A mixed portmanteau test for ARMA-GARCH model by the quasi-maximum exponential likelihood estimation approach

Zhu, Ke

31 July 2012

A MIXED PORTMANTEAU TEST FOR ARMA-GARCH MODEL BY THE QUASI-MAXIMUM EXPONENTIAL LIKELIHOOD ESTIMATION APPROACH

By Ke Zhu*

Chinese Academy of Sciences

This paper investigates the joint limiting distribution of the residual autocorrelation functions and the absolute residual autocorrelation functions of ARMA-GARCH model. This leads a mixed portmanteau test for diagnostic checking of the ARMA-GARCH model fitted by using the quasi-maximum exponential likelihood estimation approach in Zhu and Ling (2011). Simulation studies are carried out to examine our asymptotic theory, and assess the performance of this mixed test and other two portmanteau tests in Li and Li (2008). A real example is given.

1. Introduction. Following the seminal work of Engle (1982) and Bollerslev (1986), the following ARMA-GARCH model has been widely used in economics and finance:

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

$$\varepsilon_t = \eta_t \sqrt{h_t} \quad \text{and} \quad h_t = \alpha_0 + \sum_{i=1}^{r} \alpha_i \varepsilon_{t-i}^2 + \sum_{i=1}^{s} \beta_i h_{t-i}, \tag{1.2}$$

where $\alpha_0 > 0, \alpha_i \geq 0\,(i = 1, \cdots, r), \beta_j \geq 0\,(j = 1, \cdots, s)$, and η_t is a sequence of i.i.d random variables. A fundamental problem for practitioners is to check the adequacy of model (1.1)-(1.2). The portmanteau test initially introduced by Box and Pierce (1970) is for testing the i.i.d. assumption of η_t, and has become a popular tool for diagnostic checking of model (1.1)-(1.2) usually fitted by using the Gaussian quasi-maximum likelihood estimator (QMLE) approach. For a discussion on the Gaussian QMLE of model (1.1)-(1.2), we refer to Ling and Li (1997a), Francq and Zakoïan (2004) and Ling (2007). So far there has been lots of portmanteau tests for diagnostic checking of this QMLE-type fitted model. For model (1.1) with i.i.d. errors ε_t, Ljung and Box (1978) modified the method in Box and Pierce (1970) to construct a portmanteau test which has a good performance for small sample

*Supported by the National Center for Mathematics and Interdisciplinary Sciences, CAS.

Keywords and phrases: ARMA-GARCH model, LAD estimator, mixed portmanteau test, model diagnostics, quasi-maximum exponential likelihood estimator.

case; McLeod (1978) obtained the limiting distribution of residual autocorrelations by using the martingale approach; McLeod and Li (1983) further used the square-residual autocorrelations for model checking; and many others. For model (1.1)-(1.2), the earliest work goes to Li and Mak (1994) and Ling and Li (1997b); Wong and Li (2000) studied the portmanteau test for the multivariate conditional heteroscedastic models; Wong and Ling (2005) constructed a mixed portmanteau test applied for model (1.1)-(1.2); see also Francq et al. (2005), Carbon and Francq (2011), Ling and Tong (2011) and Shao (2011) on more tests for diagnostic checking of many ARMA-type or GARCH-type models fitted by the Gaussian QMLE approach.

Although the Gaussian QMLE-type portmanteau test has achieved a great success, a necessary set-up for them is that $E\varepsilon_t^4 < \infty$ and $E\eta_t^4 < \infty$. In practice, this setup may fail, and hence the Gaussian QMLE-type portmanteau test may not be reliable in this case. Recently, the quasi-maximum exponential likelihood estimator (QMELE) of model (1.1)-(1.2) as one of the least absolute deviation (LAD) estimators has been well studied in Li and Li (2008) and Zhu and Ling (2011); see also Fan et al. (2010) and Francq et al. (2011) for other non-Gaussian QML estimations. The QMELE not only requires a weaker moment condition of ε_t and η_t and is also more robust than the Gaussian QMLE. Note that a weaker moment condition on ε_t will give us a larger permissible parameter of model (1.2). Therefore, it is worthwhile to construct a portmanteau test for model (1.1)-(1.2) fitted by the QMELE method especially when $E\eta_t^4 = \infty$. Li and Li (2005) first accomplished it for model (1.2) by using the LAD estimation method in Peng and Yao (2003) (see also Berkes et al. (2003) and Francq and Zakoïan (2010)), and they further investigated two other portmanteau tests Q_r and Q_a for ARFIMA-GARCH model in Li and Li (2008) in the context of QMELE approach.

In this paper, under the condition that $E\varepsilon_t^2 < \infty$ and $E\eta_t^2 < \infty$, we first obtain the joint limiting distribution of the residual autocorrelation functions and the absolute residual autocorrelation functions, where the residual is obtained from model (1.1)-(1.2) fitted by the QMELE approach. Based on this, we propose a mixed portmanteau test statistic for model (1.1)-(1.2). Via some simplifications, we can show that our mixed portmanteau test nests two portmanteau tests Q_r and Q_a in Li and Li (2008), and hence it is useful when the fitted model has disparity in both the conditional mean and the conditional variance; see also Wong and Ling (2005). Simulation studies are carried out to examine our asymptotic theory, and assess the performance of our mixed test and two portmanteau tests Q_r and Q_a in Li and Li (2008). A real example is given as well.

This paper is organized as follows. Section 2 derives our main results and hence the mixed portmanteau test. Section 3 reports the simulation results. A real example is provided in Section 4. The proofs are presented in the Appendix. Throughout the paper, some symbols are conventional. A' is the transpose of matrix A. $o_p(1)$ ($O_p(1)$) denotes a sequence of random numbers converging to zero (bounded) in probability. $\to_d$ denotes convergence in distribution. $I(\cdot)$ is an indicator function.

