Contents lists available at ScienceDirect

Econometrics and Statistics

journal homepage: www.elsevier.com/locate/ecosta

An Automatic Portmanteau Test For Nonlinear Dependence

Charisios Grivas*

Birkbeck University of London, Malet Street, London, WC1E 7HX, United Kingdom

ARTICLE INFO

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 the type I error without being sensitive with respect to the number of autocorrelations used. In addition, the automatic test is found to have higher power in simulations when compared to the standard portmanteau 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 Box and Pierce (1970)), or its finite-sample correction developed by Ljung and Box (1978). 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 Box and Jenkins (1990), 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 Brockwell and Davis (2009) (page 312).

However, as Granger C. W. J. (1978) observed, the autocorrelation function of the squared series can be useful in identifying non-linearity 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, McLeod and Li (1983) proposed a modification of the Ljung–Box test based on the squared series or squared residuals from a fitted model. As Luukkonen et al. (1988) demonstrate, the McLeod-Li($\mathcal{ML}$) test 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 Psaradakis and Vávra (2019)). In fact, Luukkonen et al. (1988) call it an ARCH test to "remind the reader that it is an ARCH test". 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.

---

* Corresponding author.
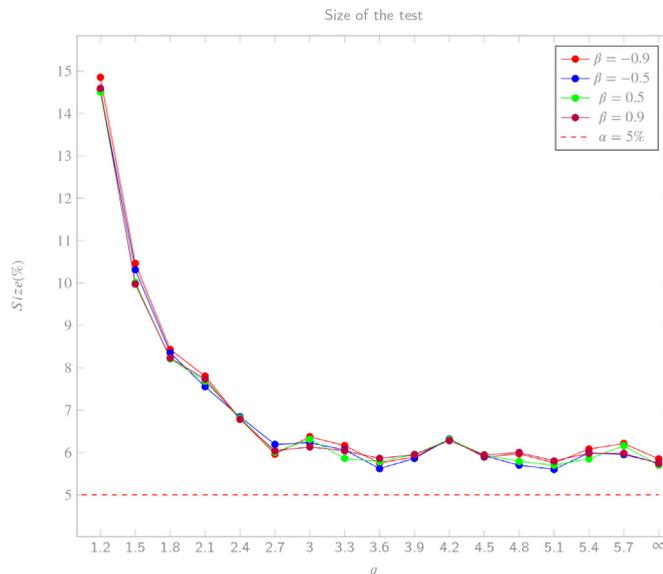 *E-mail address:* Charisios.grivas@gmail.com

Fig. 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 $T = 500$.

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 Escanciano and Lobato (2009), 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 Escanciano and Lobato (2009), 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.

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 $\gamma(j) = \text{cov}\left(y_t^2, y_{t-j}^2\right)$, with $\bar{\delta} = \frac{1}{T} \sum_{t=1}^T y_t^2$. The portmanteau test of McLeod and Li (1983) 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 $T \to \infty$.
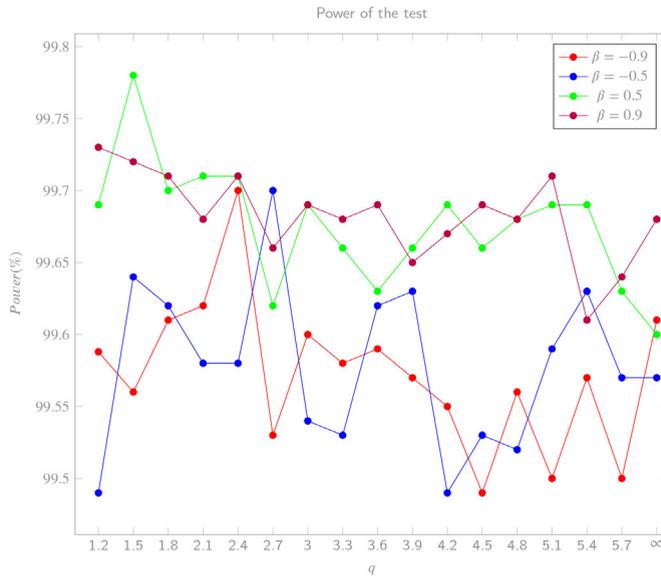
Fig. 2. Rejection frequencies at 5% level for the $\mathcal{AML}$ test on the residuals of an $AR(1)$ model, where the error follows a EGARCH, and for different values of constant q. Simulations are based on $T = 500$.

