Bootstrapping the portmanteau tests in weak auto-regressive moving average models

Zhu, Ke

Chinese Academy of Sciences

6 February 2015

BOOTSTRAPPING THE PORTMANTEAU TESTS IN WEAK AUTO-REGRESSIVE MOVING AVERAGE MODELS

By Ke Zhu

Chinese Academy of Sciences

This paper uses a random weighting (RW) method to bootstrap the critical values for the Ljung-Box/Monti portmanteau tests and weighted Ljung-Box/Monti portmanteau tests in weak ARMA models. Unlike the existing methods, no user-chosen parameter is needed to implement the RW method. As an application, these four tests are used to check the model adequacy in power GARCH models. Simulation evidence indicates that the weighted portmanteau tests have the power advantage over other existing tests. A real example on S&P 500 index illustrates the merits of our testing procedure. As one extension work, the block-wise RW method is also studied.

1. Introduction. After the seminal work of Box and Pierce (1970) and Ljung and Box (1978), the portmanteau test has been popular for diagnostic checking in the following ARMA(p, q) model:

$$y_t = \sum_{i=1}^{p} \phi_i y_{t-i} + \sum_{i=1}^{q} \psi_i \varepsilon_{t-i} + \varepsilon_t, \tag{1.1}$$

where ε_t is the error term with $E\varepsilon_t = 0$. Conventionally, we say that model (1.1) is weak when $\{\varepsilon_t\}$ is an uncorrelated sequence, and that model (1.1) is strong when $\{\varepsilon_t\}$ is an i.i.d. sequence; see, e.g., Francq and Zakoïan (1998) and Francq, Roy, and Zakoïan (2005). In the earlier decades, many portmanteau tests are proposed for strong ARMA models, and among them, the most famous ones are the Box-Pierce statistic Q_m and Ljung-Box statistic $\widetilde{Q}_m$ which are defined respectively by

$$Q_m = n \sum_{k=1}^{m} \hat{\rho}_k^2 \quad \text{and} \quad \widetilde{Q}_m = n(n+2) \sum_{k=1}^{m} \frac{\hat{\rho}_k^2}{n-k}, \tag{1.2}$$

where n is the length of the series $\{y_t\}$, m is a fixed positive integer, and $\hat{\rho}_k$ is the residuals autocorrelation at lag k. When $E\varepsilon_t^2 < \infty$, the standard analysis in Box and Pierce (1970) or McLeod (1978) showed that the limiting distribution of Q_m or $\widetilde{Q}_m$ can be approximated by $\chi^2_{m-(p+q)}$ in strong ARMA models, although $\widetilde{Q}_m$ usually has a better finite sample performance for small n. In view of this, people more or less regard that the limiting distribution of Q_m or $\widetilde{Q}_m$ is independent to the data structure (or asymptotically pivotal in other words), and hence both test statistics can be easily

Keywords and phrases: Bootstrap method; Portmanteau test; Power GARCH models; Random weighting approach; Weak ARMA models; Weighted portmanteau test.

implemented in practice. Recently, Fisher and Gallagher (2012) proposed a weighted Ljung-Box statistic $\overline{Q}_m$ for strong ARMA models based on the trace of the square of the m-th order autocorrelation matrix, where

$$\overline{Q}_m = n(n+2)\sum_{k=1}^{m}\frac{(m-k+1)}{m}\frac{\hat{\rho}_k^2}{n-k}. \tag{1.3}$$

Compared with $\widetilde{Q}_m$, this new weighted Ljung-Box statistic is numerically more stable over m, and its limiting distribution can be approximated by a gamma distribution. For more discussions on portmanteau tests in strong ARMA models, we refer to McLeod and Li (1983), Peña and Rodríguez (2002, 2006), Li (2004), and many others.

Although the aforementioned portmanteau tests have been well applied in industries, their asymptotic properties are valid only for strong ARMA models; see, e.g., Romano and Thombs (1996). Till now, more and more empirical studies have revealed that the independence structure of error terms $\{\varepsilon_t\}$ in model (1.1) may be restrictive especially for economic and financial series. For instance, Bollerslev (1986) used an AR(4)-GARCH$(1,1)$ model to study the GNP series in U.S.; Franses and Van Dijk (1996) studied several stock market indexes by AR(1)-GJR$(1,1)$ models; Zhu and Ling (2011) found that a MA(3)-GARCH$(1,1)$ model is necessary to fit the world oil prices; see also Tsay (2005) for more empirical evidence. Moreover, Francq and Zakoïan (1998) and Francq, Roy, and Zakoïan (2005) indicated that many nonlinear models have an ARMA representation with dependent error terms. Thus, it is meaningful to consider the diagnostic checking for weak ARMA models.

So far, a huge literature has focused on testing the dependence of time series. Next, we briefly review some of important results. For the observable series, Deo (2000) constructed some spectral tests of the martingale difference hypothesis; Lobato (2001) proposed an asymptotically pivotal test to detect whether the observable series is uncorrelated; Escanciano and Velasco (2006) extended the method in Durlauf (1991) to testing the martingale difference hypothesis, taking into account both linear and nonlinear dependence; Escanciano and Lobato (2009) derived a data-driven portmanteau test which is applicable in the conditionally heteroskedastic series but with no need to choose m. For the residual series, Hong and Lee (2005) considered a generalized spectral test which is valid for semi-strong ARMA models; Escanciano (2006) tested the martingale difference hypothesis by introducing a parametric family of functions as in Stinchcombe and White (1998); moreover, based on some marked residual processes, the Cramér-von Mises test and Kolmogrove-Smirnov test were studied in Escanciano (2007) towards the same goal. However, none of aforementioned tests can be applied to the residuals of weak ARMA models.

To solve this problem, Romano and Thombs (1996) and Horowitz, Lobato, Nankervis, and Savin (2006) used the block-wise bootstrap to estimate the asymptotic covariance matrix of autocorrelations for the observable series. Francq, Roy, and Zakoïan (2005) used the VAR method in Berk (1974) to estimate the asymptotic covariance matrix of $\widetilde{Q}_m$ in weak ARMA models. However, both estimation procedures have the disadvantage of requiring the selection of some user-chosen parameters as indicated by Kuan and Lee (2006). Moreover, they used the self-normalization method to propose a robust M test which is asymptotically pivotal and independent to user-chosen parameters. This robust M test is applicable for testing serial correlations in weak AR models. Recently, Shao (2011a) proved the validity of the spectral test in Hong (1996) for detecting serial correlations in weak ARMA models. Although the spectral test is consistent, it still requires the selection of kernel function and its related bandwidth.

In this paper, we bootstrap the critical values for $\widetilde{Q}_m$ and $\overline{Q}_m$ in weak ARMA models by a random weighting (RW) method. Unlike the existing methods, no user-chosen parameter is needed to implement this new method. Particularly, it is applicable when ε_t is a martingale difference (i.e., $E(\varepsilon_t|\mathcal{F}_{t-1}) = 0$ with $\mathcal{F}_t = \sigma(\varepsilon_s; s \leq t)$). Hence, the RW method is valid when ε_t has the often observed ARCH-type structure, i.e.,

$$\varepsilon_t = \eta_t \sqrt{h_t}, \tag{1.4}$$

where η_t is i.i.d. innovation, and $h_t \in \mathcal{F}_{t-1}$ is the conditional variance of ε_t. Meanwhile, we can show that the bootstrapped critical values for $\widetilde{Q}_m$ and $\overline{Q}_m$ are valid for Monti's (1994) portmanteau test $\widetilde{M}_m$ and weighted Monti portmanteau test $\overline{M}_m$ in Fisher and Gallagher (2012), respectively, where both $\widetilde{M}_m$ and $\overline{M}_m$ are based on a vector of residuals partial autocorrelations. As a byproduct, the standard deviation of m-th residuals (partial) autocorrelation can be calculated from our RW approach. Hence, the classical plot of residuals (partial) autocorrelations with the corresponding significance bounds is available for weak ARMA models. As an application, we then use all four portmanteau tests to check the adequacy of power GARCH models. Simulation studies are carried out to assess the finite-sample performance of all tests. A real example on S&P 500 index illustrates the merits of our testing procedure. As one extension work, a block-wise RW approach is also proposed to bootstrap the critical values for all tests, and its validity is justified.

This paper is organized as follows. Section 2 proposes the RW approach to bootstrap the critical values for four portmanteau tests. Section 3 applies these tests to check the model adequacy in power GARCH models. Simulation results are reported in Section 4. A real example is given in Section 5. An extension work on the block-wise RW approach is provided in Section 6. Concluding remarks are offered in Section 7. All of proofs and some additional simulation results are given in the on-line supplementary document.

2. Random weighting approach.

Denote $\theta = (\phi_1, \cdots, \phi_p, \psi_1, \cdots, \psi_q)'$ be the unknown parameter of model (1.1). Let θ_0 be the true value of θ and the parameter space Θ be a compact subset of $\mathcal{R}^s$, where $\mathcal{R} = (-\infty, \infty)$ and $s = p + q$. We make the following two assumptions:

Assumption 2.1. *θ_0 is an interior point in Θ and for each $\theta \in \Theta$, $\phi(z) \triangleq 1 - \sum_{i=1}^{p} \phi_i z^i \neq 0$ and $\psi(z) \triangleq 1 + \sum_{i=1}^{q} \psi_i z^i \neq 0$ when $|z| \leq 1$, and $\phi(z)$ and $\psi(z)$ have no common root with $\phi_p \neq 0$ or $\psi_q \neq 0$.*

Assumption 2.2. *$\{y_t\}$ is strictly stationary with $E|y_t|^{4+2\nu} < \infty$ and*

$$\sum_{k=0}^{\infty} \{\alpha_y(k)\}^{\nu/(2+\nu)} < \infty$$

for some $\nu > 0$, where $\{\alpha_y(k)\}$ is the sequence of strong mixing coefficients of $\{y_t\}$.

Assumption 2.1 is the condition for the stationarity, invertibility and identifiability of model (1.1). Assumption 2.2 from Francq, Roy, and Zakoïan (2005) gives the necessary conditions to prove our asymptotic results, and it is satisfied by many weak ARMA models, e.g.,

$$\begin{aligned}
&(a)\ y_t = a y_{t-1} + \varepsilon_t \ \text{ and } \ \varepsilon_t = \eta_t \eta_{t-1};\\
&(b)\ y_t = a \varepsilon_{t-1} + \varepsilon_t \ \text{ and } \ \varepsilon_t = \eta_t \eta_{t-1};\\
&(c)\ y_t = \eta_t + b y_{t-1} \eta_{t-2}, \text{which admits a weak MA(3) representation:}\\
&\qquad y_t = \varepsilon_t + c\varepsilon_{t-3} \text{ as shown in Francq and Zakoïan (1998),}
\end{aligned}$$

where $\eta_t \sim N(0,1)$ is i.i.d. white noise, $|a| < 1$, $|b| < 1$, and $\alpha_y(k)$ in each aforementioned model tends to zero exponentially fast.

Next, given the observations $\chi_n \triangleq \{y_1, \cdots, y_n\}$, we consider the least squares estimator (LSE) $\hat{\theta}_n$ of θ_0, defined by

$$\hat{\theta}_n = \arg\min_{\Theta} \tilde{L}_n(\theta), \text{ where } \tilde{L}_n(\theta) = \frac{1}{n} \sum_{t=1}^{n} \tilde{\varepsilon}_t^2(\theta) \triangleq \frac{1}{n} \sum_{t=1}^{n} \tilde{l}_t(\theta),$$

and $\tilde{\varepsilon}_t(\theta)$ is defined recursively by

$$\tilde{\varepsilon}_t(\theta) = y_t - \sum_{i=1}^{p} \phi_i y_{t-i} - \sum_{i=1}^{q} \psi_i \tilde{\varepsilon}_{t-i}(\theta)$$

with $\tilde{\varepsilon}_0(\theta) = \tilde{\varepsilon}_{-1}(\theta) = \cdots = \tilde{\varepsilon}_{-q+1}(\theta) = y_0 = y_{-1} = \cdots = y_{-p+1} = 0$.

Denote $\hat{\varepsilon}_t \triangleq \tilde{\varepsilon}_t(\hat{\theta}_n)$ be the residual of model (1.1). Then, the residuals autocorrelation at lag k is

$$\hat{\rho}_k = \frac{\sum_{t=k+1}^{n} \hat{\varepsilon}_t \hat{\varepsilon}_{t-k}}{\sum_{t=1}^{n} \hat{\varepsilon}_t^2}.$$

Theorem 2.1 below gives the limiting distributions of $\widetilde{Q}_m$ in (1.2) and $\overline{Q}_m$ in (1.3), and under the same conditions, its proof (omitted here) follows directly from the one for Theorem 2 in Francq, Roy, and Zakoïan (2005). In addition, the limiting distribution of $\widetilde{Q}_m$ in Theorem 2.1 is the same as the one in Theorem 2 of Francq, Roy, and Zakoïan (2005), and it reduces to the one in Li (1992, p.436) when model (1.1) is strong.

To state Theorem 2.1 precisely, some notations are needed. Define $\varepsilon_t(\theta)$ recursively by

$$\varepsilon_t(\theta) = y_t - \sum_{i=1}^{p} \phi_i y_{t-i} - \sum_{i=1}^{q} \psi_i \varepsilon_{t-i}(\theta),$$

i.e., $\varepsilon_t(\theta) = \psi^{-1}(B)\phi(B)y_t$, where B is the back shift operator. Then,

$$\frac{\partial \varepsilon_t(\theta)}{\partial \theta} = (\nu_{t-1}(\theta), \cdots, \nu_{t-p}(\theta), u_{t-1}(\theta), \cdots, u_{t-q}(\theta))',$$

where $\nu_t(\theta) = -\phi^{-1}(B)\varepsilon_t(\theta)$ and $u_t(\theta) = \psi^{-1}(B)\varepsilon_t(\theta)$. Particularly, $\varepsilon_t(\theta_0) = \varepsilon_t$ and

$$\frac{\partial \varepsilon_t}{\partial \theta} \triangleq \frac{\partial \varepsilon_t(\theta_0)}{\partial \theta} = \sum_{i=1}^{\infty} \varepsilon_{t-i}\lambda_i,$$

where $\lambda_i = (-\phi^*_{i-1}, \cdots, -\phi^*_{i-p}, \psi^*_{i-1}, \cdots, \psi^*_{i-q})'$, and coefficients ϕ^*_i and ψ^*_i are from expansions: $\phi^{-1}(z) = \sum_{i=0}^{\infty} \phi^*_i z^i$ and $\psi^{-1}(z) = \sum_{i=0}^{\infty} \psi^*_i z^i$ with $\phi^*_i = \psi^*_i = 0$ for $i < 0$.

Moreover, let $\sigma^2 = E\varepsilon_t^2$ and define the following matrixes:

$$\begin{aligned}
\mathcal{I} &= 4 \lim_{n\to\infty} \frac{1}{n} E\left(\sum_{t=1}^{n}\sum_{t'=1}^{n} \varepsilon_t \varepsilon_{t'} \frac{\partial \varepsilon_t}{\partial \theta}\frac{\partial \varepsilon_{t'}}{\partial \theta'}\right) = 4\sum_{i=1}^{\infty}\sum_{j=1}^{\infty} \Gamma(i,j)\lambda_i\lambda_j',\\
\mathcal{J} &= 2 \lim_{n\to\infty} \frac{1}{n} E\left(\sum_{t=1}^{n} \frac{\partial \varepsilon_t}{\partial \theta}\frac{\partial \varepsilon_t}{\partial \theta'}\right) = 2\sigma^2 \sum_{i=1}^{\infty} \lambda_i\lambda_i',\\
\mathcal{R}_m &= (R_{ij})_{1\le i,\, j\le m} \ \text{ with } R_{ij} = \sigma^{-4}\Gamma(i,j),\\
\mathcal{S}_m &= (s_1, \cdots, s_m) \ \text{ with } \ s_k = \sum_{i=1}^{\infty} R_{ik}\lambda_i,
\end{aligned}$$

$\Lambda_m = (\lambda_1, \cdots, \lambda_m)$, and

$$\Sigma = \mathcal{R}_m + \Lambda_m' \mathcal{J}^{-1}\mathcal{I}\mathcal{J}^{-1}\Lambda_m - 2\sigma^2\Lambda_m'\mathcal{J}^{-1}\mathcal{S}_m - 2\sigma^2\mathcal{S}_m'\mathcal{J}^{-1}\Lambda_m, \tag{2.1}$$

where

$$\Gamma(l, l') = \sum_{h=-\infty}^{\infty} E\left(\varepsilon_t \varepsilon_{t-l}\varepsilon_{t+h}\varepsilon_{t+h-l'}\right). \tag{2.2}$$

Here, $\Gamma(l, l')$ is meaningful by Assumptions 2.1-2.2 and Lemma A.1 in Francq, Roy, and Zakoïan (2005), and it satisfies $\Gamma(l, l') = \Gamma(-l, -l') = \Gamma(-l', -l)$. Note that Σ is the same as $\Sigma_{\hat{\rho}_m}$ in Theorem 2 of Francq, Roy, and Zakoïan (2005).

THEOREM 2.1. *Suppose that Assumptions 2.1-2.2 hold and model (1.1) is weak and correctly specified. Then, (i)*

$$\sqrt{n}\hat{\rho} \to_d N(0, \Sigma) \quad \textit{and} \quad W(\sqrt{n}\hat{\rho}) \to_d N(0, W\Sigma W) \quad \textit{as } n \to \infty,$$

where $\to_d$ denotes the convergence in distribution, $\hat{\rho} = (\hat{\rho}_1, \cdots, \hat{\rho}_m)'$, the matrix Σ is defined in (2.1), and

$$W = diag\left\{1, \sqrt{(m-1)/m}, \sqrt{(m-2)/m}, \cdots, \sqrt{1/m}\right\}$$

is a diagonal matrix; (ii) furthermore, it follows that

$$\widetilde{Q}_m \to_d \sum_{i=1}^{m} \tilde{\xi}_{mi} Z_i^2 \quad and \quad \overline{Q}_m \to_d \sum_{i=1}^{m} \bar{\xi}_{mi} Z_i^2 \quad as\ n \to \infty,$$

where $Z_1, \cdots, Z_m$ are independent $N(0,1)$ variables, $\tilde{\xi}_m = (\tilde{\xi}_{m1}, \cdots, \tilde{\xi}_{mm})'$ is the eigenvalues vector of the matrix Σ, and $\bar{\xi}_m = (\bar{\xi}_{m1}, \cdots, \bar{\xi}_{mm})'$ is the eigenvalues vector of the matrix $W\Sigma W$.

