Residual-Based Diagnostics for Conditional Heteroscedasticity Models

Y.K. Tse

Department of Economics
National University of Singapore
Singapore 119260

Email: ecstseyk@nus.edu.sg

Revised: February 2001

Abstract: We examine the residual-based diagnostics for univariate and multivariate conditional heteroscedasticity models. The tests are based on the parameter estimates of an autoregression with the squared standardized residuals or the cross products of the standardized residuals as dependent variables. As the regression involves estimated regressors the standard distribution theories of the ordinary least squares estimates do not apply. We provide the asymptotic variance of the regression estimates. Diagnostic statistics, which are asymptotically distributed as χ^2, are constructed. A Monte Carlo experiment is conducted to investigate the finite-sample properties of the residual-based tests for both univariate and multivariate models. The results show that the residual-based diagnostics provide useful checks for model adequacy in both univariate and multivariate cases.

Key Words: conditional heteroscedasticity, Lagrange multiplier test, Monte Carlo experiment, portmanteau statistic, residual-based diagnostic

1 Introduction

Since the seminal paper by Engle (1982) on the autoregressive conditional heteroscedasticity (ARCH) models, many alternative models have been suggested for time series with time-varying variance. Bollerslev's (1986) generalized ARCH (GARCH) model is a natural extension, with lagged conditional variances introduced as explanatory variables in the conditional-variance equation. Further extensions were suggested by Nelson (1991) (exponential GARCH (EGARCH) model), Higgins and Bera (1992) (nonlinear ARCH (NARCH) model), Glosten, Jagannathan and Runkle (1993) (an asymmetric model commonly called the GJR model), Ding, Granger and Engle (1993) (asymmetric power ARCH (APARCH) model) and Zakoian (1994) (threshold ARCH (TARCH) model). Hentschel (1995) proposed a model that encompasses many existing models. For surveys of the developments and applications of these models, see Bollerslev, Chou and Kroner (1992), Bera and Higgins (1993) and Bollerslev, Engle and Nelson (1994).

The success of the conditional heteroscedasticity models in fitting univariate time series has motivated many researchers to extend these models to the multivariate case. Examples of multivariate conditional heteroscedasticity models are the vech-representation form due to Bollerslev, Engle and Wooldridge (1988), the constant-correlation multivariate GARCH (CC-MGARCH) model due to Bollerslev (1990) and the BEKK (named after Baba, Engle, Kraft and Kroner) model due to Engle and Kroner (1995). Within the vech-representation family, the diagonal form, which we shall denote as the DVR model, has been applied in many empirical works.

As empirical researchers are equipped with various conditional heteroscedasticity models, the checking of the adequacy of a fitted model becomes an important issue for model selection. Generally, misspecified models result in inconsistency and loss of efficiency in the estimated parameters. It should be noted that in many financial econometric models the conditional-variance equation plays a major role. For example, the systematic risk as measured by beta depends on the (conditional) second moments of the asset returns, and so is the minimum-variance hedge ratio. Reliable estimates and inference of these quantities depend on well-specified conditional heteroscedasticity models.

Diagnostics for conditional heteroscedasticity models applied in the literature can be divided into three categories: portmanteau tests of the Box-Pierce-Ljung type, Lagrange multiplier (LM) tests and residual-based diagnostics. The Box-Pierce-Ljung portmanteau statistic is perhaps the most widely used diagnostic. It is readily computable from the standardized residuals and has been applied in many empirical works for model diagnostics (see, for example, the papers by Bollerslev (1990), Baillie and Myers (1991) and Karolyi (1995)). Although it has been noted that the portmanteau statistics do not have an asymptotic χ^2 distribution, many authors, nonetheless, apply the χ^2 distribution as an approximation. Li and Mak (1994) pointed out the lack of rigor in this approach and derived the asymptotic distribution of the portmanteau statistics in the univariate case.[1] Ling and Li (1997) further developed this work and derived the asymptotic distribution of the portmanteau statistics in the multivariate case. The Ling-Li statistic is based on the serial correlation coefficients of the transformed vector of residuals. The recent test suggested by Hong and Shehadeh (1999) is a portmanteau test based on the spectral density function rather than the serial correlation coefficients. Tse and Zuo (1998) and Tse and Tsui (1999) reported some Monte Carlo results on the finite-sample distributions of the Li-Mak test and the Ling-Li test, respectively.

The LM test has an advantage over the portmanteau test due to its efficiency when the alternative hypothesis is correct. The test, however, requires the specification of an alternative, and it may not have good power against other alternatives. Also, the calculation of the test statistic depends on the alternative and thus makes this approach

[1]The problem lies in the fact that estimated residuals are used to calculate the portmanteau statistics. When only the conditional mean is estimated, substituting the estimated residuals for the unobservable residuals has no effect on the asymptotic distribution of the diagnostic statistic as far as tests for conditional heteroscedasticity are concerned, as pointed out by Engle (1982). However, if a conditional-variance equation is fitted, substituting the estimated residuals for the unobserved residuals will change the asymptotic distribution of the diagnostic statistic, as pointed out by Li and Mak (1994).

less attractive as a general diagnostic tool. For the applications of LM tests to conditional heteroscedasticity models, see Bollerslev, Engle and Wooldridge (1988), Engle and Ng (1993) and Engle and Kroner (1995). Bera and Higgins (1992) suggested a diagnostic of the ARCH models against the NARCH alternatives based on the LM principle. Tse (2000) proposed a test for constant correlations in a multivariate GARCH model using the LM approach.[2] Lundbergh and Terasvirta (1998) provided a unified framework for testing univariate conditional heteroscedasticity models based on the LM principle. Some results on the equivalence between the LM and the portmanteau tests in certain cases can be found in Ling and Li (1997) and Lundbergh and Terasvirta (1998).[3]

Like the portmanteau tests, residual-based diagnostics have no specific alternative. General model adequacy is investigated using the residuals. The diagnostics involve running artificial regressions and testing for the statistical significance of the regression parameters. To a certain extent, the form of the regression depends on a particular type of model inadequacy the researcher wants to investigate. Extensive discussions of this approach can be found in Pagan and Hall (1983) and Wooldridge (1990). To test for the adequacy of the conditional-variance structure, lagged squared standardized residuals may be used, as suggested by Bollerslev (1990). As the regressors are estimated, the usual ordinary least squares (OLS) result does not apply. Empirical research, however, typically adopts the usual OLS procedure as an approximation. In the multivariate case, the Monte Carlo results reported by Tse and Tsui (1999) showed that the use of the OLS inference procedure grossly under-rejects the null hypothesis of model adequacy.

In this paper we present the asymptotic distributions of the residual-based diagnostics for the conditional heteroscedasticity models. Both univariate and multivariate models are considered. In the multivariate case we propose to examine the squared standardized residuals as well as the cross products of the standardized residuals. Diagnostic statistics

[2] Bera and Kim (1996) suggested a test for constant correlation in a bivariate GARCH model.

[3] Lundbergh and Terasvirta (1998) pointed out that the portmanteau test of model adequacy is equivalent to the LM test of the standardised residuals being independently and identically distributed against the alternative of ARCH.

based on the correct asymptotic variance of the OLS regression parameter estimates are constructed. We examine the finite-sample properties of the residual-based diagnostics using Monte Carlo methods. Our results show that the residual-based diagnostics provide a useful check for model adequacy.

