2	2.4	8.1	7.9	7.6	3.0	3.3	3.4	6.0
A_3	5	6.5	5.2	10.0	7.8	8.7	8.4	3.5	6.0	3.5
	10	6.2	5.5	9.3	7.4	8.1	7.9	3.7	5.7	3.6
A_4	5	3.7	3.6	8.1	3.0	8.3	6.5	6.4	6.6	3.4
	10	4.2	2.2	7.7	3.2	7.9	6.2	6.0	6.3	3.6

Table 2: Empirical sizes, for 5% nominal tests, based on models $A_1 - A_4$, n=300.

Model	Lags	Asymptotic distribution			MC			RWB		
		Q_{11}	Q_{11}^{w}	C_m	Q_{11}	Q_{11}^{w}	C_m	Q_{11}	Q_{11}^{w}	C_m
		Gaussian distribution								
A_1	5	5.2	3.6	5.3	3.8	5.7	3.4	6.2	4.1	6.1
	10	5.4	3.3	4.6	4.0	5.3	3.6	6.0	4.2	5.9
A_2	5	4.2	3.6	5.4	6.5	6.1	5.7	5.5	3.9	3.7
	10	5.1	2.1	3.9	6.0	5.8	5.4	5.3	4.0	3.8
A_3	5	5.5	5.9	6.0	3.5	6.1	3.4	4.0	5.9	3.9
	10	5.5	4.5	4.7	3.6	5.9	3.5	4.3	5.7	4.2

$\boldsymbol{A_4}$	5	5.0	5.5	4.8	5.9	3.9	6.4	5.6	4.0	5.5
	10	5.9	2.9	3.9	5.5	4.0	6.2	5.2	4.2	5.3
		Student-t distribution								
$\boldsymbol{A_1}$	5	3.6	2.8	6.6	6.4	3.3	3.3	3.6	3.5	5.7
	10	5.3	2.9	6.0	6.1	3.4	3.5	3.7	3.6	5.4
$\boldsymbol{A_2}$	5	4.6	3.2	7.7	6.5	6.3	6.4	6.3	6.2	3.4
	10	3.9	2.9	6.2	6.2	6.0	6.1	6.1	6.0	3.6
$\boldsymbol{A_3}$	5	4.2	4.9	8.3	6.7	3.3	3.3	5.8	3.8	5.7
	10	5.8	4.1	6.7	6.3	3.4	3.5	5.6	4.0	5.4
$\boldsymbol{A_4}$	5	5.3	3.6	7.4	3.1	3.4	6.5	5.7	6.2	6.0
	10	4.4	2.4	6.7	3.3	3.6	6.2	5.5	5.8	5.6

Table 3: Empirical sizes, for 5% nominal tests, based on models $A_1 - A_4$, n=500.

Model	**Lags**	**Asymptotic distribution**			**MC**			**RWB**		
		Q_{11}	Q_{11}^{w}	C_m	Q_{11}	Q_{11}^{w}	C_m	Q_{11}	Q_{11}^{w}	C_m
		Gaussian distribution								
$\boldsymbol{A_1}$	5	5.2	4.7	5.2	3.9	4.4	4.4	5.0	4.3	5.6
	10	5.1	4.0	4.9	4.1	4.5	4.7	5.2	4.4	5.3
$\boldsymbol{A_2}$	5	5.7	4.2	5.8	3.8	5.8	6.2	5.5	5.2	4.9
	10	4.2	2.6	4.1	4.0	5.5	6.0	5.2	4.8	5.0
$\boldsymbol{A_3}$	5	5.0	5.3	5.6	6.3	4.4	3.9	5.4	5.4	5.7
	10	5.3	4.7	4.2	6.0	4.6	4.1	5.1	5.0	5.3
$\boldsymbol{A_4}$	5	4.0	3.9	5.4	4.1	4.2	4.1	5.3	4.1	5.3
	10	4.4	2.7	3.7	4.4	4.3	4.3	4.9	4.2	5.0
		Student-t distribution								
$\boldsymbol{A_1}$	5	5.3	4.5	5.8	3.9	5.9	6.1	4.4	4.1	6.1
	10	5.5	4.5	5.6	4.1	5.5	5.7	4.5	4.3	5.8
$\boldsymbol{A_2}$	5	5.4	3.8	6.8	6.4	5.9	6.3	4.4	5.6	3.7
	10	5.5	3.4	6.2	6.2	5.7	6.0	4.7	5.4	3.8
$\boldsymbol{A_3}$	5	5.7	6.3	6.1	6.4	3.5	3.5	4.2	5.9	4.4
	10	5.7	4.8	6.0	6.2	3.7	3.6	4.3	5.6	4.7
$\boldsymbol{A_4}$	5	4.6	4.6	6.9	6.2	5.8	6.1	5.9	4.2	4.4
	10	5.0	2.4	6.6	5.8	5.6	5.9	5.5	4.4	4.7

Tables 1, 2, and 3, show that the asymptotic distribution approach can distort the size of the portmanteau test, especially for small sample sizes. Results indicate that as the sample size is increased, the test values get closer to the respective significant level. Similarly, for the same sample sizes, the observed sizes approached the nominal size as lags increased from $m = 5$ to $m = 10$. Results showed that the test statistics based on Monte-Carlo (MC) technique provide better approximate to the significant levels than based on the asymptotic approach. Moreover, the Randomly Weighted Bootstrap (RWB) were more improved and closer to the significant value as compared to both Monte-Carlo (MC) and asymptotic approaches. In all the cases, it's also observed that the Gaussian results were closer to the empirical size than using student-t distribution.

I also evaluated the empirical type I error rates of the test statistics McLeod-Li (1983), Q_{22}, Fisher and Gallagher (2012), Q_{22}^{W}, Li-Mak (1994), L_b, and Mahdi and Fisher (2021), C_m, based on the four nonlinear $AR(p) - GARCH(b, a)$ models studied by Carlos Velasco and Xuexin Wang (2014):

$\boldsymbol{B_1}$: $AR(2) - ARCH(1)$: $z_t = 0.5z_{t-1} + 0.2z_{t-2} + \varepsilon_t, \quad \varepsilon_t = \sigma_t \varepsilon_t, \quad \sigma_t^2 = 0.1 + 0.4\varepsilon_{t-1}^2$

$\boldsymbol{B_2}$: $AR(1) - ARCH(2)$: $z_t = 0.5z_{t-1} + \varepsilon_t, \quad \varepsilon_t = \sigma_t \varepsilon_t, \quad \sigma_t^2 = 0.1 + 0.4\varepsilon_{t-1}^2 + 0.2\varepsilon_{t-2}^2$

$\boldsymbol{B_3}$: $AR(2) - ARCH(2)$: $z_t = 0.5z_{t-1} + 0.2z_{t-2} + \varepsilon_t, \quad \varepsilon_t = \sigma_t \varepsilon_t, \quad \sigma_t^2 = 0.1 + 0.4\varepsilon_{t-1}^2 + 0.2\varepsilon_{t-2}^2$

$\boldsymbol{B_4}$: $AR(2) - GARCH(1,1)$: $z_t = 0.5z_{t-1} + 0.2z_{t-2} + \varepsilon_t, \quad \varepsilon_t = \sigma_t \varepsilon_t, \quad \sigma_t^2 = 0.1 + 0.4\varepsilon_{t-1}^2 + 0.5\sigma_{t-1}^2$

The results are shown in Tables 4-6.

Table 4: Empirical sizes, for 5% nominal tests, based on models $B_1 - B_4$, n=100.

Model	Lags	Asymptotic distribution				MC				RWB			
		Q_{22}	Q_{22}^{W}	L_b	C_m	Q_{22}	Q_{22}^{W}	L_b	C_m	Q_{22}	Q_{22}^{W}	L_b	C_m
		Gaussian distribution											
$\boldsymbol{B_1}$	5	1.7	1.4	1.4	4	6.5	6.6	3.4	6.6	6.5	3.5	6.5	3.6
	10	2.1	1.6	1.7	2.9	6.2	6.2	3.6	6.3	6.3	3.7	6.0	3.8
$\boldsymbol{B_2}$	5	1.4	0.9	1.2	3.5	6.7	3.4	6.3	6.6	3.6	3.6	3.6	5.9
	10	2.3	1	1.8	2.2	6.4	3.6	6.0	6.4	3.8	3.8	3.7	5.5
$\boldsymbol{B_3}$	5	1.4	0.6	1.2	3.3	6.2	3.4	6.0	3.4	6.4	6.1	3.6	5.8
	10	2.2	1.1	1.8	2.2	5.8	3.5	5.6	3.5	6.0	5.9	3.9	5.5
$\boldsymbol{B_4}$	5	2.1	2.1	1.8	3.7	3.4	6.1	6.5	3.5	6.4	3.9	3.6	3.6
	10	2.7	1.9	2.2	3	3.6	5.8	6.0	3.7	6.1	4.1	3.8	3.8
		Student-t distribution											
$\boldsymbol{B_1}$	5	2.7	1.6	2.4	10.1	8.5	6.6	6.9	6.7	3.1	6.6	6.0	6.5
	10	2.7	2.6	2.1	8.8	7.9	6.1	6.4	6.3	3.2	6.2	5.7	6.2
$\boldsymbol{B_2}$	5	1.8	0.8	1.8	9.5	8.6	2.9	2.8	2.9	3.1	6.3	6.8	3.2
	10	2.1	1.8	1.7	7.8	8.3	3.1	2.9	3.1	3.2	5.9	6.4	3.4
$\boldsymbol{B_3}$	5	2.1	0.6	1.7	10.4	3.3	3.0	8.5	2.8	3.1	3.4	6.1	5.9
	10	2.5	1.6	1.7	9.4	3.5	3.2	8.0	3.0	3.3	3.6	5.9	5.5
$\boldsymbol{B_4}$	5	2.4	1.6	1.9	8.3	3.0	3.2	3.2	3.0	3.5	6.6	5.9	6.8
	10	2.9	1.9	2.2	7.4	3.2	3.3	3.4	3.2	3.6	6.4	5.5	6.5

Table 5: Empirical sizes, for 5% nominal tests, based on models $B_1 - B_4$, n=300.