REMARK 2.1. *By using a consistent estimate $\hat{\Sigma}$ of Σ, it may be more natural to construct the test statistic $n\hat{\rho}'\hat{\Sigma}^{-1}\hat{\rho}$, which has an asymptotic distribution χ^2_m from Theorem 2.1. This test statistic is calculable if Σ is non singular. However, as shown in Francq, Roy, and Zakoïan (2005), Σ may be (or close to) singular in some cases. To overcome this difficulty, Francq, Roy, and Zakoïan (2005) used the spectral decomposition to express the limiting distribution of $\widetilde{Q}_m$, while Duchesne and Francq (2008) used the generalized inverses and $\{2\}$-inverses of Σ to construct the portmanteau test. In this paper, both methods are valid, and we follow the method in Francq, Roy, and Zakoïan (2005) for simplicity.*

REMARK 2.2. *When y_t follows a strong ARMA model, McLeod (1978) proved that the limiting distribution of $\widetilde{Q}_m$ can be approximated by a $\chi^2_{m-(p+q)}$ distribution, and Fisher and Gallagher (2012) proved that the limiting distribution of $\overline{Q}_m$ can be approximated by a $Gamma(a,b)$ distribution with*

$$a = \frac{3}{4}\frac{m(m+1)^2}{2m^2+3m+1-6m(p+q)} \quad and\ b = \frac{2}{3}\frac{2m^2+3m+1-6m(p+q)}{m(m+1)}. \tag{2.3}$$

However, simulation studies in Section 4 showed that these approximations will lead to an over-rejection problem in weak ARMA models.

From Theorem 2.1, we know that both statistics $\widetilde{Q}_m$ and $\overline{Q}_m$ are applicable when Σ can be consistently estimated. This is usually accomplished by either the kernel-based method in Newey and West (1987) or the VAR method in Francq, Roy, and Zakoïan (2005). However, both methods require picking up some user-chosen parameters, and hence it turns out that the testing results are sensitive to these user-chosen parameters; see, e.g., Kuan and Lee (2006). In this section, we now introduce a random weighting (RW) approach to bootstrap the critical values for $\widetilde{Q}_m$ and $\overline{Q}_m$. This novel method gives an estimator of Σ, but it avoids selecting any user-chosen parameter. Its detailed procedure is as follows:

step 1. Generate a sequence of positive i.i.d. random weights $\{w_1^*, \cdots, w_n^*\}$, independent of the data, from a common distribution with mean and variance both equal to 1, and then calculate $\hat{\theta}_n^*$ via

$$\hat{\theta}_n^* = \arg\min_{\Theta} \tilde{L}_n^*(\theta), \text{ where } \tilde{L}_n^*(\theta) = \frac{1}{n}\sum_{t=1}^{n} w_t^* \tilde{l}_t(\theta);$$

step 2. Calculate $\Delta = W^*[\sqrt{n}(\hat{\rho}^* - \hat{\rho})]$ with $\hat{\rho}^* = (\hat{\rho}_1^*, \cdots, \hat{\rho}_m^*)'$, where

$$\hat{\rho}_k^* = \frac{\sum_{t=k+1}^{n} w_t^* \hat{\varepsilon}_t^* \hat{\varepsilon}_{t-k}^*}{\sum_{t=1}^{n} \hat{\varepsilon}_t^{*2}}$$

with $\hat{\varepsilon}_t^* = \tilde{\varepsilon}_t(\hat{\theta}_n^*)$, and the diagonal matrix

$$W^* = \begin{cases} I_m \text{ or } \widetilde{W}^* & \text{for } \widetilde{Q}_m \\ W \text{ or } \overline{W}^* & \text{for } \overline{Q}_m \end{cases}, \tag{2.4}$$

where $I_m \in \mathcal{R}^{m\times m}$ is the identity matrix,

$$\widetilde{W}^* = diag\left\{\sqrt{\frac{n+2}{n-1}}, \sqrt{\frac{n+2}{n-2}}, \cdots, \sqrt{\frac{n+2}{n-m}}\right\},$$

and

$$\overline{W}^* = diag\left\{\sqrt{\frac{n+2}{n-1}}, \sqrt{\frac{n+2}{n-2}\frac{m-1}{m}}, \cdots, \sqrt{\frac{n+2}{n-m}\frac{1}{m}}\right\};$$

step 3. Repeat steps 1-2 J times to get $\{\Delta_{(1)}, \cdots, \Delta_{(J)}\}$, and then compute its sample variance-covariance matrix Σ^*; furthermore, simulate N i.i.d. random samples $\{z_1^{(j)}, \cdots, z_m^{(j)}\}_{j=1}^N$ from the multivariate normal distribution $N(0, I_m)$, and then calculate the sequence $\{K^{(j)}\}_{j=1}^N$ via

$$K^{(j)} = \sum_{i=1}^{m} \xi_{mi}^* z_i^{(j)},$$

where $(\xi_{m1}^*, \cdots, \xi_{mm}^*)'$ is the eigenvalues vector of the matrix Σ^*;

step 4. Choose α upper percentage of $\{K^{(j)}\}_{j=1}^N$ as the critical value of $\widetilde{Q}_m$ or $\overline{Q}_m$ at any given significance level α.

Particularly, when $p = q = 0$, we set $\hat{\varepsilon}_t = \hat{\varepsilon}_t^* = y_t$ for all t in step 2. The p-value of $\widetilde{Q}_m$ or $\overline{Q}_m$ can be calculated by $\#(K^{(j)} > \widetilde{Q}_m)/N$ or $\#(K^{(j)} > \overline{Q}_m)/N$, respectively. We now offer some remarks on the RW method. First, the RW method as a variant of the traditional wild bootstrap in Wu (1986) was initially proposed by Jin, Ying, and Wei (2001). So far, it has been widely used for statistical inference in regression and time series models, when the asymptotic covariance matrix of certain estimator can not be directly estimated; see, e.g., Chen, Ying, Zhang, and Zhao (2008), Chen, Guo, Lin, and Ying (2010), Li, Leng, and Tsai (2014), Li, Li, and Tsai (2014), and Zhu and Ling (2015). Second, the random weight w_t^* can from the standard exponential distribution as used in Section 4, the Bernoulli distribution such that

$$P\left(w_t^* = \frac{3-\sqrt{5}}{2}\right) = \frac{1+\sqrt{5}}{2\sqrt{5}} \text{ and } P\left(w_t^* = \frac{3+\sqrt{5}}{2}\right) = \frac{\sqrt{5}-1}{2\sqrt{5}},$$

and many others. Although theoretically it is unknown how sensitive is the RW method to the distribution of w_t^*, simulation studies in the on-line supplementary document imply that our testing results are robust to the distribution of w_t^*. Third, although I_m and $\widetilde{W}^*$ (or W and $\overline{W}^*$) are close to each other when n is large, we expect that the finite sample performance of $\widetilde{Q}_m$ (or $\overline{Q}_m$) based on $\widetilde{W}^*$ (or $\overline{W}^*$) may be better for small n.

To justify the validity of the bootstrapped critical value from the RW method, two more assumptions are needed:

ASSUMPTION 2.3. *$E(w_t^*)^{2+\kappa_0} < \infty$ for some $\kappa_0 > 0$.*

ASSUMPTION 2.4. *For $l, l' \geq 1$,*

$$\sum_{h=-\infty, h\neq 0}^{\infty} E\left(\varepsilon_t \varepsilon_{t-l} \varepsilon_{t+h} \varepsilon_{t+h-l'}\right) = 0.$$

Assumption 2.3 gives one mild condition for w_t^*. Assumption 2.4 restricts the dependent structure of ε_t, and it is equivalent to say that the infinite series $\Gamma(l, l')$ in (2.2) collapses to just one summand with the index $h = 0$ (i.e., $E(\varepsilon_t^2 \varepsilon_{t-l} \varepsilon_{t-l'})$). Particularly, when ε_t is a martingale difference including the special case as in (1.4), we know that Assumption 2.4 holds. To capture the dependence structure of ε_t beyond Assumption 2.4, a block-wise technique is necessary, and we will study it in Section 6.

We now are ready to state our main result:

THEOREM 2.2. *Suppose that Assumptions 2.1-2.4 hold and model (1.1) is weak and correctly specified. Then, conditional on χ_n,*

$$\sqrt{n}(\hat{\rho}^* - \hat{\rho}) \to_d N(0, \Sigma) \quad \textit{in probability}$$

as $n \to \infty$, where Σ is defined in (2.1).

REMARK 2.3. *By Theorem 2.2, it follows that conditional on χ_n,*

$$\widetilde{W}^*[\sqrt{n}(\hat{\rho}^* - \hat{\rho})] \to_d N(0, \Sigma) \textit{ and } \overline{W}^*[\sqrt{n}(\hat{\rho}^* - \hat{\rho})] \to_d N(0, W\Sigma W)$$

in probability as $n \to \infty$. In view of this, when J is large, the matrix Σ^* in step 3 of RW method provides a good estimator of Σ (or $W\Sigma W$) for $\widetilde{Q}_m$ (or $\overline{Q}_m$), and hence our bootstrapped critical values from the RW method are asymptotically valid.

In the end, we show that our RW method can be also used for the portmanteau test $\widetilde{M}_m$ in Monti (1994) and the weighted portmanteau test $\overline{M}_m$ in Fisher and Gallagher (2012), where

$$\widetilde{M}_m = n(n+2)\sum_{k=1}^{m}\frac{\hat{\pi}_k^2}{n-k} \quad \text{and} \quad \overline{M}_m = n(n+2)\sum_{k=1}^{m}\frac{(m-k+1)}{m}\frac{\hat{\pi}_k^2}{n-k}$$

with $\hat{\pi}_k$ being the residuals partial autocorrelation at lag k. Here, $\hat{\pi}_k$ is calculated by the Durbin-Levinson algorithm as follows:

$$\hat{\pi}_k = \frac{\hat{\rho}_k - (\hat{\rho}_1, \cdots, \hat{\rho}_{k-1})\hat{R}_{k-1}^{-1}(\hat{\rho}_{k-1}, \cdots, \hat{\rho}_1)'}{1 - (\hat{\rho}_1, \cdots, \hat{\rho}_{k-1})\hat{R}_{k-1}^{-1}(\hat{\rho}_1, \cdots, \hat{\rho}_{k-1})'},$$

where $\hat{R}_k = (\hat{\rho}_{|i-j|})_{i,j=1,2,\cdots,k}$ is the Toeplitz matrix. As for Theorem 2.1, we can easily show that $\sqrt{n}\hat{\pi}_k = \sqrt{n}\hat{\rho}_k + o_p(1)$, and hence the following theorem is straightforward:

THEOREM 2.3. *Suppose that the conditions in Theorem 2.1 hold. Then, (i)*

$$\sqrt{n}\hat{\pi} \to_d N(0, \Sigma) \quad \text{and} \quad W(\sqrt{n}\hat{\pi}) \to_d N(0, W\Sigma W) \quad \text{as } n \to \infty,$$

where $\hat{\pi} = (\hat{\pi}_1, \cdots, \hat{\pi}_m)'$; (ii) furthermore, it follows that $\widetilde{M}_m$ and $\overline{M}_m$ have the same liming distributions as $\widetilde{Q}_m$ and $\overline{Q}_m$, respectively.

REMARK 2.4. *Besides the portmanteau tests aforementioned, the kernel-based spectral test H_n in Hong (1996) has also been widely used for testing series correlations of $\{\varepsilon_t\}$ in model (1.1), where*

$$H_n = \sum_{j=1}^{n-1}[k(j/p_n)]^2\hat{\rho}_j^2$$

with $k(\cdot)$ being the kernel function and $p_n(>0)$ being the bandwidth. When $k(\cdot)$ satisfies Assumption 2.1 in Shao (2011a) and p_n satisfies the condition that $\log n = o(p_n)$ and $p_n = o(n^{1/2})$, Shao (2011a) proved that if model (1.1) is weak and correctly specified,

$$\frac{nH_n - p_nC(k)}{\sqrt{2p_nD(k)}} \to_d N(0,1) \quad \text{as } n \to \infty, \tag{2.5}$$

where $C(k) = \int_0^\infty [k(x)]^2 dx$ and $D(k) = \int_0^\infty [k(x)]^4 dx$. From (2.5), we know that the limiting null distribution of H_n depends on $k(\cdot)$ and p_n. Note that unlike the lag m in the portmanteau test, it is difficult to interpret p_n in H_n. Simulation studies in Section 4 below imply that the size/power performance of H_n is very sensitive to the choice of p_n. Thus, H_n may be of limited value in applications, without a data-driven/computer intensive selection of p_n. When $p = q = 0$ in model (1.1), Lee and Hong (2001) used the cross-validation procedure of Beltrao and Bloomfield (1987) to select p_n; when model (1.1) is strong, Hong and Lee (2003) studied the plug-in data-driven method as in Hong (1999) but still requiring the selection of a preliminary user-chosen bandwidth

$\bar{p}$; when ε_t in model (1.1) is a martingale difference with conditional heteroscedasticity of unknown form, Hong and Lee (2005) re-studied this plug-in method; however, when model (1.1) is weak, how to construct and theoretically study a good data-driven method for H_n is still an open question, and we leave it for future study.

Note that the limiting distribution of $\sqrt{n}\hat{\pi}$ is the same as the one of $\sqrt{n}\hat{\rho}$. Thus, when model (1.1) is strong, the limiting distribution of $\widetilde{M}_m$ (or $\overline{M}_m$) can be approximated by a $\chi^2_{m-(p+q)}$ (or Gamma(a,b)) distribution as in Remark 2.2, but when model (1.1) is weak, the bootstrapped critical values for $\widetilde{Q}_m$ and $\overline{Q}_m$ should be used for $\widetilde{M}_m$ and $\overline{M}_m$, respectively. In addition, it is worth noting that the standard deviation of $\hat{\rho}_i$ (or $\hat{\pi}_i$), denoted by sd_i, can be calculated by

$$\mathrm{sd}_i = \sqrt{\Sigma^*_{ii}/n},$$

where Σ^*_{ii} is the i-th main diagonal entry of Σ^* for $i = 1, 2, \cdots, m$. Consequently, the classical plot of residuals (partial) autocorrelations with the corresponding significance bounds are available in weak ARMA models, where the significance lower and upper bounds for $\hat{\rho}_i$ (or $\hat{\pi}_i$) at level α are $\Phi^{-1}[(1-\alpha)/2]\mathrm{sd}_i$ and $\Phi^{-1}[(1+\alpha)/2]\mathrm{sd}_i$, respectively. Here, $\Phi^{-1}(\cdot)$ is the inverse of the cdf of $N(0,1)$ distribution. As shown in Romano and Thombs (1996), the traditional significance bounds for strong ARMA models are misleading when $\{\varepsilon_t\}$ in model (1.1) are dependent. Our significance bounds generated from the RW method are now valid for weak ARMA models, and hence they are practically important. We will illustrate this by one real example in Section 5.

3. Model checking in Power GARCH models. Since Engle (1982), ARCH-type models have become the most important tools to modeling economic and financial data sets. The power GARCH model as one of ARCH-type models describes that the power of conditional volatility depends on the past information and changes over time. Specifically, the power GARCH(r,s) (denoted by PGARCH(r,s)) model is defined as

$$y_t = \eta_t \sqrt{h_t} \text{ and } h_t^{\frac{\delta}{2}} = w + \sum_{i=1}^{r} \alpha_i |y_{t-i}|^{\delta} + \sum_{i=1}^{s} \beta_i h_{t-i}^{\frac{\delta}{2}}, \tag{3.1}$$

where $\delta > 0$, $w > 0$, $\alpha_i \geq 0$, $\beta_i \geq 0$, and η_t is i.i.d. innovation with mean zero and variance one. Model (3.1) as a special case of the asymmetric power GARCH models introduced by Ding, Granger, and Engle (1993) was first proposed by Higgins and Bera (1992). When $\delta = 2$, model (3.1) reduces to the GARCH model in Bollerslev (1986), and when $\delta = 1$, it reduces to the absolute value GARCH model in Taylor (1986) and Schwert (1989). For more discussions on ARCH-type models, we refer to Bollerslev, Chou, and Kroner (1992) and Francq and Zakoïan (2010).

So far, PGARCH models have been widely applied in practice; see, e.g., McKenzie and Mitchell (2002), Giot and Laurent (2004), and the references therein. However, except GARCH model (i.e., $\delta = 2$), the diagnostic checking tools for model (3.1) are rare. In the sequel, we briefly review some related works. Li and Mak (1994) proposed a portmanteau test to detect the adequacy of GARCH models based on the squared residuals $\{\hat{\eta}_t^2\}$; Carbon and Francq (2011) extended their work to detect the adequacy of asymmetric PGARCH models; Li and Li (2005) considered the model checking in GARCH models by using the absolute residuals $\{|\hat{\eta}_t|\}$; see also Li and Li (2008) and Chen and Zhu (2015) for other portmanteau tests in GARCH models. In this section, we apply $\widetilde{Q}_m$, $\overline{Q}_m$, $\widetilde{M}_m$, and $\overline{M}_m$ to check the adequacy of model (3.1).

