# On Diagnostic Checking Time Series Models with Portmanteau Test Statistics Based on Generalized Inverses and {2}-Inverses

Pierre Duchesne[1] and Christian Francq[2]

[1] Département de mathématiques et statistique, Université de Montréal
C.P. 6128 Succursale Centre-Ville, Montréal, Québec, H3C 3J7, Canada,
*duchesne@dms.umontreal.ca*

[2] Université Lille 3, EQUIPPE-GREMARS, BP 60149, 59653, Villeneuve d'Ascq, cedex, France, *christian.francq@univ-lille3.fr*

**Abstract.** A class of univariate time series models is considered, which allows general specifications for the conditional mean and conditional variance functions. After deriving the asymptotic distributions of the residual autocorrelations based on the standardized residuals, portmanteau test statistics are studied. If the asymptotic covariance of a vector of fixed length of residual autocorrelations is non singular, portmanteau test statistics could be defined, following the approach advocated by Li (1992). However, assuming the invertibility of the asymptotic covariance of the residual autocorrelations may be restrictive, and, alternatively, the popular Box-Pierce-Ljung test statistic may be recommended. In our framework, that test statistic converges in distribution to a weighted sum of chi-square variables, and the critical values can be found using Imhof's (1961) algorithm. However, Imhof's algorithm may be time consuming. In view of this, we investigate in this article the use of generalized inverses and {2}-inverses, in order to propose new test statistics with asymptotic chi-square distributions, avoiding the need to implement Imhof's algorithm. In a small simulation study, the following test statistics are compared: Box-Pierce-Ljung test statistic, the test statistic based on the proposal of Li (1992), and the new test statistics relying on generalized inverses and {2}-inverses.

**Keywords:** Conditional heteroscedasticity, diagnostic checking, generalized inverses, portmanteau test statistics, residual autocorrelations

## 1 Introduction

Let $\{Y_t\}$ be a stationary stochastic process. We consider the following univariate time series model:

$$Y_t = m_{\boldsymbol{\theta}_0}(Y_{t-1}, Y_{t-2}, \dots) + \sigma_{\boldsymbol{\theta}_0}(Y_{t-1}, Y_{t-2}, \dots)\eta_t, \qquad (1)$$

where $\boldsymbol{\theta}_0$ denotes a $s$ dimensional vector of unknown parameters belonging to a subset $\Theta$, where $\Theta \subset \mathbb{R}^s$. The error process $\{\eta_t\}$ is an independent and identically distributed (iid) sequence of random variables with mean zero and unit variance. It is assumed that the random variable $\eta_t$ is independent of

$\{Y_{t-i}, i > 0\}$. The nonlinear model (1) represents a very general class of time series models with a general specification for the error term. It includes the classical autoregressive moving-average (ARMA) time series model, with possible [general] conditional heteroskedasticity ([G]ARCH) in the error process, and also nonlinear models, such as threshold autoregressive models (TAR), self-exciting TAR models (SETAR), and smooth versions of TAR models. Tong (1990) and Granger and Teräsvirta (1993) provide surveys of univariate nonlinear models.

Let $Y_t$, $t = 1, \ldots, n$ be a finite realization of the stochastic process $\{Y_t\}$. An important practical aspect is to validate an adjusted model such as (1), using estimation procedures such as quasi-maximum likelihood (QML) and nonlinear least squares (NLS) methods (the latter being obtained by assuming a constant conditional variance). Klimko and Nelson (1978) investigated general properties of conditional least squares estimators in univariate nonlinear time series. See also Potscher and Prucha (1997) and Taniguchi and Kakizawa (2000), amongst others.

Residual autocorrelations have been found useful for checking model adequacy of many time series models (see, e.g., Li (2004)). In view of this fact, we first derive the asymptotic distributions of the residual autocorrelations based on the standardized residuals. As an application of that result, portmanteau test statistics are studied. If the asymptotic covariance of a vector of fixed length of residual autocorrelations is non singular, portmanteau test statistics could be defined, following the approach advocated by Li (1992). However, assuming the invertibility of the asymptotic covariance of the residual autocorrelations may be somewhat restrictive. For example, in validating an ARMA model with an iid error term, it is well-known that the asymptotic covariance matrix of a vector of fixed length of residual autocorrelations is approximatively idempotent, with rank $n - p - q$, where $p$ and $q$ correspond to the autoregressive and moving average orders, respectively. On the other hand, if model (1) represents a nonlinear time series model, such as the TAR model considered in Li (1992), then, under some conditions, the asymptotic covariance matrix is expected to be non-singular. See also Li (2004, pp. 79-80). For a given model, the precise conditions which guarantee the invertibility of the asymptotic covariance matrix may be hard to obtain. Alternatively, the popular Box-Pierce-Ljung test statistic may be recommended (see Li (2004), amongst others). In our framework, this test statistic converges in distribution to a weighted sum of chi-square variables, where, in practice, the weights are determined with the data (see Francq, Roy and Zakoïan (2005) for general results in the context of ARMA models with weak errors). Interestingly, the range of applicability of Box-Pierce-Ljung test statistic appears to be more general, in the sense that if the asymptotic covariance matrix is non-singular, then all weights are strickly positive. However, contrary to the test procedure of Li (1992, 2004), the Box-Pierce-Ljung test statistic is still appropriate in linear time series models: for an AR(1) time series model, say, one weight in the weighted sum of chi-square variables is identically equal to zero, and the others weights are strictly positives. In practice, the critical values of the Box-Pierce-Ljung test statistic can be found using Imhof's algorithm. Even today, it may be still time consuming to implement this algorithm, since, to the best of our knowledge, Imhof's algorithm is not actually available in popular softwares such as S-PLUS or R.

Since the asymptotic covariance matrix of a vector of fixed length of residual autocorrelations may be essentially singular in linear time series models, and, under certain assumptions, invertible in non-linear time series models, we investigate here the use of generalized inverses, such as the Moore-Penrose inverse, and also of $\{2\}$-inverses of that covariance matrix. This leads us to propose new portmanteau test statistics with asymptotic chi-square distributions. These new test statistics avoid the need to implement Imhof's algorithm. In a small simulation study, the following test statistics are compared with respect to level and power: Box-Pierce-Ljung test statistic, the test statistic based on the proposal of Li (1992), a new test statistic relying on the Moore-Penrose inverse, and several new proposals relying on $\{2\}$-inverses. The rest of the paper is organized as follows. In Section 2, we derive the asymptotic distribution of the residual autocorrelations. Classical portmanteau test statistics are discussed in Section 3. In Section 4, modified test statistics are presented. A small simulation study is conducted in Section 5.

