The University of Manchester

MANCHESTER 1824

Economics
Discussion Paper Series
EDP-1109

An investigation of parametric tests of CCC assumption

Wasel Shadat

Chris Orme

February 2011

Economics
School of Social Sciences
The University of Manchester
Manchester M13 9PL

An Investigation of Parametric Tests of the CCC Assumption*

Wasel B Shadat†‡ and Chris D Orme§
University of Manchester

February 11, 2011

Abstract

An asymptotically valid Conditional Moment (CM) testing procedure of the CCC assumption of the MGARCH model is proposed considering both full QMLE (FQMLE) and partial or two-stage QMLE (PQMLE) framework. A "new" and easily programmable expression for the expected Hessian is provided for FQMLE. The OPG and robust to non-normality versions of the test statistics are derived. The Tse (2000) OPG-type LM test of the CCC assumption is analyzed within our CM framework and a new robust version of this test is proposed. An extensive Monte Carlo investigation demonstrates good size and power properties. The OPG versions suffer from size distortion under non-normality whereas robust versions perform better.

JEL classification: C12, C32
Keywords: Multivariate GARCH models; Constant Conditional Correlation; Conditional Moment tests, Monte Carlo experiment

1 Introduction

Applied researchers have increasingly been using the conditional correlation approach to model multivariate volatility through a multivariate GARCH (MGARCH) model. Although the Dynamic Conditional Correlation (DCC) model is by far the most popular specification among applied researchers, a good number of empirical research applies the Constant Conditional Correlation (CCC) model; see for example, Bollerslev (1990), Kroner and Claessens (1991), Kroner and Sultan

*We are grateful for the insightful comments of Ralf Becker, Alastair Hall and Len Gill, which greatly improved the exposition of this paper. The standard disclaimer applies.
†Address correspondence to Wasel Shadat, Economics, School of Social Science, University of Manchester, Manchester M13 9PL, UK; e-mail: wbshadat@yahoo.com.
‡The first author's research is part of his PhD thesis and was supported by a Commonwealth Scholarship and Fellowship Plan and a Manchester School Award; both of which are gratefully acknowledged.
§Economics, School of Social Science, University of Manchester, Manchester M13 9PL, UK

(1991), Kroner and Sultan (1993), Park and Switzer(1995) and Lien and Tse (1998). Due to the simplicity and computational advantages of the (CCC) model compared to that of the DCC model, on the one hand, but the restrictiveness of the CCC assumption on the other, testing the adequacy of the CCC-GARCH model is very important both from practical and theoretical point of view. The validity of the CCC assumption remains an empirical question. However, only a few tests of this assumption have been proposed in the literature.

To test the CCC assumption, Bollerslev (1990) suggested some diagnostics applying Ljung-Box portmanteau test statistics based on the cross-products of the standardized residuals obtained from the CCC-GARCH model. The idea is that if the CCC assumption is valid, then these crossproducts should also be serially uncorrelated. He found that the standardized residuals are uncorrelated in case of five European countries' monthly exchange rate and suggested that this provided evidence of constancy of the correlations. However, serially uncorrelated standardized residuals implies they are linearly independent over time and does not guarantee that the conditional correlations are constant over time. Further, critical values for this test procedure were based on a χ^2 distribution whereas Li and Mak (1994) pointed out that the portmanteau statistic is not asymptotically χ^2 and the use of a χ^2 approximation is inappropriate. Bollerslev (1990) used another diagnostic based on an artificial regression involving the products of the standardized residuals. In this case, however, there are usually no sufficient guidelines as to the choice of regressors in the artificial regression. Furthermore, the optimality of portmanteau and residual based tests is not established. Therefore, there remains the question of how powerful these tests are against dynamic conditional correlation.

Longin and Solnik (1995) suggested another test by taking pairs of variables at a time; explicitly specifying the conditional correlation as a function of potential sources of deviation from constant correlation and then testing the significance of the associated parameters.[1] However, their alternative correlation specification is not guaranteed to be bounded by -1 and 1 (i.e. $|\rho| \leq 1$). This would appear to be a crucial defect. In their empirical application with monthly excess returns of stock markets of seven major countries from 1960 to 1990, they considered three sources of deviation: a time trend, the presence of threshold and influence of related economic variables (dividend yields and interest rates) and found that the correlation was increased over time and related to dividend yields and interest rates implying the rejection of the CCC hypothesis.

Bera and Kim (2002) developed a test of a bivariate CCC-GARCH model against the alternative that the correlation coefficient is random (over time). This test is an Information Matrix (IM) test (White, 1982) in the form of an LM or score test of random variation in correlation parameter ρ; see Chesher (1984) and Cox (1983). The null hypothesis of this score test is that the variance

[1] For example, they specify the conditional covariance between two assets as follows:

$$h_{12t} = (\rho_0 + \rho_1 x_{1t} + \cdots + \rho_r x_{rt}) \sqrt{h_{11t} h_{22t}}$$

where h_{iit} is the conditional variance of i^{th} asset $i = 1, 2$, x_{it}'s are possible sources of deviation. The CCC assumption corresponds to the null $H_0 : \rho_1 = \cdots = \rho_r = 0$.

of the parameter of interest is zero and the test checks the local behavior of the log-likelihood function close to the null of no parameter variation. It does not check the CCC assumption directly. Secondly, this test is not robust to non-normality. Thirdly, this test is derived for bivariate case only, limiting its applicability in high dimensional cases. Finally, the IM test assesses several features of the model. Bera and Kim (2002, p.182) also recognize the fact that "ability of the IM test principle to check various feature of the underlying model might be viewed as a drawback rather than an advantage".

However, all the above-mentioned tests are not specifically designed for testing CCC assumption and in practice they may not be very helpful to address this issue. Tse (2000) proposed a LM test of the CCC assumption. This is a multivariate test in a true sense and, among applied workers, the most widely used test of CCC assumption until now (see, for example,Tse (2000), Lien, Tse and Tsui (2002), Andreou and Ghysels (2003), Lee (2006), Aslanidis, Osborn and Sensier (2008) among others). This test involves the Full QMLE (FQMLE) approach i.e. simultaneous estimation of the volatility and correlation parameters under the null of CCC. Therefore it might not be robust to GARCH misspecifications in individual volatility equations. Moreover, Tse uses the OPG version of the LM test which is based on the normality assumption; therefore it may demonstrate relatively poor finite sample properties and may not be robust under non-normality (see, for example, Davidson and MacKinnon, 1983; Bera and McKenzie, 1986; Chesher and Spady, 1991). Finally the time varying alternative specification of correlation matrix as presented by Tse is not necessarily a positive definite matrix for all t. For this reason Silvennoinen and Teräsvirta (2008) interpreted this test as a general misspecification test. In a recent paper, Nakatani and Teräsvirta (2009) proposed a LM test for volatility interaction where the null model is CCC GARCH model against the alternative of Extended CCC (ECCC) Garch model.

Nevertheless it is evident that the field of testing CCC assumption is relatively under-developed compared to other aspects of the MGARCH literature. The aim of this study is to put forward some alternative asymptotically valid testing strategies of the CCC assumption. Firstly, we present and review a conditional moment (CM) testing framework based on the FQMLE of null CCC model. However, in practice while estimating a MGARCH model adopting the conditional correlation approach (both constant and dynamic, but particularly for the dynamic one) researchers use a two-step or Partial QMLE (PQMLE) approach; where in the first stage the volatility parameters are estimated using univariate GARCH specification for individual variables and the correlation parameters are estimated using the volatility parameter estimates obtained in the first stage (see, Engle and Sheppard, 2008; Hafner, Dijk and Franses, 2005; Billio, Caporin and Gobbo, 2006; among others). There appears to be no testing approach of CCC assumption available in the literature which allows partial estimation. The implication of this is one has to first estimate FQMLE of the null CCC model in order to test the null CCC assumption; and if the null is rejected the researcher needs to use DCC specification which generally use two-step estimation procedure. Again, there is a well-developed literature which deals with the specification testing for UGARCH models and their asymptotic properties.[2] These two facts motivate us to develop asymptotically valid CM tests of the CCC assumption based on two-step estimation and utilizing UGARCH results. The second contribution of this research is to devise a simple test after PQML estimation. Thirdly, both the OPG and robust versions of the tests are developed. The proposed tests (both FQMLE and PQMLE) are easy to implement and demonstrate satisfactory size and good power properties in the simulation experiments. Fourthly, we derive a "new" expression for the average Hessian of the CCC GARCH regression model which is easy to programme. Finally, we have analyzed Tse's LM test within our CM testing framework and suggested a robust version of this which demonstrate superior size properties under non-normality.

The rest of this paper is organized in the following way. The conditional correlation approach for MGARCH model specification with the estimation framework is presented in Section 2. In Section 3, a class of parametric tests with their asymptotic properties is described. An analysis of Tse's LM test is presented in the next Section. Section 5 provides some Monte Carlo evidence and Section 6 concludes. The proof of lemmas, propositions and theorems are relegated to Appendix. Throughout we make use of the following notations: $\mathrm{E}_0(.)$ and $\mathrm{E}_{t-1}(.)$ denote the expectation with respect to true parameter value and conditional on previous history up to $t-1$ respectively; $\otimes$ and $\odot$ denote the Kronecker and Hadamard product respectively; $\mathrm{vech}(.)$ and $\mathrm{vecl}(.)$ denote the operator that stacks the lower triangular portion of a $(N \times N)$ matrix as a $\left(\frac{N(N+1)}{2} \times 1\right)$ vector and the strictly lower triangular portion of a $(N \times N)$ matrix as a $\left(\frac{N(N-1)}{2} \times 1\right)$ vector respectively; $I_N = \{\delta_{ik}\}$, is the identity matrix of order N where δ_{ik} is the kronecker delta; $\iota_K' = (1, 1, ..., 1)$, is $(1 \times K)$ vector of ones and $\mathcal{J}_K = \iota_K \iota_K'$, is the $(K \times K)$ matrix of ones.

2 The Null Constant Conditional Correlation Model

Suppose we are interested in the $(N \times 1)$ time-series vector $\{y_t\} = (y_{1t}, \cdots, y_{Nt})'$ and $\mathcal{F}_{t-1} = \sigma\left(W_t', W_{t-1}', \cdots\right)$ is the $\sigma-$field generated by the past information

[2] For example, Lundbergh and Terasvirta (2002) proposed a parametric Lagrange multiplier (LM) type tests of no ARCH effect in standardized errors, linearity, and parameter constancy. Testing for leverage effect developed by Engle and Ng (1993) is widely used in empirical finance. Bollerslev (1986) presented another LM-type test for testing a GARCH model against a higher order GARCH model. One important work in this field is of Halunga and Orme's (2009) unifying parametric testing framework based on the CM principal which takes into account the asymptotically non-negligible estimation effect from the conditional mean parameters. This is the major point of departure of the Halunga and Orme's (2009) test with that of the abovementioned tests. They demonstrated that these tests are asymptotically invalid in the regression context and may have low power. A Monte Carlo study also showed better empirical power properties of their proposed test than those of Engle and Ng (1993) and Lundbergh and Teräsvirta (2002).

up to and including time $t-1$. We consider the following CCC-GARCH specification to model this series:

$$\begin{aligned}
y_t &= m\left(W_t;\varphi\right)+\varepsilon_t \;\; t=1,...,T \\
\varepsilon_t &= H_t^{1/2}\left(\varpi\right)\xi_t \\
H_t &= D_t \Gamma D_t \\
D_t &= diag(h_{11t}^{1/2},.....,h_{NNt}^{1/2}) \qquad (1)
\end{aligned}$$

where $\varphi' = (\varphi_1', \cdots, \varphi_N')$; $\varphi_i \in \Psi \subset \Re^K$ is a $(NK \times 1)$ vector of conditional mean parameters and W_t' is the $(N \times NK)$ data matrix of the t-th observation; $H_t^{1/2}\left(\varpi\right)$ is a $(N \times N)$ positive definite matrix such that $H_t = Var\left(\varepsilon_t|\mathcal{F}_{t-1}\right)$ and ϖ is the vector of unknown parameters which includes conditional mean parameter φ as well (for notational convenience, we drop ϖ in $H_t^{1/2}\left(\varpi\right)$), D_t is a $(N \times N)$ diagonal matrix of conditional standard deviation, $\Gamma = [\rho_{ij}]$ is a time invariant symmetric positive definite conditional correlation matrix with $\rho_{ii} = 1, i = 1, ..., N$. $m\left(W_t;\varphi\right)$ can possibly be nonlinear and W_t contains current and lagged exogenous variables, and lagged dependent variables. However, for simplicity of exposition, we assume a linear specification for the conditional mean function i.e. $m\left(W_t;\varphi\right) = W_t'\varphi$ so that the conditional mean function becomes $y_t = W_t'\varphi + \varepsilon_t \;\; t = 1, ..., T$.The stochastic sequence $\{\xi_t\}$ is an i.i.d. process with $\mathrm{E}(\xi_t) = 0$ and $Var(\xi_t) = \mathrm{E}(\xi_t\xi_t') = I_N$. We further assume that given the $\sigma-$field generated by the past information up to and including time $t-1$, $\mathcal{F}_{t-1} = \sigma\left(W_t', W_{t-1}', \cdots\right)$, the error $\{\varepsilon_t, \mathcal{F}_{t-1}\}$ is a MDS.

With these assumptions,

$$\mathrm{E}\left[\varepsilon_t|\mathcal{F}_{t-1}\right] = 0; \text{ and } \mathrm{E}\left[\varepsilon_t\varepsilon_t'|\mathcal{F}_{t-1}\right] = H_t = \begin{cases} h_{it} & i = j \\ h_{it}^{1/2}h_{jt}^{1/2}\rho_{ij} & i \neq j \end{cases};$$

Also note that $corr\left[\varepsilon_{it}, \varepsilon_{jt}|\mathcal{F}_{t-1}\right] = \rho_{ij} = \dfrac{h_{ij,t}}{\sqrt{h_{it}}\sqrt{h_{jt}}}$ and $\mathrm{E}\left[\varepsilon_t\varepsilon_{t-j}'|\mathcal{F}_{t-1}\right] = \mathrm{E}\left[\varepsilon_t|\mathcal{F}_{t-1}\right]\varepsilon_{t-j}' = 0$, almost surely, for all $j \geq 1$. The CCC models uses the following classical decomposition of H_t to achieve a parsimonious way to model H_t (compared to direct modelling approach):

$$H_t = D_t \Gamma D_t \qquad (2)$$

where $h_{it},\ i = 1, ..., N$ can be defined by any univariate GARCH model and $\Gamma_t = [\rho_{ijt}]$ is a symmetric positive definite matrix with $\rho_{iit} = 1,\ i = 1, ..., N$. (2) implies that the diagonal elements of the conditional covariance matrix are simply the conditional variances while the off-diagonal elements are $h_{ijt} = h_{it}^{1/2}h_{jt}^{1/2}\rho_{ij}, \quad i \neq j, 1 \leq i, j \leq N$.

Here we assume that each $h_{it},\ i = 1, ..., N$ has a GARCH (p, q) specification

$$h_{it} = \eta_i' s_{i,t-1} = \alpha_{i0} + \sum_{k=1}^{q} \alpha_{ik}\varepsilon_{i,t-k}^2 + \sum_{j=1}^{p} \beta_{i,j}h_{i,t-j} \qquad (3)$$

Denoting $h_t = (h_{1t}, \cdots, h_{Nt})'$, we can write

$$h_t = a_0 + \sum_{k=1}^{q} \widetilde{A}_k \overrightarrow{\varepsilon}_{t-k} + \sum_{j=1}^{p} \widetilde{B}_j h_{t-j}$$

where $\widetilde{A}_k$ and.$\widetilde{B}_j$ are both $(N \times N)$ diagonal matrix and a_0 and $\overrightarrow{\varepsilon}_t = \left(\varepsilon_{1t}^2, \cdots, \varepsilon_{Nt}^2\right)'$ are $(N \times 1)$ vector.

In conditional correlation MGARCH models, standardized errors, play a crucial role. We shall term the three types of standardized errors that will appear in subsequent analysis as standardized errors, fully standardized errors and Tse's modified errors and defined as:

$$\begin{aligned}
\zeta_t &= D_t^{-1}\varepsilon_t;\ \mathrm{E}\left[\zeta_t|\mathcal{F}_{t-1}\right] = 0, \mathrm{E}\left[\zeta_t\zeta_t'|\mathcal{F}_{t-1}\right] = \Gamma. && (4)\\
\xi_t &= H_t^{-1/2}\varepsilon_t;\ \mathrm{E}\left[\xi_t|\mathcal{F}_{t-1}\right] = 0;\ \mathrm{E}\left[\xi_t\xi_t'|\mathcal{F}_{t-1}\right] = I_N. && (5)\\
\varepsilon_t^* &= \Gamma^{-1}\zeta_t = \Gamma^{-1}D_t^{-1}\varepsilon_t;\ \mathrm{E}\left[\varepsilon_t^*|\mathcal{F}_{t-1}\right] = 0, \mathrm{E}\left[\varepsilon_t^*\varepsilon_t^{*\prime}|\mathcal{F}_{t-1}\right] = \Gamma^{-1}. && (6)
\end{aligned}$$

2.1 Full QMLE (FQMLE) Estimation Framework

We will start by defining and introducing some notations which will be useful when deriving the expressions for scores and expected Hessian. Without loss of generality we assume that each variable correspond to a parameter vector of same dimension, i.e. for the $i-$th variable, $i = 1, \cdots, N$, define $\theta_i = (\varphi_i', \eta_i')' \subset \Re^{K+K'}$ with $\varphi_i \subset \Re^K$ (corresponding to conditional mean function) and $\eta_i \subset \Re^{K'}$ (corresponding to volatility function). Hence, $\theta = \left(\theta_1', ..., \theta_N'\right)' \subset \Re^{N\left(K+K'\right)}$ is the parameter vector consisting of conditional mean and volatility parameters for N variables and $\rho \subset \Re^{\frac{N(N-1)}{2}}$ is the vector of distinct correlation parameters. Then define the collection of all parameters $\varpi = \left(\theta', \rho'\right)' \in \Theta \subset \Re^{N'}$ where $N' = N\left(K + K'\right) + \frac{N(N-1)}{2}$.[3]

For the i^{th} variable, define $\underset{(T\times K)}{F_i}$, $\underset{(T\times K)}{C_i}$ and $\underset{(T\times K')}{X_i}$ with rows $f_{it}' = \dfrac{w_{it}'}{\sqrt{h_{it}}}$; $c_{it}' = \dfrac{1}{h_{it}}\dfrac{\partial h_{it}}{\partial \varphi_i'}$ and $x_{it}' = \dfrac{1}{h_{it}}\dfrac{\partial h_{it}}{\partial \eta_i'}$ respectively. Then define

$$\begin{aligned}
\underset{(NT\times NK)}{F} &= diag\left(F_i\right),\\
\underset{(N\times NK)}{F_t'} &= diag\left(f_{it}'\right) \text{ for } t = 1, \cdots, T.
\end{aligned}$$

In a similar way, define C, X, C_t' and X_t' matrices. It will be useful to define $\underset{(N\times N)}{E_t} = diag\left(\zeta_{it}\right); \Gamma_A = I_N + \left(\Gamma^{-1} \odot \Gamma\right)$ and Γ^{-1} has a typical element ρ^{ij}. Finally, let ρ^k be the k^{th} column of Γ^{-1}; define $\Gamma^k = \Gamma^{-1} diag\left(\tau_k\right),$ where $\tau_k = \{\delta_{ik}\}, \ (N \times 1), \ i = 1, ..., N$; i.e. Γ^k be the $(N \times N)$ matrix of zeros,

[3]For example, AR(1)-Bivariate CCC specification with GARCH (1,1) model for individual volatility have we have $N = 2$, $K = 2$, $K' = 3$ and $N' = 11$.

except for column k which is ρ^k. Define the following two $(N \times N)$ symmetric matrices:

$$\begin{aligned} P_k &= \Gamma^k + \left(\Gamma^k\right)' \\ \Gamma_{km} &= \rho^k \left(\rho^m\right)' + \rho^m \left(\rho^k\right)'. \end{aligned}$$

2.1.1 The Score, Hessian and limit distribution of the FQMLE

Under the assumption of conditional normality, define the average log-likelihood function as $L_T^*(\varpi) = \dfrac{1}{T} \sum l_t^*(\theta, \rho)$, where $l_t^* \equiv l_t^*(\theta, \rho)$ is the quasi-conditional log-likelihood per observation, t, (ignoring any constant terms) which can be written as,

$$l_t^* = -\frac{1}{2} \ln |\Gamma| - \frac{1}{2} \sum_{j=1}^{N} \ln h_{jt} - \frac{1}{2} \zeta_t' \Gamma^{-1} \zeta_t \tag{7}$$

where, $H_t \equiv H_t(\varpi)$, $D_t \equiv D_t(\theta)$.[4] The parameter estimates can be obtained by quasi maximum likelihood (QML) method:

$$\hat{\varpi} = \arg\max_{\varpi} \sum_{t=1}^{T} l_t^*.$$

Assuming $L_T^*(\varpi) = T^{-1} \sum_{t=1}^{T} l_t^*(\theta, \rho)$ is at least twice continuously differentiable, define the average score for CCC model $G_T^*(\varpi) = T^{-1} \sum_{t=1}^{T} g_t^*(\varpi)$ where $g_t^*(\varpi) = \left(\dfrac{\partial l_t^*}{\partial \theta'}, \dfrac{\partial l_t^*}{\partial \rho'} \right)' = \left(\dfrac{\partial l_t^*}{\partial \varphi'}, \dfrac{\partial l_t^*}{\partial \eta'}, \dfrac{\partial l_t^*}{\partial \rho'} \right)'$ and S^* as a $(T \times N')$ matrix with rows $g_t^{*\prime}(\varpi)$.Using the similar notation, define the Hessian of the log-likelihood function for observation t as $\mathcal{H}_t^*(\varpi) = \frac{\partial^2 l_t^*}{\partial \varpi \partial \varpi'} = \frac{\partial g_t^*(\varpi)}{\partial \varpi'}$. The expression for $g_t^*(\varpi)$ is provided in Lemma 1.

Lemma 1 *The score vector for observation t of (7), $g_t^*(\varpi) = \left(\dfrac{\partial l_t^*}{\partial \theta'}, \dfrac{\partial l_t^*}{\partial \rho'} \right)' = \left(\dfrac{\partial l_t^*}{\partial \varphi'}, \dfrac{\partial l_t^*}{\partial \eta'}, \dfrac{\partial l_t^*}{\partial \rho'} \right)'$ is given by*

$$\begin{aligned} \frac{\partial l_t^*}{\partial \varphi} &= F_t \Gamma^{-1} \zeta_t + \frac{1}{2} C_t \left\{ E_t \Gamma^{-1} \zeta_t - \iota_N \right\} = F_t \varepsilon_t^* + \frac{1}{2} C_t \left\{ E_t \varepsilon_t^* - \iota_N \right\} \\ \frac{\partial l_t^*}{\partial \eta} &= \frac{1}{2} X_t \left\{ E_t \Gamma^{-1} \zeta_t - \iota_N \right\} = \frac{1}{2} X_t \left\{ E_t \varepsilon_t^* - \iota_N \right\} \\ \frac{\partial l_t^*}{\partial \rho_{ij}} &= vecl\left(M_t\right) = m_{ij,t}, \quad j < i = 2, ..., N \quad (\textit{with } \rho_{ii} \equiv 1) \end{aligned} \tag{8}$$

where , E_t, F_t, C_t and X_t defined earlier and $M_t = \{m_{ij,t}\} = \Gamma^{-1} \left(\zeta_t \zeta_t' - \Gamma\right) \Gamma^{-1}$.