Model	Lags	Asymptotic distribution				MC				RWB			
		Q_{22}	Q_{22}^{w}	L_b	C_m	Q_{22}	Q_{22}^{w}	L_b	C_m	Q_{22}	Q_{22}^{w}	L_b	C_m
		Gaussian distribution											
$\boldsymbol{B_1}$	5	2.7	1.9	2.3	5.1	6.3	6.4	3.6	3.5	4.4	3.7	4.1	4.4
	10	2.7	2.4	2.4	4.2	6.1	6.0	3.8	3.6	4.6	3.8	4.3	4.6
$\boldsymbol{B_2}$	5	1.7	1.0	1.6	4.4	3.9	5.8	3.8	3.4	5.9	6.1	4.3	6.2
	10	2.4	1.5	2.2	3.4	4.1	5.6	4.0	3.5	5.7	5.9	4.4	6.0
$\boldsymbol{B_3}$	5	1.7	0.7	1.6	4.7	6.0	3.6	3.5	6.4	6.0	4.1	5.7	6.3
	10	2.2	1.4	1.8	2.9	5.6	3.7	3.6	6.1	5.8	4.4	5.5	5.9
$\boldsymbol{B_4}$	5	3	2.4	2.9	4.9	3.6	5.8	6.1	3.8	5.7	4.2	3.7	5.6
	10	3.3	2.9	3.0	3.5	3.7	5.4	5.9	4.0	5.5	4.4	3.9	5.4
		Student-t distribution											
$\boldsymbol{B_1}$	5	4.1	2.8	3.7	8.1	3.5	6.1	6.8	3.5	6.3	5.7	6.1	6.1
	10	5.4	4.3	5.1	7.5	3.6	5.7	6.5	3.6	6.0	5.4	5.7	5.9
$\boldsymbol{B_2}$	5	2.8	1.9	2.9	7.5	6.4	3.5	3.6	3.3	6.2	6.2	3.6	6.5
	10	4.9	3.1	4.2	6.9	6.0	3.6	3.8	3.5	6.0	5.8	3.9	6.3
$\boldsymbol{B_3}$	5	2.9	1.8	2.7	8.6	3.3	6.0	3.1	3.2	6.5	5.8	3.5	6.3
	10	4.5	2.7	4	7.7	3.5	5.6	3.2	3.4	6.2	5.5	3.6	6.0
$\boldsymbol{B_4}$	5	3.8	2.4	3.3	7	3.5	6.7	3.4	3.3	3.5	3.4	3.5	3.5
	10	3.6	2.9	3.1	6.5	3.7	6.4	3.6	3.4	3.7	3.5	3.7	3.6

Table 6: Empirical sizes, for 5% nominal tests, based on models $B_1 - B_4$, n=500.

Model	Lags	Asymptotic distribution				MC				RWB			
		Q_{22}	Q_{22}^{w}	L_b	C_m	Q_{22}	Q_{22}^{w}	L_b	C_m	Q_{22}	Q_{22}^{w}	L_b	C_m
		Gaussian distribution											
$\boldsymbol{B_1}$	5	2.6	1.4	2.6	5.7	3.7	4.4	3.7	3.8	4.6	5.4	5.1	4.7
	10	3.0	2.1	2.8	4.3	3.8	4.7	4.0	4.0	4.8	5.2	4.9	4.8
$\boldsymbol{B_2}$	5	2.0	1.2	2.0	4.6	4.4	5.7	5.6	4.5	4.3	4.8	5.1	5.0
	10	2.7	1.9	2.6	3.8	4.5	5.3	5.2	4.6	4.6	5.0	4.8	5.2
$\boldsymbol{B_3}$	5	2.0	0.9	1.9	4.6	5.6	5.9	5.9	3.7	4.2	4.5	5.5	5.9
	10	2.7	1.7	2.2	3.7	5.3	5.6	5.5	3.9	4.3	4.6	5.3	5.6
$\boldsymbol{B_4}$	5	2.4	1.8	2.5	5.2	5.9	4.0	3.8	6.0	5.7	4.2	4.4	4.3
	10	3.5	2.3	3.2	4.0	5.7	4.2	3.9	5.7	5.5	4.5	4.6	4.6
		Student-t distribution											
$\boldsymbol{B_1}$	5	5.4	3.6	5.3	6.5	3.7	3.8	6.1	6.2	4.3	5.9	3.8	6.1
	10	5.8	4.9	5.6	6.4	4.0	4.1	5.9	5.8	4.6	5.5	4.0	5.9
$\boldsymbol{B_2}$	5	3.9	2.3	3.7	6.6	3.9	5.7	6.1	6.0	4.3	5.8	5.8	4.5
	10	5.0	3.6	4.7	6.0	4.1	5.5	5.7	5.6	4.5	5.5	5.4	4.8
$\boldsymbol{B_3}$	5	3.9	2.5	3.8	7.1	5.8	6.2	3.5	3.4	4.2	5.8	3.9	5.6
	10	5.2	3.9	4.7	6.7	5.6	5.8	3.6	3.6	4.4	5.6	4.1	5.4
$\boldsymbol{B_4}$	5	3.3	2.1	3.3	6.5	3.9	3.9	6.0	6.3	3.8	5.9	3.8	6.2
	10	3.8	3.0	3.5	5.9	4.2	4.2	5.7	6.1	3.9	5.5	4.0	5.8

Tables 4-6 show a similar results as in Tables 1-3. I noticed that as the sample size increases the test results approach to the respective significant level in both Gaussian and non-Gaussian (t-distribution) cases. Similarly, for the same sample sizes, the observed sizes approached the nominal size as lags increased from $m = 5$ to $m = 10$. The non-Gaussian (t distribution) test results were farther away from the respective significant values compared to the Gaussian test results. The test results showed that auto and cross-correlation were closer to the significant levels as compared to both McLeod-Li, weighted McLeod-Li, and Mak-Li test statistics. Mak-Li test results provides more closer results compared to McLeod-Li results. Results show that the test statistics based on Monte-Carlo (MC) technique provide better approximate to the significant levels than based on the asymptotic approach. Moreover, the Randomly Weighted Bootstrap (RWB) were more improved and closer to the significant value as compared to both Monte-Carlo (MC) and asymptotic approaches.

4.2 Power study

In this section, I compared the power of aforthmentioned tests for detecting nonlinearity in fitted linear ARMA models. In my simulation study, I generated data from Gauusian distribution, skewed normal distribution with skewness parameters $\{-2, -1.5, -1, -0.5, 0.5, 1, 1.5, 2\}$ and student's t-distribution with $\{3, 6, 9, 12\}$ degrees of freedom. The power of the tests are calculated based on the 5% significance level at lags $m = 5, 10$, and different sample sizes. For the skewed normal and student's t distributions, I caculated the power by averaging the relative rejection frequncies across the different parameters used in the corresponding models.

Testing linearity in linear time series models

For testing nonlinearity, I simulate data according to the following nonlinear models $\boldsymbol{C}_1 - \boldsymbol{C}_4$ (see Models 15-18 in Psaradakis and Vávra (2019)):

$\boldsymbol{C}_1$: Bilinear $z_t = 0.4\varepsilon_{t-1} - 0.3z_{t-1} + (0.8 + 0.5z_{t-1})\varepsilon_{t-1} + \varepsilon_t$

$\boldsymbol{C}_2$: Bilinear $z_t = 0.5\varepsilon_t - (0.4 - 0.4\varepsilon_{t-1})z_{t-1}$

$\boldsymbol{C}_3$: Nonlinear $z_t = 0.8\varepsilon_{t-2}^2 + \varepsilon_t$

$\boldsymbol{C}_4$: Nonlinear $z_t = -0.3\varepsilon_{t-1} + (0.2 + 0.4\varepsilon_{t-1} - 0.25\varepsilon_{t-2})\varepsilon_{t-2} + \varepsilon_t$

The powers of the portmanteau test ($Q_{11}, Q_{11}^{w},\ C_m$) are calculated based on the three techniques (asymptotic distribution, Monte-Carlo significance test, and random weighted Bootstrapping) when false models AR(p) are fitted to these models, where the order p is selected based on the Bayesian information criterion (BIC) from the set of orders $\{0, 1, \dots, \lfloor 8\sqrt[4]{n/100} \rfloor\}$, where $\lfloor x \rfloor$ denotes the largest intgeger less than or equal to x.

The results are shown in Tables 7-8.

Table 7: Empirical power (for 5% significant level) for testing the neglected nonlinearity in AR(p) models fitted to data generated based on the models $C_1 - C_4$, n=100.

Model	Lags	Asymptotic distribution			MC			RWB		
		Q_{11}	Q_{11}^{w}	C_m	Q_{11}	Q_{11}^{w}	C_m	Q_{11}	Q_{11}^{w}	C_m
		Gaussian distribution								
$\boldsymbol{C_1}$	5	97.4	98.5	98.0	92.0	93.0	94.0	93.0	94.0	95.0
	10	92.5	97.0	97.8	90.2	91.1	93.1	90.2	93.1	93.1
$\boldsymbol{C_2}$	5	54.8	58.7	72.0	92.0	94.0	95.0	95.0	94.0	97.0
	10	56.3	55.2	68.3	91.1	92.1	93.1	94.1	91.2	94.1
$\boldsymbol{C_3}$	5	19.5	24.1	84.5	92.0	92.0	95.0	97.0	95.0	95.0
	10	16.1	20.4	80.7	89.2	91.1	93.1	94.1	93.1	93.1
$\boldsymbol{C_4}$	5	31.2	38.3	38.0	91.0	94.0	96.0	96.0	95.0	97.0
	10	23.7	33.0	36.0	88.3	91.2	95.0	95.0	93.1	96.0

Model	Lags	Asymptotic distribution			MC			RWB		
		Q_{11}	Q_{11}^{w}	C_m	Q_{11}	Q_{11}^{w}	C_m	Q_{11}	Q_{11}^{w}	C_m
		Skewed normal distribution								
$\boldsymbol{C_1}$	5	96.8	98.0	95.9	89.0	91.0	94.0	91.0	94.0	92.0
	10	90.3	96.0	94.6	87.2	88.3	92.1	89.2	91.2	91.1
$\boldsymbol{C_2}$	5	54.2	60.9	84.5	85.0	90.0	92.0	90.0	94.0	93.0
	10	44.6	57.7	80.7	83.3	88.2	89.2	88.2	91.2	91.1
$\boldsymbol{C_3}$	5	25.0	30.4	87.1	86.0	92.0	92.0	90.0	92.0	94.0
	10	16.4	25.6	87.0	84.3	90.2	89.2	87.3	91.1	92.1
$\boldsymbol{C_4}$	5	33.5	40.2	85.0	88.0	90.0	92.0	92.0	92.0	92.0
	10	22.7	34.5	81.4	86.2	87.3	91.1	89.2	89.2	89.2

		Student t-distribution								
C_1	5	94.7	96.9	98.0	80.0	88.0	92.0	88.0	90.0	92.0
	10	88.9	93.7	96.9	79.2	87.1	91.1	85.4	89.1	89.2
C_2	5	60.3	74.3	81.2	80.0	88.0	90.0	89.0	91.0	92.0
	10	46.1	65.2	78.9	77.6	85.4	88.2	86.3	90.1	90.2
C_3	5	16.2	21.4	94.3	77.0	88.0	90.0	89.0	92.0	93.0
	10	12.6	16.5	93.0	76.2	87.1	87.3	86.3	89.2	91.1
C_4	5	37.9	44.0	47.4	83.0	86.0	92.0	87.0	91.0	93.0
	10	24.4	37.8	46.0	82.2	83.4	90.2	86.1	89.2	90.2

Table 8: Empirical power (for 5% significant level) for testing the neglected nonlinearity in AR(p) models fitted to data generated based on the models $C_1 - C_4$, n=300.