## 2 Asymptotic distribution of the residual autocorrelations

Consider model (1). The first and second conditional moments are given by:

$$\begin{aligned} m_t(\boldsymbol{\theta}_0) &:= m_{\boldsymbol{\theta}_0}(Y_{t-1}, Y_{t-2}, \dots) = E(Y_t \mid Y_{t-1}, Y_{t-2}, \dots), \\ \sigma_t^2(\boldsymbol{\theta}_0) &:= \sigma^2_{\boldsymbol{\theta}_0}(Y_{t-1}, Y_{t-2}, \dots) = \mathrm{Var}(Y_t \mid Y_{t-1}, Y_{t-2}, \dots), \end{aligned}$$

respectively. Given the time series data $Y_1, \dots, Y_n$, and the initial values $Y_0 = y_0, Y_{-1} = y_{-1}, \dots$, at any $\boldsymbol{\theta} \in \Theta$ the conditional moments $m_t(\boldsymbol{\theta})$ and $\sigma_t^2(\boldsymbol{\theta})$ can be approximated by the measurable functions defined by $\tilde{m}_t(\boldsymbol{\theta}) = m_{\boldsymbol{\theta}}(Y_{t-1}, \dots, Y_1, y_0, \dots)$ and $\tilde{\sigma}_t^2(\boldsymbol{\theta}) = \sigma^2_{\boldsymbol{\theta}}(Y_{t-1}, \dots, Y_1, y_0, \dots)$, respectively. A natural choice for the initial values is to specify $Y_i = 0$ for all $i \le 0$. A QML estimator of $\boldsymbol{\theta}_0$ is defined as any measurable solution $\hat{\boldsymbol{\theta}}_n$ of

$$\hat{\boldsymbol{\theta}}_n = \arg \inf_{\boldsymbol{\theta} \in \Theta} \tilde{Q}_n(\boldsymbol{\theta}),$$

where $\tilde{Q}_n(\boldsymbol{\theta}) = n^{-1} \sum_{t=1}^n \tilde{\ell}_t$ and $\tilde{\ell}_t = \tilde{\ell}_t(\boldsymbol{\theta}) = (Y_t - \tilde{m}_t)^2 / \tilde{\sigma}_t^2 + \log \tilde{\sigma}_t^2$. It can be shown that the QML estimator is consistent and asymptotically normal under Assumption A.

**Assumption A**: *(i)* $\Theta$ *represents a compact set and the functions* $\boldsymbol{\theta} \to \tilde{m}_t(\boldsymbol{\theta})$ *and* $\boldsymbol{\theta} \to \tilde{\sigma}_t^2(\boldsymbol{\theta}) > 0$ *are continuous; (ii)* $\{Y_t\}$ *corresponds to a non anticipative strictly stationary and ergodic solution of (1); (iii)* $E \log^- \sigma_t^2(\boldsymbol{\theta}) <$

*$\infty$ for all $\boldsymbol{\theta} \in \Theta$, and $E \log^+ \sigma_t^2(\boldsymbol{\theta}_0) < \infty$; (iv) $\sup_{\boldsymbol{\theta} \in \Theta} \left| \ell_t - \tilde{\ell}_t \right| \to 0$ a.s. as $t \to \infty$, where $\ell_t(\boldsymbol{\theta}) = (Y_t - m_t)^2/\sigma_t^2 + \log \sigma_t^2$; (v) if $\boldsymbol{\theta} \neq \boldsymbol{\theta}_0$ then $m_t(\boldsymbol{\theta}) \neq m_t(\boldsymbol{\theta}_0)$ or $\sigma_t^2(\boldsymbol{\theta}) \neq \sigma_t^2(\boldsymbol{\theta}_0)$ with non zero probability; (vi) $\boldsymbol{\theta}_0$ belongs to the interior $\overset{\circ}{\Theta}$ of $\Theta$; (vii) $\boldsymbol{\theta} \to m_t(\boldsymbol{\theta})$ and $\boldsymbol{\theta} \to \sigma_t(\boldsymbol{\theta})$ admit continuous third order derivatives, and*

$$E \sup_{\boldsymbol{\theta} \in \Theta} \left| \frac{\partial^3 \ell_t(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}_i \partial \boldsymbol{\theta}_j \partial \boldsymbol{\theta}_k} \right| < \infty \quad \forall i, j, k.$$

*(viii) The moments $\mu_i = E\eta_t^i$, $i \leq 4$, and the information matrices $\boldsymbol{I} = E[\{\partial \ell_t(\boldsymbol{\theta}_0)/\partial \boldsymbol{\theta}\}\{\partial \ell_t(\boldsymbol{\theta}_0)/\partial \boldsymbol{\theta}'\}]$ and $\boldsymbol{J} = E\{\partial^2 \ell_t(\boldsymbol{\theta}_0)/\partial \boldsymbol{\theta} \partial \boldsymbol{\theta}'\}$ exist. Furthermore, $\boldsymbol{I}$ and $\boldsymbol{J}$ are supposed to be non singular.*

Write $a \overset{c}{=} b$ when $a = b + c$. Under Assumption A:

$$\sqrt{n}(\hat{\boldsymbol{\theta}}_n - \boldsymbol{\theta}_0) \overset{o_P(1)}{=} -\boldsymbol{J}^{-1} \frac{1}{\sqrt{n}} \sum_{t=1}^{n} \boldsymbol{Z}_t \overset{\mathcal{L}}{\to} \mathcal{N}\left(\boldsymbol{0}, \boldsymbol{\Sigma}_{\hat{\boldsymbol{\theta}}_n}\right) \tag{2}$$

as $n \to \infty$, where $\boldsymbol{\Sigma}_{\hat{\boldsymbol{\theta}}_n} := \boldsymbol{J}^{-1} \boldsymbol{I} \boldsymbol{J}^{-1}$ and

$$\boldsymbol{Z}_t = -2 \frac{\eta_t}{\sigma_t} \frac{\partial m_t(\boldsymbol{\theta}_0)}{\partial \boldsymbol{\theta}} + \left\{1 - \eta_t^2\right\} \frac{1}{\sigma_t^2} \frac{\partial \sigma_t^2(\boldsymbol{\theta}_0)}{\partial \boldsymbol{\theta}}.$$

Following the current practice, the information matrices $\boldsymbol{I}$ and $\boldsymbol{J}$ are consistently estimated by their empirical counterparts, that is by the formula $\hat{\boldsymbol{I}} = n^{-1} \sum_{t=1}^{n} \{\partial \tilde{\ell}_t(\hat{\boldsymbol{\theta}}_n)/\partial \boldsymbol{\theta}\}\{\partial \tilde{\ell}_t(\hat{\boldsymbol{\theta}}_n)/\partial \boldsymbol{\theta}'\}$ and $\hat{\boldsymbol{J}} = n^{-1} \sum_{t=1}^{n} \partial^2 \tilde{\ell}_t(\hat{\boldsymbol{\theta}}_n)/\partial \boldsymbol{\theta} \partial \boldsymbol{\theta}'$, respectively.

Define the following standardized residuals:

$$\hat{\eta}_t = \frac{Y_t - \tilde{m}_t(\hat{\boldsymbol{\theta}}_n)}{\tilde{\sigma}_t(\hat{\boldsymbol{\theta}}_n)}, \quad t = 1, \ldots, n.$$

Portmanteau test statistics based on the autocorrelations of the residuals are routinely performed for model adequacy checking. In order to derive the asymptotic distribution of the residual autocorrelations, some additional notations are needed. Let

$$\eta_t(\boldsymbol{\theta}) = \frac{Y_t - m_t(\boldsymbol{\theta})}{\sigma_t(\boldsymbol{\theta})}, \quad \tilde{\eta}_t(\boldsymbol{\theta}) = \frac{Y_t - \tilde{m}_t(\boldsymbol{\theta})}{\tilde{\sigma}_t(\boldsymbol{\theta})},$$

so that $\eta_t = \eta_t(\boldsymbol{\theta}_0)$ and $\hat{\eta}_t = \tilde{\eta}_t(\hat{\boldsymbol{\theta}}_n)$. For any fixed integer $m \geq 1$, let

$$\boldsymbol{\gamma}_m = (\gamma(1), \ldots, \gamma(m))', \quad \boldsymbol{\rho}_m = (\rho(1), \ldots, \rho(m))',$$

where, for $\ell \geq 0$,

$$\gamma(\ell) = \frac{1}{n} \sum_{t=1}^{n-\ell} \eta_t \eta_{t+\ell} \quad \text{and} \quad \rho(\ell) = \frac{\gamma(\ell)}{\gamma(0)}.$$

