An Automatic Portmanteau Test For Nonlinear Dependence

Grivas, Charisios

Birkbeck University

19 December 2021

An Automatic Portmanteau Test For Nonlinear Dependence

Charisios Grivas[a]

[a]*Birkbeck University of London, Malet Street, London, WC1E 7HX, United Kingdom*

ARTICLE INFO

Keywords:
ARMA time series;Akaike's AIC
Schwarz's BIC; Portmanteau test;
Data-driven test

ABSTRACT

A data-driven version of a portmanteau test for detecting nonlinear types of statistical dependence is considered. An attractive feature of the proposed test is that it properly controls type I error without depending on the number of lags. In addition, the automatic test is found to have higher power in simulations when compared to the McLeod and Li test, for both raw data and residuals.

1. Introduction

One of the most popular approaches for detecting serial correlation is based on the Box–Pierce portmanteau test (see [2]), or its finite-sample correction developed by [3]. The test statistic is simply the sample size times the sum of squares of the first p sample autocorrelations. A problem that practitioners often face is how to correctly specify ad hoc the order of lags p for the autocorrelation in a way that will properly control the probability of type I error while having a high power.

An attractive application of portmanteau tests is on the residuals of fitted autoregressive moving average (ARMA) models. Following the methodology of [4], once a model has been fitted, the econometrician performs a number of tests on the residuals to check the adequacy of the model. This process of checking the residuals for any remaining serial correlation continues until the resulting residuals contain no detectable additional structure. In practice, portmanteau tests are more useful for disqualifying unsatisfactory models from consideration than for selecting the best-fitting model among closely competing candidates, as pointed out by [5] (page 312).

However, as [6] observed, the autocorrelation function of the squared series can be useful in identifying nonlinearity in time series. In particular, even when the series appears not to be autocorrelated, the squared series could be autocorrelated, hence revealing some form of nonlinear dependence. Motivated by this observation, [1] proposed a modification of the Ljung–Box test based on the squared series or squared residuals from a fitted model. As [7] demonstrate, the McLeod–Li test[1] is particularly successful in the presence of autoregressive conditional heteroskedasticity (ARCH) and stochastic volatility, the test being asymptotically equivalent to a Lagrange Multiplier test for ARCH (see also [8]).[2] Although the test generally has good control of the probability type I error, this is dependent on the number of autocorrelations p used to construct the test being carefully chosen.

The aim of this article is to propose a data-driven version of the McLeod–Li test that deals with the problem of selecting the number of autocorrelations p in an effective manner. Building upon the seminal work by [9], we let the data determine the order p, by automatically adapting to the order of the serial correlation present. Under the null hypothesis of independent, identically distributed data (i.i.d) data, the proposed test statistic follows asymptotically a chi-square distribution with one degree of freedom. Under the alternative hypothesis of non-i.i.d. data, the test chooses p depending on the serial correlation present and is consistent. An attractive feature of the automatic version of the McLeod–Li test, is that its size properties are not dependent on the choice of the p value.

Adapting the methodology of [9], our paper contributes to the literature in two distinct ways. First, by considering squared data instead of raw data, we provide a means of identifying nonlinear dependence in a time series. Second, and more importantly, we extend the procedure to squared residuals from a fitted model, thus providing a general diagnostic test for (second-order) nonlinear dependence.

⋆ Corresponding author :Charisios.grivas@gmail.com
ORCID(s):
[1] Henceforth, McLeod-Li and $\mathcal{ML}$ will be used interchangeably.
[2] In fact, [7] call it an ARCH test to "remind the reader that it is an ARCH test".

The paper is organized as follows. In Section 2, we briefly present an overview of the McLeod–Li test, introduce the proposed test for both raw data and residuals, and establish their asymptotic properties. In Section 3, we explore the finite-sample behaviour of the test by means of Monte Carlo simulations. Finally, some concluding remarks and suggestions for possible future research are given in Section 4.

2. The McLeod and Li Portmanteau Test

In this section, we first revisit the McLeod–Li test for raw data and propose an automatic version of the test. We then consider the case of residuals from a fitted model.

2.1. Preliminaries

Consider a strictly stationary time series $\{y_t\}_{t=1}^T$ with $\mathbb{E}[y_t^4] < \infty$. For any $0 \leq j \leq T-1$, let $\hat{\gamma}(j) = \frac{1}{T-j}\sum_{t=1+j}^{T}\left(y_t^2 - \bar{\delta}\right)\left(y_{t-j}^2 - \bar{\delta}\right)$ be the sample analogue of $\gamma(j) = \operatorname{cov}\left(y_t^2, y_{t-j}^2\right)$, with $\bar{\delta} = \frac{1}{T}\sum_{t=1}^{T} y_t^2$. The portmanteau test of [1] is based on the statistic:

$$\mathcal{ML}(p) = T(T+2)\sum_{j=1}^{p}\frac{\hat{\rho}^2(j)}{T-j}, \tag{1}$$

where $\hat{\rho}(j) = \hat{\gamma}(j)/\hat{\gamma}(0)$ is the lag-j sample autocorrelation of the squared series and p is the maximum lag order (number of autocorrelations) specified ad hoc by the econometrician. Under the null hypothesis that $\{y_t\}_{t=1}^T$ is i.i.d., $\rho(j) = \gamma(j)/\gamma(0) = 0$ for all $j \geq 1$ and $\mathcal{ML}(p)$ has a $\chi^2(p)$ asymptotic distribution as $T \to \infty$.

2.2. Automatic Portmanteau Test

In their novel paper, [9] proposed a modification of the Ljung–Box portmanteau test that allows the data to automatically determine the number of autocorrelations on the basis of an information criterion. The proposed test statistic is the maximum value of the portmanteau test statistics penalized by a term that is an increasing function of the number of autocorrelations.

Analogously to [10] and [9], our automatic version of the McLeod–Li test is based on the statistic:

$$\mathcal{AML} = \mathcal{ML}(\tilde{p}),$$

where

$$\tilde{p} = \min\{p : 1 \leq p \leq d, \mathcal{L}_p = \max_{1\leq u\leq d}\mathcal{L}_u\},$$

and

$$\mathcal{L}_u = \mathcal{ML}(u) - \pi(u,T,q).$$

Here, d serves as a fixed upper bound, q is a fixed positive number, to be defined later, and $\pi(u,T,q)$ is a penalty function that takes the form

$$\pi(u,T,q) = \begin{cases} u\log T, & \text{if } \max_{1\leq j\leq d}\sqrt{T}|\hat{\rho}(j)| \leq \sqrt{q\log T}, \\ 2u, & \text{if } \max_{1\leq j\leq d}\sqrt{T}|\hat{\rho}(j)| > \sqrt{q\log T}. \end{cases} \tag{2}$$

Note that in (2) the penalty function involves a switching rule between the Bayesian information criterion (BIC) criterion ([11]) and the Akaike information criterion (AIC) ([12]). This combination of the two criteria is desirable since BIC is able to properly control type I error and is more powerful when there is first-order serial correlation in the data. When higher-order serial correlation is present, AIC yields more powerful tests.

Our first theorem establishes the asymptotic null distribution of the automatic McLeod–Li test.

Theorem 1. *If $\{y_t\}_{t=1}^T$ is i.i.d with $\mathbb{E}[y_t^4] < \infty$, then $\mathcal{AML}$ has a $\chi^2(1)$ asymptotic distribution as $T \to \infty$.*

The next theorem shows that the test is consistent against a fixed alternative under which $\rho(j) \neq 0$ for some $j \geq 1$.

Theorem 2. *If $\{y_t\}_{t=1}^T$ is strictly stationary and ergodic with $\mathbb{E}[y_t^4] < \infty$, then the test based on $\mathcal{AML}$ is pointwise consistent, as $T \to \infty$, against the alternative $H_a^K : \rho(1) = \cdots = \rho(K-1) = 0$, $\rho(K) \neq 0$ for $1 \leq K \leq d$.*

The proofs of the theorems are in the Appendix A.