Model	**Lags**	**Asymptotic distribution**			**MC**			**RWB**		
		Q_{11}	Q_{11}^{w}	C_m	Q_{11}	Q_{11}^{w}	C_m	Q_{11}	Q_{11}^{w}	C_m
		Gaussian distribution								
C_1	5	100.0	100.0	100.0	93.0	95.0	96.0	96.0	95.0	98.0
	10	100.0	100.0	100.0	91.1	94.1	94.1	94.1	92.2	95.1
C_2	5	67.1	76.0	99.9	93.0	94.0	95.0	95.0	96.0	100.0
	10	64.3	76.2	100.0	91.1	93.1	94.1	93.1	93.1	99.0
C_3	5	59.2	67.7	100.0	94.0	94.0	96.0	94.0	95.0	99.0
	10	45.3	59.9	100.0	92.1	92.1	94.1	92.1	94.1	98.0
C_4	5	74.1	80.5	76.5	92.0	95.0	95.0	96.0	95.0	98.0
	10	63.7	75.9	76.5	91.1	94.1	92.2	94.1	93.1	96.0
		Skewed normal distribution								
C_1	5	100.0	100.0	100.0	90.0	94.0	94.0	93.0	95.0	97.0
	10	100.0	100.0	100.0	87.3	91.2	92.1	91.1	92.2	95.1
C_2	5	97.1	98.6	100.0	92.0	92.0	94.0	96.0	94.0	96.0
	10	95.0	97.5	100.0	91.1	91.1	93.1	93.1	91.2	94.1
C_3	5	66.7	74.3	100.0	92.0	94.0	95.0	97.0	94.0	97.0
	10	52.6	67.1	100.0	91.1	92.1	94.1	94.1	91.2	95.1
C_4	5	84.1	87.6	100.0	91.0	94.0	94.0	97.0	96.0	95.0
	10	75.1	84.7	100.0	89.2	93.1	93.1	95.1	95.0	92.2

Model	Lags	Asymptotic distribution			MC			RWB		
		Q_{11}	Q_{11}^{w}	C_m	Q_{11}	Q_{11}^{w}	C_m	Q_{11}	Q_{11}^{w}	C_m
		Student t-distribution								
C_1	5	100.0	100.0	100.0	85.0	90.0	94.0	92.0	92.0	94.0
	10	100.0	100.0	100.0	84.2	89.1	93.1	90.2	89.2	93.1
C_2	5	98.8	99.4	99.9	86.0	90.0	93.0	90.0	92.0	94.0
	10	96.8	99.9	99.9	85.1	87.3	92.1	89.1	90.2	91.2
C_3	5	33.0	39.1	100.0	86.0	91.0	94.0	92.0	92.0	92.0
	10	27.1	34.4	100.0	85.1	88.3	92.1	90.2	91.1	90.2
C_4	5	87.4	89.8	89.9	88.0	90.0	92.0	91.0	94.0	93.0
	10	79.7	88.1	90.0	85.4	87.3	91.1	89.2	93.1	91.1

Tables 7 and 8 show that the Fisher and Gallagher (2012), Q_{11}^{w}, test statistic almost with a power that is higher than Ljung-Box (1978), Q_{11}, and Mahdi and Fisher (2021), Cm. I have also found that for the same sample size, as the lags value increase from $m = 5$ to $m = 10$, the power gets lesser. Results also show that as the sample size increases from $n = 100$ to 300 the power value is also improved. In addition, the powers achieved by the Monte-Carlo appoach are higher than those achieved by the asymptotic method. Moreover, the nonparametric Randomly Weighted Bootstrap provided higher power in comparision with the other two methods.

In general, the power of the Mahdi and Fisher (2021) statistic based on the asymptotic distribution is the least compared with the other tests. On the other hand, the power based on Fisher and Gallagher (2012) is better than Ljung-Box (1978). In almost all cases, the Mahdi and Fisher statistic based on the Randomly Weighted Bootstrap approach attains a better results compared with the other test statistics**.**

Testing the AR-ARCH models

I examine the power of the Q_{22}, Q_{22}^{w}, L_b, and C_m portmanteau test statistics for discriminating the mean and conditional variance parts.

The powers of these test statistics are calculated based on the three techniques (asymptotic distribution, Monte-Carlo significance test, and random weighted Bootstrapping) when false models $AR(p_1) - ARCH(p_2)$ are fitted to the models $\boldsymbol{B}_1 - \boldsymbol{B}_4$:

$\boldsymbol{B_1}$: $AR(2) - ARCH(1)$: $z_t = 0.5z_{t-1} + 0.2z_{t-2} + \varepsilon_t$, $\quad \varepsilon_t = \sigma_t \varepsilon_t$, $\quad \sigma_t^2 = 0.1 + 0.4\varepsilon_{t-1}^2$,

$\boldsymbol{B_2}$: $AR(1) - ARCH(2)$: $z_t = 0.5z_{t-1} + \varepsilon_t$, $\varepsilon_t = \sigma_t\varepsilon_t$, $\sigma_t^2 = 0.1 + 0.4\varepsilon_{t-1}^2 + 0.2\varepsilon_{t-2}^2$,

$\boldsymbol{B_3}$: $AR(2) - ARCH(2)$: $z_t = 0.5z_{t-1} + 0.2z_{t-2} + \varepsilon_t$, $\varepsilon_t = \sigma_t\varepsilon_t$, $\sigma_t^2 = 0.1 + 0.4\varepsilon_{t-1}^2 + 0.2\varepsilon_{t-2}^2$,

$\boldsymbol{B_4}$: $AR(2) - GARCH(1,1)$: $z_t = 0.5z_{t-1} + 0.2z_{t-2} + \varepsilon_t$, $\varepsilon_t = \sigma_t\varepsilon_t$, $\sigma_t^2 = 0.1 + 0.4\varepsilon_{t-1}^2 + 0.5\sigma_{t-1}^2$,

where the orders p_1 and p_2 are selected based on the Bayesian information criterion (BIC). These are the models studied by Carlos Velasco and Xuexin Wang (2014) that I also include in my simulation study for estimating the empirical size of the portmanteau statistics.

The results are shown in Tables $9 - 11$.

Table 9: Empirical power (for 5% significant level) for testing the adequacy of fitted model under linear models ($B_1 - B_4$), n=100.

Model	**Lags**	**Asymptotic distribution**				**MC**				**RWB**			
		Q_{22}	Q_{22}^W	L_b	C_m	Q_{22}	Q_{22}^W	L_b	C_m	Q_{22}	Q_{22}^W	L_b	C_m
		Gaussian distribution											
$\boldsymbol{B_1}$	5	9.2	10.1	18.8	22.3	16.7	10.3	26.4	25.9	26.0	19.6	32.1	30.1
	10	8.4	9.5	16.7	20.1	11.0	11.1	19.3	28.6	13.6	19.8	20.1	36.3
$\boldsymbol{B_2}$	5	8.0	11.8	18.0	20.0	11.1	19.2	21.7	21.4	18.2	22.4	23.4	23.9
	10	8.9	13.3	10.7	22.9	13.0	19.2	11.3	32.2	18.0	27.7	18.1	33.1
$\boldsymbol{B_3}$	5	39.6	40.0	48.1	52.3	48.9	47.3	52.3	55.5	54.7	48.5	56.8	58.4
	10	38.0	39.8	40.2	50.0	40.4	48.9	49.3	51.5	49.9	51.5	57.4	59.1
$\boldsymbol{B_4}$	5	10.9	14.4	18.2	16.2	11.7	22.2	20.5	23.1	14.3	27.0	29.7	32.0
	10	9.9	10.0	15.5	17.1	11.6	16.5	24.9	18.7	13.0	19.3	27.1	28.1

		Skewed normal distribution											
$\boldsymbol{B_1}$	5	18.7	10.2	18.0	27.0	24.1	12.1	26.0	33.0	27.6	21.1	35.6	42.8
	10	16.6	16.6	15.4	23.2	20.7	20.2	21.9	26.0	29.6	29.5	24.6	27.7
$\boldsymbol{B_2}$	5	10.3	18.2	23.7	37.3	11.4	20.3	31.7	37.7	11.9	25.4	34.1	40.7
	10	10.0	20.4	25.5	33.8	20.0	24.9	33.2	43.4	24.2	25.4	41.2	47.5
$\boldsymbol{B_3}$	5	65.7	70.3	68.6	77.7	66.3	76.1	74.2	87.5	67.5	81.4	77.8	88.8
	10	60.5	66.1	65.9	73.1	68.5	74.4	72.6	76.8	74.7	79.9	80.1	85.5
$\boldsymbol{B_4}$	5	13.6	16.3	19.0	18.8	15.3	20.0	22.7	27.4	17.1	26.5	24.2	33.9
	10	11.5	14.9	17.2	19.3	13.0	15.2	19.8	27.3	21.6	20.4	29.4	34.4
		Student t-distribution											
$\boldsymbol{B_1}$	5	16.2	17.0	16.9	27.7	21.1	24.9	25.8	31.4	29.9	29.7	32.5	33.3
	10	14.6	17.7	13.5	24.0	18.1	25.8	20.2	32.5	25.8	30.1	27.1	42.2
$\boldsymbol{B_2}$	5	38.2	25.8	26.6	37.4	43.9	26.0	26.8	44.9	45.5	26.6	31.2	48.3
	10	34.4	27.9	23.9	34.9	41.6	30.4	30.4	36.3	47.6	32.5	34.8	37.2
$\boldsymbol{B_3}$	5	56.1	67.0	56.1	77.8	63.3	76.8	56.4	80.3	67.1	85.7	65.7	84.1
	10	44.3	60.8	53.6	74.9	46.7	69.1	61.9	79.5	50.6	73.8	66.1	83.0
$\boldsymbol{B_4}$	5	15.7	19.9	21.0	20.2	17.1	24.9	30.9	28.7	20.2	29.7	35.8	34.6
	10	13.7	18.5	22.3	22.0	15.4	26.1	28.4	23.4	15.4	31.6	33.0	25.6

Table 10: Empirical power (for 5% significant level) for testing the adequacy of fitted model under linear models ($B_1 - B_4$), n=300.

