A Lagrange Multiplier Test for Testing the Adequacy of the Constant Conditional Correlation GARCH Model

Paul Catani*
Hanken School of Economics, FIN-00101 Helsinki, Finland

Timo Teräsvirta†
CREATES, Aarhus University, DK-8210 Aarhus V, Denmark

Meiqun Yin‡
Beijing International Studies University, 100024, Beijing, China

August 31, 2016

Abstract

A Lagrange multiplier test for testing the parametric structure of a constant conditional correlation generalized autoregressive conditional heteroskedasticity (CCC-GARCH) model is proposed. The test is based on decomposing the CCC-GARCH model multiplicatively into two components, one of which represents the null model, whereas the other one describes the misspecification. A simulation study shows that the test has good finite sample properties. We compare the test with other tests for misspecification of multivariate GARCH models. The test has high power against alternatives where the misspecification is in the GARCH parameters and is superior to other tests. The test is not greatly affected by misspecification in the conditional correlations and is therefore well suited for considering misspecification of GARCH equations.

JEL Codes: C32, C52, C58

Keywords: constant conditional correlation, LM test, misspecification testing, modelling volatility, multivariate GARCH

Acknowledgements. The first author acknowledges financial support from The Society of Swedish Literature in Finland. The research of the second author has been supported by CREATES - Centre for Research in Econometric Analysis of Time Series (DNRF78), funded by the Danish National Research Foundation. The third author has received support from the China National Social Science Fund, Project No. 13BGL042. Material from this paper has been presented at the 7th International Conference on Computational and Financial Econometrics (CFE 2013), 14-16 December 2013, London. We wish to thank Christian Conrad and James Davidson for useful remarks. Special thanks go to the referee for careful reading of the paper. Errors and shortcomings in this work remain our own responsibility.

*Hanken School of Economics, email: paul.catani@hanken.fi
†CREATES, Aarhus University, email: tterasvirta@creates.au.dk
‡Beijing International Studies University, email: yinmeiqun@bisu.edu.cn

1 Introduction

Multiple GARCH models have become an important tool in forecasting volatility of portfolios. There are several classes of multivariate GARCH models, beginning with the general Vector GARCH model of Bollerslev, Engle & Woodridge (1988). This model is even 'too general' in the sense that conditional covariance matrices generated by this model are positive definite with probability less than one. Following this first attempt at joint modelling of conditional variances and covariances using the GARCH approach, the main goal of econometricians has been to develop models whose parametric structure would guarantee positive definiteness of the conditional covariance matrix. Two classes of such models have become quite popular. The first one is the so-called BEKK-GARCH model discussed by Engle & Kroner (1995), and the second one is the family of conditional correlation models. The basic model nested in the other members of this family is the Constant Conditional Correlation GARCH (CCC-GARCH) model by Bollerslev (1990). For information about these and other multivariate GARCH models, see Bauwens, Laurent & Rombouts (2006) and Silvennoinen & Teräsvirta (2015).

In this paper the focus is on conditional correlation GARCH models. While they are frequently fitted to financial time series, testing the parametric structure of the GARCH equations in them has not been very common. Our aim is to derive a portmanteau test for testing misspecification of the GARCH structure of these models. The predecessor of our test is the portmanteau test of Ling & Li (1997) who generalised the univariate test of Li & Mak (1994) to a multivariate situation. Their test is not restricted to conditional correlation GARCH models, but by a suitable choice of the conditional covariance matrix it becomes a misspecification test of the GARCH equations in the CCC-GARCH model.

Nakatani & Teräsvirta (2009) derived a test of the CCC-GARCH model against the Extended CCC-GARCH model of Jeantheau (1998). In their Lagrange multiplier (LM-) test the alternative to the GARCH equations is the model with GARCH equations that contains lags of squared errors and conditional variances from other GARCH equations. Our aim is to derive a general portmanteau test in the spirit of Ling & Li (1997) such that the alternative to the GARCH equations is more general than in the test of Nakatani & Teräsvirta (2009). It is based on decomposing the conditional variance equations in the CCC-GARCH model multiplicatively into two components, one of which represents the null model, whereas the other one describes the misspecification. The inspiration comes from the univariate 'no ARCH in GARCH' test in Lundbergh & Teräsvirta (2002). This leads to a portmanteau test that is more general than that of Ling & Li (1997).

A practical question in applying tests of the GARCH structure of the CCC-GARCH model is whether these tests also have power against misspecification of the correlation structure. This will be investigated by simulation. There are also tests of the correlation structure of the CCC-GARCH model. Tse (2000) derived a portmanteau-type test against the alternative that the conditional correlations are not constant over time. Silvennoinen & Teräsvirta (2009) constructed an LM test against the Smooth Transition Conditional Correlation GARCH (STCC-GARCH) model. The question then is whether tests of constant conditional correlations in turn have power against misspecification in the GARCH equations. In this paper this problem is investigated by simulating the test of Tse (2000). His test can be viewed as a portmanteau-type test without a specific alternative to constant correlations.

It would be useful to test the adequacy of GARCH equations when the estimated model is a time-varying conditional correlation model such as the DCC-GARCH model of Engle (2002), the STCC-GARCH model, or the Markov-switching CC-GARCH model of Pelletier (2006). The difficulty is, however, that asymptotic normality of the maximum likelihood estimators of the parameters of these models has not been rigorously proven. For an illuminating discussion, see Engle & Kelly (2012). The corresponding proof exists for the CCC-GARCH model, see Ling & McAleer (2003), which is why that model constitutes the null hypothesis for the test derived in this paper.

The plan of the paper is as follows. In section 2 the CCC-GARCH process is defined and we present the decomposition of the conditional variance equations which our test is based upon. In section 3 we give the first and second order partial derivatives of the quasi-log-likelihood function of the decomposed CCC-GARCH model. The LM test is derived in section 4 and section 5 contains a bivariate illustration of the test. The finite sample properties of the test are studied by Monte Carlo simulations in section 6. Section 7 concludes. Mathematical proofs can be found in the Appendix.

2 Model

Consider the following stochastic model of a random vector $\mathbf{y}_t$:

$$\mathbf{y}_t = \mathsf{E}\{\mathbf{y}_t|\mathcal{F}_{t-1}\} + \boldsymbol{\varepsilon}_t$$

where $\mathbf{y}_t = (y_{1t}, \ldots, y_{mt})'$ is an $m \times 1$ vector and $\mathcal{F}_{t-1}$ contains the conditioning information available at $t-1$. The m-dimensional error term $\boldsymbol{\varepsilon}_t$ is decomposed as follows:

$$\boldsymbol{\varepsilon}_t = \mathbf{D}_t\mathbf{z}_t \tag{1}$$

where

$$\mathbf{D}_t = \mathrm{diag}(h_{1t}^{1/2}, \ldots, h_{mt}^{1/2}) \tag{2}$$

is a diagonal matrix of conditional standard deviations of the elements of $\boldsymbol{\varepsilon}_t$. In what follows we assume $\mathsf{E}\{\mathbf{y}_t|\mathcal{F}_{t-1}\} = \mathbf{0}$ for simplicity and that h_{it} follows a GARCH(1,1) process

$$h_{it} = \alpha_{i0} + \alpha_{i1}\varepsilon_{i,t-1}^2 + \beta_{i1}h_{i,t-1}, \tag{3}$$

where $\alpha_{i0} > 0$, α_{i1} and β_{i1} are nonnegative, $i = 1, \ldots, m$. Furthermore, $\mathbf{z}_t \sim \mathrm{iid}(\mathbf{0}, \mathbf{P})$, where $\mathbf{P} = [\rho_{ij}]$ is a positive definite correlation matrix, i.e., $\rho_{ii} = 1$, $i = 1, \ldots, m$.

Equation (3) may be generalised to contain asymmetric or higher-order terms. From (1) we have

$$\mathbf{z}_t = (z_{1t}, \ldots, z_{mt})' = \mathbf{D}_t^{-1}\boldsymbol{\varepsilon}_t = (\varepsilon_{1t}h_{1t}^{-1/2}, \ldots, \varepsilon_{mt}h_{mt}^{-1/2})' \tag{4}$$

and equations (1) and (4) define a CCC-GARCH model. The model can be written as

$$\mathbf{h}_t = \mathbf{a}_0 + \mathbf{A}_1\boldsymbol{\varepsilon}_{t-1}^{(2)} + \mathbf{B}_1\mathbf{h}_{t-1}, \tag{5}$$

where $\boldsymbol{\varepsilon}_t^{(2)} = (\varepsilon_{1t}^2, \ldots, \varepsilon_{mt}^2)'$, $\mathbf{h}_t = (h_{1t}, \ldots, h_{mt})'$ and $\mathbf{a}_0 = (\alpha_{10}, \ldots, \alpha_{m0})'$ are $(m \times 1)$ vectors and $\mathbf{A}_1$ and $\mathbf{B}_1$ are diagonal $(m \times m)$ parameter matrices with positive diagonal elements α_{i1} and β_{i1}, $i = 1, \ldots, m$, respectively.