Assume that $\tau \triangleq (E|\eta_t|^\delta)^{1/\delta}(> 0)$ is well defined. Then, we can re-parameterize model (3.1) as

$$y_t = \eta_t^* \sqrt{h_t^*}, \;\; (h_t^*)^{\frac{\delta}{2}} = w_1^* + \sum_{i=1}^{r} \alpha_i^* |y_{t-i}|^\delta + \sum_{i=1}^{s} \beta_i (h_{t-i}^*)^{\frac{\delta}{2}}, \tag{3.2}$$

where $w_1^* = w\tau^{-\delta}$, $\alpha_i^* = \alpha_i \tau^{-\delta}$, and $\eta_t^* = \eta_t \tau^{-1}$. Moreover, by (3.2), a direct calculation gives us that

$$|y_t|^\delta = w_1^* + \sum_{i=1}^{\max(r,s)} (\alpha_i^* + \beta_i)|y_{t-i}|^\delta + v_t - \sum_{i=1}^{s} \beta_i v_{t-i}, \tag{3.3}$$

where we set $\alpha_i^* = 0$ if $i > r$ and $\beta_i = 0$ if $i > s$, and $v_t = (h_t^*)^{\frac{\delta}{2}}(|\eta_t^*|^\delta - 1)$ is a martingale difference. From (3.3), we know that if model (3.1) is well specified, the mean-adjusted series $\tilde{y}_t \triangleq |y_t|^\delta - E|y_t|^\delta$ admits a weak ARMA(max(r,s),s) representation:

$$\tilde{y}_t = \sum_{i=1}^{\max(r,s)} \tilde{\alpha}_i \tilde{y}_{t-i} + v_t - \sum_{i=1}^{s} \beta_i v_{t-i}, \tag{3.4}$$

where $\tilde{\alpha}_i = \alpha_i^* + \beta_i$. Thus, we can check the adequacy of model (3.1) by $\widetilde{Q}_m$, $\overline{Q}_m$, $\widetilde{M}_m$, and $\overline{M}_m$, which are calculated from model (3.4). Denote $\tilde{\alpha}(z) \triangleq 1 - \sum_{i=1}^{\max(r,s)} \tilde{\alpha}_i z^i$ and $\beta(z) \triangleq 1 - \sum_{i=1}^{s} \beta_i z^i$. We have the following corollary:

Corollary 3.1. *Suppose that (i) Assumption 2.1 holds with $\phi(z)$ and $\psi(z)$ being replaced by $\tilde{\alpha}(z)$ and $\beta(z)$, respectively; (ii) Assumption 2.2 holds with y_t being replaced by $\tilde{y}_t$. Then, if model (3.1) is correctly specified, the conclusions in Theorems 2.1 and 2.3 hold, where $\widetilde{Q}_m$, $\overline{Q}_m$, $\widetilde{M}_m$, and $\overline{M}_m$ are calculated from model (3.4).*

We now offer some remarks of Corollary 3.1. First, it is worth noting that a sufficient condition for condition (ii) in Corollary 3.1 can be found in Carrasco and Chen (2002). Second, since model (3.4) is not a strong ARMA model but a weak ARMA model with a martingale difference error term, the critical values of these four tests should be obtained by the RW method. Third, it is necessary to point out that our portmanteau tests are infeasible to check the adequacy of some asymmetric ARCH-type models, which do not have a weak ARMA representation like (3.4).

4. Simulation study. In this section, we examine the finite-sample performance of $\widetilde{Q}_m$, $\overline{Q}_m$, $\widetilde{M}_m$ and $\overline{M}_m$. For $\widetilde{Q}_m$, we denote it by $\widetilde{Q}_m^{(1)}$, when we choose its critical value to be the upper percentage of $\chi^2_{m-(p+q)}$ as in strong ARMA models, while we denote it by $\widetilde{Q}_m^{(2)}$ (or $\widetilde{Q}_m^{(3)}$), when we choose its critical value to be the one from the RW method with $W^* = I_m$ (or $\widetilde{W}^*$ as in (2.4)). For $\overline{Q}_m$, we denote it by $\overline{Q}_m^{(1)}$, when we choose its critical value to be the upper percentage of Gamma(a, b) as in (2.3), while we denote it by $\overline{Q}_m^{(2)}$ (or $\overline{Q}_m^{(3)}$), when we choose its critical value to be the one from the RW method with $W^* = W$ (or $\overline{W}^*$ as in (2.4)). Likewise, we can define $\widetilde{M}_m^{(i)}$ and $\overline{M}_m^{(i)}$ for $i = 1, 2, 3$. In all calculations, we set the significance level $\alpha = 5\%$ and the bootstrap sample size $J = 500$, and generate random weights from the standard exponential distribution. In addition, more simulation studies are given in the on-line supplementary document.

4.1. *Study on fitted weak ARMA models.* In this subsection, we assess the finite-sample performance of $\widetilde{Q}_m$, $\overline{Q}_m$, $\widetilde{M}_m$ and $\overline{M}_m$ on testing the model adequacy in weak ARMA models. As a comparison, we also consider the spectral test in Hong (1996) and the robust M test (denoted by KL_m) in Kuan and Lee (2006). For Hong's kernel-based spectral test H_n, we denote it by $H_B^{(1)}$, $H_B^{(2)}$, and $H_B^{(3)}$, when we use the Bartlett kernel with the bandwidth $p_n = \lfloor 3n^{0.3} \rfloor$, $\lfloor 9n^{0.3} \rfloor$, and $\lfloor 12n^{0.3} \rfloor$, respectively (see, e.g., Kuan and Lee (2006)). Likewise, we can define $H_D^{(i)}$, $H_P^{(i)}$, and $H_Q^{(i)}$ for $i = 1, 2, 3$, when we use the Daniell, Parzen, and Quadratic kernels, respectively.

In size simulations, we generate 1000 replications of sample size $n = 100$ and 500 from the following model:

$$y_t = 0.9y_{t-1} + \varepsilon_t, \ \varepsilon_t = \eta_t \sqrt{h_t} \ \text{ and } \ h_t = 1 + 0.4\varepsilon_{t-1}^2, \tag{4.1}$$

where $\{\eta_t\}$ is a sequence of i.i.d. $N(0, 1)$ random variables. Table 1 reports the empirical sizes for all tests. From this table, our findings are as follows:

(ai) $\widetilde{Q}_m^{(1)}$, $\overline{Q}_m^{(1)}$, $\widetilde{M}_m^{(1)}$ or $\overline{M}_m^{(1)}$ has a severe over-rejection problem for each m, and hence the critical values taken for strong ARMA models are not valid for weak ARMA models.

(aii) When n is small, the sizes of KL_m become more conservative as m becomes larger, while the sizes of $\widetilde{Q}_m^{(2)}$, $\overline{Q}_m^{(2)}$, $\widetilde{M}_m^{(2)}$ or $\overline{M}_m^{(2)}$ tend to be larger than their nominal ones in general. In this case, the sizes of $\widetilde{Q}_m^{(3)}$, $\overline{Q}_m^{(3)}$, $\widetilde{M}_m^{(3)}$ or $\overline{M}_m^{(3)}$ are close to their nominal ones, although they tend to be conservative for large m.

(aiii) When n is large, the sizes of KL_m still have a severe conservative problem for large m. In this case, the sizes of $\widetilde{Q}_m^{(2)}$, $\overline{Q}_m^{(2)}$, $\widetilde{M}_m^{(2)}$ or $\overline{M}_m^{(2)}$ are close to their nominal ones for each m, while the sizes of $\widetilde{Q}_m^{(3)}$, $\overline{Q}_m^{(3)}$, $\widetilde{M}_m^{(3)}$ or $\overline{M}_m^{(3)}$ are slightly conservative for large m.

(aiv) For Hong's spectral test, its size performance is quite sensitive to the choice of the kernel function and related bandwidth p_n. To look for further evidence, Figure

TABLE 1

Empirical sizes (×100) of all tests based on model (4.1).

		Tests												
m	n	$\widetilde{Q}_m^{(1)}$	$\widetilde{Q}_m^{(2)}$	$\widetilde{Q}_m^{(3)}$	$\overline{Q}_m^{(1)}$	$\overline{Q}_m^{(2)}$	$\overline{Q}_m^{(3)}$	$\widetilde{M}_m^{(1)}$	$\widetilde{M}_m^{(2)}$	$\widetilde{M}_m^{(3)}$	$\overline{M}_m^{(1)}$	$\overline{M}_m^{(2)}$	$\overline{M}_m^{(3)}$	KL_m
2	100	19.7	5.4	5.1	20.6	6.1	5.4	19.1	6.0	5.8	21.2	6.0	5.8	4.0
	500	23.4	5.5	5.4	24.9	5.8	5.8	23.8	5.3	5.2	25.3	5.9	5.9	5.4
6	100	11.9	6.0	4.8	13.0	6.3	5.0	11.6	5.7	4.3	13.5	6.3	5.4	1.7
	500	14.7	5.6	5.5	17.3	5.0	5.0	15.0	5.5	5.3	16.9	4.9	4.7	5.3
18	100	8.5	7.5	3.1	8.7	6.3	3.4	6.4	6.4	1.7	8.0	4.7	2.8	1.4
	500	10.6	5.9	5.0	12.0	5.0	5.0	9.9	5.2	4.4	10.7	4.7	4.3	1.8
20	100	8.4	7.4	3.5	8.6	6.6	3.4	6.2	5.9	2.4	7.2	5.0	2.9	1.1
	500	11.0	6.0	5.3	11.1	5.2	4.8	9.2	4.8	4.1	10.1	4.8	4.0	1.5

	Tests											
n	$H_B^{(1)}$	$H_B^{(2)}$	$H_B^{(3)}$	$H_D^{(1)}$	$H_D^{(2)}$	$H_D^{(3)}$	$H_P^{(1)}$	$H_P^{(2)}$	$H_P^{(3)}$	$H_Q^{(1)}$	$H_Q^{(2)}$	$H_Q^{(3)}$
100	5.7	4.2	3.3	6.6	4.1	3.4	6.3	4.7	4.4	5.3	3.5	2.5
500	11.2	9.1	7.9	13.7	11.2	10.2	12.4	10.3	9.1	12.0	8.0	6.7

1 plots the sizes of H_B, H_D, H_P, and H_Q in the cases that $p_n = 2, 3, \cdots, \lfloor n/2 \rfloor$. As a comparison, the sizes of the portmanteau tests $\widetilde{Q}_m^{(i)}$, $\overline{Q}_m^{(i)}$, $\widetilde{M}_m^{(i)}$, and $\overline{M}_m^{(i)}$ (for $i = 2, 3$) are also plotted in the cases that $m = 2, 3, \cdots, 20$ when $n = 100$ or $m = 2, 3, \cdots, 50$ when $n = 500$. From Figure 1, we find that for every choice of the kernel function, the sizes of the spectral test can suffer a severe over-rejection problem under a very wide range of p_n, especially when n is large, while the sizes of all portmanteau tests have a robust size performance over m. This indicates that a good data-driven method for selecting p_n is very important for the application of Hong's spectral tests.

Therefore, all of these imply that the portmanteau tests along with the bootstrapped critical values from the RW method have a good size performance, while the size performance of the kernel-based spectral test in Hong (1996) heavily depends on the choice of the kernel function and related bandwidth. Particularly, we recommend $\widetilde{Q}_m^{(2)}$, $\overline{Q}_m^{(2)}$, $\widetilde{M}_m^{(2)}$ or $\overline{M}_m^{(2)}$ for large n, and $\widetilde{Q}_m^{(3)}$, $\overline{Q}_m^{(3)}$, $\widetilde{M}_m^{(3)}$ or $\overline{M}_m^{(3)}$ for small n in applications.

Next, in power simulations, we generate 1000 replications of sample size $n = 100$ and 500 from the following model:

$$y_t = 0.9y_{t-1} + 0.2\varepsilon_{t-1} + \varepsilon_t, \ \varepsilon_t = \eta_t\sqrt{h_t} \ \text{ and } \ h_t = 1 + 0.4\varepsilon_{t-1}^2, \tag{4.2}$$

where $\{\eta_t\}$ is a sequence of i.i.d. $N(0, 1)$ random variables. For each replication, we fit it by an AR(1) model, and then use all tests to check the adequacy of this fitted model. Table 2 reports the empirical power for all tests. Since the sizes of $\widetilde{Q}_m^{(1)}$, $\overline{Q}_m^{(1)}$, $\widetilde{M}_m^{(1)}$, $\overline{M}_m^{(1)}$, $H_B^{(i)}$, $H_D^{(i)}$, $H_P^{(i)}$, and $H_Q^{(i)}$ (for $i = 1, 2, 3$) are severely distorted in Table 1, the power of these tests in Table 2 has been adjusted by the size-correction method in Francq, Roy, and Zakoïan (2005, p.541) (that is, for each test T_n with a severe size distortion, its critical value used for model (4.2) is chosen as the α upper percentage of $\{T_{ni}\}_{i=1}^{1000}$, where T_{ni} is the value of T_n for the i-th replication from model (4.1)). From Table 2, our findings are as follows:

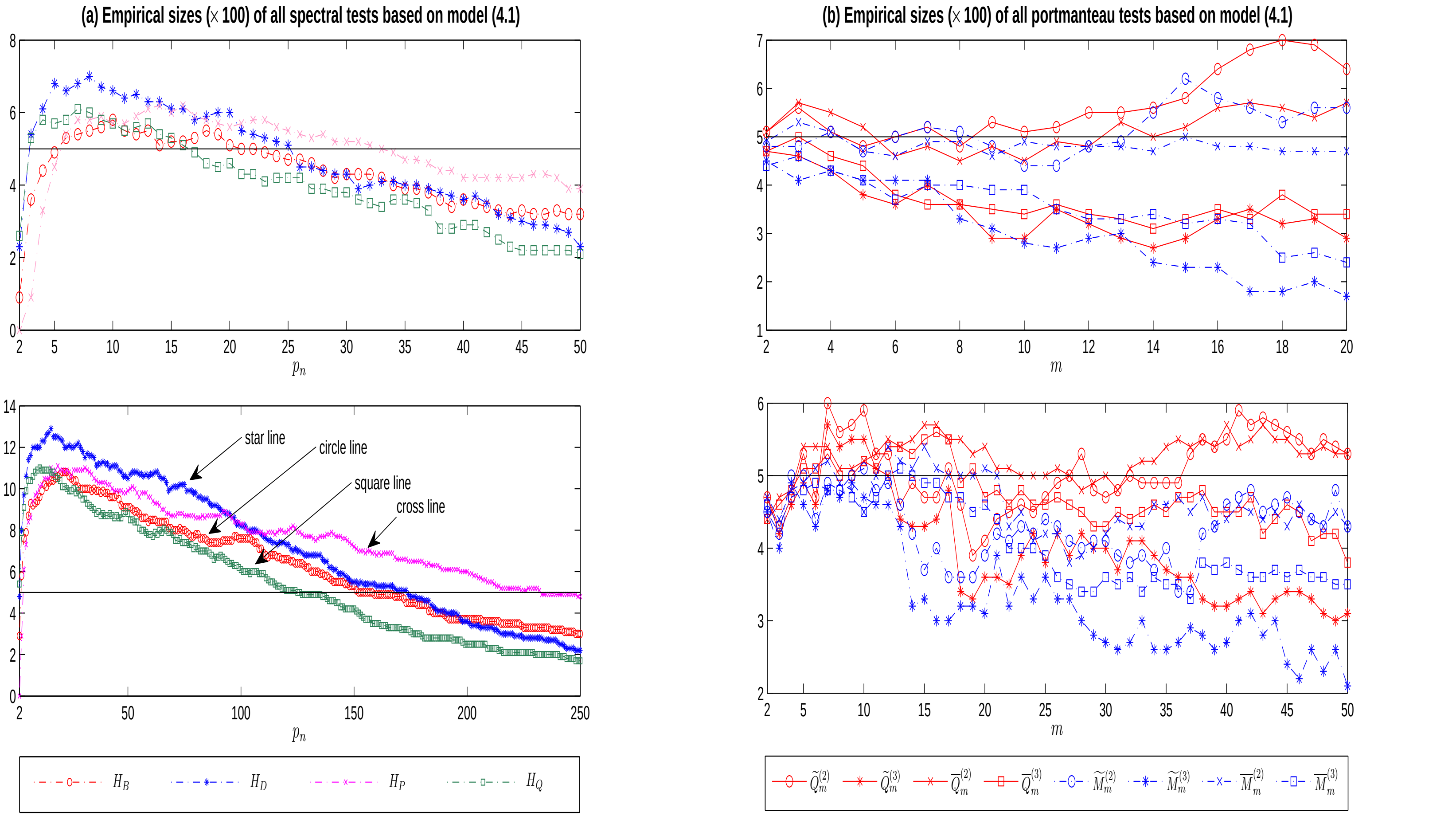

FIG 1. *The empirical sizes ($\times 100$) of all spectral tests (part (a)) and all portmanteau tests (part (b)) based on model (4.1) for the cases that $n = 100$ (upper panel) and $n = 500$ (lower panel). Here, the solid line stands for the significance level $\alpha = 5\%$.*

TABLE 2
Empirical power (×100) of all tests based on model (4.2).

		Tests												
m	n	$\widetilde{Q}_m^{(1)}$	$\widetilde{Q}_m^{(2)}$	$\widetilde{Q}_m^{(3)}$	$\overline{Q}_m^{(1)}$	$\overline{Q}_m^{(2)}$	$\overline{Q}_m^{(3)}$	$\widetilde{M}_m^{(1)}$	$\widetilde{M}_m^{(2)}$	$\widetilde{M}_m^{(3)}$	$\overline{M}_m^{(1)}$	$\overline{M}_m^{(2)}$	$\overline{M}_m^{(3)}$	KL_m
2	100	23.5	25.4	23.3	28.2	28.6	27.8	26.6	30.1	28.5	29.1	31.8	30.7	14.7
	500	85.0	80.0	79.8	87.8	83.4	83.2	86.8	82.2	82.1	89.2	84.7	84.4	56.2
6	100	17.0	17.1	14.3	24.3	24.0	22.6	21.2	20.9	17.5	27.6	28.0	25.5	4.5
	500	74.7	71.2	70.2	83.3	79.8	79.4	78.0	73.7	73.0	85.6	81.9	81.8	36.6
18	100	14.9	17.2	10.4	18.1	17.5	14.1	16.0	14.0	8.0	23.7	18.9	14.1	1.7
	500	61.7	58.0	55.1	73.2	71.7	70.7	66.1	59.2	56.5	77.1	73.7	72.5	14.9
20	100	14.8	17.3	10.1	18.0	18.1	13.6	15.8	13.6	6.0	23.1	18.4	13.0	0.9
	500	59.4	56.4	54.3	72.5	71.2	69.5	61.8	57.5	54.2	76.5	72.5	71.0	12.3

	Tests											
n	$H_B^{(1)}$	$H_B^{(2)}$	$H_B^{(3)}$	$H_D^{(1)}$	$H_D^{(2)}$	$H_D^{(3)}$	$H_P^{(1)}$	$H_P^{(2)}$	$H_P^{(3)}$	$H_Q^{(1)}$	$H_Q^{(2)}$	$H_Q^{(3)}$
100	25.1	20.1	17.7	23.4	17.8	16.3	26.8	19.8	19.2	23.4	16.4	16.3
500	72.6	59.5	57.2	68.8	53.9	52.0	74.6	60.9	55.7	67.9	52.9	49.2

(bi) As we expected, the power of all tests becomes large as n increases, and the power of all portmanteau tests becomes small as m increases. Moreover, the power performance of Ljung-Box-type portmanteau tests and Monti-type portmanteau tests is generally comparable, while the weighted portmanteau test is more powerful than the corresponding un-weighted one, and this power advantage grows as m increases.