[4] Note that we make use of asterisk (*) to differentiate joint log-likelihood from the univariate GARCH log-likelihood.

The FQMLE $\hat{\varpi}' = \left(\hat{\theta}', \hat{\rho}'\right)$ satisfies $G^{*}_{T}(\hat{\varpi}) = 0$. Bollerslev and Wooldridge (1992) showed that under regularity conditions the conditional heteroskedasticity FQML estimators are consistent and asymptotically normal. However, they did not verify whether the regularity conditions hold for specific MGARCH model. Jeantheau (1998) gave conditions for strong consistency of FQMLE for MGARCH and verified the conditions for extended CCC (ECCC) model. Comte and Lieberman (2003) proved the strong consistency and asymptotic normality of QMLE (both when initial state is stationary or fixed) for the BEKK MGARCH specification which requires the finiteness of the moments of the non-Gaussian process ε_t up to order 8 i.e. $\mathrm{E}\left[\varepsilon_{i,t}^{8}\right] < \infty$, $i = 1, \cdots, , N$. Ling and McAleer (2003) presented a theoretical framework for a class of vector ARMA-GARCH models with ECCC specification for the conditional heteroskedasticity and their conditions require $\mathrm{E}\left[\varepsilon_{i,t}^{6}\right] < \infty$, $i = 1, \cdots, , N$. Since the CCC model is nested within this class, we can make use of the following results. Following Ling and McAleer (2003) and Nakatani and Teräsvirta (2009), to ensure the asymptotic normality of QMLE $\hat{\varpi}$ we assume that the followings to hold:

Assumption 2.1 *The elements* (y_{it}, W'_{it}) *are strictly stationary and ergodic for all* $i = 1, \cdots, N$*; and* $m(W_{it}; \varphi_i)$ *is continuous and* $\mathcal{F}_{t-1}$*-measurable for all* $\varphi_i \in \Psi \subset \Re^{K}$.

Assumption 2.2 *The spectral radius* $\varsigma(\Gamma)$ *has a positive lower bound over the parameter space* Θ *which is a compact subset of the Euclidean space such that* ϖ_0 *lie in the interior of* Θ*. In addition each element of* a_0 *has a positive lower and upper bounds over* Θ.

Assumption 2.3 *All the roots of* $\det\left(I_N - \sum_{k=1}^{q} \widetilde{A}_k x^k - \sum_{j=1}^{p} \widetilde{B}_j x^j\right)$ *lie outside the unit circle.*

Assumption 2.4 *The identifiability conditioned presented in Jeantheau (1998) are satisfied.*

Assumption 2.5 $\mathrm{E}\left[\varepsilon_{i,t}^{6}\right] < \infty$, $i = 1, \cdots, N$.

Assumption 2.6 $\mathrm{plim}_{T\to\infty} \frac{1}{T} \sum_{t=1}^{T} \mathcal{H}^{*}_{t}(\varpi)$ *exists and finite for all* $\varpi \in \Theta$ *such that the* $N \times N$ *nonrandom matrix*

$$J^{*}_{\varpi\varpi} = -\,\mathrm{E}_0\left[\mathcal{H}^{*}_{t}(\varpi_0)\right] = \operatorname*{plim}_{T\to\infty} -\frac{1}{T} \sum_{t=1}^{T} \mathcal{H}^{*}_{t}(\varpi_0).$$

Theorem 1 *Given these assumptions,* $\hat{\varpi} \xrightarrow{p} \varpi_0$ *and*

$$\sqrt{T}(\hat{\varpi} - \varpi_0) \xrightarrow{d} N\left(0, J^{*-1}_{\varpi\varpi} \Sigma^{*}_{GG} J^{*-1}_{\varpi\varpi}\right)$$

where $J^{*}_{\varpi\varpi} = -\,\mathrm{E}_0\left[\mathcal{H}^{*}_{t}(\varpi_0)\right]$ *and* $\mathbf{\Sigma}^{*}_{GG} = \mathrm{E}_0\left[g^{*}_{0t} g^{*\prime}_{0t}\right]$ *are both finite and positive definite and* $\mathrm{E}_0[.]$ *denotes expectation evaluated at the true parameter values* $\varpi_0 = \left(\theta'_0,\ \rho'_0\right)'$.

The matrix $J^*_{\varpi\varpi}$ is the negative of the expected Hessian while Σ^*_{GG} is the expectation of the outer product of the score vector both evaluated at ϖ_0 and the later is often called the population information matrix. Moreover, if $\xi_t \sim N(0, I_N)$, then $\Sigma^*_{GG} = J^*_{\varpi\varpi}$ and the asymptotic covariance matrix reaches to the Cramer-Rao lower bound i.e. Σ^{*-1}_{GG}. Note that by the consistency of the QMLE $\hat{\varpi}$, $J^*_{\varpi\varpi}$ can be consistently estimated by $\hat{J}^*_{\varpi\varpi T} = -\frac{1}{T}\sum_{t=1}^{T}\mathcal{H}^*_t(\hat{\varpi}) = -\frac{1}{T}\sum_{t=1}^{T}\frac{\partial^2 l_t}{\partial\varpi\partial\varpi'}\Big|_{\varpi=\hat{\varpi}}$. Note that by definition $H_t(\varpi_0) = \mathrm{E}_0(\varepsilon_t\varepsilon_t'|\mathcal{F}_{t-1})$ which implies that it would be computationally easier to work with $\widetilde{\mathcal{H}}^*_t(\varpi_0) = \mathrm{E}_0[\mathcal{H}^*_t(\varpi_0)|\,\mathcal{F}_{t-1}]$ (say); as under conditional expectation operator a number of terms in $\frac{\partial^2 l_t}{\partial\varpi\partial\varpi'}$ cancel when evaluated at ϖ_0. Further by the law of iterated expectation we have $J^*_{\varpi\varpi} = -\mathrm{E}_0[\mathrm{E}_0[\mathcal{H}^*_t(\varpi_0)|\,\mathcal{F}_{t-1}]]$ and a simpler estimate of $J^*_{\varpi\varpi}$ is obtained as $\hat{J}^*_{\varpi\varpi T} = -\frac{1}{T}\sum_{t=1}^{T}\widetilde{\mathcal{H}}^*_t(\hat{\varpi})$.[5] The Hessian can be derived with reference to Nakatani and Teräsvirta (2009, 2008) who provide the general expression of $\widetilde{\mathcal{H}}^*_t(\varpi)$ for ECCC-GARCH model. However these authors derive the expression assuming a known or zero conditional mean. Besides, they did not specify any particular D_t.

The following Lemma 2 provides a *new* expression for $\widetilde{\mathcal{H}}^*_t(\varpi_0)$ in the regression context, which considers the conditional mean function and GARCH (p,q) specification for individual conditional variances in D_t.While Lemma 3 provides the expression for $\hat{J}^*_{\varpi\varpi T}$ which will be required in our tests discussed in the next section.

Lemma 2 $\widetilde{\mathcal{H}}^*_t(\varpi_0) = \mathrm{E}_0[\mathcal{H}^*_t(\varpi_0)|\,\mathcal{F}_{t-1}]$ *where*

$$\widetilde{\mathcal{H}}^*_t(\varpi) = \begin{bmatrix} \widetilde{\mathcal{H}}^*_{\varphi\varphi} & \widetilde{\mathcal{H}}^*_{\varphi\eta} & \widetilde{\mathcal{H}}^*_{\varphi\rho} \\ \mathcal{H}^{*\prime}_{\varphi\eta} & \widetilde{\mathcal{H}}^*_{\eta\eta} & \widetilde{\mathcal{H}}^*_{\eta\rho} \\ \widetilde{\mathcal{H}}^{*\prime}_{\varphi\rho} & \widetilde{\mathcal{H}}^{*\prime}_{\eta\rho} & \widetilde{\mathcal{H}}^*_{\rho\rho} \end{bmatrix}.$$

and the typical (i,j)-th block of $\widetilde{\mathcal{H}}^*_{\varphi\varphi}$, $\widetilde{\mathcal{H}}^*_{\varphi\eta}$, $\widetilde{\mathcal{H}}^*_{\varphi\rho}$, $\widetilde{\mathcal{H}}^*_{\eta\eta}$, $\widetilde{\mathcal{H}}^*_{\eta\rho}$ and $\widetilde{\mathcal{H}}^*_{\rho\rho}$; $i,j = 1,\cdots,N$ are given as, respectively:

$$\begin{aligned}
\widetilde{\mathcal{H}}^*_{\varphi_i\varphi_j} &= -\rho^{ij} f_{it} f'_{jt} - \frac{1}{4}\left(\delta_{ij} + \rho^{ij}\rho_{ij}\right) c_{it} c'_{jt}; \\
\widetilde{\mathcal{H}}^*_{\varphi_i\eta_j} &= -\frac{1}{4}\left(\delta_{ij} + \rho^{ij}\rho_{ij}\right) c_{it} x'_{jt}; \\
\widetilde{\mathcal{H}}^*_{\varphi_i\rho_{ij}} &= -\frac{1}{2}\delta_{jk}\rho^{ik} c_{kt} - \frac{1}{2}\delta_{ik}\rho^{jk} c_{kt};\ i > j \\
\widetilde{\mathcal{H}}^*_{\eta_i\eta_j} &= -\frac{1}{4}\left(\delta_{ij} + \rho^{ij}\rho_{ij}\right) x_{it} x'_{jt}. \\
\widetilde{\mathcal{H}}^*_{\eta_i\rho_{ij}} &= -\frac{1}{2}\delta_{jk}\rho^{ik} x_{kt} - \frac{1}{2}\delta_{ik}\rho^{jk} x_{kt};\ i > j \\
\widetilde{\mathcal{H}}^*_{\rho_{ij}\rho_{km}} &= -\rho^{ik}\rho^{jm} - \rho^{im}\rho^{jk};\ i > j
\end{aligned}$$

[5] Although both $\hat{J}^*_{\varpi\varpi T} = -\frac{1}{T}\sum_{t=1}^{T}\mathcal{H}^*_t(\hat{\varpi})$ and $\hat{J}^*_{\varpi\varpi T} = -\frac{1}{T}\sum_{t=1}^{T}\widetilde{\mathcal{H}}^*_t(\hat{\varpi})$ are asymptotically equivalent; in finite sample their performance may vary (see Hafner and Herwartz, 2008).

Lemma 3 *For QMLE* $\hat{\varpi}$, $J^*_{\varpi\varpi} - \hat{J}^*_{\varpi\varpi T} = o_p(1)$; *and* $\hat{J}^*_{\varpi\varpi T} = -\frac{1}{T}\sum_{t=1}^{T} \widetilde{\mathcal{H}}^*_t(\hat{\varpi})$ *has the form*

$$\hat{J}^*_{\varpi\varpi T} = \begin{bmatrix} \hat{J}^*_{\varphi\varphi T} & \hat{J}^*_{\varphi\eta T} & \hat{J}^*_{\varphi\rho T} \\ \hat{J}^{*\prime}_{\varphi\eta T} & \hat{J}^*_{\eta\eta T} & \hat{J}^*_{\eta\rho T} \\ \hat{J}^{*\prime}_{\varphi\rho} & \hat{J}^{*\prime}_{\eta\rho T} & \hat{J}^*_{\rho\rho T} \end{bmatrix} = T^{-1}\sum_{t=1}^{T} \left[\begin{array}{ccc|} \dfrac{\partial^2 l_t^*}{\partial\varphi\partial\varphi'} & \dfrac{\partial^2 l_t^*}{\partial\varphi\partial\eta'} & \dfrac{\partial^2 l_t^*}{\partial\varphi\partial\rho'} \\ \dfrac{\partial^2 l_t^*}{\partial\eta\partial\varphi'} & \dfrac{\partial^2 l_t^*}{\partial\eta\partial\eta'} & \dfrac{\partial^2 l_t^*}{\partial\eta\partial\rho'} \\ \dfrac{\partial^2 l_t^*}{\partial\rho\partial\varphi'} & \dfrac{\partial^2 l_t^*}{\partial\rho\partial\eta'} & \dfrac{\partial^2 l_t^*}{\partial\rho\partial\rho'} \end{array} \mathcal{F}_{t-1} \right]_{\varpi=\hat{\varpi}}$$

where

$$\begin{aligned}
\hat{J}^*_{\varphi\varphi T} &= \frac{1}{T}\left[\hat{F}'\left(\hat{\Gamma}^{-1}\otimes I_T\right)\hat{F} + \frac{1}{4}\hat{C}'\left(\hat{\Gamma}_A\otimes I_T\right)\hat{C}\right] \longrightarrow J^*_{\varphi\varphi} \\
\hat{J}^*_{\varphi\eta T} &= \frac{1}{4T}\hat{C}'\left(\hat{\Gamma}_A\otimes I_T\right)\hat{X} \longrightarrow J^*_{\varphi\eta} \\
\hat{J}^*_{\varphi\rho T} &= \frac{1}{2T}\hat{C}'\left(I_N\otimes\iota_T'\right)\hat{P} \longrightarrow J^*_{\varphi\rho} \\
\hat{J}^*_{\eta\eta T} &= \frac{1}{4T}\hat{X}'\left(\hat{\Gamma}_A\otimes I_T\right)\hat{X} \longrightarrow J^*_{\eta\eta} \\
\hat{J}^*_{\eta\rho T} &= \frac{1}{2T}\hat{X}'\left(I_N\otimes\iota_T'\right)\hat{P} \longrightarrow J^*_{\eta\rho} \\
\hat{J}^*_{\rho\rho T} &= \hat{\widetilde{P}} \longrightarrow J^*_{\varphi\varphi}
\end{aligned}$$

where $\underset{N\times\frac{N(N-1)}{2}}{P}$ *has rows* $p_k' = vecl(P_k)'$, $k = 1, ..., N$ *and* $\underset{\frac{N(N-1)}{2}\times\frac{N(N-1)}{2}}{\widetilde{P}}$ *has columns* $\widetilde{p}_{km} = vecl(\Gamma_{km})$, $m = 1, ..., N-1$, $k = m+1, ..., N$ *(k changes more quickly than m) while* F', C', X', Γ_A, P_k *and* Γ_{km} *are defined at the onset of this section.*

2.2 Partial (or Two-step) QMLE (PQMLE) Estimation

Because of the structure of log-likelihood of the conditional correlation model, a simplified two step estimation procedure can be implemented as suggested by Engle (2002), which involves (at the first step) separate estimation of the N univariate GARCH models to get the volatility estimates, and then using these obtain the correlation parameter estimates. Such a procedure is consistent, but asymptotically inefficient when compared with the FQMLE procedure. This partial estimation technique is mostly useful for the DCC models due to the complexity of the estimation procedure, but can be used for the CCC model.

Note that (7) can be expressed as the sum of two components; $l_t^*(\theta,\rho) = \sum_{j=1}^{N} l_t^V(\theta_j) + l_t^C(\theta,\rho)$ where $\sum_{j=1}^{N} l_t^V(\theta_j) = -\dfrac{1}{2}\sum_{j=1}^{N}\left\{\ln h_{jt} + h_{jt}^{-1}\varepsilon_{jt}^2\right\}$ represents conditional log-likelihood contributions for N separate GARCH(p,q) models which is functionally independent of ρ, and $l_t^C(\theta,\rho) = -\dfrac{1}{2}\ln|\Gamma| - \frac{1}{2}\zeta_t'\Gamma^{-1}\zeta_t + \frac{1}{2}\zeta_t'\zeta_t$ contains the correlation structure. Two step estimation is then pursued as follows:

1. Obtain $\hat{\theta}_j = \arg\max_{\theta_j} \sum_{t=1}^{T} l_t^v(\theta_j), \quad j = 1, ..., N$ by QML applying to univariate GARCH(p, q) specification for individual variables.[6] Then construct standardized residuals as $\hat{\zeta}_{jt} = \hat{h}_{jt}^{-1/2}\hat{\varepsilon}_{jt}$, and $l_t^C(\hat{\theta}, \rho) = -\frac{1}{2}\ln|\Gamma| - \frac{1}{2}\hat{\zeta}_t'\Gamma^{-1}\hat{\zeta}_t + \frac{1}{2}\hat{\zeta}_t'\hat{\zeta}_t = k_t - \frac{1}{2}\ln|\Gamma| - \frac{1}{2}\hat{\zeta}_t'\Gamma^{-1}\hat{\zeta}_t$ where k_t is a constant as far as ρ is concerned.

2. obtain $\hat{\rho} = \arg\max_{\rho} \sum_{t=1}^{T} l_t^C(\hat{\theta}, \rho)$, which satisfies the score equations $\sum_{t=1}^{T}\left(\hat{\varepsilon}_{it}^*\hat{\varepsilon}_{jt}^* - \hat{\rho}^{ij}\right) = 0$, $j < i$, with $\hat{\varepsilon}_t^* = \left\{\hat{\varepsilon}_{jt}^*\right\} = \hat{\Gamma}^{-1}\hat{\zeta}_t$.

Hence the PQMLE $\hat{\varpi} = \left(\hat{\theta}', \hat{\rho}'\right)' = \left(\hat{\theta}_1', \cdots \hat{\theta}_N', \hat{\rho}'\right)'$ can be obtained from the above two steps. Note that to avoid notational complexity we use "hat" to denote both FQMLE and PQMLE; this should not make any confusion as later, while deriving the test statistics, notational differences will clearly distinguish the estimation procedure employed. Hafner and Herwartz (2008) provides an analytical expression for the variance of the two-step QMLE for both the CCC and DCC models.

As noted by Engle (2002), the correlation matrix Γ_t is also the conditional covariance matrix of standardized errors i.e. $\mathrm{E}\left[\zeta_t\zeta_t'|F_{t-1}\right] = \Gamma$. Although, the scores for ρ obtained in second step is not equal to $\sum_{t=1}^{T}\left(\hat{\zeta}_{it}\hat{\zeta}_{jt} - \hat{\rho}_{ij}\right) = 0, j < i$, Bollerslev's (1990) pointed out that a suitable reparameterization ensures that $T^{-1}\sum_{t=1}^{T}\left(\frac{\hat{\varepsilon}_t^2}{\hat{h}_t} - 1\right) = 0$ so that $\sum_{t=1}^{T}\left(\hat{\zeta}_{it}\hat{\zeta}_{jt} - \hat{\rho}_{ij}\right) = \sum_{t=1}^{T}\left(\hat{\varepsilon}_{it}^*\hat{\varepsilon}_{jt}^* - \hat{\rho}^{ij}\right) = 0$, $j < i$. Therefore, we can use $\hat{\rho}_{ij} = \frac{1}{T}\sum_{t=1}^{T}\hat{\zeta}_{it}\hat{\zeta}_{jt}$, $j < i$ as a consistent estimator for ρ_{ij}. However, noting that in finite sample sample covariance matrix of standardized residuals will never be a correlation matrix, as the diagonal will not be exactly equal (though very close) to 1, another option is to use the usual correlation estimator i.e.

$$\hat{\rho}_{ij}^* = \frac{\sum_{t=1}^{T}\hat{\zeta}_{it}\hat{\zeta}_{jt}}{\sqrt{\sum_{t=1}^{T}\hat{\zeta}_{it}^2\sum_{t=1}^{T}\hat{\zeta}_{jt}^2}};\; i, j = 1, \cdots, N$$

which is a linear (one-to-one) transformation of $\hat{\rho}_{ij} = \frac{1}{T}\sum_{t=1}^{T}\hat{\zeta}_{it}\hat{\zeta}_{jt}$, $j < i$. In the literature both versions are used to estimate the correlation parameters. For testing of the CCC assumption we only need $\hat{\rho}_{ij}$ $i \neq j$; and score tests are invariant to linear transformation of parameter space (see Dagenais and Dufour, 1991). Hence in this paper when developing the asymptotic theory we will use $\hat{\rho} = \frac{1}{T}\hat{\zeta}_t\hat{\zeta}_t'$ as the PQML estimator of ρ.

[6] For the scores, see e.g. Halunga and Orme (2009).

3 A Class of Asymptotically Valid CM Test Procedures

In this section, we develop a class of asymptotically valid parametric testing procedures, along with the first order asymptotic distribution results, of the CCC assumption that are derived from the conditional moment (CM) principle. If both individual GARCH specifications and CCC assumption is correct, then the definition of standardized residuals, given in (4), provides the moment condition corresponding to CCC assumption i.e. $\mathrm{E}\left[\zeta_t\zeta_t' - \Gamma|\mathcal{F}_{t-1}\right] = 0$. Note that, the diagonal elements of $\left(\zeta_t\zeta_t' - \Gamma\right)$ correspond to the individual GARCH (or volatility) specifications *whereas the off-diagonal elements correspond to the CCC assumption*. Also due to the symmetry of $\left(\zeta_t\zeta_t' - \Gamma\right)$, there are $\frac{N(N+1)}{2}$ independent restrictions in this moment condition; hence we can write these distinct moment restrictions as:

$$\mathrm{E}\left[\mathrm{vech}\left(\zeta_t\zeta_t' - \Gamma|\mathcal{F}_{t-1}\right)\right] = 0. \tag{9}$$

If we are interested in testing simply the CCC assumption leaving the individual GARCH specifications aside, then we need to consider the strictly lower triangular portion of $\left(\zeta_t\zeta_t' - \Gamma\right)$, i.e.

$$\mathrm{E}\left[\mathrm{vecl}\left(\zeta_t\zeta_t' - \Gamma|\mathcal{F}_{t-1}\right)\right] = 0. \tag{10}$$

The parametric misspecification tests of the conditional correlation models can be constructed by considering either (9) or (10). If the test is based on (9), which will be referred as *Full CM (FCM) test*, it can be treated as a joint misspecification test of the complete MGARCH specification as this would also pick any misspecification in individual volatility specifications with that of the correlation specification. On the other hand if the underlying moment restriction of the test is (10), we will refer the test as *CCC CM (CCM) test.*

Therefore, a joint parametric misspecification test of the CCC and individual volatility assumptions might be constructed as test of the following null moment restriction :

$$\mathrm{E}\left[\mathrm{vech}\left(\zeta_t\zeta_t' - \Gamma\right) \otimes r_t\left(\varpi_0\right)\right] = 0 \tag{11}$$

where $r_t\left(\varpi_0\right)$ be a F_{t-1} measurable test variables. To test this null, the generic CM test indicator is constructed as

$$\hat{M}_T^j = \frac{1}{T}\sum_{t=1}^{T}\left(\hat{v}_t \otimes \hat{r}_t\right) = \frac{1}{T}\sum_{t=1}^{T}\hat{m}_t^j \tag{12}$$

where the superscript j denote joint testing of CCC and individual volatility specifications and $\hat{v}_t = \mathrm{vech}\left(\hat{\zeta}_t\hat{\zeta}_t' - \hat{\Gamma}\right)$ where "hats" denote that everything is evaluated at the consistent null parameter estimator (either FQMLE or PQMLE), $\hat{\varpi} = \left(\hat{\theta}', \hat{\rho}'\right)'$. Similarly, a misspecification test of the CCC assumption, *only*, can be conducted by testing the moment restriction:

$$\mathrm{E}\left[\mathrm{vecl}\left(\zeta_t\zeta_t' - \Gamma\right) \otimes r_t\left(\varpi_0\right)\right] = 0 \tag{13}$$

The corresponding CM test indicator would have the following form:

$$\hat{M}_T^c = \frac{1}{T}\sum_{t=1}^{T}\left(\hat{\upsilon}_t^* \otimes \hat{r}_t\right) = \frac{1}{T}\sum_{t=1}^{T}\hat{m}_t^c \tag{14}$$

where $\hat{\upsilon}_t^* = \text{vecl}\left(\hat{\zeta}_t\hat{\zeta}_t' - \hat{\Gamma}\right)$ and superscript c denote testing of only CCC assumption. It is to be noted here that (14) is simply a subset of (12).