In order to construct a misspecification test for the CCC-GARCH model (1), we assume that $\mathbf{z}_t = \mathbf{G}_t\mathbf{u}_t$, where

$$\mathbf{G}_t = \mathrm{diag}(g_{1t}^{1/2}, \ldots, g_{mt}^{1/2}) \tag{6}$$

with

$$g_{it} = 1 + \sum_{j=1}^{r} \zeta_{ij} z_{i,t-j}^2, \tag{7}$$

and $\mathbf{u}_t = (u_{1t}, \ldots, u_{mt})' = (\varepsilon_{1t} h_{1t}^{-1/2} g_{1t}^{-1/2}, \ldots, \varepsilon_{mt} h_{mt}^{-1/2} g_{mt}^{-1/2})' \sim \text{iid}(\mathbf{0}, \mathbf{P})$. Then (1) can be written as follows:

$$\boldsymbol{\varepsilon}_t = \mathbf{D}_t \mathbf{G}_t \mathbf{u}_t \tag{8}$$

and (8) can be regarded as an 'ARCH nested in GARCH' model. For the univariate case, see Lundbergh & Teräsvirta (2002) and for another definition of g_{it}, in which g_{it} is a deterministic positive-valued function, see Amado & Teräsvirta (2013).

Let $\boldsymbol{\zeta} = (\boldsymbol{\zeta}_1', ..., \boldsymbol{\zeta}_m')'$ be an $mr \times 1$ matrix where $\boldsymbol{\zeta}_i = (\zeta_{i1}, ..., \zeta_{ir})', i = 1, ..., m$, is an $r \times 1$ vector. Our misspecification test consists of testing

$$\text{H}_0\text{: } \boldsymbol{\zeta} = \mathbf{0} \text{ or } \mathbf{G}_t \equiv \mathbf{I} \tag{9}$$

in the model (7). Thus under H_0, $\{\boldsymbol{\varepsilon}_t\}$ follows a CCC-GARCH model, and the alternative implies that there is dynamic structure unaccounted for in this model, because none of the sequences $\{z_{i,t}\}$ is a sequence of independent random variables.

3 The log-likelihood function and its partial derivatives

3.1 The log-likelihood function

First, we introduce some notation. Let $\mathbf{0}_m$ be an $m \times 1$ null vector, $\mathbf{0}_{mn}$ an $mn \times 1$ null vector, $\mathbf{1}_m$ an $m \times 1$ vector of ones, $\mathbf{I}_m$ an $m \times m$ identity matrix, and $\mathsf{diag}(\mathbf{a})$ a diagonal matrix whose diagonal elements are the elements of vector $\mathbf{a}$. In order to derive the Lagrange Multiplier statistic for testing the null hypothesis (9), we need the log-likelihood function of the model and its first two partial derivatives. Under the null hypothesis, we assume that $\{\boldsymbol{\varepsilon}_t\}$ is a sequence of vector white noise with $\mathsf{E}\boldsymbol{\varepsilon}_t = \mathbf{0}_m$ and the conditional covariance matrix $\boldsymbol{\Sigma}_t = \mathbf{D}_t \mathbf{P} \mathbf{D}_t$. Let $\boldsymbol{\omega} = (\boldsymbol{\omega}_1', ..., \boldsymbol{\omega}_m')'$ be a $3m$-dimensional vector where $\boldsymbol{\omega}_i = (\alpha_{i0}, \alpha_{i1}, \beta_{i1})'$, $i = 1, ..., m$, and $\boldsymbol{\rho} = \mathsf{vecl}(\mathbf{P}) = (\rho_{12}, ..., \rho_{1m}, \rho_{23}, ..., \rho_{2m}, ..., \rho_{m-1,m})'$ be an $m(m-1)/2$-dimensional vector. Furthermore, let $\boldsymbol{\zeta} = (\boldsymbol{\zeta}_1', ..., \boldsymbol{\zeta}_m')'$ be an mr-dimensional vector such that $\boldsymbol{\zeta}_i = (\zeta_{i1}, ..., \zeta_{ir})', i = 1, ..., m$, is an $r \times 1$ vector, and finally, set $\boldsymbol{\theta} = (\boldsymbol{\omega}', \boldsymbol{\rho}', \boldsymbol{\zeta}')'$. Thus, the quasi-log-likelihood of the CCC-GARCH model for observation t takes the form of the Gaussian log-likelihood:

$$\begin{aligned} l_t(\boldsymbol{\theta}) &= -(1/2)\sum_{i=1}^{m} \ln h_{it} - (1/2)\sum_{i=1}^{m} \ln g_{it} - (1/2)\ln|\mathbf{P}| - (1/2)\mathbf{u}_t'\mathbf{P}^{-1}\mathbf{u}_t \\ &= -\ln|\mathbf{D}_t| - \ln|\mathbf{G}_t| - (1/2)\ln|\mathbf{P}| - (1/2)\mathbf{u}_t'\mathbf{P}^{-1}\mathbf{u}_t. \end{aligned} \tag{10}$$

Maximising

$$L_T(\boldsymbol{\theta}) = \sum\nolimits_{t=1}^{T} l_t(\boldsymbol{\theta})$$

with respect to $\boldsymbol{\theta}$ yields the quasi maximum likelihood estimator (QMLE) $\hat{\boldsymbol{\theta}}$.

To ensure asymptotic normality of the QMLE, we make the following assumptions:
Assumption 1 (Stationarity). *Roots of* $\mathsf{det}(\mathbf{I}_m - \mathbf{A}_1 x - \mathbf{B}_1 x)$ *lie outside the unit circle.*
Assumption 2. *The parameter space* Θ *is a compact subspace of Euclidean space; the matrix* $\mathbf{P}$ *is a finite and positive definite symmetric matrix, with the elements on the main diagonal being 1 and the largest absolute eigenvalue of the matrix* $\mathbf{P}$ *having a positive lower bound over*

Θ; *each* α_{i1} *and* β_{i1} *is nonnegative,* $i = 1, ..., m$, *and each element of* $\{\alpha_{i0}, i = 1, ..., m\}$ *has positive lower and upper bounds over* Θ. *Furthermore, if* $\beta_{i1} > 0$, *then* $\alpha_{i1} > 0$, $i = 1, ..., m$.

Assumption 3 (Identifiability). *The formulation at the true parameter value* $\boldsymbol{\theta}_0$ *of the CCC-GARCH-model is minimal.*

For definition of minimality, see Jeantheau (1998).

Assumption 4. $\mathsf{E}|\varepsilon_{it}^6| < \infty, i = 1, ..., m.$

Under Assumption 1 the CCC-GARCH(1,1) model has a unique weakly stationary solution. Furthermore the model is also strictly stationary and ergodic (see Jeantheau (1998) and Ling & McAleer (2003)).

Jeantheau (1998) shows that under Assumption 3 the model is identifiable. Define $\mathbf{B}(\mathsf{L}) = \mathbf{I}_m - \mathbf{B}_1\mathsf{L}$ and $\mathbf{A}(\mathsf{L}) = \mathbf{A}_1\mathsf{L}$ where L is the lag operator. Sufficient conditions for Assumption 3 to hold are:

- $\mathsf{det}(\mathbf{A}(\mathsf{L})) \neq 0$ and $\mathsf{det}(\mathbf{B}(\mathsf{L})) \neq 0$.
- $\mathbf{A}(\mathsf{L})$ and $\mathbf{B}(\mathsf{L})$ are left coprime.
- $\mathbf{A}(\mathsf{L})$ or $\mathbf{B}(\mathsf{L})$ is column reduced.

$\mathbf{A}(\mathsf{L})$ and $\mathbf{B}(\mathsf{L})$ are left coprime if any of the greatest common left divisors, $\mathbf{D}$, of $\mathbf{A}(\mathsf{L})$ and $\mathbf{B}(\mathsf{L})$ are unimodular. $\mathbf{D}$ is unimodular if $\mathsf{det}(\mathbf{D})$ is not equal to zero and if it is independent of the lag operator L. Furthermore, the polynomial matrix $\mathbf{A}(\mathsf{L})$ or $\mathbf{B}(\mathsf{L})$ is column reduced if $\mathsf{det}(\mathbf{A}_1) \neq 0$ or $\mathsf{det}(\mathbf{B}_1) \neq 0$, respectively. See Jeantheau (1998) for details and proof.

Assumptions 2 and 4 are crucial for the proof of asymptotic normality of the QMLE, see Ling & McAleer (2003).