(bii) When n is small, $\widetilde{Q}_m^{(2)}$, $\overline{Q}_m^{(2)}$, $\widetilde{M}_m^{(2)}$ or $\overline{M}_m^{(2)}$ is more powerful than $\widetilde{Q}_m^{(3)}$, $\overline{Q}_m^{(3)}$, $\widetilde{M}_m^{(3)}$ or $\overline{M}_m^{(3)}$, respectively, while this power advantage disappears as n becomes large. This is probably because $\widetilde{Q}_m^{(3)}$, $\overline{Q}_m^{(3)}$, $\widetilde{M}_m^{(3)}$ or $\overline{M}_m^{(3)}$ has a conservative size for small n. Moreover, the adjusted portmanteau test $\widetilde{Q}_m^{(1)}$, $\overline{Q}_m^{(1)}$, $\widetilde{M}_m^{(1)}$ or $\overline{M}_m^{(1)}$ has a similar power as $\widetilde{Q}_m^{(i)}$, $\overline{Q}_m^{(i)}$, $\widetilde{M}_m^{(i)}$ or $\overline{M}_m^{(i)}$, respectively, for $i = 1, 2$, especially when n is large. This is reasonable because all of them are based on the same statistic, and the size-correction method becomes more accurate when n is larger.

(biii) For KL_m, its power is not satisfactory when n is small and m is large. Like portmanteau tests, its power becomes smaller as m becomes larger.

(biv) For Hong's spectral test, its power performance varies significantly in terms of kernel function and related bandwidth. To look for further evidence, Figure 2 plots the power of H_B, H_D, H_P, and H_Q, and as a comparison, the power of the portmanteau tests $\widetilde{Q}_m^{(i)}$, $\overline{Q}_m^{(i)}$, $\widetilde{M}_m^{(i)}$, and $\overline{M}_m^{(i)}$ (for $i = 2, 3$) is also plotted in this figure. Here, the choices of p_n and m are the same as those for Figure 1. From Figure 2, we find that H_B and H_P have a similar power performance, and they are more powerful than H_D and H_Q, especially when n is large. Moreover, we can see that the power of all spectral tests generally tends to be smaller as p_n becomes larger. Compared with weighted portmanteau tests, the spectral tests exhibit a dominated power advantage only when p_n is close to 2 and m is greater than 30 in the case that $n = 500$; however, when $p_n > 50$ in the case that $n = 500$, H_D and H_Q are much less powerful than the weighted portmanteau tests. On the contrary, since the power of un-weighted portmanteau tests is not stable over m, the spectral tests with an appropriate p_n (e.g., $p_n < 10$ in the

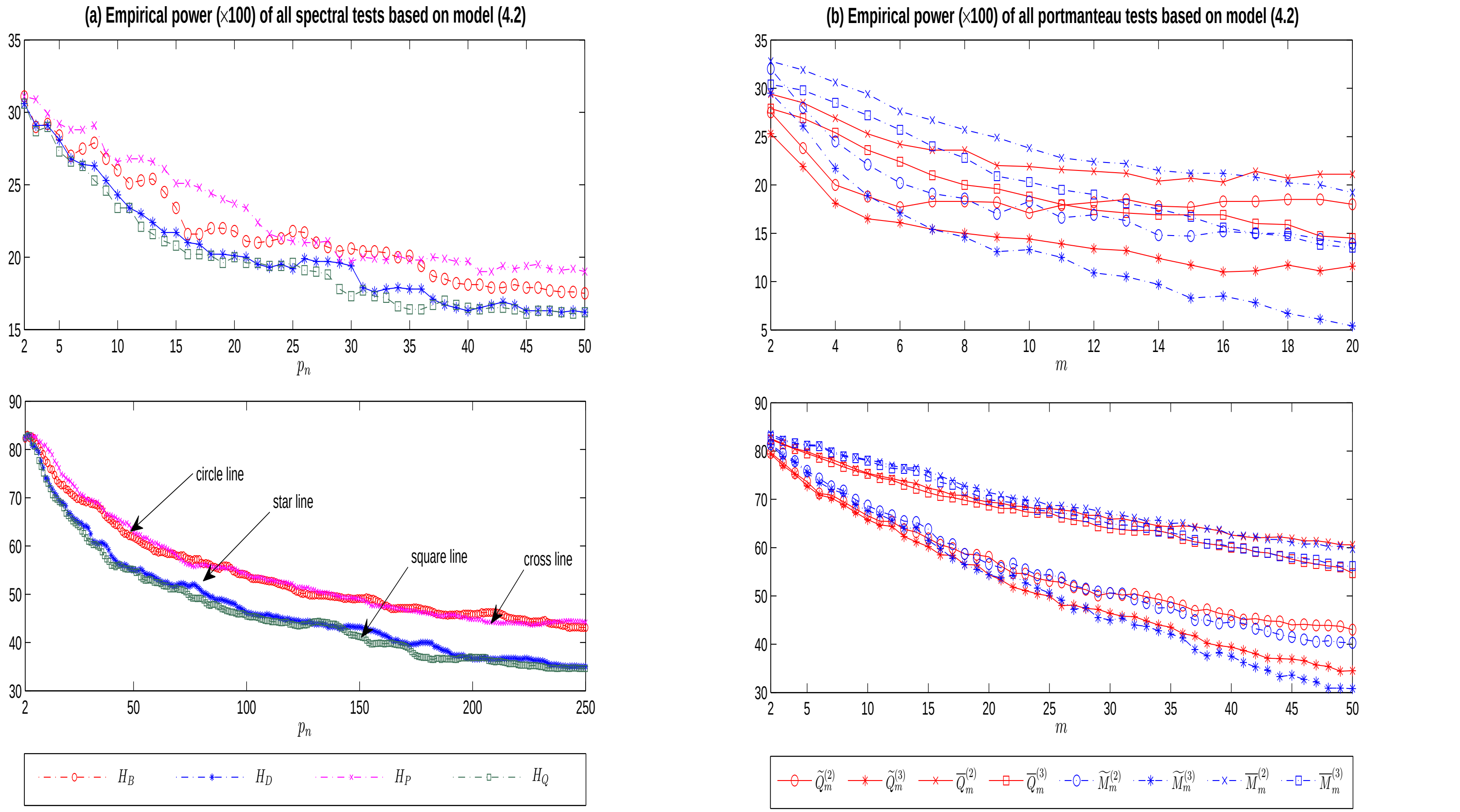

FIG 2. *The empirical power (×100) of all spectral tests (part (a)) and all portmanteau tests (part (b)) based on model (4.2) for the cases that $n = 100$ (upper panel) and $n = 500$ (lower panel).*

case that $n = 100$ and $p_n < 50$ in the case that $n = 500$) can be much more powerful than the un-weighted portmanteau tests with a large m.

Overall, based on the bootstrapped critical values, the portmanteau tests, especially the weighted ones, give us a good indication in diagnostic checking of weak ARMA models, while the size/power performance of the spectral test is sensitive to the choice of kernel function and related bandwidth, and hence the spectral test urgently calls for a good data-driven method for selecting the bandwidth.

4.2. *Study on fitted PGARCH models.* In this subsection, we examine the finite-sample performance of $\widetilde{Q}_m$, $\overline{Q}_m$, $\widetilde{M}_m$ and $\overline{M}_m$ on testing the adequacy of PGARCH models. As a comparison, we also consider the portmanteau test Q_m^{CF} in Carbon and Francq (2011). In size simulations, we generate 1000 replications of sample size $n = 1000$ from the following model:

$$y_t = \eta_t \sqrt{h_t}, \quad h_t^{\frac{\delta}{2}} = 0.2 + 0.2|y_{t-1}|^\delta, \tag{4.3}$$

where $\{\eta_t\}$ is a sequence of i.i.d. $N(0,1)$ random variables. Table 3 reports the results for the size study. From Table 3, we find that in general, the sizes of all tests are close to their nominal ones, while the sizes of $\widetilde{Q}_m$, $\overline{Q}_m$, $\widetilde{M}_m$ and $\overline{M}_m$ seem to be conservative when both δ and m are large.

TABLE 3
Empirical sizes ($\times 100$) of all tests based on model (4.3).

		Tests				
δ	m	$\widetilde{Q}_m$	$\overline{Q}_m$	$\widetilde{M}_m$	$\overline{M}_m$	Q_m^{CF}
2.5	6	4.9	5.5	4.9	5.7	2.7
	12	2.6	4.0	2.8	4.3	3.1
	24	1.7	2.1	2.0	2.3	2.6
	32	0.9	1.6	1.5	1.9	3.3
2.0	6	3.5	4.8	3.9	4.6	2.6
	12	3.3	4.3	3.3	4.3	2.9
	24	2.1	2.6	2.1	2.7	4.0
	32	2.3	2.3	2.3	2.4	3.7
1.0	6	7.3	7.2	7.1	6.5	6.6
	12	6.5	7.3	6.3	6.5	8.3
	24	5.4	6.6	4.8	6.2	6.7
	32	5.8	5.9	5.1	5.5	7.4
0.5	6	5.4	5.6	5.4	5.6	3.5
	12	6.3	5.0	6.0	5.2	4.1
	24	5.2	6.1	5.4	5.3	4.1
	32	5.0	5.7	5.2	5.8	3.4

Next, in power simulations, we generate 1000 replications of sample size $n = 1000$ from the following model:

$$y_t = \eta_t \sqrt{h_t}, \quad h_t^{\frac{\delta}{2}} = 0.2 + 0.2|y_{t-1}|^\delta + 0.2|y_{t-2}|^\delta, \tag{4.4}$$

where $\{\eta_t\}$ is a sequence of i.i.d. $N(0,1)$ random variables. For each replication, we fit its mean-adjusted series $\tilde{y}_t$ by an AR(1) model as in (3.4), and then we use all portmanteau tests to check the adequacy of this fitted model. Table 4 reports the results for the power study, where the critical values for $\widetilde{Q}_m$, $\overline{Q}_m$, $\widetilde{M}_m$ or $\overline{M}_m$ are chosen as for $\widetilde{Q}_m^{(2)}$, $\overline{Q}_m^{(2)}$, $\widetilde{M}_m^{(2)}$ or $\overline{M}_m^{(2)}$, respectively. From Table 4, we find that when δ is small, the weighted tests $\overline{Q}_m$ and $\overline{M}_m$ are more powerful than others. However, when δ is large, Q_m^{CF} is the best performing test statistic. Overall, our weighted portmanteau tests have a good performance especially when δ is small.

TABLE 4

Empirical power ($\times 100$) of all tests based on model (4.4).

		Tests				
δ	m	$\widetilde{Q}_m$	$\overline{Q}_m$	$\widetilde{M}_m$	$\overline{M}_m$	Q_m^{CF}
2.5	6	51.2	55.9	51.1	54.9	97.1
	12	45.4	51.9	44.3	50.8	94.8
	24	38.0	46.8	34.8	45.2	89.4
	32	33.7	43.8	30.9	41.9	86.1
2.0	6	74.7	79.2	73.1	78.8	94.7
	12	65.8	75.2	65.5	74.3	89.6
	24	56.9	68.5	54.6	68.3	82.0
	32	50.6	65.8	48.6	65.0	77.6
1.0	6	95.4	98.2	95.7	98.1	94.9
	12	90.2	96.0	90.4	96.5	91.6
	24	80.9	92.8	80.5	92.6	81.6
	32	74.0	90.5	72.4	91.4	77.4
0.5	6	98.3	99.7	98.5	99.7	95.5
	12	95.6	98.9	95.6	98.9	90.6
	24	88.0	97.5	88.2	97.4	82.2
	32	84.5	96.7	83.1	96.2	77.9

5. A real example. In this section, we revisit the real example on S&P 500 index in Francy, Roy, and Zakoïan (2005). The data sample taken from January 3, 1979 to December 31, 2001 has in total 5808 observations, and its log-return is denoted by $\{y_t\}_{t=1}^{5807}$. First, we use $\widetilde{Q}_m$, $\overline{Q}_m$, $\widetilde{M}_m$ and $\overline{M}_m$ to check whether $\{y_t\}$ is a sequence of white noises. The p-value is calculated either as in strong ARMA models (denoted by p-value1) or via the RW method with J = 500 (denoted by p-value2). All testing results are reported in Table 5, from which we know that at the significance level 5%, the strong white noise hypothesis is rejected, while the weak white noise hypothesis is not rejected. This is consistent to the findings in Francy, Roy, and Zakoïan (2005).

Next, we want to check whether a PGARCH(1, 1) model in (3.1) with $\delta = 2.0$, 1.0 or 0.5 is adequate to fit $\{y_t\}$. Denote by $\tilde{y}_t = |y_t|^\delta - E|y_t|^\delta$ the mean-adjusted series. As in (3.4), we obtain the following three fitted models:

$$\text{Model A}: \tilde{y}_t = 0.8287\tilde{y}_{t-1} - 0.7236v_{t-1} + v_t \text{ with } \sigma_{v_t}^2 = 5.26 \times 10^{-7} \text{ and } \delta = 2.0;$$

$$\text{Model B}: \tilde{y}_t = 0.9753\tilde{y}_{t-1} - 0.8979v_{t-1} + v_t \text{ with } \sigma_{v_t}^2 = 4.99 \times 10^{-5} \text{ and } \delta = 1.0;$$

$$\text{Model C}: \tilde{y}_t = 0.9920\tilde{y}_{t-1} - 0.9513v_{t-1} + v_t \text{ with } \sigma_{v_t}^2 = 1.30 \times 10^{-3} \text{ and } \delta = 0.5.$$

For models A-C, we use $\widetilde{Q}_m$, $\overline{Q}_m$, $\widetilde{M}_m$ and $\overline{M}_m$ to check their adequacy. As a comparison, we also use Q_m^{CF} to check the adequacy of these fitted models. All testing results are reported in Table 6. From this table, we find that at the significance level 5%, all tests suggest that models A and B are adequate, while except Q_m^{CF}, the other tests indicate that model C is not adequate. To look for further evidence, Figure 3 plots the residuals autocorrelations and partial autocorrelations for models A-C. From Figure 3, we know that the 1st, 5th, 22th, 25th and 33th residuals autocorrelations or partial autocorrelations in model C all exceed the 95% significance bounds. Thus, it suggests that model C is not adequate. Moreover, since model A has a smaller sum of squared errors than model B, we prefer to fit $\{y_t\}$ by a PGARCH$(1,1)$ model with $\delta = 2$ (i.e., GARCH$(1,1)$ model). Finally, it is worth noting that although the conclusions drawn from $\widetilde{Q}_m$, $\overline{Q}_m$, $\widetilde{M}_m$ and $\overline{M}_m$ are the same, the p-value2s of $\overline{Q}_m$ or $\overline{M}_m$ tend to be more stable over all lags for model C. In view of this, we recommend the weighted test $\overline{Q}_m$ or $\overline{M}_m$ to practitioners.

Table 5
Testing results for the adequacy of a white noise on S&P 500 index.

		m						
		6	12	18	24	30	36	42
$\widetilde{Q}_m$	statistic	25.58	36.07	38.49	46.62	64.49	78.77	87.71
	p-value1[†]	0.0003	0.0003	0.0033	0.0037	0.0003	0.0000	0.0000
	p-value2[§]	0.4758	0.5283	0.6874	0.6990	0.5292	0.4634	0.4518
$\overline{Q}_m$	statistic	20.64	25.17	28.98	32.25	37.65	43.73	49.37
	p-value1	0.0000	0.0000	0.0001	0.0002	0.0002	0.0001	0.0000
	p-value2	0.3508	0.4805	0.5494	0.5967	0.5817	0.5719	0.5477
$\widetilde{M}_m$	statistic	25.06	36.39	38.78	46.63	63.75	76.68	86.08
	p-value1	0.0003	0.0003	0.0033	0.0037	0.0003	0.0000	0.0000
	p-value2	0.4881	0.5219	0.6833	0.6990	0.5441	0.4977	0.4776
$\overline{M}_m$	statistic	20.40	24.87	28.89	32.30	37.72	43.47	48.85
	p-value1	0.0000	0.0000	0.0001	0.0002	0.0002	0.0001	0.0000
	p-value2	0.3578	0.4871	0.5507	0.5941	0.5783	0.5754	0.5545

[†] p-values taken as in strong ARMA models.
[§] p-values bootstrapped by the RW method with $W^* = I_m$ (or W) for $\widetilde{Q}_m$ and $\widetilde{M}_m$ (or $\overline{Q}_m$ and $\overline{M}_m$).

6. Block-wise random weighting approach. Although the bootstrapped critical values based on the RW method are valid for weak ARMA models, the condition in Assumption 2.4 seems to be slightly restrictive in practice. In this section, we propose a block-wise RW method to obtain the bootstrapped critical values for all portmanteau tests. Our bootstrapped critical values from the block-wise RW method are valid without posing Assumption 2.4, but it requires a selection of block size as many block-wise bootstrap methods in the literature; see, e.g., Romano and Thombs (1996), Horowitz, Lobato, Nankervis, and Savin (2006), and Shao (2011b).