Note that $\boldsymbol{\gamma}_m \overset{o_P(1)}{=} n^{-1}\sum_{t=1}^n \boldsymbol{\Upsilon}_t\boldsymbol{\Upsilon}_t'$ where $\boldsymbol{\Upsilon}_t = \eta_t\boldsymbol{\eta}_{t-1:t-m}$ and $\boldsymbol{\eta}_{t-1:t-m} = (\eta_{t-1},\dots,\eta_{t-m})'$. In view of (2), the central limit theorem applied to the martingale difference $\left\{(\boldsymbol{Z}_t',\boldsymbol{\Upsilon}_t')';\sigma(\eta_u,\, u\leq t)\right\}$ implies the following asymptotic distribution:

$$\sqrt{n}\begin{pmatrix}\hat{\boldsymbol{\theta}}_n-\boldsymbol{\theta}_0\\ \gamma_m\end{pmatrix} \xrightarrow{\mathcal{L}} \mathcal{N}\left\{\mathbf{0},\begin{pmatrix}\boldsymbol{\Sigma}_{\hat{\boldsymbol{\theta}}_n} & \boldsymbol{\Sigma}_{\hat{\boldsymbol{\theta}}_n\boldsymbol{\Upsilon}_m}\\ \boldsymbol{\Sigma}'_{\hat{\boldsymbol{\theta}}_n\boldsymbol{\Upsilon}_m} & I_m\end{pmatrix}\right\}, \tag{3}$$

where $I_m$ denotes the identity matrix of order $m$ and:

$$\begin{aligned}\boldsymbol{\Sigma}_{\hat{\boldsymbol{\theta}}_n\boldsymbol{\Upsilon}_m} &= -\boldsymbol{J}^{-1}E\boldsymbol{Z}_t\boldsymbol{\Upsilon}_t',\\ &= 2\boldsymbol{J}^{-1}E\frac{1}{\sigma_t}\frac{\partial m_t(\boldsymbol{\theta}_0)}{\partial\boldsymbol{\theta}}\boldsymbol{\eta}'_{t-1:t-m} + \boldsymbol{J}^{-1}\mu_3E\frac{1}{\sigma_t^2}\frac{\partial\sigma_t^2(\boldsymbol{\theta}_0)}{\partial\boldsymbol{\theta}}\boldsymbol{\eta}'_{t-1:t-m}.\end{aligned}$$

We now turn to the residual autocorrelation function $\hat{\rho}(\cdot)$ obtained by replacing $\eta_t$ by $\hat{\eta}_t$ in $\rho(\cdot)$. Similarly, define the function $\hat{\gamma}(\cdot)$ and the vectors $\hat{\boldsymbol{\gamma}}_m$ and $\hat{\boldsymbol{\rho}}_m$. A Taylor expansion of the function $\boldsymbol{\theta}\mapsto n^{-1}\sum_{t=1}^{n-\ell}\eta_t(\boldsymbol{\theta})\eta_{t+\ell}(\boldsymbol{\theta})$ around $\hat{\boldsymbol{\theta}}_n$ and $\boldsymbol{\theta}_0$ gives $\hat{\gamma}(\ell) \overset{o_P(1)}{=} \gamma(\ell)+\boldsymbol{c}_\ell'(\hat{\boldsymbol{\theta}}_n-\boldsymbol{\theta}_0)$, where

$$\begin{aligned}\boldsymbol{c}_\ell &= E\eta_{t-\ell}\frac{\partial\eta_t}{\partial\boldsymbol{\theta}}(\boldsymbol{\theta}_0) = -E\frac{\eta_{t-\ell}}{\sigma_t(\boldsymbol{\theta}_0)}\left\{\frac{\partial m_t}{\partial\boldsymbol{\theta}}(\boldsymbol{\theta}_0)+\frac{\eta_t}{2\sigma_t(\boldsymbol{\theta}_0)}\frac{\partial\sigma_t^2}{\partial\boldsymbol{\theta}}(\boldsymbol{\theta}_0)\right\}\\ &= -E\frac{\eta_{t-\ell}}{\sigma_t(\boldsymbol{\theta}_0)}\frac{\partial m_t}{\partial\boldsymbol{\theta}}(\boldsymbol{\theta}_0).\end{aligned}$$

Using (3) and the notation $\boldsymbol{C}_m = (\boldsymbol{c}_1\ \boldsymbol{c}_2\ \cdots\ \boldsymbol{c}_m)$, it follows that:

$$\sqrt{n}\hat{\gamma}_m \xrightarrow{\mathcal{L}} \mathcal{N}\left\{\mathbf{0},\boldsymbol{\Sigma}_{\hat{\gamma}_m}\right\}$$

as $n\to\infty$, where $\boldsymbol{\Sigma}_{\hat{\gamma}_m} = I_m+\boldsymbol{C}_m'\boldsymbol{\Sigma}_{\hat{\boldsymbol{\theta}}_n}\boldsymbol{C}_m+\boldsymbol{C}_m'\boldsymbol{\Sigma}_{\hat{\boldsymbol{\theta}}_n\boldsymbol{\Upsilon}_m}+\boldsymbol{\Sigma}'_{\hat{\boldsymbol{\theta}}_n\boldsymbol{\Upsilon}_m}\boldsymbol{C}_m$. Some simplifications are possible. First, we note that $\boldsymbol{\Sigma}_{\hat{\boldsymbol{\theta}}_n\boldsymbol{\Upsilon}_m} = -2\boldsymbol{J}^{-1}\boldsymbol{C}_m + \mu_3\boldsymbol{J}^{-1}\boldsymbol{D}_m$, where $\boldsymbol{D}_m = (\boldsymbol{d}_1\ \boldsymbol{d}_2\ \cdots\ \boldsymbol{d}_m)$ and $\boldsymbol{d}_\ell = E\eta_{t-\ell}\sigma_t^{-2}\partial\sigma_t^2(\boldsymbol{\theta}_0)/\partial\boldsymbol{\theta}$. Thus, the asymptotic covariance matrix can be written as:

$$\begin{aligned}\boldsymbol{\Sigma}_{\hat{\gamma}_m} = I_m &+ \boldsymbol{C}_m'\boldsymbol{J}^{-1}\boldsymbol{I}\boldsymbol{J}^{-1}\boldsymbol{C}_m - 4\boldsymbol{C}_m'\boldsymbol{J}^{-1}\boldsymbol{C}_m\\ &+\mu_3\left(\boldsymbol{C}_m'\boldsymbol{J}^{-1}\boldsymbol{D}_m+\boldsymbol{D}_m'\boldsymbol{J}^{-1}\boldsymbol{C}_m\right).\end{aligned} \tag{4}$$