3.2 The score and the information matrix of the log-likelihood function

In this section we define the first and second partial derivatives of (10). Let $\mathbf{q}_t(\boldsymbol{\theta}) = \partial l_t(\boldsymbol{\theta})/\partial\boldsymbol{\theta}$ be the score vector for observation t, and let

$$\mathbf{q}(\boldsymbol{\theta}) = (1/T)\sum_{t=1}^{T}\mathbf{q}_t(\boldsymbol{\theta}) \tag{11}$$

be the average score. The $3m + m(m-1)/2 + mr$-dimensional score vector for the observation t of (10) has the following form

$$\mathbf{q}_t(\boldsymbol{\theta}) = (\frac{\partial l_t(\boldsymbol{\theta})}{\partial\boldsymbol{\omega}'}, \frac{\partial l_t(\boldsymbol{\theta})}{\partial\boldsymbol{\rho}'}, \frac{\partial l_t(\boldsymbol{\theta})}{\partial\boldsymbol{\zeta}'})'$$

where, see Nakatani & Teräsvirta (2009),

$$\frac{\partial l_t(\boldsymbol{\theta})}{\partial\boldsymbol{\omega}} = -\nabla\mathbf{D}_t\mathsf{vec}\left(\mathbf{D}_t^{-1} - \frac{1}{2}\mathbf{D}_t^{-1}\mathbf{G}_t^{-1}\boldsymbol{\varepsilon}_t\boldsymbol{\varepsilon}_t'\mathbf{G}_t^{-1}\mathbf{M}_t^{-1} - \frac{1}{2}\mathbf{M}_t^{-1}\mathbf{G}_t^{-1}\boldsymbol{\varepsilon}_t\boldsymbol{\varepsilon}_t'\mathbf{G}_t^{-1}\mathbf{D}_t^{-1}\right) \tag{12}$$

and

$$\frac{\partial l_t(\boldsymbol{\theta})}{\partial\boldsymbol{\rho}} = -\frac{1}{2}\nabla\mathbf{P}\mathsf{vec}\left(\mathbf{P}^{-1} - \mathbf{P}^{-1}\mathbf{D}_t^{-1}\mathbf{G}_t^{-1}\boldsymbol{\varepsilon}_t\boldsymbol{\varepsilon}_t'\mathbf{G}_t^{-1}\mathbf{D}_t^{-1}\mathbf{P}^{-1}\right) \tag{13}$$

with $\mathbf{M}_t = \mathbf{D}_t\mathbf{P}\mathbf{D}_t$, $\mathbf{H}_t = \mathbf{G}_t\mathbf{M}_t\mathbf{G}_t$, $\nabla\mathbf{D}_t = \partial\mathsf{vec}(\mathbf{D}_t)/\partial\boldsymbol{\omega}'$ and $\nabla\mathbf{P} = \partial\mathsf{vec}(\mathbf{P})/\partial\boldsymbol{\rho}'$.

The following lemma gives the first-order partial derivative of the log-likelihood function with respect to $\boldsymbol{\zeta}$.

Lemma 1 *The bottom block of the score vector has the following form:*

$$\frac{\partial l_t(\boldsymbol{\theta})}{\partial \boldsymbol{\zeta}} = -\nabla \mathbf{G}_t \mathsf{vec}\left(\mathbf{G}_t^{-1} - \frac{1}{2}\mathbf{H}_t^{-1}\boldsymbol{\varepsilon}_t\boldsymbol{\varepsilon}_t'\mathbf{G}_t^{-1} - \frac{1}{2}\mathbf{G}_t^{-1}\boldsymbol{\varepsilon}_t\boldsymbol{\varepsilon}_t'\mathbf{H}_t^{-1} \right) \tag{14}$$

where $\nabla \mathbf{G}_t = \partial \mathsf{vec}(\mathbf{G}_t)'/\partial \boldsymbol{\zeta}$. *Under* H_0,

$$\frac{\partial l_t(\boldsymbol{\theta})}{\partial \boldsymbol{\zeta}} = -\nabla \mathbf{G}_t \mathsf{vec}(\mathbf{I} - \frac{1}{2}\mathbf{M}_t^{-1}\boldsymbol{\varepsilon}_t\boldsymbol{\varepsilon}_t' - \frac{1}{2}\boldsymbol{\varepsilon}_t\boldsymbol{\varepsilon}_t'\mathbf{M}_t^{-1}).$$

Proof. See the Appendix.

From (12), (13) and Lemma 1 it follows that under the null hypothesis, the average score vector has the form

$$\begin{aligned} \mathbf{q}(\boldsymbol{\theta}) &= \frac{1}{T}\sum_{t=1}^{T}\left[\frac{\partial l_t(\boldsymbol{\theta})}{\partial \boldsymbol{\omega}'}, \frac{\partial l_t(\boldsymbol{\theta})}{\partial \boldsymbol{\rho}'}, \frac{\partial l_t(\boldsymbol{\theta})}{\partial \boldsymbol{\zeta}'}\right]' \\ &= \frac{1}{T}\sum_{t=1}^{T}\left[\begin{array}{c} \nabla \mathbf{D}_t \mathsf{vec}(\mathbf{D}_t^{-1} - \frac{1}{2}\mathbf{D}_t^{-1}\boldsymbol{\varepsilon}_t\boldsymbol{\varepsilon}_t'\mathbf{M}_t^{-1} - \frac{1}{2}\mathbf{M}_t^{-1}\boldsymbol{\varepsilon}_t\boldsymbol{\varepsilon}_t'\mathbf{D}_t^{-1}) \\ -\nabla \mathbf{P}\mathsf{vec}(\mathbf{P}^{-1} - \mathbf{P}^{-1}\mathbf{D}_t^{-1}\boldsymbol{\varepsilon}_t\boldsymbol{\varepsilon}_t'\mathbf{D}_t^{-1}\mathbf{P}^{-1}) \\ -\nabla \mathbf{G}_t \mathsf{vec}(\mathbf{I} - \frac{1}{2}\mathbf{M}_t^{-1}\boldsymbol{\varepsilon}_t\boldsymbol{\varepsilon}_t' - \frac{1}{2}\boldsymbol{\varepsilon}_t\boldsymbol{\varepsilon}_t'\mathbf{M}_t^{-1}) \end{array}\right]. \end{aligned} \tag{15}$$

The population information matrix is

$$\mathbf{I}(\boldsymbol{\theta}_0) = (1/T)\mathsf{E}\mathbf{q}(\boldsymbol{\theta}_0)\mathbf{q}(\boldsymbol{\theta}_0)' = \mathsf{E}\mathbf{q}_t(\boldsymbol{\theta}_0)\mathbf{q}_t(\boldsymbol{\theta}_0)' \tag{16}$$

where $\boldsymbol{\theta}_0$ is the true parameter and $\mathbf{q}_t(\boldsymbol{\theta}_0)$ is $\mathbf{q}_t(\boldsymbol{\theta})$ evaluated at $\boldsymbol{\theta}_0$. The negative of the expected Hessian evaluated at $\boldsymbol{\theta}_0$ equals

$$\mathbf{J}(\boldsymbol{\theta}_0) = -(1/T)\mathsf{E}\sum_{t=1}^{T}\frac{\partial^2 l_t(\boldsymbol{\theta}_0)}{\partial \boldsymbol{\theta}\partial \boldsymbol{\theta}'}. \tag{17}$$

The Hessian for observation t has the form

$$\mathbf{J}_t(\boldsymbol{\theta}) = \frac{\partial^2 l_t(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}\partial \boldsymbol{\theta}'} = \left[\begin{array}{ccc} \mathbf{J}_{\omega\omega t}(\boldsymbol{\theta}) & \mathbf{J}_{\omega\rho t}(\boldsymbol{\theta}) & \mathbf{J}_{\omega\zeta t}(\boldsymbol{\theta}) \\ \mathbf{J}_{\rho\omega t}(\boldsymbol{\theta}) & \mathbf{J}_{\rho\rho t}(\boldsymbol{\theta}) & \mathbf{J}_{\rho\zeta t}(\boldsymbol{\theta}) \\ \mathbf{J}_{\zeta\omega t}(\boldsymbol{\theta}) & \mathbf{J}_{\zeta\rho t}(\boldsymbol{\theta}) & \mathbf{J}_{\zeta\zeta t}(\boldsymbol{\theta}) \end{array}\right] = \left[\begin{array}{ccc} \frac{\partial^2 l_t(\boldsymbol{\theta})}{\partial \boldsymbol{\omega}\partial \boldsymbol{\omega}'} & \frac{\partial^2 l_t(\boldsymbol{\theta})}{\partial \boldsymbol{\omega}\partial \boldsymbol{\rho}'} & \frac{\partial^2 l_t(\boldsymbol{\theta})}{\partial \boldsymbol{\omega}\partial \boldsymbol{\zeta}'} \\ \frac{\partial^2 l_t(\boldsymbol{\theta})}{\partial \boldsymbol{\omega}\partial \boldsymbol{\rho}'} & \frac{\partial^2 l_t(\boldsymbol{\theta})}{\partial \boldsymbol{\rho}\partial \boldsymbol{\rho}'} & \frac{\partial^2 l_t(\boldsymbol{\theta})}{\partial \boldsymbol{\zeta}\partial \boldsymbol{\rho}'} \\ \frac{\partial^2 l_t(\boldsymbol{\theta})}{\partial \boldsymbol{\omega}\partial \boldsymbol{\zeta}'} & \frac{\partial^2 l_t(\boldsymbol{\theta})}{\partial \boldsymbol{\rho}\partial \boldsymbol{\zeta}'} & \frac{\partial^2 l_t(\boldsymbol{\theta})}{\partial \boldsymbol{\zeta}\partial \boldsymbol{\zeta}'} \end{array}\right] \tag{18}$$