2. Main results. Let $\theta = (\gamma', \delta')'$ be the unknown parameter of model (1.1)-(1.2) and its true value be θ_0, where $\gamma = (\mu, \phi_1, ..., \phi_p, \psi_1, ..., \psi_q)'$ and $\delta = (\alpha_0, ..., \alpha_r, \beta_1, ..., \beta_s)'$. Given the observations $\{y_n, ..., y_1\}$ and the initial values $Y_0 \equiv \{y_0, y_{-1}, ...\}$, we can rewrite the parametric model (1.1)-(1.2) as

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

$$\eta_t(\theta) = \varepsilon_t(\gamma)/\sqrt{h_t(\theta)} \text{ and } h_t(\theta) = \alpha_0 + \sum_{i=1}^{r} \alpha_i \varepsilon_{t-i}^2(\gamma) + \sum_{i=1}^{s} \beta_i h_{t-i}(\theta).$$

Here, $\eta_t(\theta_0) = \eta_t$, $\varepsilon_t(\gamma_0) = \varepsilon_t$ and $h_t(\theta_0) = h_t$. The parameter space is $\Theta = \Theta_\gamma \times \Theta_\delta$, where $\Theta_\gamma \subset R^{p+q+1}$, $\Theta_\delta \subset R_0^{r+s+1}$, $R = (-\infty, \infty)$ and $R_0 = [0, \infty)$. Assume that Θ_γ and Θ_δ are compact and θ_0 is an interior point in Θ. Denote $\alpha(z) = \sum_{i=1}^{r} \alpha_i z^i$, $\beta(z) = 1 - \sum_{i=1}^{s} \beta_i z^i$, $\phi(z) = 1 - \sum_{i=1}^{p} \phi_i z^i$ and $\psi(z) = 1 + \sum_{i=1}^{q} \psi_i z^i$. We introduce the following basic assumptions:

ASSUMPTION 2.1. *For each $\theta \in \Theta$, $\phi(z) \neq 0$ and $\psi(z) \neq 0$ 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. *For each $\theta \in \Theta$, $\alpha(z)$ and $\beta(z)$ have no common root, $\alpha(1) \neq 1, \alpha_r + \beta_s \neq 0$, and $\sum_{i=1}^{s} \beta_i < 1$.*

Assumption 2.1 implies the stationarity, invertibility and identifiability of model (1.1), and Assumption 2.2 is the identifiability condition for model (1.2). Following Zhu and Ling (2011), the quasi-maximum exponential likelihood estimator (QMELE) of θ_0, denoted by $\hat{\theta}_n \triangleq (\hat{\gamma}_n', \hat{\delta}_n')'$, is defined as

$$\hat{\theta}_n = \arg\min_{\Theta} L_n(\theta), \;\; L_n(\theta) = \frac{1}{n} \sum_{t=1}^{n} \left[\log \sqrt{h_t(\theta)} + \frac{|\varepsilon_t(\gamma)|}{\sqrt{h_t(\theta)}} \right];$$

see also Li and Li (2008). The QMELE viewed as a LAD-type estimator is an efficient estimator when η_t follows a double exponential distribution. In order to obtain the asymptotic theory of $\hat{\theta}_n$, we need further three assumptions:

ASSUMPTION 2.3. *The median of η_t is equal to zero, $E|\eta_t| = 1$, $var(\eta_t) = \sigma_0^2 < \infty$ and the probability density function $f(x)$ of η_t satisfying $f(0) > 0$ and $\sup_{x \in R} f(x) < \infty$, is continuous at zero.*

ASSUMPTION 2.4. *ε_t is strictly stationary and ergodic with $E\varepsilon_t^2 < \infty$.*

ASSUMPTION 2.5. $\sqrt{n}(\hat{\theta}_n - \theta_0) = O_p(1)$.

Assumption 2.3 is a general set-up for the LAD-type estimator; see Peng and Yao (2003), Li and Li (2008) and Zhu and Ling (2011). As shown in Bollerslev (1986), the necessary and sufficient condition of Assumption 2.4 is

$$E\eta_t^2 \sum_{i=1}^{r} \alpha_i + \sum_{i=1}^{s} \beta_i < 1.$$

Following Li and Li (2008), we presume that $\hat{\theta}_n$ has a $\sqrt{n}$-convergence rate in Assumption 2.5 to simplify our proof. A sufficient condition of Assumption 2.5 is that $E|\varepsilon_t|^3 < \infty$ as shown in Zhu and Ling (2011). Denote the residual $\hat{\eta}_t \triangleq \eta_t(\hat{\theta}_n)$. Then, the lag-$l$ residual autocorrelation function can be defined as

$$\hat{\rho}_l^* = \frac{\sum_{t=l+1}^{n} (\hat{\eta}_t - \bar{\eta}_{0n})(\hat{\eta}_{t-l} - \bar{\eta}_{0n})}{\sum_{t=1}^{n} (\hat{\eta}_t - \bar{\eta}_{0n})^2},$$

where $\bar{\eta}_{0n} = n^{-1} \sum_{t=1}^{n} \hat{\eta}_t$. Note that $\hat{\theta}_n - \theta_0 = o_p(1)$ by Assumption 2.5. Under Assumptions 2.1-2.4, by Theorem 3.1 in Ling and McAleer (2003) and the dominated convergence theorem, we can show that

$$\bar{\eta}_{0n} = m_0 + o_p(1) \quad \text{and} \quad \frac{1}{n} \sum_{t=1}^{n} (\hat{\eta}_t - \bar{\eta}_{0n})^2 = \sigma_0^2 + o_p(1), \tag{2.1}$$

where $m_0 = E\eta_t$, and hence theoretically we only need to consider

$$\hat{\rho}_l = \frac{1}{n\sigma_0^2} \sum_{t=l+1}^{n} (\hat{\eta}_t - m_0)(\hat{\eta}_{t-l} - m_0).$$