Since $\hat{\gamma}(0) = 1+o_P(1)$, the asymptotic distribution of the residual autocorrelations follows easily:

$$\sqrt{n}\hat{\boldsymbol{\rho}}_m \xrightarrow{\mathcal{L}} \mathcal{N}\left\{\mathbf{0},\boldsymbol{\Sigma}_{\hat{\boldsymbol{\rho}}_m}\right\}, \tag{5}$$

where $\boldsymbol{\Sigma}_{\hat{\boldsymbol{\rho}}_m} = \boldsymbol{\Sigma}_{\hat{\gamma}_m}$. One can define empirical estimates $\hat{\boldsymbol{C}}_m$ and $\hat{\boldsymbol{D}}_m$ by replacing $\boldsymbol{c}_\ell$ and $\boldsymbol{d}_\ell$ in $\boldsymbol{C}_m$ and $\boldsymbol{D}_m$ by

$$\hat{\boldsymbol{c}}_\ell = -\frac{1}{n}\sum_{t=\ell+1}^{n}\frac{\hat{\eta}_{t-\ell}}{\sigma_t(\hat{\boldsymbol{\theta}}_n)}\frac{\partial m_t}{\partial\boldsymbol{\theta}}(\hat{\boldsymbol{\theta}}_n) \quad\text{and}\quad \hat{\boldsymbol{d}}_\ell = \frac{1}{n}\sum_{t=\ell+1}^{n}\frac{\hat{\eta}_{t-\ell}}{\sigma_t^2(\hat{\boldsymbol{\theta}}_n)}\frac{\partial\sigma_t^2}{\partial\boldsymbol{\theta}}(\hat{\boldsymbol{\theta}}_n).$$

We then obtain an estimator $\hat{\boldsymbol{\Sigma}}_{\hat{\boldsymbol{\rho}}_m}$ of $\boldsymbol{\Sigma}_{\hat{\boldsymbol{\rho}}_m}$ by replacing $\mu_3$ and the matrices $\boldsymbol{I}$, $\boldsymbol{J}$, $\boldsymbol{C}_m$ and $\boldsymbol{D}_m$ by their empirical counterparts in (4).

## 3 Classical portmanteau tests

### 3.1 Box-Pierce-Ljung test statistics

For checking the adequacy of an ARMA$(p, q)$ model, it is customary to employ the so-called portmanteau tests, such as the Box-Pierce-Ljung test statistic $Q_m^{BPL} = n(n+2)\sum_{i=1}^{m} \hat{\rho}^2(i)/(n-i)$. When diagnosing ARMA models, the null hypothesis of an ARMA$(p, q)$ model is rejected at the nominal level $\alpha$ when $Q_m^{BPL} > \chi^2_{m-(p+q)}(1-\alpha)$, where $m > p+q$ and $\chi^2_\ell(1-\alpha)$ denotes the $(1-\alpha)$-quantile of a $\chi^2$ distribution with $\ell$ degrees of freedom.

More generally, when the conditional mean $m_{\boldsymbol{\theta}_0}(\cdot)$ and the conditional variance $\sigma_{\boldsymbol{\theta}_0}(\cdot)$ are well specified in (1), the residual autocorrelations $\hat{\rho}(h)$ are expected to be close to zero for all $h \neq 0$. Therefore, it is natural to reject the null hypothesis $H_0$ that the data generating process (DGP) is the model (1) when $\|\sqrt{n}\hat{\boldsymbol{\rho}}_m\|^2$ is larger than a certain critical value. More precisely, (5) shows that, under the null hypothesis $H_0$ of model adequacy:

$$Q_m^{BPL} \xrightarrow{\mathcal{L}} \sum_{i=1}^{m} \lambda_i \mathrm{Z}_i^2 \quad \text{as } n \to \infty, \tag{6}$$

where $\mathrm{Z}_1, \ldots, \mathrm{Z}_m$ correspond to independent $\mathcal{N}(0,1)$ random variables and $\lambda_1, \ldots, \lambda_m$ represent the eigenvalues of $\boldsymbol{\Sigma}_{\hat{\boldsymbol{\rho}}_m}$. For an ARMA$(p, q)$ model with iid errors, it is shown in McLeod (1978) that the $p+q$ smallest eigenvalues $\lambda_i$ are close to zero and that the other eigenvalues are equal to one. Thus, we obtain a $\chi^2_{m-s}$ approximation, where $s = p+q$ is the number of estimated parameters, for the asymptotic distribution of $Q_m^{BPL}$ when the DGP is an ARMA $(p, q)$ model with iid errors.

When the errors are uncorrelated but not independent, and when the ARMA coefficients are estimated by least squares, it is shown in Francq, Roy and Zakoïan (2005) that the asymptotic distribution of $Q_m^{BPL}$ is poorly approximated by the chi-square distribution $\chi^2_{m-s}$.

In this paper, the framework is different from the one considered in Francq, Roy and Zakoïan (2005): here, a more general model is permitted than the classical ARMA model. However, in the present set-up, the error process $\{\eta_t\}$ is assumed to be iid and the error term in (1) represents a martingale difference sequence.

It is clear that all the eigenvalues of the matrix $\boldsymbol{\Sigma}_{\hat{\boldsymbol{\rho}}_m}$ are positive and that, when $\mu_3 = 0$ and $m > s$, at least $m - s$ of its eigenvalues are equal to one. When $\mu_3 = 0$ and $m > s$, we then have $\lim_{n\to\infty} P(Q_m^{BPL} > x) \geq P(\chi^2_{m-s} > x)$. Consequently, the test statistic defined by the critical region $\{Q_m^{BPL} > \chi^2_{m-s}(1-\alpha)\}$ is expected to be liberal at the nominal level $\alpha$. In the sequel, this test statistic will be referred to as the $\chi^2_{m-s}$-based ($\mathrm{BPL}_{\chi^2_{m-s}}$) Box-Pierce-Ljung portmanteau test statistic.

It is possible to evaluate the distribution of the Gaussian quadratic form in (6) by means of Imhof's algorithm. Following Francq, Roy and Zakoïan

(2005), one can thus propose a modified portmanteau test statistic based on the following steps: 1) compute the eigenvalues $\hat{\lambda}_1, \ldots, \hat{\lambda}_m$ of a consistent estimator $\hat{\boldsymbol{\Sigma}}_{\hat{\boldsymbol{\rho}}_m}$ of $\boldsymbol{\Sigma}_{\hat{\boldsymbol{\rho}}_m}$, 2) evaluate the $(1-\alpha)$-quantile $c_\alpha(\hat{\lambda}_1, \ldots, \hat{\lambda}_m)$ of $\sum_{i=1}^{m} \hat{\lambda}_i Z_i^2$ using Imhof's algorithm, 3) reject the null that the DGP is (1) when $n\hat{\boldsymbol{\rho}}_m' \hat{\boldsymbol{\rho}}_m \geq c_\alpha(\hat{\lambda}_1, \ldots, \hat{\lambda}_m)$. For further reference, this test statistic will be referred to as the Imhof-based ($\mathrm{BPL}_{Imhof}$) Box-Pierce-Ljung portmanteau test statistic. Compared to the $\mathrm{BPL}_{\chi^2_{m-s}}$ method, the $\mathrm{BPL}_{Imhof}$ version is asymptotically more accurate (since the $\chi^2_{m-s}$ distribution is only a crude approximation of the true asymptotic distribution), but the $\mathrm{BPL}_{Imhof}$ test statistic is relatively more involved to implement since an estimator of $\boldsymbol{\Sigma}_{\hat{\boldsymbol{\rho}}_m}$ is required, and Imhof's algorithm must be implemented, which is relatively complicated and may be time consuming. Section 4 below proposes alternatives to $\mathrm{BPL}_{\chi^2_{m-s}}$ and $\mathrm{BPL}_{Imhof}$ portmanteau test statistics.