Model	**Lags**	**Asymptotic distribution**				**MC**				**RWB**			
		Q_{22}	Q_{22}^{w}	L_b	C_m	Q_{22}	Q_{22}^{w}	L_b	C_m	Q_{22}	Q_{22}^{w}	L_b	C_m
		Gaussian distribution											
$\boldsymbol{B_1}$	5	22.9	29.4	20.1	45.1	32.6	31.2	26.2	49.8	41.5	39.6	34.2	52.5
	10	19.2	23.4	15.9	47.4	21.1	28.7	15.9	49.4	23.2	30.4	21.1	56.6
$\boldsymbol{B_2}$	5	20.1	22.8	19.1	25.5	24.6	28.1	24.9	35.1	31.7	37.3	28.8	36.7
	10	16.3	21.9	14.7	27.3	23.7	31.7	23.1	31.5	28.6	36.2	31.6	33.1
$\boldsymbol{B_3}$	5	56.6	51.2	68.1	72.3	64.4	57.0	74.1	80.5	65.2	66.0	77.4	86.4
	10	58.6	48.8	63.3	77	67.7	53.9	71.8	80.2	73.1	63.9	81.3	80.7
$\boldsymbol{B_4}$	5	17.3	18.6	24.8	23.8	22.2	26.3	34.1	27.0	24.8	29.1	42.3	27.9
	10	16.2	19.2	27.3	24.7	24.1	21.9	34.6	34.3	33.0	26.1	36.3	40.9

		Skewed normal distribution											
$\boldsymbol{B_1}$	5	37.6	25.6	25.5	53.9	43.0	26.2	29.8	58.6	44.8	30.2	30.2	63.6
	10	35.1	26	23.9	54.9	43.1	33.7	30.6	60.2	45.1	37.4	36.9	62.2
$\boldsymbol{B_2}$	5	32.3	20.2	24.7	50	40.3	21.7	34.7	55.4	46.3	21.7	44.4	56.6
	10	30.9	26.6	22.7	48.5	33.1	26.9	31.1	54.5	40.2	27.9	33.7	64.4
$\boldsymbol{B_3}$	5	85.3	80.3	78	87.3	89.9	86.5	80.5	95.2	95.1	94.8	87.1	104.7
	10	82.2	78.9	75.7	86.6	90.1	80.9	83.3	96.3	90.8	83.6	91.4	99.1
$\boldsymbol{B_4}$	5	25.3	27.7	30.7	31	34.5	29.3	37.4	35.1	36.3	29.7	40.5	39.0
	10	25.6	24.3	32.9	30.2	26.5	32.6	34.3	31.5	27.4	38.6	38.7	33.7
		Student t-distribution											
$\boldsymbol{B_1}$	5	28.2	30.2	30.8	59	29.9	31.9	33.3	65.7	34.8	33.5	38.3	66.2
	10	26.9	28	31.4	58.8	28.5	32.5	31.6	61.4	36.0	33.8	39.2	66.4
$\boldsymbol{B_2}$	5	78.2	65.8	66.6	77.4	79.5	70.8	76.0	84.2	80.1	80.2	85.4	93.9
	10	74.4	67.9	63.9	74.9	78.2	76.7	67.1	81.1	87.7	77.1	71.9	82.5
$\boldsymbol{B_3}$	5	60.1	67.6	76.9	88.8	68.1	74.1	77.6	98.2	69.6	79.7	82.2	101.8
	10	69.7	67.2	68.4	84	72.5	76.0	77.2	88.4	73.9	80.0	84.3	90.2
$\boldsymbol{B_4}$	5	28	29.6	36.4	33.9	35.5	37.3	45.7	43.5	36.8	43.8	51.3	51.7
	10	26.9	30.4	34.7	32.7	27.6	36.6	43.8	42.1	32.6	41.1	44.5	51.4

Table 11: Empirical power (for 5% significant level) for testing the adequacy of fitted model under linear models ($B_1 - B_4$), n=500.

Model	Lags	Asymptotic distribution				MC				RWB			
		Q_{22}	Q_{22}^{w}	L_b	C_m	Q_{22}	Q_{22}^{w}	L_b	C_m	Q_{22}	Q_{22}^{w}	L_b	C_m
		Gaussian distribution											
$\boldsymbol{B_1}$	5	33.3	39.6	39	55.9	33.8	45.2	46.2	56.5	43.5	51.0	47.1	62.7
	10	29.8	43	36.8	53.8	31.0	50.8	37.1	63.8	35.0	59.8	45.5	65.0
$\boldsymbol{B_2}$	5	33.3	39.6	39	55.9	33.5	49.6	39.0	58.7	37.0	52.9	44.5	63.8
	10	29.8	43	36.8	53.8	34.9	44.1	40.5	60.7	35.9	54.0	43.5	61.5
$\boldsymbol{B_3}$	5	38.1	35.9	33	87.8	44.0	42.3	40.1	94.7	48.6	48.9	50.0	100.0
	10	34	32.1	28.5	90.1	38.5	34.4	36.5	98.5	42.3	40.5	46.0	100.0
$\boldsymbol{B_4}$	5	28	29.6	35.4	34.5	29.3	30.3	38.5	39.8	37.8	40.1	46.0	40.6
	10	27.8	31.6	34.5	35.7	32.1	36.7	38.2	42.7	34.4	42.3	42.5	43.3

		Skewed normal distribution											
B_1	5	40.1	45.6	35.1	63.6	41.5	53.4	40.3	72.5	49.3	61.6	48.8	76.9
	10	46.3	42.1	33.3	64.2	47.7	52.0	35.4	68.3	56.6	58.4	43.2	77.4
B_2	5	37.6	38.7	33.2	58.9	41.1	48.1	40.5	67.1	44.7	55.7	48.2	68.6
	10	36.9	39.9	30.8	61.1	46.9	44.6	32.9	70.1	48.0	51.1	42.7	71.8
B_3	5	89.7	80.8	88.9	98.9	97.4	88.5	94.9	100.0	100.0	93.3	100.0	100.0
	10	86.7	88.8	85.9	95.9	90.3	89.8	95.5	100.0	98.6	91.0	100.0	100.0
B_4	5	30	39.9	41.5	40.2	35.5	46.0	45.4	45.7	38.8	48.2	46.7	49.8
	10	29.9	40.9	38.3	41.7	30.3	44.5	45.8	49.5	40.1	50.2	46.2	58.4
		Student t-distribution											
B_1	5	39.2	41.2	33.8	70.6	43.1	50.9	37.8	74.6	51.6	51.8	39.6	76.1
	10	36.9	48	31.4	70	42.2	53.7	37.7	77.5	51.9	56.3	44.9	82.0
B_2	5	48.1	49.9	44.4	68.8	54.1	55.1	51.4	74.5	61.6	63.6	61.1	80.5
	10	43.9	53.3	47	69.9	52.7	62.2	53.6	73.0	56.4	70.1	58.8	76.1
B_3	5	89.5	89.1	89.9	100	99.0	95.7	91.6	100.0	100.0	95.8	99.0	100.0
	10	91.1	80.9	88.9	100	92.7	88.3	90.1	100.0	94.9	98.1	92.5	100.0
B_4	5	34.7	40	42.6	42.2	43.4	46.6	49.0	45.9	49.3	48.9	50.7	49.7
	10	33.9	38.2	41.7	40.5	41.3	44.9	45.1	42.8	48.9	51.1	45.7	45.8

From the Tables 9, 10, and 11 it was observed that the in most of the cases Mahdi and Fisher (2021), Cm, results were higher than both Ljung-Box (1978), Q_{11} and Li-Mak (1994), $L_{b.}$ It's also observed that Random Weighted Bootstrap (RWB) tests results were higher than both, Asymptotic and Monte Carlo Technique results. Monte Carlo results were found out to be higher than Asymptotic results. All these observations were for all the sample sizes, i.e., n = 100, 300, and 500. Also it is noted that the higher sample sizes yielded the higher power whether using Asymptotic, Monte Carlo or Random Weighted Bootstrap (RWB) tests approach.

CHAPTER 5: APPLICATION

In the section I applied the aforementioned test statistics and on a real data. Real data represented the log returns of the Qatar Ntional Bank (QNB) from January 2/2019 through December 30/2020 which involves data before and after Covid-19 pandemic. The log-returns data, which was calculated as the difference of log natural of the current value and the log natural of the previous value. When we compare two different data then we can use either ratio or differences. In Time-series we use ratios. When the ratios of two quantities are calculated then it is either grater than 1, 1, or between 0 and 1 (inclusive). So when the logs are calculated of the ratios then if the ratio is greater than 1 then its log value will be positive but can be very smaller than the actual ratio itself, if the ratio equals 1 then its log value will be 0 and negative otherwise. It is also convenient to plot the log ratios as compared to the actual values. The data I used can be obtaimed from the online web source *https://finance.yahoo.com/quote/QNBK.QA?p=QNBK.QA*.

Table 12: The Descriptive analysis of QNB log-returns

Mean	Median	Var	Std. Dev	Min	Max	Range	Skewness	Kurtosis
-0.00013	*0.0000*	*0.0002*	*0.0153*	*-0.1053*	*0.0607*	*0.1660*	*-0.6261*	*6.8611*

I first calculated the basic descriptive statistics of the log-returns data and results are given in Table 12 above. The log-returns values were negatively skewed with the value of skewness as -0.6261. Regarding the peakedness, the kurtosis score 6.8611 which is larger than 3. The graph also shows that the log returns are skewed to the left and it has an excess Kurtosis. There is an extreme negative spike in March 2020 which is the representation of the beginning Coronavirus pandemic (COVID-19). These results with the histogram graph shown in Figure 1 a suggests that the log-returns has a leptokurtic distribution (not Gausian).