2.3. Residual Portmanteau Test

We now consider the case where the adequacy of a fitted model is checked by using a portmanteau test based on the autocorrelations of the squared residuals. The econometrician observes a finite stretch of data $\{y_t\}_{t=1}^T$ from a stochastic process with mean μ satisfying:

$$y_t = \mu + \sum_{j=0}^{\infty} c(\beta, j)\epsilon_{t-j}, \tag{3}$$

where $c(\beta, j)$ are real weights, assumed to be known functions of an unknown finite-dimensional vector of parameters β and satisfying $\sum_{j=0}^{\infty} |c(\beta, j)| < \infty$ and $c(\beta, 0) = 1$, and $\{\epsilon_t\}_{t=1}^{\infty}$ are strictly stationary white noise errors with $\mathbb{E}[\epsilon_t] = 0$ and $\mathbb{E}[|\epsilon_t|^s] < \infty$ for some $s \geq 8$. A well-known special case of (3) are ARMA(r_1, r_2) processes, for which the weights $c(\beta, j)$ satisfy $\sum_{j=0}^{\infty} c(\beta, j)z^j = \vartheta(z)/\varphi(z)$ for all complex $|z| \leq 1$, where $\vartheta(\cdot)$ and $\varphi(\cdot)$ are polynomials of degree r_1 and r_2, respectively, having no common roots and no roots inside or on the unit circle.[3]

Following [1], the autocorrelation function of the squared errors $\{\epsilon_t^2\}$ can be useful in identifying nonlinear dependence. The presence of such nonlinear dependence is of importance since it can be an indication of misspecification of the fitted model, or can, if taken into consideration, lead to improved forecast accuracy ([6]).

The errors $\{\epsilon_t\}$ are, of course, unobservable in practice and hence one must use residuals $\{\hat{\epsilon}_t\}_{t=1}^T$ in their place.

Given a consistent estimator $\hat{\theta} = (\hat{\beta}, \hat{\mu})$ of $\theta = (\beta, \mu)$, and assuming (3) is invertible, residuals can be computed as (e.g., [13]):

$$\hat{\epsilon}_t = y_t - \hat{\mu} - \sum_{j=1}^{t-1} c\left(\hat{\beta}, j\right)(y_{t-j} - \hat{\mu}). \tag{4}$$

Suitable estimators of θ can be obtained by quasi-maximum likelihood, instrumental variables, or least-squares methods (see, e.g., [14] and the references therein). Then, given a prespecified fixed $p \geq 1$, a test for nonlinear dependence in $\{\epsilon_t\}$ may be based on the McLeod–Li portmanteau statistic:

$$\mathcal{ML}^*(p) = T(T+2)\sum_{j=1}^{p} \frac{\hat{\rho}_{\hat{\epsilon}\hat{\epsilon}}^2(j)}{T-j} \tag{5}$$

where $\hat{\rho}_{\hat{\epsilon}\hat{\epsilon}}(j) = \hat{\gamma}_{\hat{\epsilon}\hat{\epsilon}}(j)/\hat{\gamma}_{\hat{\epsilon}\hat{\epsilon}}(0)$, $\hat{\gamma}_{\hat{\epsilon}\hat{\epsilon}}(j) = \frac{1}{T-j}\sum_{t=1+j}^{T}\left(\hat{\epsilon}_t^2 - \bar{\delta}_{\hat{\epsilon}}\right)\left(\hat{\epsilon}_{t-j}^2 - \bar{\delta}_{\hat{\epsilon}}\right)$ for $0 \leq j \leq T-1$, and $\bar{\delta}_{\hat{\epsilon}} = \frac{1}{T}\sum_{t=1}^{T}\hat{\epsilon}_t^2$. Under suitable regularity conditions (see Assumption 1 below), $\mathcal{ML}^*(p)$ has a $\chi^2(p)$ asymptotic distribution as $T \to \infty$ when $\{\epsilon_t\}$ are i.i.d. with $\mathbb{E}[\epsilon_t^8] < \infty$ (see [8]).

Instead of relying on an ad hoc choice for the number of estimated autocorrelations p in (5), we suggest employing an automatic version of the test, analogous to that introduced in the previous subsection. The test statistic is $\mathcal{AML}^* = \mathcal{ML}^*(\tilde{p})$, with $\tilde{p}$ determined in the same manner as before but now using a penalty function $\pi(u, T, q)$ that takes the form:

$$\pi(u, T, q) = \begin{cases} u \log T, & \text{if } \max \sqrt{T}|\hat{\rho}_{\hat{\epsilon}\hat{\epsilon}}(j)| \leq \sqrt{q \log T}, \\ 2u, & \text{if } \max_{1 \leq j \leq d} \sqrt{T}|\hat{\rho}_{\hat{\epsilon}\hat{\epsilon}}(j)| > \sqrt{q \log T}. \end{cases} \tag{6}$$

To establish the asymptotic properties of the automatic test, we will need the following assumption.

[3]This is the case considered in [1].

Assumption 1. *(a) $C(z) = \sum_{j=0}^{\infty} c(\beta, j) z^j$ is analytic and without zeros inside and on the unit circle, and differentiable with respect to β; (b) $\hat{\theta}$ is $\sqrt{T}$-consistent for θ; (c) $\sqrt{T}\left|\partial \tilde{\gamma}_{\epsilon\epsilon}(j)/\partial\theta\right| = O_p(1)$ for $0 \leq j \leq T-1$, where $\tilde{\gamma}_{\epsilon\epsilon}(j) = \frac{1}{T-j}\sum_{t=1+j}^{T}\left(\epsilon_t^2 - \bar{\delta}_\epsilon\right)\left(\epsilon_{t-j}^2 - \bar{\delta}_\epsilon\right)$ and $\bar{\delta}_\epsilon = \frac{1}{T}\sum_{t=1}^{T}\epsilon_t^2$.*

These conditions, which are similar to those of [8], ensure that the residuals in (4) are well defined and the estimated autocorrelations $\hat{\rho}_{\hat{\epsilon}\hat{\epsilon}}(\cdot)$ are consistent and asymptotically normal under the null hypothesis of i.i.d. errors. In the ARMA case mentioned earlier, with parameters estimated by conventional methods (see, e.g., [5, Ch. 8]), all the requirements of the assumption are satisfied.

The asymptotic properties of the automatic test are given in the next two theorems. Here, $\rho_{\epsilon\epsilon}(j) = \gamma_{\epsilon\epsilon}(j)/\gamma_{\epsilon\epsilon}(0)$ for $j \geq 0$, where $\gamma_{\epsilon\epsilon}(j) = \text{cov}\left(\epsilon_t^2, \epsilon_{t-j}^2\right)$.

Theorem 3. *If $\{\epsilon_t\}_{t=1}^{\infty}$ is i.i.d. with $\mathbb{E}[\epsilon_t^8] < \infty$ and Assumption 1 holds, then $\mathcal{AML}^*$ has a $\chi^2(1)$ asymptotic distribution as $T \to \infty$.*

Theorem 4. *If Assumption 1 holds and $\hat{\rho}_{\hat{\epsilon}\hat{\epsilon}}(j) = \rho_{\epsilon\epsilon}(j) + o_p(1)$ for all $0 \leq j \leq T-1$, then the test based on $\mathcal{AML}^*$ is consistent, as $T \to \infty$, against the fixed alternative $H_a^K : \rho_{\epsilon\epsilon}(1) = \cdots = \rho_{\epsilon\epsilon}(K-1) = 0$, $\rho_{\epsilon\epsilon}(K) \neq 0$ for $1 \leq K \leq d$.*

The proofs of the theorems are in the Appendix A.