Next, the lag-l absolute residual autocorrelation function can be defined as

$$\hat{r}_l^* = \frac{\sum_{t=l+1}^{n} (|\hat{\eta}_t| - \bar{\eta}_{1n})(|\hat{\eta}_{t-l}| - \bar{\eta}_{1n})}{\sum_{t=1}^{n} (|\hat{\eta}_t| - \bar{\eta}_{1n})^2},$$

where $\bar{\eta}_{1n} = n^{-1} \sum_{t=1}^{n} |\hat{\eta}_t|$. Under Assumptions 2.1-2.4, by a similar argument as for (2.1), we can show that

$$\bar{\eta}_{1n} = E|\eta_t| + o_p(1) = 1 + o_p(1) \quad \text{and} \quad \frac{1}{n} \sum_{t=1}^{n} (|\hat{\eta}_t| - \bar{\eta}_{1n})^2 = \sigma_1^2 + o_p(1),$$

where $\sigma_1^2 = var(|\eta_t|)$, and hence theoretically we only need to consider

$$\hat{r}_l = \frac{1}{n\sigma_1^2} \sum_{t=l+1}^{n} (|\hat{\eta}_t| - 1)(|\hat{\eta}_{t-l}| - 1).$$

Let ρ_l and r_l be $\hat{\rho}_l$ and $\hat{r}_l$, respectively, when $\hat{\eta}_t$ is replaced by η_t. Denote $\hat{\rho} = (\hat{\rho}_1, \cdots, \hat{\rho}_M)'$ and $\hat{r} = (\hat{r}_1, \cdots, \hat{r}_M)'$. Similarly define ρ and r. Furthermore, let $m_1 = E[\eta_t(|\eta_t| - 1)]$, and denote the matrixes

$$V = \begin{pmatrix} I_M & 0 & \sigma_0^{-2} X_\rho \Sigma^{-1} \\ 0 & I_M & \sigma_1^{-2} X_r \Sigma^{-1} \end{pmatrix}, \tag{2.2}$$

$$\Omega = \begin{pmatrix} I_M & m_1^2 I_M/(\sigma_0^2\sigma_1^2) & -X_\rho/(2\sigma_0^2) + m_1 X_\rho^*/(4\sigma_0^2) \\ * & I_M & -X_r/2 + m_0 X_r^*/\sigma_1^2 \\ * & * & W \end{pmatrix}, \tag{2.3}$$

where Ω is a symmetric block matrix, I_M is a $M \times M$ identity matrix, $X_\rho = (X_{\rho 1}, \cdots, X_{\rho M})'$, $X_\rho^* = (X_{\rho 1}^*, \cdots, X_{\rho M}^*)'$, $X_r = (X_{r1}, \cdots, X_{rM})'$, $X_r^* = (X_{r1}^*, \cdots, X_{rM}^*)'$, $\Sigma = E\left[U_t(\theta_0) J_1 U_t'(\theta_0)\right]$ and $W = E\left[U_t(\theta_0) J_2 U_t'(\theta_0)\right]$ with $J_1 = diag\{f(0), 1/4\}$,

$$J_2 = \begin{pmatrix} 1/4 & -\sqrt{2}m_0/8 \\ -\sqrt{2}m_0/8 & \sigma_1^2/8 \end{pmatrix}, \quad U_t(\theta) = \left(\frac{1}{\sqrt{h_t(\theta)}} \frac{\partial \varepsilon_t(\gamma)}{\partial \theta}, \frac{1}{\sqrt{2}h_t(\theta)} \frac{\partial h_t(\theta)}{\partial \theta} \right)$$

and

$$X_{\rho l} = E\left[\frac{\eta_{t-l} - m_0}{\sqrt{h_t}} \frac{\partial \varepsilon_t(\gamma_0)}{\partial \theta} \right], \qquad X_{\rho l}^* = E\left[\frac{\eta_{t-l} - m_0}{h_t} \frac{\partial h_t(\theta_0)}{\partial \theta} \right],$$

$$X_{rl} = E\left[\frac{1 - |\eta_{t-l}|}{2h_t} \frac{\partial h_t(\theta_0)}{\partial \theta} \right], \qquad X_{rl}^* = E\left[\frac{1 - |\eta_{t-l}|}{2\sqrt{h_t}} \frac{\partial \varepsilon_t(\gamma_0)}{\partial \theta} \right],$$

for any integer $l \geq 1$. We are now ready to give our main result on the joint limiting distribution of $\hat{\rho}$ and $\hat{r}$ in the following theorem:

THEOREM 2.1. *If Assumptions 2.1-2.5 hold, then*

$$\sqrt{n} \begin{pmatrix} \hat{\rho} \\ \hat{r} \end{pmatrix} \to_d N(0, V\Omega V') \quad \text{as } n \to \infty,$$

where V and Ω are defined as in (2.2)-(2.3), respectively.