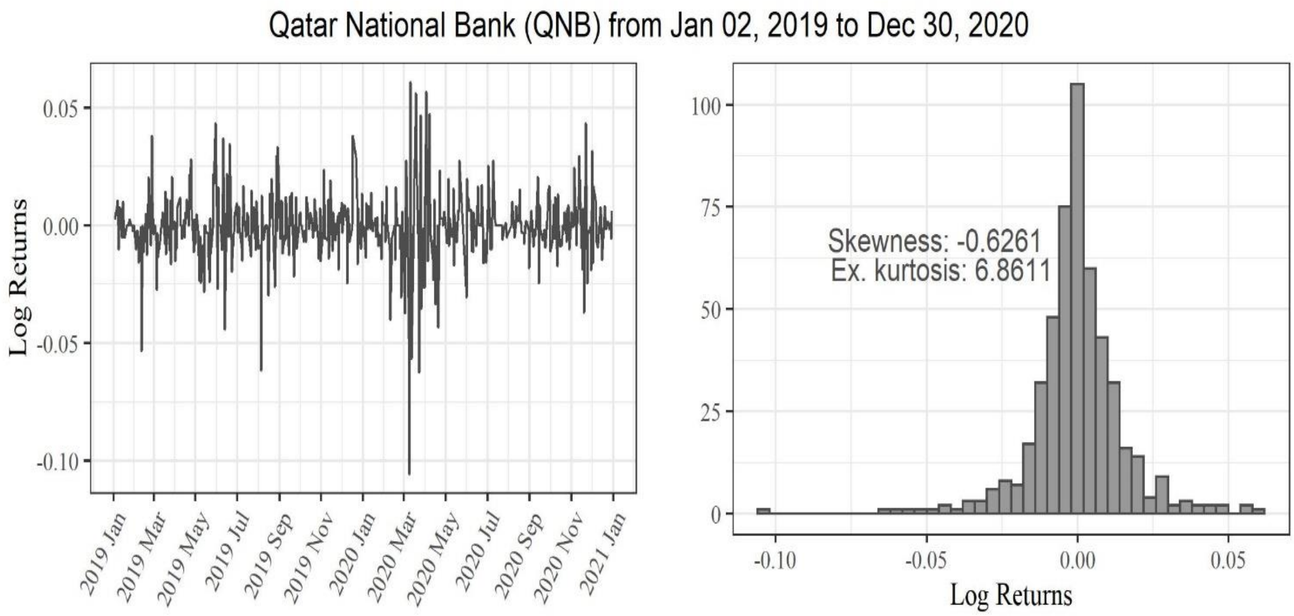

Figure 1: Qatar National Bank (QNB) log-returns distribution and long returns histogram.

The Kolmogorov-Smirnov and Shapiro-Wilk tests have significance values. In both cases the significance values are less than 0.05. The Shapiro-Wilk normality test for the log-returns data is W = 0.9024 with a p-value almost zero. Also, the Q-Q plot of log-returns shows the values are deviated away from the normal line and I can conclude that the log-returns were skewed (see Figure 2).

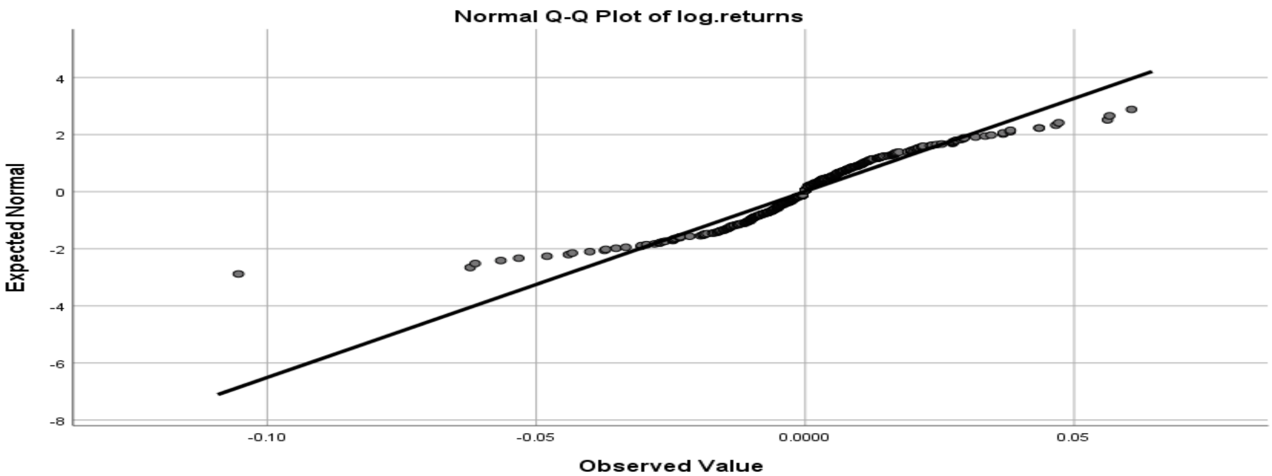

Figure 2: Q-Q plot of Qatar national Bank (QNB) log-returns.

I first fitted several ARIMA models and select the best model based on the Bayesian Information Criterion (BIC). I found the best model is the Moving Average of order 2, MA(2),

$$z_t = \varepsilon_t - 0.0024\varepsilon_{t-1} - 0.0972\varepsilon_{t-2}.$$

I then tested the adequacy in the fitted model using the test statistics of Ljung-Box (Q_{11}) and Fisher-Gallagher (Q_{11}^{W}), and Mahdi and Fisher (C_m) based on the asymptotic distribution, Monte-Carlo, and Randomly Weighted Bootstrap approaches. The results are given in Table 13.

.

Lags	Asymptotic distribution			MC			RWB		
	Q_{11}	Q_{11}^{w}	C_m	Q_{11}	Q_{11}^{w}	C_m	Q_{11}	Q_{11}^{w}	C_m
5	0.84	1.00	8.19×10^{-8}	0.98	0.99	7.59×10^{-9}	0.88	0.81	7.04×10^{-8}
10	0.08	0.66	9.250378×10^{-13}	0.1828	0.58	8.12×10^{-12}	0.12	0.33	5.11×10^{-14}

Table 13: P-value – test adequacy of fitted MA(2) model by checking the dependency in the residuals

I found that in asymptotic distribution using the Q_{11} and Q_{11}^{w} suggest that the MA(2) model is good as all the p-values are high. On the other hand, when I tested using the C_m, the model is not good as the p-vales are approximately equal to zero. I found that the Monte-Carlo and the Randomly Weighted Bootstrap still do not catch the inadequacy using the Q_{11} and Q_{11}^{w} as the p-values still are very high. On the other hand, the Randomly Weighted Bootstrap of the test statistic C_m suggests that there is an inadequacy in the fitted model.

The autocorrelation function (ACF) and partial autocorrelation function (PACF) in Figure 3 suggest that the log-returns might exhibit an autoregressive conditional heteroskedasticity (ARCH) and that MA(2) might not be a good model to fit the data. Thus, I applied the aforementioned tests on the squared residuals and test for ARCH effect. All tests suggest that the ARMA is not a good model to fit the data due to ARCH effect as all p-values less than 5%.

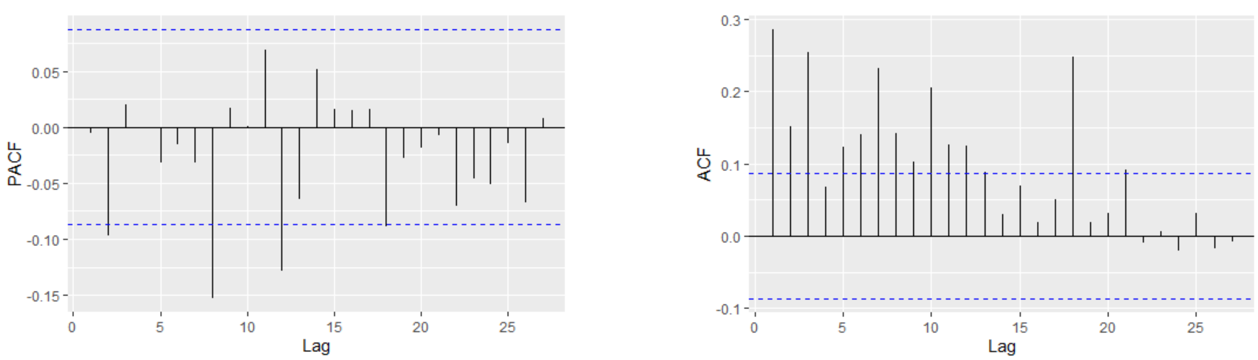

Figure 3: Patrtial auto correlation function (PACF) and auto correlation function (ACF) of square residuals for Qatar National Bank (QNB) data log-returns.

After that, I fitted some $\mathrm{ARMA}-\mathrm{ARCH}$ models and select that best based on the BIC criteria. The BIC suggest that the $\mathrm{MA}(1)-\mathrm{ARCH}(1)$ model might be a good candidate to fit the data. The selected $\mathrm{MA}(1)-\mathrm{ARCH}(1)$ is

$$z_t = -0.00618\varepsilon_t,$$

$$\text{where } \ \varepsilon_t = \sigma_t \xi_t, \quad \sigma_t^2 = 0.05\varepsilon_{t-1}^2.$$

I applied the test statistics McLeod-Li (1983), Q_{22}, Fisher and Gallagher (2012), Q_{22}^{W}, Li-Mak (1994), L_b, and Mahdi and Fisher (2021), C_m on the squared residuals of the $\mathrm{MA}(1)-\mathrm{ARCH}(1)$ model.

Table 14: P-value – test adequacy of fitted MA(1)-ARCH(1) model.

Lags	**Asymptotic distribution**				**MC**				**RWB**			
	Q_{22}	Q_{22}^{w}	L_b	C_m	Q_{22}	Q_{22}^{w}	L_b	C_m	Q_{22}	Q_{22}^{w}	L_b	C_m
5	0.86	0.70	0.97	2.5×10^{-6}	0.86	0.71	0.86	2.2×10^{-6}	0.85	0.71	0.68	1.9×10^{-6}
10	0.84	0.91	0.89	1.1×10^{-6}	0.84	0.92	0.72	9.1×10^{-7}	0.82	0.93	0.62	8.2×10^{-7}

Table 14 shows that the portmanteau tests based on $Q_{22}, Q_{22}^{W}, L_b,$ and C_m using both parametric and non parametric approaches. The test statistics of Q_{22} , Q_{22}^{W}, and $L_b$based on the three approaches, the asymptotic distribution, Monte-Carlo, and Randomly Weighted Bootstrap methods suggest that the model is good, whereas the $C_m$suggests that the model is not a good fit.