Note that the assumption that $\mathbb{E}[\epsilon_t^8] < \infty$ is a standard requirement for $\mathcal{ML}$-type tests (e.g., [1], [8]) and ensures that sample autocovariances of squared errors have a well-defined asymptotic distribution when the errors are not autocorrelated. It may be possible to relax the moment condition somewhat along the lines of [15], although this would typically require imposing stronger restrictions on the parameters of the data-generating process. A different approach that has been considered in the literature relies on using the absolute value of a time series or log-squared time series in the construction of $\mathcal{ML}$ test statistics as these only require the existence of the fourth moment. These approaches, however, are beyond the scope of the article and the interested reader is referred to [16], and the references therein, for some discussion.

It is also worth remarking that methods based on automatic model selection typically entail biases, as shown by [17], and the sampling distributions of estimators and test statistics may be affected as a result. Ideally, any such effects should not be ignored post estimation. When the McLeod–Li test, or its automatic version, are applied to the residuals of a particular model to test for possible nonlinear dependence, the diagnostic testing may affect subsequent inference. However, a number of strategies have been proposed to deal with this issue, as pointed out by [18]. For example, [19] suggest a model selection method based on hierarchical models. They develope a model selection strategy limiting the selection to ARMA$(r, r-1)$ candidate models and show that this does not lead to a lower-quality model being selected even in finite samples. On the other hand, [20] consider bootstrap as a way of incorporating model uncertainty into inference. The bootstrap, with model selection applied independently to each resample, allows inference without conditioning on a single selected model.

2.4. Linearity Test

A time series $\{y_t\}$ is sometimes considered to be linear if it admits a representation such as (3) with respect to an i.i.d. sequence $\{\epsilon_t\}$. This is the notion of linearity considered in [1], [21], [22] and [23], among others.[4] Within this framework, deviations from the i.i.d. assumption about the errors $\{\epsilon_t\}$ are viewed as evidence of nonlinear behaviour of $\{y_t\}$. Hence, one may employ tests based on the $\mathcal{ML}^*$ or $\mathcal{AML}^*$ statistics constructed from the residuals of a suitable approximation to (3) as general portmanteau tests for detecting deviations from linearity. [25], for example, use a first-order autoregressive model to obtain the required residuals, while [1] rely on autoregressive models the order of which is estimated by means of information criteria.

3. Simulations

In this section, Monte Carlo experiments are carried out to investigate the finite-sample performance of the $\mathcal{AML}$ test. The main objective is to compare the properties of the automatic $\mathcal{ML}$ to the classical $\mathcal{ML}$, in the presence and

[4] This is not, of course, the only characterization of non-linearity considered in the literature. For example, [24] considers a different characterization based on the properties of one-step-ahead linear predictors.

absence of nonlinear serial dependence in the errors of a model. The case of raw data is provided in the supplement, to conserve space.

In all simulations[5] in this section, 10,000 independent artificial time series y_t of length 100+T with $T \in \{50, 100, 200, 500\}$ are generated , but only the last T observations for each series are used. We consider p/T = 2.5%, 5%, 7% and 10% for the original test. The nominal level[6] considered, for comparing the empirical significance levels of $\mathcal{AML}$ and $\mathcal{ML}$, is $\alpha = 5\%$. We set $d = 75$, as in [9], except when $T = 50$ where we use $d = 25$,and $q = 3.6$ following our simulations in the next section. Note that in both [9] and [10] q is set to 2.4.

3.1. Choice of q

As discussed in section 2 above, expression (2) involves a fixed parameter q which provides a switching rule between the BIC and AIC criteria. In this subsection, we provide evidence to support our choice of $q = 3.6$ used in some of the Monte Carlo simulations, in a similar manner to [9] and [10]. Specifically, we estimate the empirical size of the test by Monte Carlo simulation for a sample of $T = 500$. For all the simulations in this subsection, $\mathcal{AML}$ is applied to the residuals obtained from an AR(1) model with a coefficient $\beta \in \{\pm 0.5, \pm 0.9\}$,[7] for eight different values of q, ($q = 1.2,\ 1.5,\ 1.8,\ 2.4,\ 2.7,\ 3,\ \infty$).

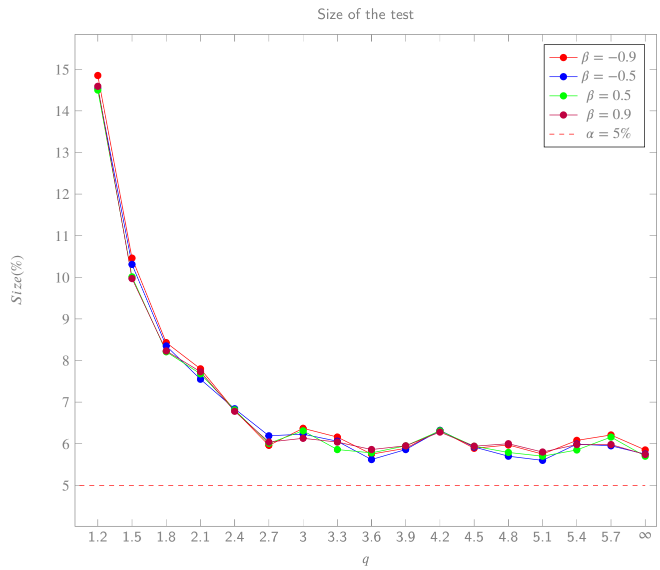

Figure 1: Rejection frequencies at 5% level for the $\mathcal{AML}$ test on the residuals of an $AR(1)$ model, where the error follows a standard normal distribution, and for different values of constant q. Simulations are based on $T = 500$.

Figure 1 above, presents graphically the rejection probabilities under the null, where the errors are generated from a standard normal distribution. Figure 1 shows that for $q > 3$ the rejection probability becomes relative flat near 5% and hence the value of $q = 3.6$ suffices to properly control for type I error.

We also provide some further justification for our choice for q, under the alternative hypothesis. Specifically, in Figure 2 below, we report the rejection frequencies for the automatic test under the alternative for $T = 500$. We employ an EGARCH(1,1) model, which we will revisit in subsection 3.2, as model **N.5**. Simulations based on a stochastic volatility model provided similar results and are hence omitted. Again, Figure 2 shows that $q = 3.6$ is a reasonable choice with high power, that does not vary much.

[5] All simulations were performed using a 2.7 GHz Intel Core i5 with a 8 GB DDR3 RAM. The code was written and executed in JuliaPro-1.4.1.0-1.

[6] Simulation results for tests of nominal levels 0.01 and 0.10 are not reported, due to space constraints, but are available upon request.

[7] Similar results are obtained for other values of the β parameter but are not provided due to space constraints.

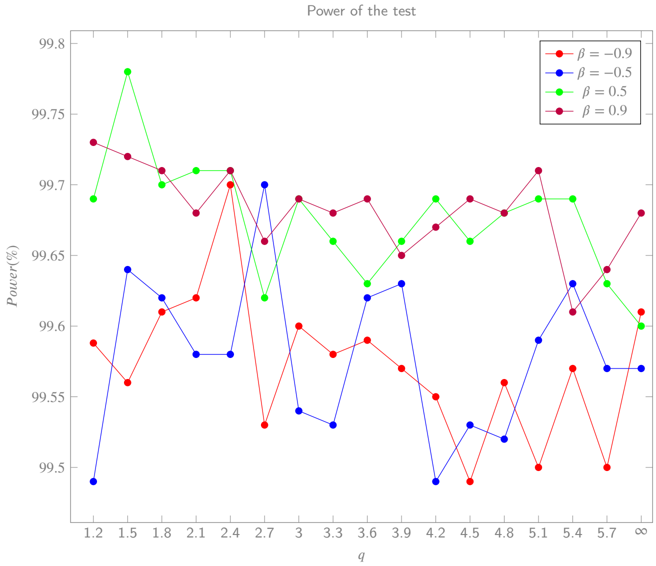

Figure 2: Rejection frequencies at 5% level for the $\mathcal{AML}$ test on the residuals of an $AR(1)$ model, where the error follows a EGARCH, and for different values of constant q. Simulations are based on $T = 500$.