PROOF. See the Appendix. □

In practice, the initial values Y_0 are unknown, and can be replaced by any constants. This will not affect our asymptotic result in Theorem 2.1; see Ling (2007) and Zhu (2011). Unless stated otherwise, we set the initial values $Y_0 \equiv 0$. Given the observations $\{y_n, \cdots, y_1\}$, we can estimate the matrixes V and Ω by their sample means V_n and Ω_n, respectively. Here, we estimate $f(0)$ by the normal kernel estimator with bandwidth chosen as in Fan and Yao (2003, page 201). Under Assumptions 2.1-2.4, by a similar argument as for (2.1), we can show that $\hat{V}_n = V + o_p(1)$ and $\hat{\Omega}_n = \Omega + o_p(1)$. Thus, from Theorem 2.1, the following corollary is straightforward.

Corollary 2.1. *If Assumptions 2.1-2.5 hold, then*

$$Q_m(M) \triangleq n \begin{pmatrix} \hat{\rho} \\ \hat{r} \end{pmatrix}' (\hat{V}_n \hat{\Omega}_n \hat{V}_n')^{-1} \begin{pmatrix} \hat{\rho} \\ \hat{r} \end{pmatrix} \to_d \chi^2(2M) \quad \text{as } n \to \infty.$$

We call Q_m in Corollary 2.1 as the mixed portmanteau test statistic, and use it for diagnostic checking of the ARMA-GARCH model fitted by the QMELE approach. We now offer some reduced form of Q_m. By a direct calculation, from (2.2)-(2.3), we have

$$V\Omega V' = \begin{pmatrix} S_1 & S_2 \\ S_2' & S_3 \end{pmatrix},$$

where

$$S_1 = I_M - \frac{1}{\sigma_0^4} X_\rho \Sigma^{-1} X_\rho' + \frac{m_1}{4\sigma_0^4}\left(X_\rho \Sigma^{-1} X_\rho^{*'} + X_\rho^* \Sigma^{-1} X_\rho'\right) + \frac{1}{\sigma_0^4} X_\rho \Sigma^{-1} W \Sigma^{-1} X_\rho',$$

$$S_2 = \frac{m_1^2}{\sigma_0^2\sigma_1^2} I_M + \frac{m_0}{\sigma_0^2\sigma_1^2} X_\rho \Sigma^{-1} X_r^{*'} + \frac{m_1}{4\sigma_0^2\sigma_1^2} X_\rho^* \Sigma^{-1} X_r' - \left(\frac{1}{2\sigma_0^2} + \frac{m_1}{2\sigma_0^2\sigma_1^2}\right) X_\rho \Sigma^{-1} X_r' + \frac{1}{\sigma_0^2\sigma_1^2} X_\rho \Sigma^{-1} W \Sigma^{-1} X_r',$$

$$S_3 = I_M - \frac{1}{\sigma_1^2} X_r \Sigma^{-1} X_r' + \frac{m_0}{\sigma_1^4}\left(X_r \Sigma^{-1} X_r^{*'} + X_r^* \Sigma^{-1} X_r'\right) + \frac{1}{\sigma_1^4} X_r \Sigma^{-1} W \Sigma^{-1} X_r'.$$

Particularly, if $\eta_t \sim Laplace(0,1)$, we have $m_0 = m_1 = 0$ and $W = \Sigma/2$. In this case, the limiting distributions of $\hat{\rho}$ and $\hat{r}$ are independent, and

$$Q_m = n\hat{\rho}\left[I_M - \frac{1}{2}\hat{X}_{\rho n}\hat{\Sigma}_n^{-1}\hat{X}_{\rho n}'\right]^{-1}\hat{\rho}' + n\hat{r}\left[I_M + \frac{1}{2}\hat{X}_{rn}\hat{\Sigma}_n^{-1}\hat{X}_{rn}'\right]^{-1}\hat{r}'.$$

However, the limiting distributions of $\hat{\rho}$ and $\hat{r}$ are not independent in general, and only holds for some special cases. For instance, when $X_r \approx 0$, $m_0 \approx 0$ and $m_1 \approx 0$, Q_m can be approximated by

$$Q_{1m} \approx n\hat{\rho}\left[I_M - \frac{1}{\hat{\sigma}_{0n}^4}\hat{X}_{\rho n}\hat{\Sigma}_n^{-1}\hat{X}_{\rho n}' + \frac{1}{\hat{\sigma}_{0n}^4}\hat{X}_{\rho n}\hat{\Sigma}_n^{-1}\hat{W}_n\hat{\Sigma}_n^{-1}\hat{X}_{\rho n}'\right]^{-1}\hat{\rho}' + n\sum_{l=1}^{M}\hat{r}_l^2,$$

which nests the test statistic Q_r in Li and Li (2008) as the first term; when $X_\rho \approx 0$, $m_0 \approx 0$ and $m_1 \approx 0$, Q_m can be approximated by

$$Q_{2m} \approx n\sum_{l=1}^{M}\hat{\rho}_l^2 + n\hat{r}\left[I_M - \frac{1}{\hat{\sigma}_{1n}^2}\hat{X}_{rn}\hat{\Sigma}_n^{-1}\hat{X}_{rn}' + \frac{1}{\hat{\sigma}_{1n}^4}\hat{X}_{rn}\hat{\Sigma}_n^{-1}\hat{W}_n\hat{\Sigma}_n^{-1}\hat{X}_{rn}'\right]^{-1}\hat{r}',$$

which nests the test statistic Q_a in Li and Li (2008) as the second term. Specifically, for the pure ARMA model with a symmetric η_t, we have $X_r = 0$, $m_0 = m_1 = 0$ and $W = [4f(0)]^{-1}\Sigma$, and hence Q_m can be simplified as