2.2. Automatic Portmanteau Test

In their novel paper, Escanciano and Lobato (2009) 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 Inglot and Ledwina (2006) and Escanciano and Lobato (2009), 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 (Schwarz, 1978) and the Akaike information criterion (AIC) (Akaike, 1974). 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. The threshold $\sqrt{q \log T}$ employed here is an adaptation of the standard solution for a white noise sequence of *i.i.d* $N(0, 1)$ random variables; see Inglot and Ledwina (2006) (page 126-127) for a motivation of such a data-driven rule.

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* $T \rightarrow \infty$.

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* $T \rightarrow \infty$*, against the alternative* $H_a^K : \rho(1) = \cdots = \rho(K-1) = 0$, $\rho(K) \neq 0$ *for* $1 \leq K \leq d$.
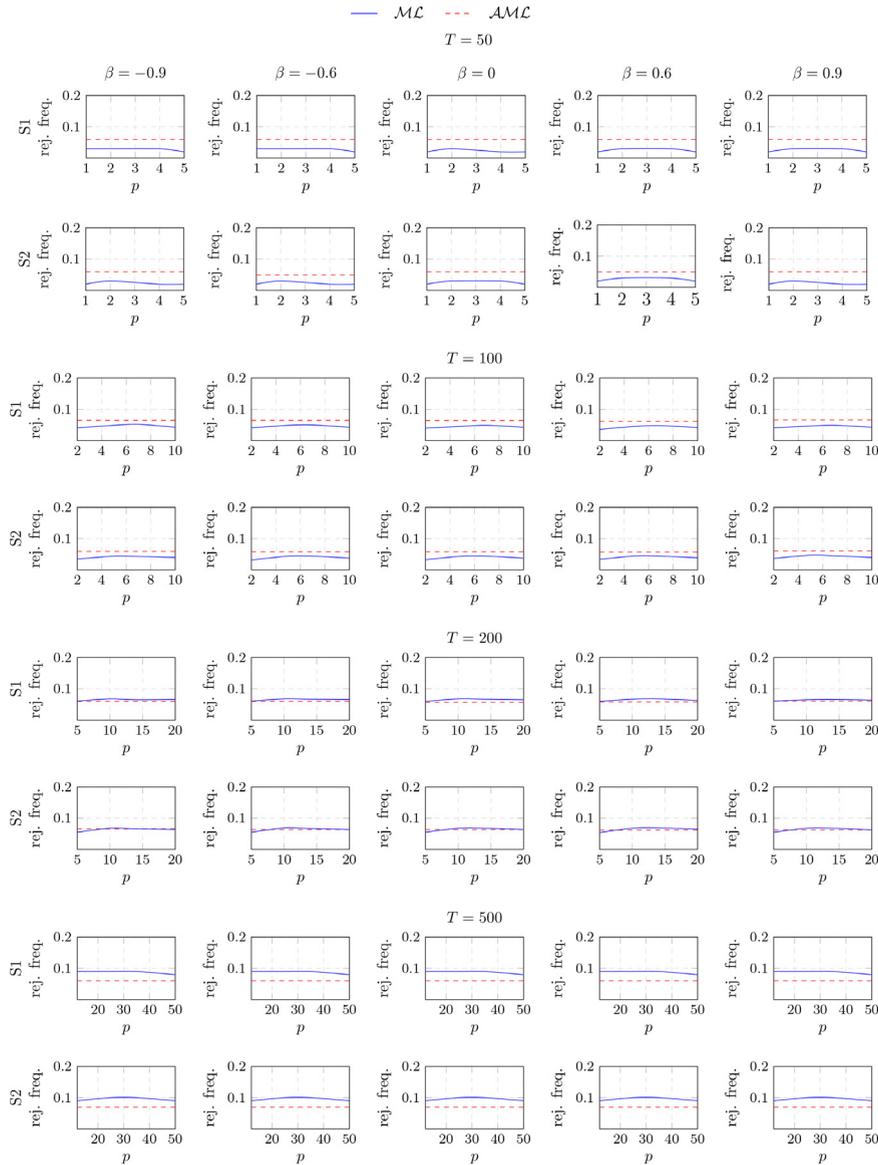
The proofs of the theorems are in the Appendix.