The plan of the rest of the paper is as follows. In Section 2 we present the results for the asymptotic distributions of the residual-based diagnostic tests for the univariate as well as multivariate conditional heteroscedasticity models. Section 3 reports the Monte Carlo results of the finite-sample distributions of the residual-based diagnostics. We consider a variety of univariate and multivariate conditional heteroscedasticity models. Both the size and the power of the diagnostics are studied. Finally, we give some concluding remarks in Section 4.

2 Residual-Based Diagnostic Tests

Residual-based diagnostics are constructed to test for certain residual patterns implied by the deviation of the fitted model from its underlying assumptions. Pagan and Hall (1983) and Wooldridge (1990) provided comprehensive discussions of residual-based tests. These tests may be designed to diagnose particular types of model misspecifications, including serial correlation, heteroscedasticity, constancy of coefficients, nonnormality, simultaneity and so on. In this paper, our concern is conditional heteroscedasticity in time series models.[4]

[4]See also Wooldridge (1991) for some results on testing for conditional variances. Our results differ from Wooldridge's result in the following aspects. First, Wooldridge's results are for univariate time series only, although extension to the multivariate case can be developed following the theorems in Wooldridge (1990). Second, Wooldridge's method assumes only consistency for the parameter estimates. Although his method is robust to certain model misspecifications, there may be loss in power compared to tests that exploit the asymptotic efficiency of the MLE (not necessarily assuming normality in the errors). As pointed out by Wooldridge (1990, p.26) the statistics computed under the two approaches may not be asymptotically equivalent.

2.1 The Univariate Case

Consider a univariate time series $\{X_t\}$, for $t = 1, ..., T$, with conditional heteroscedasticity generated by the following equations:

$$X_t - \mu_t = \varepsilon_t, \tag{1}$$

$$\varepsilon_t = \sigma_t \, \eta_t, \tag{2}$$

where η_t are independently and identically distributed with mean zero and variance 1, and μ_t and σ_t^2 are, respectively, the conditional mean and variance of X_t based on the information set Φ_{t-1} at time $t-1$. This framework incorporates many ARCH and GARCH type of conditional heteroscedasticity models applied in the literature. Furthermore, μ_t may be nonlinear functions of past observations and/or dependent on weakly exogenous variables.

Let θ be the parameter vector of the model and $\hat{\theta}$ be the maximum likelihood estimator (MLE) of θ.[5] Here we assume that θ is with N elements and contains all the parameters appearing in μ_t and σ_t. We assume that the usual regularity conditions hold,[6] so that $\sqrt{T}\,(\hat{\theta} - \theta) \stackrel{D}{\rightarrow} N(0, G)$, where $\stackrel{D}{\rightarrow}$ denotes convergence in distribution. Assuming correct error specification, G can be estimated either by the inverse of the Hessian matrix or the cross-product of the first derivative of the likelihood function. To allow for misspecification in the model, however, the robust quasi-MLE of the variance matrix, as proposed by Bollerslev and Wooldridge (1992) may be used.

Denoting $\hat{\sigma}_t^2$ as σ_t^2 evaluated at $\hat{\theta}$ and $\hat{\varepsilon}_t$ as the estimated residual, we define the standardized residual as $\hat{\eta}_t = \hat{\varepsilon}_t / \hat{\sigma}_t$. Following Pagan and Hall (1983) and Bollerslev (1990), model diagnostics can be conducted using the standardized residuals. Noting that the conditional expectations of the squared standardized residuals tend to 1, we run a regression of $\hat{\eta}_t^2 - 1$ on some information variables and examine the statistical significance

[5] In the literature, the most commonly assumed likelihood functions are probably based on errors being normal or Student's t.

[6] See, for example, Bollerslev and Wooldridge (1992) and White (1994) for the details.

of the regression parameters. The lagged standardized residuals are natural regressors to use. Thus, denoting $\hat{d}_t = (\hat{\eta}^2_{t-1}, ..., \hat{\eta}^2_{t-M})'$, we consider the regression

$$\hat{\eta}^2_t - 1 = \hat{d}'_t\, \delta + \xi_t, \tag{3}$$

where δ is an M-vector of regression parameters. We denote the OLS estimator of δ by $\hat{\delta}$.

As $\hat{d}_t$ consists of estimated regressors, the inference procedure based on the usual OLS results is invalid. This point was stressed by Pagan and Hall (1983) in a broader context. The appropriate procedure is to correct for the asymptotic variance of the OLS estimate. Pierce (1982) established a general result for the asymptotic variance matrix of a test statistic which involves substituting estimates for unknown parameters.[7] To apply Pierce's result we denote δ^* as the (unobservable) OLS estimate when $\eta^2_t - 1$ is regressed on $d_t = (\eta^2_{t-1}, ..., \eta^2_{t-M})'$. Following Pierce (1982) we require the following assumptions: (A1) δ^* and $\hat{\theta}$ are asymptotically jointly normally distributed, and (A2) $\hat{\theta}$ is asymptotically efficient. Theorem 1 below provides the asymptotic distribution of $\hat{\delta}$.[8]

Theorem 1: If equations (1) and (2) specify the correct model for the univariate time series $\{X_t\}$, and assumptions A1 and A2 hold, then $\sqrt{T}\,\hat{\delta} \xrightarrow{D} N(0,\, L^{-1}\Omega L^{-1})$, where

$$L = \operatorname{plim}\,(\frac{1}{T}\sum d_t\, d'_t), \tag{4}$$

$$\Omega = cL - QGQ', \tag{5}$$

with

$$Q = \operatorname{plim}\,(\frac{1}{T}\sum d_t\, \frac{\partial \eta^2_t}{\partial \theta'}) \tag{6}$$

and $c = \mathrm{E}\{(\eta^2_t - 1)^2\}$.

In empirical applications, c, L and Q may be estimated by $\hat{c} = \{\sum(\hat{\eta}^2_t - 1)^2\}/T$, $\hat{L} = (\sum \hat{d}_t\, \hat{d}'_t)/T$, and $\hat{Q} = \{\sum \hat{d}_t\,(\partial\hat{\eta}^2_t/\partial\theta')\}/T$, respectively.[9] Thus, when there is no

[7] Bera and Zuo (1996) applied Pierce's result to derive diagnostic tests for ARCH models. See also Bera and Kim (1996) for similar applications.

[8] The Appendix provides a summary of the result due to Pierce (1982) and a proof of Theorem 1.

[9] Note that $\partial\hat{\eta}^2_t/\partial\theta'$ denotes $\partial\eta^2_t/\partial\theta'$ evaluated at $\theta = \hat{\theta}$. Similar interpretations are adopted for other derivatives used below.

misspecification in the model, $T\hat{\delta}'\hat{L}\hat{\Omega}^{-1}\hat{L}\hat{\delta}$ is asymptotically distributed as a χ^2_M, where $\hat{\Omega} = \hat{c}\hat{L} - \hat{Q}\hat{G}\hat{Q}'$. The derivatives in $\hat{Q}$ may be computed using numerical methods. Alternatively, for GARCH type of models recursive formulae as given by Fiorentini, Calzolari and Panattoni (1996) and Tse (2000) may be used.