$$Q^*_{1m} = n\hat{\rho}\left[I_M - \frac{1}{\hat{\sigma}^4_{0n}}\left(1 - \frac{1}{4\hat{f}(0)}\right)\hat{X}_{\rho n}\hat{\Sigma}^{-1}_n\hat{X}'_{\rho n}\right]^{-1}\hat{\rho}' + n\sum_{l=1}^{M}\hat{r}^2_l;$$

for the pure GARCH model with a symmetric η_t, we have $X_\rho = 0$, $m_0 = m_1 = 0$ and $W = \sigma_1^2\Sigma/2$, and hence Q_m can be simplified as

$$Q^*_{2m} = n\sum_{l=1}^{M}\hat{\rho}^2_l + n\hat{r}\left[I_M - \frac{1}{2\hat{\sigma}^2_{1n}}\hat{X}_{rn}\hat{\Sigma}^{-1}_n\hat{X}'_{rn}\right]^{-1}\hat{r}'.$$

The test statistic Q^*_{1m} is useful for diagnostic checking of ARMA model fitted by the LAD estimation method, while Q^*_{2m} is useful for diagnostic checking of GARCH model fitted by the QMELE method, especially if there is a further evidence to show that $E\eta_t^4 = \infty$.

3. Simulation. In this section, we first examine the asymptotic result in Theorem 2.1. We generate 1000 replications of sample size $n = 200$ and 400 from model (3.1) and fit each replication by using the QMELE method:

$$\left\{\begin{array}{l} y_t = 0.2y_{t-1} + \varepsilon_t, \quad \varepsilon_t = \eta_t\sqrt{h_t}, \\ h_t = 0.01 + 0.2\varepsilon^2_{t-1} + 0.2h_{t-1}, \end{array}\right. \tag{3.1}$$

where η_t is chosen to be re-scaled $N(0,1)$, $Laplace(0,1)$, t_3 and $SN(0,0,1)$, respectively, such that it satisfies $median(\eta_t) = 0$ and $E|\eta_t| = 1$. Here, $SN(\xi,\omega,\alpha)$ stands for the skew normal distribution, which has the probability density function

$$\frac{1}{\omega\pi}e^{-\frac{(x-\xi)^2}{2\omega^2}}\int_{-\infty}^{\alpha\left(\frac{x-\xi}{\omega}\right)} e^{-\frac{t^2}{2}}dt;$$

see Azzalini (1985). The asymptotic standard deviations of the residual autocorrelations $\hat{\rho}$, and the absolute residual autocorrelations $\hat{r}$, are calculated according to Theorem 2.1 with $M = 6$. Table 1 lists the sample standard deviations (SD) and the average estimated asymptotic standard deviations (AD) of $\hat{\rho}$ and $\hat{r}$, for lags 1, 2, 3 and 6. From Table 1, we can see that all pairs of AD and SD are close to each other for n as small as 200. As n increases from 200 to 400, all of the SDs and ADs become smaller.

Next, we compare the finite sample performance of our mixed test Q_m with those of two portmanteau tests Q_r and Q_a in Li and Li (2008). We choose our null model

Table 1
SDs and ADs (×10) for model (3.1)

			Lags							
			Residual				Absolute residual			
η_t	n		1	2	3	6	1	2	3	6
Normal	200	SD	0.525	0.700	0.684	0.685	0.218	0.539	0.634	0.694
		AD	0.582	0.696	0.698	0.700	0.209	0.483	0.635	0.688
	400	SD	0.390	0.517	0.498	0.501	0.140	0.354	0.444	0.496
		AD	0.400	0.495	0.497	0.495	0.134	0.347	0.445	0.489
Laplace	200	SD	0.552	0.690	0.691	0.688	0.266	0.536	0.615	0.663
		AD	0.545	0.694	0.689	0.686	0.246	0.498	0.620	0.668
	400	SD	0.388	0.495	0.497	0.491	0.186	0.398	0.444	0.489
		AD	0.389	0.491	0.493	0.490	0.171	0.366	0.440	0.483
t_3	200	SD	0.568	0.701	0.659	0.659	0.312	0.553	0.603	0.633
		AD	0.559	0.676	0.663	0.660	0.279	0.479	0.582	0.621
	400	SD	0.407	0.478	0.499	0.472	0.215	0.380	0.436	0.459
		AD	0.394	0.479	0.481	0.470	0.198	0.350	0.419	0.448
SN	200	SD	0.544	0.672	0.684	0.665	0.259	0.569	0.629	0.676
		AD	0.565	0.700	0.706	0.706	0.221	0.479	0.644	0.698
	400	SD	0.381	0.488	0.491	0.483	0.150	0.379	0.444	0.488
		AD	0.394	0.495	0.499	0.500	0.144	0.347	0.450	0.495

as

$$\begin{cases} y_t = 0.2y_{t-1} + \varepsilon_t, \quad \varepsilon_t = \eta_t\sqrt{h_t}, \\ h_t = 0.01 + \alpha\varepsilon_{t-1}^2, \end{cases} \tag{3.2}$$

and use the following three alternative models to study the powers for all tests:

$$\begin{cases} y_t = 0.2y_{t-1} + 0.3y_{t-2} + \varepsilon_t, \quad \varepsilon_t = \eta_t\sqrt{h_t}, \\ h_t = 0.01 + \alpha\varepsilon_{t-1}^2, \end{cases} \tag{3.3}$$

$$\begin{cases} y_t = 0.2y_{t-1} + \varepsilon_t, \quad \varepsilon_t = \eta_t\sqrt{h_t}, \\ h_t = 0.01 + \alpha\varepsilon_{t-1}^2 + 0.3\varepsilon_{t-2}^2, \end{cases} \tag{3.4}$$