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

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), as considered in McLeod and Li (1983), 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.

Following McLeod and Li (1983), 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 mis-
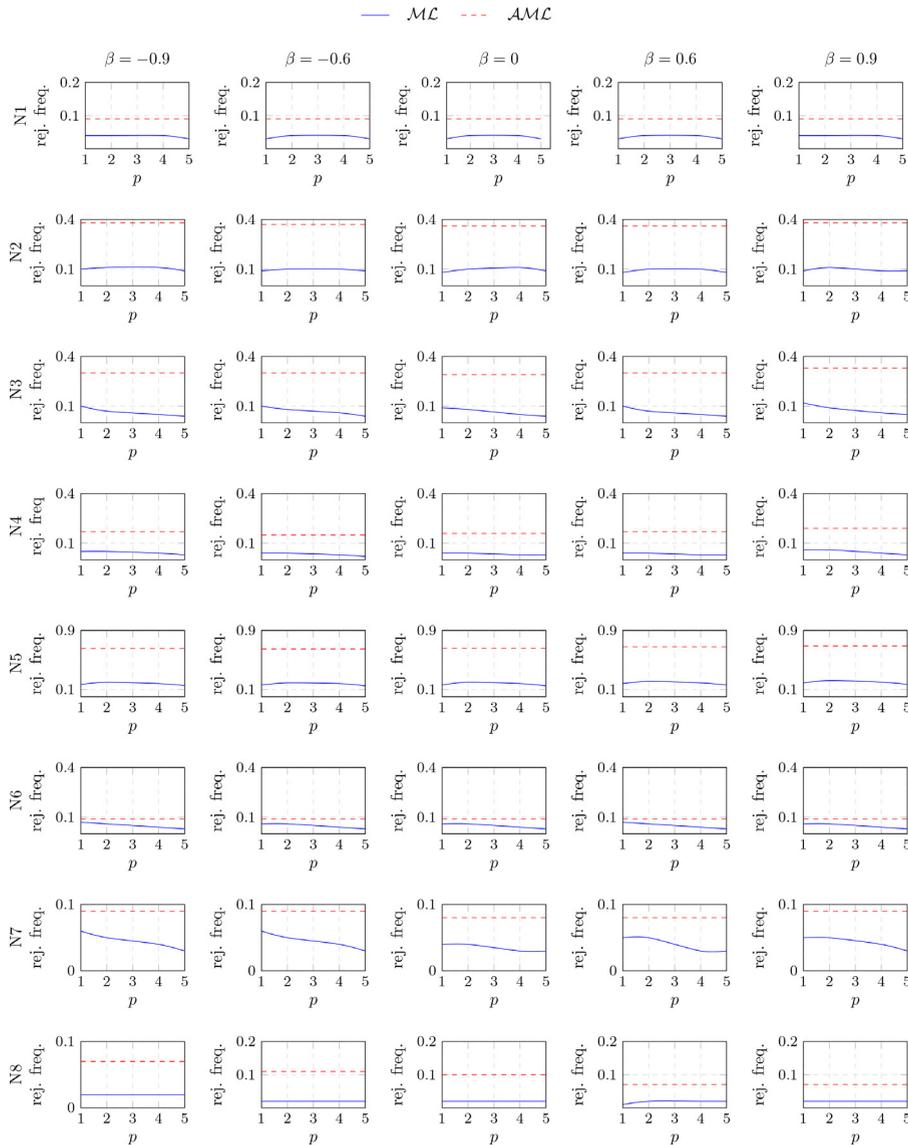
Fig. 4. Rejection frequencies of $\mathcal{ML}$ and $\mathcal{AML}$ for the residuals under the alternative for $T = 50$ at 5% singificance level.

specification of the fitted model, or can, if taken into consideration, lead to improved forecast accuracy (Granger C. W. J., 1978).