Following the arguments of Ling and Li (1997) and Lundbergh and Terasvirta (1998) it can be shown that the residual-based diagnostic is asymptotically equivalent to the portmanteau statistic as well as the LM statistic of no ARCH in the standardized errors against ARCH(M). This result is particularly clear when we use the sample moment of $\hat{d}_t$ as an estimate of L so that the residual-based statistic becomes $(\sum \hat{d}_t\hat{v}_t)'\hat{\Omega}^{-1}(\sum \hat{d}_t\hat{v}_t)/T$, where $\hat{v}_t = \hat{\eta}_t^2 - 1$. As the vector of autoregressive coefficients of the standardized residuals are asymptotically equivalent to $(\sum \hat{d}_t\hat{v}_t)/(cT)$ in distribution under the null, the residual-based statistic is asymptotically equivalent to the portmanteau statistic. However, two points should be made here. First, although the tests are asymptotically equivalent under model adequacy, they may differ under model misspecification. That is, the power of the tests may differ. Second, the performance of the tests may differ in finite samples.

2.2 The Multivariate Case

In this subsection notations are redefined to cater for multivariate observations. Thus, $X_t = (X_{t1}, ..., X_{tK})'$ denote a K-vector of observations generated by the following equations:

$$X_t - \mu_t = \varepsilon_t, \tag{7}$$

$$\varepsilon_t = V_t^{1/2}\zeta_t, \tag{8}$$

where ε_t and μ_t are K-vectors of residual and conditional mean, respectively, $V_t = \{\sigma_{tij}\}$ is the conditional variance matrix of $\varepsilon_t = (\varepsilon_{t1}, ..., \varepsilon_{tK})'$, and $\zeta_t = (\zeta_{t1}, ..., \zeta_{tK})'$ are independently and identically distributed random variables with mean zero and variance I_K (the $K \times K$ identity matrix). We standardize ε_{ti} to obtain $\eta_{ti} = \varepsilon_{ti}/\sigma_{tii}^{1/2}$ for $i = 1, ..., K$. Note that the variance matrix of $\eta_t = (\eta_{t1}, ..., \eta_{tK})'$ has unity in the diagonal but is not equal to the identity matrix. Thus, the components of η_t are generally correlated. Again we let θ be the N-vector of the parameters of the model and $\hat{\theta}$ be the MLE of θ. We assume that under certain regularity conditions, $\sqrt{T}\,(\hat{\theta} - \theta) \xrightarrow{D} N(0,\, G)$.

Let $\hat{V}_t = \{\hat{\sigma}_{tij}\}$ be the estimated conditional variance matrix, $\hat{\varepsilon}_t = (\hat{\varepsilon}_{t1}, ..., \hat{\varepsilon}_{tK})'$ be the estimated residual and $\hat{\eta}_t = (\hat{\eta}_{t1}, ..., \hat{\eta}_{tK})'$ be the standardized residual with $\hat{\eta}_{ti} = \hat{\varepsilon}_{ti}/\hat{\sigma}_{tii}^{1/2}$. We also denote $\rho_{tij} = \sigma_{tij}/(\sigma_{tii}\sigma_{tjj})^{1/2}$ as the conditional correlation and $\hat{\rho}_{tij} = \hat{\sigma}_{tij}/(\hat{\sigma}_{tii}\hat{\sigma}_{tjj})^{1/2}$ as its estimated value. Residual-based diagnostics can be conducted on the squared standardized residuals and the cross products of the standardized residuals. Extending the results on the univariate case, we run regressions of $\hat{\eta}_{ti}^2 - 1$, for $i = 1, ..., K$, and $\hat{\eta}_{ti}\,\hat{\eta}_{tj} - \hat{\rho}_{tij}$, for $1 \leq i < j \leq K$, on some information variables. Again, the lagged squared standardized residuals and the lagged cross products of the standardized residuals are natural candidates. Thus, denoting $\hat{d}_{ti} = (\hat{\eta}_{t-1,i}^2, ..., \hat{\eta}_{t-M,i}^2)'$ and $\hat{d}_{tij} = (\hat{\eta}_{t-1,i}\,\hat{\eta}_{t-1,j}, ..., \hat{\eta}_{t-M,i}\,\hat{\eta}_{t-M,j})'$, we consider the following regressions:[10]

$$\hat{\eta}_{ti}^2 - 1 = \hat{d}_{ti}'\,\delta_i + \xi_{ti}, \qquad i = 1, ..., K, \tag{9}$$

$$\hat{\eta}_{ti}\,\hat{\eta}_{tj} - \hat{\rho}_{tij} = \hat{d}_{tij}'\,\delta_{ij} + \xi_{tij}, \qquad 1 \leq i < j \leq K, \tag{10}$$

where δ_i and δ_{ij} are M-vectors of regression parameters. We further define $d_{ti} = (\eta_{t-1,i}^2, ..., \eta_{t-M,i}^2)'$ and $d_{tij} = (\eta_{t-1,i}\,\eta_{t-1,j}, ..., \eta_{t-M,i}\,\eta_{t-M,j})'$, which are the unobservable counterparts of $\hat{d}_{ti}$ and $\hat{d}_{tij}$, respectively. The following theorems provide the asymptotic distributions of the OLS estimators $\hat{\delta}_i$ and $\hat{\delta}_{ij}$ of δ_i and δ_{ij}, respectively (see the Appendix for the proof).

Theorem 2: If equations (7) and (8) specify the correct model for the multivariate time series $\{X_t\}$, and assumptions A1 and A2 hold, then $\sqrt{T}\,\hat{\delta}_i \xrightarrow{D} N(0,\, L_i^{-1}\Omega_i L_i^{-1})$,

[10]Note that M is generic and does not imply that the regressions represented by (9) and (10) are of the same order.

where[11]

$$\begin{aligned} L_i &= \text{plim}(\frac{1}{T}\sum d_{ti}\, d'_{ti}), \qquad (11) \\ \Omega_i &= c_i L_i - Q_i G Q'_i, \qquad (12) \end{aligned}$$

with

$$Q_i = \text{plim}(\frac{1}{T}\sum d_{ti}\,\frac{\partial \eta^2_{ti}}{\partial \theta'}), \qquad (13)$$

and $c_i = \text{E}\{(\eta^2_{ti} - 1)^2\}$.

To compute a diagnostic we replace c_i, L_i, and Q_i by their sample estimates, denoted by hats, so that $\hat{c}_i = \{\sum(\hat{\eta}^2_{ti} - 1)^2\}/T$, $\hat{L}_i = (\sum \hat{d}_{ti}\,\hat{d}'_{ti})/T$ and $\hat{Q}_i = \{\sum \hat{d}_{ti}\,(\partial\hat{\eta}^2_{ti}/\partial\theta')\}/T$. The test statistic can then be calculated as $T\hat{\delta}'_i\hat{L}_i\hat{\Omega}_i^{-1}\hat{L}_i\hat{\delta}_i$, which is asymptotically distributed as a χ^2_M.