Example 1 *For example, in the bivariate case,*

$$\mathrm{E}\left[\zeta_t\zeta_t' - \Gamma|\mathcal{F}_{t-1}\right] = \mathrm{E}\left[\left.\begin{pmatrix}\zeta_{1t}^2 & \zeta_{1t}\zeta_{2t}\\ \zeta_{1t}\zeta_{2t} & \zeta_{2t}^2\end{pmatrix} - \begin{pmatrix}1 & \rho\\ \rho & 1\end{pmatrix}\right|\mathcal{F}_{t-1}\right] = 0.$$

Then (9) becomes

$$\mathrm{E}\left[\left.\left(\zeta_{1t}^2 - 1,\ \zeta_{1t}\zeta_{2t} - \rho,\ \zeta_{2t}^2 - 1\right)'\right|\mathcal{F}_{t-1}\right] = 0.$$

The first and third components refer to individual GARCH equations while second one corresponds to the CCC assumption. Therefore, in this case (10) is simply $\mathrm{E}\left[\zeta_{1t}\zeta_{2t} - \rho|\mathcal{F}_{t-1}\right] = 0$. *Subsequently for a F_{t-1} measurable test variable $r_t(\varpi_0)$ (11) becomes:*

$$\mathrm{E}\begin{bmatrix}\zeta_{1t}^2 - 1\\ \zeta_{1t}\zeta_{2t} - \rho\\ \zeta_{2t}^2 - 1\end{bmatrix} \otimes r_t(\varpi_0) = 0. \tag{15}$$

and (13) becomes: $\mathrm{E}\left[\left(\zeta_{1t}\zeta_{2t} - \rho\right) \otimes r_t(\varpi_0)\right] = 0.$

To develop asymptotically valid tests of CCC hypothesis we need to establish the limit distributions of the test indicator vector$\sqrt{T}\hat{M}_T$. Both FQMLE and PQMLE approaches are considered while deriving the test statistics and their asymptotic distributions. We illustrate the procedure of constructing the test statistics considering both Gaussian and non-Gaussion distribution of the fully standardized error process, ξ_t. In case of non-normally distributed ξ_t we develop a *non-normality robust* procedure in the similar spirit of Wooldridge (1990).[7] When ξ_t follows a normal distribution the generalized IM inequality holds (see, e.g. Newey 1985) and the outer product of gradient (OPG) covariance matrix estimator can be employed in deriving the test statistics.

3.1 Case 1: Tests based on FQMLE

The test indicator under consideration is $\hat{M}_{FT} \equiv M_{FT}(\hat{\varpi}) = T^{-1}\sum_{t=1}^{T}\hat{m}_{Ft}$, where subscript F represent FQMLE case; with $\mathrm{E}_0\left[m_{Ft}\right] = 0$. Define the $(T \times r)$ matrix R with rows m_{Ft}'. Hereafter we will use the notation $G_{0T}^* = G_T^*(\varpi_0)$, $M_{0FT} \equiv M_{FT}(\varpi_0)$, etc. where ϖ_0 denotes the true parameter values and $\hat{M}_{FT}$, $\hat{J}_{\varpi\varpi T}^{*-1}$ etc. to denote evaluation at $\hat{\varpi}$. We assume sufficient regularity to satisfy the following central limit theorem :

[7] Similar approach was employed by Halunga and Orme (2009).

Proposition 1 $\sqrt{T}\begin{bmatrix} M_{0FT} \\ G_{0T}^{*} \end{bmatrix} = T^{-1/2}\begin{bmatrix} \sum_{t=1}^{T} m_{0Ft} \\ \sum_{t=1}^{T} g_{0t}^{*} \end{bmatrix} \xrightarrow{d} N\left(0, \Sigma^{*}\right),$

where $\Sigma^{*} = \begin{bmatrix} \Sigma_{MM} & \Sigma_{MG}^{*} \\ \Sigma_{GM}^{*} & \Sigma_{GG}^{*} \end{bmatrix}$,
$\Sigma_{MM} = \text{plim}_{T\to\infty} T^{-1}\sum_{t=1}^{T} m_{0Ft}m_{0Ft}'$,
$\Sigma_{GG}^{*} = \text{plim}_{T\to\infty} T^{-1}\sum_{t=1}^{T} g_{0t}^{*}g_{0t}^{*\prime}$; and
$\Sigma_{MG}^{*} = \text{plim}_{T\to\infty} T^{-1}\sum_{t=1}^{T} m_{0Ft}g_{0t}^{*\prime}$.

Remark 1 *Assuming a suitable ULLN, Σ_{MM} might be consistently estimated, for example, by $T^{-1}\sum_{t=1}^{T}\hat{m}_{Ft}\hat{m}_{Ft}'$; but see also Halunga and Orme (2009).*

We then have the following result.

Theorem 2 *Given $\hat{\varpi} \xrightarrow{p} \varpi_0$, the CLT stated in Proposition 1 and a suitable ULLN,*

$$\sqrt{T}\hat{M}_{FT} \xrightarrow{d} N\left(0, V\right)$$

where
$V = A^{*}\Sigma^{*}A^{*\prime}$,
$\Sigma^{*} = \begin{bmatrix} \Sigma_{MM} & \Sigma_{MG}^{*} \\ \Sigma_{GM}^{*} & \Sigma_{GG}^{*} \end{bmatrix}$, *(see Proposition 1) and*
$A^{*} = \left[I_r : -J_{M\varpi}^{*}J_{\varpi\varpi}^{*-1}\right]$, *with* $J_{\varpi\varpi}^{*} = -\underset{T\longrightarrow\infty}{\text{plim}}\frac{1}{T}\left[\sum_{t=1}^{T}\mathcal{H}_t^{*}\left(\varpi_0\right)\right]$, $J_{M\varpi}^{*} = -\underset{T\longrightarrow\infty}{\text{plim}}\left[\frac{\partial M_{0Ft}}{\partial\varpi'}\right]$, *and I_r is the identity matrix of rank $r = rank\left(\mathbf{\Sigma}_{MM}\right)$.*

Remark 2 *Note that the variance-covariance matrix V can be written as*

$$V = \Sigma_{MM} - \Sigma_{MG}^{*}J_{\varpi\varpi}^{*-1}J_{M\varpi}^{*\prime} - J_{M\varpi}^{*}J_{\varpi\varpi}^{*-1}\Sigma_{GM}^{*} + J_{M\varpi}^{*}J_{\varpi\varpi}^{*-1}\Sigma_{GG}^{*}J_{\varpi\varpi}^{*-1}J_{M\varpi}^{*\prime}.$$

From the preceding result, the general form of the CCC misspecification test statistic based on FQMLE is the quadratic form

$$T_F = T\hat{M}_{FT}'\hat{V}_T^{-1}\hat{M}_{FT} \qquad (16)$$

under the null which has a χ_r^2 limiting distribution, where $\hat{V}_T$ is any consistent estimator for V i.e. $\hat{V}_T = V + o_p\left(1\right)$.

3.1.1 Case 1a: Robust FQMLE test

To construct a robust (to non-normality) test statistics we need a consistent estimator $\hat{V}_T^{r} = \hat{A}^{*}\hat{\Sigma}^{*}\hat{A}^{*\prime}$ where the superscript r signifies the robust estimator; for which we require $\hat{J}_{M\varpi T}^{*} = -T^{-1}\sum\frac{\partial m_{Ft}}{\partial\hat{\varpi}'}$, $\hat{J}_{\varpi\varpi T}^{*}$ and $\hat{\Sigma}^{*}$. For $\hat{\varpi}$, we can construct $\hat{J}_{M\varpi T}^{*}$ using the results provided in Proposition 2.

Proposition 2 *It can be shown that for $i = 1, \cdots N$.*

$$\mathrm{E}_0\left[\frac{\partial\left(\zeta_{it}^2 - 1\right)}{\partial\theta'} r_t\right] = -\operatorname*{plim}_{T\longrightarrow\infty}\frac{1}{T}\left[R^{*\prime}(0, \ldots, Z_i, \cdots, 0)\right];$$

and for $i \neq j, j < i = 2, \cdots, N$.

$$\begin{aligned} \mathrm{E}_0\left[\frac{\partial(\zeta_{it}\zeta_{jt} - \rho_{ij})}{\partial\theta'} r_t\right] &= -\frac{1}{2}\rho_0 \operatorname*{plim}_{T\longrightarrow\infty}\frac{1}{T}\left\{R^{*\prime}(0, \cdots, Z_i, \cdots, Z_j, \cdots 0)\right\} \\ \mathrm{E}_0\left[\frac{\partial(\zeta_{it}\zeta_{jt} - \rho_{ij})}{\partial\rho'} r_t\right] &= -\operatorname*{plim}_{T\longrightarrow\infty}\frac{1}{T}(0, \ldots, 1', \cdots, 0)R^*. \end{aligned}$$

where Z_i is $(T \times k_i)$ matrix having rows $z_{it}' = (c_{it}', x_{it}')$ and R^* having rows r_t', if r_t is a vector of test variables, or R^* is a vector with typical element r_t if r_t is a scalar.

Example 2 *Again consider $N = 2$ Then the full moment condition given in (15); so we have:*

$$\hat{J}_{M\varpi T}^* = \frac{1}{T}\left[\begin{array}{c} \hat{R}^{*\prime}(\hat{Z}_1, 0, 0) \\ \frac{1}{2}\hat{\rho}\hat{R}^{*\prime}(\hat{Z}_1, \hat{Z}_2), \iota_T'\hat{R}^* \\ \hat{R}^{*\prime}(0, \hat{Z}_2, 0) \end{array}\right]$$

Clearly, for only CCC moment condition $\hat{J}_{M\varpi T}^* = \left\{\frac{1}{2T}\hat{\rho}\hat{R}^{*\prime}(\hat{Z}_1, \hat{Z}_2), \iota_T'\hat{R}^*\right\}$.

Now using the next lemma a robust and consistent estimator $\hat{V}_T^r$ can be obtained.

Lemma 4 *Under suitable assumptions, $\hat{\Sigma}_T^* - \Sigma^* = o_p(1)$ and $\hat{A}^* - A^* = o_p(1)$ where*

$$\begin{aligned} \hat{\Sigma}_T^* &= \left[\begin{array}{cc} \hat{\Sigma}_{MM} & \hat{\Sigma}_{MG}^* \\ \hat{\Sigma}_{GM}^* & \hat{\Sigma}_{GG}^* \end{array}\right] = \frac{1}{T}\begin{bmatrix} \hat{R}'\hat{R} & \hat{S}^{*\prime}\hat{R} \\ \hat{R}'\hat{S}^* & \hat{S}^{*\prime}\hat{S}^* \end{bmatrix} \\ \hat{A}^* &= \left[I_r : -\hat{J}_{M\varpi T}^*\hat{J}_{\varpi\varpi T}^{*-1}\right] \end{aligned}$$

where R and S^* are matrices with rows m_{Ft}' and $g_t^{*\prime}(\varpi)$ respectively; $\hat{J}_{\varpi\varpi T}^*$ and $\hat{J}_{M\varpi T}^*$ are constructed using Lemma 3 and Proposition 2 respectively; all evaluated at $\hat{\varpi}$. Therefore we have,

$$\hat{V}_T^r = \hat{\Sigma}_{MM} - \hat{\Sigma}_{MG}^*\hat{J}_{\varpi\varpi T}^{*-1}\hat{J}_{M\varpi T}^{*\prime} - \hat{J}_{M\varpi T}^*\hat{J}_{\varpi\varpi T}^{*-1}\hat{\Sigma}_{GM}^* + \hat{J}_{M\varpi T}^*\hat{J}_{\varpi\varpi T}^{*-1}\hat{\Sigma}_{GG}^*\hat{J}_{\varpi\varpi T}^{*-1}\hat{J}_{M\varpi T}^{*\prime}.$$

Tests that are based on this estimator $\hat{V}_T^r$ will be referred as robust FQMLE test and will be denoted as $T_F^{(r)}$.

3.1.2 Case 1(b) OPG FQMLE test

However, if $\xi_t \sim N(0, I_N)$, then $\Sigma^*_{GG} = J^*_{\varpi\varpi}$ and V reduces to $V = \Sigma_{MM} - \Sigma^*_{MG}\Sigma^{*-1}_{GG}\Sigma^*_{GM}$. The following lemma provides an expression for the consistent estimator of V when the normality assumption holds.

Lemma 5 *Under suitable assumptions and $\xi_t \sim N(0, I_N)$, V can be consistently estimated by*

$$\hat{V}_T = \hat{\Sigma}_{MM} - \hat{\Sigma}^*_{MG}\hat{\Sigma}^{*-1}_{GG}\hat{\Sigma}^*_{GM} = \frac{1}{T}\hat{W}^{*\prime}\hat{W}^*$$

where

$$\begin{aligned}
\hat{\Sigma}_{MM} &= T^{-1}\sum_{t=1}^{T}\hat{m}_{Ft}\hat{m}'_{Ft} = T^{-1}\hat{R}'\hat{R}, \\
\hat{\Sigma}^*_{GG} &= T^{-1}\sum_{t=1}^{T}\hat{g}^*_t\hat{g}^{*\prime}_t = T^{-1}\hat{S}^{*\prime}\hat{S}^*, \\
\hat{\Sigma}^*_{MG} &= T^{-1}\sum_{t=1}^{T}\hat{m}_{Ft}\hat{g}^{*\prime}_t = T^{-1}\hat{S}^{*\prime}\hat{R}, \\
\hat{W}^* &= \hat{B}^*\hat{A}^{*\prime} \\
\hat{A}^* &= \left[I_r : -\hat{\Sigma}^*_{MG}\hat{\Sigma}^{*-1}_{GG}\right], \text{ and} \\
B^* &= [R,\ S^*]
\end{aligned}$$

where R and S^* are matrices with rows m'_{Ft} and $g^{*\prime}_t(\varpi)$ respectively evaluated at $\hat{\varpi}$.

In this case, the test statistic (16) has a convenient OPG (Outer Product of Gradient) form. To see this, note that

$$\hat{W}^* = \hat{B}^*\hat{A}^{*\prime} = \hat{R} - \hat{S}^*\left(\hat{S}^{*\prime}\hat{S}^*\right)^{-1}\hat{S}^{*\prime}\hat{R}$$

Exploiting the FOC that $\hat{S}^{*\prime}\iota_T \equiv 0 \Longrightarrow \hat{W}^{*\prime}\iota_T \equiv \hat{R}'\iota_T$, hence an alternative form of the test statistic under normality is give by:

$$T_F = \iota'_T\hat{R}\left(\hat{W}^{*\prime}\hat{W}^*\right)^{-1}\hat{R}'\iota_T = \iota'_T\hat{W}^*\left(\hat{W}^{*\prime}\hat{W}^*\right)^{-1}\hat{W}^{*\prime}\iota_T \tag{17}$$

where ι_T is the $(T \times 1)$ column vector of ones. (17) can be interpreted as $T - RSS$ where RSS is the residual sum of squares from the regression of ι_T on $\hat{W}$. Note that, this test can be constructed easily by defining $\hat{U}^* = \left(\hat{R}, \hat{S}^*\right)$, then

$$T_F = \iota'_T\hat{U}^*\left(\hat{U}^{*\prime}\hat{U}^*\right)^{-1}\hat{U}^{*\prime}\iota_T \tag{18}$$

and can be obtained as $T - RSS$ from a regression of ι_T on $\hat{U}^*$.

3.1.3 Summary: FQMLE

From the above results, for each of the FCM and CCM test statistics, based on the FQMLE $\hat{\varpi}$, we have two versions namely, robust and OPG i.e.

1. Robust (to non-normality) FCM test:

$$T_F^{j(r)} = T\hat{M}_{FT}^{j\prime}\left(\hat{V}_T^{j(r)}\right)^{-1}\hat{M}_{FT}^{j} \tag{19}$$

2. OPG FCM test

$$T_F^{j} = T\hat{M}_{FT}^{j\prime}\left(\hat{V}_T^{j}\right)^{-1}\hat{M}_{FT}^{j} \tag{20}$$

3. Robust (to non-normality) CCM test

$$T_F^{c(r)} = T\hat{M}_{FT}^{c\prime}\left(\hat{V}_T^{c(r)}\right)^{-1}\hat{M}_{FT}^{c} \tag{21}$$

4. OPG CCM test

$$T_F^{c} = T\hat{M}_{FT}^{c\prime}\left(\hat{V}_T^{c}\right)^{-1}\hat{M}_{FT}^{c} \tag{22}$$

where the robust variance estimator $\hat{V}_T^{j(r)}$ and $\hat{V}_T^{c(r)}$ can be obtained using lemma (4) while OPG test statistics are constructed using the artificial regression as given in (18) with appropriate test indicators.

3.2 Case 2: Tests based on PQMLE

Define the test indicator under investigation as:

$$\hat{M}_{PT} \equiv M_{PT}(\hat{\theta},\hat{\rho}) = \frac{1}{T}\sum_{t=1}^{T} m_{Pt}(\hat{\theta},\hat{\rho})$$

with $\mathrm{E}_0\left[m_{pt}\right] = 0$; where subscript P represents the PQMLE case i.e. the correlation parameters are estimated by $\hat{\rho}_{ij} = \frac{1}{T}\sum_{t=1}^{T}\hat{\zeta}_{it}\hat{\zeta}_{jt}$, $j < i$. We will establish the results considering the bivariate case for the ease of exposition so that $\theta' = \left(\theta_1', \theta_2'\right)$; the results can be generalized to higher dimensional cases in an obvious way.

Let $L_{iT}(\theta_i) = \dfrac{1}{T}\sum_{t=1}^{T} l_{it}(\theta_i)$ be the average log-likelihood of univariate GARCH models for the i-th variable where $l_{it}(\theta_i) = -\frac{1}{2}\left[\ln\left(h_{it}\right) + \dfrac{\varepsilon_{it}^2}{h_{it}}\right]$ (ignoring constants). Define $G_i(\theta_i) = \dfrac{1}{T}\sum_{t=1}^{T} g_{it}(\theta_i) = \dfrac{1}{T}\sum_{t=1}^{T}\dfrac{\partial l_{it}(\theta_i)}{\partial\theta_i}$, and

$$J_{\theta\theta} = \begin{bmatrix} J_1(\theta_1) & 0 \\ 0 & J_2(\theta_2) \end{bmatrix}$$

where $J_i(\theta_i) \equiv J_i = -\mathrm{E}_0\left[\frac{\partial^2 l_{it}(\theta_i)}{\partial\theta_i\partial\theta_i'}\right] = -\operatorname{plim}_{T\to\infty}\frac{1}{T}\sum_{t=1}^{T}\frac{\partial^2 l_{it}(\theta_i)}{\partial\theta_i\partial\theta_i'}$. The block diagonal structure of $J_{\theta\theta}$ follows from the fact that in PQMLE framework, the univariate GARCH estimating equations are functionally independent i.e. $G_i(\theta_i),\ i = 1,2$ are functionally independent. Also define $Q = (Q_1, Q_2)$ and $S = (S_1, S_2)$ where Q_i and S_i are both $(T \times k_i)$ matrix, $k_i = K + K_1$, with rows $g_{it}'(\theta_i) = \dfrac{\partial l_{it}(\theta_i)}{\partial\theta_i'}$ and $g_t^{*\prime}(\theta_i) = \dfrac{\partial l_t^*(\theta,\rho)}{\partial\theta_i'}$, $i = 1,2$ respectively.[8] Also, as before, R be the $(T \times r)$ matrix but now with rows m_{Pt}'.

The separate limit distributions of $\sqrt{T}\left(\hat{\theta}_i - \theta_{i0}\right) = J_i(\theta_{i0})^{-1}\sqrt{T}G_i(\theta_{i0}) + o_p(1)$, for true parameter values θ_{i0}, $(k_i \times 1)$, $i = 1,2$, are essentially given in Halunga and Orme (1990, Theorem 1). We have, $\sqrt{T}G(\theta_0) \xrightarrow{d} N(0, \Sigma_{GG})$, where, $G_{0T} = (G_1(\theta_{10})', G_2(\theta_{20})')'$, and

$$\Sigma_{GG} = \operatorname*{plim}_{T\to\infty}\frac{1}{T}Q'Q = \operatorname*{plim}_{T\to\infty}\frac{1}{T}\begin{bmatrix} Q_1'Q_1 & Q_1'Q_2 \\ Q_2'Q_1 & Q_2'Q_2 \end{bmatrix}.$$

Now to test the significance of the test indicator, $\hat{M}_{PT}$, the limit joint distribution needs to take account of the estimation effect from correlation parameter. We can ignore this estimation effect from ρ, which will eventually lead to relatively simple to construct asymptotically valid tests, if we can impose the following condition:

Condition 1 $\sqrt{T}M_T(\hat{\theta},\hat{\rho}) = \sqrt{T}\hat{M}_T(\hat{\theta},\rho_0) + o_p(1)$.

This implies that the effect of estimating ρ using the first step estimator $\hat{\theta}' = \left(\hat{\theta}_1', \hat{\theta}_2'\right)$ can be ignored (asymptotically). Although it seems a very restricted condition, in our case this condition can easily be met by using a centered (i.e. demeaned) test variable $\left(\hat{r}_t - \widehat{\overline{r}}\right)$ and thereby transforming the test indicator $\hat{M}_{PT}$ functionally independent of ρ.

Example 3 *For example, in the bivariate context consider the only correlation test indicator which is given by*

$$\begin{aligned} \hat{M}_{PT}^{c} &= \frac{1}{T}\sum_{t=1}^{T}\left[\hat{\zeta}_{1t}\hat{\zeta}_{2t} - \hat{\rho}\right]\hat{r}_t \\ &= \frac{1}{T}\sum_{t=1}^{T}\hat{\zeta}_{1t}\hat{\zeta}_{2t}\left(\hat{r}_t - \widehat{\overline{r}}\right) = \frac{1}{T}\sum_{t=1}^{T}\left(\hat{\zeta}_{1t}\hat{\zeta}_{2t} - \hat{\rho}\right)\left(\hat{r}_t - \widehat{\overline{r}}\right) \end{aligned} \tag{23}$$

where $\widehat{\overline{r}} = \frac{1}{T}\sum_{t=1}^{T}\hat{r}_t$ *and* $\hat{\rho} = \frac{1}{T}\sum_{t=1}^{T}\hat{\zeta}_{1t}\hat{\zeta}_{2t}$.

[8] That is, Q_i is the matrix having rows univariate GARCH scores $g_{it}'(\theta_i)$ while the rows of S_i contains the FQMLE scores $g_t^{*\prime}(\theta_i)$, given in Lemma 1, corresponding to the conditional mean and volatility parameters θ_i barring correlation parameter ρ.

This simple demeaning trick produce algebraically equivalent test indicators, but will not involve $\hat{\rho}$. In other words, $\hat{M}_{PT}$ is simply a function of $\left(\hat{\theta}_1', \hat{\theta}_2'\right)$ and does not involve $\hat{\rho}$ and hence allow us to deduce the limit distribution with Condition 1.