### 3.2 An example

The autoregressive (AR) and the autoregressive conditional heteroscedastic (ARCH) models are among the most widely used models for the conditional mean and conditional variance. We combine the simplest versions of these two models to obtain the AR(1)-ARCH(1) model:

$$\begin{cases} Y_t = a_0 Y_{t-1} + \epsilon_t, \\ \epsilon_t = \sigma_t \eta_t, \quad \sigma_t^2 = \omega_0 + \alpha_0 \epsilon_{t-1}^2. \end{cases} \tag{7}$$

Under very general assumptions, Assumption **A** holds true (see Francq and Zakoïan (2004), who discuss Assumption **A** in the framework of ARMA-GARCH models). The unknown parameter is $\boldsymbol{\theta}_0 = (a_0, \omega_0, \alpha_0)$. In order to be able to compute explicitly the information matrices $\boldsymbol{I}$ and $\boldsymbol{J}$, we assume $\alpha_0 = 0$. We then have:

$$\frac{\partial \ell_t(\boldsymbol{\theta}_0)}{\partial \boldsymbol{\theta}} = -2\frac{\eta_t}{\sqrt{\omega_0}} \begin{pmatrix} Y_{t-1} \\ 0 \\ 0 \end{pmatrix} + (1-\eta_t^2)\frac{1}{\omega_0} \begin{pmatrix} 0 \\ 1 \\ \epsilon_{t-1}^2 \end{pmatrix}$$

and

$$\frac{\partial^2 \ell_t(\boldsymbol{\theta}_0)}{\partial \boldsymbol{\theta} \partial \boldsymbol{\theta}'} = \begin{pmatrix} \frac{2Y_{t-1}^2}{\omega_0} & \frac{2\eta_t Y_{t-1}}{\omega_0^{3/2}} & \frac{2\eta_t Y_{t-1}\epsilon_{t-1}^2}{\omega_0^{3/2}} + \frac{2(\eta_t^2-1)\epsilon_{t-1}Y_{t-2}}{\omega_0} \\ \cdot & \frac{2\eta_t^2-1}{\omega_0^2} & \frac{(2\eta_t^2-1)\epsilon_{t-1}^2}{\omega_0^2} \\ \cdot & \cdot & \frac{(2\eta_t^2-1)\epsilon_{t-1}^4}{\omega_0^2} \end{pmatrix}.$$

Thus we have:

$$\boldsymbol{I} = \mathrm{Var}\, \frac{\partial \ell_t(\boldsymbol{\theta}_0)}{\partial \boldsymbol{\theta}} = \begin{pmatrix} \frac{4}{1-a_0^2} & 0 & 2\mu_3^2 \\ 0 & \frac{\mu_4-1}{\omega_0^2} & \frac{\mu_4-1}{\omega_0} \\ 2\mu_3^2 & \frac{\mu_4-1}{\omega_0} & \mu_4(\mu_4-1) \end{pmatrix},$$

$$\boldsymbol{J} = E\frac{\partial^2 \ell_t(\boldsymbol{\theta}_0)}{\partial \boldsymbol{\theta} \partial \boldsymbol{\theta}'} = \begin{pmatrix} \frac{2}{1-a_0^2} & 0 & 0 \\ 0 & \frac{1}{\omega_0^2} & \frac{1}{\omega_0} \\ 0 & \frac{1}{\omega_0} & \mu_4 \end{pmatrix}$$

and

$$\boldsymbol{\Sigma}_{\hat{\boldsymbol{\theta}}_n} = \begin{pmatrix} 1-a_0^2 & -\frac{\omega_0\mu_3^2(1-a_0^2)}{\mu_4-1} & \frac{\mu_3^2(1-a_0^2)}{\mu_4-1} \\ -\frac{\omega_0\mu_3^2(1-a_0^2)}{\mu_4-1} & \omega_0^2\mu_4 & -\omega_0 \\ \frac{\mu_3^2(1-a_0^2)}{\mu_4-1} & -\omega_0 & 1 \end{pmatrix}.$$

Note that, when $\eta_t \sim \mathcal{N}(0,1)$, we have $\boldsymbol{I} = 2\boldsymbol{J}$ and $\boldsymbol{\Sigma}_{\hat{\boldsymbol{\theta}}_n} = 2\boldsymbol{J}^{-1}$. We also have

$$\boldsymbol{C}_m = -\begin{pmatrix} 1 & a_0 & \dots & a_0^{m-1} \\ & 0_{2\times m} & & \end{pmatrix}, \quad \boldsymbol{D}_m = \begin{pmatrix} & 0_{2\times m} \\ \mu_3 & 0'_{m-1} \end{pmatrix}.$$

Note that $\boldsymbol{D}'_m \boldsymbol{J}^{-1}\boldsymbol{C}_m = \boldsymbol{0}$, $\boldsymbol{J}^{-1}\boldsymbol{C}_m = \boldsymbol{C}_m(1-a_0^2)/2$ and $\boldsymbol{I}\boldsymbol{C}_m = \boldsymbol{C}_m 4/(1-a_0^2) + 2\mu_3^3\boldsymbol{C}_m^*$, where $\boldsymbol{C}_m^*$ is obtained by permuting the rows 1 and 3 of $\boldsymbol{C}_m$, so that $\boldsymbol{C}'_m\boldsymbol{C}_m^* = 0$. It follows that

$$\boldsymbol{\Sigma}_{\hat{\boldsymbol{\gamma}}_m} = I_m - (1-a_0^2)\boldsymbol{C}'_m\boldsymbol{C}_m.$$

When $m$ is large or $a_0$ is close to 0, $\boldsymbol{\Sigma}_{\hat{\boldsymbol{\gamma}}_m} \simeq I_m - \boldsymbol{C}'_m\left(\boldsymbol{C}_m\boldsymbol{C}'_m\right)^{-1}\boldsymbol{C}_m$ is close to a projection matrix with $m-1$ eigenvalues equal to 1, and one eigenvalue equals to 0. Therefore, in this particular situation where $\alpha_0 = 0$, the asymptotic distribution of the Box-Pierce-Ljung test statistics can be approximated by a $\chi^2_{m-1}$ distribution. Note that this is not the approximation usually employed in the ARMA case, namely the $\chi^2_{m-s}$ where $s$ is the number of estimated parameters.

### 3.3 Test statistic based on a proposal of Li (1992)

Assume $\boldsymbol{\Sigma}_{\hat{\boldsymbol{\rho}}_m}$ to be non-singular. A natural approach considered by Li (1992, 2004) in non-linear time series with independent errors consists to define the following test statistic:

$$Q_m^{INV} = n\hat{\boldsymbol{\rho}}'_m\hat{\boldsymbol{\Sigma}}^{-1}_{\hat{\boldsymbol{\rho}}_m}\hat{\boldsymbol{\rho}}_m, \tag{8}$$

which follows asymptotically a $\chi^2_m$ distribution under the null hypothesis $H_0$ that the DGP satisfies (1), provided $\hat{\boldsymbol{\Sigma}}_{\hat{\boldsymbol{\rho}}_m}$ corresponds to a consistent estimator of the nonsingular matrix $\boldsymbol{\Sigma}_{\hat{\boldsymbol{\rho}}_m}$.