Theorem 3: If equations (7) and (8) specify the correct model for the multivariate time series $\{X_t\}$, and assumptions A1 and A2 hold, then $\sqrt{T}\,\hat{\delta}_{ij} \xrightarrow{D} N(0,\, L^{-1}_{ij}\Omega_{ij}L^{-1}_{ij})$, where

$$\begin{aligned} L_{ij} &= \text{plim}(\frac{1}{T}\sum d_{tij}\, d'_{tij}) \qquad (14) \\ \Omega_{ij} &= c_{ij} L_{ij} - Q_{ij} G Q'_{ij} \qquad (15) \end{aligned}$$

with

$$Q_{ij} = \text{plim}\{\frac{1}{T}\sum d_{tij}\frac{\partial(\eta_{ti}\,\eta_{tj} - \rho_{tij})}{\partial\theta'}\}, \qquad (16)$$

and $c_{ij} = \text{E}\{(\eta_{ti}\,\eta_{tj} - \rho_{tij})^2\}$.

To compute a diagnostic we replace c_{ij}, L_{ij}, and Q_{ij} by their sample estimates, denoted by hats, so that $\hat{c}_{ij} = \{\sum(\hat{\eta}_{ti}\hat{\eta}_{tj} - \hat{\rho}_{tij})^2\}/T$, $\hat{L}_{ij} = (\sum \hat{d}_{tij}\,\hat{d}'_{tij})/T$ and $\hat{Q}_{ij} =$

[11] In the multivariate case, A1 refers to the assumption that the (unobservable) OLS estimates of δ_i and δ_{ij} (when the *actual* residuals are used in the regression) are asymptotically jointly normally distributed with the MLE of θ.

$[\sum \hat{d}_{tij}\{\partial(\hat{\eta}_{ti}\hat{\eta}_{tj}-\hat{\rho}_{tij})/\partial\theta')\}]/T$. A diagnostic test statistic may be calculated as $T\hat{\delta}'_{ij}\hat{L}_{ij}\hat{\Omega}_{ij}^{-1}\hat{L}_{ij}\hat{\delta}_{ij}$, which is asymptotically distributed as a χ^2_M.

It should be noted that equations (3), (9) and (10) are just some possible forms of artificial regression. There are other possibilities. For example, lagged values of $\hat{\eta}_{ti}$ may be added to equation (9) as regressors to test for asymmetric effects. Also, lagged values of $\hat{\eta}^2_{tj}$ may be included in the equation as regressors to test for volatility spill-over.. As long as the regressors contain only lagged values, the martingale difference property is not affected. Thus, the formulae for the asymptotic variance of the OLS estimates remain applicable with suitable definition of the vector $\hat{d}_{ti}$.[12]

We conclude this section by making the caveat that we have proposed a battery of tests which are statistically dependent. Thus, there is the issue of controlling the size of multiple tests. Unlike the Ling-Li test, which is a general diagnostic for multivariate conditional heteroscedasticity, the diagnostics we propose for the multivariate case examine separate aspects of possible model misspecification. Notwithstanding the problem of multiple test, examining a battery of diagnostics provides additional information about the likely source of misspecification and thus may provide clues as to how the model may be reformulated if a diagnostic is found to be significant.[13]

3 Some Monte Carlo Results

In this section we report the results of a Monte Carlo experiment on the finite-sample distributions of the diagnostics suggested in Section 2. Subsection 3.1 discusses the results of the univariate case, while Subsection 3.2 discusses the results of the multivariate case.

[12] I am indebted to an anonymous referee for this suggestion.

[13] Hendry (1995, pp.490–491 and chapter 15) discusses the issues of multiple testing and data mining. When the tests are independent, it is easy to control the overall size of the diagnostics. Though the tests suggested here are not independent, our Monte Carlo results showed that the correlations between the diagnostics when there is no misspecification are very low.

3.1 The Univariate Case

We consider three data generating processes (DGP), denoted by M1, M2 and M3. These are low-order ARCH and GARCH processes. M1 is an ARCH(1) process, M2 is an ARCH(2) process and M3 is an GARCH(1, 1) process. The parameters of the DGP are given by the following conditional-variance equations:

$$\text{M1} \quad : \quad \sigma_t^2 = 0.2 + 0.6\varepsilon_{t-1}^2, \tag{17}$$

$$\text{M2} \quad : \quad \sigma_t^2 = 0.2 + 0.4\varepsilon_{t-1}^2 + 0.4\varepsilon_{t-2}^2, \tag{18}$$

$$\text{M3} \quad : \quad \sigma_t^2 = 0.2 + 0.6\sigma_{t-1}^2 + 0.2\varepsilon_{t-1}^2. \tag{19}$$

The conditional mean μ_t of each DGP is assumed to be zero. Given a DGP we generate samples of T observations and fit a conditional heteroscedasticity model to the data. The estimated model (EM) considered are ARCH(1), ARCH(2) and GARCH(1, 1). Various diagnostics are then calculated. The sample estimates defined in Section 2 are used to compute the test statistics. Also, the robust quasi-MLE variance matrix of $\hat{\theta}$ is used. Thus, the diagnostics are robust with respect to nonnormality in the residuals.[14] To investigate the effects of incorrect error specification on the test statistics we generate errors that are (i) normally distributed, and (ii) distributed as t_8 (properly standardized to give unit variance). However, in all cases the MLE are estimated assuming errors are normal.[15] All computations in the Monte Carlo experiment are coded in GAUSS supplemented by the application module Constrained Maximum Likelihood.

We denote RBM as the adjusted (with the correct asymptotic variance) residual-based diagnostic with M lagged regressors, PORM as the Li-Mak portmanteau statistic with

[14] The portmanteau statistic of Li and Mak (1994) is also robustified, and so is the Ling-Li (1997) statistic reported in the next subsection.

[15] When the errors are normally distributed there is no model misspecification and the asymptotic theory of the test statistics should apply. When the errors are distributed as t_8, the MLE assuming normal errors will not be efficient and Pierce's result will not apply. How this misspecification will affect the asymptotic distribution of the test statistic is unknown. Bollerslev and Wooldridge (1992) argued that the loss of efficiency in the quasi-MLE with symmetric t-distributed errors is small. In this case, we may expect the residual-based diagnostic to be robust against t-distributed errors.

M lagged autocorrelation coefficients and OLSM as the residual-based diagnostic using the unadjusted OLS variance. All test statistics are compared against the χ^2 critical value. We let $T = 200$, 500 and 1000, and estimate the empirical size and power of the diagnostics based on Monte Carlo sample size of 1000.

It has been noted that the validity of the diagnostics does not depend on M being large. Indeed, the Monte Carlo results of Tse and Zuo (1997) showed that the power of the Li-Mak test drops when M is large. Thus, we consider low values of $M = 1, 2, 3$ and 4. We first examine the empirical size of the test when the correct model is estimated for each of the DGP. We set the nominal size to be 5 percent. The empirical size of the three types of diagnostics are summarized in Table 1.

It can be seen that the OLS test grossly under-rejects the null hypothesis of model adequacy. As the correction term QGQ', which is positive semidefinite, is dropped from the variance matrix of $\sqrt{T}\hat{\delta}$, the variance matrix has been overstated. This results in a smaller test statistic and causes the OLS test to under-reject the null hypothesis.[16] The RB and POR tests generally have quite reliable size, although there is a slight tendency for these tests to over-reject rather than under-reject. Rather remarkably, these diagnostics give good empirical size even for relatively small sample size of 200. It is also clear that the RB and POR tests are quite robust to nonnormal errors. The empirical sizes when the errors are t-distributed are quite similar to the case when the errors are normal.