Remark 3 *Note that since* $\frac{1}{T}\sum_{t=1}^{T}\left(\hat{\zeta}_{it}^2 - 1\right) \neq 0$*;*

$$\frac{1}{T}\sum_{t=1}^{T}\left(\hat{\zeta}_{it}^2 - 1\right)\hat{r}_t \neq \frac{1}{T}\sum_{t=1}^{T}\hat{\zeta}_{it}^2\left(\hat{r}_t - \widehat{\overline{r}}\right).$$

However, Condition 1 does apply to the full CM test indicator M_{PT}^j *since* $\dfrac{1}{T}\sum_{t=1}^{T}\left(\hat{\zeta}_{it}^2 - 1\right)\hat{r}_t,\ i = 1, 2$ *does not involve* $\hat{\rho}$*. Hence, it is not necessary to demean the test variable for the elements of test indicator* $\hat{M}_{PT}^j$ *which do not involve* $\hat{\rho}$*. However, hereafter for simplicity we will consider demeaned test variables for all elements of* $\hat{M}_{PT}^j$*; i.e.*

$$\hat{M}_{PT}^j = \frac{1}{T}\sum_{t=1}^{T}\begin{pmatrix} \left[\hat{\zeta}_{1t}^2 - 1\right]\left(\hat{r}_t - \widehat{\overline{r}}\right) \\ \left(\hat{\zeta}_{1t}\hat{\zeta}_{2t} - \hat{\rho}\right)\left(\hat{r}_t - \widehat{\overline{r}}\right) \\ \left[\hat{\zeta}_{2t}^2 - 1\right]\left(\hat{r}_t - \widehat{\overline{r}}\right) \end{pmatrix} = \frac{1}{T}\sum_{t=1}^{T}\left\{\hat{v}_t \otimes \left(\hat{r}_t - \widehat{\overline{r}}\right)\right\} \tag{24}$$

To derive the asymptotic distribution, we assume the following central limit theorem to hold:

Proposition 3 *Under suitable regularity conditions,*

$$\sqrt{T}\begin{bmatrix} M_{0PT} \\ G_{0T} \end{bmatrix} = \frac{1}{\sqrt{T}}\begin{bmatrix} \sum_{t=1}^{T} m_{0Pt} \\ \sum_{t=1}^{T} g_{0t} \end{bmatrix} \xrightarrow{d} N(0, \Sigma)$$

where $\Sigma = \begin{bmatrix} \mathbf{\Sigma}_{MM} & \mathbf{\Sigma}_{MG} \\ \mathbf{\Sigma}_{GM} & \mathbf{\Sigma}_{GG} \end{bmatrix}$ *with*

$$\begin{aligned} \Sigma_{MM} &= \operatorname*{plim}_{T\to\infty} T^{-1}\sum_{t=1}^{T} m_{0Pt}m_{0Pt}' = \operatorname*{plim}_{T\to\infty}\frac{1}{T}R'R, \\ \Sigma_{GG} &= \operatorname*{plim}_{T\to\infty} T^{-1}\sum_{t=1}^{T} g_{0t}g_{0t}', \text{ and} \\ \Sigma_{MG} &= \operatorname*{plim}_{T\to\infty} T^{-1}\sum_{t=1}^{T} m_{0t}g_{0t}' = \operatorname*{plim}_{T\to\infty}\frac{1}{T}R'Q. \end{aligned}$$

The above arguments enable us to construct asymptotically valid test from the first stage estimates of $\hat{\theta}$ only. The following theorem gives the limit distribution of $\sqrt{T}\hat{M}_{PT}$ under condition (1).

Theorem 3 *Given* $\hat{\varpi} \xrightarrow{p} \varpi_0$, *the CLT stated in proposition (3) and a suitable ULLN;* $\sqrt{T}\hat{M}_{PT} \xrightarrow{d} N(0, V_1)$ *where*
$V_1 = A_1 \Sigma A_1'$,
$\Sigma = \begin{bmatrix} \Sigma_{MM} & \Sigma_{MG} \\ \Sigma_{GM} & \Sigma_{GG} \end{bmatrix}$ *see Proposition (3),*
$A_1 = \left[I_r, \;\; -J_{M\theta} \times J_{\theta\theta}^{-1}\right]$ *with*
$J_{\theta\theta} = \begin{bmatrix} J_1(\theta_{10}) & 0 \\ 0 & J_2(\theta_{20}) \end{bmatrix}$, $J_{M\theta} = -\operatorname*{plim}_{T \longrightarrow \infty} \left[\frac{\partial M_{0Pt}}{\partial \theta'}\right]$ *and* I_r *is the identity matrix of rank* $r = rank\left(\mathbf{\Sigma}_{MM}\right)$.

From this result, the general form of the CCC misspecification test using PQMLE has the quadratic form

$$T_P = T\hat{M}_{PT}'\hat{V}_{1T}^{-1}\hat{M}_{PT} \tag{25}$$

under the null which has a χ_r^2 limiting distribution, where $\hat{V}_{1T}$ is any consistent estimator for V_1 i.e. $\hat{V}_{1T} = V_1 + o_p(1)$. We want to stress here that to get asymptotically valid test statistic, one has to use $\hat{\upsilon}_t \otimes \left(\hat{r}_t - \widehat{\overline{r}}\right)$ (for FCM) and $\left(\hat{\zeta}_{1t}\hat{\zeta}_{2t} - \hat{\rho}\right)(\hat{r}_t - \overline{r})$ (for CCM) rather than $\zeta_{1t}\zeta_{2t}(r_t - \overline{r})$ when constructing the test indicator $\hat{M}_{PT}$. This has no effect on the numerator of the test statistic, as $\sum_{t=1}^{T}\left[\hat{\zeta}_{1t}\hat{\zeta}_{2t} - \hat{\rho}\right]\hat{r}_t = \sum_{t=1}^{T}\left(\hat{\zeta}_{1t}\hat{\zeta}_{2t} - \hat{\rho}\right)\left(\hat{r}_t - \widehat{\overline{r}}\right) = \sum_{t=1}^{T}\hat{\zeta}_{1t}\hat{\zeta}_{2t}\left(\hat{r}_t - \widehat{\overline{r}}\right)$, but gives us the right expression for the asymptotic variance estimate.

To construct asymptotically valid test statistics we need a consistent expression of V_1. Similar to FQMLE case, we will consider both the robust and OPG version in the following.

3.2.1 Case 2a: Robust PQMLE test

To construct a robust (to non-normality) test of (25), first note that $\hat{J}_{M\theta T}$ can be obtained using the results of Proposition 2, but corresponding to θ only and replacing r_t by $\left(\hat{r}_t - \widehat{\overline{r}}\right)$, the demeaned test variables. Let us define $\overline{R}^*$ as a matrix having rows $(r_t - \overline{r}_t)'$ if r_t is a vector of test variables, or as a vector with typical element $(r_t - \overline{r}_t)$ in case of scalar r_t.

Example 4 *In bivariate case with full moment condition, we have:*

$$\hat{J}_{M\varpi T}^{*} = \frac{1}{T}\begin{bmatrix} \widehat{\overline{R}}^{*\prime}(\hat{Z}_1, 0) \\ \frac{1}{2}\hat{\rho}\widehat{\overline{R}}^{*\prime}(\hat{Z}_1, \hat{Z}_2) \\ \widehat{\overline{R}}^{*\prime}(0, \hat{Z}_2) \end{bmatrix} \tag{26}$$

and for only correlation moment condition it becomes $\hat{J}_{M\varpi T}^{*} = \frac{1}{2T}\hat{\rho}\widehat{\overline{R}}^{*\prime}(\hat{Z}_1, \hat{Z}_2)$.

Besides, from Halunga and Orme (1990, Lemma 1), for i^{th} variable we have:

$$\hat{J}_i = \frac{1}{2T}\begin{bmatrix} \hat{C}_i'\hat{C}_i & \hat{C}_i'\hat{X}_i \\ \hat{X}_i'\hat{C}_i & \hat{X}_i'\hat{X}_i \end{bmatrix} + \frac{1}{T}\begin{bmatrix} \hat{F}_i'\hat{F}_i & 0 \\ 0 & 0 \end{bmatrix}, \tag{27}$$

where F_i, C_i and X_i have rows $f'_{it} = \frac{w'_{it}}{\sqrt{h_{it}}}$, $c'_{it} = \frac{1}{h_{it}}\frac{\partial h_{it}}{\partial \varphi_i}$ and $x'_{it} = \frac{1}{h_{it}}\frac{\partial h_{it}}{\partial \eta_i}$; all evaluated at PQMLE $\hat{\theta}'_i = \left(\hat{\varphi}'_i, \hat{\eta}'_i\right)$.

Combining the above two results, the next lemma provides an expression for the robust consistent variance estimator $\hat{V}^r_{1T}$ for bivariate case which can be generalized in an obvious way.

Lemma 6 *Suppose $\hat{M}^j_{PT}$ is the joint CM test indicator (partial QMLE case) for $N = 2$, a robust (to non-normality) consistent estimator of V_1 is given by*

$$\hat{V}^r_{1T} = \frac{1}{T}\hat{W}^{r\prime}\hat{W}^r = \frac{1}{T}\hat{A}^r_1\hat{B}'\hat{B}A^{r\prime}_1$$

where

$W^r = BA^{r\prime}_1$, $\hat{B} = \left[\hat{R},\ \hat{Q}_1,\ \hat{Q}_2\right]$*; i.e. $\hat{B}$ has rows* $\left(\hat{m}'_{Pt},\ \hat{g}'_{1t},\ \hat{g}'_{2t}\right)$,

$\hat{A}^r_1 = \left[I_r :\ -\hat{J}^*_{M\varpi T}\hat{J}^{-1}_{\theta\theta}\right]$, *with* $\hat{J}_{\theta\theta} = \begin{bmatrix} \hat{J}_1 & 0 \\ 0 & \hat{J}_2 \end{bmatrix}$; *$\hat{J}_i$ is obtained from (27), $\hat{J}^*_{M\varpi T}$ is given in (26) and I_r is the identity matrix of rank $r = rank\left(\mathbf{\Sigma}_{MM}\right)$.*

Remark 4 *For $N = 2$, we can write $\hat{A}^r_1$ and $\hat{W}^r$ as*

$$\begin{aligned}
\hat{A}^r_1 &= \left[I_r :\ -\frac{1}{T}\begin{bmatrix} \widehat{\overline{R}}^{*\prime}(\hat{Z}_1, 0) \\ \hat{\rho}\widehat{\overline{R}}^{*\prime}(\hat{Z}_1, \hat{Z}_2) \\ \widehat{\overline{R}}^{*\prime}(0, \hat{Z}_2) \end{bmatrix}\begin{bmatrix} \hat{J}^{-1}_1 & 0 \\ 0 & \hat{J}^{-1}_2 \end{bmatrix}\right] \\
&= \left[I_r :\ -\frac{1}{T}\begin{bmatrix} \widehat{\overline{R}}^{*\prime}\hat{Z}_1\hat{J}^{-1}_1 & 0 \\ \hat{\rho}\widehat{\overline{R}}^{*\prime}\hat{Z}_1\hat{J}^{-1}_1 & \hat{\rho}\widehat{\overline{R}}^{*\prime}\hat{Z}_2\hat{J}^{-1}_2 \\ 0 & \widehat{\overline{R}}^{*\prime}\hat{Z}_2\hat{J}^{-1}_2 \end{bmatrix}\right]
\end{aligned}$$

and

$$\hat{W}^r = \hat{R} - \frac{1}{T}\left(\hat{Q}_1\hat{J}^{-1}_1\hat{Z}'_1\widehat{\overline{R}}^*, \frac{1}{2}\hat{\rho}\left(\hat{Q}_1\hat{J}^{-1}_1\hat{Z}'_1 + \hat{Q}_2\hat{J}^{-1}_2\hat{Z}'_2\right)\widehat{\overline{R}}^*, \hat{Q}_2\hat{J}^{-1}_2\hat{Z}'_2\widehat{\overline{R}}^*\right)$$

respectively.

Tests that are based on this estimator $\hat{V}^r_{1T}$ will be referred as robust PQMLE test and will be denoted as $T^{(r)}_P$. Note that exploiting the FOC $\hat{Q}'_i\iota_T \equiv 0$ for $\hat{\theta}_i$, $i = 1, 2$, we have $\hat{W}^{r\prime}\iota_T \equiv \hat{R}'\iota_T$, where ι_T is the $(T \times 1)$ column vector of ones implying

$$T^{(r)}_p = T\hat{M}'_{PT}\hat{V}^{r-1}_{1T}\hat{M}_{PT} = \iota'_T\hat{W}^r\left(\hat{W}^{r\prime}\hat{W}^r\right)^{-1}\hat{W}^{r\prime}\iota_T \tag{28}$$

which can be obtained as $T - RSS$ from a regression of ι_T on $\hat{W}^r$.

3.2.2 Case 2(b): OPG PQMLE test

The following lemma provides an expression for a consistent estimator of the variance-covariance matrix V_1 under normality.

Lemma 7 *Assuming that the specification of the log-likelihood for the joint estimation of parameters in section 3.1 is correct (i.e. $\xi_t \sim N(0, I_N)$); $\hat{A}_1\hat{\Sigma}\hat{A}_1' - V_1 = o_p(1)$ where*

$$\hat{A}_1 = \left[I_r : -\tfrac{1}{T}\hat{R}'\hat{S} \begin{pmatrix} \left(\frac{1}{T}\hat{Q}_1'\hat{S}_1\right)^{-1} & 0 \\ 0 & \left(\frac{1}{T}\hat{Q}_2'\hat{S}_2\right)^{-1} \end{pmatrix} \right],$$

$\hat{\Sigma} = \frac{1}{T}\hat{B}'\hat{B}$ *and*

$\hat{B} = \left[\hat{R},\ \hat{Q}_1,\ \hat{Q}_2\right]$ *i.e. $\hat{B}$ has rows $(\hat{m}_{Pt}',\ \hat{g}_{1t}',\ \hat{g}_{2t}')$*

Hence, V_1 can be consistently estimated by $\hat{V}_{1T} = \dfrac{1}{T}\hat{W}'\hat{W}$ where

$\hat{W} = \hat{B}\hat{A}_1' = \hat{R} - \hat{Q}_1\left(\hat{S}_1'\hat{Q}_1\right)^{-1}\hat{S}_1'\hat{R} - \hat{Q}_2\left(\hat{S}_2'\hat{Q}_2\right)^{-1}\hat{S}_2'\hat{R}.$

Again, the FOC $\hat{Q}_i'\iota_T \equiv 0$ implies $\hat{W}'\iota_T \equiv \hat{R}'\iota_T$ so that

$$T_p = T\hat{M}_{PT}'\hat{V}_{1T}^{-1}\hat{M}_{PT} = \iota_T'\hat{W}\left(\hat{W}'\hat{W}\right)^{-1}\hat{W}'\iota_T \qquad (29)$$

which can be obtained as $T - RSS$ from a regression of ι_T on $\hat{W}$.

3.2.3 Summary: PQMLE

Hence, in case of PQMLE, we again consider the following four test statistics:

5. Robust (to non-normality) FCM test:

$$T_p^{j(r)} = T\hat{M}_{pT}^{j(r)\prime}\left(\hat{V}_{1T}^{j(r)}\right)^{-1}\hat{M}_{pT}^{j(r)} \qquad (30)$$

6. OPG FCM test

$$T_p^{j} = T\hat{M}_{pT}^{j\prime}\left(\hat{V}_{1T}^{j}\right)^{-1}\hat{M}_{pT}^{j} \qquad (31)$$

7. Robust (to non-normality) CCM test

$$T_p^{c(r)} = T\hat{M}_{pT}^{c(r)\prime}\left(\hat{V}_{1T}^{c(r)}\right)^{-1}\hat{M}_{pT}^{c(r)} \qquad (32)$$

8. OPG CCM test

$$T_p^{c} = T\hat{M}_{pT}^{c\prime}\left(\hat{V}_{1T}^{c}\right)^{-1}\hat{M}_{pT}^{c} \qquad (33)$$

where the robust variance estimator ($\hat{V}_{1T}^{j(r)}$ and $\hat{V}_{1T}^{c(r)}$) and OPG variance estimator ($\hat{V}_{1T}^{j}$ and $\hat{V}_{1T}^{c}$) can be obtained using Lemma 6 and Lemma 7 respectively.

4 Analysis of Tse's LM test

Tse (2000) proposed a LM test for the multivariate CCC-GARCH model against the alternative that the correlation are changing as functions of the previous standardized residuals, having the form

$$\rho_{ijt} = \rho_{ij} + \tau_{ij} y_{i,t-1} y_{j,t-1} \text{ or } \Gamma_t = \Gamma + \Delta \odot y_{t-1} y'_{t-1} \tag{34}$$

where Δ is a symmetric parameter matrix with the leading diagonal elements equal to zero. Note that (34) does not define a particular alternative to CCC as Γ_t is not necessarily a positive definite matrix for all t. Therefore, Silvennoinen and Teräsvirta (2008) interpreted this as a general misspecification test. And the null hypothesis is $H_0 : \Delta = 0$ or $H_0 : \text{vecl}(\Delta) = 0$. Note that τ_{ij}, $1 \leq i < j \leq N$ are $\frac{N(N-1)}{2}$ additional parameters in the extended model. Under this setting Tse proposed the following statistic:

$$\begin{aligned} LM_T &= \widehat{\widehat{s}}' \left(\widehat{\widehat{S}}' \widehat{\widehat{S}} \right)^{-1} \widehat{\widehat{s}} && (35) \\ &= \iota'_T \widehat{\widehat{S}} \left(\widehat{\widehat{S}}' \widehat{\widehat{S}} \right)^{-1} \widehat{\widehat{S}}' \iota_T && (36) \end{aligned}$$

where $\widehat{\widehat{s}}$ is the $\left(\left(N' + \frac{N(N-1)}{2} \right) \times 1 \right)$ score vector, $\widehat{\widehat{S}}$ is $\left(T \times N' + \frac{N(N-1)}{2} \right)$ matrix, with rows of partial derivatives of the log likelihood function and ι_T is the $(T \times 1)$ column vector of ones. [9] Note that (36) can be interpreted as T times R^2, where R^2 is the uncentered coefficient of determination of the regression of ι_T on $\widehat{\widehat{S}}$. Under the usual regularity conditions LM_T is asymptotically distributed as $\chi^2_{\frac{N(N-1)}{2}}$.

It is informative to note that this LM_T can be interpreted as a test of moment condition $\text{E}\left[\text{vecl}\left(\varepsilon^*_t \varepsilon^{*\prime}_t - \Gamma^{-1} | \mathcal{F}_{t-1}\right)\right] = 0$ where ε^*_t is the transformed standardized errors as given in (6). This test is based on the FQMLE approach and can not be implemented directly within the PQMLE framework. We can, however, modify the test indicator in such a way so that the testing procedure based on PQMLE developed in previous section can be employed. Again the procedure will be demonstrated in the bivariate context.

In the bivariate case, $\varepsilon^*_t = (\varepsilon^*_{1t}, \varepsilon^*_{2t})' = \Gamma^{-1} \zeta_t = \dfrac{1}{1-\rho^2} \begin{bmatrix} \zeta_{1t} - \rho \zeta_{2t} \\ \zeta_{2t} - \rho \zeta_{1t} \end{bmatrix}$; and

[9] For the full expressions of the first partial derivatives of the likelihood function l_t with respect to the model parameters, readers are referred to Tse (2000).

the implicit null of CCC is $\mathrm{E}\left[\varepsilon_{1t}^{*}\varepsilon_{2t}^{*}+\dfrac{\rho}{1-\rho^{2}}|\mathcal{F}_{t-1}\right]=0$.Note that

$$\begin{aligned}
\varepsilon_{1t}^{*}\varepsilon_{2t}^{*}+\frac{\rho}{1-\rho^{2}} &= \frac{1}{\left(1-\rho^{2}\right)^{2}}\left[\left\{\zeta_{1t}-\rho\zeta_{2t}\right\}\left\{\zeta_{2t}-\rho\zeta_{1t}\right\}+\rho\left(1-\rho^{2}\right)\right] \\
&= \frac{1}{\left(1-\rho^{2}\right)^{2}}\left[-\rho\left(\zeta_{1t}^{2}-1\right)+\left(1+\rho^{2}\right)\left(\zeta_{1t}\zeta_{2t}-\rho\right)-\rho\left(\zeta_{2t}^{2}-1\right)\right] \\
&= \frac{1}{\left(1-\rho^{2}\right)^{2}}\pi' v_{t}
\end{aligned}$$

where $\pi'(\rho)\equiv\pi'=\left(-\rho,\left(1+\rho^{2}\right),-\rho\right),\ v_{t}'=\left(\zeta_{1t}^{2}-1,\ \zeta_{1t}\zeta_{2t}-\rho,\ \zeta_{2t}^{2}-1\right).$

4.1 FQMLE case

Assuming that r_t is a scalar and ignoring the irrelevant factor of proportionality, $1/\left(1-\rho^{2}\right)^{2}$, define Tse's "modified" indicator as

$$\hat{M}_{FT}^{t}=\frac{1}{T}\sum_{t=1}^{T}\hat{m}_{Ft}^{t}\left(\hat{\varpi}\right)=\frac{1}{T}\sum_{t=1}^{T}\hat{\pi}'\hat{v}_{t}\hat{r}_{t}=\frac{1}{T}\sum_{t=1}^{T}\hat{\pi}'\hat{m}_{Ft}^{j}\left(\hat{\varpi}\right) \tag{37}$$

where the superscript t represents Tse's indicator and $\hat{m}_{Ft}^{j}\left(\hat{\varpi}\right)$ is the contribution of t^{th} observation to the test indicator for FCM test $\hat{M}_{FT}^{j}$, all evaluated at FQMLE $\hat{\varpi}$.

Corollary 1 *From Theorem (2), $\sqrt{T}\hat{M}_{FT}^{j}\xrightarrow{d}N\left(0,V^{j}\right)$; hence $\sqrt{T}\hat{M}_{FT}^{t}\xrightarrow{d}N\left(0,V^{t}\right)$ where $V^{t}=\pi'V^{j}\pi$.*

Then, an equivalent procedure to Tse's LM test, LM_T can be obtained applying the CM testing framework developed in section 3.1 by using $\hat{M}_{FT}^{t}$ and constructing the OPG version of the test as:

$$T_{F}^{t}=T\hat{M}_{FT}^{t\prime}\left(\hat{V}_{T}^{t}\right)^{-1}\hat{M}_{FT}^{t}=T\hat{M}_{FT}^{t\prime}\left(\hat{\pi}'\hat{V}^{j}\hat{\pi}\right)^{-1}\hat{M}_{FT}^{t} \tag{38}$$

where $\hat{V}^{j}=\hat{\Sigma}_{MM}-\hat{\Sigma}_{MG}^{*}\hat{\Sigma}_{GG}^{*-1}\hat{\Sigma}_{GM}^{*}$ comes from OPG FCM test given in (20).

As noted earlier that Tse assumes the generalized IM equality to hold while developing his OPG version of LM test which may not be robust under non-normality. Using the robust variance estimator as given in Lemma (4), we can now robustify this LM test i.e.

$$T_{F}^{t(r)}=T\hat{M}_{FT}^{t\prime}\left(\hat{V}_{T}^{t(r)}\right)^{-1}\hat{M}_{FT}^{t}=T\hat{M}_{FT}^{t\prime}\left(\hat{\pi}\hat{V}_{T}^{j(r)}\hat{\pi}'\right)^{-1}\hat{M}_{FT}^{t} \tag{39}$$

where $\hat{V}_{T}^{t(r)}=\hat{\pi}\hat{V}_{T}^{j(r)}\hat{\pi}'$ and $\hat{V}_{T}^{j(r)}$ is the variance-covariance matrix defined in (19).

4.2 PQMLE case

For PQMLE, to obtain asymptotically valid test statistic of the CCC assumption, we apply the demeaning technique so that the estimation effect from ρ asymptotically negligible (i.e. condition (1) holds) in the following way:

$$\begin{aligned} \hat{M}_{PT}^{t} &= \frac{1}{T}\sum_{t=1}^{T}\hat{m}_{Pt}^{t}\left(\hat{\theta},\hat{\rho}\right)=\frac{1}{T}\sum_{t=1}^{T}\left(\hat{\varepsilon}_{1t}^{*}\hat{\varepsilon}_{2t}^{*}+\frac{\hat{\rho}}{1-\hat{\rho}^{2}}\right)\left(\hat{r}_{t}-\widehat{\overline{r}}\right) \\ &= \frac{1}{T}\sum_{t=1}^{T}\hat{\pi}'\hat{v}_{t}\left(\hat{r}_{t}-\widehat{\overline{r}}\right)=\frac{1}{T}\sum_{t=1}^{T}\hat{\pi}'\hat{m}_{Pt}^{j}\left(\hat{\theta},\hat{\rho}\right) \end{aligned} \tag{40}$$

where the last equality follows from (24).