3.2. Level and Power of Tests based on Residuals

In this subsection, we examine the application of the test on residuals. Following [1], we consider the residuals of an AR(1) model, that is:

$$y_t = \beta y_{t-1} + \epsilon_t \tag{7}$$

where $\beta \in \{0, \pm 0.6, \pm 0.9\}$. We estimate that model by least squares and allow for a constant. Under the null, the errors $\{\epsilon_t\}$ are an i.i.d. sequence, following:

S.1 a standard normal distribution,

S.2 a t-distribution with 9 degrees of freedom.[8]

The latter choice is made so that the error term satisfies the finite eigth moment condition. Under the alternative, we allow the errors to have one of the structures defined below.

Under the alternative, the following data-generating processes(DGPs) are used in the simulations for the error term:

N.1 $\epsilon_t = v_t \sigma_t$ where $\sigma_t^2 = 0.001 + 0.05\epsilon_{t-1}^2 + 0.90\sigma_{t-1}^2$,

N.2 $\epsilon_t = v_t \exp(\sigma_t^2)$, where $\sigma_t^2 = 0.936\sigma_{t-1}^2 + 0.32u_t$,

N.3 $\epsilon_t = v_t v_{t-1}$,

N.4 $\epsilon_t = v_{t-2} v_{t-1}(v_{t-2} + v_t + 1)$,

N.5 $\epsilon_t = v_t \sigma_t$ where $\log \sigma_t^2 = 0.001 + 0.5|v_{t-1}| - 0.2v_{t-1} + 0.95 \log \sigma_{t-1}^2$,

N.6 $\epsilon_t = -0.5\epsilon_{t-1} I(\epsilon_{t-1} \leq 1) + 0.4\epsilon_{t-1} I(\epsilon_{t-1} > 1) + u_t$,

N.7 $\epsilon_t = -0.5\epsilon_{t-1}\{1 - G(\epsilon_{t-1})\} + 0.4\epsilon_{t-1} G(\epsilon_{t-1}) + u_t$,

N.8 $\epsilon_t = 0.8u_{t-2}^3 + u_t$.

[8] A log-normal distribution was also considered and provided similar results.

In all the DGPs above, $\{u_t\}$ and $\{v_t\}$ are i.i.d. standard normal random variables independent of each other, $I(A)$ is the indicator of event A, and $G(x) = 1/(1 + e^{-x})$ is the logistic distribution function. The DGPs cover a wide variety of nonlinear processes often encountered in economics and finance. Models **N.1**-**N.4** are taken from [26] and represent a GARCH process, a stochastic volatility process, an 1-dependent process, and an uncorrelated non-martingale-difference process, respectively. **N.5** is an EGARCH process taken from [9]. The remaining three DGPs are taken from [8]; they represent a threshold AR (TAR) [**N.6**], a smooth-transition AR [**N.7**], and a nonlinear MA (NLMA) [**N.8**].[9]

The Monte Carlo rejection frequencies of the conventional $\mathcal{ML}$ test and its automatic version, $\mathcal{AML}$ (at nominal level 5%) are shown in Figure 3-Figure 7. Under the null, the $\mathcal{AML}$ has empirical size close to the nominal level while the size of $\mathcal{ML}$ varies substantially. Figure 3 presents the empirical size of the tests under the null of i.i.d standard normal errors and t-distributed errors. The empirical size of the $\mathcal{ML}$ test seems to be quite close to the nominal level only for $T = 100$, and varies a lot depending on the choice of p, while the bigger the sample the more it deviates from 5%. The empirical size of the $\mathcal{AML}$, on the other hand, does not differ significantly from the nominal level regardless of the sample size T. It is worth mentioning that the $\mathcal{AML}$ experiences small size distortions for a 1% significance level for small samples. On the other hand, the automatic test seems to have proper size for significance levels of 10% independently of the sample size. Nevertheless, we argue that for sample sizes typically encountered in finance, the automatic test has size close to nominal at all significance levels.

[9]Bilinear models were also considered and provided similar results.

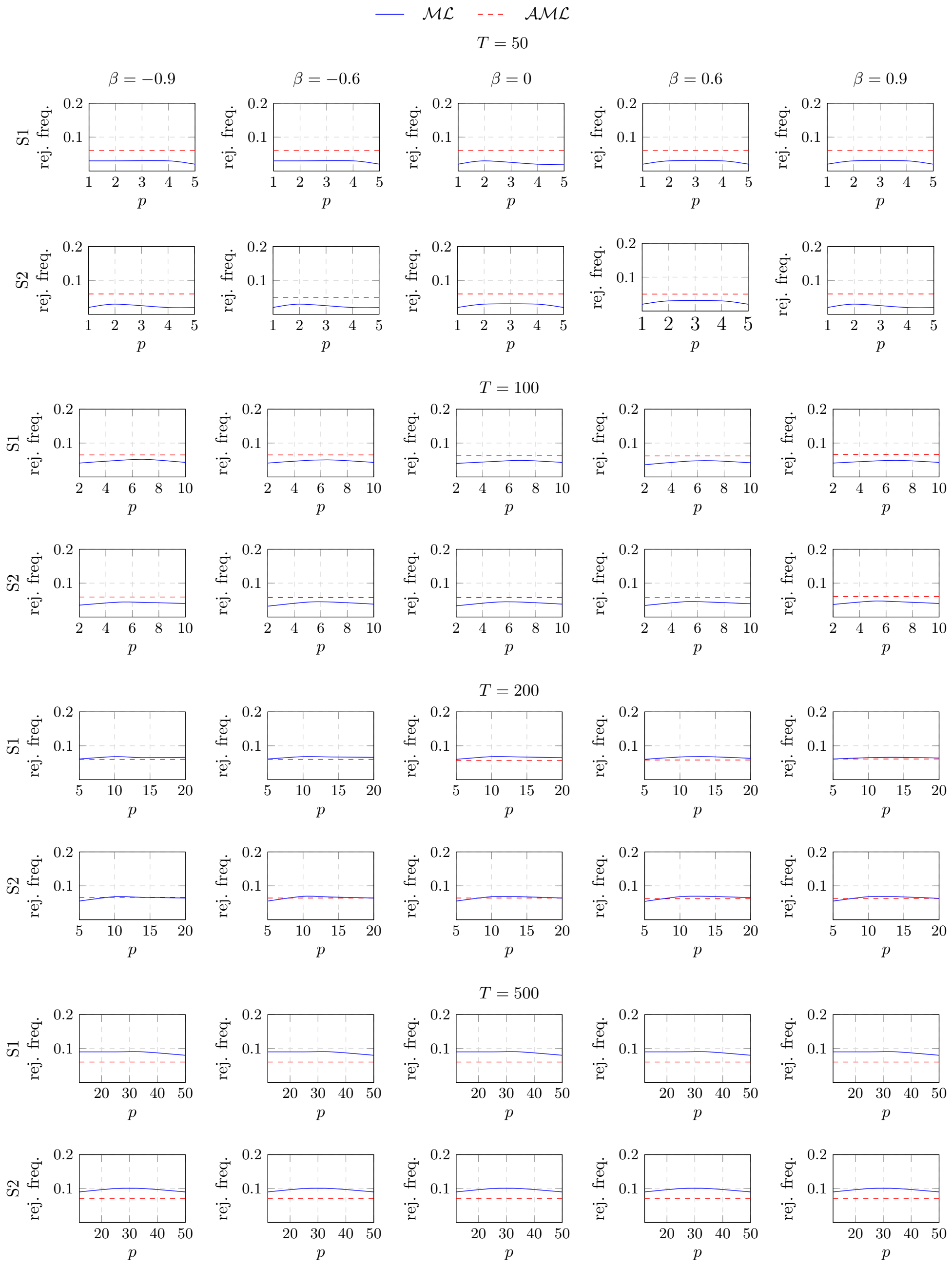

Figure 3: Rejection frequencies of $\mathcal{ML}$ and $\mathcal{AML}$ for the residuals under the null, at 5% singificance level and for different sample sizes.