$$\begin{cases} y_t = 0.2y_{t-1} + 0.3y_{t-2} + \varepsilon_t, \quad \varepsilon_t = \eta_t\sqrt{h_t}, \\ h_t = 0.01 + \alpha\varepsilon_{t-1}^2 + 0.3\varepsilon_{t-2}^2, \end{cases} \tag{3.5}$$

where η_t is chosen to be re-scaled $N(0,1)$, $Laplace(0,1)$, t_3 and $SN(0,0,1)$ respectively. In order to make sure that $(E\eta_t^2)\alpha \approx 1$, we take $\alpha = 0.6, 0.48, 0.38$ and 0.6 for these four re-scaled distributions, respectively. Based on these choices of α, we generate 1000 replications of sample size $n = 200$ and 400 from each model and fit each replication by an AR(1)-ARCH(1) model with the QMELE method. The significance level $\underline{\alpha} = 0.05$ and $M = 6$. The empirical power and sizes of these tests are reported in Table 2. Their sizes correspond to the results for model (3.2).

From Table 2, it is clear that the sizes of these tests are close to their nominal ones, and the power of them is generally as expected. First, all the powers become large as

n increases. Second, Q_r and Q_a only have the power to detect the mis-specification in the conditional mean and the conditional variance, respectively, while Q_m is feasible to detect both of them, and hence it is more robust. Third, when the mis-specification happens only in the conditional mean (or the conditional variance), Q_m is less powerful than Q_r (or Q_a). Forth, all the conclusions are invariant regardless of the distribution of η_t. Overall, our mixed test Q_m can give us a good indication in diagnostic checking of ARMA-GARCH model fitted by the QMELE method, especially when the sample size is large.

TABLE 2
Empirical size and power ($\times 100$) *for* Q_m, Q_r *and* Q_a

		Models											
			(3.2)			(3.3)			(3.4)			(3.5)	
η_t	n	Q_m	Q_r	Q_a	Q_m	Q_r	Q_a	Q_m	Q_r	Q_a	Q_m	Q_r	Q_a
Normal	200	5.0	4.5	4.1	48.7	69.5	4.8	26.7	6.7	46.2	58.9	43.4	48.6
	400	4.6	4.6	4.7	93.3	97.5	12.7	76.4	8.8	90.8	97.7	79.5	93.0
Laplace	200	4.2	3.3	4.8	50.9	65.2	6.6	18.1	5.8	28.2	49.3	42.8	34.4
	400	5.8	5.1	4.5	92.2	95.0	11.7	54.6	11.5	73.9	93.7	75.0	77.7
t_3	200	4.5	4.7	4.1	53.9	62.5	5.5	15.5	4.3	27.1	45.3	40.7	28.1
	400	5.7	4.3	5.6	88.3	92.1	10.3	48.4	6.9	63.8	88.6	71.7	68.6
SN	200	4.8	4.1	5.7	48.2	68.0	6.1	24.1	5.9	44.1	57.1	44.0	46.0
	400	5.0	4.7	5.5	93.4	97.3	14.6	75.8	11.3	89.6	96.5	78.8	90.1

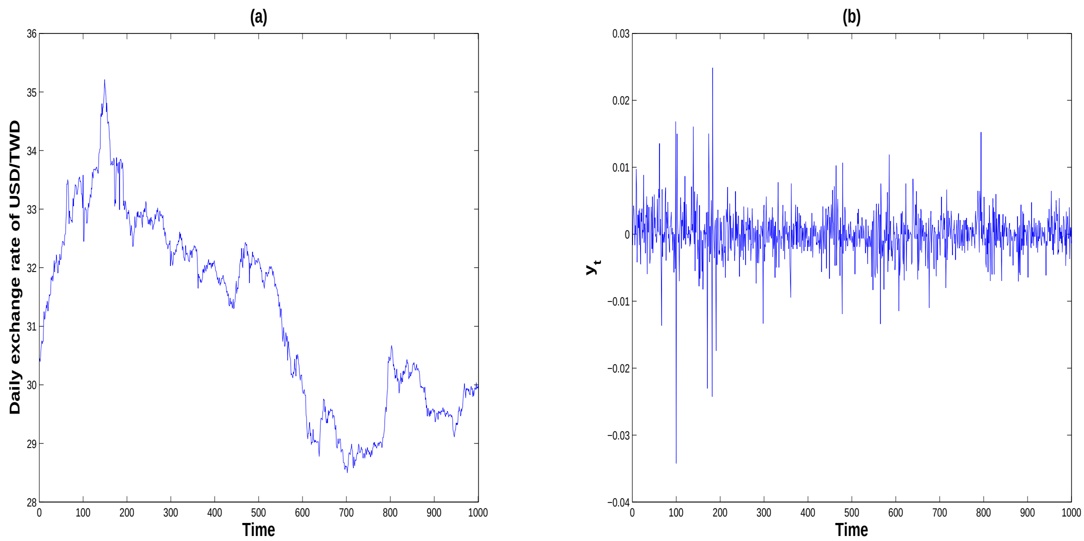

FIG 1. *(a) the daily exchange rate of USD/TWD from July 25, 2008 to July 20, 2012 and (b) its log return.*

4. A real example. In this section, we study the daily exchange rate of United States Dollars (USD) to New Taiwan Dollars (TWD) from July 25, 2008 to July 20, 2012, which has in total 1001 observations; see Figure 1 (a). Its log return, denoted by $\{y_t\}_{t=1}^{1000}$ is plotted in Figure 1 (b). Next, we fit $\{y_t\}$ by the following four different models:

AR(1)-ARCH(1) (model A), AR(1)-GARCH(1,1) (model B),
ARMA(1,2)-GARCH(1,1) (model C), ARMA(2,2)-GARCH(1,2) (model D).