The detailed procedure to bootstrap the critical values by the block-wise RW method is the same as the one in steps 1-4 in Section 2, except replacing step 1 by step $1'$ as follows:

step $1'$. Set a block size b_n such that $1 \leq b_n < n$, and denote the blocks by $B_s =$

Table 6
Testing results for the adequacy of a PGARCH(1, 1) model on S&P 500 index.

			m						
			6	12	18	24	30	36	42
$\delta = 2.0$	$\widetilde{Q}_m$	statistic	87.44	96.42	97.39	99.61	105.6	110.3	111.7
		p-value1†	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
		p-value2§	0.4063	0.4358	0.4416	0.4483	0.4432	0.4428	0.4246
	$\overline{Q}_m$	statistic	40.60	67.95	77.56	82.86	86.67	90.39	93.35
		p-value1	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
		p-value2	0.4349	0.4293	0.4319	0.4431	0.4452	0.4475	0.4281
	$\widetilde{M}_m$	statistic	86.51	106.20	108.45	110.80	115.93	121.02	123.28
		p-value1	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
		p-value2	0.4115	0.3841	0.3845	0.3929	0.3911	0.3903	0.3677
	$\overline{M}_m$	statistic	40.77	71.77	83.64	90.31	94.79	98.91	102.26
		p-value1	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
		p-value2	0.4325	0.4009	0.3934	0.3995	0.3983	0.3993	0.3811
	Q_m^{CF}	statistic	1.44	7.79	10.44	16.53	23.62	29.08	56.45
		p-value	0.9635	0.8014	0.9166	0.8680	0.7889	0.7864	0.0673
$\delta = 1.0$	$\widetilde{Q}_m$	statistic	22.90	39.57	46.11	67.54	76.99	94.13	101.4
		p-value1	0.0001	0.0000	0.0001	0.0000	0.0000	0.0000	0.0000
		p-value2	0.4497	0.4093	0.4174	0.2739	0.2839	0.2091	0.2151
	$\overline{Q}_m$	statistic	9.15	20.58	28.30	34.78	42.27	50.10	56.98
		p-value1	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
		p-value2	0.6445	0.5258	0.4703	0.4507	0.4044	0.3553	0.3248
	$\widetilde{M}_m$	statistic	22.72	38.80	45.40	67.75	75.87	88.27	95.44
		p-value1	0.0001	0.0000	0.0001	0.0000	0.0000	0.0000	0.0000
		p-value2	0.4544	0.4244	0.4309	0.2717	0.2966	0.2672	0.2768
	$\overline{M}_m$	statistic	9.09	20.71	28.14	34.59	42.16	49.26	55.37
		p-value1	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
		p-value2	0.6474	0.5219	0.4745	0.4556	0.4069	0.3723	0.3497
	Q_m^{CF}	statistic	1.83	1.98	2.02	2.10	2.11	2.13	2.23
		p-value	0.9350	0.9994	1.0000	1.0000	1.0000	1.0000	1.0000
$\delta = 0.5$	$\widetilde{Q}_m$	statistic	18.64	22.86	25.07	39.42	49.92	59.99	65.88
		p-value1	0.0009	0.0113	0.0687	0.0126	0.0066	0.0039	0.0061
		p-value2	0.0204	0.0847	0.2465	0.0962	0.0571	0.0417	0.0515
	$\overline{Q}_m$	statistic	10.97	15.82	18.62	21.58	26.85	31.92	36.36
		p-value1	0.0000	0.0009	0.0063	0.0149	0.0092	0.0062	0.0054
		p-value2	0.0339	0.0405	0.0703	0.1090	0.0784	0.0637	0.0555
	$\widetilde{M}_m$	statistic	19.45	23.08	25.30	40.35	50.74	61.14	67.90
		p-value1	0.0009	0.0113	0.0687	0.0126	0.0066	0.0039	0.0061
		p-value2	0.0172	0.0820	0.2385	0.0837	0.0515	0.0351	0.0389
	$\overline{M}_m$	statistic	11.30	16.23	18.96	22.01	27.45	32.56	37.10
		p-value1	0.0000	0.0009	0.0063	0.0149	0.0092	0.0062	0.0054
		p-value2	0.0295	0.0344	0.0634	0.0985	0.0679	0.0566	0.0475
	Q_m^{CF}	statistic	0.0927	0.0930	0.0931	0.0934	0.0935	0.0936	0.0942
		p-value	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000

† p-values taken as in strong ARMA models.

§ p-values bootstrapped by the RW method with $W^* = I_m$ (or W) for $\widetilde{Q}_m$ and $\widetilde{M}_m$ (or $\overline{Q}_m$ and $\overline{M}_m$).

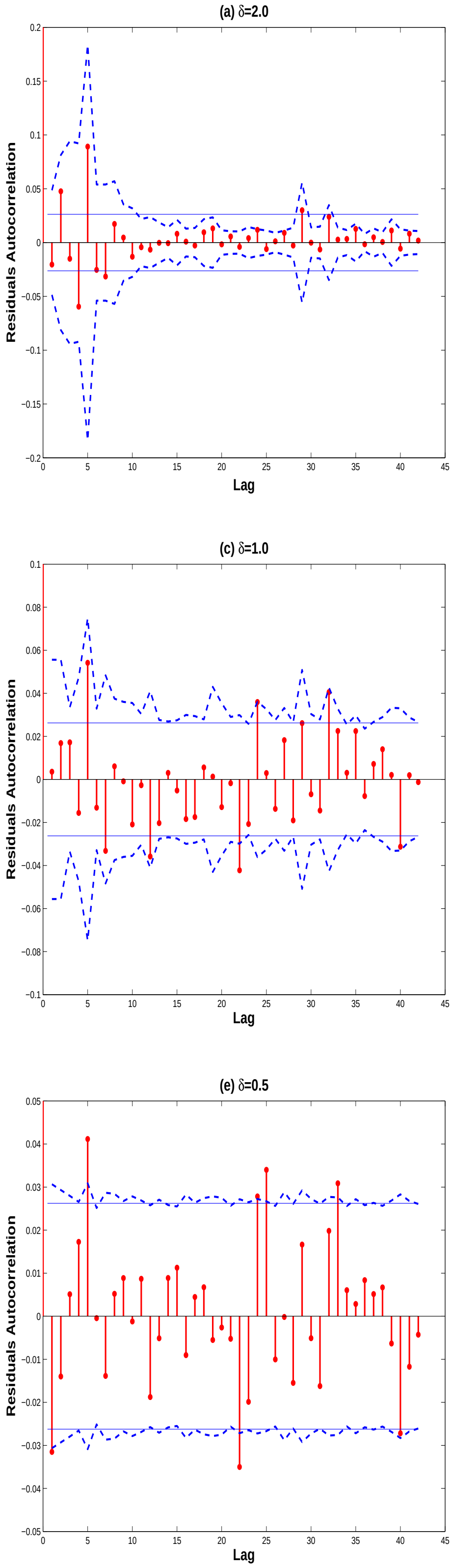

FIG 3. *The residuals autocorrelations and partial autocorrelations for models A-C. The solid lines are 95% significance bounds under the strong ARMA model. The dashed lines are 95% significance bounds under the weak ARMA model.*

$\{(s-1)b_n+1, \cdots, sb_n\}$ for $s = 1, \cdots, L_n$, where $L_n = n/b_n$ is assumed to be an integer for the convenience of presentation; furthermore, generate a sequence of positive i.i.d. random variables $\{\delta_1, \cdots, \delta_{L_n}\}$, independent of the data, from a common distribution with mean and variance both equal to 1, and then define the random weights $w_t^* = \delta_s$, if $t \in B_s$, for $t = 1 \cdots, n$; finally, calculate $\hat{\theta}_n^*$ via

$$\hat{\theta}_n^* = \arg\min_{\Theta} \tilde{L}_n^*(\theta), \quad \text{where } \tilde{L}_n^*(\theta) = \frac{1}{n}\sum_{t=1}^{n} w_t^* \tilde{l}_t(\theta).$$

Clearly, the block-wise RW method is a natural extension of the RW method, and the validity of the bootstrapped critical value from the block-wise RW method is justified by the following theorem:

THEOREM 6.1. *Suppose that (i) Assumptions 2.1-2.3 hold and model (1.1) is weak and correctly specified; (ii)* $E|y_t|^{8+4\nu} < \infty$ *for some* $\nu > 0$ *and* $\lim_{k\to\infty} k^2[\alpha_y(k)]^{\nu/(2+\nu)} = 0$*; and (iii)* $\lim_{n\to\infty} b_n = \infty$ *and* $\lim_{n\to\infty} b_n/n^{1/3} = 0$*. Then, conditional on* χ_n*,*

$$\sqrt{n}(\hat{\rho}^* - \hat{\rho}) \to_d N(0, \Sigma) \quad \textit{in probability}$$

as $n \to \infty$*, where* Σ *is defined in (2.1).*

The proof of Theorem 6.1 is given in the Appendix. Here, condition (ii) poses some additional requirements on y_t, and condition (iii) gives some restrictions on the block-size b_n. Both of them are necessary for the proof. From Theorem 6.1, we know that the performance of our tests along with their critical values from the block-wise RW method depends on the user-chosen parameter b_n. This may be the price we pay for not assuming Assumption 2.4. Simulation studies in the on-line supplementary document show that our testing results are not sensitive to the user-chosen parameter b_n, while the size performance of Hong's (1996) kernel-based spectral test is not robust to the choices of the bandwidth p_n. Till now, how to select the optimal b_n under certain "criterion" is unknown. This is a familiar problem with all blocking methods. The heuristic work in Hall, Horowitz, and Jing (1995), Politis, Romano, and Wolf (1999), and Sun (2014) may be extended in this case, and we leave it for future study.

7. Concluding remarks. In this paper, by using the RW method, we bootstrap the critical values for Ljung-Box/Monti portmanteau tests $\widetilde{Q}_m/\widetilde{M}_m$ and weighted Ljung-Box/Monti portmanteau tests $\overline{Q}_m/\overline{M}_m$ in weak ARMA models. Unlike the existing methods (e.g., kernel-based method, VAR method, or block bootstrap method), the easy-to-implement RW method requires no user-chosen parameter. Thus, it overcomes a drawback in the existing methods that the testing results are sensitive to the choices of user-chosen parameters. As an application, we further use our portmanteau tests to check the adequacy of PGARCH models. Simulation studies reveal that (i) the weighted portmanteau tests $\overline{Q}_m/\overline{M}_m$ along with the critical values from the RW method have the power advantage over the un-weighted ones in general; (ii) the sizes of all portmanteau tests are robust to the choices of the lag m, while the sizes of the kernel-based spectral tests in Hong (1996) are sensitive to the choices of the bandwidth p_n; and (iii) the weighted portmanteau tests $\overline{Q}_m/\overline{M}_m$ can be significantly more powerful than the kernel-based spectral tests with an inappropriate choice of p_n. As one extension work, we also propose a block-wise RW method to bootstrap the critical values for all portmanteau tests, and its validity is justified.

Finally, we suggest three future subjects, which may lead to some better specification tests. First, as in Escanciano and Lobato (2009), it is of interest to consider the case that m is not fixed but optimally chosen by the data set. If it is possible, a more powerful testing procedure should be expected. Second, till now, less is known to choose the optimal weight matrix W (in some sense) such that the corresponding weighted Ljung-Box or Monti portmanteau test has the best performance among all weighted portmanteau tests in strong ARMA models. We may expect that the merits of this optimal weighted portmanteau test still hold in weak ARMA models. Third, since all portmanteau tests still need a selection of m, they can not detect serial correlations beyond lag m. Hence, it is a practical demand to study the Cramér-von Mises spectral test (e.g., Shao (2011b)) which can detect serial correlations at all lags.

Acknowledgement. I am grateful to the two referees, the Associate Editor and the Joint Editor I. Van Keilegom for very constructive suggestions and comments, leading to a substantial improvement in the presentation and the elimination of two errors in the previous manuscript. This work is partially supported by National Natural Science Foundation of China (No.11201459), President Fund of Academy of Mathematics and Systems Science, CAS (No.2014-cjrwlzx-zk), and Key Laboratory of RCSDS, CAS.

References.

[1] Beltrao, K. and Bloomfield, P. (1987) Determining the bandwidth of a kernel spectrum estimate. *Journal of Time Series Analysis* **8**, 21-38.

[2] Berk, K.N. (1974) Consistent autoregressive spectral estimates. *Annals of Statistics* **2**, 489-502.

[3] Bollerslev, T. (1986) Generalized autoregressive conditional heteroskedasticity. *Journal of Econometrics* **31**, 307-327.

[4] Bollerslev, T., Chou, R.Y., and Kroner, K.F. (1992) ARCH modeling in finance: A review of the theory and empirical evidence. *Journal of Econometrics* **52**, 5-59.

[5] Box, G.E.P. and Pierce, D.A. (1970) Distribution of the residual autocorrelations in autoregressive integrated moving average time series models. *Journal of the American Statistical Association* **65**, 1509-1526.

[6] Carbon, M. and Francq, C. (2011) Portmanteau goodness-of-fit test for asymmetric power GARCH models. *Austrian Journal of Statistics* **40**, 55-64.

[7] Carrasco, M. and Chen, X. (2002) Mixing and moment properties of various GARCH and stochastic volatility models. *Econometric Theory* **18**, 17-39.

[8] Chen, M. and Zhu, K. (2015) Sign-based portmanteau test for ARCH-type models with heavy-tailed innovations. Forthcoming in *Journal of Econometrics*.

[9] Chen, K., Ying, Z., Zhang, H., and Zhao, L. (2008) Analysis of least absolute deviation. *Biometrika* **95**, 107-122.

[10] Chen, K., Guo, S., Lin, Y., and Ying, Z. (2010) Least absolute relative error estimation. *Journal of the American Statistical Association* **105**, 1104-1112.

[11] Deo, R.S. (2000) Spectral tests of the martingale hypothesis under conditional heteroskedasticity. *Journal of Econometrics* **99**, 291-315.

[12] Ding, Z., Granger, C.W.J. and Engle, R.F. (1993) A long memory property of stock market returns and a new model. *Journal of Empirical Finance* **1**, 83-106.

[13] Duchesne, P. and Francq, C. (2008) On diagnostic checking time series models with portmanteau test statistics based on generalized inverses and {2}-inverses. *COMPSTAT 2008, Proceedings in Computational Statistics*, (ed.) P. Brito, Heidelberg: Physica-Verlag, pp. 143-154.

[14] Durlauf, S.N. (1991) Spectral-based testing of the martingale hypothesis. *Journal of Econometrics* **50**, 355-376.

[15] Engle, R.F. (1982) Autoregressive conditional heteroskedasticity with estimates of variance of U.K. inflation. *Econometrica* **50**, 987-1008.

[16] Escanciano, J.C. (2006) Goodness-of-fit tests for linear and non-linear time series models. *Journal of the American Statistical Association* **101**, 531-541.

[17] Escanciano, J.C. (2007) Model checks using residual marked emprirical processes. *Statistica Sinica* **17**, 115-138.

[18] Escanciano, J.C. and Velasco, C. (2006) Generalized spectral tests for the martingale difference hypothesis. *Journal of Econometrics* **134**, 151-185.

[19] Escanciano, J.C. and Lobato, I.N. (2009) An automatic portmanteau test for serial correlation. *Journal of Econometrics* **151**, 140-149.

[20] Fisher, T.J. and Gallagher, C.M. (2012) New weighted portmanteau statistics for time series goodness of fit testing. *Journal of the American Statistical Association* **107**, 777-787.

[21] Francq, C., Roy, R., and Zakoïan, J.M. (2005) Diagostic checking in ARMA models with uncorrelated errors. *Journal of the American Statistical Association* **100**, 532-544.

[22] Francq, C. and Zakoïan, J.M. (1998) Estimating linear representations of nonlinear processes. *Journal of Statistical Planning and Inference* **68**, 145-165.

[23] Francq, C. and Zakoïan, J.M. (2010) GARCH Models: Structure, Statistical Inference and Financial Applications. Wiley, Chichester, UK.

[24] Franses, P.H. and Van Dijk, R. (1996) Forecasting stock market volatility using (non-linear) Garch models. *Journal of Forecasting* **15**, 229-235.

[25] Giot, P. and Laurent, S. (2004) Modelling daily value-at-risk using realized volatility and Arch type models. *Journal of Empirical Finance* **11**, 379-398.

[26] Hall, P., Horowitz, J.L., and Jing, B.-Y. (1995) On blocking rules for the bootstrap with dependent data. *Biometrica* **82**, 561-574.

[27] Higgins, M.L. and Bera, A.K. (1992) A class of nonlinear ARCH models. *International Economic Review* **33**, 137-158.

[28] Hong, Y. (1996) Consistent testing for serial correlation of unknown form. *Econometrica* **64**, 837-864.

[29] Hong, Y. (1999) Hypothesis testing in time series via the empirical characteristic function: a generalized spectral density approach. *Journal of the American Statistical Association* **94**, 1201-1220.

[30] Hong, Y. and Lee, T.H. (2003) Diagnostic checking for adequacy of nonlinear time series models. *Econometric Theory* **19**, 1065-1121.

[31] Hong, Y. and Lee, Y.-J. (2005) Generalized spectral testing for conditional mean models in time series with conditional heteroskedasticity of unknown form. *Review of Economic Studies* **72**, 499-541.

[32] Horowitz, J.L., Lobato, I.N., Nankervis, J.C., and Savin, N.E. (2006) Bootstrapping the Box-Pierce Q test: A robust test of uncorrelatedness. *Journal of Econometrics* **133**, 841-862.

[33] Jin, Z., Ying, Z. and Wei, L.J. (2001) A simple resampling method by perturbing the minimand. *Biometrika* **88**, 381-390.

[34] Kuan, C.-M. and Lee, W.-M. (2006) Robust M tests without consistent estimation of the asymptotic covariance matrix. *Journal of the American Statistical Association* **101**, 1264-1275.

[35] Lee, J. and Hong, Y. (2001) Testing for serial correlation of unknown form using wavelet methods. *Econometric Theory* **17**, 386-423.

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

[37] Li, W.K. (2004) Diagnostic checks in time series. Chapman& Hall/CRC.

[38] Li, G., Leng, C. and Tsai, C.-L. (2014) A hybrid bootstrap approach to unit root tests. *Journal of Time Series Analysis* **35**, 299-321.

[39] Li, G. and Li, W.K. (2005) Diagnostic checking for time series models with conditional heteroscedasticity estimated by the least absolute deviation approach. *Biometrika* **92**, 691-701.

[40] Li, G. and Li, W.K. (2008). Least absolute deviation estimation for fractionally integrated autoregressive moving average time series models with conditional heteroscedasticity. *Biometrika* **95**, 399-414.

[41] Li, G., Li, Y. and Tsai, C.-L. (2014) Quantile correlations and quantile autoregressive modeling. Forthcoming in *Journal of the American Statistical Association.*

[42] Li, W.K. and Mak, T.K. (1994) On the squared residual autocorrelations in non-linear time series with conditional heteroscedasticity. *Journal of Time Series Analysis* **15**, 627-636.

[43] Ljung, G.M. and Box, G.E.P. (1978) On a measure of lack of fit in time series models. *Biometrika* **65**, 297-303.

[44] Lobato, I.N. (2001) Testing that a dependent process is uncorrelated. *Journal of the American Statistical Association* **96**, 1066-1076.

[45] McKenzie, M. and Mitchell, H. (2002) Generalised asymmetric power arch modeling of exchange rate volatility. *Applied Financial Economics* **12**, 555-564.

[46] McLeod, A.I. and Li, W.K. (1983) Diagnostic checking ARMA time series models using squaredresidual autocorrelations. *Journal of Time Series Analysis* **4**, 269-273.

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

[48] Monti, A.C. (1994) A proposal for a residual autocorrelation test in linear models. *Biometrika* **81**, 776-780.

[49] Newey, W.K. and West, K.D. (1987) A simple positive semi-definite heteroskedasticity and autocorrelation consistent covariance matrix. *Econometrica* **55**, 703-708.

[50] Peña, D. and Rodríguez, J. (2002) A powerful portmanteau test of lack of fit for time series. *Journal of the American Statistical Association* **97**, 601-610.

[51] Peña, D. and Rodríguez, J. (2006), The Log of the determinant of the autocorrelation matrix for testing goodness of fit in time series. *Journal of Statistical Planning and Inference* **136**, 2706-2718.

[52] Romano, J.L. and Thombs, L.A. (1996) Inference for autocorrelations under weak assumptions. *Journal of the American Statistical Association* **91**, 590-600.