In Figure 4-Figure 7 we show the rejection frequencies under the alternative. For virtually all DGP, and across all sample sizes, the $\mathcal{AML}$ outperforms the $\mathcal{ML}$. The automatic test has higher rejection frequencies than the $\mathcal{ML}$ not only for ARCH models but also for a variety of different nonlinear models. Also, the value of the β does not seem to affect the size and the power of the tests.

Figure 4 below depicts the rejection frequencies for a sample size of $T = 50$. The $\mathcal{AML}$ test statistic provides higher power for all DGP, with the most obvious case for **N.5**, while the power of the $\mathcal{ML}$ depends somewhat on the choice of p.

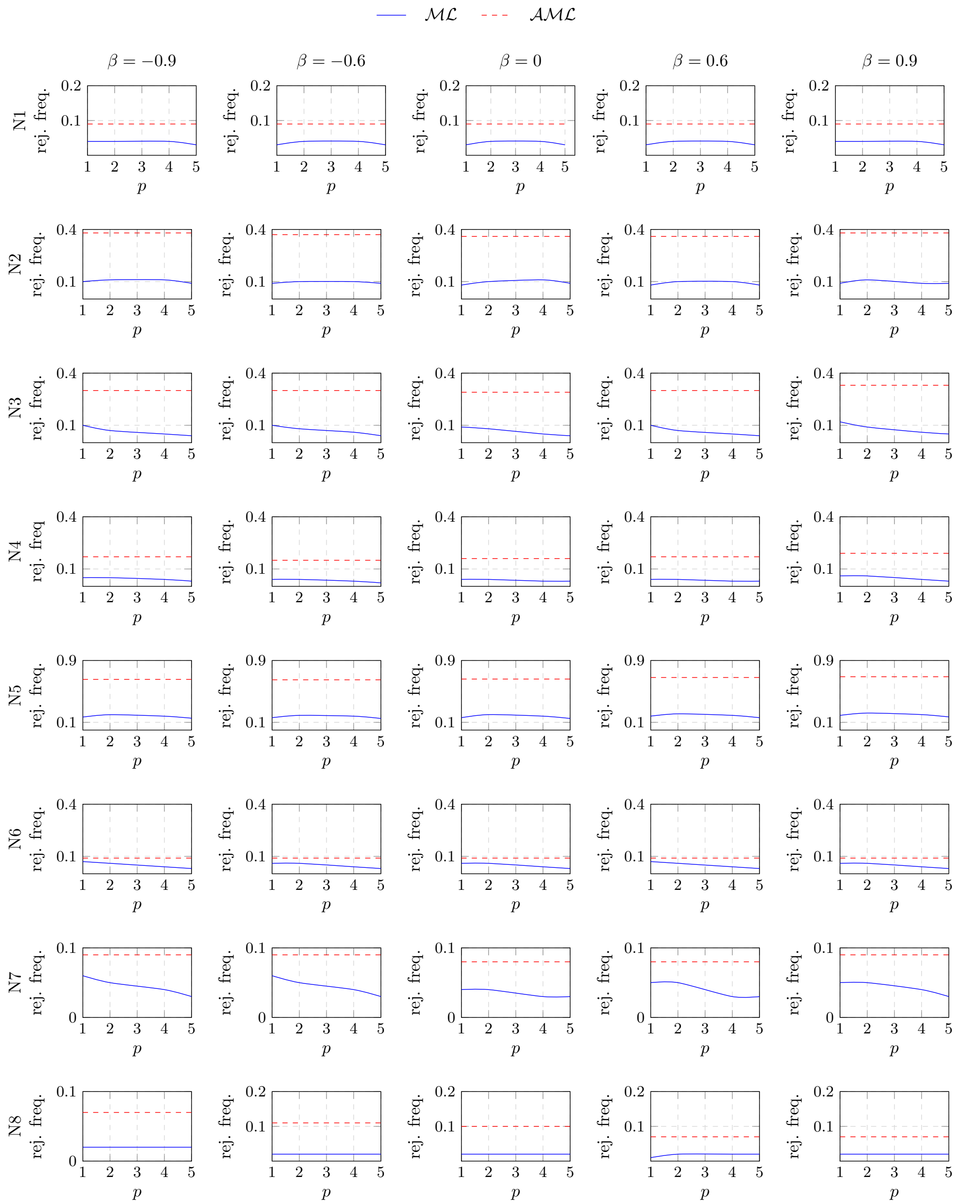

Figure 4: Rejection frequencies of $\mathcal{ML}$ and $\mathcal{AML}$ for the residuals under the alternative for $T = 50$ at 5% singificance level.

Figure 5 shows the rejection frequencies for $T = 100$. Again the $\mathcal{AML}$ provides higher power compared to $\mathcal{ML}$. The sensitivity of the $\mathcal{ML}$ to the choice of p is more apparent here. Note the increase in power of the automatic test for **N.5** and **N.4**. Similarly as above, the value of the β does not seem to affect the power of the tests.