where the expression for the upper left-hand 2×2 block can be found in Nakatani & Teräsvirta (2009). The information matrix for observation t under the null hypothesis is given by

$$\mathbf{I}(\boldsymbol{\theta}_0) = -\mathsf{E}[\mathbf{J}_t(\boldsymbol{\theta}_0)] = -\left[\begin{array}{ccc} \mathbf{I}_{\omega\omega}(\boldsymbol{\theta}_0) & \mathbf{I}_{\omega\rho}(\boldsymbol{\theta}_0) & \mathbf{I}_{\omega\zeta}(\boldsymbol{\theta}_0) \\ \mathbf{I}_{\omega\rho}'(\boldsymbol{\theta}_0) & \mathbf{I}_{\rho\rho}(\boldsymbol{\theta}_0) & \mathbf{I}_{\rho\zeta}(\boldsymbol{\theta}_0) \\ \mathbf{I}_{\omega\zeta}'(\boldsymbol{\theta}_0) & \mathbf{I}_{\rho\zeta}'(\boldsymbol{\theta}_0) & \mathbf{I}_{\zeta\zeta}(\boldsymbol{\theta}_0) \end{array}\right] = -\mathbf{J}(\boldsymbol{\theta}_0) \tag{19}$$

where $\mathbf{I}_{\omega\omega}(\boldsymbol{\theta}_0)$, $\mathbf{I}_{\omega\rho}(\boldsymbol{\theta}_0)$ and $\mathbf{I}_{\rho\rho}(\boldsymbol{\theta}_0)$ are defined in Nakatani & Teräsvirta (2009). Let $\nabla \mathbf{D}_t = \partial \mathsf{vec}(\mathbf{D}_t)/\partial \boldsymbol{\omega}$, $\nabla \mathbf{P} = \partial \mathsf{vec}(\mathbf{P})/\partial \boldsymbol{\rho}$ and $\nabla \mathbf{G}_t = \partial \mathsf{vec}(\mathbf{G}_t)/\partial \boldsymbol{\zeta}$. The following lemma gives the remaining second partial derivatives of the log-likelihood function (10) for observation t and their conditional expectations $\mathbf{J}_{\zeta\omega t}(\boldsymbol{\theta})$, $\mathbf{J}_{\zeta\rho t}(\boldsymbol{\theta})$ and $\mathbf{J}_{\zeta\zeta t}(\boldsymbol{\theta})$ under the null hypothesis $\mathbf{G}_t = \mathbf{I}$.

Lemma 2 *The second partial derivatives* $\partial^2 l_t(\boldsymbol{\theta})/\partial\boldsymbol{\omega}\partial\boldsymbol{\zeta}'$, $\partial^2 l_t(\boldsymbol{\theta})/\partial\boldsymbol{\rho}\partial\boldsymbol{\zeta}'$ *and* $\partial^2 l_t(\boldsymbol{\theta})/\partial\boldsymbol{\zeta}\partial\boldsymbol{\zeta}'$ *of the log-likelihood function (10) for observation* t *and their conditional expectations* $\mathbf{J}_{\zeta\omega t}(\boldsymbol{\theta})$, $\mathbf{J}_{\zeta\rho t}(\boldsymbol{\theta})$ *and* $\mathbf{J}_{\zeta\zeta t}(\boldsymbol{\theta})$ *assuming* $\mathbf{G}_t = \mathbf{I}$ *are as follows:*

$$\begin{aligned}\frac{\partial^2 l_t(\boldsymbol{\theta})}{\partial\boldsymbol{\omega}\partial\boldsymbol{\zeta}'} &= -\frac{1}{2}\nabla\mathbf{D}_t(\mathbf{D}_t\mathbf{u}_t\mathbf{u}_t'\mathbf{P}^{-1}\mathbf{D}_t^{-1}\otimes\mathbf{G}_t^{-1}\mathbf{D}_t^{-1}+\mathbf{D}_t\mathbf{u}_t\mathbf{u}_t'\otimes\mathbf{G}_t^{-1}\mathbf{M}_t^{-1}\\ &\quad +\mathbf{G}_t^{-1}\mathbf{M}_t^{-1}\otimes\mathbf{D}_t\mathbf{u}_t\mathbf{u}_t'+\mathbf{G}_t^{-1}\mathbf{D}_t^{-1}\otimes\mathbf{D}_t\mathbf{u}_t\mathbf{u}_t'\mathbf{P}^{-1}\mathbf{D}_t^{-1})\nabla\mathbf{P}' \end{aligned} \tag{20}$$

$$\begin{aligned}\frac{\partial^2 l_t(\boldsymbol{\theta})}{\partial\boldsymbol{\rho}\partial\boldsymbol{\zeta}'} &= -\frac{1}{2}\nabla\mathbf{P}\left\{\mathbf{D}_t\mathbf{u}_t\mathbf{u}_t'\mathbf{P}^{-1}\otimes\mathbf{G}_t^{-1}\mathbf{D}_t^{-1}\mathbf{P}^{-1}\right.\\ &\quad \left.+\mathbf{G}_t^{-1}\mathbf{D}_t^{-1}\mathbf{P}^{-1}\otimes\mathbf{D}_t\mathbf{u}_t\mathbf{u}_t'\mathbf{P}^{-1}\right\}\nabla\mathbf{G}_t' \end{aligned} \tag{21}$$

and

$$\begin{aligned}\frac{\partial^2 l_t(\boldsymbol{\theta})}{\partial\boldsymbol{\zeta}\partial\boldsymbol{\zeta}'} &= -\left\{\left[(\mathsf{vec}(\mathbf{G}_t^{-1})'\otimes\mathbf{I})\right]-\frac{1}{2}\left[\mathsf{vec}(\mathbf{G}_t^{-1}\mathbf{M}_t^{-1}\mathbf{D}_t\mathbf{u}_t\mathbf{u}_t'\mathbf{D}_t)'\otimes\mathbf{I}\right]\right.\\ &\quad \left.-\frac{1}{2}\left[\mathsf{vec}(\mathbf{G}_t^{-1}\mathbf{M}_t^{-1}\mathbf{D}_t\mathbf{u}_t\mathbf{u}_t'\mathbf{D}_t)'\otimes\mathbf{I}\right]\right\}\frac{\partial^2\mathsf{vec}(\mathbf{G}_t)'}{\partial\boldsymbol{\zeta}\partial\boldsymbol{\zeta}'}\\ &\quad +\frac{1}{2}\nabla\mathbf{G}_t\left\{2(\mathbf{G}_t^{-1}\otimes\mathbf{G}_t^{-1})-\mathbf{D}_t\mathbf{u}_t\mathbf{u}_t'\mathbf{D}_t\otimes\mathbf{H}_t^{-1}-\mathbf{G}_t^{-1}\otimes\mathbf{G}_t^{-1}\mathbf{M}_t^{-1}\mathbf{D}_t\mathbf{u}_t\mathbf{u}_t'\mathbf{D}_t\right.\\ &\quad -\mathbf{D}_t\mathbf{u}_t\mathbf{u}_t'\mathbf{D}_t\mathbf{M}_t^{-1}\mathbf{G}_t^{-1}\otimes\mathbf{G}_t^{-1}-\mathbf{G}_t^{-1}\otimes\mathbf{D}_t\mathbf{u}_t\mathbf{u}_t'\mathbf{D}_t\mathbf{M}_t^{-1}\mathbf{G}_t^{-1}\\ &\quad \left.-\mathbf{G}_t^{-1}\mathbf{M}_t^{-1}\mathbf{D}_t\mathbf{u}_t\mathbf{u}_t'\mathbf{D}_t\otimes\mathbf{G}_t^{-1}-\mathbf{H}_t^{-1}\otimes\mathbf{D}_t\mathbf{u}_t\mathbf{u}_t'\mathbf{D}_t\right\}\nabla\mathbf{G}_t'. \end{aligned} \tag{22}$$

Their conditional expectations noting that $\mathsf{E}\mathbf{u}_t\mathbf{u}_t' = \mathbf{P}$ *and assuming* $\mathbf{G}_t = \mathbf{I}$ *in (20)-(22) are*

$$\mathbf{J}_{\omega\zeta t} = \frac{1}{2}\nabla\mathbf{D}_t\left\{\mathbf{I}\otimes\mathbf{D}_t^{-1}+\mathbf{P}\mathbf{D}_t\otimes\mathbf{M}_t^{-1}+\mathbf{M}_t^{-1}\otimes\mathbf{P}\mathbf{D}_t+\mathbf{D}_t^{-1}\otimes\mathbf{I}\right\}\nabla\mathbf{G}_t' \tag{23}$$

$$\mathbf{J}_{\rho\zeta t} = \frac{1}{2}\nabla\mathbf{P}\left\{\mathbf{D}_t\otimes\mathbf{P}^{-1}\mathbf{D}_t^{-1}+\mathbf{P}^{-1}\mathbf{D}_t^{-1}\otimes\mathbf{D}_t\right\}\nabla\mathbf{G}_t' \tag{24}$$

$$\mathbf{J}_{\zeta\zeta t} = \frac{1}{2}\nabla\mathbf{G}_t\left\{2(\mathbf{I}\otimes\mathbf{I})+\mathbf{M}_t\otimes\mathbf{M}_t^{-1}+\mathbf{M}_t^{-1}\otimes\mathbf{M}_t\right\}\nabla\mathbf{G}_t'. \tag{25}$$

Proof. See the Appendix.