[53] Politis, D.N., Romano, J., and Wolf, M. (1999) *Subsampling* Springer-Verla, New York.

[54] Schwert, G.W. (1989) Why does stock market volatility change over time? *Journal of Finance* **45**, 1129-1155.

[55] Shao, X. (2011a). Testing for white noise under unknown dependence and its applications to diagnostic checking for time series models. *Econometric Theory* **27**, 312-343.

[56] Shao, X. (2011b) A bootstrap-assisted spectral test of white noise under unknown dependence.

Journal of Econometrics **162**, 213-224.

[57] Stinchcombe, M. and White, H. (1998) Consistent specification testing with nuisance parameters present only under the alternative. *Econometric Theory* **14**, 295-325.

[58] Sun, Y. (2014) Let's fix it: Fixed-b asymptotics versus small-b asymptotics in heteroskedasticity and autocorrelation robust inference. *Journal of Econometrics* **178**, 659-677.

[59] Taylor, S. (1986) Modelling Financial Time Series. Wiley, New York.

[60] Tsay, R.S. (2005) Analysis of Financial Time Series (2nd ed.). New York: John Wiley&Sons, Incorporated.

[61] Wu, C.F.J. (1986) Jacknife, bootstrap and other resampling methods in regression analysis (with discussion). *Annals of Statistics* **14**, 1261-1350.

[62] Zhu, K. and Ling, S. (2011) Global self-weighted and local quasi-maximum exponential likelihood estimators for ARMA-GARCH/IGARCH models. *Annals of Statistics* **39**, 2131-2163.

[63] Zhu, K. and Ling, S. (2015) LADE-based inference for ARMA models with unspecified and heavy-tailed heteroscedastic noises. Forthcoming in *Journal of the American Statistical Association.*

Chinese Academy of Sciences
Institute of Applied Mathematics
Haidian District, Zhongguancun
Beijing, China
E-mail: kzhu@amss.ac.cn

BOOTSTRAPPING THE PORTMANTEAU TESTS IN WEAK AUTO-REGRESSIVE MOVING AVERAGE MODELS (SUPPLEMENTARY DOCUMENT)

By Ke Zhu

Chinese Academy of Sciences

In this supplementary material, we give some simulation studies and all of the proofs for the paper in Appendix.

1. Additional simulation studies. In this section, we give some simulation studies, which are useful but not reported in the paper. In Subsection 1.1, we check the sensitivity of the distribution of random weights to our portmanteau tests by simulation. In Subsection 1.2, we examine the finite-sample performance of our portmanteau tests as in Section 4 of the paper by some additional simulations. In Subsection 1.3, we assess the finite-sample performance of the block-wise RW method.

1.1. *Sensitivity of the distribution of random weights.* In simulation studies of the paper, we generate random weights $\{w_t^*\}$ from the standard exponential distribution. In this subsection, we check the sensitivity of the distribution of $\{w_t^*\}$ to our portmanteau tests by generating $\{w_t^*\}$ from the Bernoulli distribution such that

$$P\left(w_t^* = \frac{3-\sqrt{5}}{2}\right) = \frac{1+\sqrt{5}}{2\sqrt{5}} \text{ and } P\left(w_t^* = \frac{3+\sqrt{5}}{2}\right) = \frac{\sqrt{5}-1}{2\sqrt{5}}.$$

Table 1
Empirical sizes $(\times 100)$ *of all tests based on model (4.1) in the paper.*

		Tests							
m	n	$\widetilde{Q}_m^{(2)}$	$\widetilde{Q}_m^{(3)}$	$\overline{Q}_m^{(2)}$	$\overline{Q}_m^{(3)}$	$\widetilde{M}_m^{(2)}$	$\widetilde{M}_m^{(3)}$	$\overline{M}_m^{(2)}$	$\overline{M}_m^{(3)}$
2	100	5.8	5.2	5.6	5.3	5.3	4.8	5.4	5.1
	500	4.3	4.2	4.8	4.3	4.2	4.2	4.7	4.7
6	100	5.0	4.1	5.2	4.5	4.1	3.5	4.6	3.7
	500	3.9	3.7	4.3	4.2	3.5	3.3	3.7	3.6
18	100	6.0	3.4	5.2	4.0	4.2	1.7	3.8	2.6
	500	3.9	3.4	3.4	3.2	3.2	3.0	3.4	3.3
20	100	6.5	3.4	5.5	3.8	4.4	2.0	3.8	2.5
	500	4.0	3.7	3.4	3.0	3.1	3.0	3.0	2.8

Tables 1-2 are obtained in the same way as Tables 1-2 in the paper. From these four tables, we can see that our testing results are robust to the distribution of w_t^*.

1.2. *Some additional simulations.* In this subsection, we examine the finite-sample performance of our portmanteau tests as in Section 4 of the paper by some additional simulations. Hereafter, we generate $\{w_t^*\}$ from the standard exponential distribution.

TABLE 2
Empirical power (×100) *of all tests based on model (4.2) in the paper.*

m	n	$\widetilde{Q}_m^{(2)}$	$\widetilde{Q}_m^{(3)}$	$\overline{Q}_m^{(2)}$	$\overline{Q}_m^{(3)}$	$\widetilde{M}_m^{(2)}$	$\widetilde{M}_m^{(3)}$	$\overline{M}_m^{(2)}$	$\overline{M}_m^{(3)}$
2	100	25.6	24.3	28.4	27.2	29.4	28.2	30.7	29.2
	500	82.1	81.8	85.0	84.6	84.0	83.5	85.6	85.6
6	100	17.3	14.9	23.4	21.8	21.1	18.2	25.3	23.8
	500	74.5	74.2	80.8	80.7	77.4	76.8	82.3	82.2
18	100	17.2	11.1	19.2	13.8	14.7	7.9	18.8	14.4
	500	61.7	60.0	73.8	73.2	62.2	60.5	75.9	74.8
20	100	17.1	9.8	19.4	13.1	14.6	6.5	18.5	13.6
	500	60.4	58.4	72.9	71.9	61.4	58.9	74.7	73.7

First, we study it for fitted weak ARMA models. In size simulations, we generate 1000 replications of sample size $n = 100$ and 500 from the following model:

$$y_t = 0.9\varepsilon_{t-1} + \varepsilon_t,\ \varepsilon_t = \eta_t\eta_{t-1}, \tag{1.1}$$

where $\{\eta_t\}$ is a sequence of i.i.d. $N(0,1)$ random variables. Table 3 reports the empirical sizes for all tests except the robust M test, which is not valid for MA models. From Table 3, our findings are the same as those from Table 1 in the paper. Moreover, Figure 1 below plots the empirical sizes of all spectral tests and portmanteau tests, and our findings from this figure are the same as those from Figure 1 in the paper.

TABLE 3
Empirical sizes (×100) *of all tests based on model (1.1).*

m	n	$\widetilde{Q}_m^{(1)}$	$\widetilde{Q}_m^{(2)}$	$\widetilde{Q}_m^{(3)}$	$\overline{Q}_m^{(1)}$	$\overline{Q}_m^{(2)}$	$\overline{Q}_m^{(3)}$	$\widetilde{M}_m^{(1)}$	$\widetilde{M}_m^{(2)}$	$\widetilde{M}_m^{(3)}$	$\overline{M}_m^{(1)}$	$\overline{M}_m^{(2)}$	$\overline{M}_m^{(3)}$
2	100	17.7	5.5	4.9	20.2	6.4	5.9	18.1	6.7	6.1	21.1	6.3	6.2
	500	29.6	6.2	6.1	29.6	6.6	6.5	25.7	6.6	6.4	29.6	7.2	6.9
6	100	8.9	4.6	3.8	10.6	5.9	5.0	10.0	5.2	4.8	11.1	6.4	5.2
	500	14.5	5.1	4.9	17.7	6.3	6.3	14.5	5.6	5.5	17.2	6.4	6.3
18	100	7.2	5.3	2.1	6.7	5.2	3.3	6.3	4.8	2.4	7.2	5.1	3.5
	500	10.2	3.6	3.3	11.7	5.2	4.7	9.6	4.3	4.0	11.7	4.8	4.7
20	100	6.1	5.8	2.1	6.9	5.3	2.7	5.0	4.7	2.2	6.8	5.1	3.0
	500	9.6	4.4	3.7	11.7	5.1	4.8	9.2	4.6	4.2	11.1	4.9	4.6
		$H_B^{(1)}$	$H_B^{(2)}$	$H_B^{(3)}$	$H_D^{(1)}$	$H_D^{(2)}$	$H_D^{(3)}$	$H_P^{(1)}$	$H_P^{(2)}$	$H_P^{(3)}$	$H_Q^{(1)}$	$H_Q^{(2)}$	$H_Q^{(3)}$
	100	5.6	3.8	3.0	6.3	3.8	2.8	6.1	4.6	3.5	5.4	3.2	2.4
	500	11.2	7.6	7.1	13.0	9.0	7.9	11.8	8.5	7.8	10.0	6.7	5.4

In power simulations, we generate 1000 replications of sample size $n = 100$ and 500 from the following model:

$$y_t = 0.2y_{t-1} + 0.9\varepsilon_{t-1} + \varepsilon_t,\ \varepsilon_t = \eta_t\eta_{t-1}, \tag{1.2}$$

where $\{\eta_t\}$ is a sequence of i.i.d. $N(0,1)$ random variables. For each replication, we fit it by a MA(1) model, and then use all tests to check the adequacy of this fitted model. Table 4 reports the empirical power for all tests after the size-correction method in Francq, Roy, and Zakoïan (2005, p.541). From Table 4, our findings are the same as

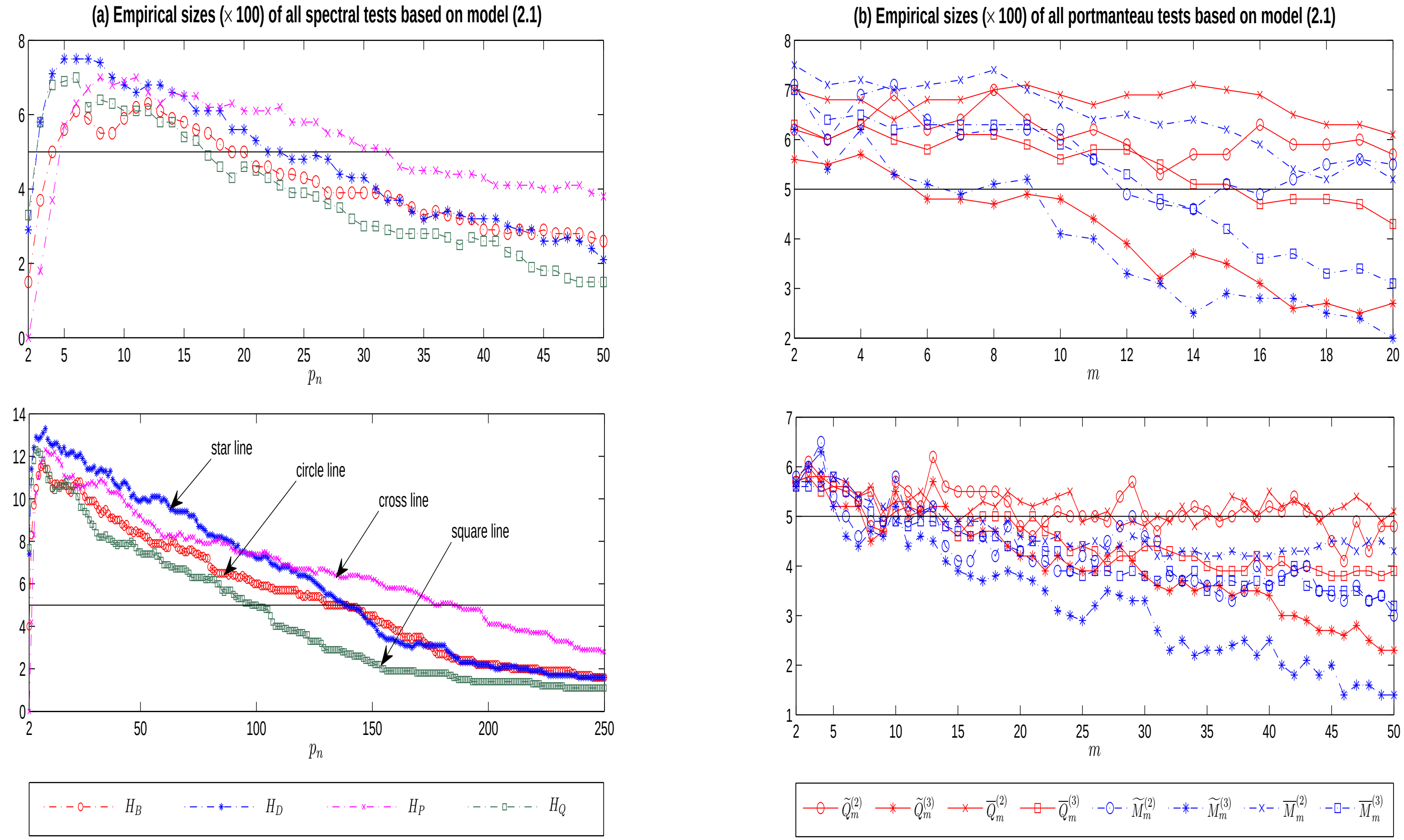

FIG 1. *The empirical sizes (×100) of all spectral tests (part (a)) and all portmanteau tests (part (b)) based on model (1.1) for the cases that $n = 100$ (upper panel) and $n = 500$ (lower panel). Here, the solid line stands for the significance level $\alpha = 5\%$.*

those from Table 2 in the paper. Moreover, Figure 2 below plots the empirical power of all spectral tests and portmanteau tests. From this figure, our findings are the same as those from Figure 2 in the paper, except that the spectral tests generally are more powerful than the portmanteau tests in the case that $n = 100$.

TABLE 4
Empirical power ($\times 100$) *of all tests based on model (1.2).*

		Tests											
m	n	$\widetilde{Q}_m^{(1)}$	$\widetilde{Q}_m^{(2)}$	$\widetilde{Q}_m^{(3)}$	$\overline{Q}_m^{(1)}$	$\overline{Q}_m^{(2)}$	$\overline{Q}_m^{(3)}$	$\widetilde{M}_m^{(1)}$	$\widetilde{M}_m^{(2)}$	$\widetilde{M}_m^{(3)}$	$\overline{M}_m^{(1)}$	$\overline{M}_m^{(2)}$	$\overline{M}_m^{(3)}$
2	100	27.3	19.9	18.9	29.0	20.1	18.8	24.7	16.7	15.2	25.7	18.4	17.6
	500	74.1	76.1	75.6	75.5	76.5	76.3	72.7	74.4	74.0	75.7	75.7	75.4
6	100	19.7	15.6	13.4	24.4	18.7	16.9	18.2	13.5	12.1	22.4	16.3	14.6
	500	68.4	68.8	68.5	71.1	74.6	74.1	67.0	67.3	66.9	70.3	72.3	71.8
18	100	16.5	14.6	9.9	20.7	16.9	13.1	12.8	10.4	7.0	18.8	13.2	9.4
	500	59.1	57.0	55.4	69.8	67.7	67.4	56.6	53.9	52.3	68.1	65.9	65.3
20	100	16.6	15.1	9.3	20.6	16.4	12.7	14.4	10.5	6.0	18.6	12.8	9.0
	500	56.2	55.5	53.9	68.9	67.1	65.8	54.5	52.3	49.8	67.4	65.3	64.1
		Tests											
		$H_B^{(1)}$	$H_B^{(2)}$	$H_B^{(3)}$	$H_D^{(1)}$	$H_D^{(2)}$	$H_D^{(3)}$	$H_P^{(1)}$	$H_P^{(2)}$	$H_P^{(3)}$	$H_Q^{(1)}$	$H_Q^{(2)}$	$H_Q^{(3)}$
	100	27.2	22.0	20.4	23.6	19.5	18.6	29.7	24.0	21.9	23.5	19.6	18.3
	500	69.2	63.7	59.0	69.3	57.6	54.6	71.3	65.0	61.9	68.4	56.2	53.4

Second, we study it for fitted PGARCH models. In size simulations, we generate 1000 replications of sample size $n = 1000$ from the following model:

$$y_t = \eta_t \sqrt{h_t}, \quad h_t^{\frac{\delta}{2}} = 0.001 + 0.08|y_{t-1}|^{\delta} + 0.9 h_{t-1}^{\frac{\delta}{2}}, \tag{1.3}$$

where $\{\eta_t\}$ is a sequence of i.i.d. $N(0,1)$ random variables. Table 5 reports the results for the size study. From Table 5, our findings are the same as those from Table 3 in the paper.