Table 2 summarizes the results of the power consideration of the tests. As our set-up has three DGPs and three EMs, there are a total of nine combinations. To investigate the power of the tests we consider the combinations for which the DGP is not nested within the EM. Thus, denoting the combinations by {DGP, EM}, we consider the four combinations: {M2, ARCH(1)}, {M2, GARCH(1, 1)}, {M3, ARCH(1)} and {M3, ARCH(2)}.[17] As the OLS test does not have the correct empirical size it is excluded from the power comparison.

[16] I am indebted to an anonymous referee for this point.

[17] Apart from the cases considered in Table 1, other combinations for which the DGP is nested within the EM are {M1, ARCH(2)} and {M1, GARCH(1, 1)}. These combinations are not considered.

It can be seen that the RB and POR tests are quite similar. The empirical power of the tests appears to be the lowest when $M = 1$. Otherwise, there seems to be no clear-cut choice among $M = 2$, 3 or 4. It is noted that the power of the tests drops when the true residuals are t-distributed. This may be due to the loss of efficiency in the quasi-MLE which assumes normal errors. The tests give quite low power when the DGP is M3 and the EM is ARCH(2). Thus, the missing σ_{t-1}^2 term in the conditional-variance equation of the GARCH(1, 1) model is picked up by the lagged ε_t^2 terms in the fitted equation. Overall, the results suggest that the properly constructed residual-based test provides a useful diagnostic for conditional heteroscedasticity.

3.2 The Multivariate Case

Recently Tse and Tsui (1999) examined the performance of several model diagnostics for multivariate conditional heteroscedasticity models. They compared the Ling-Li portmanteau test with the (uncorrected) equation-by-equation portmanteau test and (uncorrected) residual-based test in a Monte Carlo experiment. They found that the Ling-Li test may have very weak power under certain circumstances. The uncorrected residual-based test grossly under-rejects the null, while the uncorrected portmanteau test based on the cross products of the standardized residuals may provide a useful diagnostic.[18]

The estimation of multivariate conditional heteroscedasticity models is computationally more tricky than the univariate models. The main difficulty lies in controlling the conditional variance matrix to be positive definite in each iteration. There are a number of alternative forms of multivariate conditional heteroscedasticity models in the literature. Of these models, the CC-MGARCH model appears to have the best convergence property. While the BEKK model is deemed to provide a positive definite conditional variance matrix regardless of the parameter values, our experience is that the convergence

[18] Although the Tse-Tsui study showed that the asymptotic χ^2 approximation works well for the portmanteau statistics the correct asymptotic distribution of the test has not been established. It can be seen, however, that the asymptotic distributions of the portmanteau statistics based on the standardised residuals of individual equations can be developed along the line in Section 2.

of the BEKK model is extremely slow. Although the experiment started with investigating the CC-MGARCH model and the BEKK model as the EM, we decided to abort the BEKK model as we proceeded. In this paper we report the Monte Carlo results with the CC-MGARCH(1, 1) model as the only EM.

We consider the following DGPs:

1. CC-MGARCH Model:

$$\begin{pmatrix} \sigma_{t11} \\ \sigma_{t22} \end{pmatrix} = \begin{pmatrix} 0.2 + 0.8\,\sigma_{t-1,11} + 0.1\,\varepsilon^2_{t-1,1} \\ 0.2 + 0.8\,\sigma_{t-1,22} + 0.1\,\varepsilon^2_{t-1,2} \end{pmatrix}, \tag{20}$$

$$\sigma_{t12} = 0.5\sqrt{\sigma_{t11}\,\sigma_{t22}} \tag{21}$$

2. BEKK(D) Model:[19]

$$V_t = \begin{pmatrix} 0.2 & 0.1 \\ 0.1 & 0.2 \end{pmatrix} + \begin{pmatrix} 0.8 & 0 \\ 0 & 0.8 \end{pmatrix} V_{t-1} \begin{pmatrix} 0.8 & 0 \\ 0 & 0.8 \end{pmatrix} \tag{22}$$

$$+ \begin{pmatrix} 0.4 & 0 \\ 0 & 0.4 \end{pmatrix} \varepsilon_{t-1}\,\varepsilon'_{t-1} \begin{pmatrix} 0.4 & 0 \\ 0 & 0.4 \end{pmatrix}. \tag{23}$$

3. BEKK Model:

$$V_t = \begin{pmatrix} 0.2 & 0.1 \\ 0.1 & 0.2 \end{pmatrix} + \begin{pmatrix} 0.4 & 0.1 \\ 0.1 & 0.4 \end{pmatrix} V_{t-1} \begin{pmatrix} 0.4 & 0.1 \\ 0.1 & 0.4 \end{pmatrix} \tag{24}$$

$$+ \begin{pmatrix} 0.4 & 0.1 \\ 0.1 & 0.4 \end{pmatrix} \varepsilon_{t-1}\,\varepsilon'_{t-1} \begin{pmatrix} 0.4 & 0.1 \\ 0.1 & 0.4 \end{pmatrix}. \tag{25}$$

4. DVR Model:

$$\begin{pmatrix} \sigma_{t11} \\ \sigma_{t12} \\ \sigma_{t22} \end{pmatrix} = \begin{pmatrix} 0.2 + 0.6\,\sigma_{t-1,11} + 0.2\,\varepsilon^2_{t-1,1} \\ 0.1 + 0.4\,\sigma_{t-1,12} + 0.1\,\varepsilon_{t-1,1}\,\varepsilon_{t-1,2} \\ 0.2 + 0.6\,\sigma_{t-1,22} + 0.2\,\varepsilon^2_{t-1,2} \end{pmatrix}. \tag{26}$$

Table 3 summarizes the results of the empirical size (when the DGP is CC-MGARCH) and power of the diagnostics. LLM denotes the Ling-Li statistic with lag-M autocorrelation coefficients. R1-M, R2-M and R3-M denote the residual-based statistics with,

[19] This acronym denotes the diagonal form of the BEKK model.

respectively, $\hat{\eta}_{t1}^2 - 1$, $\hat{\eta}_{t2}^2 - 1$ and $\hat{\eta}_{t1}\hat{\eta}_{t2} - \hat{\rho}$ as the dependent variables in the artificial regression.[20] We consider $M = 1, 2, 3$ and 4, $T = 500$ and 1000, and use Monte Carlo samples of 1000. Rejection frequencies at nominal size of 5 percent are recorded. It can be seen that the tests have good empirical size for the sample sizes considered. There may be signs of over-rejection for $M = 4$. Otherwise, the nominal size appears to be accurate. Rather remarkably the nominal size is quite reliable even for the case when the true errors are t-distributed.[21] Clearly, the R3 test (based on $\hat{\eta}_{t1}\hat{\eta}_{t2} - \hat{\rho}$) represents the test with the best power against the alternatives considered. For the DGP considered, the Ling-Li test has quite weak power. This reinforces the results of Tse and Tsui (1999). Similarly we can see that the residual-based diagnostic based on $\hat{\eta}_{ti}^2 - 1$ also have rather weak power. Once again, there is a drop in power when the true errors are t-distributed, probably due to the loss of efficiency in the quasi-MLE. Overall, the R3 test provides a useful check for the model adequacy of multivariate conditional heteroscedasticity models.