4 The LM test statistic

When Assumptions 1-4 hold the asymptotic null distribution of the maximum likelihood estimator $\widetilde{\boldsymbol{\theta}}_T$ is given by

$$\sqrt{T}(\widetilde{\boldsymbol{\theta}}_T-\boldsymbol{\theta}_0)\xrightarrow{D}\mathcal{N}(\mathbf{0},\mathbf{J}^{-1}(\boldsymbol{\theta}_0)\mathbf{I}(\boldsymbol{\theta}_0)\mathbf{J}^{-1}(\boldsymbol{\theta}_0))$$

see Ling & McAleer (2003). If $\mathbf{z}_t \sim \text{iid}\mathcal{N}(\mathbf{0},\mathbf{P})$**,** the information matrix $\mathbf{I}(\boldsymbol{\theta}_0) = -\mathsf{E}\mathbf{J}(\boldsymbol{\theta}_0)$ and the asymptotic covariance matrix reduces to $\mathbf{I}^{-1}(\boldsymbol{\theta}_0)$. Ling & McAleer (2003) showed that under H_0, $\mathbf{I}(\boldsymbol{\theta}_0)$ and $\mathbf{J}(\boldsymbol{\theta}_0)$ can be consistently estimated by

$$\mathbf{I}(\widetilde{\boldsymbol{\theta}}_T) = \frac{1}{T}\sum_{t=1}^{T}\mathbf{q}_t(\widetilde{\boldsymbol{\theta}})\mathbf{q}_t(\widetilde{\boldsymbol{\theta}})' \tag{26}$$

where $\mathbf{q}_t(\widetilde{\boldsymbol{\theta}})$ is $\mathbf{q}_t(\boldsymbol{\theta})$ evaluated at $\boldsymbol{\theta} = \widetilde{\boldsymbol{\theta}}$, and

$$\mathbf{J}(\widetilde{\boldsymbol{\theta}}_T) = -\frac{1}{T}\sum_{t=1}^{T}\frac{\partial^2 l_t(\widetilde{\boldsymbol{\theta}}_T)}{\partial\boldsymbol{\theta}\partial\boldsymbol{\theta}'} \tag{27}$$

respectively. See also Nakatani & Teräsvirta (2009).

Let $\widetilde{\boldsymbol{\theta}} = (\widetilde{\boldsymbol{\omega}}', \widetilde{\boldsymbol{\rho}}', \widetilde{\boldsymbol{\zeta}}')'$ be the QML estimator of $\boldsymbol{\theta}_0$ under the null hypothesis. The average score evaluated at $\widetilde{\boldsymbol{\theta}}$ equals

$$\mathbf{q}(\widetilde{\boldsymbol{\theta}}) = (\mathbf{q}'_\omega(\widetilde{\boldsymbol{\theta}}), \mathbf{q}'_\rho(\widetilde{\boldsymbol{\theta}}), \mathbf{q}'_\zeta(\widetilde{\boldsymbol{\theta}}))' = (\mathbf{0}'_{3m}, \mathbf{0}'_{m(m-1)/2}, \mathbf{q}'_{\boldsymbol{\zeta}}(\widetilde{\boldsymbol{\theta}}))' \tag{28}$$

where

$$\mathbf{q}_{\boldsymbol{\zeta}}(\widetilde{\boldsymbol{\theta}}) = -\frac{1}{T}\sum_{t=1}^{T}\{\nabla\mathbf{G}_t \mathsf{vec}(\mathbf{I} - \frac{1}{2}\mathbf{M}_t^{-1}\boldsymbol{\varepsilon}_t\boldsymbol{\varepsilon}'_t - \frac{1}{2}\boldsymbol{\varepsilon}_t\boldsymbol{\varepsilon}'_t\mathbf{M}_t^{-1})\} \tag{29}$$

is the relevant (nonzero) block in the LM test statistic. The corresponding block of the information matrix evaluated under H_0 equals

$$\mathbf{I}^{\zeta\zeta}(\widetilde{\boldsymbol{\theta}}_T) = \{\mathbf{I}_{\zeta\zeta}(\widetilde{\boldsymbol{\theta}}_T) - \mathbf{I}_{\zeta\cdot}(\widetilde{\boldsymbol{\theta}}_T)\mathbf{I}_{\cdot\cdot}^{-1}(\widetilde{\boldsymbol{\theta}}_T)\mathbf{I}_{\cdot\zeta}(\widetilde{\boldsymbol{\theta}}_T)\}^{-1}$$

where

$$\mathbf{I}_{\zeta\cdot}(\widetilde{\boldsymbol{\theta}}_T) = \begin{bmatrix} \mathbf{I}_{\zeta\omega}(\widetilde{\boldsymbol{\theta}}_T) & \mathbf{I}_{\zeta\rho}(\widetilde{\boldsymbol{\theta}}_T) \end{bmatrix} = \mathbf{I}'_{\cdot\zeta}(\widetilde{\boldsymbol{\theta}}_T)$$

and

$$\mathbf{I}_{\cdot\cdot}(\widetilde{\boldsymbol{\theta}}_T) = \begin{bmatrix} \mathbf{I}_{\omega\omega}(\widetilde{\boldsymbol{\theta}}_T) & \mathbf{I}_{\omega\rho}(\widetilde{\boldsymbol{\theta}}_T) \\ \mathbf{I}_{\rho\omega}(\widetilde{\boldsymbol{\theta}}_T) & \mathbf{I}_{\rho\rho}(\widetilde{\boldsymbol{\theta}}_T) \end{bmatrix}.$$

We now state our main result:

Theorem 1 *(the LM test statistic) Assume that* $\mathbf{z}_t \sim iid(\mathbf{0}, \mathbf{P})$ *and that Assumptions 1-4 hold. Under* H_0: $\boldsymbol{\zeta} = \mathbf{0}$ *or* $\mathbf{G}_t = \mathbf{I}$, *the LM statistic*

$$LM_\zeta = T\mathbf{q}(\widetilde{\boldsymbol{\theta}})'\mathbf{I}^{-1}(\widetilde{\boldsymbol{\theta}}_T)\mathbf{q}(\widetilde{\boldsymbol{\theta}}_T) = T\mathbf{q}_\zeta(\widetilde{\boldsymbol{\theta}})'\mathbf{I}^{\zeta\zeta}(\widetilde{\boldsymbol{\theta}}_T)\mathbf{q}_{\boldsymbol{\zeta}}(\widetilde{\boldsymbol{\theta}}_T) \tag{30}$$

where $\widetilde{\boldsymbol{\theta}}$ *is a consistent estimator of* $\boldsymbol{\theta}_0$ *under* H_0, *has an asymptotic* χ^2 *distribution with* mr *degrees of freedom. When* $\mathbf{I}(\boldsymbol{\theta}_0) = -\mathsf{E}\mathbf{J}(\boldsymbol{\theta}_0)$, $\mathbf{I}^{\zeta\zeta}(\widetilde{\boldsymbol{\theta}}_T)$ *can be replaced by the corresponding block of* $-\mathbf{J}^{-1}(\widetilde{\boldsymbol{\theta}}_T)$.

If $\mathbf{D}_t = \mathrm{diag}(\alpha_{01}, ..., \alpha_{0m})$, LM_ζ becomes a test of no conditional heteroskedasticity against CCC-ARCH. This test statistic is a special case of the constant error covariance matrix test derived by Eklund & Teräsvirta (2007).

Corollary 2 *Under the sequence of local alternatives* H_1: $\boldsymbol{\zeta}_{0T} = \boldsymbol{\delta}/T^{1/2}$, *where* $\boldsymbol{\delta} = (\mathbf{0}', \mathbf{0}', \boldsymbol{\delta}'_\zeta)'$, *the LM statistic (30) has an asymptotic noncentral* χ^2 *distribution with* mr *degrees of freedom and noncentrality parameter*

$$\lambda = \boldsymbol{\delta}'_\zeta\mathbf{I}^{\zeta\zeta}(\boldsymbol{\theta}_0)\boldsymbol{\delta}_\zeta$$

where $\boldsymbol{\delta}_\zeta$ *is a fixed* mr*-vector.*

Proof. See the Appendix.

If $\boldsymbol{\delta}_\zeta = \delta_\zeta\mathbf{1}_{mr}$, where the mr-vector $\mathbf{1}_{mr} = (1, ..., 1)'$ and $\delta_\zeta \neq 0$ is a scalar, all parameters under the alternative deviate from 0 equally much at each T, and the noncentrality parameter simplifies to $\lambda = \delta_\zeta^2\mathbf{1}'_{mr}\mathbf{I}^{\zeta\zeta}(\boldsymbol{\theta}_0)\mathbf{1}_{mr}$.