Also note that since $\frac{1}{T}\sum_{t=1}^{T}\left(\hat{\varepsilon}_{1t}^{*}\hat{\varepsilon}_{2t}^{*}+\frac{\hat{\rho}}{1-\hat{\rho}^{2}}\right)\neq 0$ unless FQMLE is employed, $\hat{M}_{PT}^{t}\neq\frac{1}{T}\sum_{t=1}^{T}\left(\hat{\varepsilon}_{1t}^{*}\hat{\varepsilon}_{2t}^{*}+\frac{\hat{\rho}}{1-\hat{\rho}^{2}}\right)\hat{r}_{t}$. Now using the following corollary and the results of preceding section, we can construct asymptotically valid tests of the CCC assumption employing Tse's modified indicator $\hat{M}_{PT}^{t}$ based on PQMLE.

Corollary 2 *From theorem (3), $\sqrt{T}\hat{M}_{PT}^{j}\xrightarrow{d}N\left(0,V_{1}^{j}\right)$; hence $\sqrt{T}\hat{M}_{PT}^{t}\xrightarrow{d}N\left(0,V_{1}^{t}\right)$ where $V_{1}^{t}=\pi'V_{1}^{j}\pi$.*

In particular, the OPG and robust test statistics with $\hat{M}_{PT}^{t}$ can be constructed easily by:

$$T_{P}^{t}=T\hat{M}_{PT}^{t\prime}\left(\hat{V}_{1T}^{t}\right)^{-1}\hat{M}_{PT}^{t}=T\hat{M}_{PT}^{t\prime}\left(\hat{\pi}'\hat{V}_{1T}^{j}\hat{\pi}\right)^{-1}\hat{M}_{PT}^{t} \tag{41}$$

and

$$T_{p}^{t(r)}=T\hat{M}_{PT}^{t\prime}\left(\hat{V}_{1T}^{t(r)}\right)^{-1}\hat{M}_{PT}^{t}=T\hat{M}_{PT}^{t\prime}\left(\hat{\pi}'\hat{V}_{1T}^{j(r)}\hat{\pi}\right)^{-1}\hat{M}_{PT}^{t} \tag{42}$$

where $\hat{V}_{1T}^{j(r)}$ and $\hat{V}_{1T}^{j}$ are given in (30) and (31) respectively.

Alternatively, $\hat{V}_{1T}^{t}$ can be constructed as $\hat{V}_{1T}^{t}=\frac{1}{T}\hat{W}^{t\prime}\hat{W}^{t}$ where $\hat{W}^{t}=\hat{B}\hat{A}_{1}'\hat{\pi}$, and $\hat{B}$ and $\hat{A}_{1}$ are defined as before. Similarly, $\hat{V}_{1T}^{t(r)}=\frac{1}{T}\hat{W}^{t(r)\prime}\hat{W}^{t(r)}$ where $\hat{W}^{t(r)}=\hat{B}\hat{A}_{1}^{r\prime}\hat{\pi}$. And then the test statistics are obtained as $T-RSS$ from a regression of ι_{T} on $\hat{W}^{t}$ or $\hat{W}^{t(r)}$.

Remark 5 *Notice that if we demean test indicator $\hat{M}_{PT}^{j}$ only for the components which involve $\hat{\rho}$ then, as noted before, OPG-FCM and robust FCM procedure will be asymptotically valid but (41) or (42) can not be employed. In other words, it is necessary to demean all elements of the moment condition to obtain valid test statistics based on $\hat{M}_{PT}^{t}$.*

Test indicator	Tests
$\hat{M}_T^j$	T_F^j, $T_F^{j(r)}$, T_P^j and $T_p^{j(r)}$
$\hat{M}_T^c$	T_F^c, $T_F^{c(r)}$, T_P^c and $T_p^{c(r)}$
$\hat{M}_T^t$	T_F^t, $T_F^{t(r)}$, T_P^t and $T_p^{t(r)}$

Table 1: Various test indicatots and tests considered in the simulation

5 Monte Carlo Evidence

In this section, we present Monte Carlo evidence on finite sample size and power performance of the 12 tests defined in (19)-(22), (30)-(33), (38), (39), (41) and (42). To recapitulate, we consider three test indicators FCM $\left(\hat{M}_T^j\right)$, CCM $\left(\hat{M}_T^c\right)$ and Tse's "modified" indicator $\left(\hat{M}_T^t\right)$; each having four versions (FQMLE OPG, FQMLE robust, PQMLE OPG and PQMLE robust). Table 1 displays various test indicators and associated test statistics under consideration.

The parameter values for the null and alternative DGPs are taken from the existing literature (e.g. Engle and Ng, 1993; Tse, 2000; Lundbergh and Teräsvirta, 2002; Halunga and Orme, 2009). For each experiment, three series of 1200, 900 and 700 data realizations were generated with the first 200 observations being discarded to avoid initialization effects, yielding sample sizes of $T = 1000$, 700 and 300 respectively. Each model is replicated and estimated, 10,000 times (for size experiments) and 2000 times (for robustness to non-normality and power experiments), both by FQMLE and PQMLE. Next, the above mentioned 12 test statistics are calculated. For this simulation study, we consider the product of the 1-period lagged standardized residuals as the scalar test variable, i.e. $\hat{r}_t = \zeta_{1,t-1}\zeta_{2,t-1}$ to calculate all 12 test statistics.

5.1 Size

To assess the size properties of the tests we consider a bivariate AR(1)-CCC-GARCH (1,1) data generating process (DGP) as our null model i.e.[10]

$$\begin{aligned}
y_{it} &= \varphi_{i0} + \varphi_{i1} y_{i,t-1} + \varepsilon_{it}, \quad i = 1,2 \\
Var\left(\varepsilon_t | \mathcal{F}_{t-1}\right) &= H_t \Rightarrow \mathrm{E}\left[\varepsilon_{it}^2 | \mathcal{F}_{t-1}\right] = h_{ii,t}, \; \varepsilon_t = H_t^{1/2}\left(\varpi\right)\xi_t; \; \xi_t \sim N(0, I) \\
h_{ii,t} &= \alpha_{i0} + \alpha_{i1}\varepsilon_{i,t-1}^2 + \beta_{i1} h_{i,t-1} \\
H_t &= D_t \Gamma D_t; \; D_t = \begin{bmatrix} \sqrt{h_{11t}} & 0 \\ 0 & \sqrt{h_{22t}} \end{bmatrix} \text{ and } \Gamma = \begin{bmatrix} 1 & \rho \\ \rho & 1 \end{bmatrix} \qquad (43)
\end{aligned}$$

Four experiments are considered and the corresponding true parameter vectors are presented in Table (2). E1 and E2 represent models with relatively

[10] We also consider CCC-GARCH (1,1) DGP (i.e. assuming zero or known conditional mean) to evaluate size property. However, to save space and due to the qualitative similarity we will discuss only AR(1)-CCC GARCH (1,1) results.

high persistence ($\alpha + \beta = 0.95$ for E1 and $\alpha + \beta = 0.90$ for E2) while E3 and E4 correspond to relatively low persistence model ($\alpha + \beta = 0.70$ for both E3 and E4). On the other hand, E1 and E3 represent high correlation models and E2 and E4 represent low correlation models. Hence E1, E2, E3 and E4 represent high-persistent-high-correlation, high-persistent-low-correlation, low-persistent-high-correlation and low-persistent-low-correlation specification respectively.

Table (3) reports the actual rejection frequencies when the null of CCC is true and $\xi_t \sim N(0, I)$. The results are reported for a nominal size of 5%. It can be seen that for low correlation DGPs (E2 and E4), the empirical sizes for all test statistics, except OPG FCM (T_F^j and T_P^j), are very close to the nominal size of 5%. Although the OPG version based on FQMLE of other two tests (T_F^t and T_F^c) slightly overrejects when $T = 500$, size property improves as T increases. Interestingly, PQMLE and robust version of all these tests demonstrate better performance even in small sample. In case of experiments with high correlation, particularly with high persistence volatility (E1), all FQMLE-OPG tests (T_F^t, T_F^c and T_F^j) are slightly oversized; robust version of these statistics, however, corrects this size distortion. Tests based on $\hat{M}_T^t$ perform comparatively better for high correlation case.

Our finding that size performance depends on correlation but volatility persistence does not have much impact on rejection frequencies, are in line with that of Tse (2000). He reports "correlations seem to play a role in determining the rate of convergence to the nominal size. Models with low correlations are less subject to over-rejection in small samples....the persistence of the conditional variance does not have much effect.. " (Tse, 2000 pp:115).

In summary, tests with Tse's modified indicator based on PQMLE (i.e. T_P^t and $T_P^{t(r)}$) provide the most reliable size property; robust versions, in general, perform better than OPG; OPG-FCM tests are slightly oversized and all test statistics perform better in low correlation experiments.

5.2 Effect of Non-normality

Table (4) presents the actual rejection frequencies when the null of CCC is true and $\xi_t \sim t(6)$, $\xi_t \sim t(8)$ and $\xi_t \sim t(10)$. The inclusion of $t(6)$ offers some evidence on the robustness of the procedure to violations of the underlying moment assumptions (cf. Assumption 2.5). First thing to observe that all OPG-FQMLE tests (T_F^t, T_F^c and T_F^j) overrejects the null for both high and low correlation models, but more severe in high correlation models. Particularly, note that Tse's LM test (T_F^t) is sensitive to the departure from normality assumption. Interestingly, the OPG-PQMLE tests T_P^t and T_P^c demonstrate robust size performance under non-normality. The robust version of FQMLE tests reduce the overrejection rate considerably and in fact $T_F^{c(r)}$ and $T_F^{c(r)}$ are slightly undersized for low persistent-low correlation model with $t(8)$ and $t(10)$ errors. The empirical size of robust tests based on Tse's indicator (particularly $T_P^{t(r)}$) and CCM tests, in general, are close to nominal level of 5% while all versions of the FCM tests show unreliable size property (in general, they are oversized).

	E1	E2	E3	E4
$\varphi_1' = (\varphi_{10}, \varphi_{11})$	$(1.00, 0.10)$			
$\eta_1' = (\alpha_{10}, \alpha_{11}, \beta_{11})$	$(0.01, 0.15, 0.80)$	$(0.01, 0.15, 0.80)$	$(0.40, 0.40, 0.30)$	$(0.40, 0.40, 0.30)$
$\varphi_2' = (\varphi_{20}, \varphi_{21})$	$(1.00, 0.50)$			
$\eta_2' = (\alpha_{20}, \alpha_{21}, \beta_{21})$	$(0.05, 0.20, 0.70)$	$(0.05, 0.20, 0.70)$	$(0.20, 0.50, 0.20)$	$(0.20, 0.50, 0.20)$
ρ	0.80	0.20	0.80	0.20

Table 2: True parameter values for size simulation

	DGP:E1			DGP:E2			DGP:E3			DGP:E4		
	$T=1000$	$T=700$	$T=500$	$T=1000$	$T=700$	$T=500$	$T=1000$	$T=700$	$T=500$	$T=1000$	$T=700$	$T=500$
T_F^t	6.49	6.84	7.42	5.54	5.51	6.31	6.41	6.83	7.00	5.63	5.54	6.65
$T_F^{t(r)}$	5.61	5.37	5.51	4.75	4.33	4.70	5.28	5.47	5.18	4.70	4.46	4.88
T_P^t	5.25	5.02	5.13	5.07	4.65	5.30	4.99	4.80	4.59	5.08	4.89	5.57
$T_P^{t(r)}$	5.01	4.79	4.83	4.66	4.32	4.68	5.18	5.04	4.69	4.70	4.39	4.91
T_F^c	6.43	6.99	6.96	5.76	5.63	6.37	6.37	6.99	7.08	5.74	5.51	6.47
$T_F^{c(r)}$	5.64	5.90	5.58	4.91	4.52	4.68	5.59	5.78	5.71	4.86	4.54	4.78
T_P^c	5.60	6.10	6.05	5.29	4.99	5.67	4.51	4.56	4.13	5.32	4.99	5.89
$T_P^{c(r)}$	5.70	6.37	6.12	4.90	4.50	4.66	5.85	6.49	6.23	4.84	4.52	4.73
T_F^j	7.80	9.02	8.99	6.66	6.81	7.53	7.57	8.75	8.78	6.80	6.83	7.82
$T_F^{j(r)}$	5.85	5.96	5.12	4.92	4.42	4.15	5.29	5.74	5.05	4.76	4.32	4.01
T_P^j	7.07	6.96	6.75	5.71	5.71	6.21	5.73	6.23	5.56	6.21	6.06	6.69
$T_P^{j(r)}$	5.44	5.54	4.96	4.86	4.47	4.16	5.26	5.89	5.08	4.77	4.26	4.03

Table 3: Empirical size with normal errors

	DGP:E1			DGP:E2			DGP:E3			DGP:E4		
	T=1000	T=700	T=500	T=1000	T=700	T=500	T=1000	T=700	T=500	T=1000	T=700	T=500
						$t(6)$						
T_F^t	11.05	12.15	12.65	7.70	7.35	7.40	11.40	10.65	11.80	7.30	8.50	9.10
$T_F^{t(r)}$	7.10	6.70	6.90	5.70	4.65	4.10	7.50	6.25	6.40	4.85	5.00	5.20
T_P^t	5.75	5.70	5.75	6.45	5.70	5.25	5.60	4.90	5.60	5.15	6.15	6.65
$T_P^{t(r)}$	5.40	5.50	5.20	5.65	4.45	4.25	6.05	5.10	5.25	4.85	4.95	5.40
T_F^c	11.60	12.30	10.75	7.80	6.90	7.15	10.05	10.65	12.45	7.00	8.45	9.05
$T_F^{c(r)}$	7.80	7.30	5.70	5.20	4.30	4.15	6.35	6.95	6.85	4.80	5.40	5.55
T_P^c	8.05	7.75	6.80	6.45	5.50	5.75	5.70	5.55	4.65	5.80	6.75	7.50
$T_P^{c(r)}$	8.05	7.55	6.45	5.25	4.30	4.20	7.45	7.55	7.05	4.70	5.45	5.70
T_F^j	16.50	16.55	17.05	11.40	12.15	11.85	15.85	15.10	17.55	9.95	11.05	12.55
$T_F^{j(r)}$	8.05	6.10	6.05	4.65	3.55	3.85	6.45	5.10	5.25	3.45	3.65	3.45
T_P^j	10.45	10.85	9.55	8.80	8.75	8.90	11.05	9.55	8.95	8.35	8.80	10.00
$T_P^{j(r)}$	6.10	4.95	4.35	4.65	3.75	4.20	6.50	4.90	3.90	3.75	3.70	3.45
						$t(8)$						
T_F^t	8.80	9.25	9.90	6.10	6.10	7.20	8.70	9.55	11.90	7.05	7.95	10.05
$T_F^{t(r)}$	6.15	5.80	6.25	4.80	4.15	4.30	6.65	5.40	7.50	5.10	4.95	5.70
T_P^t	5.30	5.65	4.40	5.35	4.75	4.95	5.20	5.05	5.75	5.65	6.30	7.65
$T_P^{t(r)}$	5.30	5.40	4.05	4.65	4.05	4.25	5.45	4.70	5.90	5.25	4.90	5.65
T_F^c	10.05	9.90	10.25	5.80	6.75	6.70	7.80	9.20	10.05	6.15	7.65	8.80
$T_F^{c(r)}$	7.10	7.05	7.00	4.30	4.85	4.05	6.05	6.50	6.25	4.60	5.40	6.20
T_P^c	7.55	6.80	7.40	4.85	5.70	5.20	4.00	5.30	5.05	5.30	6.60	7.85
$T_P^{c(r)}$	7.65	7.25	7.85	4.20	4.80	4.00	5.80	7.35	7.50	4.45	5.20	6.00
T_F^j	12.60	12.75	14.55	8.70	8.90	10.05	12.95	13.55	15.55	7.80	9.75	11.10
$T_F^{j(r)}$	7.50	6.30	5.55	5.20	3.95	3.80	6.15	5.95	7.20	3.75	4.15	4.05

	DGP:E1			DGP:E2			DGP:E3			DGP:E4		
	T=1000	T=700	T=500	T=1000	T=700	T=500	T=1000	T=700	T=500	T=1000	T=700	T=500
T_P^j	9.45	8.95	8.00	7.15	7.15	7.20	8.65	8.95	8.40	6.45	7.85	9.40
$T_P^{j(r)}$	5.85	5.25	4.05	5.25	4.10	4.15	4.95	5.60	6.00	3.60	4.05	4.15
						$t(10)$						
T_F^t	8.95	8.45	9.30	5.65	5.50	6.85	8.25	8.45	9.40	7.15	7.50	7.20
$T_F^{t(r)}$	6.70	5.55	5.80	4.35	3.95	4.15	6.15	6.05	6.25	5.90	5.55	4.55
T_P^t	6.20	5.00	5.20	4.65	4.55	4.65	5.75	4.90	5.80	6.15	6.35	5.55
$T_P^{t(r)}$	6.45	5.05	5.50	4.45	4.05	4.05	5.25	5.05	5.35	5.90	5.50	4.50
T_F^c	8.30	8.45	10.30	5.55	6.25	5.95	7.80	7.30	8.85	6.35	7.30	6.55
$T_F^{c(r)}$	6.05	5.75	7.05	3.85	4.50	3.95	6.35	5.00	5.65	4.90	5.05	4.40
T_P^c	7.45	6.35	7.75	4.60	5.55	4.75	4.95	4.55	4.65	5.55	6.60	5.55
$T_P^{c(r)}$	6.80	6.20	7.20	3.85	4.60	3.95	7.05	6.25	6.05	4.85	5.20	4.20
T_F^j	11.95	11.90	14.25	7.70	7.60	9.55	11.10	11.20	13.35	8.35	9.70	9.40
$T_F^{j(r)}$	7.50	6.40	6.00	4.40	3.00	4.15	6.25	5.95	5.50	4.10	3.80	3.70
T_P^j	8.05	8.35	8.30	6.00	5.70	7.40	8.70	8.00	8.75	7.25	7.95	7.60
$T_P^{j(r)}$	6.35	5.40	5.25	4.35	3.10	4.10	5.95	4.80	5.30	4.10	4.00	3.55

Table 4: Empirical size under non-normality (with t errors)

5.3 Impact of univariate volatility misspecification

We consider four experiments (M1, M2, M3 and M4) in the regression context to investigate the effect of misspecification in the univariate GARCH model when the true correlation structure is constant. M1 and M3 has low correlation ($\rho = 0.20$) and M2 and M4 follow high correlation ($\rho = 0.80$) structure. The conditional mean parameters are the same as in the size experiments. For both M1 and M2, the univariate volatility specification for first variable is given by high persistence GARCH (1,1) model i.e. $h_{11,t} = 0.01 + 0.15\varepsilon_{1,t-1}^2 + 0.80h_{1,t-1}$ while the second variable follows the EGARCH(1,1) model of Nelson (1991) with parameter values considered by Engle and Ng (1993) and Halunga and Orme (2009) in their simulation study: $\log(h_{22,t}) = -0.23 + 0.9\log(h_{2,t-1}) + 0.25\left[\left|\xi_{t-1}\right| - 0.3\xi_{t-1}\right]$. On the other hand, in experiment M3 and M4, we assume that both variables are subject to volatility spillover (i.e. ECCC model) in the following way:

$$h_{ii,t} = \alpha_{i0} + \alpha_{i1}\varepsilon_{i,t-1}^2 + \beta_{ii}h_{i,t-1} + \beta_{ij}h_{j,t-1};\ i = 1, 2 \text{ and } i \neq j$$

with

$$\begin{aligned} (\alpha_{10}, \alpha_{11}, \beta_{11}, \beta_{12}) &= (0.01, 0.15, 0.80, 0.02) \text{ and} \\ (\alpha_{20}, \alpha_{21}, \beta_{22}, \beta_{21}) &= (0.05, 0.20, 0.70, 0.03)\,. \end{aligned}$$

Table(5) reports the results of the simulation study based on 2000 replications where the data is generated with normal errors and the nominal level of significance is set to 5%. It can be observed that the tests are robust to volatility spillover case (i.e. M3 and M4). On the other hand these tests seem to be non-robust with GARCH-EGARCH-High correlation specification (M2), all tests overreject the null of CCC. It is to be noted here that due to the fact that FCM test indicator involve the volatility moment condition, these tests expectedly display the power to pick the misspecification. For M3 (GARCH-EGARCH-low correlation), the tests, except FCM, are not that much sensitive to univariate conditional variance misspecification.