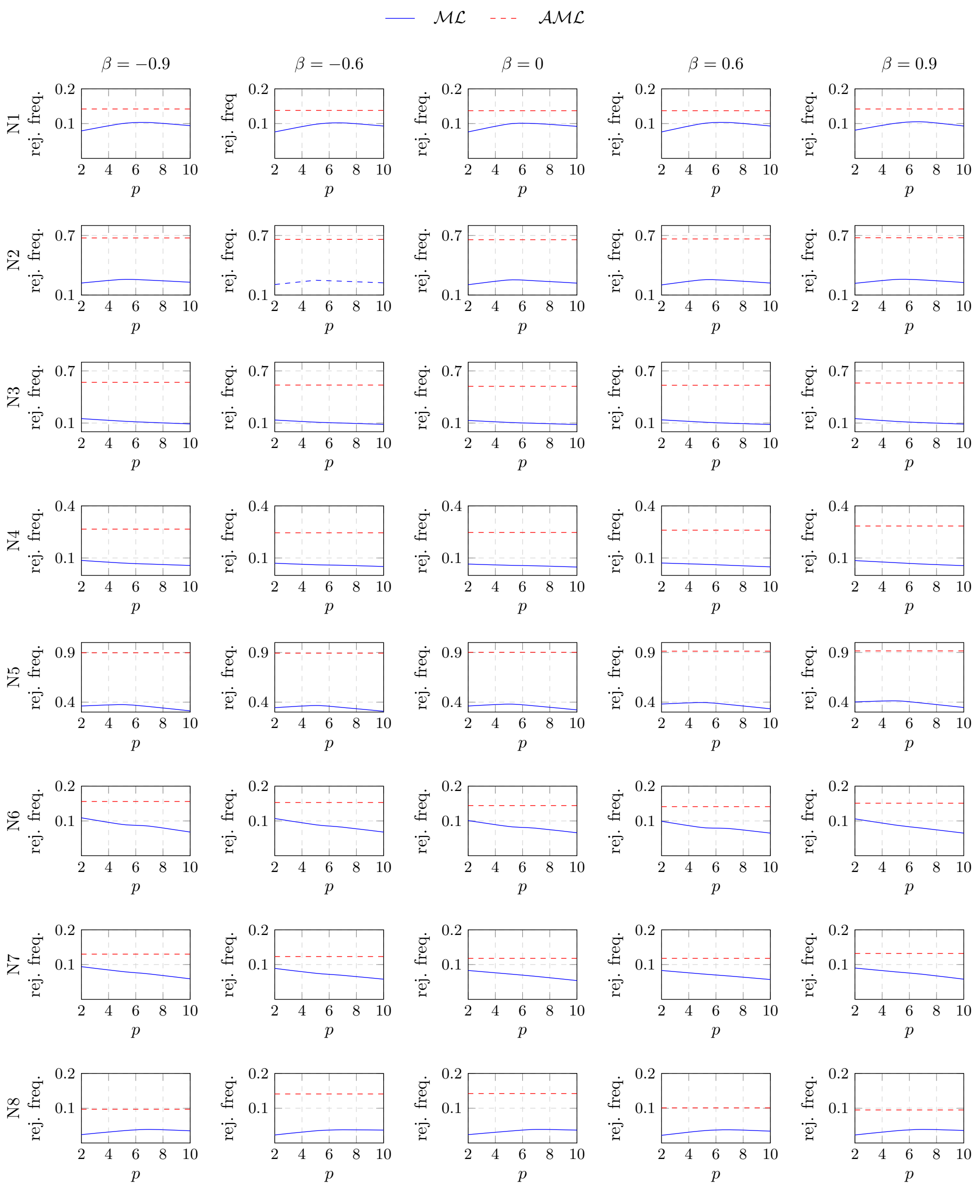

Figure 5: Rejection frequencies of $\mathcal{ML}$ and $\mathcal{AML}$ for the residuals under the alternative, for $T = 100$ and at 5% singificance level.

Moving on to sample sizes that are most common in finance, we see that our main conclusions remain unaffected as seen in Figure 6 and Figure 7. Starting from Figure 6 and a sample size of $T = 200$ we notice that the $\mathcal{AML}$ has higher power than the $\mathcal{ML}$, while the latter seems to have a large sensitivity to the parameter p.

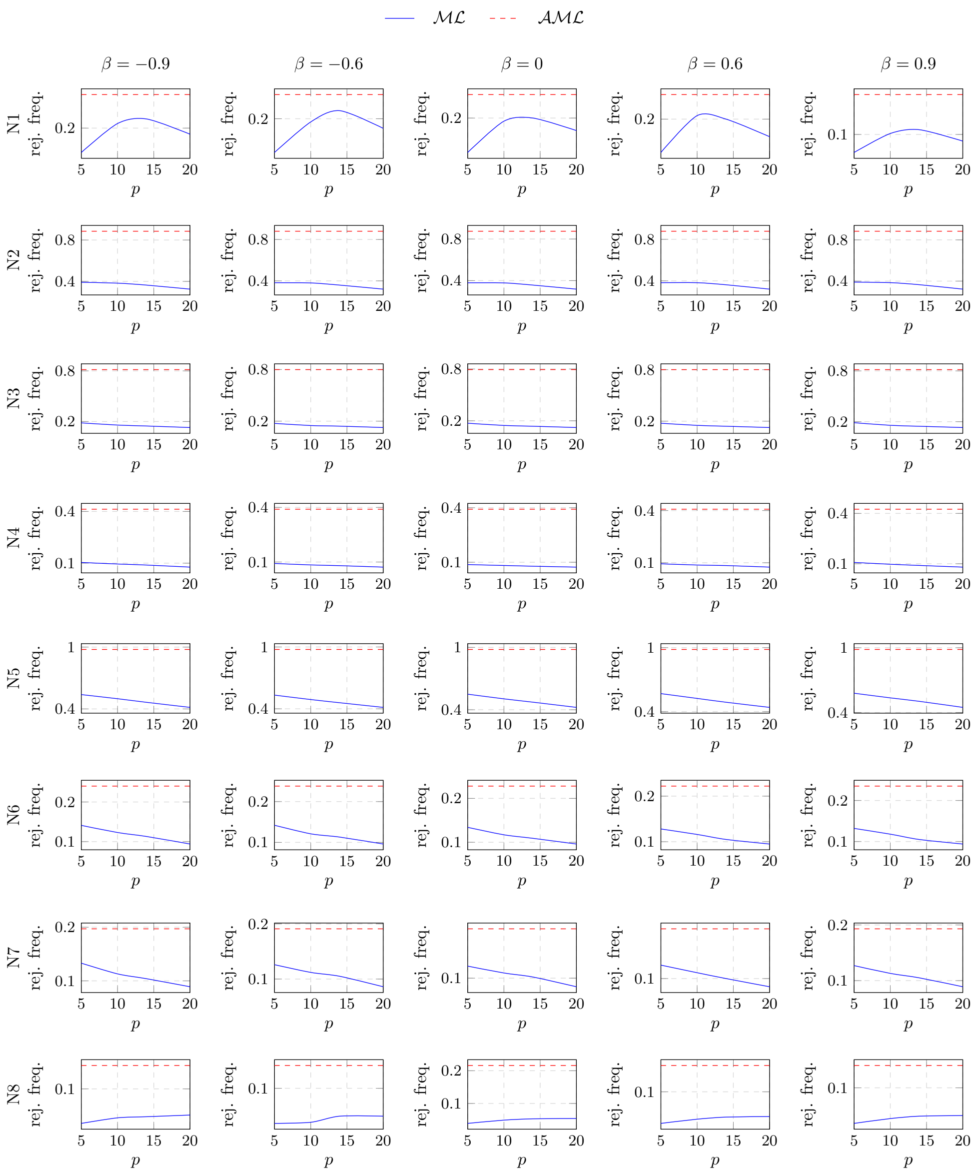

Figure 6: Rejection frequencies of $\mathcal{ML}$ and $\mathcal{AML}$ for the residuals under the alternative, for $T = 200$ and at 5% singificance level.

For example, for **N.1**, the power of the $\mathcal{ML}$ is stable and around 0.2 for values of p close to 10 but it goes below 0.2 for values less or greater than 10. Similarly, for **N.5**-**N.7** the $\mathcal{ML}$ has a power around 0.15 when $p = 5$ while its power diminishes for other values of p. Finally, similar results are obtained for $T = 500$ as depicted in Figure 7.