5 Bivariate illustration

In this section we discuss the bivariate case, $m = 2$. Then $\boldsymbol{\omega} = (\boldsymbol{\omega}_1', \boldsymbol{\omega}_2')'$, where $\boldsymbol{\omega}_i = (\alpha_{i0}, \alpha_{i1}, \beta_{i1})'$ for $i = 1, 2$, $\boldsymbol{\zeta} = (\boldsymbol{\zeta}_1', \boldsymbol{\zeta}_2')'$, and

$$h_{it} = \alpha_{i0} + \alpha_{i1}\varepsilon_{i,t-1}^2 + \beta_{i1}h_{i,t-1},\ i = 1, 2.$$

The block of the score vector corresponding to the parameter $\boldsymbol{\zeta}$ in Lemma 1 becomes

$$\mathbf{q}_{\boldsymbol{\zeta}}(\widetilde{\boldsymbol{\theta}}) = -\frac{1}{T}\sum_{t=1}^{T}\begin{bmatrix} \widetilde{\mathbf{v}}_{11t}^0 \left(1 - \frac{1}{(1-\rho^2)h_{1t}h_{2t}}\right)\varepsilon_{1t}^2 h_{2t} - \rho\varepsilon_{1t}\varepsilon_{2t}\sqrt{h_{1t}h_{2t}} \\ \widetilde{\mathbf{v}}_{22t}^0 \left(1 - \frac{1}{(1-\rho^2)h_{1t}h_{2t}}\right)\varepsilon_{2t}^2 h_{1t} - \rho\varepsilon_{1t}\varepsilon_{2t}\sqrt{h_{1t}h_{2t}} \end{bmatrix}$$

where $\widetilde{\mathbf{v}}_{ijt}^0 = \partial\sqrt{g_{it}}/\partial\boldsymbol{\zeta}_j = (1/2)\widetilde{\mathbf{z}}_{jt}^{(2)}$ estimated under H_0 and $\widetilde{\mathbf{z}}_{jt}^{(2)} = (\widetilde{z}_{jt-1}^2, ..., \widetilde{z}_{jt-r}^2)'$ for $i, j = 1, 2$, and ρ is the conditional correlation between ε_{1t} and ε_{2t}.

The block of the maximum likelihood estimated information matrix corresponding to $\boldsymbol{\zeta}$ in Theorem 1 equals

$$\mathbf{J}_{\zeta\zeta}(\widetilde{\boldsymbol{\theta}}) = \tilde{\mathbf{J}}_{\zeta\zeta} - \begin{bmatrix} \tilde{\mathbf{J}}_{\zeta\omega} & \tilde{\mathbf{J}}_{\zeta\rho} \end{bmatrix}_{2r\times 7}\begin{bmatrix} \tilde{\mathbf{J}}_{\omega\omega} & \tilde{\mathbf{J}}_{\omega\rho} \\ \tilde{\mathbf{J}}_{\rho\omega} & \tilde{\mathbf{J}}_{\rho\rho} \end{bmatrix}_{7\times 7}^{-1}\begin{bmatrix} \tilde{\mathbf{J}}_{\omega\zeta} \\ \tilde{\mathbf{J}}_{\rho\zeta} \end{bmatrix}_{7\times 2r}$$

where

$$\begin{bmatrix} \tilde{\mathbf{J}}_{\omega\omega} & \tilde{\mathbf{J}}_{\omega\rho} \\ \tilde{\mathbf{J}}_{\rho\omega} & \tilde{\mathbf{J}}_{\rho\rho} \end{bmatrix} = \frac{1}{4T}\sum_{t=1}^{T}\begin{bmatrix} \left(1+\frac{1}{1-\rho^2}\right)\widetilde{\mathbf{k}}_{\omega\omega t}\widetilde{\mathbf{k}}_{\omega\omega t}' & -\frac{\rho^2}{1-\rho^2}\widetilde{\mathbf{k}}_{\omega\omega t}\widetilde{\mathbf{k}}_{\rho\rho t}' & -\frac{2\rho}{1-\rho^2}\widetilde{\mathbf{k}}_{\omega\omega t} \\ -\frac{\rho^2}{1-\rho^2}\widetilde{\mathbf{k}}_{\rho\rho t}\widetilde{\mathbf{k}}_{\omega\omega t}' & \left(1+\frac{1}{1-\rho^2}\right)\widetilde{\mathbf{k}}_{\rho\rho t}\widetilde{\mathbf{k}}_{\rho\rho t}' & -\frac{2\rho}{1-\rho^2}\widetilde{\mathbf{k}}_{\rho\rho t} \\ -\frac{2\rho}{1-\rho^2}\widetilde{\mathbf{k}}_{\omega\omega t}' & -\frac{2\rho}{1-\rho^2}\widetilde{\mathbf{k}}_{\rho\rho t}' & \frac{4(1+\rho^2)}{(1-\rho^2)^2} \end{bmatrix} \quad (31)$$

$$\begin{bmatrix} \tilde{\mathbf{J}}_{\omega\zeta} \\ \tilde{\mathbf{J}}_{\rho\zeta} \end{bmatrix} = \frac{1}{2T}\sum_{t=1}^{T}\begin{bmatrix} \left(1+\frac{1}{1-\rho^2}\right)\widetilde{\mathbf{k}}_{\omega\omega t}\widetilde{\mathbf{v}}_{11t}^{0\prime} & -\frac{\rho^2}{1-\rho^2}\widetilde{\mathbf{k}}_{\omega\omega t}'\widetilde{\mathbf{v}}_{22t}^{0\prime} \\ -\frac{\rho^2}{1-\rho^2}\widetilde{\mathbf{k}}_{\rho\rho t}\widetilde{\mathbf{v}}_{11t}^{0\prime} & \left(1+\frac{1}{1-\rho^2}\right)\widetilde{\mathbf{k}}_{\rho\rho t}\widetilde{\mathbf{v}}_{22t}^{0\prime} \\ -\frac{2\rho}{1-\rho^2}\widetilde{\mathbf{v}}_{11t}^{0} & -\frac{2\rho}{1-\rho^2}\widetilde{\mathbf{v}}_{22t}^{0} \end{bmatrix} \quad (32)$$

and

$$\tilde{\mathbf{J}}_{\zeta\zeta} = \frac{1}{T}\sum_{t=1}^{T}\begin{bmatrix} \left(1+\frac{1}{1-\rho^2}\right)\widetilde{\mathbf{v}}_{11t}^{0}\widetilde{\mathbf{v}}_{11t}^{0\prime} & -\frac{\rho^2}{1-\rho^2}\widetilde{\mathbf{v}}_{\omega\omega t}^{0}\widetilde{\mathbf{v}}_{22t}^{0\prime} \\ -\frac{\rho^2}{1-\rho^2}\widetilde{\mathbf{v}}_{22t}^{0}\widetilde{\mathbf{v}}_{11t}^{0\prime} & \left(1+\frac{1}{1-\rho^2}\right)\widetilde{\mathbf{v}}_{22t}^{0}\widetilde{\mathbf{v}}_{22t}^{0\prime} \end{bmatrix}.$$

In (31) and (32) $\widetilde{\mathbf{k}}_{ijt} = \widetilde{h}_{it}^{-1}\partial\widetilde{h}_{it}/\partial\boldsymbol{\omega}_j$ estimated under H_0, where $\partial\widetilde{h}_{it}/\partial\boldsymbol{\omega}_j = \widetilde{\mathbf{x}}_{jt} + \widetilde{\beta}_{ii}\partial\widetilde{h}_{it-1}/\partial\boldsymbol{\omega}_j$, and $\widetilde{\mathbf{x}}_{jt} = (1, \varepsilon_{jt-1}^2, \widetilde{h}_{jt-1})'$ for $i, j = 1, 2$. Furthermore $\widetilde{h}_{it}$ is h_{it} estimated under H_0. Following the suggestion by Fiorentini, Calzolari & Panattoni (1996), we use the following initial values for the recursions:

$$\widetilde{\mathbf{x}}_{j0} = (1, \frac{1}{T}\sum_{t=1}^{T}\varepsilon_{jt}^2, \frac{1}{T}\sum_{t=1}^{T}\varepsilon_{jt}^2)'$$

and $\partial\widetilde{h}_{i0}/\partial\boldsymbol{\omega}_j = \mathbf{0}$.

Under H_0, the LM test statistic (30) has an asymptotic χ^2 distribution with $2r$ degrees of freedom.