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., Kreiss (1991)):

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

Suitable estimators of θ can be obtained by quasi-maximum likelihood, instrumental variables, or least-squares methods (see, e.g., Kuersteiner (2001) 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:

$$\mathcal{ML}^*(p) = T(T + 2) \sum_{j=1}^{p} \frac{\hat{\rho}_{\hat{\epsilon}\hat{\epsilon}}^2(j)}{T - j}, \tag{5}$$
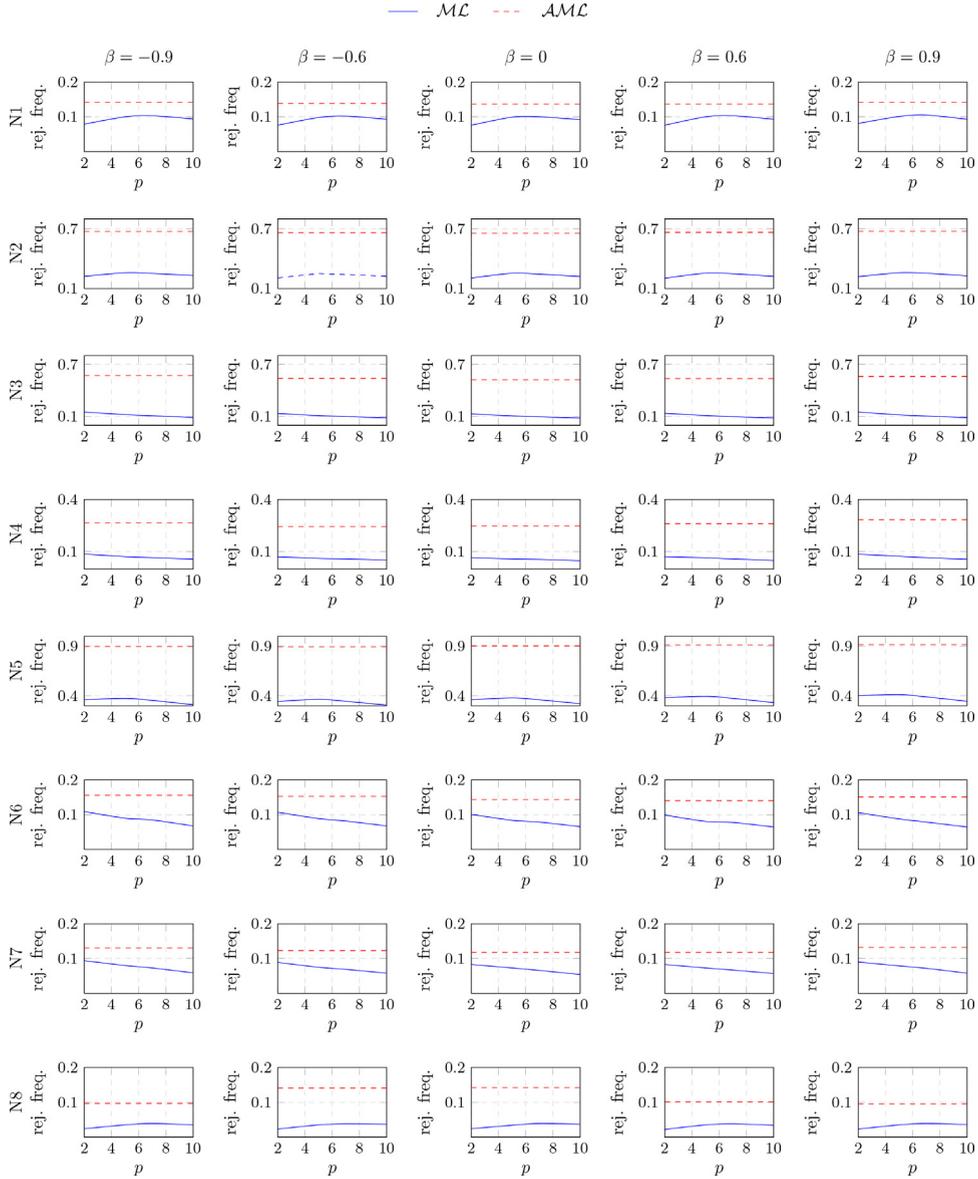
Fig. 5. Rejectionfrequencies of $\mathcal{ML}$ and $\mathcal{AML}$ for the residuals under thealternative, for $T = 100$ and at 5% singificance level.