Models A-D are estimated by using the QMELE procedure. The estimators $\hat{\theta}_n$ with their standard deviations (sd) are given in Table 3. To check the adequacy of these models, Table 3 reports the values of $Q_m(M), Q_r(M)$ and $Q_a(M)$ with $M = 6$ and 12 for each fitted model.

TABLE 3
Results for all fitted model

	Models							
	Model A		Model B		Model C		Model D	
Parameters	$\hat{\theta}_n$	sd	$\hat{\theta}_n$	sd	$\hat{\theta}_n$	sd	$\hat{\theta}_n$	sd
ϕ_1	-0.0998	0.0346	-0.0675	0.0321	-0.9383	0.0470	-1.0094	0.2940
ϕ_2							-0.6437	0.2051
ψ_1					0.8468	0.0555	0.9512	0.3024
ψ_2					-0.1011	0.0321	0.6123	0.2078
α_0 [a]	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
α_1	0.1048	0.0375	0.0659	0.0267	0.0588	0.0199	0.0655	0.0322
α_2								
β_1			0.7105	0.1058	0.8084	0.0551	0.2270	0.3312
β_2							0.5425	0.2991
$(Q_m(6), Q_m(12))$ [b]	(33.19, 43.45)		(25.72, 34.68)		(25.48, 32.38)		(18.48, 27.08)	
$(Q_a(6), Q_a(12))$ [c]	(10.35, 18.85)		(11.36, 17.43)		(11.37, 16.54)		(8.63, 14.96)	
$(Q_r(6), Q_r(12))$ [c]	(21.42, 26.81)		(14.22, 17.96)		(11.64, 14.81)		(8.62, 11.50)	

[a] The estimator $\hat{\alpha}_{0n}$ and its sd are less than 10^{-4} for each model.
[b] The 95% upper percentages for $Q_m(6)$ and $Q_m(12)$ are 21.03 and 36.42, respectively.
[c] The 95% upper percentages for $Q_a(6)$ (or $Q_r(6)$) and $Q_a(12)$ (or $Q_r(12)$) are 12.59 and 21.03, respectively.

From Table 3, we first find that model A is not adequate to fit $\{y_t\}$, and this can be detected by Q_m and Q_r but not Q_a. Second, Q_m indicates that model B and C are not adequate, but Q_r and Q_a can not detect them. Third, all of tests indicate that model D is adequate to fit $\{y_t\}$. Moreover, the estimators of $E\eta_t^2(\alpha_1 + \alpha_2) + (\beta_1 + \beta_2)$, $skewness(\eta_t)$ and $Kurtosis(\eta_t)$ in model D are 0.91, -0.48 and 9.38, respectively. This implies that $E\varepsilon_t^2 < \infty$ and η_t is skewed and heavy-tailed. Hence, an ARMA(2,2)-GARCH(1,2) model, based on the QMELE procedure, should be adequate for us to fit $\{y_t\}$.

APPENDIX

In this Appendix, we give the proof of Theorem 2.1, which relies on a key lemma below.

LEMMA A.1. *Suppose that Assumptions 2.1-2.4 hold. Then, the matrix $V\Omega V'$ is positive definite, where V and Ω are defined as in (2.2) and (2.3), respectively.*

PROOF. It suffices to prove that Ω is positive definite. Let

$$D_t = -\frac{sgn(\eta_t)}{2\sqrt{h_t}}\frac{\partial \varepsilon_t(\gamma_0)}{\partial \theta} + \frac{|\eta_t|-1}{4h_t}\frac{\partial h_t(\theta_0)}{\partial \theta}, \tag{A.1}$$

where $sgn(\eta_t) = I(\eta_t > 0) - I(\eta_t < 0)$. Then, we know that $\Omega = E[v_t v_t']$, where

$$v_t = (\tilde{\rho}_{1t}, \cdots, \tilde{\rho}_{Mt}, \tilde{r}_{1t}, \cdots, \tilde{r}_{Mt}, D_t')'$$

with

$$\tilde{\rho}_{it} = \frac{(\eta_t - m_0)(\eta_{t-i} - m_0)}{\sigma_0^2} \quad \text{and} \quad \tilde{r}_{it} = \frac{(|\eta_t|-1)(|\eta_{t-i}|-1)}{\sigma_1^2}$$

for $i = 1, \cdots, M$. Let $c = (c_\rho', c_r', c_D')' \triangleq (c_{\rho r}', c_D')'$ be a vector of constants. If $c'v_t = 0$, then

$$c_{\rho r}'(\tilde{\rho}_{1t}, \cdots, \tilde{\rho}_{Mt}, \tilde{r}_{1t}, \cdots, \tilde{r}_{Mt})' = -c_D' D_t \text{ almost surely (a.s.).} \tag{A.2}$$

We now consider three cases. First, suppose that $c_{\rho r} \not\equiv 0$ and $c_D \not\equiv 0$. By (A.2), it follows that $D_t \in \sigma(\eta_{t-i}; 0 \le i \le M)$. However, by Assumption 2.2, h_t is non-degenerate, and hence it depends on $\sigma(\eta_{t-i}; i \ge M+1)$. This entails that D_t depends on $\sigma(\eta_{t-i}; i \ge M+1)$. So it is a contradiction due to the independence of $\{\eta_t\}$.