Lastly, I tried several models to improve the fit and obtained the final model as $\mathrm{GARCH}(3{,}3)$:

$$\sigma_t^2 = 0.00004 + 0.14262\varepsilon_{t-1}^2 + 0.33143\varepsilon_{t-2}^2 + 0.03266\varepsilon_{t-3}^2 + \ 0.000001\sigma_{t-1}^2 + 0.36862\,\sigma_{t-3}^2$$

I tested the adequacy of this fitted model using the $Q_{22}, Q_{22}^{W}, L_b,$ and C_m based on the three parametric and non parametric methods.

The results in Table 15, show that the protmanteau tests based on parametric and nonparametric approaches suggest that the fitted GARCH (3,3) model is found out to be the best.

Table 15: P-value – test adequacy of fitted GARCH(3,3) model.

Lags	Asymptotic distribution				MC				RWB			
	Q_{22}	Q_{22}^{w}	L_b	C_m	Q_{22}	Q_{22}^{w}	L_b	C_m	Q_{22}	Q_{22}^{w}	L_b	C_m
5	0.77	0.80	0.42	0.05	0.76	0.78	0.42	0.05	0.72	0.77	0.42	0.05
10	0.94	0.90	0.85	0.07	0.92	0.87	0.84	0.07	0.91	0.83	0.81	0.07

CONCLUSION

In this thesis, I conducted a comparison study between the parametric and nonparametric portmanteau tests that are used in linear and nonlinear time series models. The parametric portmanteau tests are used based on the asymptotic distributions, whereas the nonparametric portmanteau tests are used based on the Monte-Carlo and Randomly Weighted Bootstrap (RWB) approaches.

To estimate the size of the test statistics, I generated data based on 1000 replications using different ARMA and GARCH models. For the Monte-Carlo case, I replicated each simulation 500 times. In each case, I fitted the true model and calculated the Type-I error at lags 5 and 10 with sample sizes n=100,300, and 500 to cover small, moderate, and large sample sizes based on Gaussian and non-Gaussian distributions. For the non Gaussian distribution, I considered the t-distribution for degrees of freedom $\{3,6,9,12\}$, the skewed normal for skewness $\{-2, -1.5, -1, -0.5, 0.5, 1, 1.5, 2\}$, and skewed t-distribution for skewness $\{-2, -1.5, -1, -0.5, 0.5, 1, 1.5, 2\}$ and degrees of freedom $\{3,6,9,12\}$. The skewness and degrees of freedom parameters were used to cover a range of values that are sufficiently representative of mild asymmetry and heavy-tailed leptokurtosis distributions in several financial time series.

I found that as the sample size increases (for samples sizes increases from 100 through 300 to 500) the test values get closer to the respective significant level. Similarly, Similarly, for the same sample sizes, the observed sizes approached the nominal size as lags increased from $m = 5$ to $m = 10$.. As compared to Ljung and Box (1978), McLeod and Li (1983), Fisher and Gallagher (2012), Li and Mak (1994), and Mahdi and Fisher (2021), the nonparametric Monte-Carlo and Randomly Weighted Bootstrap approachs provide more accurate results compared with the parametric approach.

I also compared the power of the aforementioned tests in detecting the absence of linearity and nonlinearity assumption as well as in testing for the adequacy in fitted ARMA and GARCH models. In each case, I fitted a false model and then I calculated the relative frequency of the number of rejections. In almost all scenarios, I found that the test statistics based on the asymptotic distribution (parametric approach) provided the least power compared with the nonparametric Monte-Carlo and Randomly Weighted Bootstrap approaches. In all cases, my simulation results suggested that the Randomly Weighted Bootstrap approach provided the higher power, and Mahdi and Fisher test is almost always provided the higher power. Unlike the Emperical Size tests, for the same sample sizes, the observed power decreases as lags increased from $m = 5$ to $m = 10$.

I finished this thesis by applying these techniques on real-world data using the log-returns of Qatar National Band, QNB, data, from January 2/2019 through December 30/2020 was used which involves data before and after the COVID-19 pandemic. I found that all tests based on all methods (asymptotic distribution, Monte-Carlo, and Randomly Weighted Bootstrap) reject the null hypothesis that the ARMA model can fit the data accurately.

After that, I fitted some Garch models and found that the $\mathrm{GARCH}(3,3)$ model with a skewed t-distribution is a reasonable model that can fit the log-returns data well. My final model was:

$$\sigma_t^2 = 0.00004 + 0.14262\varepsilon_{t-1}^2 + 0.33143\varepsilon_{t-2}^2 + 0.03266\varepsilon_{t-3}^2 + 0.000001\sigma_{t-1}^2 + 0.36862\,\sigma_{t-3}^2$$

Then, I applied the McLeod-Li, Fisher-Gallagher, Li- Mak, and Mahdi-Fisher on the squared residuals of this fitted model using the parametric and nonparametric methods. I found all test statistics have failed to detect inadequacy of this fitted model, except the Mahdi-Fisher statistic based on the Randomly Weighted Bootstrap. This suggests that the GARCH(3,3) model might be inaccurate and need to be improved.

I concluded that the nonparametric portmanteau tests based on Bootstrapping and Monte-Carlo techniques are recommended to be used in diagnosis checks of the time series model.

I focused my attention on studying the portmanteau tests that can be used with ARMA/GARCH models, but one can extend this study to test for goodness-of-fit in seasonal ARMA/GARCH models (see Mahdi (2016)). I would suggest some studies to be done on observing how the power changes as emperical size changes from 1% to 5%.

Appendix

Tables 16-21 below show thae results of the empirical Type I errors based on the models $\boldsymbol{A_1 - A_4}$ and $\boldsymbol{B_1 - B_4}$ by averaging of the relative rejection frequncies across the Skewed Normal and Students' t distribution with skewness $\{-2, -1.5, -1, -0.5, 0.5, 1, 1.5, 2\}$ and degress of freedom $\{3,6,9,12\}$. In each case, I calculate the empirical size corresponding to a nominal size of 1%.

Table 16: Empirical sizes, for 1% nominal tests, based on models $A_1 - A_4$, n=100.

Model	**Lags**	**Asymptotic distribution**			**MC**			**RWB**		
		Q_{11}	Q_{11}^{w}	C_m	Q_{11}	Q_{11}^{w}	C_m	Q_{11}	Q_{11}^{w}	C_m
		Gaussian distribution								
$\boldsymbol{A_1}$	5	1.9	2.7	1.9	1.4	1.2	1.2	0.7	1.3	0.8
	10	1.7	1.9	1.2	1.3	1.1	1.2	0.8	1.2	0.8
$\boldsymbol{A_2}$	5	1.1	2.7	1.7	0.7	1.3	1.3	0.8	0.7	0.8
	10	0.8	0.6	0.9	0.7	1.2	1.2	0.8	0.7	0.8
$\boldsymbol{A_3}$	5	2.2	3.1	1.9	1.2	1.3	0.7	0.7	1.2	1.2
	10	2.3	2.2	1.4	1.2	1.2	0.7	0.8	1.1	1.2
$\boldsymbol{A_4}$	5	1.5	3.0	2.0	0.6	0.7	0.6	0.7	0.7	1.2
	10	0.8	0.6	1.1	0.6	0.7	0.6	0.8	0.8	1.2
		Skewed normal distribution								
$\boldsymbol{A_1}$	5	0.7	1.1	2.0	1.5	1.3	1.3	0.8	1.3	0.8
	10	3.2	2.8	2.8	1.4	1.2	1.3	0.8	1.3	0.8
$\boldsymbol{A_2}$	5	1.1	0.6	1.8	0.7	1.4	1.4	0.8	0.7	0.8
	10	1.3	1.5	2.2	0.8	1.3	1.3	0.8	0.8	0.9
$\boldsymbol{A_3}$	5	1.0	0.6	2.6	1.3	1.3	0.7	0.8	1.2	1.3
	10	1.6	1.4	1.8	1.3	1.3	0.8	0.8	1.2	1.2
$\boldsymbol{A_4}$	5	2.5	2.6	1.4	0.6	0.7	0.7	0.8	0.8	1.3
	10	2.5	2.5	1.4	0.7	0.7	0.7	0.8	0.8	1.2
		Student t-distribution								
$\boldsymbol{A_1}$	5	0.7	1.1	2.7	1.6	1.4	1.4	0.8	1.4	0.9
	10	3.3	2.2	3.4	1.5	1.3	1.3	0.9	1.4	0.9
$\boldsymbol{A_2}$	5	2.5	2.3	0.3	0.8	1.4	1.5	0.9	0.8	0.9
	10	1.1	2.2	1.9	0.9	1.4	1.5	0.9	0.9	0.9
$\boldsymbol{A_3}$	5	4.5	2.1	2.1	1.4	1.4	0.7	0.9	1.3	1.4
	10	4.7	1.1	2.4	1.4	1.4	0.8	0.9	1.3	1.3
$\boldsymbol{A_4}$	5	3.3	0.6	4.8	0.7	0.8	0.7	0.9	0.9	1.4
	10	2.4	3.3	2.1	0.7	0.8	0.7	0.9	0.9	1.3

Table 17: Empirical sizes, for 1% nominal tests, based on models $A_1 - A_4$, n=300.