	DGP:M1			DGP:M2			DGP:M3			DGP:M4		
	$T=1000$	$T=700$	$T=500$	$T=1000$	$T=700$	$T=500$	$T=1000$	$T=700$	$T=500$	$T=1000$	$T=700$	$T=500$
T_F^t	7.75	8.60	9.21	19.50	20.55	21.36	5.60	6.45	5.65	6.25	7.10	7.70
$T_F^{t(r)}$	4.70	5.05	5.21	14.95	12.95	11.57	4.70	5.00	4.50	5.25	5.90	5.55
T_P^t	5.60	5.95	6.32	10.55	9.65	9.05	5.10	5.65	5.00	4.50	5.15	4.60
$T_P^{t(r)}$	4.80	5.15	5.54	9.25	8.55	8.33	4.70	5.10	4.45	4.95	5.20	4.30
T_F^c	6.45	6.60	6.87	14.85	15.15	15.35	5.70	6.25	6.45	6.90	7.70	7.00
$T_F^{c(r)}$	5.65	4.85	4.41	13.05	12.45	12.05	4.90	5.00	4.70	6.05	6.80	5.65
T_P^c	6.45	6.00	5.70	12.60	12.85	13.67	5.30	5.50	5.75	5.65	6.55	5.90
$T_P^{c(r)}$	5.75	4.90	4.77	14.50	14.20	14.20	4.85	4.90	4.80	5.90	6.60	5.90
T_F^j	12.70	11.95	11.45	61.70	51.80	45.42	6.60	7.05	6.70	9.05	9.40	10.55
$T_F^{j(r)}$	9.05	7.25	6.15	52.50	38.40	29.30	4.70	4.25	4.00	7.00	6.15	7.25
T_P^j	12.35	10.85	9.85	41.65	34.00	28.90	5.30	6.10	5.80	7.50	7.35	7.45
$T_P^{j(r)}$	9.15	7.20	6.10	33.85	27.00	22.45	4.70	4.25	4.20	5.55	6.20	5.55

Table 5: The effect of misspecified univariate GARCHmodel

	P1	P2	P3	P4	P5
$\widetilde{\alpha}$	0.05	0.10	0.15	-	-
$\widetilde{\beta}$	0.90	0.85	0.80	-	-
$\overline{Q}$	$\begin{bmatrix} 1.00 & 0.60 \\ 0.60 & 1.00 \end{bmatrix}$			-	-
A_B	-	-	-	$\begin{bmatrix} 0.30 & 0.10 \\ 0.10 & 0.30 \end{bmatrix}$	$\begin{bmatrix} 0.40 & 0.20 \\ 0.20 & 0.40 \end{bmatrix}$
B_B	-	-	-	$\begin{bmatrix} 0.60 & 0.20 \\ 0.20 & 0.60 \end{bmatrix}$	$\begin{bmatrix} 0.40 & 0.20 \\ 0.20 & 0.40 \end{bmatrix}$
C_B	-	-	-	$\begin{bmatrix} 0.20 & 0.10 \\ 0.10 & 0.20 \end{bmatrix}$	$\begin{bmatrix} 0.20 & 0.04 \\ 0.04 & 0.20 \end{bmatrix}$

Table 6: True parameter values for power simulation

5.4 Power Simulation

To examine the power of these tests we consider two types MGARCH models with time varying correlations again in the regression context. The AR(1) specification for the conditional mean function introduced for size simulation is retained. First we assume that the true DGP for conditional variance matrix H_t follows Engle's (2002) DCC-GARCH(1,1) model as follows:

$$\begin{aligned} y_{it} &= \varphi_{i0} + \varphi_{i1} y_{i,t-1} + \varepsilon_{it}, \quad i = 1, 2 \\ Var\left(\varepsilon_t | \mathcal{F}_{t-1}\right) &= H_t \Rightarrow \mathrm{E}\left[\varepsilon_{it}^2 | \mathcal{F}_{t-1}\right] = h_{ii,t}, \\ h_{ii,t} &= \alpha_{i0} + \alpha_{i1}\varepsilon_{i,t-1}^2 + \beta_{i1} h_{i,t-1} \\ H_t &= D_t \Gamma_t D_t; \ D_t = \begin{bmatrix} \sqrt{h_{11t}} & 0 \\ 0 & \sqrt{h_{22t}} \end{bmatrix} \\ \Gamma_t &= \left(I \odot Q_t\right)^{-1/2} Q_t \left(I \odot Q_t\right)^{-1/2} = diag(Q_t)^{-1/2} Q_t diag(Q_t)^{-1/2} \\ Q_t &= (1 - \widetilde{\alpha} - \widetilde{\beta})\overline{Q} + \widetilde{\alpha}\zeta_{t-1}\zeta_{t-1}' + \widetilde{\beta} Q_{t-1}; \end{aligned} \tag{44}$$

Secondly, we consider BEKK model of Engle and Kroner (1995) as the true DGP for conditional variance matrix H_t

$$\begin{aligned} y_{it} &= \varphi_{i0} + \varphi_{i1} y_{i,t-1} + \varepsilon_{it}, \quad i = 1, 2 \\ Var\left(\varepsilon_t | \mathcal{F}_{t-1}\right) &= H_t \\ H_t &= C_B + A_B'\left(\varepsilon_{t-1}\varepsilon_{t-1}'\right) A_B + B_B' H_{t-1} B_B \end{aligned} \tag{45}$$

Five experiments are considered; P1, P2 and P3 with DCC DGP and remaining two (P4 and P5) with BEKK DGP. The true parameter values for conditional mean functions of size simulation experiment are maintained for all DGPs. Also for DCC DGPs, high persistence individual volatility specification for both variables, as given in E1 and E2, is retained (i.e. $\eta_1' = (0.01, 0.15, 0.80)$ and $\eta_2' = (0.05, 0.20, 0.70)$). The remaining true parameter vectors are given in Table (6). The parameter values for BEKK models are taken from Tse (2000).

Table (7) and Table (8) presents the power results with 2000 replications for DCC and BEKK DGPs respectively where the nominal size is again 5%. The data is generated assuming normality. The average of the estimated correlation parameter and true range of correlations in the simulated sample (as a measure of time variability) are presented in the last panels.

It can be seen when the true DGP is DCC, P3 has the largest variability in correlations followed by P2 and P1 i.e. variability increases as $\widetilde{\alpha}$ increases and $\widetilde{\beta}$ decreases. In general, the tests based on Tse's indicator is found to have higher power in all three DCC experiments. However, as the variability in correlation decreases power decreases. The FCM tests also show nice power property. It is to be noted that even with $T = 500$, both Tse and FCM tests show high power especially in P2 and P3. But CCM tests lack power considerably, particularly for P1. The OPG-FQMLE tests show greater power; however using robust and PQMLE versions do not cost much power. In case of BEKK DGP the conclusion is quite similar to DCC models. P5 has larger variability in correlation than P4 and the tests also oblige the fact. All tests show excellent power for P5; CCM tests, however, lacks power for P3.

	DGP: P1			DGP: P2			DGP: P3		
	$T = 1000$	$T = 700$	$T = 500$	$T = 1000$	$T = 700$	$T = 500$	$T = 1000$	$T = 700$	$T = 500$
T_F^t	51.15	39.45	31.90	96.35	88.40	79.85	99.90	99.50	96.50
$T_F^{t(r)}$	48.45	36.00	27.70	95.45	86.30	75.25	99.85	99.10	95.50
T_P^t	47.60	35.65	27.40	94.85	84.25	71.90	99.70	98.70	93.30
$T_P^{t(r)}$	47.10	35.10	26.50	95.05	85.15	73.50	99.85	99.10	94.95
T_F^c	16.65	12.20	9.85	60.35	45.75	33.75	92.05	80.00	69.10
$T_F^{c(r)}$	14.70	9.65	6.95	57.15	42.10	29.15	91.15	76.65	63.25
T_P^c	15.60	10.85	7.70	55.70	41.00	30.05	90.55	76.10	62.65
$T_P^{c(r)}$	13.80	9.55	6.55	54.55	38.85	28.00	90.10	75.05	61.65
T_F^j	40.35	30.90	27.25	92.20	81.40	70.15	99.60	97.95	94.20
$T_F^{j(r)}$	34.40	24.20	18.85	89.30	74.70	60.10	99.55	96.90	89.60
T_P^j	36.05	26.65	20.15	88.60	73.95	57.50	99.35	95.65	87.25
$T_P^{j(r)}$	33.60	24.25	18.20	88.80	73.10	57.15	99.60	96.80	89.00
Estimated Correlation (Average)									
FQMLE	0.593	0.595	0.596	0.575	0.580	0.666	0.552	0.558	0.557
PQMLE	0.590	0.596	0.591	0.570	0.573	0.571	0.545	0.550	0.547
Range of correlation in simulated sample									
Average	0.568	0.534	0.503	1.056	1.001	0.947	1.400	1.344	1.286
Max	0.961	0.930	0.876	1.457	1.506	1.455	1.711	1.727	1.728
Min	0.369	0.305	0.279	0.679	0.548	0.462	0.937	0.727	0.579

Table 7: Empirical Power with DCC DGP

	DGP: P4			DGP: P5		
	$T = 1000$	$T = 700$	$T = 500$	$T = 1000$	$T = 700$	$T = 500$
T_F^t	85.50	73.95	64.80	99.85	98.45	95.30
$T_F^{t(r)}$	83.10	69.30	57.00	99.80	97.55	93.35
T_P^t	91.10	76.95	62.80	99.95	99.10	96.05
$T_P^{t(r)}$	89.10	73.75	59.40	99.80	98.75	94.10
T_F^c	72.45	56.75	42.55	100.00	99.50	97.80
$T_F^{c(r)}$	68.35	49.90	34.10	99.95	99.15	95.35
T_P^c	25.35	16.20	10.10	100.00	98.45	94.75
$T_P^{c(r)}$	22.85	15.25	9.65	99.95	97.80	92.25
T_F^j	81.25	66.95	55.10	99.95	99.20	95.75
$T_F^{j(r)}$	74.75	56.85	42.25	99.80	97.60	88.80
T_P^j	81.25	62.40	46.15	100.00	98.95	94.70
$T_P^{j(r)}$	74.80	53.55	38.35	99.85	97.55	88.55
Estimated Correlation (Average)						
FQMLE	0.839	0.839	0.840	0.616	0.617	0.617
PQMLE	0.831	0.831	0.831	0.595	0.596	0.596
Range of correlation in simulated sample						
Average	0.247	0.239	0.231	0.668	0.653	0.637
Max	0.349	0.347	0.321	0.829	0.824	0.802
Min	0.191	0.189	0.178	0.567	0.557	0.537

Table 8: Empirical Power with BEKK DGP

6 Concluding Remarks

In this paper, we propose a set of asymptotically valid CM tests of testing the CCC hypothesis for MGARCH model. We consider tests considering both FQMLE and particularly PQMLE framework for CCC models and tests with the latter is nonexistent in the literature. Moreover, the robust and OPG versions of these tests are developed. These tests are very easy to implement. We also analyze and accommodate Tse's (2000) LM test, which is a OPG type test, and consider a robust version of it. We examine the finite sample performance of these asymptotically valid tests.

Monte Carlo experiments indicate that in general all tests have desirable size property and robust version perform better than OPG version. It is found that the correlation parameter has a significant impact on empirical size of these tests (low correlation is associated with better size property); the size is, however, not affected by the degree of univariate volatility persistence. The robust versions demonstrate better size than OPG tests in case of the departure from normality assumption of true error; particularly all OPG-FQMLE tests are oversized. Interestingly, PQMLE based tests exhibit more robustness compared to FQMLE tests. Besides, when the assumption of the null model is violated by assuming misspecified univariate volatility structure but maintaining CCC assumption, the size of these tests are not affected by volatility spillover effect; however when one equation is misspecified and true correlation is high all tests overreject the null of the CCC assumption. The rejection rate is higher in case of FCM tests expectedly; as by construction these tests consider the individual volatility moment conditions as well.

The power of these tests depends on the variability of the true correlation parameter and it is found that tests based on Tse's modified indicator and FCM show excellent power, even in models with less dispersed correlations. The CCM tests, in general, show lower power and particularly in models with less dispersed correlations have limited power. In terms of power there is very little to choose between OPG and robust; and between FQMLE and PQMLE.

To sum up, testing correlation constancy depends on the true correlation parameter and no significant difference is observed whether one use FQMLE and PQMLE approach. The robust versions manifest better size under non-normality. The FCM tests check the individual volatility along with CCC assumption; hence can be treated as a general diagnostic test. The CCM test has desirable size properties, but lacks power under certain DGPs. Tse's LM test, which is a OPG-FQMLE type, has good size and power properties but is sensitive to the departure from normality while its OPG-PQMLE display impressive robustness maintaining the high power performance. The robust version of Tse-FQMLE, however, has empirical size close to the nominal level under non-normality.

The tests here is derived for to check CCC assumption which in many situations is not a realistic or reasonable one. It is therefore is of interest to devise test of time varying correlation. In practice, two-stage estimation approach is almost always applied to estimate time varying correlation model indicating to develop a testing framework based on PQMLE approach. However,in this case correlation is not a scalar and the simple demeaning technique, that we have used to derive the PQMLE tests in this paper, is not possible and we need to consider the estimation effect emerging from correlation parameter to derive the limit distribution of the test statistics. Such extensions, however, left for future research.

Appendices

A Proofs

For the CCC we have,

$$corr\left[\varepsilon_{it},\varepsilon_{jt}|\mathcal{F}_{t-1}\right]=\mathrm{E}\left[\left.\frac{\varepsilon_{it}\varepsilon_{jt}}{\sqrt{h_{it}}\sqrt{h_{jt}}}\right|\mathcal{F}_{t-1}\right]=\rho_{ij}.$$

We have the following definitions:

1. $\zeta_{it}=\varepsilon_{it}/\sqrt{h_{it}}$ is $iid\ (0,1)$, for $t=1,...,T$, with $\mathrm{E}\left[\zeta_{it}\zeta_{jt}|\mathcal{F}_{t-1}\right]=\rho_{ij}$; or $\mathrm{E}\left[\zeta_t\zeta_t'|\mathcal{F}_{t-1}\right]=\Gamma=\left\{\rho_{ij}\right\}$, $(N\times N)$, where $\zeta_t=\{\zeta_{it}\}$, $(N\times 1)$, and $\rho_{ii}\equiv 1$, $\Gamma=\Gamma'$ is symmetric and positive definite.
2. Let $\Gamma^{-1}=\left\{\rho^{ij}\right\}$, so that $\sum_{m=1}^{N}\rho^{im}\rho_{mj}=\delta_{ij}$, the Kronecker Delta, i.e. $\delta_{ii}=1$, $\delta_{ij}=0$, $i\neq j$.
3. $\varepsilon_{it}^*=\sum_{m=1}^{N}\rho^{im}\zeta_{mt}$, so that $\varepsilon_t^*=\{\varepsilon_{it}^*\}=\Gamma^{-1}\zeta_t$.
4. $f_{it}=w_{it}/\sqrt{h_{it}}$, $c_{it}=\dfrac{1}{h_{it}}\dfrac{\partial h_{it}}{\partial\varphi_i}$, $x_{it}=\dfrac{1}{h_{it}}\dfrac{\partial h_{it}}{\partial\eta_i}$.

Then, in addition to the properties of ζ_{it} listed in (1), we have the following:

$$\begin{aligned}
\mathrm{E}\left[\varepsilon_{it}^*|\mathcal{F}_{t-1}\right] &= 0\\
\mathrm{E}\left[\varepsilon_t^*\varepsilon_t^{*\prime}|\mathcal{F}_{t-1}\right] &= \Gamma^{-1}=\left\{\rho^{ij}\right\}=\left\{\mathrm{E}\left[\varepsilon_{it}^*\varepsilon_{jt}^*|\mathcal{F}_{t-1}\right]\right\}\\
\mathrm{E}\left[\varepsilon_t^*\zeta_t'|\mathcal{F}_{t-1}\right] &= I_N=\left\{\mathrm{E}\left[\varepsilon_{it}^*\zeta_{jt}|\mathcal{F}_{t-1}\right]\right\}
\end{aligned}$$

so that, in particular, $\mathrm{E}\left[\varepsilon_{it}^*\zeta_{jt}|\mathcal{F}_{t-1}\right]=\delta_{ij}$.

A.1 Proof of Lemma 1

To construct the expected Hessian matrix conditional on $\mathcal{F}_{t-1}$, we first obtain the second partial derivatives given F_{t-1}. Since $\varepsilon_t'H_t^{-1}\varepsilon_t=\zeta_t'\Gamma^{-1}\zeta_t=\zeta_t'\varepsilon_t^*$ we can write the likelihood function (7) as:

$$\begin{aligned}
l_t^* &= -\frac{1}{2}\ln|\Gamma|-\frac{1}{2}\sum_{j=1}^{N}\ln h_{jt}-\frac{1}{2}\zeta_t'\Gamma^{-1}\zeta_t\\
&= -\frac{1}{2}\ln|\Gamma|-\frac{1}{2}\sum_{j=1}^{N}\ln h_{jt}-\frac{1}{2}\sum_{j=1}^{N}\zeta_{jt}\varepsilon_{jt}^*.
\end{aligned}$$

Using the results of Lemma 1, and noting that $\varepsilon_{it}^*=\sum_{j=1}^{N}\rho^{ij}\zeta_{jt}$ and $\rho^{ij}=\rho^{ji}$; we have $\dfrac{\partial\varepsilon_{jt}^*}{\partial\varphi_i}=-\rho^{ij}\left(f_{it}+\dfrac{1}{2}\zeta_{it}c_{it}\right)$, and $\dfrac{\partial\varepsilon_{jt}^*}{\partial\eta_i}=-\dfrac{1}{2}\rho^{ij}\zeta_{it}x_{it}$. Then we have:

$$
\begin{aligned}
\frac{\partial l_t^*}{\partial \varphi_i} &= -\frac{1}{2}c_{it} + \frac{1}{2}\left(f_{it} + \frac{1}{2}\zeta_{it}c_{it}\right)\varepsilon_{it}^* - \frac{1}{2}\sum_{j=1}^{N}\zeta_{jt}\frac{\partial \varepsilon_{jt}^*}{\partial \varphi_i} \\
&= f_{it}\varepsilon_{it}^* + \frac{1}{2}c_{it}\left(\zeta_{it}\varepsilon_{it}^* - 1\right). \qquad (46)
\end{aligned}
$$

Similarly,

$$
\frac{\partial l_t^*}{\partial \eta_i} = \frac{1}{2}\left(\zeta_{it}\varepsilon_{it}^* - 1\right)x_{it}. \qquad (47)
$$

Finally, as in Tse (2000, p.113) we have:

$$
\frac{\partial l_t^*}{\partial \rho_{ij}} = \varepsilon_{it}^*\varepsilon_{jt}^* - \rho^{ij}, \quad i > j. \qquad (48)
$$

To see how (48) is derived, note from Magnus & Neudecker (p. 178) we have:

$$
\begin{aligned}
\mathrm{d}\ln|\Gamma| &= |\Gamma|^{-1}\,\mathrm{d}\,|\Gamma| = tr\left(\Gamma^{-1}\,\mathrm{d}\,\Gamma\right) \\
&= \sum_{k=1}^{N}\sum_{m=1}^{N}\rho^{ij}\left(\mathrm{d}\,\rho_{ji}\right) = \sum_{k=1}^{N}\sum_{m=1}^{N}\rho^{ij}\left(\mathrm{d}\,\rho_{ij}\right) \\
&= \sum_{k=1}^{N}\sum_{m\neq 1}^{N}\rho^{ij}\left(\mathrm{d}\,\rho_{ij}\right) = 2\sum_{k=2}^{N}\sum_{m=1}^{k-1}\rho^{ij}\left(\mathrm{d}\,\rho_{ij}\right)
\end{aligned}
$$

where $\Gamma = \left\{\rho_{ij}\right\}$, $\Gamma^{-1} = \left\{\rho^{ij}\right\}$, and where (in line 2) we have used the fact that $\rho_{kk} \equiv 1$ and that $\rho_{km} = \rho_{mk}$. Thus

$$
\frac{\partial \ln|\Gamma|}{\partial \rho_{ij}} = 2\rho^{ij}, \quad i > j.
$$

Furthermore, since $\varepsilon_{kt}^* = \sum_{m=1}^{N}\rho^{km}\zeta_{mt}$ we can write

$$
\frac{\partial \varepsilon_{kt}^*}{\partial \rho_{ij}} = \sum_{m=1}^{N}\frac{\partial \rho^{km}}{\partial \rho_{ij}}\zeta_{mt}.
$$

Note from Magnus & Neudecker (1999, pp. 183) we have (the differential) $\mathrm{d}\,\Gamma^{-1} = -\Gamma^{-1}\left(\mathrm{d}\,\Gamma\right)\Gamma^{-1}$. Then looking at the elements we see (with $\Gamma = \left\{\rho_{ij}\right\}$,

$\Gamma^{-1} = \left\{\rho^{ij}\right\})$

$$
\begin{aligned}
\mathrm{d}\,\rho^{km} &= -\sum_{r=1}^{N}\sum_{s=1}^{N}\rho^{kr}\left(\mathrm{d}\,\rho_{rs}\right)\rho^{sm} \\
&= -\sum_{r=1}^{N}\sum_{s\neq r}^{N}\rho^{kr}\left(\mathrm{d}\,\rho_{rs}\right)\rho^{sm} \\
&= -\sum_{r=2}^{N}\sum_{s=1}^{r-1}\rho^{kr}\left(\mathrm{d}\,\rho_{rs}\right)\rho^{sm} - \sum_{s=2}^{N}\sum_{r=1}^{s-1}\rho^{kr}\left(\mathrm{d}\,\rho_{rs}\right)\rho^{sm} \\
&= -\sum_{r=2}^{N}\sum_{s=1}^{r-1}\rho^{kr}\left(\mathrm{d}\,\rho_{rs}\right)\rho^{sm} - \sum_{r=2}^{N}\sum_{s=1}^{s-1}\rho^{ks}\left(\mathrm{d}\,\rho_{rs}\right)\rho^{rm}
\end{aligned}
$$

where (in line 2) we have used the fact that $\rho_{rr} \equiv 1$ and (in line 4) that $\rho_{rs} = \rho_{sr}$. Thus,

$$
\begin{aligned}
\frac{\partial \rho^{km}}{\partial \rho_{ij}} &= -\rho^{ki}\rho^{jm} - \rho^{kj}\rho^{im} \\
&= -\rho^{ki}\rho^{mj} - \rho^{kj}\rho^{mi}.
\end{aligned}
$$

Then,

$$
\begin{aligned}
\frac{\partial \varepsilon_{kt}^{*}}{\partial \rho_{ij}} &= \sum_{m=1}^{N}\frac{\partial \rho^{km}}{\partial \rho_{ij}}\zeta_{mt} \\
&= -\sum_{m=1}^{N}\left(\rho^{ki}\rho^{jm} + \rho^{kj}\rho^{im}\right)\zeta_{mt} \\
&= -\rho^{ki}\varepsilon_{jt}^{*} - \rho^{kj}\varepsilon_{it}^{*}.
\end{aligned}
$$

Now we have,

$$
\begin{aligned}
\frac{\partial l_t^{*}}{\partial \rho_{ij}} &= -\rho^{ij} - \frac{1}{2}\sum_{k=1}^{N}\zeta_{kt}\frac{\partial \varepsilon_{kt}^{*}}{\partial \rho_{ij}} \\
&= \frac{1}{2}\sum_{k=1}^{N}\zeta_{kt}\left(\rho^{ki}\varepsilon_{jt}^{*} + \rho^{kj}\varepsilon_{it}^{*}\right) - \rho^{ij} \\
&= \frac{1}{2}\left(\varepsilon_{jt}^{*}\sum_{k=1}^{N}\rho^{ik}\zeta_{kt} + \varepsilon_{it}^{*}\sum_{k=1}^{N}\rho^{jk}\zeta_{kt}\right) - \rho^{ij} \\
&= \frac{1}{2}\left(\varepsilon_{jt}^{*}\varepsilon_{it}^{*} + \varepsilon_{it}^{*}\varepsilon_{jt}^{*}\right) - \rho^{ij} \\
&= \varepsilon_{it}^{*}\varepsilon_{jt}^{*} - \rho^{ij}.
\end{aligned}
$$

A.2 Proof of Lemma 2

Recall that $\mathrm{E}\left[\varepsilon_{it}^{*}|\mathcal{F}_{t-1}\right]=0$, $\mathrm{E}\left[\zeta_{it}\varepsilon_{it}^{*}|\mathcal{F}_{t-1}\right]=1$ and $\mathrm{E}\left[\varepsilon_{it}^{*}\varepsilon_{jt}^{*}|\mathcal{F}_{t-1}\right]=\rho^{ij}$, so that each of the above scores has zero mean. Note that $\partial f_{it}/\partial\varphi_i$, $\partial f_{it}/\partial\eta_i$, $\partial c_{it}/\partial\varphi_i$, $\partial c_{it}/\partial\eta_i$, $\partial x_{it}/\partial\varphi_i$, and $\partial x_{it}/\partial\eta_i$ are all $\mathcal{F}_{t-1}$ measurable. Also, $\Gamma=\left\{\rho_{ij}\right\}$ and $\Gamma^{-1}=\left\{\rho^{ij}\right\}$, so that $\sum_{m=1}^{N}\rho^{im}\rho_{mj}=\delta_{ij}$,where $\delta_{ii}=1$, $\delta_{ij}=0$, $i\neq j$. And $\varepsilon_t^{*}=\{\varepsilon_{it}^{*}\}=\Gamma^{-1}\zeta_t$; hence $\varepsilon_{it}^{*}=\sum_{m=1}^{N}\rho^{im}\zeta_{mt}$.