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 Psaradakis and Vávra (2019)).

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.

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) $\sqrt{T}|\partial \tilde{\gamma}_{\epsilon\epsilon}(j)/\partial \theta| = O_p(1)$ 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 $\bar{\delta}_{\epsilon} = \frac{1}{T}\sum_{t=1}^{T}\epsilon_t^2$.
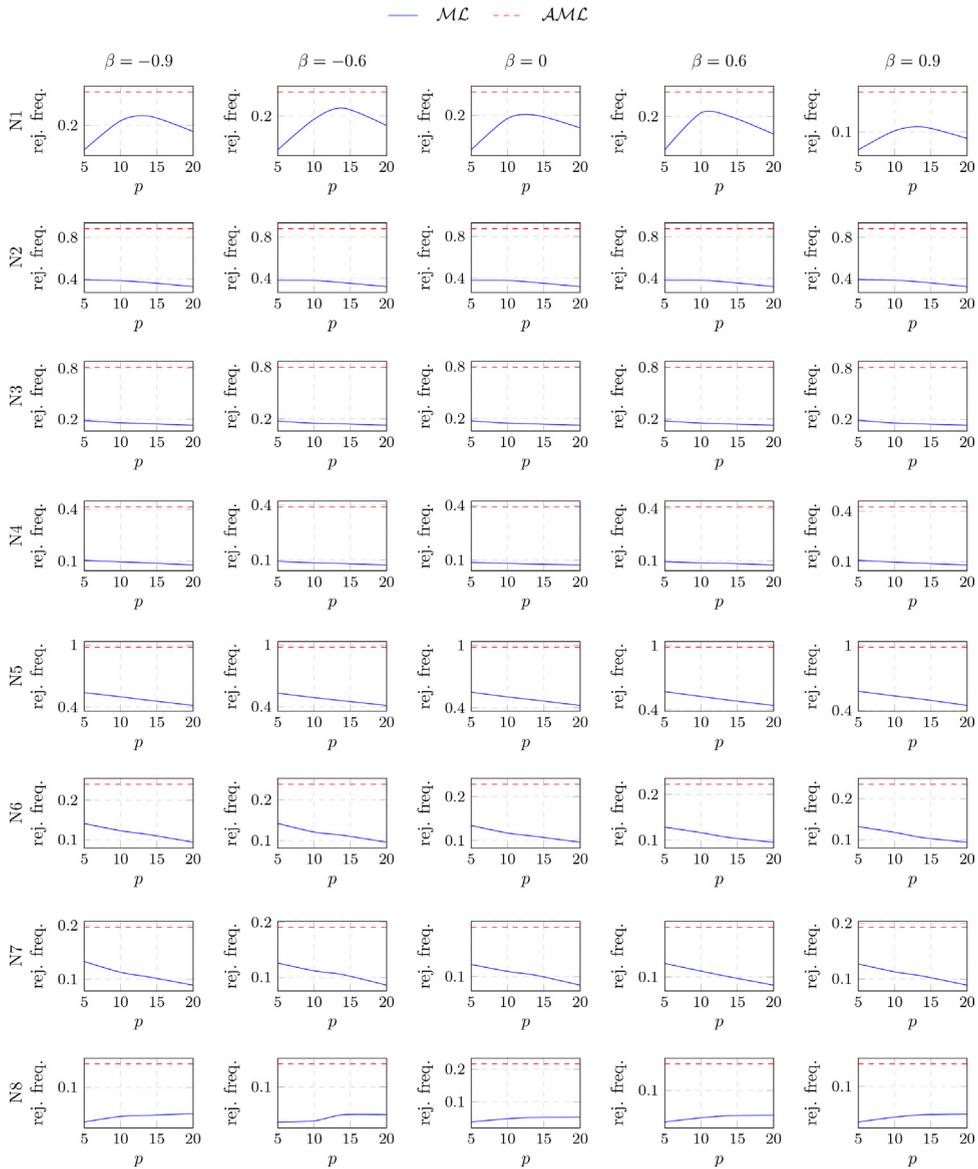
Fig. 6. Rejectionfrequencies of $\mathcal{ML}$ and $\mathcal{AML}$ for the residuals under the alternative, for $T = 200$ and at 5% singificance level.