4 Conclusions

We have provided the asymptotic distributions of the residual-based diagnostics for the conditional heteroscedasticity models. Both univariate and multivariate models are considered. In the univariate case we consider the artificial regression of the squared standardized residual on its lagged values. In the multivariate case we propose to examine the squared standardized residuals as well as the cross products of the standardized residuals. Diagnostic statistics based on the correct asymptotic variance of the OLS regression parameter estimate are constructed. We examine the finite-sample properties of the residual-based diagnostics using Monte Carlo methods. In the univariate case we find that the residual-based diagnostics have similar performance versus the Li-Mak portmanteau test.

[20]Refer to Subsection 2.2 for the notations. As the EM assumes constant correlations and bivariate time series are considered, the notation for the dependent variable in the regression involving the cross products of the standardised residuals can be simplified to $\hat{\eta}_{t1}\hat{\eta}_{t2} - \hat{\rho}$.

[21]The over-rejection appears to be more serious for the Ling-Li test, especially for the case of t-distributed errors.

In the multivariate case, the residual-based diagnostics based on the cross products of the standardized residuals provide tests with the appropriate empirical size and good power against the alternatives considered. The residual-based diagnostics are easy to compute and would be useful tools for checking the adequacy of conditional heteroscedasticity models.

Acknowledgment: Research support from the National University of Singapore Academic Research Grant RP-3981003 is gratefully acknowledged. I am indebted to Professor Richard Smith (the Editor) and two anonymous referees for their many helpful comments. Any remaining errors are, of course, mine only.

Appendix

To prove Theorem 1 we first note that

$$\sqrt{T}\,\hat{\delta} = (\frac{1}{T}\sum \hat{d}_t\,\hat{d}_t')^{-1}(\frac{1}{\sqrt{T}}\sum \hat{d}_t\hat{v}_t). \tag{I.1}$$

As $\text{plim}\{(\sum \hat{d}_t\,\hat{d}_t')/T\} = \text{plim}\{(\sum d_t\,d_t')/T\} = L$, it is sufficient to derive the asymptotic distribution of $(\sum \hat{d}_t\hat{v}_t)/\sqrt{T}$. We denote $\lambda_T = (\sum d_t v_t)/T$ and $\hat{\lambda}_T = (\sum \hat{d}_t\hat{v}_t)/T$, where $v_t = \eta_t^2 - 1$, and make the following assumptions. Firstly, assumption A1 states that[22]

$$\begin{bmatrix} \sqrt{T}\lambda_T \\ \sqrt{T}(\hat{\theta}-\theta) \end{bmatrix} \xrightarrow{D} N(0, \begin{bmatrix} V_{11} & V_{12} \\ V_{21} & V_{22} \end{bmatrix}). \tag{I.2}$$

Thus, $\sqrt{T}\lambda_T$ and $\sqrt{T}(\hat{\theta}-\theta)$ are jointly asymptotically normally distributed. Secondly assumption A2 states that $\hat{\theta}$ is asymptotically efficient. Under these assumptions, Pierce's (1982) argument shows that[23]

$$\sqrt{T}\hat{\lambda}_T \xrightarrow{D} N(0, V_{11} - BV_{22}B'), \tag{I.3}$$

where $B = \lim_{T\to\infty}\text{E}(\partial\lambda_T/\partial\theta')$. As $V_{22} = G$, to prove Theorem 1 it is sufficient to show that $V_{11} = cL$ and $B = Q$. Since $\{d_t v_t\}$ is a martingale difference sequence, from the martingale central limit theorem (see White (1994, Theorem 5.23)) we have $V_{11} = cL$. Now

$$\text{E}(\frac{\partial\lambda_T}{\partial\theta'}) = \text{E}[\{\sum \frac{\partial d_t}{\partial\theta'}(\eta_t^2 - 1) + \sum d_t(\frac{\partial\eta_t^2}{\partial\theta'})\}\frac{1}{T}] \tag{I.4}$$

$$= \text{E}\{\sum d_t(\frac{\partial\eta_t^2}{\partial\theta'})\frac{1}{T}\}, \tag{I.5}$$

after taking iterative expectations. Thus, $B = Q$.

Theorems 2 and 3 can be proved similarly. For Theorem 2, we define $\lambda_T = (\sum d_{ti}v_{ti})/T$ and $\hat{\lambda}_T = (\sum \hat{d}_{ti}\hat{v}_{ti})/T$, where $v_{ti} = \eta_{ti}^2 - 1$ and $\hat{v}_{ti} = \hat{\eta}_{ti}^2 - 1$. By similar assumptions and

[22] The validity of this assumption can be verified following Lemma 3.3 of Ling and Li (1997) under suitable regularity conditions.

[23] Note that V_{12} does not play a role in the asymptotic distribution of $\sqrt{T}\hat{\lambda}_T$.

argument as above, $V_{11} = c_i L_i$, and

$$\begin{aligned} \mathrm{E}(\frac{\partial \lambda_T}{\partial \theta'}) &= \mathrm{E}[\{\sum \frac{\partial d_{ti}}{\partial \theta'}(\eta_{ti}^2 - 1) + \sum d_{ti}(\frac{\partial \eta_{ti}^2}{\partial \theta'})\}\frac{1}{T}] \qquad (I.6)\\ &= \mathrm{E}\{\sum d_{ti}(\frac{\partial \eta_{ti}^2}{\partial \theta'})\frac{1}{T}\}, \qquad (I.7)\end{aligned}$$

so that $B = Q_i$. For Theorem 3, we define $\lambda_T = (\sum d_{tij} v_{tij})/T$ and $\hat{\lambda}_T = (\sum \hat{d}_{tij}\hat{v}_{tij})/T$, where $v_{tij} = \eta_{ti}\eta_{tj} - \rho_{tij}$ and $\hat{v}_{tij} = \hat{\eta}_{ti}\hat{\eta}_{tj} - \hat{\rho}_{tij}$. By the martingale central limit theorem, $V_{11} = c_{ij} L_{ij}$. Also,

$$\begin{aligned} \mathrm{E}(\frac{\partial \lambda_T}{\partial \theta'}) &= \mathrm{E}[\{\sum \frac{\partial d_{tij}}{\partial \theta'}(\eta_{ti}\eta_{tj} - \rho_{tij}\ ) + \sum d_{ti}\frac{\partial(\eta_{ti}\eta_{tj} - \rho_{tij}\ )}{\partial \theta'}\}\frac{1}{T}] \qquad (I.8)\\ &= \mathrm{E}[\{\sum d_{ti}\frac{\partial(\eta_{ti}\eta_{tj} - \rho_{tij}\ )}{\partial \theta'}\}\frac{1}{T}], \qquad (I.9)\end{aligned}$$

by taking iterative expectations. Thus, $B = Q_{ij}$ and Theorem 3 holds.

References

[1] Baillie, R.T. and R.J. Myers (1991). Bivariate GARCH estimation of the optimal commodity futures hedge, *Journal of Applied Econometrics 6*, 109–124.

[2] Bera, A.K. and M.L. Higgins (1992). A test for conditional heteroscedasticity in time series models, *Journal of Time Series Analysis 13*, 501–519.

[3] Bera, A.K. and M.L. Higgins (1993). A survey of ARCH models: Properties, estimation and testing, *Journal of Economic Surveys 7*, 305–366.