However, as suggested by the example of the preceding section, the matrix $\boldsymbol{\Sigma}_{\hat{\boldsymbol{\rho}}_m}$ is not invertible in the ARMA case and conditions which guaranty the invertibility of that asymptotic matrix seem difficult to find. If $\boldsymbol{\Sigma}_{\hat{\boldsymbol{\rho}}_m}$ is singular, this invalidates the asymptotic $\chi^2_m$ distribution. In practice, numerical instability is expected in the computation of $Q_m^{INV}$ when $\boldsymbol{\Sigma}_{\hat{\boldsymbol{\rho}}_m}$ is singular.

In the next section, we investigate the use of several generalized inverses of the matrix $\hat{\boldsymbol{\Sigma}}_{\hat{\boldsymbol{\rho}}_m}$. Basic results on generalized inverses are reviewed in the next section.

# 4 Modified portmanteau tests using generalized inverses and {2}-inverses

## 4.1 Generalized inverses and {2}-inverses

A generalized inverse ($g$-inverse) of a matrix $\boldsymbol{\Sigma}$ is a matrix $\tilde{\boldsymbol{\Sigma}}$ satisfying $\boldsymbol{\Sigma}\tilde{\boldsymbol{\Sigma}}\boldsymbol{\Sigma} = \boldsymbol{\Sigma}$. Usually this condition is the first of the four conditions defining the (unique) Moore-Penrose inverse of $\boldsymbol{\Sigma}$, and $\tilde{\boldsymbol{\Sigma}}$ is called a $\{1\}$-inverse (Getson and Hsuan (1988)). On the other hand, a $\{2\}$-inverse of $\boldsymbol{\Sigma}$ is any matrix $\boldsymbol{\Sigma}^*$ satisfying the second relation defining the Moore-Penrose inverse of $\boldsymbol{\Sigma}$, that is $\boldsymbol{\Sigma}^*\boldsymbol{\Sigma}\boldsymbol{\Sigma}^* = \boldsymbol{\Sigma}^*$. When both requirements are satisfied, the resulting matrix is sometimes called a reflexive $g$-inverse or a $\{1,2\}$-inverse (Rao (1973, p. 25)). Let $\boldsymbol{\Sigma} \neq \mathbf{0}$ be a positive semidefinite symmetric matrix of order $m$, with eigenvalues $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_m \geq 0$. The spectral decomposition of $\boldsymbol{\Sigma}$ is $\boldsymbol{\Sigma} = \boldsymbol{P}\boldsymbol{\Lambda}\boldsymbol{P}' = \sum_{i=1}^m \lambda_i \boldsymbol{v}_i \boldsymbol{v}_i'$, where $\boldsymbol{\Lambda} = \mathrm{diag}(\lambda_1, \ldots, \lambda_m)$ and the columns $\boldsymbol{v}_1, \ldots, \boldsymbol{v}_m$ of the matrix $\boldsymbol{P}$ constitute an orthonormal basis of $\mathbb{R}^m$. If $\lambda_{m-s} > 0$ and $\lambda_{m-s+1} = \cdots = \lambda_m = 0$, then the matrix $\boldsymbol{\Sigma}^- = \boldsymbol{P}\boldsymbol{\Lambda}^-\boldsymbol{P}'$ where $\boldsymbol{\Lambda}^- = \mathrm{diag}(\lambda_1^{-1}, \ldots, \lambda_{m-s}^{-1}, \mathbf{0}_s')$ is the Moore-Penrose inverse (or pseudo-inverse) of $\boldsymbol{\Sigma}$. For $k = 1, \ldots, m-s$, let the matrix $\boldsymbol{\Sigma}^{-_k} = \boldsymbol{P}\boldsymbol{\Lambda}^{-_k}\boldsymbol{P}'$ where $\boldsymbol{\Lambda}^{-_k} = \mathrm{diag}(\lambda_1^{-1}, \ldots, \lambda_k^{-1}, \mathbf{0}_{m-k}')$. The matrix $\boldsymbol{\Lambda}^{-_k}$ is always a $\{2\}$-inverse, but this is not a $g$-inverse of $\boldsymbol{\Sigma}$ when $k < m-s$. Now suppose that $\boldsymbol{Z} \sim \mathcal{N}(\mathbf{0}_m, \boldsymbol{\Sigma})$. Then, using natural notations, we have $\boldsymbol{\Lambda}^{-_k 1/2}\boldsymbol{P}'\boldsymbol{Z} \sim \mathcal{N}\left\{\mathbf{0}_m, \mathrm{diag}(\mathbf{1}_k', \mathbf{0}_{m-k}')\right\}$ and thus $\boldsymbol{Z}'\boldsymbol{\Sigma}^{-_k}\boldsymbol{Z} = \|\boldsymbol{\Lambda}^{-_k 1/2}\boldsymbol{P}'\boldsymbol{Z}\|^2 \sim \chi_k^2$. Now suppose that $\boldsymbol{Z}_n \stackrel{\mathcal{L}}{\to} \mathcal{N}(\mathbf{0}_m, \boldsymbol{\Sigma})$ and $\boldsymbol{\Sigma}_n \to \boldsymbol{\Sigma}$ in probability, as $n \to \infty$. For $k = 1, \ldots, \mathrm{rank}(\boldsymbol{\Sigma})$, the matrix $\boldsymbol{\Sigma}^{-_k}$ exists and can be approximated by $\boldsymbol{\Sigma}_n^{-_k}$, for all large enough $n$ (Note, however, that the matrices $\boldsymbol{\Sigma}^{-_k}$ and $\boldsymbol{\Sigma}_n^{-_k}$ are not unique, because they depend on the particular choice of the orthonormal basis in the decompositions $\boldsymbol{\Sigma} = \boldsymbol{P}\boldsymbol{\Lambda}\boldsymbol{P}'$ and $\boldsymbol{\Sigma}_n = \boldsymbol{P}_n\boldsymbol{\Lambda}_n\boldsymbol{P}_n'$). Using the continuity property of the eigenvalues and eigenprojections (see Tyler (1981)), it can be shown that:

$$\boldsymbol{Z}_n'\boldsymbol{\Sigma}_n^{-_k}\boldsymbol{Z}_n \stackrel{\mathcal{L}}{\to} \chi_k^2, \qquad \forall k \leq \mathrm{rank}(\boldsymbol{\Sigma}). \tag{9}$$

The condition $k \leq \mathrm{rank}(\boldsymbol{\Sigma})$ appears to be essential: for example, take:

$$\boldsymbol{\Sigma}_n = \boldsymbol{P}_n\boldsymbol{\Lambda}_n\boldsymbol{P}_n', \qquad \boldsymbol{\Lambda}_n = \begin{pmatrix} 1 & 0 \\ 0 & \frac{1}{n} \end{pmatrix}, \ \boldsymbol{P}_n = \begin{pmatrix} \sqrt{\frac{1}{n}} & -\sqrt{\frac{n-1}{n}} \\ \sqrt{\frac{n-1}{n}} & \sqrt{\frac{1}{n}} \end{pmatrix},$$

and

$$\boldsymbol{Z}_n \sim \mathcal{N}\left\{\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} \frac{c}{n} & 0 \\ 0 & 1 \end{pmatrix}\right\}, \quad c \geq 0.$$

Then $\boldsymbol{Z}_n \stackrel{\mathcal{L}}{\to} \mathcal{N}(\mathbf{0}_2, \boldsymbol{\Sigma})$ and $\boldsymbol{\Sigma}_n \to \boldsymbol{\Sigma}$ with $\boldsymbol{\Sigma} = \mathrm{diag}(0,1)$, but $\boldsymbol{Z}_n'\boldsymbol{\Sigma}_n^{-_2}\boldsymbol{Z}_n \sim (c + \frac{c}{n^2} - \frac{c}{n})Z_1^2 + (2 - \frac{1}{n})Z_2^2 \stackrel{\mathcal{L}}{\not\to} \chi_2^2$. In the next subsection, test statistics will be constructed, relying on an appropriate estimator of the Moore-Penrose inverse, and on estimators of the $\{2\}$-inverses considered in this section.