Model	**Lags**	**Asymptotic distribution**			**MC**			**RWB**		
		Q_{11}	Q_{11}^{w}	C_m	Q_{11}	Q_{11}^{w}	C_m	Q_{11}	Q_{11}^{w}	C_m
		Gaussian distribution								
A_1	5	0.7	1.0	1.5	1.3	1.2	1.2	0.7	1.2	0.7
	10	1.2	0.8	1.6	1.3	1.1	1.1	0.7	1.2	0.8
A_2	5	0.6	2.0	2.3	0.7	1.3	1.3	0.7	0.6	0.8
	10	1.5	0.6	1.7	0.7	1.1	1.2	0.8	0.7	0.8
A_3	5	1.0	1.4	2.5	1.2	1.2	0.6	0.7	1.2	1.2
	10	1.5	1.1	1.8	1.1	1.2	0.7	0.8	1.1	1.1
A_4	5	1.1	2.6	2.2	0.6	0.6	0.6	0.7	0.7	1.2
	10	1.9	1.0	1.8	0.6	0.7	0.6	0.8	0.7	1.2
		Skewed normal distribution								
A_1	5	1.2	1.7	1.8	1.4	1.3	1.3	0.7	1.4	0.8
	10	3.0	2.6	2.6	1.4	1.2	1.2	0.8	1.3	0.9
A_2	5	1.2	0.6	1.9	0.7	1.4	1.4	0.8	0.7	0.8
	10	1.3	1.5	2.3	0.8	1.2	1.3	0.8	0.8	0.8
A_3	5	1.0	0.6	2.7	1.3	1.3	0.7	0.8	1.2	1.4
	10	1.6	1.5	1.8	1.2	1.2	0.7	0.8	1.2	1.2
A_4	5	2.5	2.7	1.5	0.6	0.7	0.6	0.8	0.8	1.3
	10	2.6	2.6	1.5	0.7	0.7	0.7	0.8	0.8	1.3
		Student t-distribution								
A_1	5	0.9	0.9	3.3	1.5	1.3	1.4	0.8	1.5	0.8
	10	4.1	2.4	4.2	1.5	1.2	1.3	0.8	1.4	0.9
A_2	5	2.5	2.4	0.3	0.8	1.5	1.5	0.9	0.7	0.9
	10	1.1	2.2	2.0	0.8	1.3	1.4	0.9	0.8	0.9
A_3	5	4.7	2.1	2.2	1.4	1.4	0.7	0.8	1.3	1.4
	10	4.9	1.0	2.4	1.3	1.3	0.8	0.9	1.3	1.3
A_4	5	3.4	0.6	5.0	0.7	0.7	0.7	0.8	0.8	1.4
	10	2.4	3.4	2.1	0.7	0.8	0.7	0.9	0.8	1.4

Table 18: Empirical sizes, for 1% nominal tests, based on models $A_1 - A_4$, n=500.

Model	**Lags**	**Asymptotic distribution**			**MC**			**RWB**		
		Q_{11}	Q_{11}^{w}	C_m	Q_{11}	Q_{11}^{w}	C_m	Q_{11}	Q_{11}^{w}	C_m
		Gaussian distribution								
A_1	5	0.7	1.1	1.6	1.3	1.1	1.2	0.7	1.2	0.7
	10	1.2	0.8	1.4	1.3	1.1	1.1	0.7	1.2	0.8
A_2	5	0.7	1	2	0.7	1.2	1.3	0.7	0.6	0.8
	10	1.2	0.8	1.6	0.7	1.1	1.2	0.8	0.7	0.8
A_3	5	1.1	1.3	2.2	1.2	1.2	0.6	0.7	1.1	1.2
	10	1.5	1.1	1.6	1.1	1.1	0.7	0.7	1.1	1.1
A_4	5	1.1	2.4	1.9	0.6	0.6	0.6	0.7	0.7	1.2
	10	1.8	1	1.5	0.6	0.7	0.6	0.7	0.7	1.1
		Skewed normal distribution								
A_1	5	0.5	1.3	2.5	1.4	1.2	1.3	0.7	1.3	0.8
	10	3.2	3	3.4	1.3	1.2	1.2	0.8	1.3	0.8
A_2	5	0.8	1.0	1.0	0.7	1.3	1.4	0.8	0.7	0.8
	10	1.7	1.8	1.1	0.8	1.2	1.3	0.8	0.7	0.8
A_3	5	1.5	2.5	1.4	1.3	1.3	0.7	0.8	1.2	1.3
	10	1.5	1.6	1.3	1.2	1.3	0.7	0.8	1.2	1.2
A_4	5	1.8	0.8	0.7	0.6	0.7	0.6	0.8	0.7	1.3
	10	1.7	0.7	0.7	0.7	0.7	0.7	0.8	0.8	1.2
		Student t-distribution								
A_1	5	1	1.7	2.8	1.5	1.3	1.4	0.8	1.4	0.8
	10	3.6	3.7	4	1.4	1.3	1.3	0.8	1.4	0.9
A_2	5	0.5	0.8	1.2	0.7	1.4	1.4	0.8	0.7	0.9
	10	2.0	1.2	1.3	0.8	1.3	1.3	0.9	0.8	0.9
A_3	5	2.1	1.8	1.4	1.3	1.4	0.7	0.8	1.3	1.4
	10	1.4	1.6	1.4	1.3	1.3	0.8	0.8	1.3	1.3
A_4	5	1.9	1.0	0.8	0.7	0.8	0.7	0.9	0.8	1.3
	10	2.2	1.9	0.8	0.7	0.7	0.7	0.9	0.8	1.3

Table 19: Empirical sizes, for 1% nominal tests, based on models $B_1 - B_4$*, n=100.*

Model	Lags	Asymptotic distribution				MC				RWB			
		Q_{22}	Q_{22}^{W}	L_b	C_m	Q_{22}	Q_{22}^{W}	L_b	C_m	Q_{22}	Q_{22}^{W}	L_b	C_m
		Gaussian distribution											
$\boldsymbol{B_1}$	5	0.6	0.1	0.3	1.0	1.2	0.7	1.3	1.3	1.3	0.8	1.2	0.7
	10	1.0	0.4	0.8	0.6	1.1	0.7	1.2	1.2	1.2	0.8	1.1	0.7
$\boldsymbol{B_2}$	5	0.5	0.2	0.4	0.9	1.3	1.3	0.6	1.3	1.1	0.7	0.7	0.7
	10	1.0	0.4	0.5	0.6	1.3	1.2	0.7	1.3	1.1	0.7	0.7	0.8
$\boldsymbol{B_3}$	5	0.6	0.2	0.4	0.7	0.6	1.2	1.3	1.3	1.1	0.7	0.8	0.7
	10	0.8	0.4	0.6	0.6	0.6	1.1	1.3	1.2	1.1	0.8	0.8	0.7
$\boldsymbol{B_4}$	5	0.5	0.5	0.5	1.0	1.2	0.7	1.2	0.6	1.2	0.7	0.7	0.8
	10	1.3	0.6	1.1	0.5	1.2	0.7	1.2	0.6	1.1	0.7	0.7	0.8
		Skewed normal distribution											
$\boldsymbol{B_1}$	5	0.4	0.3	0.4	1.8	1.0	0.5	1.5	1.5	1.4	0.9	1.0	0.6
	10	0.5	0.5	0.4	1.6	1.0	0.6	1.5	1.0	1.1	0.7	0.9	0.6
$\boldsymbol{B_2}$	5	0.3	0.1	0.3	1.8	1.1	1.5	0.7	1.5	1.3	0.6	0.6	0.6
	10	0.4	0.3	0.3	1.3	1.1	1.0	0.8	1.1	1.2	0.6	0.9	0.9
$\boldsymbol{B_3}$	5	0.3	0.1	0.3	1.7	0.7	1.1	1.1	1.0	1.3	0.8	0.6	0.6
	10	0.4	0.3	0.3	1.7	0.7	1.0	1.1	1.4	1.3	0.7	1.0	0.6
$\boldsymbol{B_4}$	5	0.4	0.3	0.3	1.4	1.4	0.6	1.4	0.5	1.4	0.6	0.6	0.9
	10	0.5	0.3	0.4	1.2	1.0	0.8	1.3	0.7	1.3	0.6	0.6	0.9
		Student t-distribution											
$\boldsymbol{B_1}$	5	0.5	0.2	0.5	2.2	1.1	0.4	1.2	1.7	1.2	1.0	1.2	0.5
	10	0.5	0.4	0.3	1.4	0.9	0.7	1.6	1.1	1.2	0.6	0.7	0.7
$\boldsymbol{B_2}$	5	0.4	0.2	0.4	2.0	1.2	1.7	0.8	1.3	1.2	0.5	0.7	0.7
	10	0.3	0.3	0.4	1.1	0.9	0.9	0.6	1.3	1.0	0.7	0.7	1.0
$\boldsymbol{B_3}$	5	0.3	0.1	0.3	1.9	0.8	1.3	1.3	1.2	1.1	0.7	0.5	0.7
	10	0.4	0.3	0.3	1.4	0.6	0.8	1.2	1.2	1.4	0.7	0.8	0.5
$\boldsymbol{B_4}$	5	0.4	0.2	0.4	1.5	1.2	0.5	1.5	0.4	1.2	0.7	0.5	1.1
	10	0.6	0.4	0.5	1.4	0.8	0.7	1.1	0.8	1.2	0.5	0.5	1.0

Table 20: Empirical sizes, for 1% nominal tests, based on models $B_1 - B_4$, n=300.

Model	Lags	Asymptotic distribution				MC				RWB			
		Q_{22}	Q_{22}^{w}	L_b	C_m	Q_{22}	Q_{22}^{w}	L_b	C_m	Q_{22}	Q_{22}^{w}	L_b	C_m
		Gaussian distribution											
$\boldsymbol{B_1}$	5	0.5	0.2	0.5	1.7	1.2	1.1	0.8	1.3	1.1	1.2	0.9	0.8
	10	0.4	0.5	0.3	1.1	1.1	1.1	0.8	1.2	1.1	1.1	0.9	0.9
$\boldsymbol{B_2}$	5	0.3	0.2	0.3	1.9	0.8	0.7	0.7	1.2	1.1	0.9	1.2	0.8
	10	0.2	0.2	0.2	1.1	0.8	0.7	0.8	1.2	1.1	0.9	1.1	0.8
$\boldsymbol{B_3}$	5	0.3	0.2	0.3	1.6	1.2	1.2	0.7	0.7	0.9	0.8	1.2	0.8
	10	0.2	0.3	0.4	1	1.2	1.1	0.7	0.7	0.9	0.8	1.2	0.8
$\boldsymbol{B_4}$	5	0.6	0.4	0.6	1.6	1.1	0.7	1.2	0.8	1.1	1.2	0.7	1.1
	10	0.4	0.3	0.4	1.2	1.1	0.7	1.2	0.8	1.0	1.1	0.8	1.0
		Skewed normal distribution											
$\boldsymbol{B_1}$	5	0.7	0.7	0.6	1.4	1.1	1.4	0.8	0.8	0.9	1.1	0.6	0.6
	10	1.3	0.7	0.9	1.8	1.4	0.9	0.8	0.6	0.9	0.9	0.9	0.9
$\boldsymbol{B_2}$	5	0.5	0.3	0.7	1.3	1.2	1.5	1.0	1.5	1.4	0.6	0.9	0.6
	10	0.8	0.5	0.9	1.6	1.0	1.4	1.4	1.4	1.2	0.9	0.7	0.6
$\boldsymbol{B_3}$	5	0.7	0.3	0.6	1.9	1.4	1.1	0.7	1.2	0.6	0.9	1.1	0.8
	10	1.0	0.6	0.9	1.3	1.2	1.4	0.5	1.1	0.9	0.7	1.4	0.8
$\boldsymbol{B_4}$	5	0.6	0.4	0.7	1.6	0.5	1.1	1.1	0.5	1.5	1.0	1.5	1.0
	10	0.6	0.5	0.5	1.5	0.6	1.3	1.4	0.8	1.4	1.0	1.4	1.3
		Student t-distribution											
$\boldsymbol{B_1}$	5	0.9	0.5	0.8	1.8	0.8	1.7	1.0	1.0	1.1	0.8	0.4	0.5
	10	1.6	0.5	0.6	1.3	1.8	0.7	1.0	0.8	1.1	1.1	1.1	0.7
$\boldsymbol{B_2}$	5	0.6	0.4	0.8	1.6	1.5	1.9	1.3	1.2	1.0	0.4	0.7	0.4
	10	1.0	0.6	1.2	1.1	0.8	1.9	1.0	1.1	0.9	1.1	0.6	0.4
$\boldsymbol{B_3}$	5	0.8	0.2	0.4	2.4	1.1	0.8	0.5	0.9	0.5	1.2	0.8	1.0
	10	1.2	0.5	0.7	1.6	0.9	1.0	0.7	1.4	0.6	0.9	1.1	1.1
$\boldsymbol{B_4}$	5	0.8	0.5	0.5	2.1	0.7	1.3	0.8	0.7	1.9	0.8	1.1	1.3
	10	0.8	0.6	0.7	1.2	0.7	1.5	1.7	0.9	1.0	1.2	1.1	1.0