[4] Bera, A.K. and S. Kim (1996). Testing constancy of correlation with an application to international equity returns. Mimeo. University of Illinois, Urbana-Champaign.

[5] Bera, A.K. and X.L. Zuo (1996). Specification test for a linear regression model with ARCH process, *Journal of Statistical Planning and Inference 50*, 283–308.

[6] Bollerslev, T. (1986). Generalized autoregressive conditional heteroscedasticity, *Journal of Econometrics 31*, 307–327.

[7] Bollerslev, T. (1990). Modelling the coherence in short-run nominal exchange rates: A multivariate generalized ARCH model, *Review of Economics and Statistics 72*, 498–505.

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

[9] Bollerslev, T., R.F. Engle, and D.B. Nelson (1994). ARCH Models. In R.F. Engle and D.L. McFadden, (Eds.), *Handbook of Econometrics, Volume 4*, pp. 2961–3038. Amsterdam: North-Holland.

[10] Bollerslev, T., R.F. Engle, and J.M. Wooldridge (1988). A capital asset pricing model with time-varying covariances, *Journal of Political Economy 96*, 116–131.

[11] Bollerslev, T. and J.M. Wooldridge (1992). Quasi-maximum likelihood estimation and inference in dynamic models with time-varying covariances, *Econometric Reviews 1*, 143–173.

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

[13] Engle, R.F. (1982). Autoregressive conditional heteroscedasticity with estimates of the variance of U.K. inflation, *Econometrica 50*, 987–1008.

[14] Engle, R.F. and K.F. Kroner (1995). Multivariate simultaneous generalized ARCH, *Econometric Theory 11*, 122–150.

[15] Engle, R.F. and V.K. Ng (1993). Measuring and testing the impact of news on volatility, *Journal of Finance 48*, 1749–1778.

[16] Fiorentini, G, G. Calzolari, and L. Panattoni (1996). Analytic derivatives and the computation of GARCH estimates, *Journal of Applied Econometrics 11*, 399–417.

[17] Glosten, L.R., R. Jagannathan, and D.E. Runkle (1993). On the relation between expected value and the volatility of the nominal excess return on stocks, *Journal of Finance 48*, 1779–1801.

[18] Hendry, D.F. (1995). *Dynamic Econometrics*. Oxford, UK: Oxford University Press.

[19] Hentschel, L. (1995). All in the family: Nesting symmetric and asymmetric GARCH models, *Journal of Financial Economics 39*, 71–104.

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

[21] Hong, Y. and R.D. Shehadeh (1999). A new test for ARCH effects and Its finite-sample performance, *Journal of Business and Economic Statistics 17*, 91–108.

[22] Karolyi, G.A. (1995). A multivariate GARCH model of international transmissions of stock returns and volatility: The case of the United States and Canada, *Journal of Business and Economic Statistics 13*, 11–25.

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

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

[25] Lundbergh, S. and T. Terasvirta (1998). Evaluating GARCH models. Mimeo. Stockholm School of Economics.

[26] Nelson, D.B. (1991) "Conditional heteroscedasticity in asset returns: A new approach, *Econometrica 59*, 347–370.

[27] Pagan, A.R. and A.D. Hall (1983). Diagnostic tests as residual analysis, *Econometric Reviews 2*, 159–218.

[28] Pierce, D.A. (1982). The asymptotic effect of substituting estimators for parameters in certain types of statistics, *Annals of Statistics 10*, 475–478.

[29] Tse, Y.K. (2000). A test for constant correlations in a multivariate GARCH model, *Journal of Econometrics 98*, 107–127.

[30] Tse, Y.K. and A.K.C. Tsui (1999). A note on diagnosing multivariate conditional heteroscedasticity models, *Journal of Time Series Analysis 20*, 679–691.

[31] Tse, Y.K. and X.L. Zuo (1997). Testing for conditional heteroscedasticity: Some Monte Carlo results, *Journal of Statistical Computation and Simulation 58*, 237–253.

[32] White, H. (1994). *Estimation, Inference, and Specification Analysis.* New York: Cambridge University Press.

[33] Wooldridge, J.M. (1990). A unified approach to robust, regression-based specification tests, *Econometric Theory 6*, 17–43.

[34] Wooldridge, J.M. (1991). On the application of robust, regression-based diagnostics to models of conditional means and conditional variances, *Journal of Econometrics 47*, 5–6.

[35] Zakoian, J.M. (1994). Threshold heteroscedastic models, *Journal of Economic Dynamics and Control 18*, 931–955.

Table 1: Empirical Size of Diagnostics for Univariate Conditional Heteroscedasticity

Error Distribution	DGP	EM	T	RB1	RB2	RB3	RB4	PR1	PR2	PR3	PR4	OLS1	OLS2	OLS3	OLS4
$N(0,1)$	M1	ARCH(1)	200	7.2	6.5	7.0	6.8	7.2	6.0	6.6	6.2	0.0	0.8	1.9	2.0
			500	7.3	6.4	6.0	6.6	7.2	6.2	6.0	6.5	0.0	1.4	1.8	2.3
			1000	5.1	4.8	6.3	5.8	5.1	4.8	6.2	6.1	0.0	1.1	1.8	1.4
	M2	ARCH(2)	200	5.1	7.0	6.3	6.3	5.1	7.2	6.3	6.1	0.0	0.1	1.0	1.4
			500	6.4	6.4	8.3	8.0	6.4	7.0	8.4	8.3	0.0	0.0	0.6	1.0
			1000	6.2	7.4	6.8	7.4	6.2	7.3	6.8	7.0	0.0	0.0	0.6	1.1
	M3	GARCH(1, 1)	200	2.9	4.5	7.6	7.8	2.7	4.5	6.0	6.4	0.0	0.9	1.8	1.7
			500	3.4	5.4	6.8	7.4	3.4	5.7	6.3	7.0	0.0	1.1	1.5	1.7
			1000	3.8	4.9	6.7	6.9	3.8	4.8	6.1	6.6	0.0	1.2	1.3	1.5
t_8	M1	ARCH(1)	200	5.1	5.9	6.2	6.3	5.0	5.7	6.4	6.1	0.4	1.7	2.1	2.7
			500	6.2	6.3	6.7	6.3	6.2	5.7	6.7	6.5	1.2	2.2	2.9	2.8
			1000	3.8	5.4	5.7	5.7	3.8	5.4	5.5	5.6	0.7	1.9	2.5	3.3
	M2	ARCH(2)	200	4.8	6.2	5.8	6.8	4.8	5.9	5.5	5.5	0.5	0.7	1.4	2.0
			500	5.0	5.7	6.5	7.4	5.0	5.8	6.5	7.3	0.5	0.8	1.7	1.9
			1000	6.2	5.9	5.4	5.1	6.2	5.5	5.1	5.5	0.6	0.3	1.1	1.2
	M3	GARCH(1, 1)	200	2.1	4.6	6.7	7.6	2.1	4.6	6.1	7.3	0.2	1.0	1.2	1.8
			500	2.5	4.7	6.3	8.0	2.5	4.8	6.0	7.4	0.3	2.1	2.8	3.5
			1000	3.8	5.4	7.3	8.3	3.8	5.4	7.0	8.1	0.3	2.0	2.4	2.5

Test Statistics: RB1–RB4, PR1–PR4, OLS1–OLS4.