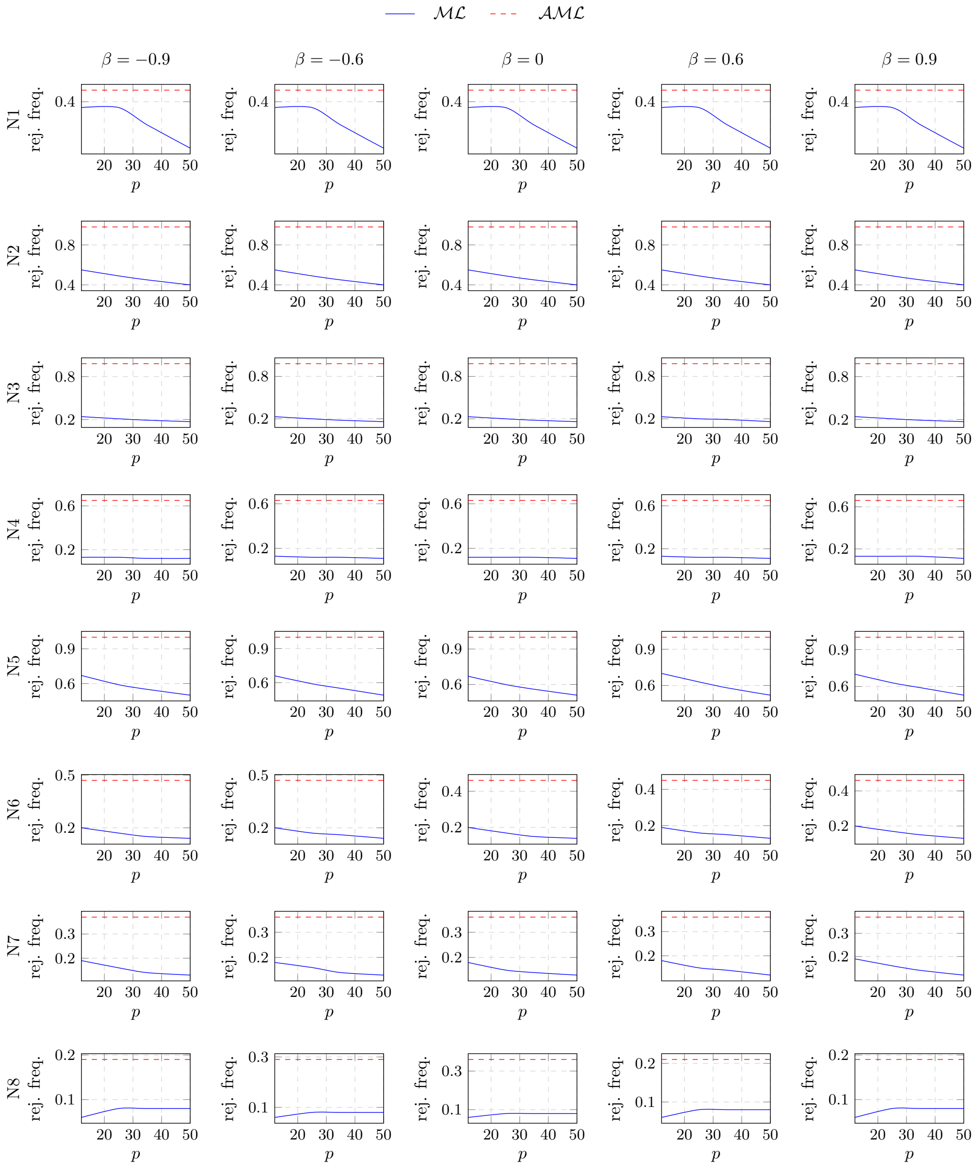

Figure 7: Rejection frequencies of $\mathcal{ML}$ and $\mathcal{AML}$ for the residuals under the alternative, for $T = 500$ and at 5% singificance level.

4. Summary and Conclusion

In this article, we introduced a data-driven version of the McLeod–Li portmanteau test for detecting nonlinear dependence. The proposed test is easy to implement, has a chi-square asymptotic distribution and, most importantly, it properly controls the probability of type I error for sample sizes that are common in applications. The simulation results for both raw data and residuals indicate good size and power properties in finite samples.

Further research could focus on choosing the parameter q in a data-driven way too. This could be done for example by means of bootstrapping or subsampling. Other interesting extensions of the automatic procedure would be to tests based on cross-correlations of the type considered in [8], and tests for multivariate time series and comparisons with the procedures developed in [27].

Funding

This work was supported by the Department of Economics, Mathematics and Statistics, Birkbeck, University of London.

Acknowledgements

The authors wish to thank Zacharias Psaradakis, Walter Beckert, Ron Smith, Marian Vavra, Weining Wang, Karolos Arapakis and Nikos Paltalidis for helpful comments and discussions. The authors are also grateful to an Associate Editor and two anonymous referees for many helpful comments and suggestions that greatly improved an earlier version of the paper.

Declarations Of Interest

None.

A. Appendix

Proof of Theorem 1: Define

$$p_{BIC} = min\{m : 1 \le m \le d; L_{BIC}(m) \ge L_{BIC}(h), h = 1, 2, ..., d\},$$

where $L_{BIC}(p) = ML(p) - p\, logT$. We need to prove that:

$$lim_{T\to\infty} P\left(\tilde{p} = p_{BIC}\right) = 1, \tag{8}$$

and

$$lim_{T\to\infty} P(p_{BIC} = 1) = 1, \tag{9}$$

Under the null of $i.i.d$ data, $\sqrt{T}\hat{\rho}$ has a $N(0, I)$ asymptotic distribution, where $\hat{\rho} = (\hat{\rho}(1), ..., \hat{\rho}(p))$ and I is the identity matrix.

Now, consider the event

$$A_T(q) = \{\, max_{1\le j\le d}\sqrt{T}|\hat{\rho}(j)| > \sqrt{q\, log\, T}\}$$

and assume $q \ge 2$. We have $\max_{1\le j\le d}\sqrt{T}|\hat{\rho}(j)| = O_p(1)$ under the null, so that $P\left(A_T(q)\right) \to 0$, which implies that (8) holds. We also have

$$P\left(p_{BIC} = 1\right) = 1 - \sum_{i=2}^{d} P\left(p_{BIC} = i\right) \ge 1 - \sum_{i=2}^{d} P\left(L_{BIC}(i) \ge L_{BIC}(1)\right) \tag{10}$$

and,

$$P\left(L_{BIC}(i) \ge L_{BIC}(1)\right) \le P\left(ML \ge (i-1)\, log\, T\right)$$

and since under the null ML is $O_p(1)$ we conclude that (9) holds. Therefore, Theorem 1 follows from an application of Lindeberg-Lévy CLT for $i.i.d$ random variables(see [28]27.2, p.359). □

Proof of Theorem 2: Define

$$p_{AIC} = min\{m : 1 \leq m \leq d; L_{AIC}(m) \geq L_{AIC}(h), h = 1, 2, ..., d\}$$

where $L_{AIC}(p) = ML(p) - 2p$. We need to prove that:

$$\lim_{T\to\infty} P\left(\tilde{p} = p_{AIC}\right) = 1 \tag{11}$$

and

$$\lim_{T\to\infty} P\left(p_{AIC} \geq K\right) \to 1 \tag{12}$$

Consider the event

$$A_T(q) = \{\max_{1\leq j\leq d} \sqrt{T}|\hat{\rho}(j)| < \sqrt{q\ log\ T}\}$$

Then for $K \leq d$ we have $\hat{\rho}(K) \to \rho(K) \neq 0$ by the Ergodic Theorem and

$$P\left(A_T(q)\right) \leq P\left(\sqrt{T}|\hat{\rho}(K)| < \sqrt{q\ log\ T}\right) \to 0$$

and so (11) holds. Now for (12), we have:

$$\begin{aligned} P\left(p_{AIC} = k\right) &\leq P\left(L_{AIC}(k) \geq L_{AIC}(K)\right) \\ &\leq P(|ML(k)| \geq\ 2(k-K) + |ML(K)|) \to 0 \end{aligned}$$

which follows directly from an application of the Ergodic Theorem and so (12) also holds. Hence, $\forall\ M > 0$

$$\begin{aligned} P(ML(\tilde{p}) \leq M) &= P(ML(\tilde{p}) \leq M \cap \tilde{p} \geq K) + o(1) \\ &\leq P\left(|\hat{\rho}(K)|^2 \leq M\right) + o(1) \\ &= o(1) \end{aligned}$$

and so $ML(\tilde{p}) \to \infty$, asymptotically and hence the test is consistent against $H_a^K, \forall K \leq d$ □

Proof of Theorem 3: By a similar argument as in the proof of Theorem 1, under the null, and and by defining

$$p_{BIC} = min\{m : 1 \leq m \leq d; L_{BIC}(m) \geq L_{BIC}(h), h = 1, 2, ..., d\}$$

where $L_{BIC}(p) = ML(p) - p\ logT$ and $ML(p)$ refers now to the statistic based on residuals we obtain the following:

$$lim_{T\to\infty} P(\tilde{p} = p_{BIC}) = 1 \tag{13}$$

$$lim_{T\to\infty} P(p_{BIC} = 1) = 1 \tag{14}$$

. □

Proof of Theorem 4: Define

$$p_{AIC} = min\{m : 1 \leq m \leq d; L_{AIC}(m) \geq L_{AIC}(h), h = 1, 2, ..., d\}$$

where $L_{AIC}(p) = ML(p) - 2p$. Under the alternative, we prove that:

$$\lim_{T\to\infty} P(\tilde{p} = p_{AIC}) = 1 \tag{15}$$

and,

$$\lim_{T\to\infty} P(p_{AIC} \geq K) \to 1 \tag{16}$$

Now consider the event

$$A_T(q) = \{ \max_{1\leq j\leq d} \sqrt{T}|\hat{\rho}_{\hat{\epsilon}\hat{\epsilon}}(j)| < \sqrt{q\ log(T)}\} \tag{17}$$