TABLE 5
Empirical sizes ($\times 100$) *of all tests based on model (1.3).*

		Tests				
δ	m	$\widetilde{Q}_m$	$\overline{Q}_m$	$\widetilde{M}_m$	$\overline{M}_m$	Q_m^{CF}
2.5	6	2.8	2.6	2.2	2.6	2.2
	12	1.3	2.6	1.0	1.6	3.5
	24	1.1	1.5	0.5	0.5	6.0
	32	0.9	0.9	0.1	0.4	6.2
2.0	6	2.7	2.3	2.1	1.8	3.3
	12	2.5	2.2	0.9	1.7	3.1
	24	2.1	2.3	0.6	0.6	7.6
	32	1.3	2.0	0.3	0.7	8.0
1.0	6	3.3	3.5	3.4	3.2	4.8
	12	5.0	3.6	3.7	3.3	6.4
	24	4.1	3.9	3.9	3.1	7.0
	32	4.0	3.8	3.3	2.8	6.1
0.5	6	4.2	4.0	4.4	4.1	4.6
	12	5.5	4.3	4.4	4.3	4.6
	24	4.9	4.9	3.9	4.5	4.6
	32	5.7	4.5	4.4	4.1	3.9

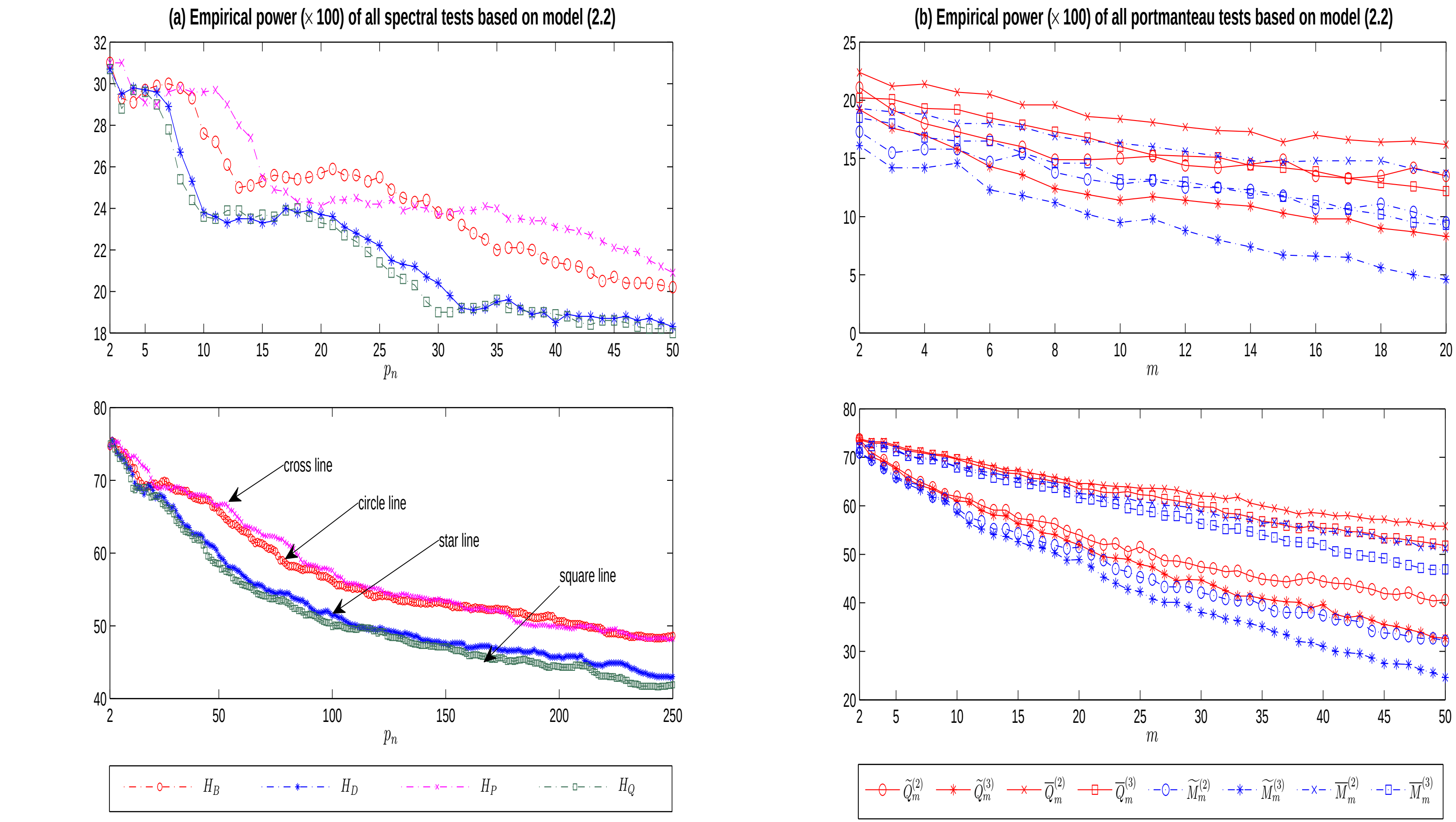

FIG 2. *The empirical power ($\times 100$) of all spectral tests (part (a)) and all portmanteau tests (part (b)) based on model (1.2) for the cases that $n = 100$ (upper panel) and $n = 500$ (lower panel).*

In power simulations, we generate 1000 replications of sample size $n = 1000$ from the following model:

$$y_t = \eta_t \sqrt{h_t}, \;\; h_t^{\frac{\delta}{2}} = 0.001 + 0.08|y_{t-1}|^{\delta} + 0.2|y_{t-2}|^{\delta} + 0.6h_{t-1}^{\frac{\delta}{2}}, \tag{1.4}$$

where $\{\eta_t\}$ is a sequence of i.i.d. $N(0,1)$ random variables. For each replication, we fit its mean-adjusted series $\tilde{y}_t \triangleq |y_t|^{\delta} - E|y_t|^{\delta}$ by an ARMA(1,1) model, and then use all portmanteau tests to check the adequacy of this fitted model. Table 6 reports the results for the power study. From Table 6, our findings are the same as those from Table 4 in the paper.

TABLE 6
Empirical power ($\times 100$) *of all tests based on model (1.4).*

		Tests				
δ	m	$\tilde{Q}_m$	$\overline{Q}_m$	$\tilde{M}_m$	$\overline{M}_m$	Q_m^{CF}
2.5	6	9.5	14.3	8.3	11.9	69.0
	12	7.8	10.2	6.2	8.3	33.1
	24	7.2	8.8	4.8	6.2	26.8
	32	6.6	8.2	3.8	5.8	24.3
2.0	6	23.5	31.2	20.1	28.9	75.0
	12	18.3	25.9	15.5	21.5	35.9
	24	16.9	20.6	12.2	17.6	28.3
	32	15.5	19.3	10.9	15.9	24.7
1.0	6	65.6	78.1	64.8	77.8	73.6
	12	53.6	70.0	53.9	69.5	58.4
	24	44.4	60.6	42.6	59.0	43.8
	32	39.9	56.4	36.6	55.3	38.7
0.5	6	84.6	91.7	84.3	91.7	75.9
	12	72.3	87.2	71.8	86.8	61.0
	24	58.6	77.4	57.6	77.9	44.8
	32	51.7	72.5	49.9	72.2	36.0

1.3. *Simulations on the block-wise RW method.* In this subsection, we assess the finite-sample performance of our portmanteau tests with critical values bootstrapped from the block-wise RW method. In size simulations, we generate 1000 replications of sample size $n = 100$ and 500 from the following model:

$$y_t = 0.9\varepsilon_{t-1} + \varepsilon_t, \;\; \varepsilon_t = \eta_t^2 \eta_{t-1}, \tag{1.5}$$

where $\{\eta_t\}$ is a sequence of i.i.d. $N(0,1)$ random variables. For the block-wise RW method, we choose $b_n = n^{1/5}$ or $2n^{1/5}$. This leads $b_n = 2$ or 5 for $n = 100$, and $b_n = 3$ or 7 for $n = 500$. Table 7 reports the empirical sizes for all tests. From Table 7, our findings are the same as those from Table 1 in the paper, except that (i) the size performance of four portmanteau tests is robust the choice of block size b_n; and (ii) the sizes of all portmanteau tests tend to be conservative when m is large.

In power simulations, we generate 1000 replications of sample size $n = 100$ and 500 from the following model:

$$y_t = 0.2y_{t-1} + 0.9\varepsilon_{t-1} + \varepsilon_t, \;\; \varepsilon_t = \eta_t^2 \eta_{t-1}, \tag{1.6}$$

where $\{\eta_t\}$ is a sequence of i.i.d. $N(0,1)$ random variables. For each replication, we fit it by a MA(1) model, and then use all tests to check the adequacy of this fitted model. Table 8 reports the empirical power for all tests after the size-correction method in Francq, Roy, and Zakoïan (2005, p.541). From Table 8, our findings are the same as those from Table 2 in the paper, except that (i) the power performance of four portmanteau tests is robust to the choices of block size b_n; (ii) the power of four portmanteau tests is relatively low when m is large and n is small; and (iii) the spectral tests generally are more powerful than the portmanteau tests, especially when n is small.

Overall, our portmanteau tests with critical values boostrapped from the block-wise RW method give us a good indication in diagnostic checking of weak ARMA models, especially for small m.

TABLE 7
Empirical sizes ($\times 100$) *of all tests based on model (1.5).*

			Tests											
m	n	b_n	$\widetilde{Q}_m^{(1)}$	$\widetilde{Q}_m^{(2)}$	$\widetilde{Q}_m^{(3)}$	$\overline{Q}_m^{(1)}$	$\overline{Q}_m^{(2)}$	$\overline{Q}_m^{(3)}$	$\widetilde{M}_m^{(1)}$	$\widetilde{M}_m^{(2)}$	$\widetilde{M}_m^{(3)}$	$\overline{M}_m^{(1)}$	$\overline{M}_m^{(2)}$	$\overline{M}_m^{(3)}$
2	100	2	20.3	5.0	4.0	25.8	6.4	5.9	20.5	6.1	5.4	25.7	7.2	6.8
		5	21.2	5.7	5.4	25.8	6.0	5.8	23.0	6.2	5.9	25.9	6.7	6.1
	500	3	31.9	5.9	5.7	36.3	6.6	6.4	32.4	6.0	5.8	36.2	6.8	6.8
		7	30.6	5.6	5.5	34.4	6.1	6.0	30.7	5.6	5.6	34.9	6.4	6.3
6	100	2	9.6	5.2	4.0	13.2	5.3	4.5	9.9	5.7	3.9	13.4	5.3	4.7
		5	10.8	4.9	4.4	13.3	5.4	4.8	11.3	5.7	4.8	14.1	6.1	5.5
	500	3	18.5	4.6	4.5	23.5	5.1	4.9	18.3	4.1	4.1	23.4	5.2	5.0
		7	18.3	2.8	2.7	21.1	4.0	3.8	17.9	2.8	2.7	22.2	4.4	4.4
18	100	2	5.3	4.1	2.0	6.3	4.4	2.8	5.2	3.6	2.3	7.0	3.8	2.2
		5	5.9	4.8	3.0	8.5	4.6	3.1	6.2	4.8	3.3	8.5	4.2	3.1
	500	3	11.6	2.4	2.0	14.4	3.2	3.0	12.1	2.3	2.0	14.7	3.6	3.3
		7	13.1	1.6	1.3	15.6	2.1	2.0	12.5	1.7	1.3	15.9	2.1	2.0
20	100	2	4.7	4.1	2.0	6.0	4.3	2.8	5.2	4.3	2.3	6.5	3.7	2.1
		5	6.1	4.9	3.4	8.2	4.9	2.9	6.0	5.4	3.5	7.9	4.2	3.3
	500	3	11.4	2.3	2.0	14.1	3.0	2.7	11.3	2.2	2.0	14.2	3.2	2.8
		7	12.4	1.7	1.4	14.7	1.9	1.9	12.2	1.7	1.2	15.5	2.3	2.2
			Tests											
			$H_B^{(1)}$	$H_B^{(2)}$	$H_B^{(3)}$	$H_D^{(1)}$	$H_D^{(2)}$	$H_D^{(3)}$	$H_P^{(1)}$	$H_P^{(2)}$	$H_P^{(3)}$	$H_Q^{(1)}$	$H_Q^{(2)}$	$H_Q^{(3)}$
	100		8.3	5.2	3.5	8.9	4.9	3.4	10.2	6.8	5.5	8.1	3.8	3.3
	500		15.5	9.9	9.2	17.3	11.8	10.0	17.2	11.6	10.1	14.5	9.5	7.5

APPENDIX: PROOFS

In this appendix, we give the proof of Theorems 2.2 and 6.1 in the paper. Denote by E^* the expectation conditional on χ_n; by $o_p^*(1)(O_p^*(1))$ a sequence of random variables converging to zero (bounded) in probability conditional on χ_n. Then, we are ready to give three lammas. The first one from Lemma A.4 of Ling (2007) makes the initial values $\{y_s; s \leq 0\}$ ignorable. The second one is directly taken from Francq, Roy, and Zakoïan (2005). The third one is crucial to prove Theorems 2.2 and 6.1.

TABLE 8
Empirical power (×100) of all tests based on model (1.6).

m	n	b_n	$\widetilde{Q}_m^{(1)}$	$\widetilde{Q}_m^{(2)}$	$\widetilde{Q}_m^{(3)}$	$\overline{Q}_m^{(1)}$	$\overline{Q}_m^{(2)}$	$\overline{Q}_m^{(3)}$	$\widetilde{M}_m^{(1)}$	$\widetilde{M}_m^{(2)}$	$\widetilde{M}_m^{(3)}$	$\overline{M}_m^{(1)}$	$\overline{M}_m^{(2)}$	$\overline{M}_m^{(3)}$
2	100	2	24.2	13.8	12.9	26.5	14.4	13.9	19.2	11.4	10.4	23.9	13.5	12.5
		5	25.1	13.9	13.4	27.1	15.0	13.7	20.4	11.7	11.0	24.3	14.2	13.1
	500	3	61.5	58.1	57.8	61.4	59.3	59.2	58.9	55.3	55.0	60.3	56.8	56.6
		7	59.7	52.4	52.2	59.0	55.1	54.8	57.1	49.2	48.7	57.8	53.1	52.8
6	100	2	17.0	8.5	6.9	23.6	10.7	9.8	15.7	8.1	6.3	19.1	9.5	8.5
		5	18.5	8.7	8.0	24.0	11.6	10.7	16.8	8.3	6.7	20.2	10.1	9.2
	500	3	59.2	49.5	48.4	61.1	56.6	56.4	56.8	46.7	46.2	58.9	54.4	53.9
		7	57.2	43.8	43.4	60.0	49.7	49.6	54.2	40.7	40.2	57.4	46.5	45.8
18	100	2	12.4	5.6	3.2	17.7	7.5	4.9	12.9	4.5	2.7	17.3	5.2	3.8
		5	11.6	6.4	4.5	19.7	7.4	6.2	11.7	5.3	3.8	17.3	5.3	3.8
	500	3	49.9	31.4	30.1	54.8	45.7	44.6	48.4	27.9	26.7	53.6	42.5	41.8
		7	48.1	27.0	24.6	54.0	39.0	38.3	45.6	23.6	21.4	53.0	36.8	35.8
20	100	2	12.8	5.9	3.0	17.2	7.2	4.7	13.1	4.6	2.2	16.7	5.1	3.3
		5	11.9	6.1	4.7	19.3	7.2	5.8	12.0	5.3	4.0	16.9	6.2	4.9
	500	3	48.7	30.3	27.8	54.5	43.7	42.8	46.5	25.4	23.9	53.0	40.1	39.3
		7	46.9	23.9	22.1	52.9	36.9	36.0	44.6	20.6	19.2	51.7	35.2	34.0
			$H_B^{(1)}$	$H_B^{(2)}$	$H_B^{(3)}$	$H_D^{(1)}$	$H_D^{(2)}$	$H_D^{(3)}$	$H_P^{(1)}$	$H_P^{(2)}$	$H_P^{(3)}$	$H_Q^{(1)}$	$H_Q^{(2)}$	$H_Q^{(3)}$
	100		23.1	18.8	17.9	22.4	16.9	17.4	23.4	19.7	18.0	21.8	16.7	16.5
	500		67.5	59.0	56.3	65.8	55.0	51.6	69.7	59.6	57.6	65.2	54.7	49.3

LEMMA A.1. *Suppose that Assumption 2.1 holds. Then,*

$$(i) \sup_{\Theta} |\varepsilon_t(\theta) - \tilde{\varepsilon}_t(\theta)| = O(\rho^t)\xi_{\rho 0};$$
$$(ii) \sup_{\Theta} \left\| \frac{\partial \varepsilon_t(\theta)}{\partial \theta} - \frac{\partial \tilde{\varepsilon}_t(\theta)}{\partial \theta} \right\| = O(\rho^t)\xi_{\rho 0};$$
$$(iii) \sup_{\Theta} \left\| \frac{\partial \varepsilon_t^2(\theta)}{\partial \theta \partial \theta'} - \frac{\partial \tilde{\varepsilon}_t^2(\theta)}{\partial \theta \partial \theta'} \right\| = O(\rho^t)\xi_{\rho 0},$$

for some $\rho \in (0,1)$, where $\xi_{\rho t} = 1 + \sum_{i=0}^{\infty} \rho^i |y_{t-i}|$.

LEMMA A.2. *Suppose that the conditions in Theorem 2.1 hold. Then,*

$$\sqrt{n}\hat{\rho} = \sigma^{-2}\left(\sqrt{n}\Gamma_m\right) + \Lambda_m'(\sqrt{n}\mathcal{J}^{-1}Y_n) + o_p(1),$$

where $Y_n = -2/n \sum_{t=1}^n \varepsilon_t(\partial \varepsilon_t/\partial \theta)$ and $\Gamma_m = (\gamma_1, \cdots, \gamma_m)'$ with $\gamma_k = n^{-1}\sum_{t=k+1}^n \varepsilon_t \varepsilon_{t-k}$.

LEMMA A.3. *Suppose that the conditions in Theorem 2.2 or 6.1 hold. Then,*

$$\sqrt{n}\hat{\rho}^* = \sigma^{-2}\left(\sqrt{n}\Gamma_m^*\right) + \Lambda_m'(\sqrt{n}\mathcal{J}^{-1}Y_n^*) + o_p^*(1),$$

where $Y_n^* = -2/n \sum_{t=1}^n w_t^* \varepsilon_t(\partial \varepsilon_t/\partial \theta)$ and $\Gamma_m^* = (\gamma_1^*, \cdots, \gamma_m^*)'$ with $\gamma_k^* = n^{-1}\sum_{t=k+1}^n w_t^*$ $\varepsilon_t \varepsilon_{t-k}$.