Notes: DGP is the data generating process. EM is the estimated model. RBM is the residual-based diagnostic using the correct variance matrix. PRM is the Li-Mak portmanteau test. OLSM is the residual-based diagnostic using the OLS variance. We consider $M = 1, 2, 3$ and 4. The figures in the table are the empirical frequency of rejection in percentage. The nominal size of the tests is 5 percent. The estimation is based on Monte Carlo runs of 1000.

Table 2: Empirical Power of Diagnostics for Univariate Conditional Heteroscedasticity

			Test Statistics															
			$N(0,1)$ errors								t_8 errors							
DGP	EM	T	RB1	RB2	RB3	RB4	PR1	PR2	PR3	PR4	RB1	RB2	RB3	RB4	PR1	PR2	PR3	PR4
M2	ARCH(1)	200	15.7	69.3	71.4	69.8	15.6	70.5	69.9	69.3	10.6	55.7	58.1	55.7	10.6	57.1	56.2	55.7
		500	35.0	94.9	96.7	97.3	35.0	95.3	96.6	97.5	18.7	87.7	89.5	88.8	18.6	88.7	88.9	89.0
		1000	55.6	97.8	98.0	98.5	55.4	97.9	98.0	98.5	33.7	96.4	97.5	97.8	33.6	96.4	97.3	97.8
M2	GARCH(1, 1)	200	3.2	25.7	23.6	22.0	2.9	27.4	24.2	21.4	1.8	19.1	16.3	15.8	1.6	19.9	16.8	15.8
		500	14.3	59.4	53.1	48.3	14.1	62.1	56.3	51.0	2.3	35.2	30.5	30.4	2.3	36.6	32.3	30.3
		1000	40.5	88.1	84.8	80.6	40.5	89.0	86.3	83.5	11.0	58.5	52.9	50.0	10.8	60.1	56.0	51.5
M3	ARCH(1)	200	10.0	17.2	23.0	23.5	9.9	18.5	22.3	23.3	7.1	15.7	19.8	20.3	7.0	16.4	19.0	19.6
		500	15.7	42.7	57.1	58.5	15.5	43.3	55.3	59.5	6.8	27.2	40.6	41.1	6.8	28.7	39.0	40.0
		1000	24.0	69.1	84.4	84.8	24.0	70.6	83.1	85.1	10.9	47.0	64.1	66.8	10.7	48.7	64.1	66.8
M3	ARCH(2)	200	6.3	8.2	9.2	7.8	6.3	8.2	8.3	6.7	4.7	6.1	6.7	7.3	4.7	5.8	6.0	7.1
		500	7.9	10.3	10.0	9.4	7.8	10.1	9.9	9.1	5.4	7.2	7.8	8.4	5.4	6.9	7.6	8.6
		1000	6.2	10.0	12.0	12.0	6.0	9.7	11.4	11.8	3.8	5.8	8.4	10.0	3.7	5.5	8.5	9.5

Notes: DGP is the data generating process. EM is the estimated model. RBM is the residual-based diagnostic using the correct variance matrix. PRM is the Li-Mak portmanteau test. We consider $M = 1, 2, 3$ and 4. The figures in the table are the empirical frequency of rejection in percentage. The nominal size of the tests is 5 percent. The estimation is based on Monte Carlo runs of 1000.

Table 3: Empirical Size and Power of Diagnostics for Multivariate Conditional Heteroscedasticity

Error Distribution	DGP	T	LL1	LL2	LL3	LL4	R1-1	R1-2	R1-3	R1-4	R2-1	R2-2	R2-3	R2-4	R3-1	R3-2	R3-3	R3-4
$N(0,1)$	CC-MGARCH	500	6.1	6.1	7.5	8.2	4.4	5.7	6.0	7.2	4.0	5.5	6.0	7.2	5.6	6.0	6.5	7.5
		1000	5.1	5.2	5.5	6.7	6.8	6.2	6.9	7.7	5.9	5.7	5.8	6.0	5.1	5.0	5.9	6.5
	BEKK(D)	500	9.7	16.5	19.6	20.8	8.5	14.8	18.6	20.1	9.4	12.8	16.9	17.9	52.4	64.8	70.8	72.0
		1000	14.5	20.2	23.7	24.4	10.7	16.2	21.3	26.1	10.5	16.0	21.3	24.6	82.3	90.7	91.9	93.2
	BEKK	500	8.5	9.9	8.8	10.0	35.5	43.8	42.5	40.5	37.2	43.7	40.6	37.4	79.0	81.0	78.3	74.5
		1000	9.4	9.8	9.2	10.2	59.0	63.9	60.7	58.4	60.3	65.7	63.4	60.5	96.5	97.0	95.8	93.8
	DVR	500	6.7	7.4	7.9	9.2	3.5	4.6	5.5	7.0	4.6	5.5	5.8	8.1	26.8	23.9	22.5	20.8
		1000	5.1	6.1	7.0	8.1	4.1	5.3	5.7	6.8	4.7	5.3	5.1	6.5	53.7	51.0	45.6	40.8
t_8	CC-MGARCH	500	5.6	6.0	8.8	11.0	3.2	5.0	6.4	7.2	3.1	4.9	5.8	7.8	4.9	6.7	7.1	7.4
		1000	4.6	5.5	7.1	8.6	3.9	4.5	5.8	6.6	5.0	6.0	6.9	7.8	4.8	5.9	5.6	6.6
	BEKK(D)	500	7.2	13.7	17.2	18.9	4.5	9.7	11.6	12.7	7.4	8.6	10.5	12.2	40.1	52.6	55.7	57.4
		1000	9.8	13.8	18.9	20.7	7.3	10.3	11.8	15.5	7.1	9.5	11.5	14.4	65.9	78.3	81.7	81.7
	BEKK	500	9.3	13.9	14.4	15.8	21.1	27.1	25.6	24.1	22.8	29.7	26.6	24.0	69.4	60.5	57.1	54.3
		1000	8.7	10.7	11.3	10.6	30.8	39.0	37.1	36.5	36.7	43.1	41.1	38.5	83.0	83.2	81.0	79.0
	DVR	500	5.0	8.2	10.5	11.8	2.8	4.8	5.6	7.7	3.1	5.2	6.2	7.6	21.9	22.5	20.6	19.0
		1000	5.5	6.9	7.7	8.8	3.0	5.4	3.6	6.7	4.4	5.0	8.2	7.5	42.3	39.6	35.0	31.7

Test Statistics: LL1–R3-4.

Notes: DGP is the data generating process. LLM is the Ling-Li test based on M autocorrelation coefficients. R1-M, R2-M and R3-M are the residual-based diagnostics with, respectively, $\hat{\eta}_{t1}^2 - 1$, $\hat{\eta}_{t2}^2 - 1$ and $\hat{\eta}_{t1}\hat{\eta}_{t2} - \hat{\rho}$ as the dependent variables in the artificial regression, where $M = 1$, 2, 3 and 4 stands for the order of the lagged terms taken in the regression. The figures in the table are the empirical frequency of rejection in percentage at the nominal size of 5 percent. The estimated model is CC-MGARCH. When the DGP is CC-MGARCH, the figures are the empirical size. Otherwise, the figures are the empirical power. The estimation is based on Monte Carlo runs of 1000. The model parameters can be found in Section 3.2.