By a similar argument as in Theorem 2 we can show that

$$P(A_T(q)) \to 0$$

and so (15) hold and the rest of the proof follows for (16). □

References

[1] A. I. McLeod, W. K. Li, Diagnostic checking arma time series models using squared-residual autocorrelations, Journal of Time Series Analysis 4 (4) (1983) 269–273. arXiv:https://onlinelibrary.wiley.com/doi/pdf/10.1111/j.1467-9892.1983.tb00373.x, doi:10.1111/j.1467-9892.1983.tb00373.x.
URL https://onlinelibrary.wiley.com/doi/abs/10.1111/j.1467-9892.1983.tb00373.x

[2] G. E. P. Box, D. A. Pierce, Distribution of residual autocorrelations in autoregressive-integrated moving average time series models, Journal of the American Statistical Association 65 (332) (1970) 1509–1526.
URL http://www.jstor.org/stable/2284333

[3] G. M. Ljung, G. E. P. Box, On a measure of lack of fit in time series models, Biometrika 65 (2) (1978) 297–303.
URL http://www.jstor.org/stable/2335207

[4] G. E. P. Box, G. Jenkins, Time Series Analysis, Forecasting and Control, Holden-Day, Inc., USA, 1990.

[5] P. J. Brockwell, R. A. Davis, Time series: theory and methods, Springer science & business media, 2009.

[6] A. A. P. Granger C. W. J., An introduction to bilinear time series models., Vandenhoeck & Ruprecht, Gottingen, 1978.

[7] R. Luukkonen, P. Saikkonen, T. Teräsvirta, Testing linearity in univariate time series models, Scandinavian Journal of Statistics 15 (3) (1988) 161–175.
URL http://www.jstor.org/stable/4616098

[8] Z. Psaradakis, M. Vávra, Portmanteau tests for linearity of stationary time series, Econometric Reviews 38 (2) (2019) 248–262. arXiv:https://doi.org/10.1080/07474938.2016.1261015, doi:10.1080/07474938.2016.1261015.
URL https://doi.org/10.1080/07474938.2016.1261015

[9] J. C. Escanciano, I. N. Lobato, An automatic portmanteau test for serial correlation, Journal of Econometrics 151 (2) (2009) 140 – 149, recent Advances in Time Series Analysis: A Volume Honouring Peter M. Robinson. doi:https://doi.org/10.1016/j.jeconom.2009.03.001.
URL http://www.sciencedirect.com/science/article/pii/S0304407609000773

[10] T. Inglot, T. Ledwina, Towards data driven selection of a penalty function for data driven neyman tests, Linear Algebra and its Applications 417 (2006) 124–133. doi:10.1016/j.laa.2005.10.023.

[11] G. Schwarz, Estimating the dimension of a model, Ann. Statist. 6 (2) (1978) 461–464. doi:10.1214/aos/1176344136.
URL https://doi.org/10.1214/aos/1176344136

[12] H. Akaike, A new look at the statistical model identification, IEEE Transactions on Automatic Control 19 (6) (1974) 716–723.

[13] J. Kreiss, Estimation of the distribution function of noise in stationary processes, Metrika: International Journal for Theoretical and Applied Statistics 38 (1) (1991) 285–297. doi:10.1007/BF02613623.
URL https://ideas.repec.org/a/spr/metrik/v38y1991i1p285-297.html

[14] G. M. Kuersteiner, Optimal instrumental variables estimation for arma models, Journal of Econometrics 104 (2) (2001) 359 – 405. doi:https://doi.org/10.1016/S0304-4076(01)00088-4.
URL http://www.sciencedirect.com/science/article/pii/S0304407601000884

[15] T. W. Anderson, The asymptotic distributions of autoregressive coefficients, Tech. rep., STANFORD UNIV CA DEPT OF STATISTICS (1991).

[16] A. Pérez, E. Ruiz, Properties of the Sample Autocorrelations of Nonlinear Transformations in Long-Memory Stochastic Volatility Models, Journal of Financial Econometrics 1 (3) (2003) 420–444. arXiv:https://academic.oup.com/jfec/article-pdf/1/3/420/2489457/nbg017.pdf, doi:10.1093/jjfinec/nbg017.
URL https://doi.org/10.1093/jjfinec/nbg017

[17] H. Leeb, B. M. Pötscher, Model selection and inference: Facts and fiction, Econometric Theory (2005) 21—-59.

[18] L. Fenga, D. N. Politis, Bootstrap-based arma order selection, Journal of Statistical Computation and Simulation 81 (2011) 799 – 814.

[19] P. M. T. Broersen, S. de Waele, Finite sample properties of arma order selection, IEEE Transactions on Instrumentation and Measurement 53 (3) (2004) 645–651. doi:10.1109/TIM.2004.827058.

[20] S. T. Buckland, K. P. Burnham, N. H. Augustin, Model selection: An integral part of inference, Biometrics 53 (2) (1997) 603–618.
URL http://www.jstor.org/stable/2533961

[21] A. J. Lawrance, P. A. W. Lewis, Higher-order residual analysis for nonlinear time series with autoregressive correlation structures, International Statistical Review / Revue Internationale de Statistique 55 (1) (1987) 21–35.
URL http://www.jstor.org/stable/1403268

[22] A. Berg, E. Paparoditis, D. N. Politis, A bootstrap test for time series linearity, Journal of Statistical Planning and Inference 140 (12) (2010) 3841–3857.

[23] S. Giannerini, E. Maasoumi, E. B. Dagum, Entropy testing for nonlinear serial dependence in time series, Biometrika 102 (3) (2015) 661–675. arXiv:https://academic.oup.com/biomet/article-pdf/102/3/661/17460650/asv007.pdf, doi:10.1093/biomet/asv007.
URL https://doi.org/10.1093/biomet/asv007

[24] E. J. Hannan, The asymptotic theory of linear time-series models, Journal of Applied Probability 10 (1) (1973) 130–145.
URL http://www.jstor.org/stable/3212501

[25] T.-H. Lee, H. White, C. W. Granger, Testing for neglected nonlinearity in time series models: A comparison of neural network methods and alternative tests, Journal of Econometrics 56 (3) (1993) 269–290. doi:https://doi.org/10.1016/0304-4076(93)90122-L.
URL https://www.sciencedirect.com/science/article/pii/030440769390122L

[26] J. P. Romano, L. A. Thombs, Inference for autocorrelations under weak assumptions, Journal of the American Statistical Association 91 (434) (1996) 590–600. arXiv:https://www.tandfonline.com/doi/pdf/10.1080/01621459.1996.10476928, doi:10.1080/01621459.1996.10476928.
URL https://www.tandfonline.com/doi/abs/10.1080/01621459.1996.10476928

[27] J. C. Escanciano, I. N. Lobato, L. Zhu, Automatic specification testing for vector autoregressions and multivariate nonlinear time series models, Journal of Business & Economic Statistics 31 (4) (2013) 426–437. arXiv:https://doi.org/10.1080/07350015.2013.803973, doi:10.1080/07350015.2013.803973.
URL https://doi.org/10.1080/07350015.2013.803973

[28] P. Billingsley, Probability and Measure, 3rd Edition, John Wiley and Sons, 1986.