1. Differentiating (46) with respect to φ_j, we obtain

$$\begin{aligned}\frac{\partial^2 l_t^{*}}{\partial\varphi_i\partial\varphi_j'} &= \delta_{ij}\frac{\partial f_{it}}{\partial\varphi_j'}\varepsilon_{it}^{*}+f_{it}\frac{\partial\varepsilon_{it}^{*}}{\partial\varphi_j'}+\frac{1}{2}\delta_{ij}\frac{\partial c_{it}}{\partial\varphi_j'}\left(\zeta_{it}\varepsilon_{it}^{*}-1\right)\\ &\quad+\frac{1}{2}\delta_{ij}c_{it}\varepsilon_{it}^{*}\frac{\partial\zeta_{it}}{\partial\varphi_j'}+\frac{1}{2}c_{it}\zeta_{it}\frac{\partial\varepsilon_{it}^{*}}{\partial\varphi_j'}\\ &= f_{it}\frac{\partial\varepsilon_{it}^{*}}{\partial\varphi_j'}+\frac{1}{2}\delta_{ij}c_{it}\varepsilon_{it}^{*}\frac{\partial\zeta_{it}}{\partial\varphi_j'}+\frac{1}{2}c_{it}\zeta_{it}\frac{\partial\varepsilon_{it}^{*}}{\partial\varphi_j'}+\varkappa_{1t}\end{aligned}$$

where $\varkappa_{1t}=\delta_{ij}\frac{\partial f_{it}}{\partial\varphi_j}\varepsilon_{it}^{*}+\frac{1}{2}\delta_{ij}\frac{\partial c_{it}}{\partial\varphi_j'}\left(\zeta_{it}\varepsilon_{it}^{*}-1\right)$ so that $\mathrm{E}_0\left[\varkappa_t|\mathcal{F}_{t-1}\right]_{\varpi=\varpi_0}=0$.Hence,

$$\begin{aligned}\frac{\partial^2 l_t^{*}}{\partial\varphi_i\partial\varphi_j'} &= -\rho^{ij}f_{it}\left(f_{jt}'+\frac{1}{2}\zeta_{jt}c_{jt}'\right)-\frac{1}{2}\delta_{ij}c_{it}\varepsilon_{it}^{*}\left(f_{jt}'+\frac{1}{2}\zeta_{jt}c_{jt}'\right)\\ &\quad-\frac{1}{2}\rho^{ij}c_{it}\zeta_{it}\left(f_{jt}'+\frac{1}{2}\zeta_{jt}c_{jt}'\right)+\varkappa_{1t}\\ &= -\rho^{ij}f_{it}f_{jt}'-\frac{1}{4}\delta_{ij}c_{it}c_{jt}'-\frac{1}{4}\rho^{ij}\rho_{ij}c_{it}c_{jt}'+\varkappa_{1t}^{*}\text{ (say)}\end{aligned}$$

where $\mathrm{E}_0\left[\varkappa_{1t}^{*}|\mathcal{F}_{t-1}\right]_{\varpi=\varpi_0}=0$. Thus we can write

$$\widetilde{\mathcal{H}}_{\varphi_i\varphi_j}^{*}=-\rho^{ij}f_{it}f_{jt}'-\frac{1}{4}\left(\delta_{ij}+\rho^{ij}\rho_{ij}\right)c_{it}c_{jt}'. \tag{49}$$

such that $\mathrm{E}_0\left[\widetilde{\mathcal{H}}_{\varphi_i\varphi_j}^{*}\left(\varpi_0\right)\right]=\mathrm{E}_0\left[\left.\frac{\partial^2 l_t^{*}}{\partial\varphi_i\partial\varphi_j'}\right|\mathcal{F}_{t-1}\right]_{\varpi=\varpi_0}$.

2. Differentiating (47) with respect to η_j, we obtain

$$\begin{aligned}\frac{\partial^2 l_t^{*}}{\partial\varphi_i\partial\eta_j'} &= \delta_{ij}\frac{\partial f_{it}}{\partial\eta_j'}\varepsilon_{it}^{*}+f_{it}\frac{\partial\varepsilon_{it}^{*}}{\partial\eta_j'}+\frac{1}{2}\delta_{ij}\frac{\partial c_{it}}{\partial\eta_j'}\left(\zeta_{it}\varepsilon_{it}^{*}-1\right)\\ &\quad+\frac{1}{2}\delta_{ij}c_{it}\varepsilon_{it}^{*}\frac{\partial\zeta_{it}}{\partial\eta_j'}+\frac{1}{2}c_{it}\zeta_{it}\frac{\partial\varepsilon_{it}^{*}}{\partial\eta_j'}\\ &= f_{it}\frac{\partial\varepsilon_{it}^{*}}{\partial\eta_j'}+\frac{1}{2}\delta_{ij}c_{it}\varepsilon_{it}^{*}\frac{\partial\zeta_{it}}{\partial\eta_j'}+\frac{1}{2}c_{it}\zeta_{it}\frac{\partial\varepsilon_{it}^{*}}{\partial\eta_j'}+\varkappa_{2t}\end{aligned}$$

where $\varkappa_{2t} = \delta_{ij} \dfrac{\partial f_{it}}{\partial \eta_j} \varepsilon_{it}^* + \dfrac{1}{2} \delta_{ij} \dfrac{\partial c_{it}}{\partial \eta_j'} \left(\zeta_{it} \varepsilon_{it}^* - 1 \right)$ so that $\mathrm{E}_0 \left[\varkappa_{2t} | \mathcal{F}_{t-1} \right]_{\varpi = \varpi_0} = 0$. Therefore we have,

$$\begin{aligned} \frac{\partial^2 l_t^*}{\partial \varphi_i \partial \eta_j'} &= -\frac{1}{2} \rho^{ij} f_{it} \zeta_{jt} x_{jt}' - \frac{1}{4} \delta_{ij} c_{it} \varepsilon_{it}^* \zeta_{jt} x_{jt}' - \frac{1}{4} \rho^{ij} c_{it} \zeta_{it} \zeta_{jt} x_{jt}' + \varkappa_{2t} \\ &= -\frac{1}{4} \delta_{ij} c_{it} x_{jt}' - \frac{1}{4} \rho^{ij} \rho_{ij} c_{it} x_{jt}' + \varkappa_{2t}^* \text{ (say)} \end{aligned}$$

where $\mathrm{E}_0 \left[\varkappa_{2t}^* | \mathcal{F}_{t-1} \right]_{\varpi = \varpi_0} = 0$. Thus, similarly

$$\widetilde{\mathcal{H}}_{\varphi_i \eta_j}^* = -\frac{1}{4} \left(\delta_{ij} + \rho^{ij} \rho_{ij} \right) c_{it} x_{jt}'. \tag{50}$$

3. Differentiating (48) with respect to φ_k yields

$$\begin{aligned} \frac{\partial^2 l_t^*}{\partial \rho_{ij} \partial \varphi_k'} &= \varepsilon_{jt}^* \frac{\partial \varepsilon_{it}^*}{\partial \varphi_k'} + \varepsilon_{it}^* \frac{\partial \varepsilon_{jt}^*}{\partial \varphi_k'} \\ &= -\rho^{ik} \varepsilon_{jt}^* \left(f_{kt} + \frac{1}{2} \zeta_{kt} c_{kt} \right) - \rho^{jk} \varepsilon_{it}^* \left(f_{kt} + \frac{1}{2} \zeta_{kt} c_{kt} \right) \\ &= -\frac{1}{2} \delta_{jk} \rho^{ik} c_{kt} - \frac{1}{2} \delta_{ik} \rho^{jk} c_{kt} + \varkappa_{3t} \end{aligned} \tag{51}$$

where $\mathrm{E}_0 \left[\varkappa_{3t} | \mathcal{F}_{t-1} \right]_{\varpi = \varpi_0} = 0$ and we have

$$\widetilde{\mathcal{H}}_{\varphi_i \rho_{ij}}^* = -\frac{1}{2} \delta_{jk} \rho^{ik} c_{kt} - \frac{1}{2} \delta_{ik} \rho^{jk} c_{kt}$$

Note that $i > j$, here, so that

$$\widetilde{\mathcal{H}}_{\varphi_i \rho_{ij}}^* = \begin{cases} -\frac{1}{2} \rho^{ji} c_{it} = -\frac{1}{2} \rho^{ij} c_{it}, \quad k = i > j \\ -\frac{1}{2} \rho^{ij} c_{jt}, \quad k = j < i \\ 0, \quad k \neq i, k \neq j \end{cases}$$

4. Differentiating (47) with respect to η_j, we obtain

$$\begin{aligned} \frac{\partial^2 l_t^*}{\partial \eta_i \partial \eta_j'} &= \frac{1}{2} \delta_{ij} \frac{\partial x_{it}}{\partial \eta_j'} \left(\zeta_{it} \varepsilon_{it}^* - 1 \right) + \frac{1}{2} \delta_{ij} x_{it} \varepsilon_{it}^* \frac{\partial \zeta_{it}}{\partial \eta_j'} + \frac{1}{2} x_{it} \zeta_{it} \frac{\partial \varepsilon_{it}^*}{\partial \eta_j'} \\ &= \frac{1}{2} \delta_{ij} x_{it} \varepsilon_{it}^* \frac{\partial \zeta_{it}}{\partial \eta_j'} + \frac{1}{2} x_{it} \zeta_{it} \frac{\partial \varepsilon_{it}^*}{\partial \eta_j'} + \varkappa_{4t} \\ &= -\frac{1}{4} \delta_{ij} x_{it} \varepsilon_{it}^* \zeta_{jt} x_{jt}' - \frac{1}{4} \rho^{ij} x_{it} \zeta_{it} \zeta_{jt} x_{jt}' + \varkappa_{4t} \\ &= -\frac{1}{4} \delta_{ij} x_{it} x_{jt}' - \frac{1}{4} \rho^{ij} \rho_{ij} x_{it} x_{jt}' + \varkappa_{4t} \end{aligned}$$

where $\varkappa_{4t} = \dfrac{1}{2} \delta_{ij} \dfrac{\partial x_{it}}{\partial \eta_j'} \left(\zeta_{it} \varepsilon_{it}^* - 1 \right)$ so that $\mathrm{E}_0 \left[\varkappa_{4t} | \mathcal{F}_{t-1} \right]_{\varpi = \varpi_0} = 0$. Thus

$$\widetilde{\mathcal{H}}_{\eta_i \eta_j}^* = -\frac{1}{4} \left(\delta_{ij} + \rho^{ij} \rho_{ij} \right) x_{it} x_{jt}'. \tag{52}$$

5. Differentiating (48) with respect to η_k yields

$$\begin{aligned}
\frac{\partial^2 l_t^*}{\partial \rho_{ij} \partial \eta_k'} &= \varepsilon_{jt}^* \frac{\partial \varepsilon_{it}^*}{\partial \eta_k'} + \varepsilon_{it}^* \frac{\partial \varepsilon_{jt}^*}{\partial \eta_k'} \\
&= -\frac{1}{2}\rho^{ik}\varepsilon_{jt}^*\zeta_{kt}x_{kt} - \frac{1}{2}\rho^{jk}\varepsilon_{it}^*\zeta_{kt}x_{kt} \\
&= -\frac{1}{2}\delta_{jk}\rho^{ik}x_{kt} - \frac{1}{2}\delta_{ik}\rho^{jk}x_{kt} \qquad (53)
\end{aligned}$$

Note that $i > j$, here, so that

$$\widetilde{\mathcal{H}}_{\eta_i \rho_{ij}}^* = \begin{cases} -\frac{1}{2}\rho^{ji}x_{it} = -\frac{1}{2}\rho^{ij}x_{it}, & k = i > j; \\ -\frac{1}{2}\rho^{ij}x_{jt}, & k = j < i \\ 0, & k \neq i, k \neq j. \end{cases}$$

6. Differentiating (48), $\dfrac{\partial l_t^*}{\partial \rho_{ij}} = \varepsilon_{it}^*\varepsilon_{jt}^* - \rho^{ij}$, $i > j$,with respect to ρ_{km} yields

$$\begin{aligned}
\frac{\partial^2 l_t^*}{\partial \rho_{ij} \partial \rho_{km}} &= \frac{\partial \varepsilon_{it}^*}{\partial \rho_{km}}\varepsilon_{jt}^* + \varepsilon_{it}^* \frac{\partial \varepsilon_{jt}^*}{\partial \rho_{km}} - \frac{\partial \rho^{ij}}{\partial \rho_{km}} \\
&= -\rho^{ik}\varepsilon_{mt}^*\varepsilon_{jt}^* - \rho^{im}\varepsilon_{kt}^*\varepsilon_{jt}^* - \rho^{jk}\varepsilon_{mt}^*\varepsilon_{it}^* - \rho^{jm}\varepsilon_{kt}^*\varepsilon_{it}^* \\
&\quad + \rho^{ik}\rho^{jm} + \rho^{im}\rho^{jk}
\end{aligned}$$

where we have used the previous results: $\dfrac{\partial \rho^{ij}}{\partial \rho_{km}} = -\rho^{ik}\rho^{jm} - \rho^{im}\rho^{jk}$ and $\dfrac{\partial \varepsilon_{it}^*}{\partial \rho_{km}} = -\rho^{ik}\varepsilon_{mt}^* - \rho^{im}\varepsilon_{kt}^*$. We thus obtain, using symmetry,

$$\begin{aligned}
\widetilde{\mathcal{H}}_{\rho_{ij}\rho_{km}}^* &= -\rho^{ik}\rho^{mj} - \rho^{im}\rho^{kj} - \rho^{jk}\rho^{mi} \\
&\quad - \rho^{jm}\rho^{ki} + \rho^{ik}\rho^{jm} + \rho^{im}\rho^{jk} \\
&= -\rho^{jk}\rho^{mi} - \rho^{jm}\rho^{ki} \\
&= -\rho^{ik}\rho^{jm} - \rho^{im}\rho^{jk} = \frac{\partial \rho^{ij}}{\partial \rho_{km}} \qquad (54)
\end{aligned}$$

A.3 Proof of Lemma 3

First we define the following:

$\rho = \text{vecl}(\Gamma) = \left\{\rho_{ij}\right\}$, $j = 1, ..., N-1$, $i = j+1, ..., N$ (i.e., the i subscript changes more quickly than the j subscript);

For the i^{th} variable, define $\underset{(T\times K)}{F_i}$, $\underset{(T\times K)}{C_i}$ and $\underset{(T\times K')}{X_i}$ with rows $f_{it}' = \dfrac{w_{it}'}{\sqrt{h_{it}}}$;

$c'_{it} = \dfrac{1}{h_{it}} \dfrac{\partial h_{it}}{\partial \varphi'_i}$ and $x'_{it} = \dfrac{1}{h_{it}} \dfrac{\partial h_{it}}{\partial \eta'_i}$ respectively. Then define

$$\begin{aligned}
\underset{(NT \times NK)}{F} &= diag\left(F_i\right), \\
\underset{(T \times NK)}{\tilde{F}} &= \left[F_1, F_2, ..., F_N\right] \\
\underset{(N \times NK)}{F'_t} &= diag\left(f'_{it}\right) for t = 1, \cdots, T.
\end{aligned}$$

In a similar way, define C, X, $\tilde{C}$, $\tilde{X}$. C'_t and X'_t. Also, define $\underset{(N \times N)}{E_t} = diag\left(\zeta_{it}\right)$; and the $(N \times T)$ matrices $E = \{\zeta_{it}\}$ and $E^* = \{\varepsilon^*_{it}\} = \Gamma^{-1}E$ having columns ζ_t and ε^*_t respectively. It will be useful to define $\Gamma_A = I_N + \left(\Gamma^{-1} \odot \Gamma\right)$.

Let ρ^k be the k^{th} column of Γ^{-1}; define $\Gamma^k = \Gamma^{-1} diag\left(\tau_k\right)$, where $\tau_k = \{\delta_{ik}\}, (N \times 1), i = 1, ..., N$; i.e. Γ^k be the $(N \times N)$ matrix of zeros, except for column k which is ρ^k. Define the following two $(N \times N)$ symmetric matrices:

$$\begin{aligned}
P_k &= \Gamma^k + \left(\Gamma^k\right)' \\
\Gamma_{km} &= \rho^k \left(\rho^m\right)' + \rho^m \left(\rho^k\right)'
\end{aligned}$$

Note that $\tilde{F} = \left(\iota'_N \otimes I_T\right) F$, so that, $\tilde{F}'\tilde{F} = F'\left(\iota_N \otimes I_T\right)\left(\iota'_N \otimes I_T\right) F = F'\left(\mathcal{J}_N \otimes I_T\right) F$. Now

1.
$$\begin{aligned}
\mathrm{E}_0 \left[\sum_{t=1}^{T} \frac{\partial^2 l_t}{\partial \varphi_i \partial \varphi'_j} \middle| \mathcal{F}_{t-1}\right]_{\varpi = \varpi_0} &= -\rho^{ij} \sum_{t=1}^{T} f_{it} f'_{jt} - \frac{1}{4}\left(\delta_{ij} + \rho^{ij}\rho_{ij}\right) \sum_{t=1}^{T} c_{it} c'_{jt} \\
&= -\rho^{ij} F'_i F_j - \frac{1}{4}\left(\delta_{ij} + \rho^{ij}\rho_{ij}\right) C'_i C_j
\end{aligned}$$

Hence,

$$\begin{aligned}
\mathrm{E}_0 \left[\sum_{t=1}^{T} \frac{\partial^2 l_t}{\partial \varphi \partial \varphi'} \middle| \mathcal{F}_{t-1}\right]_{\varpi = \varpi_0} &= -\left(\Gamma^{-1} \otimes J_K\right) \odot \tilde{F}'\tilde{F} - \frac{1}{4}\left(\Gamma_A \otimes J\right) \odot \tilde{C}'\tilde{C} \\
&= -\left(\Gamma^{-1} \otimes J_K\right) \odot F'\left(\mathcal{J}_N \otimes I_T\right) F \\
&\quad -\frac{1}{4}\left(\Gamma_A \otimes \mathcal{J}_K\right) \odot C'\left(J_N \otimes I_T\right) C \\
&= -F'\left(\Gamma^{-1} \otimes I_T\right) F - \frac{1}{4} C'\left(\Gamma_A \otimes I_T\right) C
\end{aligned}$$

2. Similarly,
$$\begin{aligned}
\mathrm{E}_0 \left[\sum_{t=1}^{T} \frac{\partial^2 l_t}{\partial \varphi \partial \eta'} \middle| \mathcal{F}_{t-1}\right]_{\varpi = \varpi_0} &= -\frac{1}{4}\left(\Gamma_A \otimes J_K\right) \odot \tilde{C}'\tilde{X} \\
&= -\frac{1}{4}\left(\Gamma_A \otimes J_K\right) \odot C'\left(J_N \otimes I_T\right) X \\
&= -\frac{1}{4} C'\left(\Gamma_A \otimes I_T\right) X
\end{aligned}$$

3.

$$
\begin{aligned}
\mathrm{E}_0 \left[\sum_{t=1}^{T} \frac{\partial^2 l_t}{\partial \eta \partial \eta'} \middle| \mathcal{F}_{t-1} \right]_{\varpi=\varpi_0} &= -\frac{1}{4} \left(A \otimes J_K \right) \odot \tilde{X}' \tilde{X} \\
&= -\frac{1}{4} X' \left(A \otimes I_T \right) X
\end{aligned}
$$

4. For $i > j$,

$$
\begin{aligned}
\mathrm{E}_0 \left[\sum_{t=1}^{T} \frac{\partial^2 l_t}{\partial \rho_{ij} \partial \varphi_k'} \middle| \mathcal{F}_{t-1} \right]_{\varpi=\varpi_0} &= -\frac{1}{2} \left(\delta_{ik} \rho^{jk} + \delta_{jk} \rho^{ik} \right) \iota_T' C_k \\
&= -\frac{1}{2} \left(\delta_{ik} \rho^{kj} + \rho^{ik} \delta_{kj} \right) \iota_T' C_k
\end{aligned}
$$

Then, the matrix with typical element $\rho^{ik} \delta_{kj}$ is $\rho^k \tau_k'$, where $\tau_k = \{\delta_{ik}\}$, $i = 1, ..., N.$ $(N \times N)$, ρ^k is the k^{th} column of Γ^{-1}, and e_k is the k^{th} column of I_N, $k = 1, ..., N$. Similarly, $\left(\rho^k e_k' \right)' = \left\{ \delta_{ik} \rho^{kj} \right\}$. Alternatively, let Γ^k be the $(N \times N)$ matrix of zeros, except for column k which is ρ^k, the k^{th} column of Γ^{-1}; i.e., $\Gamma^k = \Gamma^{-1} diag\left(\tau_k\right)$. Define the symmetric matrix $P_k = \Gamma^k + \left(\Gamma^k \right)'$, $p_k = vecl(P_k)$, and $R_k = \iota_T p_k'$. Then, since $i > j$,

$$
\begin{aligned}
\mathrm{E}_0 \left[\sum_{t=1}^{T} \frac{\partial^2 l_t}{\partial \rho \partial \varphi_k'} \middle| \mathcal{F}_{t-1} \right]_{\varpi=\varpi_0} &= -\frac{1}{2} \operatorname{vecl}\left(P_k \right) \iota_T' C_k \\
&= -\frac{1}{2} p_k \iota_T' C_k
\end{aligned}
$$

Collecting the k blocks together we get

$$
\mathrm{E}_0 \left[\sum_{t=1}^{T} \frac{\partial^2 l_t}{\partial \rho \partial \varphi'} \middle| \mathcal{F}_{t-1} \right]_{\varpi=\varpi_0} = -\frac{1}{2} P' \left(I_N \otimes \iota_T' \right) C
$$

where $\underset{N \times \frac{N(N-1)}{2}}{P}$ has rows $p_k' = \operatorname{vecl}(P_k)'$, $k = 1, ..., N$.

5. In a similar way to the previous result,

$$
\mathrm{E}_0 \left[\sum_{t=1}^{T} \frac{\partial^2 l_t}{\partial \rho \partial' \eta} \middle| \mathcal{F}_{t-1} \right]_{\varpi=\varpi_0} = -\frac{1}{2} P' \left(I_N \otimes \iota_T' \right) X
$$

6. Finally, $\underset{\frac{N(N-1)}{2} \times \frac{N(N-1)}{2}}{\widetilde{P}}$ has columns $\widetilde{p}_{km} = \operatorname{vecl}\left(\Gamma_{km} \right)$, $m = 1, ..., N-1$, $k = m+1, ..., N$

$$
\begin{aligned}
\mathrm{E}_0 \left[\frac{\partial^2 l_t}{\partial \rho_{ij} \partial \rho_{km}} \middle| \mathcal{F}_{t-1} \right]_{\varpi=\varpi_0} &= -\rho^{ik} \rho^{jm} - \rho^{im} \rho^{jk} \\
&= -\rho^{ik} \rho^{mj} - \rho^{im} \rho^{kj}.
\end{aligned}
$$

Now, the $(N \times N)$ matrix with typical $(i,j)^{th}$ element equal to $\rho^{ik}\rho^{mj}$ is $\rho^k (\rho^m)'$ and that with typical $(i,j)^{th}$ element equal to $\rho^{im}\rho^{kj}$ is $\rho^m \left(\rho^k\right)'$. Let $\Gamma_{km} = \rho^k (\rho^m)' + \rho^m \left(\rho^k\right)'$, and let $p_{km} = \text{vecl}\left(\Gamma_{km}\right)$. Then

$$\mathrm{E}_0 \left[\left. \frac{\partial^2 l_t}{\partial \rho \partial \rho_{km}} \right| \mathcal{F}_{t-1} \right]_{\varpi=\varpi_0} = -p_{km}$$

or

$$\begin{aligned} \mathrm{E}_0 \left[\left. \frac{\partial^2 l_t}{\partial \rho \partial \rho'} \right| F_{t-1} \right]_{\varpi=\varpi_0} &= -\widetilde{P} \\ \mathrm{E}_0 \left[\sum_{t=1}^{T} \left. \frac{\partial^2 l_t}{\partial \rho \partial \rho'} \right| \mathcal{F}_{t-1} \right]_{\varpi=\varpi_0} &= -T\widetilde{P} \end{aligned}$$

where $\widetilde{P} = \left[\widetilde{p}_{21}, \widetilde{p}_{31}, ..., \widetilde{p}_{N,N-1}\right]$, a matrix with columns $\widetilde{p}_{km}$, $m = 1, ..., N-1$, $k = m+1, ..., N$ (k changes more quickly than m).