Table 21: Empirical sizes, for 1% nominal tests, based on models $B_1 - B_4$, n=500.

Model	Lags	Asymptotic distribution				MC				RWB			
		Q_{22}	Q_{22}^{w}	L_b	C_m	Q_{22}	Q_{22}^{w}	L_b	C_m	Q_{22}	Q_{22}^{w}	L_b	C_m
		Gaussian distribution											
$\boldsymbol{B_1}$	5	0.7	0.3	0.7	1.8	0.9	1.1	1.3	1.2	1.1	1.2	1.0	0.9
	10	1.0	0.6	0.7	1.2	0.9	1.1	1.2	1.1	1.1	1.2	1.0	0.9
$\boldsymbol{B_2}$	5	0.5	0.1	0.4	2.0	1.1	1.2	0.7	0.8	1.1	1.1	0.9	1.0
	10	0.5	0.3	0.5	1.0	1.1	1.0	0.8	0.9	1.0	1.0	1.0	1.0
$\boldsymbol{B_3}$	5	0.5	0.1	0.5	2.0	1.2	1.2	1.2	1.1	0.8	0.8	1.2	0.8
	10	0.4	0.3	0.4	1.3	1.1	1.2	1.2	1.0	0.9	0.8	1.1	0.9
$\boldsymbol{B_4}$	5	0.8	0.2	0.8	2.0	0.9	1.2	0.8	0.9	1.0	1.2	1.0	1.0
	10	0.5	0.4	0.5	1.3	0.9	1.2	0.9	0.9	1.0	1.1	1.0	1.1
		Skewed normal distribution											
$\boldsymbol{B_1}$	5	1.1	0.7	1.1	1.2	0.7	0.7	1.2	1.2	1.1	1.2	0.8	1.1
	10	1.1	1.0	1.1	1.3	0.7	0.8	1.2	1.2	1.1	1.1	0.8	1.1
$\boldsymbol{B_2}$	5	0.8	0.5	0.7	1.4	1.3	1.3	0.7	1.2	0.9	1.2	0.8	0.8
	10	1.0	0.7	1.0	1.2	1.1	1.2	0.8	1.2	0.9	1.1	0.9	0.8
$\boldsymbol{B_3}$	5	0.8	0.5	0.8	1.4	0.7	1.2	0.7	1.2	1.2	1.2	0.9	1.1
	10	1.1	0.8	0.9	1.4	0.8	1.1	0.7	1.1	1.1	1.2	0.8	1.1
$\boldsymbol{B_4}$	5	0.6	0.4	0.6	1.2	1.4	1.2	1.3	0.7	0.7	1.1	1.1	1.1
	10	0.7	0.6	0.7	1.1	1.2	1.2	1.1	0.8	0.8	1.1	1.1	1.0
		Student t-distribution											
$\boldsymbol{B_1}$	5	1.3	0.6	1.3	1.5	0.6	0.8	1.1	1.4	0.9	1.4	0.9	1.0
	10	1.3	1.1	1.3	1.1	0.8	0.7	1.3	1.4	0.9	1.3	1.0	1.3
$\boldsymbol{B_2}$	5	0.7	0.4	0.8	1.6	1.0	1.5	0.6	1.0	1.1	1.4	1.0	0.7
	10	0.9	0.8	0.9	1.0	1.2	1.0	0.9	1.0	1.0	0.9	0.8	0.7
$\boldsymbol{B_3}$	5	0.8	0.6	0.9	1.1	0.6	1.4	0.6	1.3	1.4	1.0	0.7	1.3
	10	1.3	0.6	1.1	1.1	0.7	1.3	0.6	0.9	1.2	1.4	0.7	0.9
$\boldsymbol{B_4}$	5	0.5	0.5	0.7	1.1	1.6	1.1	1.1	0.8	0.7	0.9	1.2	0.9
	10	0.8	0.7	0.9	1.3	1.0	1.0	1.3	0.9	1.0	0.9	1.2	0.9

REFERENCES

Anderson, T. W. (1993). Goodness of fit tests for spectral distributions. *The Annals of Statistics*, 21(2): 830-847.

Bollerslev, T. (1986). Generalized autorregressive conditional heteroskedasticity. *Journal of Econometrics*, 31: 307-327.

Box, G. E. P., Pierce, D. A. (1970). Distribution of residual autocorrelation in autoregressive-integrated moving average time series models. *Journal of American Statistical Association*, 65:1509–1526.

Engle, R. F. (1982). Autoregressive conditional Heteroskedasticity with estimates of the variance of united kingdom ination. *Econometrica*, 50(4): 987-1007.

Fisher, T. J., Gallagher, C. M. (2012). New weighted portmanteau statistics for time series goodness of fit testing. *Journal of the American Statistical Association*, 107(498): 777–787.

Gallagher, C. M., Fisher, T. J. (2015). On weighted portmanteau tests for time-series goodness-of-fit. *Journal of Time Series Analysis*, 36(1): 67–83.

Granger, C. W. and Anderson, A. P. (1978). *An Introduction to Bilinear Time Series model*. Vandenhoeck and Ruprecht, Guttingen.

Hong, Y. (1996a). Consistent testing for serial correlation of unknown form. *Econometrica*, 46(4): 837–864.

Hong, Y. (1996b). Testing for independence between two covariance stationary time series. *Biometrika*, 83(3): 615–625.

Lawrance, A. J., Lewis, P. A. W. (1985). Modelling and residual analysis of nonlinear autoregressive time series in exponential variables. *Journal of the Royal Statistical Society B*, 47: 165–202.

Lawrance, A. J., Lewis, P. A. W. (1987). Higher-order residual analysis for nonlinear time series with autoregressive correlation structures. *International Statistical Review*, 55: 21–35.

Lee, T. (2016) Wild bootstrap ljung-box test for cross correlations of multivariate time series, *Economics Letters*, 147. 59-62.

Lin, J. W., McLeod, A. I. (2006). Improved Peňa-Rodrĭguez portmanteau test. *Computational Statistics and Data Analysis*, 51(3): 1731–1738.

Li, W. K., Mak, T. K. (1994). On the squared residual autocorrelations in non-linear time series with conditional heteroskedasticity. *Journal of Time Series Analysis*, 15(6):627–636.

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

Mahdi, E. (2011). Diagnostic Checking, Time Series and Regression. *Electronic Thesis and Dissertation Repository*. 244. https://ir.lib.uwo.ca/etd/244

Mahdi, E., McLeod, A. I. (2012). Improved multivariate portmanteau test. *Journal of Time Series Analysis*, 33(2): 211–222.

Mahdi, E. (2016). Portmanteau test statistics for seasonal serial correlation in time series models. *SpringerPlus* **5,** 1485.

Mahdi, E. (2017). Kernel-based portmanteau diagnostic test for ARMA time series models.*Cogent Mathematics*, 4(1), doi: 10.1080/23311835.2017.1296327.

Mahdi, E. (2020a). portes: An R Package for Portmanteau Tests in Time Series Models. *arXiv preprint arXiv:2005.00931*.

Mahdi, E. (2020b). New Goodness-of-Fit Tests for Time Series Models. arXiv preprint *arXiv:2008.08176.*

Mahdi, E., Fisher T. J. (2021). Bootstrapping a Powerful Mixed Portmanteau Test for Time Series. ArXiv: 2005.00971 math.ST (2020) (available at https://arxiv.org/abs/2005.00971).

McLeod, A. I. and Li, W. K. (1983). Distribution of the residual autocorrelation in multivariate ARMA time series models. *Journal of Time Series Analysis*, 4: 269–273.

Monti, A. C. (1994). A proposal for residual autocorrelation test in linear models.

Biometrika, 81: 776–780.

Peňa, D., Rodríguez, J. (2002). A powerful portmanteau test of lack of test for time series. *Journal of American Statistical Association*, 97(458): 601–610.

Peňa, D., Rodríguez, J. (2006). The log of the determinant of the autocorrelation matrix for testing goodness of fit in time series. *Journal of Statistical Planning and Inference*, 8(136): 2706–2718.

Psaradakis, Z., Vávra, M. (2019). Portmanteau tests for linearity of stationary time series. *Econometric Reviews*, 38(2): 248–262.

Rodríguez, J., Ruiz, E. (2005). A powerful test for conditional Heteroskedasticity for financial time series with highly persistent volatilites. *Statistica Sinica*, 15: 505–525.

Tong, H., Lim, K. (1980). Threshold autoregression, limit cyclesand cyclical data (with discussion). *Journal of the Royal Statistical Society B,* 42: 245-292.

Zhu, K. (2016). Bootstrapping the portmanteau tests in weak auto-regressive moving average models, *Journal of the Royal Statistical Society: Series B (Statistical Methodology),* 78 463-485. doi:https://doi.org/10.1111/rssb.12112.