### 4.2 Generalized portmanteau test statistics

Consider a consistent estimator $\hat{\boldsymbol{\Sigma}}_{\hat{\boldsymbol{\rho}}_m}$ of the matrix $\boldsymbol{\Sigma}_{\hat{\boldsymbol{\rho}}_m}$. Since $\boldsymbol{\Sigma}_{\hat{\boldsymbol{\rho}}_m}$ in (5) may be singular, one can propose:

$$Q_m^{MP} = n\hat{\boldsymbol{\rho}}_m' \hat{\boldsymbol{\Sigma}}_{\hat{\boldsymbol{\rho}}_m}^{-} \hat{\boldsymbol{\rho}}_m, \tag{10}$$

where $\hat{\boldsymbol{\Sigma}}_{\hat{\boldsymbol{\rho}}_m}^{-}$ is an estimator of the Moore-Penrose inverse of $\boldsymbol{\Sigma}_{\hat{\boldsymbol{\rho}}_m}$. At the nominal level $\alpha$, the null hypothesis $H_0$ that the DGP follows a nonlinear model of the form (1) is rejected when $Q_m^{MP} > \chi^2_{k_n}(1-\alpha)$, with $k_n$ the number of eigenvalues of $\hat{\boldsymbol{\Sigma}}_{\hat{\boldsymbol{\rho}}_m}$ larger that a certain tolerance $\epsilon$ (e.g, $\epsilon =$ `sqrt(.Machine$double.eps)` $= 1.49 \times 10^{-8}$ with the `R` software). In view of (5) and (9), test statistics relying on estimators $\hat{\boldsymbol{\Sigma}}_{\hat{\boldsymbol{\rho}}_m}^{-k}$ of the $\{2\}$-inverses $\boldsymbol{\Sigma}_{\hat{\boldsymbol{\rho}}_m}^{-k}$ introduced in Section 4.1 can be proposed. For $k \in \{1, \ldots, m\}$ fixed, they are defined by:

$$Q_m^{-k} = n\hat{\boldsymbol{\rho}}_m' \hat{\boldsymbol{\Sigma}}_{\hat{\boldsymbol{\rho}}_m}^{-k} \hat{\boldsymbol{\rho}}_m. \tag{11}$$

At the nominal level $\alpha$, the null hypothesis $H_0$ is rejected when $Q_m^{-k} > \chi^2_k(1-\alpha)$. The test statistics (10) and (11) constitute interesting alternatives to $\mathrm{BPL}_{Imhof}$; they do not require the use of Imhof's algorithm. However, the range of application of the test statistics relying on $\{2\}$-inverses is more limited because the test statistic $Q_m^{-k}$ presumes the assumption that $\mathrm{rank}(\boldsymbol{\Sigma}_{\hat{\boldsymbol{\rho}}_m}) \geq k$. The range of application of the $Q_m^{-k}$-test decreases as $k$ increases from one to $m$. The test obtained with $k = m$, which is actually that based on $Q_m^{-m} = Q_m^{INV}$, that is the test statistic proposed by Li (1992), is the most restrictive one, in the sense that the invertibility of $\boldsymbol{\Sigma}_{\hat{\boldsymbol{\rho}}_m}$ is required. On the other hand, the set of the alternatives for which the $Q_m^{-k}$-test is consistent should increase with $k$: under appropriate regularity conditions $n^{-1}Q_m^{-1} \to \lambda_1^{-1}\left(\boldsymbol{\rho}_m'\boldsymbol{v}_1\right)^2$ with probability one as $n \to \infty$. Thus, the $Q_m^{-1}$-test should not have much power against alternatives such that $\boldsymbol{\rho}_m'\boldsymbol{v}_1 = 0$. The next section provides an empirical comparison of the different test statistics.

## 5 Numerical illustrations

Here, we compare empirically the following portmanteau tests: $\mathrm{BPL}_{Imhof}$ and the liberal test statistic $\mathrm{BPL}_{\chi^2_{m-s}}$ described in Section 3.1, $\mathrm{BPL}_{\chi^2_{m-1}}$ advocated in Section 3.2, and the test statistics $Q_m^{MP}$ (with $\epsilon = 1.49 \times 10^{-8}$) and $Q_m^{-k}$ introduced in Section 4.2; $Q_m^{INV} = Q_m^{-m}$ of Section 3.3 is included in our experiments. We concentrate on the case $m = 4$, which leads to comparing eight tests. In a first set of Monte Carlo experiments, $N = 1000$ independent trajectories of the AR(1)-ARCH(1) model (7) are simulated. The lengths of the trajectories are $n = 200, 2000$. The code is written in `R` and `FORTRAN`.

Table 1 displays the empirical sizes. For the nominal level $\alpha = 5\%$, the empirical size over the $N = 1000$ independent replications should belong to

**Table 1.** Empirical size of portmanteau tests: relative frequencies (in %) of rejection of the AR(1)-ARCH(1) model (7), when the DGP follows the same model. The number of replications is $N = 1000$.

| Model | $n$ | $\alpha$ | $\mathrm{BPL}_{\chi^2_{m-1}}$ | $\mathrm{BPL}_{\chi^2_{m-s}}$ | $Q_4^{-1}$ | $Q_4^{-2}$ | $Q_4^{-3}$ | $Q_4^{-4}$ | $Q_4^{MP}$ | $\mathrm{BPL}_{Imhof}$ |
|---|---|---|---|---|---|---|---|---|---|---|
| | | 1% | 1.0 | <u>**7.3**</u> | 0.5 | 0.6 | 1.0 | 1.0 | 1.2 | 1.0 |
| I | 200 | 5% | 4.2 | <u>**24.9**</u> | 4.1 | <u>**2.7**</u> | 4.0 | <u>**2.9**</u> | 3.7 | 3.7 |
| | | 10% | 8.3 | <u>**40.1**</u> | 8.3 | 8.2 | **8.0** | <u>**4.8**</u> | <u>**6.8**</u> | **7.9** |
| | | 1% | 1.0 | <u>**9.1**</u> | 1.2 | 1.0 | 1.0 | 1.3 | 1.2 | 1.0 |
| I | 2000 | 5% | 5.9 | <u>**26.9**</u> | 5.3 | 4.8 | 5.9 | 3.6 | 6.3 | 5.9 |
| | | 10% | 10.9 | <u>**44.1**</u> | 9.1 | 9.6 | 10.9 | <u>**7.2**</u> | 9.9 | 10.7 |
| | | 1% | 1.2 | <u>**9.6**</u> | 0.6 | 0.6 | 0.5 | 0.9 | 0.9 | 0.8 |
| II | 200 | 5% | 5.1 | <u>**31.3**</u> | 4.3 | **3.4** | <u>**3.1**</u> | **3.3** | **3.3** | **3.3** |
| | | 10% | 11.7 | <u>**48.7**</u> | 8.7 | **7.8** | <u>**6.9**</u> | <u>**6.8**</u> | <u>**6.8**</u> | <u>**6.2**</u> |
| | | 1% | 1.6 | <u>**16.0**</u> | 0.8 | 1.2 | 0.9 | 0.8 | 0.8 | 0.8 |
| II | 2000 | 5% | <u>**9.8**</u> | <u>**38.9**</u> | 5.2 | 5.0 | 5.1 | 5.0 | 5.0 | 4.5 |
| | | 10% | <u>**18.0**</u> | <u>**58.9**</u> | 10.7 | 10.5 | 10.4 | **12.1** | **12.1** | **12.0** |