These conditions, which are similar to those of Psaradakis and Vávra (2019), 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., Brockwell and Davis (2009), 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 $\gamma_{\epsilon\epsilon}(j) = \mathrm{cov}\left(\epsilon_t^2, \epsilon_{t-j}^2\right)$.

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.

Note that the assumption that $\mathbb{E}[\epsilon_t^8] < \infty$ is a standard requirement for $\mathcal{ML}$-type tests (e.g., McLeod and Li (1983), Psaradakis and Vávra (2019)) and ensures that sample autocovariances of squared errors have a well-defined asymptotic
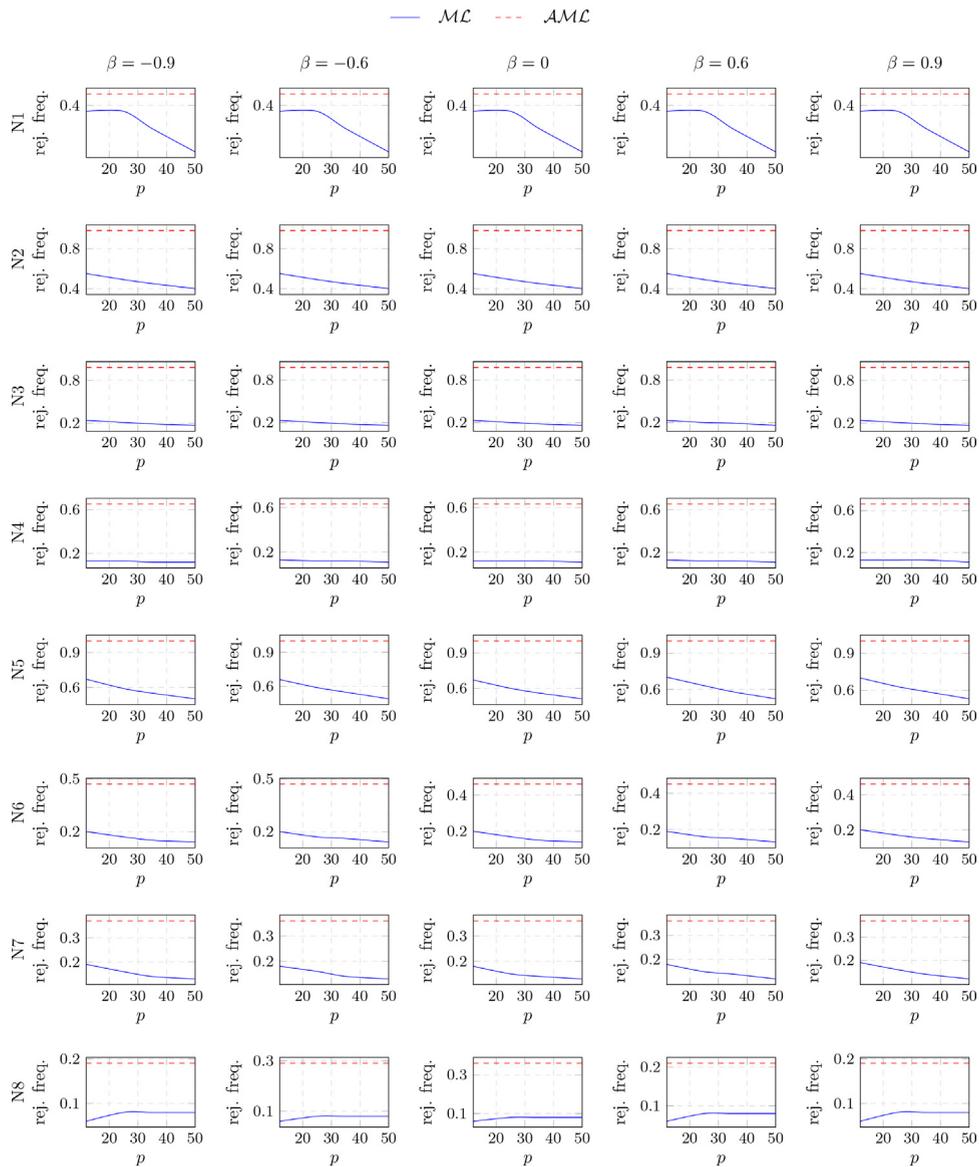
Fig. 7. Rejection frequencies of $\mathcal{ML}$ and $\mathcal{AML}$ for the residuals under the alternative, for $T = 500$ and at 5% singificance level.