A.4 Proof of Theorem 2

Proof. The test indicator under consideration is $\hat{M}_{FT} \equiv T^{-1} \sum_{t=1}^{T} \hat{m}_{Ft}$. By the consistency of $\hat{\varpi}$ we have, $\sqrt{T}\left(\hat{\varpi} - \varpi_0\right) = J_{\varpi\varpi}^{*-1} \sqrt{T} G_{0T}^{*} + o_p(1)$ where $J_{\varpi\varpi}^{*} = -\mathrm{E}_0 \left[\mathcal{H}_t^{*}\left(\varpi_0\right)\right] = \text{plim}_{T\to\infty} T^{-1} \sum_{t=1}^{T} \dfrac{\partial g_{0t}^{*}}{\partial \varpi'}$ and assuming $J_{M\varpi}^{*} = \text{plim}_{T\to\infty} J_{M\varpi T}^{*}$ (by a ULLN) where $J_{M\varpi T}^{*} = -T^{-1} \sum_{t=1}^{T} \dfrac{\partial m_{0Ft}}{\partial \varpi}$. Taking a mean value expansion of $\hat{M}_{FT}$ about $\varpi_0 = \left(\theta_0', \rho_0'\right)'$,

$$\begin{aligned} \sqrt{T}\hat{M}_{FT} &= \sqrt{T} M_{0FT} - \overline{J}_{M\varpi T}^{*} \sqrt{T}\left(\hat{\varpi} - \varpi_0\right) \\ &= \sqrt{T} M_{0FT} - J_{M\varpi}^{*} J_{\varpi\varpi}^{*-1} \sqrt{T} G_{0T}^{*} + o_p(1) \\ &= A^{*} \frac{1}{\sqrt{T}} \sum_{t=1}^{T} \begin{pmatrix} m_{0Ft} \\ g_{0t}^{*} \end{pmatrix} + o_p(1) \end{aligned}$$

where $\overline{J}_{M\varpi T}^{*} = -T^{-1} \sum_{t=1}^{T} \left. \dfrac{\partial m_{Ft}}{\partial \varpi} \right|_{\overline{\varpi}}$ and $\overline{\varpi}$ is the usual "mean value" satisfying $\overline{\varpi} = \varpi_0 + o_p(1) \Rightarrow \overline{J}_{M\varpi T}^{*} = J_{M\varpi}^{*} + o_p(1)$ and $A^{*} = \left[I_r : -J_{M\varpi}^{*} J_{\varpi\varpi}^{*-1}\right]$. Now using Proposition 1, we conclude that

$$\sqrt{T}\hat{M}_{FT} \xrightarrow{d} N(0, V)$$

where $V = A^{*}\Sigma^{*}A^{*\prime}$ with $\Sigma^{*} = \begin{bmatrix} \Sigma_{MM} & \Sigma_{MG}^{*} \\ \Sigma_{GM}^{*} & \Sigma_{GG}^{*} \end{bmatrix}$. ■

A.5 Proof of Proposition 2

Proof. The test indicator under consideration is $\hat{M}_{FT} = \dfrac{1}{T} \sum_{t=1}^{T} \left(\hat{\upsilon}_t \otimes \hat{r}_t\right) = \dfrac{1}{T} \sum_{t=1}^{T} \hat{m}_{Ft}$ where $\hat{\upsilon}_t = \text{vech}\left(\hat{\zeta}_t \hat{\zeta}_t' - \hat{\Gamma}\right)$. Define $J_{M\varpi}^{*} = \underset{T\longrightarrow\infty}{\text{plim}} J_{M\varpi T}^{*}$ where

$J^*_{M\varpi T} = -\frac{1}{T}\sum_{t=1}^{T}\frac{\partial m_{Ft}}{\partial \varpi'} = -\left[\frac{1}{T}\sum_{t=1}^{T}\frac{\partial m_{Ft}}{\partial \theta'}, \frac{1}{T}\sum_{t=1}^{T}\frac{\partial m_{Ft}}{\partial \rho'}\right], Z = (Z_1, \cdots, Z_N)$
where Z_i is $(T \times k_i)$ matrix having rows $z'_{it} = (c'_{it}, x'_{it})$ for $i = 1, \cdots N$. Also define R^* having rows r'_t, if r_t is a vector of test variables, or R^* is a vector with typical element r_t if r_t is a scalar.

Note that $\upsilon_t = \left\{\zeta_{it}^2 - 1\right\}$ if $i = j$ and $\upsilon_t = \left\{\zeta_{it}\zeta_{jt} - \rho_{ij}\right\}$ if $i \neq j$, $j < i = 2, \cdots, N$. Since ζ_{jt} are functionally independent of both φ_i and η_i, $i \neq j$, and ρ does not enter in θ; Now, we have $\frac{\partial \zeta_{it}}{\partial \varphi_i} = -f_{it} - \frac{1}{2}\zeta_{it}c_{it}$, $\frac{\partial \zeta_{it}}{\partial \eta_i} = -\frac{1}{2}\zeta_{it}x_{it}$; hence $\frac{\partial\left(\zeta_{it}^2 - 1\right)}{\partial \varphi_i} = -2\zeta_{it}\left(f_{it} + \frac{1}{2}\zeta_{it}c_{it}\right)$ and $\frac{\partial\left(\zeta_{it}^2 - 1\right)}{\partial \eta_i} = -2\zeta_{it}^2 x_{it}$ and for $i \neq j$, $j < i = 2, \cdots, N$; $\frac{\partial(\zeta_{it}\zeta_{jt} - \rho_{ij})}{\partial \varphi_i} = -\zeta_{jt}\left(f_{it} + \frac{1}{2}\zeta_{it}c_{it}\right)$, $\frac{\partial(\zeta_{it}\zeta_{jt} - \rho_{ij})}{\partial \eta_i} = -\zeta_{jt}\left(\frac{1}{2}\zeta_{it}x_{it}\right)$, $\frac{\partial(\zeta_{it}\zeta_{jt} - \rho_{ij})}{\partial \rho_{ij}} = -1$. Hence,

$$\begin{aligned}
\frac{\partial\left(\zeta_{it}^2 - 1\right)}{\partial \theta'_i} &= \left(-2\zeta_{it}\left(f'_{it} + \frac{1}{2}\zeta_{it}c'_{it}\right), -2\zeta_{it}^2 x'_{it}\right), \\
\frac{\partial(\zeta_{it}\zeta_{jt} - \rho_{ij})}{\partial \theta'_i} &= -\zeta_{jt}\left(f'_{it} + \frac{1}{2}\zeta_{it}c'_{it}, \frac{1}{2}\zeta_{it}x'_{it}\right).
\end{aligned}$$

Now, note that $\mathrm{E}\left[\zeta_{jt}^2|\mathcal{F}_{t-1}\right] = 1$, $\mathrm{E}\left[f_{it}\zeta_{jt}|\mathcal{F}_{t-1}\right] = \mathrm{E}\left[f_{it}\zeta_{it}|\mathcal{F}_{t-1}\right] = 0$, since f_{it} is $\mathcal{F}_{t-1}$ measurable and $\mathrm{E}_0\left[\zeta_{1t}\zeta_{2t}|\mathcal{F}_{t-1}\right] = \rho_0$. Therefore

$$\mathrm{E}_0\left[\frac{\partial\left(\zeta_{it}^2 - 1\right)}{\partial \theta'} r_t|\mathcal{F}_{t-1}\right] = \underset{T\longrightarrow\infty}{\mathrm{plim}}\frac{1}{T}R^{*\prime}(0, \ldots, Z_i, \cdots, 0) \tag{55}$$

and, for $i \neq j$, $j < i = 2, \cdots, N$

$$\mathrm{E}_0\left[\sum_{t=1}^{T}\frac{\partial(\zeta_{it}\zeta_{jt} - \rho_{ij})}{\partial \theta'} r_t|\mathcal{F}_{t-1}\right] = \frac{1}{2}\rho_0 \underset{T\longrightarrow\infty}{\mathrm{plim}}\frac{1}{T}\left\{R^{*\prime}(0, \ldots, Z_i, \cdots, Z_j, \cdots 0)\right\} \tag{56}$$

Finally,

$$\mathrm{E}_0\left[\frac{\partial\left(\zeta_{it}^2 - 1\right)}{\partial \rho'} r_t|\mathcal{F}_{t-1}\right] = 0 \tag{57}$$

and

$$\mathrm{E}_0\left[\frac{\partial(\zeta_{it}\zeta_{jt} - \rho_{ij})}{\partial \rho'} r_t|\mathcal{F}_{t-1}\right] = \underset{T\longrightarrow\infty}{\mathrm{plim}}\frac{1}{T}(0, \ldots, 1', \cdots, 0)R^* \tag{58}$$

■

A.6 Proof of Lemma 5

Proof. Under normality, from the generalized IM inequality (e.g. Newey, 1985), we have $J^*_{M\varpi} = \Sigma^*_{MG}$ and $J^*_{\varpi\varpi} = \Sigma^*_{GG}$ and the result follows. Further (by

ULLN) the consistent estimator of Σ^*_{MG}, Σ^*_{GG}.and Σ^*_{MM} are given by

$$\begin{aligned}
\hat{\Sigma}^*_{MG} &= T^{-1}\sum_{t=1}^{T}\hat{m}_{Ft}\hat{g}^{*\prime}_t = T^{-1}\hat{S}^{*\prime}\hat{R}, \\
\hat{\Sigma}^*_{GG} &= T^{-1}\sum_{t=1}^{T}\hat{g}^*_t\hat{g}^{*\prime}_t = T^{-1}\hat{S}^{*\prime}\hat{S}^*, \\
\hat{\Sigma}_{MM} &= T^{-1}\sum_{t=1}^{T}\hat{m}_{Ft}\hat{m}'_{Ft} = T^{-1}\hat{R}'\hat{R}.
\end{aligned}$$

Hence, $\hat{A}^* = \left[I_r : -\hat{\Sigma}^*_{MG}\hat{\Sigma}^{*-1}_{GG}\right]$. Now Define $\hat{B}^* = \left[\hat{R},\ \hat{S}^*\right]$, and $\hat{W}^* = \hat{B}^*\hat{A}^{*\prime}$ where R and $\hat{S}^*$are $(T\times r)$ and $(T\times N')$ matrices having rows $\hat{m}'_{Ft}$ and $\dfrac{\partial l^*_t}{\partial \varpi'}$; evaluated at $\hat{\varpi}$.Then V can be consistently estimated by $\hat{V}_T = \dfrac{1}{T}\hat{W}^{*\prime}\hat{W}^* = \hat{\Sigma}_{MM} - \hat{\Sigma}^*_{MG}\hat{\Sigma}^{*-1}_{GG}\hat{\Sigma}^*_{GM}$. ■

A.7 Proof of Theorem 3

Proof. Define $\hat{M}_{PT} \equiv \hat{M}$, $J_i(\theta_0) = -\operatorname{plim}_{T\to\infty}\frac{1}{T}\left[\sum_{t=1}^{T}\frac{\partial^2 l_{it}(\theta_i)}{\partial\theta_i\partial\theta_i'}\right]_{\theta=\theta_0}$; $\nabla M_i \equiv \nabla M_{PTi}(\theta_1,\theta_2) = \operatorname{plim}_{T\to\infty}\dfrac{\partial M_{0PT}}{\partial\theta_i}$, $i=1,2$. Now taking a mean value expansion of $\sqrt{T}\hat{M}$ about $\hat{\theta}=\theta_0$

$$\begin{aligned}
\sqrt{T}\hat{M} &= \sqrt{T}M(\theta_0) + \nabla M_1(\theta_{10},\theta_{20})\times J_1(\theta_{10})^{-1}\sqrt{T}G_1(\theta_{10}) \\
&\quad + \nabla M_2(\theta_{10},\theta_{20})\times J_2(\theta_{20})^{-1}\sqrt{T}G_2(\theta_{20}) + o_p(1) \\
&= \sqrt{T}M(\theta_0) - J_{M\theta}\times J^{-1}_{\theta\theta}\sqrt{T}G(\theta_0) + o_p(1) \\
&= A_1\begin{bmatrix}\sqrt{T}M(\theta_0)\\ \sqrt{T}G(\theta_0)\end{bmatrix} + o_p(1) \qquad (59)
\end{aligned}$$

where $J_{M\theta} = -\begin{bmatrix}\nabla M_1 & \nabla M_2\end{bmatrix}$, $G(\theta_0) = \begin{bmatrix}G_1(\theta_{10})\\ G_2(\theta_{20})\end{bmatrix}$, $J_{\theta\theta} = \begin{bmatrix}J_1(\theta_{10}) & 0\\ 0 & J_2(\theta_{20})\end{bmatrix}$ and $A_1 = [I_r,\ \ -J_{M\theta}\times J_{\theta\theta}]$. Thus, when the proposition (3) holds we can write that

$$\sqrt{T}\hat{M}_T \xrightarrow{d} N(0,V_1)$$

where $V_1 = A_1\Sigma A_1'$. ■

A.8 Proof of Lemma 7

Proof. Assuming that the specification of the log-likelihood for the FQML estimation of parameters is correct, we can use a generalized (conditional) IM equality which says that

$$\begin{aligned}
\mathrm{E}_0\left[\frac{\partial}{\partial\theta_i}\left(\frac{\partial l_{it}(\theta_i)}{\partial\theta_i'}\right)|\mathcal{F}_{t-1}\right] &= -\mathrm{E}_0\left[\left(\frac{\partial l_{it}(\theta_i)}{\partial\theta_i}\right)\left(\frac{\partial l_t^*(\theta,\rho)}{\partial\theta_i'}\right)\middle|\mathcal{F}_{t-1}\right] \\
\mathrm{E}_0\left[\frac{\partial m_t(\theta)}{\partial\theta_i'}\middle|\mathcal{F}_{t-1}\right] &= -\mathrm{E}_0\left[m_t(\theta)\frac{\partial l_t^*(\theta,\rho)}{\partial\theta_i'}\middle|\mathcal{F}_{t-1}\right],
\end{aligned}$$

where $\dfrac{\partial l_t^*(\theta,\rho)}{\partial\theta_i'}$ is the score for θ_i, $i=1,2.$ from the FQMLE log-likelihood.

Then $J_i = -\operatorname{plim}_{T\to\infty}\frac{1}{T}\left[\sum_{t=1}^{T}\frac{\partial^2 l_{it}(\theta_i)}{\partial\theta_i\partial\theta_i'}\right] = \operatorname{plim}_{T\to\infty}\dfrac{1}{T}Q_i'S_i$ and $J_{M\theta} = -\operatorname{plim}_{T\to\infty}\dfrac{1}{T}R'S_i$. Substituting these into (59) yields

$$\begin{aligned}
\sqrt{T}\hat{M} &= \sqrt{T}M(\theta_0) - \frac{1}{T}R'S\begin{bmatrix}\frac{1}{T}Q_1'S_1 & 0 \\ 0 & \frac{1}{T}Q_2'S_2\end{bmatrix}^{-1}\sqrt{T}G(\theta_0) + o_p(1) \\
&= A_1\begin{bmatrix}\sqrt{T}M(\theta_0) \\ \sqrt{T}G(\theta_0)\end{bmatrix} + o_p(1)
\end{aligned}$$

where $A_1 = \left[I_r : \ -R'S\begin{pmatrix}(Q_1'S_1)^{-1} & 0 \\ 0 & (Q_2'S_2)^{-1}\end{pmatrix}\right]$ and $S = [S_1,\ S_2]$. Now defining $B = [R,\ Q_1,\ Q_2]$, $\Sigma = \operatorname{plim}_{T\to\infty}B'B$. Hence, the variance-covariance matrix V_1 can be written as $V_1 = \operatorname{plim}_{T\to\infty}\dfrac{1}{T}A_1B'BA_1' = \operatorname{plim}_{T\to\infty}\dfrac{1}{T}W'W$ where $W = BA_1' = R - Q_1(S_1'Q_1)^{-1}S_1'R - Q_2(S_2'Q_2)^{-1}S_2'R$.

Now V_1 can be consistently estimated by $\hat{V}_{1T} = \dfrac{1}{T}\hat{W}'\hat{W}$, where hats denote θ_0 replaced by the individual GARCH estimators, $\hat{\theta}$, and ρ_0 replaced by the estimator $\hat{\rho} = \dfrac{1}{T}\sum_{t=1}^{T}\hat{\zeta}_{1t}\hat{\zeta}_{2t}$. ■

References

[1] Andreou, E. and Ghysels, E., 2003. Tests for breaks in the conditional co-movements of asset returns. Statistica Sinica, 13 (4), pp.1045–1074.

[2] Aslanidis, N., Osborn, D. and Sensier, M., 2008. Co-movements between US and UK stock prices: the roles of macroeconomic information and time-varying conditional correlations. Manchester Economics Discussion Paper Series 0805.

[3] Bera, A.K. and Kim, S., 2002. Testing constancy of correlation and other specifications of the BGARCH model with an application to international equity returns. Journal of Empirical Finance, 9(2), pp.171–195.

[4] Bera, A.K. and McKenzie, C.R., 1986. Alternative forms and properties of the score test. Journal of Applied Statististics, 13(1), pp.13–25.

[5] Billo, M., Caporin, M. and Gobbo, M., 2006. Flexible dynamic conditional correlation multivariate GARCH models for asset allocation. Applied Financial Economics Letters, 2(2), pp.123–130.

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

[7] Bollerslev, T., 1990. Modeling the coherence in short-run nominal exchange rates: a multivariate generalized ARCH model. Review of Economics and Statistics, 72(3), pp.498–505.

[8] Bollerslev, T. and Wooldridge, J.M., 1992. Quasi–Maximum Likelihood Estimation and Inference in Dynamic Models with Time–varying Covariances. Econometric Reviews, 11(2), pp.143–172.

[9] Chesher, A., 1984. Testing for neglected heterogeneity. Econometrica, 52(4), pp.865–872.

[10] Chesher, A. and Spady, R., 1991. Asymptotic expansions of the information matrix test statistics. Econometrica, 59(3), pp.787–816.

[11] Comte, F. and Lieberman, O., 2003. Asymptotic Theory for Multivariate GARCH Processes. Journal of Multivariate Analysis, 84(1), pp.61–84.

[12] Cox, D.R., 1983. Some remarks on overdispersion. Biometrika, 70(1), pp. 269–274.

[13] Dagenais, M.G. and Dufour, J.M., 1991. Invariance, nonlinear models, and asymptotic tests. Econometrica, 59(6), pp.1601-1615.

[14] Davidson, R. and MacKinnon, J.G., 1983. Small sample properties of alternative forms of the Lagrange multiplier test. Economics Letters, 12(3-4), pp.269–275.

[15] Engle, R.F., 2002. Dynamic conditional correlation: a simple class of multivariate GARCH models. Journal of Business and Economic Statistics, 20(3), pp.339–350.

[16] Engle, R.F. and Kroner, F.K., 1995. Multivariate simultaneous generalized ARCH. Econometric Theory, 11(1), pp.122–150.

[17] Engle, R.F and Ng, V.K., 1993. Measuring and testing the impact of news on volatility. Journal of Finance, 48(5), pp.1749–1778.

[18] Engle, R.F and Sheppard, K., 2008. Evaluating the specification of covariance models for large portfolios. Working Paper, Department of Economics, University of Oxford.

[19] Hafner, C.M., van Dijk, D. and Franses, P.H., 2005. Semi-parametric modelling of correlation dynamics. In: Fomby, T., Hill, C. and Terrell, D. eds. Advances in Econometrics, vol. 20/A: 59–103. Amsterdam: Elsevier Sciences.

[20] Hafner, C.M. and Herwartz, H., 2008. Analytical quasi maximum likelihood inference in multivariate volatility models. Metrika, 67(2), pp.219-239.

[21] Halunga. A. and Orme, C.D., 2009. First-order asymptotic theory for misspecification tests of GARCH models. Econometric Theory, 25(2), pp.364-410.

[22] Jeantheau, T., 1998. Strong consistency of estimators for multivariate ARCH models. Econometric Theory, 14(1), pp.70–86.

[23] Kroner, K.F. and S. Claessens, 1991. Optimal Dynamic Hedging Portfolios and the Currency Composition of External Debt. Journal of International Money and Finance, 10(1), pp.131-148.

[24] Kroner, K.F., and Sultan, J., 1991. Exchange rate volatility and time varying hedge ratios. In: Rhee, S.G. and Chang, R.P. eds, Pacific-Basin Capital Market Research, 2, Amsterdam: North-Holland, pp.397-412.

[25] Kroner, K.F., and Sultan, J., 1993. Time varying distribution and dynamic hedging with foreign currency futures. Journal of Financial and Quantitative Analysis, 28(4), pp.535-551.

[26] Lee, J., 2006. The comovement between output and prices: Evidence from a dynamic conditional correlation GARCH model. Economics Letters, 91(1), pp.110-116

[27] Lien, D. and Tse, Y.K., 1998. Hedging Time-Varying Downside Risk. Journal of Futures Markets, 18(2), pp.705 - 722.

[28] Lien, D., Tse, Y.K. and Tsui, A. K. C., 2002. Evaluating the hedging performance of the constant-correlation GARCH model. Applied Financial Economics, 12(11), pp.791-798.

[29] Ling, S. and McAleer, M., 2003. Asymptotic theory for a vector ARMA-GARCH model. Econometric Theory, 19(2), pp.280–310.

[30] Li, W.K. and Mak, T.K., 1994. On the squared residual autocorrelations in non-linear time series with conditional heteroskedasticity. Journal of Time Series Analysis, 15(6), pp.627–636.

[31] Longin, F. and Solnik, B., 1995. Is the correlation in international equity returns constant: 1960–1990? Journal of International Money and Finance, 14(1), pp.3–26.

[32] Lundbergh, S. and Teräsvirta, T., 2002. Evaluating GARCH models. Journal of Econometrics, 110(2), pp.417–435.

[33] Magnus, J. R. and Neudecker, H., 1999. Matrix Differential Calculus. Wiley Series in Probability and Statistics (revised edition). Wiley.

[34] Nakatani, T. and Teräsvirta, T., 2009. Testing for volatility interactions in the Constant Conditional Correlation GARCH model. Econometrics Journal, 12(1), pp.147-163.

[35] Nakatani, T. and Teräsvirta, T., 2008. Appendix to Testing for volatility interactions in the Constant Conditional Correlation GARCH model. Department of Economic Statistics, Stockholm School of Economics, available at http://swopec.hhs.se/hastef/abs/hastef0649.htm.

[36] Nelson, D.B., 1991. Conditional heteroskedasticity in asset returns: a new approach. Econometrica, 59(2), pp.349–370.

[37] Newey, W.K., 1985. Maximum likelihood specification testing and conditional moment tests. Econometrica, 53(5), pp.1047–1070.

[38] Park, T. and Switzer, L., 1995. Time-varying distribution and the optimal hedge ratios for stock index futures. Applied Financial Economics, 5(3), pp.131–137.

[39] Silvennoinen, A. and Teräsvirta, T., 2008. Multivariate GARCH Models. CREATES Research Paper 2008-6.

[40] Tse, Y.K., 2000. A test for constant correlations in a multivariate GARCH model. Journal of Econometrics, 98 (1), pp.107–127.

[41] Weiss, A. A., 1986. Asymptotic theory for ARCH models: Estimation and Testing. Econometric Theory, 2(1), pp.107-131.

[42] White, H., 1982. Maximum Likelihood Estimation of Misspecified Models. Econometrica, 50(1), pp.1-25.

[43] Wooldridge, J., 1990. A unified approach to robust, regression-based specification tests. Econometric Theory, 6(1), pp.17-43.