Proof. We only give the proof under the conditions in Theorem 6.1, because the proof under the conditions in Theorem 2.2 is similar. By the definition of $\hat{\theta}_n^*$, it is straightforward to see that

$$\hat{\theta}_n^* - \theta_0 = -\left[\frac{1}{n}\sum_{t=1}^{n} w_t^* \frac{\partial^2 \tilde{l}_t(\xi_n)}{\partial\theta\partial\theta'}\right]^{-1}\left[\frac{1}{n}\sum_{t=1}^{n}(w_t^*-1)\frac{\partial \tilde{l}_t(\theta_0)}{\partial\theta} + \frac{1}{n}\sum_{t=1}^{n}\frac{\partial \tilde{l}_t(\theta_0)}{\partial\theta}\right]$$

$$\triangleq -\left[s_{1n}\right]^{-1}\left[s_{2n}+s_{3n}\right], \tag{A.1}$$

where ξ_n lies between θ_0 and $\hat{\theta}_n^*$. Let $l_t(\theta) = \varepsilon_t^2(\theta)$. First, by Lemma A.4 in Ling (2007), the ergodic theorem and Lemma A.1, it follows that

$$E^*\|s_{1n}\| = \frac{1}{n}\sum_{t=1}^{n}\left\|\frac{\partial^2 \tilde{l}_t(\xi_n)}{\partial\theta\partial\theta'}\right\| \le \frac{1}{n}\sum_{t=1}^{n}\sup_{\theta}\left\|\frac{\partial^2 l_t(\theta)}{\partial\theta\partial\theta'}\right\| + o_p(1) = O_p(1),$$

which entails $s_{1n} = O_p^*(1)$. Next, by a direct calculation and Lemma A.1, we have

$$\begin{aligned} E^*\left[s_{2n}s_{2n}'\right] &= \frac{1}{n^2}\sum_{s=1}^{L_n}\left[\sum_{t,t'\in B_s}\frac{\partial \tilde{l}_t(\theta_0)}{\partial\theta}\frac{\partial \tilde{l}_{t'}(\theta_0)}{\partial\theta'}\right] \\ &= \frac{1}{n^2}\sum_{s=1}^{L_n}\left[\sum_{t,t'\in B_s}\frac{\partial l_t(\theta_0)}{\partial\theta}\frac{\partial l_{t'}(\theta_0)}{\partial\theta'}\right] + o_p(1) \\ &\triangleq s_{4n} + o_p(1). \end{aligned} \tag{A.2}$$

Moreover, since $E[l_t(\theta_0)/\partial\theta] = 0$, by the stationarity of $l_t(\theta)$, we can show that

$$E\left(s_{4n}\right) = \frac{b_n}{n^2}\sum_{s=1}^{L_n} var\left[\frac{1}{\sqrt{b_n}}\sum_{t\in B_s}\frac{\partial l_t(\theta_0)}{\partial\theta}\right] = \frac{1}{n}var\left[\frac{1}{\sqrt{b_n}}\sum_{t=1}^{b_n}\frac{\partial l_t(\theta_0)}{\partial\theta}\right] = O\left(\frac{1}{n}\right),$$

duo to the fact that $b_n^{-1} = o(1)$ and Lemma 3 in Francq and Zakoïan (1998). Thus, by (A.2), we have $\sqrt{n}s_{2n} = O_p^*(1)$. Moreover, since $s_{3n} = o_p(1)$ by the ergodic theorem, it follows that $\tilde{\theta}_n^* - \theta_0 = o_p^*(1)$, and consequently, by using Theorem 3.1 in Ling and McAleer (2003), it is not hard to show that

$$\frac{1}{n}\sum_{t=1}^{n}(w_t^*-1)\frac{\partial^2 \tilde{l}_t(\xi_n)}{\partial\theta\partial\theta'} = o_p^*(1) \quad\text{and}\quad \frac{1}{n}\sum_{t=1}^{n}\frac{\partial^2 \tilde{l}_t(\xi_n)}{\partial\theta\partial\theta'} = \mathcal{J} + o_p^*(1).$$

Therefore, we have $s_{1n} = \mathcal{J} + o_p^*(1)$. Note that $\sqrt{n}s_{3n} = O_p(1)$ by Lemma 3 in Francq and Zakoïan (1998) and Lemma A.1. Thus, by (A.1), it follows that

$$\sqrt{n}(\hat{\theta}_n^* - \theta_0) = \sqrt{n}\mathcal{J}^{-1}Y_n^* + o_p^*(1). \tag{A.3}$$

Now, the conclusion follows directly from (A.3), Lemma A.1 and Taylor's expansion. □

Proof of Theorem 2.2. By Lemmas A.1-A.3, it is not hard to show that

$$\sqrt{n}(\hat{\rho}^* - \hat{\rho}) = \mathcal{V}\left(\sum_{t=1}^{n} \tilde{z}_{tn}^*\right) + o_p^*(1), \tag{A.4}$$

where $\mathcal{V} = [\sigma^{-2} I_m, -2\Lambda_m' \mathcal{J}^{-1}]$ and $\tilde{z}_{tn}^* = (w_t^* - 1)\tilde{z}_t/\sqrt{n}$ with

$$\tilde{z}_t = \left(\tilde{\varepsilon}_t\tilde{\varepsilon}_{t-1}, \cdots, \tilde{\varepsilon}_t\tilde{\varepsilon}_{t-m}, \tilde{\varepsilon}_t \frac{\partial \tilde{\varepsilon}_t}{\partial \theta'}\right)'.$$

Let $c \in \mathcal{R}^{m+p+q}$ be a constant vector. We now consider the conditional distribution of $\sum_{t=1}^{n} c'\tilde{z}_{tn}^*$. First, since w_t^* is independent to y_t with $Ew_t^* = 1$, we know that

$$E^*[c'\tilde{z}_{tn}^*] = 0. \tag{A.5}$$

Second, since $var(w_t^*) = 1$, by Assumption 2.4, Lemma A.1 and the same lines as done in Francq, Roy, and Zakoïan (2005, p.543), we can show that

$$\begin{aligned}
\sum_{t=1}^{n} E^*\left[c'\tilde{z}_{tn}^*\tilde{z}_{tn}^{*\prime}c\right] &= c'\left(\frac{1}{n}\sum_{t=1}^{n}\tilde{z}_t\tilde{z}_t'\right)c \\
&= c'\left(\frac{1}{n}\sum_{t=1}^{n} z_t z_t'\right)c + o_p(1) \\
&\to c'\begin{pmatrix} \sigma^4\mathcal{R}_m & \sigma^4\mathcal{S}_m' \\ \sigma^4\mathcal{S}_m & \mathcal{I}/4 \end{pmatrix} c \triangleq c'\mathcal{O}c \qquad \text{(A.6)}
\end{aligned}$$

in probability as $n \to \infty$, where z_t is defined in the same way as $\tilde{z}_t$ with ε_t replacing $\tilde{\varepsilon}_t$. Third, we check the Lindeberg condition. Let C_0 be a positive generic constant. Since $E(w_t^*)^{2+\kappa_0} < \infty$ from Assumption 2.3, by Hölder and Markov inequalities, for all $t = 1, \cdots, n$ and any given $\eta > 0$, we have

$$\begin{aligned}
&E^*\left[(w_t^* - 1)^2 I(|c'\tilde{z}_{tn}^*| > \eta)\right] \\
&\le \left[E^*\left(|w_t^* - 1|^{2+\delta_0}\right)\right]^{\frac{2}{2+\kappa_0}} \left[E^*\left(I(|c'\tilde{z}_{tn}^*| > \eta)\right)\right]^{\frac{\kappa_0}{2+\kappa_0}} \\
&\le C_0\left[\frac{E^*|c'\tilde{z}_{tn}^*|}{\eta}\right]^{\frac{\kappa_0}{2+\kappa_0}} \\
&\le C_0\left[\mathbf{1}' \times \tilde{K}_n\right]^{\frac{\kappa_0}{2+\kappa_0}}, \qquad \text{(A.7)}
\end{aligned}$$

where $\mathbf{1} = (1, 1, \cdots, 1)' \in \mathcal{R}^{m+p+q}$ and

$$\tilde{K}_n = \left(\frac{1}{\sqrt{n}}\max_{1\le t\le n}|\tilde{\varepsilon}_t\tilde{\varepsilon}_{t+1}|, \cdots, \frac{1}{\sqrt{n}}\max_{1\le t\le n}|\tilde{\varepsilon}_t\tilde{\varepsilon}_{t+m}|, \frac{1}{\sqrt{n}}\max_{1\le t\le n}\left\|\tilde{\varepsilon}_t\frac{\partial\tilde{\varepsilon}_t}{\partial\theta'}\right\|\right)'.$$

Then, by (A.7), it follows that

$$\begin{aligned}
&\sum_{t=1}^{n} E^*\left[c'\tilde{z}_{tn}^*\tilde{z}_{tn}^{*\prime}cI(|c'\tilde{z}_{tn}^*|>\eta)\right] \\
&\quad = c'\left\{\frac{1}{n}\sum_{t=1}^{n}\tilde{z}_t\tilde{z}_t' E^*\left[(w_t^*-1)^2 I(|c'\tilde{z}_{tn}^*|>\eta)\right]\right\}c \\
&\quad \le C_0\left[\mathbf{1}'\times\tilde{K}_n\right]^{\frac{\kappa_0}{2+\kappa_0}} c'\left\{\frac{1}{n}\sum_{t=1}^{n}\tilde{z}_t\tilde{z}_t'\right\}c \\
&\quad = C_0\left[\mathbf{1}'\times K_n+o_p(1)\right]^{\frac{\kappa_0}{2+\kappa_0}} c'\left\{\frac{1}{n}\sum_{t=1}^{n}z_t z_t'+o_p(1)\right\}c, \qquad \text{(A.8)}
\end{aligned}$$

where the last equality follows directly from Lemma A.1, and K_n is defined in the same way as $\tilde{K}_n$ with ε_t replacing $\tilde{\varepsilon}_t$. Note that $K_n = o_p(1)$ and $n^{-1}\sum_{t=1}^{n} z_t z_t' = O_p(1)$ by Assumption 2.2. Thus, by (A.8), it follows that

$$\sum_{t=1}^{n} E^*\left[c'\tilde{z}_{tn}^*\tilde{z}_{tn}^{*\prime}cI(|c'\tilde{z}_{tn}^*|>\eta)\right]\to 0 \ \text{ in probability as } n\to\infty. \qquad \text{(A.9)}$$

Finally, by (A.5), (A.6) and (A.9), the Cramér-Wold device and central limit theorem in Pollard (1984, Theorem VIII.1) yield that conditional on χ_n,

$$\sum_{t=1}^{n}\tilde{z}_{tn}^* \to_d N(0,\mathcal{O}) \quad \text{in probability as } n\to\infty. \qquad \text{(A.10)}$$

Since $\mathcal{V}\mathcal{O}\mathcal{V}'=\Sigma$, the conclusion follows from (A.4) and (A.10). □

Proof of Theorem 6.1. The proof is directly from the one for Theorem 2.2, except re-proving (A.6) under the conditions of Theorem 6.1.

By Lemma A.1, it is straightforward to see that

$$\begin{aligned}
\sum_{t=1}^{n} E^*\left[c'\tilde{z}_{tn}^*\tilde{z}_{tn}^{*\prime}c\right] &= c'\left\{\frac{1}{n}\sum_{s=1}^{L_n}\left[\sum_{t,t'\in B_s}\tilde{z}_t\tilde{z}_{t'}'\right]\right\}c \\
&= c'\left\{\frac{1}{n}\sum_{s=1}^{L_n}\left[\sum_{t,t'\in B_s}z_t z_{t'}'\right]\right\}c+o_p(1) \\
&\triangleq c'Z_n c+o_p(1).
\end{aligned}$$

Clearly, $EZ_n\to\mathcal{O}$ as $n\to\infty$. Thus, (A.6) holds if we can show that

$$var(Z_n)\to 0 \text{ as } n\to\infty. \qquad \text{(A.11)}$$

Note that $Z_n\in\mathcal{R}^{\tau\times\tau}$, where $\tau=m+p+q$. For simplicity, we only prove (A.11) for $Z_n^{(\tau,\tau)}$, the last entry of Z_n, where

$$Z_n^{(\tau,\tau)}=\frac{1}{n}\sum_{s=1}^{L_n}\left[\sum_{t,t'\in B_s}\varepsilon_t\varepsilon_{t'}\frac{\partial\varepsilon_t}{\partial\psi_q}\frac{\partial\varepsilon_{t'}}{\partial\psi_q}\right].$$

By a direct calculation, we have

$$\begin{aligned} var[Z_n^{(\tau,\tau)}] &= \frac{1}{n^2}\sum_{s,s'=1}^{L_n}\sum_{t_1,t_2\in B_s}\sum_{t_1',t_2'\in B_{s'}} E\left[\varepsilon_{t_1}\varepsilon_{t_2}\varepsilon_{t_1'}\varepsilon_{t_2'}\frac{\partial\varepsilon_{t_1}}{\partial\psi_q}\frac{\partial\varepsilon_{t_2}}{\partial\psi_q}\frac{\partial\varepsilon_{t_1'}}{\partial\psi_q}\frac{\partial\varepsilon_{t_2'}}{\partial\psi_q}\right] \\ &\triangleq \frac{1}{n^2}\sum_{s,s'=1}^{L_n}\xi(s,s'). \end{aligned}$$

Rewrite

$$var[Z_n^{(\tau,\tau)}] = \frac{1}{n^2}\sum_{s=1}^{L_n}\sum_{|s'-s|\le 1}\xi(s,s') + \frac{1}{n^2}\sum_{s=1}^{L_n}\sum_{|s'-s|>1}\xi(s,s'). \tag{A.12}$$

Fort the first summand in (A.12), since $b_n = o(n^{1/3})$, it is straightforward to see that

$$\frac{1}{n^2}\sum_{s=1}^{L_n}\sum_{|s'-s|\le 1}\xi(s,s') = O\left(\frac{L_n b_n^4}{n^2}\right) = O\left(\frac{b_n^3}{n}\right) = o(1). \tag{A.13}$$

For the second summand in (A.12), since

$$\varepsilon_t = \sum_{i=1}^{\infty} c_i y_{t-i} \text{ and } \frac{\partial\varepsilon_t}{\partial\psi_q} = \sum_{k=1}^{\infty} d_k y_{t-k}$$

by Lemma 1 in Francq and Zakoïan (1998), we have

$$\begin{aligned} &\left|E\left[\varepsilon_{t_1}\varepsilon_{t_2}\varepsilon_{t_1'}\varepsilon_{t_2'}\frac{\partial\varepsilon_{t_1}}{\partial\psi_q}\frac{\partial\varepsilon_{t_2}}{\partial\psi_q}\frac{\partial\varepsilon_{t_1'}}{\partial\psi_q}\frac{\partial\varepsilon_{t_2'}}{\partial\psi_q}\right]\right| \\ =& \left|\sum_{i_1,k_1,i_2,k_2,i_1',k_1',i_2',k_2'} c_{i_1}d_{k_1}c_{i_2}d_{k_2}c_{i_1'}d_{k_1'}c_{i_2'}d_{k_2'}M(i_1,k_1,i_2,k_2,i_1',k_1',i_2',k_2')\right| \\ \le& \left[\sum_{i_1>b_n/4}+\sum_{k_1>b_n/4}+\sum_{i_2>b_n/4}+\sum_{k_2>b_n/4}+\sum_{i_1'>b_n/4}+\sum_{k_1'>b_n/4}+\sum_{i_2'>b_n/4}+\sum_{k_2'>b_n/4}\right. \\ &\left.\sum_{i_1,k_1,i_2,k_2,i_1',k_1',i_2',k_2'\le b_n/4}\right]\left|c_{i_1}d_{k_1}c_{i_2}d_{k_2}c_{i_1'}d_{k_1'}c_{i_2'}d_{k_2'}M(i_1,k_1,i_2,k_2,i_1',k_1',i_2',k_2')\right| \\ =:& \sum_{i=1}^{9} g_i, \end{aligned}$$

where $M(i_1,k_1,i_2,k_2,i_1',k_1',i_2',k_2') = E\left|y_{t_1-i_1}y_{t_1-k_1}y_{t_2-i_2}y_{t_2-k_2}y_{t_1'-i_1'}y_{t_1'-k_1'}y_{t_2'-i_2'}y_{t_2'-k_2'}\right|$. By Cauchy-Schwarz inequality, we can show that

$$\begin{aligned} &M(i_1,k_1,i_2,k_2,i_1',k_1',i_2',k_2') \\ \le& \sqrt{E\left(y_{t_1-i_1}y_{t_1-k_1}y_{t_2-i_2}y_{t_2-k_2}\right)^2 E\left(y_{t_1'-i_1'}y_{t_1'-k_1'}y_{t_2'-i_2'}y_{t_2'-k_2'}\right)^2} \le Ey_t^8 < \infty. \end{aligned}$$

Since $c_i = O(\rho^i)$ and $d_i = O(\rho^i)$ for some $\rho \in (0,1)$, it is straightforward to see that

$$g_i \le C_0\rho^{b_n/4}, \text{ for } 1 \le i \le 8.$$

Furthermore, the Davydov inequality in Davydov (1968) implies that

$$
\begin{aligned}
g_9 \le C_0 & \sum_{i_1,k_1,i_2,k_2,i_1',k_1',i_2',k_2' \le b_n/4} \|y_{t_1-i_1}y_{t_1-k_1}y_{t_2-i_2}y_{t_2-k_2}\|_{2+\nu}\|y_{t_1'-i_1'}y_{t_1'-k_1'}y_{t_2'-i_2'}y_{t_2'-k_2'}\|_{2+\nu} \\
& \times \left[\alpha_y\left(\left\lfloor \frac{b_n}{2} \right\rfloor\right)\right]^{\nu/(2+\nu)} \left|c_{i_1}d_{k_1}c_{i_2}d_{k_2}c_{i_1'}d_{k_1'}c_{i_2'}d_{k_2'}\right| \\
& \le C_0 \left(E|y_t|^{8+4\nu}\right) \left[\alpha_y\left(\left\lfloor \frac{b_n}{2} \right\rfloor\right)\right]^{\nu/(2+\nu)} \\
& \times \sum_{i_1,k_1,i_2,k_2,i_1',k_1',i_2',k_2' \le b_n/4} \left|c_{i_1}d_{k_1}c_{i_2}d_{k_2}c_{i_1'}d_{k_1'}c_{i_2'}d_{k_2'}\right| \\
& \le C_0 \left[\alpha_y\left(\left\lfloor \frac{b_n}{2} \right\rfloor\right)\right]^{\nu/(2+\nu)}.
\end{aligned}
$$

Therefore, since $\lim_{k\to\infty} k^2[\alpha_y(k)]^{\nu/(2+\nu)} = 0$, it follows that

$$
\text{(A.14)} \qquad \frac{1}{n^2}\sum_{s=1}^{L_n}\sum_{|s'-s|>1} \xi(s,s') \le O\left(\frac{L_n^2 b_n^4}{n^2}\right)\left[\rho^{b_n/4} + \left[\alpha_y\left(\left\lfloor \frac{b_n}{2} \right\rfloor\right)\right]^{\nu/(2+\nu)}\right] = o(1).
$$

By (A.12)-(A.14), we know that (A.11) holds. This completes the proof of (A.6). □

REFERENCES

[1] Davydov, Y.A. (1968) Convergence of distributions generated by stationary stochastic processes. *Theory of Probability & Its Applications* **13**, 691-696.

[2] Francq, C., Roy, R., and Zakoïan, J.M. (2005) Diagostic checking in ARMA models with uncorrelated errors. *Journal of the American Statistical Association* **100**, 532-544.

[3] Francq, C. and Zakoïan, J.M. (1998) Estimating linear representations of nonlinear processes. *Journal of Statistical Planning and Inference* **68**, 145-165.

[4] Ling, S. (2007) Self-weighted and local quasi-maximum likelihood estimators for ARMA-GARCH/IGARCH models. *Journal of Econometrics* **140**, 849-873.

[5] Ling, S. and McAleer, M. (2003) Asymptotic theory for a new vector ARMA-GARCH model. *Econometric Theory* **19**, 280-310.

[6] Pollard, D. (1984) Convergence of stochastic processes. Springer, Berlin.

Chinese Academy of Sciences
Institute of Applied Mathematics
Haidian District, Zhongguancun
Beijing, China
E-mail: kzhu@amss.ac.cn