distribution when the errors are not autocorrelated. It may be possible to relax the moment condition somewhat along the lines of Anderson (1991), 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 Pérez and Ruiz (2003), 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 Leeb and Pötscher (2005), 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 Fenga and Politis (2011). For example, Broersen and de Waele (2004) suggest a model selection method based on hierarchical models. They develop 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, Buckland et al. (1997) 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 McLeod and Li (1983), Lawrance and Lewis (1987), Berg et al. (2010) and Giannerini et al. (2015), among others. 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. Lee et al. (1993), for example, use a first-order autoregressive model to obtain the required residuals, while McLeod and Li (1983) rely on autoregressive models the order of which is estimated by means of information criteria. Note however, that this is not the only notion of non-linearity in the literature. For example, Hannan (1973) considers a different characterization based on the properties of one-step-ahead linear predictors.

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 standard $\mathcal{ML}$, in the presence and absence of nonlinear serial dependence in the errors of a model. To conserve space, we only consider the case of residuals with a slight abuse of notation, denoting as $\mathcal{AML}$ and $\mathcal{ML}$ the tests applied to the residuals instead of $\mathcal{AML}^\star$ and $\mathcal{ML}^\star$. The case of raw data is provided in the supplement.

In all simulations 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 considered, for comparing the empirical significance levels of $\mathcal{AML}$ and $\mathcal{ML}$, is $\alpha = 5\%$. To conserve space, simulation results for nominal levels 0.01 and 0.10 are not reported here but provided similar results. We set $d = 75$, as in Escanciano and Lobato (2009), except when $T = 50$ where we use $d = 25$, and $q = 3.6$ following our simulations in the next section. Note that in both Escanciano and Lobato (2009) and Inglot and Ledwina (2006) 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 Escanciano and Lobato (2009) and Inglot and Ledwina (2006). 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 $\beta \in \{\pm 0.5, \pm 0.9\}$, for eight different values of q, ($q = 1.2, 1.5, 1.8, 2.4, 2.7, 3, \infty$). Similar results are obtained for other values of the β parameter but are not provided due to space constraints.

Fig. 1 above, presents graphically the rejection probabilities under the null, where the errors are generated from a standard normal distribution. Fig. 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 Fig. 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 Section 3.2, as model N.5. Simulations based on a stochastic volatility model provided similar results and are hence omitted. Again, Fig. 2 shows that $q = 3.6$ is a reasonable choice with high power, that does not vary much.

3.2. Level and Power of Tests based on Residuals

In this subsection, we examine the application of the test on residuals. Following McLeod and Li (1983), 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.

The latter choice is made so that the error term satisfies the finite eighth moment condition. A log-normal distribution was also considered as an alternative to the t-distribution but provided similar results. 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 $\epsilon_t = 0.8u_{t-2}^3 + u_t$.

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 Romano and Thombs (1996) 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 Escanciano and Lobato (2009). The remaining three DGPs are taken from Psaradakis and Vávra (2019); they represent a threshold AR (TAR) [N.6], a smooth-transition AR [N.7], and a nonlinear MA (NLMA) [N.8]. Bilinear models were also considered but provided similar results.