6 A portmanteau test and a comparison

Ling & Li (1997) introduced a portmanteau test for testing the adequacy of the multivariate GARCH(p,q) model. They defined $\boldsymbol{\varepsilon}_t = \mathbf{V}_t^{1/2}\mathbf{z}_t$,where $\mathbf{V}_t$ is the conditional covariance matrix of $\boldsymbol{\varepsilon}_t$ and $\{\mathbf{z}_t\}$ $\sim$iid$(\mathbf{0}, \mathbf{I}_m)$,where $\mathbf{I}_m$ is an $m \times m$ identity matrix instead of a positive definite correlation matrix $\mathbf{P}$ in our model and m is the dimension of $\boldsymbol{\varepsilon}_t = (\varepsilon_{1t}, ..., \varepsilon_{mt})'$. Let

$$R_j = \mathsf{E}(\boldsymbol{\varepsilon}_t'\mathbf{V}_t^{-1}\boldsymbol{\varepsilon}_t - m)(\boldsymbol{\varepsilon}_{t-j}'\mathbf{V}_{t-j}^{-1}\boldsymbol{\varepsilon}_{t-j} - m),\ j = 0, 1, ..., r$$

be the jth autocovariance of $\boldsymbol{\varepsilon}_t'\mathbf{V}_t^{-1}\boldsymbol{\varepsilon}_t$, and set $\mathbf{R} = (R_1/R_0, ..., R_r/R_0)'$. The null hypothesis to be tested is $\mathbf{R} = \mathbf{0}$. The corresponding consistent estimators are

$$\widetilde{R}_j = \frac{1}{T}\sum_{r=j+1}^{T}(\widetilde{\boldsymbol{\varepsilon}}_t'\widetilde{\mathbf{V}}_t^{-1}\widetilde{\boldsymbol{\varepsilon}}_t - \widetilde{m})(\widetilde{\boldsymbol{\varepsilon}}_{t-j}'\widetilde{\mathbf{V}}_{t-j}^{-1}\widetilde{\boldsymbol{\varepsilon}}_{t-j} - \widetilde{m}),\ j = 0, 1, ..., r$$

where $\widetilde{m} = (1/T)\sum_{r=j+1}^{T}\widetilde{\boldsymbol{\varepsilon}}_t'\widetilde{\mathbf{V}}_t^{-1}\widetilde{\boldsymbol{\varepsilon}}_t$. Set $\widetilde{\mathbf{R}} = (\widetilde{R}_1/\widetilde{R}_0, ..., \widetilde{R}_r/\widetilde{R}_0)'$. Under the standard regularity conditions, including $\widetilde{R}_0 \xrightarrow{p} \kappa < \infty$, Ling & Li (1997) showed that under the null hypothesis,

$$\sqrt{T}\widetilde{\mathbf{R}}\xrightarrow{d}\mathcal{N}(\mathbf{0}, \mathbf{\Omega}).$$

It then follows that the portmanteau test statistic

$$\mathbf{Q}(r) = T\widetilde{\mathbf{R}}'\widetilde{\mathbf{\Omega}}^{-1}\widetilde{\mathbf{R}} \tag{33}$$

where $\widetilde{\mathbf{\Omega}}$ is a consistent plug-in estimator of $\mathbf{\Omega}$, has an asymptotic $\chi^2(r)$-distribution under $\mathbf{R} = \mathbf{0}$.

In order to better understand the difference between our test and that of Ling and Li, we shall show that the latter is also an LM test. To this end, define R_j using $\boldsymbol{\varepsilon}_t = \mathbf{D}_t\mathbf{z}_t$ and $\mathbf{V}_t^{-1} = \mathbf{D}_t^{-1}\mathbf{P}^{-1}\mathbf{D}_t^{-1}$, which gives $\boldsymbol{\varepsilon}_t'\mathbf{V}_t^{-1}\boldsymbol{\varepsilon}_t = \mathbf{z}_t'\mathbf{P}^{-1}\mathbf{z}_t$. Then

$$\frac{R_j}{R_0} = \frac{\mathsf{E}(\mathbf{z}_t'\mathbf{P}^{-1}\mathbf{z}_t - m)(\mathbf{z}_{t-j}'\mathbf{P}^{-1}\mathbf{z}_{t-j} - m)}{\mathsf{E}(\mathbf{z}_t'\mathbf{P}^{-1}\mathbf{z}_t - m)^2}.$$

The null hypothesis is unchanged: $\mathbf{R} = \mathbf{0}$.

Consider the following building-block of the Ling and Li statistic:

$$\sqrt{T}\widetilde{\mathbf{R}} = \sqrt{T}(\frac{\widetilde{R}_1}{\widetilde{R}_0}, ..., \frac{\widetilde{R}_r}{\widetilde{R}_0})'$$

where

$$\frac{\widetilde{R}_j}{\widetilde{R}_0} = \frac{(1/T)\sum_{t=r+1}^{T}(\widetilde{\mathbf{z}}_t'\widetilde{\mathbf{P}}^{-1}\widetilde{\mathbf{z}}_t - \widetilde{m})(\widetilde{\mathbf{z}}_{t-j}'\widetilde{\mathbf{P}}^{-1}\widetilde{\mathbf{z}}_{t-j} - \widetilde{m})}{(1/T)\sum_{t=1}^{T}(\widetilde{\mathbf{z}}_t'\widetilde{\mathbf{P}}^{-1}\widetilde{\mathbf{z}}_t - \widetilde{m})^2}.$$

with $\widetilde{m} = (1/T)\sum_{t=r+1}^{T}\widetilde{\mathbf{z}}_t'\widetilde{\mathbf{P}}^{-1}\widetilde{\mathbf{z}}_t$. When $\mathbf{z}_t \sim \text{iid}(\mathbf{0}, \mathbf{P})$, adapting the corresponding result in Ling & Li (1997) to our special case gives

$$(1/T)\sum_{t=1}^{T}(\widetilde{\mathbf{z}}_t'\widetilde{\mathbf{P}}^{-1}\widetilde{\mathbf{z}}_t - \widetilde{m})^2 \xrightarrow{p} \mathsf{E}(\mathbf{z}_t'\mathbf{P}^{-1}\mathbf{z}_t - m)^2 = \kappa. \tag{34}$$

When $\mathbf{z}_t \sim \text{iid}\mathcal{N}(\mathbf{0}, \mathbf{P})$, $\kappa = 2m$.

Now consider the following one-dimensional linear combination of lags of $\mathbf{z}_t'\mathbf{P}^{-1}\mathbf{z}_t$:

$$g_t^* = 1 + \sum_{j=1}^{r} \zeta_j^* \frac{\mathbf{z}_{t-j}'\mathbf{P}^{-1}\mathbf{z}_{t-j} - m}{\kappa} \tag{35}$$

and define $\mathbf{G}_t^* = g_t^* \mathbf{I}_m$. We argue that the LM test for testing the null hypothesis $\boldsymbol{\zeta}^* = (\zeta_1^*, ..., \zeta_r^*)' = \mathbf{0}$ in (35) is asymptotically equivalent to Ling and Li's test adapted to the CCC-GARCH framework.

To show this, define

$$\nabla \mathbf{G}_t^* = \frac{\partial \text{vec}(\mathbf{G}_t^*)'}{\partial \boldsymbol{\zeta}^*} = \frac{1}{\kappa}(\mathbf{z}_{t-1}'\mathbf{P}^{-1}\mathbf{z}_{t-1} - m, ..., \mathbf{z}_{t-r}'\mathbf{P}^{-1}\mathbf{z}_{t-r} - m)'\text{vec}(\mathbf{I})' = \frac{\mathbf{w}_{t-1}}{\kappa}\text{vec}(\mathbf{I})'$$

The average score vector evaluated at H_0 equals

$$\begin{aligned}
\overline{\mathbf{q}}_{\boldsymbol{\zeta}}(\widetilde{\boldsymbol{\zeta}}^*) &= -\frac{1}{\kappa T}\sum_{t=r+1}^{T} \widetilde{\mathbf{w}}_{t-1}\text{vec}(\mathbf{I})'\text{vec}(\mathbf{I} - (1/2)\widetilde{\mathbf{D}}_t^{-1}\widetilde{\mathbf{P}}^{-1}\widetilde{\mathbf{z}}_t\widetilde{\mathbf{z}}_t'\widetilde{\mathbf{D}}_t \\
&\quad -(1/2)\widetilde{\mathbf{D}}_t\widetilde{\mathbf{z}}_t\widetilde{\mathbf{z}}_t'\widetilde{\mathbf{P}}^{-1}\widetilde{\mathbf{D}}_t^{-1}) \\
&= \frac{1}{\kappa T}\sum_{t=r+1}^{T} \widetilde{\mathbf{w}}_{t-1}(\widetilde{\mathbf{z}}_t'\widetilde{\mathbf{P}}^{-1}\widetilde{\mathbf{z}}_t - m)
\end{aligned}$$

where $\widetilde{\mathbf{w}}_{t-1} = (\widetilde{w}_{t-1}, ..., \widetilde{w}_{t-r})'$ with $\widetilde{w}_{t-j} = \widetilde{\mathbf{z}}_{t-j}'\widetilde{\mathbf{P}}^{-1}\widetilde{\mathbf{z}}_{t-j} - m$, $j = 1, ..., r$. Now, Ling & Li (1997) showed that under the null hypothesis, $(1/T)\sum_{r=j+1}^{T} \widetilde{\boldsymbol{\varepsilon}}_t'\widetilde{\mathbf{V}}_t^{-1}\widetilde{\boldsymbol{\varepsilon}}_t \xrightarrow{p} \mathsf{E}\boldsymbol{\varepsilon}_t'\mathbf{V}_t^{-1}\boldsymbol{\varepsilon}_t$, as $T \to \infty$. In our special case, this implies $\widetilde{m} \xrightarrow{p} m$. Furthermore, they established that $(1/T)\sum_{r=j+1}^{T}(\widetilde{\mathbf{z}}_t'\widetilde{\mathbf{P}}_t^{-1}\widetilde{\mathbf{z}}_t - m)^2 \xrightarrow{p} \kappa$, as $T \to \infty$, which allowed them to focus on the asymptotic distribution of

$$\mathbf{R}^* = \frac{1}{T}\sum_{t=r+1}^{T} \widetilde{\mathbf{w}}_{t-1}(\widetilde{\mathbf{z}}_t'\widetilde{\mathbf{P}}^{-1}\widetilde{\mathbf{z}}_t - m) = \kappa\overline{\mathbf{q}}_{\boldsymbol{\zeta}}(\widetilde{\boldsymbol{\zeta}}^*)$$