I: Model (7) with $a_0 = 0$, $\omega_0 = 1$ and $\alpha_0 = 0$
II: Model (7) with $a_0 = 0.95$, $\omega_0 = 1$ and $\alpha_0 = 0.55$

**Table 2.** Empirical power of portmanteau tests: relative frequencies (in %) of rejection of the AR(1)-ARCH(1) model (7), when the DGP follows an AR(3)-ARCH(1) model (model III) or an AR(1)-ARCH(3) model (model IV).

| Model | $n$ | $\alpha$ | $Q_4^{-1}$ | $Q_4^{-2}$ | $Q_4^{-3}$ | $Q_4^{-4}$ | $Q_4^{MP}$ | $\mathrm{BPL}_{Imhof}$ |
|---|---|---|---|---|---|---|---|---|
| | | 1% | 20.0 | 31.2 | 38.4 | 29.2 | 32.2 | 38.3 |
| III | 200 | 5% | 36.1 | 53.2 | 65.3 | 52.6 | 57.0 | 65.8 |
| | | 10% | 46.3 | 64.2 | 75.4 | 63.4 | 67.7 | 75.3 |
| | | 1% | 35.6 | 61.0 | 80.8 | 71.4 | 73.9 | 81.2 |
| III | 400 | 5% | 49.5 | 73.4 | 92.0 | 84.6 | 86.0 | 91.9 |
| | | 10% | 58.7 | 79.3 | 95.4 | 88.7 | 89.2 | 95.2 |
| | | 1% | 1.5 | 3.5 | 4.7 | 3.4 | 4.3 | 4.4 |
| IV | 200 | 5% | 6.1 | 8.8 | 10.0 | 8.4 | 9.5 | 8.8 |
| | | 10% | 11.1 | 14.1 | 16.5 | 11.3 | 14.9 | 14.9 |
| | | 1% | 2.7 | 5.8 | 7.6 | 6.4 | 7.7 | 7.4 |
| IV | 400 | 5% | 8.3 | 13.5 | 15.3 | 11.3 | 13.8 | 14.5 |
| | | 10% | 13.9 | 20.9 | 22.7 | 16.2 | 19.9 | 21.4 |

III: $Y_t = 0.2Y_{t-3} + \epsilon_t$ where $\epsilon_t^2 = \sqrt{1 + 0.2\epsilon_{t-1}^2}\eta_t$
IV: $Y_t = 0.2Y_{t-1} + \epsilon_t$ where $\epsilon_t^2 = \sqrt{1 + 0.5\epsilon_{t-3}^2}\eta_t$

the interval [3.6%, 6.4%] with probability 95%. When the relative rejection frequencies are outside the 95% significant limits, they are displayed in bold in Table 1. When the relative rejection frequencies are outside the 99% significant limits [3.2%, 6.9%], they are underlined. It can be seen that: 1) the rejection frequency of $\mathrm{BPL}_{\chi^2_{m-s}}$ is definitely too high, 2) as expected, $\mathrm{BPL}_{\chi^2_{m-1}}$ works well when $a_0 = 0$ and $\alpha_0 = 0$, but not when $a_0 \neq 0$ or $\alpha_0 \neq 0$, 3) the empirical levels of $Q_4^{-4} = Q_4^{INV}$ are far from the nominal levels for Model I, which is explained by the singularity of $\boldsymbol{\Sigma}_{\boldsymbol{\rho}_m}$, 4) the errors of the first kind of the test statistics $Q_4^{MP}$, $Q_4^{-k}$, $k < 4$, and $\mathrm{BPL}_{Imhof}$ are well controlled when $n$ is large. Table 2 compares the empirical powers, excluding $\mathrm{BPL}_{\chi^2_{m-s}}$ and $\mathrm{BPL}_{\chi^2_{m-1}}$, which display unsatisfactory empirical levels. Note that misspecification of the conditional mean (model III) seems easier to detect than misspecification of the conditional variance (model IV). As expected, the power of $Q_m^{-k}$ is function of $k$. From Table 2, $Q_4^{-3}$ and $\mathrm{BPL}_{Imhof}$ are the most powerful portmanteau test statistics, at least in our experiments. Interestingly, $Q_4^{MP}$ offers an empirical power very close to the one of $\mathrm{BPL}_{Imhof}$, and slightly better than the one of $Q_4^{-4}$. In general, $\mathrm{BPL}_{Imhof}$ seems to be advisable in view of its good theoretical and finite sample performance, but given its computational simplicity, $Q_4^{MP}$ appears to be a close competitor.

## References

FRANCQ, C., ROY, R. and ZAKOÏAN, J.-M. (2005): Diagnostic checking in ARMA models with uncorrelated errors. *Journal of the American Statistical Association 100, 532-544.*

FRANCQ, C. and ZAKOÏAN, J.-M. (2004): Maximum likelihood estimation of pure GARCH and ARMA-GARCH processes. *Bernoulli 10, 605-637.*

GETSON, A. J. and HSUAN, F. C. (1988): *{ 2 }-Inverses and Their Statistical Application.* Lecture Notes in Statistics 47, Springer-Verlag, New York.

GRANGER, C. W. J. and TERÄSVIRTA, T. (1993): *Modelling Nonlinear Economic Relationships.* Oxford University Press, Oxford.

IMHOF, J. P. (1961): Computing the distribution of quadratic forms in normal variables. *Biometrika 48, 419-426.*

KLIMKO, L. A. and NELSON, P. I. (1978): On conditional least squares estimation for stochastic processes. *The Annals of Statistics 6, 629-642.*

LI, W. K. (2004): *Diagnostic Checks in Time Series.* Chapman & Hall/CRC, New York.

LI, W. K. (1992): On the asymptotic standard errors of residual autocorrelations in nonlinear time series modelling. *Biometrika 79, 435-437.*

MCLEOD, A. I. (1978): On the distribution of residual autocorrelations in Box-Jenkins method. *Journal of the Royal Statistical Society B 40, 296-302.*

POTSCHER, B. M. and PRUCHA, I. R. (1997): *Dynamic Nonlinear Econometric Models.* Springer, Berlin.

RAO, C. R. (1973): *Linear Statistical Inference and Its Applications.* Wiley, New York.

TANIGUCHI, M. and KAKIZAWA, Y. (2000): *Asymptotic Theory of Statistical Inference for Time Series.* Springer, New York.

TONG, H. (1990): *Non-linear Time Series: A Dynamical System Approach.* Oxford University Press, Oxford.

TYLER, D. E. (1981): Asymptotic inference for eigenvectors. *The Annals of Statistics 9, 725-736.*