The Monte Carlo rejection frequencies of the conventional $\mathcal{ML}$ test and its automatic version, $\mathcal{AML}$ (at nominal level 5%) are shown in Figs. 3 -7. Under the null, the $\mathcal{AML}$ has empirical size close to the nominal level while the size of $\mathcal{ML}$ varies substantially. Fig. 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.

In Figs. 4-7 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.

Fig. 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.

Fig. 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.

Moving on to sample sizes that are most common in finance, we see that our main conclusions remain unaffected as seen in Figs. 6 and 7. Starting from Fig. 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.

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 Fig. 7.

4. Summary and Conclusion

A data-driven version of the McLeod-Li portmanteau test for detecting nonlinear dependence is proposed. 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 Psaradakis and Vávra (2019), and tests for multivariate time series and comparisons with the procedures developed in Escanciano et al. (2013).

Funding

Declaration of Competing Interest

None.

Appendix A

Proof of Theorem 1. Define

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

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

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

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

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

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

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

and,

$$P(L_{BIC}(i) \geq L_{BIC}(1)) \leq P(ML \geq (i-1)log \, T)$$

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 Billingsley (1986) 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 $L_{AIC}(p) = ML(p) - 2 \, p$. We need to prove that:

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

and

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

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

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

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

$$P(p_{AIC} = k) \leq P(L_{AIC}(k) \geq L_{AIC}(K))$$
$$\leq P(|ML(k)| \geq 2(k-K) + |ML(K)|) \to 0$$

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

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

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

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

Anderson, T.W., 1991. The asymptotic distributions of autoregressive coefficients. Technical Report. STANFORD UNIV CA DEPT OF STATISTICS.

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

Billingsley, P., 1986. Probability and Measure, third John Wiley and Sons.

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

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

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

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

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

Escanciano, J.C., Lobato, I.N., 2009. An automatic portmanteau test for serial correlation. Journal of Econometrics 151 (2), 140–149. https://doi.org/10.1016/j.jeconom.2009.03.001. Recent Advances in Time Series Analysis: A Volume Honouring Peter M. Robinson

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

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

Giannerini, S., Maasoumi, E., Dagum, E.B., 2015. Entropy testing for nonlinear serial dependence in time series. Biometrika 102 (3), 661–675. doi:10.1093/biomet/asv007.

Granger CWJ, A.A.P., 1978. An introduction to bilinear time series models. Vandenhoeck & Ruprecht, Gottingen.

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

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

Kreiss, J., 1991. Estimation of the distribution function of noise in stationary processes. Metrika: International Journal for Theoretical and Applied Statistics 38 (1), 285–297. doi:10.1007/BF02613623.

Kuersteiner, G.M., 2001. Optimal instrumental variables estimation for arma models. Journal of Econometrics 104 (2), 359–405. doi:10.1016/S0304-4076(01)00088-4.

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

Lee, T.-H., White, H., Granger, C.W., 1993. Testing for neglected nonlinearity in time series models: A comparison of neural network methods and alternative tests. Journal of Econometrics 56 (3), 269–290. doi:10.1016/0304-4076(93)90122-L.

Leeb, H., Pötscher, B.M., 2005. Model selection and inference: Facts and fiction. Econometric Theory 21–59.

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

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

McLeod, A.I., Li, W.K., 1983. Diagnostic checking arma time series models using squared-residual autocorrelations. Journal of Time Series Analysis 4 (4), 269–273. doi:10.1111/j.1467-9892.1983.tb00373.x.

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

Pérez, A., Ruiz, E., 2003. Properties of the Sample Autocorrelations of Nonlinear Transformations in Long-Memory Stochastic Volatility Models. Journal of Financial Econometrics 1 (3), 420–444. doi:10.1093/jjfinec/nbg017.

ARTICLE IN PRESS

Romano, J.P., Thombs, L.A., 1996. Inference for autocorrelations under weak assumptions. Journal of the American Statistical Association 91 (434), 590–600. doi:10.1080/01621459.1996.10476928.

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