Second, suppose that $c_{\rho r} \not\equiv 0$ and $c_D \equiv 0$. In this case, by (A.2), we have

$$E\left[c_{\rho r}'(\tilde{\rho}_{1t}, \cdots, \tilde{\rho}_{Mt}, \tilde{r}_{1t}, \cdots, \tilde{r}_{Mt})' | \eta_{t-2}, \cdots, \eta_{t-M}\right] = 0,$$

which implies that $c_{\rho 1}\tilde{\rho}_{1t} + c_{r1}\tilde{r}_{1t} = 0$ a.s. According to the independence of $\{\eta_t\}$, we know that $c_{\rho 1} = c_{r1} = 0$. Similarly, we can show that $c_{\rho i} = c_{ri} = 0$ for $i = 2, \cdots, M$. Thus, $c_{\rho r} \equiv 0$, which is contradicted with our presumption on $c_{\rho r}$.

Third, suppose that $c_{\rho r} \equiv 0$ and $c_D \not\equiv 0$. In this case, by (A.2) and using the same argument as for property (ii) in Francq and Zakoïan (2004, page 628), we can show that $c_D \equiv 0$, and hence it is a contradiction.

Therefore, from (A.2), we can only obtain $c_{\rho r} \equiv 0$ and $c_D \equiv 0$. This implies that Ω is positive definite. □

PROOF OF THEOREM 2.1: First, from equation (3) in Li and Li (2008), we have

$$\sqrt{n}(\hat{\theta}_n - \theta_0) = \frac{\Sigma^{-1}}{\sqrt{n}} \sum_{t=1}^{n} D_t + o_p(1), \tag{A.3}$$

where D_t is defined as in (A.1). Next, under Assumptions 2.1-2.5, by a similar argument as for (2.1), we can show that for any integer $l \geq 1$,

$$\sqrt{n}\hat{\rho}_l = \sqrt{n}\rho_l + \sigma_0^{-2} X'_{\rho l}\sqrt{n}(\hat{\theta}_n - \theta_0) + o_p(1),$$
$$\sqrt{n}\hat{r}_l = \sqrt{n}r_l + \sigma_1^{-2} X'_{rl}\sqrt{n}(\hat{\theta}_n - \theta_0) + o_p(1).$$

This follows that

$$\sqrt{n}\hat{\rho} = \sqrt{n}\rho + \sigma_0^{-2} X_{\rho}\sqrt{n}(\hat{\theta}_n - \theta_0) + o_p(1), \tag{A.4}$$
$$\sqrt{n}\hat{r} = \sqrt{n}r + \sigma_1^{-2} X_{r}\sqrt{n}(\hat{\theta}_n - \theta_0) + o_p(1). \tag{A.5}$$

By (A.3)-(A.5), we have $\sqrt{n}\,(\hat{\rho}', \hat{r}')' = VZ_n + o_p(1)$, where

$$Z_n = \sqrt{n}\left[\rho', r', \frac{1}{n}\sum_{t=1}^{n} D_t'\right]'.$$

Finally, the conclusion holds by Lemma A.1 and the martingale central limit theorem. □

REFERENCES

[1] Azzalini, A. (1985) A class of distributions which includes the normal ones. *Scandinavian Journal of Statistics* **12**, 171-178.

[2] Berkes, L., Horváth, L. and Kokoszka, P. (2003) Asymptotics for GARCH squared residual correlations. *Econometric Theory* **19**, 515-540.

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

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

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

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

[7] Fan, J., Qi, L. and Xiu, D. (2010) Quasi maximum likelihood estimation of GARCH models with heavy-tailed likelihoods. Working paper. Princeton University.

[8] Fan, J. and Yao, Q. (2003) Nonlinear time series: Nonparametric and parametric methods. Springer, New York.

[9] Francq, C., Lepage, G. and Zakoïan, J.M. (2011) Two-stage non-Gaussian QML estimation of GARCH models and testing the efficiency of the Gaussian QMLE. *Journal of Econometrics* **165**, 246-257.

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

[11] Francq, C. and Zakoïan, J.M. (2004) Maximum likelihood estimation of pure GARCH and ARMA-GARCH processes. *Bernoulli* **10**, 605-637.

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

[13] 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.

[14] 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.

[15] 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.

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

[17] Ling, S. and Li, W.K. (1997a) Fractional autoregressive integrated moving-average time series conditional hereroskedasticity. *Journal of American Statistical Association* **92**, 1184-1194.

[18] Ling, S. and Li, W.K. (1997b) Diagnostic checking of nonlinear multivariate time series with multivariate arch errors. *Journal of Time Series Analysis* **18**, 447-464.

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

[20] Ling, S. and Tong, H. (2011) Score based goodness-of-fit tests for time series. *Statistica Sinica* **21**, 1807-1829.

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

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

[23] 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.

[24] Peng, L. and Yao, Q.W. (2003) Least absolute deviations estimation for ARCH and GARCH models. *Biometrika* **90**, 967-975.

[25] Shao, X. (2011) Testing for white noise under unknown dependence and its applications to diagnostic checking for time series models. *Econometric Theory.* To appear.

[26] Wong, H. and Li, W.K. (2002) Detecting and diagnostic checking multivariate conditional heteroscedastic time series models. *Annals of the Institute Statistical Mathematics* **54**, 45-59.

[27] Wong, H. and Ling, S. (2005) Mixed portmanteau tests for time series. *Journal of Time Series Analysis* **26**, 569-579.

[28] Zhu, K. (2011) On the LAD estimation and likelihood ratio test for time series models. Thesis Dissertion. Hong Kong University of Science and Technology.

[29] 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.

Chinese Academy of Sciences
Institute of Applied Mathematics
Haidian District, Zhongguancun
Bei Jing, China
E-mail: zkxaa@ust.hk