Random variables $\sqrt{T}\mathbf{R}^*$ and $\kappa\sqrt{T}\overline{\mathbf{q}}_{\boldsymbol{\zeta}}(\widetilde{\boldsymbol{\zeta}}^*)$ are thus identical and so have the same distribution. Since by $\widetilde{m} \xrightarrow{p} m$ and (34) $\sqrt{T}\mathbf{R}^*/\kappa$ has the same asymptotic distribution as $\sqrt{T}\widetilde{\mathbf{R}}$, one is able to conclude that $\sqrt{T}\widetilde{\mathbf{R}}$ and $\sqrt{T}\overline{\mathbf{q}}_{\boldsymbol{\zeta}}(\widetilde{\boldsymbol{\zeta}}^*)$ have the same asymptotic distribution under their null hypotheses. It follows that the test statistics of these hypotheses, $T\mathbf{q}_{\boldsymbol{\zeta}}(\widetilde{\boldsymbol{\zeta}}^*)'\boldsymbol{\Omega}_{\boldsymbol{\zeta}}^{-1}\mathbf{q}_{\boldsymbol{\zeta}}(\widetilde{\boldsymbol{\zeta}}^*)$, where $\boldsymbol{\Omega}_{\boldsymbol{\zeta}}$ is the asymptotic covariance matrix of $\sqrt{T}\mathbf{q}_{\boldsymbol{\zeta}}(\widetilde{\boldsymbol{\zeta}}^*)$, and $T\widetilde{\mathbf{R}}'\boldsymbol{\Omega}^{-1}\widetilde{\mathbf{R}}$, have the same asymptotic null distribution. This shows that if misspecification of the GARCH equations is characterised by a number of lags of $\mathbf{z}_t'\mathbf{P}^{-1}\mathbf{z}_t$ and assumed to be exactly the same for all m equations, the resulting LM-test is asymptotically equivalent to the test of Ling & Li (1997).

Our test may therefore viewed as one in which we relax the restrictions inherent in Ling and Li's test by letting the assumed misspecification vary from one equation to the next. It can also be seen as a multivariate extension of the LM test of no remaining ARCH in GARCH by Lundbergh & Teräsvirta (2002). They proved their test is asymptotically equivalent to the portmanteau test by Li & Mak (1994). When $m = 1$, our LM test and Ling and Li's portmanteau test collapse into the Lundbergh & Teräsvirta (2002) and the Li & Mak (1994) test, respectively. If $m = 1$ and the conditional variance is constant, Ling and Li's test reduces to the one by McLeod & Li (1983) and ours to the no ARCH test of Engle (1982).

7 A simulation study

We study the size and power properties of the test statistic LM_ζ by simulation. The power of LM_ζ is considered in situations in which the GARCH equations are misspecified and in situations in which the alternative is a model with time-varying correlations. Our test is constructed for situations in which the GARCH equations may be misspecified. Nevertheless, it is interesting to know whether it may also reveal misspecification in the conditional correlation structure. We compare the power of the test to the power of the portmanteau test of Ling & Li (1997), the LM-test of constant conditional correlations of Tse (2000) and the general test of correct specification of a parametric conditional covariance model by Long, Su & Ullah (2011).

Tse & Tsui (1999) study the power of Ling & Li's test in testing the adequacy of a multivariate model for conditional heteroskedasticity. They find that the test has low power in most cases where the conditional correlation structure of the true model differs from the estimated one.

The LM test of constant conditional correlations by Tse (2000), denoted LMC following the original article, is based on assuming time-varying correlations, defined as

$$\rho_{ijt} = \rho_{ij} + \delta_{ij}\varepsilon_{i,t-1}\varepsilon_{j,t-1}, \qquad 1 \leq i < j \leq m,$$

where δ_{ij} are additional parameters under the alternative hypothesis. The null hypothesis is H_0: $\delta_{ij} = 0$ for $1 \leq i < j \leq m$, and the test statistic is given by

$$LMC = \boldsymbol{\iota}_T' \widetilde{\mathbf{S}}(\widetilde{\mathbf{S}}'\widetilde{\mathbf{S}})^{-1}\widetilde{\mathbf{S}}'\boldsymbol{\iota}_T$$

where $\boldsymbol{\iota}_T$ is a $T \times 1$ vector of ones, $\widetilde{\mathbf{S}}$ is the $T \times m$ matrix of partial derivatives $\partial l_t/\partial\boldsymbol{\theta}'$ evaluated under H_0 and $\boldsymbol{\theta}$ is the vector of parameters in the model under the alternative hypothesis. Under H_0, LMC has an asymptotic χ^2-distribution with $m(m-1)/2$ degrees of freedom.

The general misspecification test of Long et al. (2011) is based on a semiparametric estimate of the covariance matrix

$$\widehat{\mathbf{H}}_{sp,t} = \widehat{\mathbf{H}}_{p,t}^{1/2}\widehat{\mathbf{G}}_{np,t}(\mathbf{x}_t)\widehat{\mathbf{H}}_{p,t}^{1/2},$$

where $\widehat{\mathbf{H}}_{p,t}$ is the parametric covariance matrix estimator, $\widehat{\mathbf{G}}_{np,t}(\mathbf{x}_t)$ is a nonparametric estimator of $\mathsf{E}\{\mathbf{z}_t\mathbf{z}_t'|\mathcal{F}_{t-1}\}$ and $\mathbf{x}_t = \boldsymbol{\varepsilon}_{t-1}$. Long et al. (2011) show that if the parametric model is correctly specified then $\mathbf{G}_{np,t}(\mathbf{x}_t) = \mathbf{I}_m$. The alternative to the null hypothesis H_0: $\mathbf{G}_{np,t}(\mathbf{x}_t) = \mathbf{I}_m$ is H_1: $\Pr(\mathbf{G}_{np,t}(\mathbf{x}_t) = \mathbf{I}_m) < 1$. The test statistic, $\mathsf{Sup}T_n^*$, is calculated using a Gaussian kernel and a suitable bandwidth and the p-values are obtained by a wild bootstrap procedure. For details see Long et al. (2011).

7.1 Size

The size of LM_ζ is simulated for five different CCC-GARCH$(1,1)$ models at sample sizes $T = 1000, 2500, 5000$ and 10000 and dimensions $m = 2$ and 5. The nominal size of the tests is 5%. The data are generated from the five bivariate CCC-GARCH$(1,1)$ models used in Nakatani & Teräsvirta (2009). The DGPs are reported in Table 1. DGP 1 has moderate persistence in volatility, while DGPs 2 and 3 represent models with high persistence and DGPs 4 and 5 models with low persistence in volatility. The correlation is low ($\rho = 0.3$) in DGPs 1, 3 and 5 and high ($\rho = 0.9$) in DGPs 2 and 4. All simulations have been performed in R (R Core Team (2013)) using the ccgarch package by Nakatani (2013).

We simulate both two- and five-dimensional models. When simulating the latter models the DGPs are extensions of the former models. For example, in the two-dimensional case

Table 1: Parameter values for the DGPs used in size simulations.

	DGP 1	DGP 2	DGP 3	DGP 4	DGP 5
$\mathbf{A}_1$	$\begin{bmatrix} 0.1 & 0 \\ 0 & 0.2 \end{bmatrix}$	$\begin{bmatrix} 0.04 & 0 \\ 0 & 0.05 \end{bmatrix}$	$\begin{bmatrix} 0.04 & 0 \\ 0 & 0.05 \end{bmatrix}$	$\begin{bmatrix} 0.1 & 0 \\ 0 & 0.2 \end{bmatrix}$	$\begin{bmatrix} 0.1 & 0 \\ 0 & 0.2 \end{bmatrix}$
$\mathbf{B}_1$	$\begin{bmatrix} 0.8 & 0 \\ 0 & 0.7 \end{bmatrix}$	$\begin{bmatrix} 0.95 & 0 \\ 0 & 0.9 \end{bmatrix}$	$\begin{bmatrix} 0.95 & 0 \\ 0 & 0.9 \end{bmatrix}$	$\begin{bmatrix} 0.45 & 0 \\ 0 & 0.6 \end{bmatrix}$	$\begin{bmatrix} 0.45 & 0 \\ 0 & 0.6 \end{bmatrix}$
ρ	0.3	0.9	0.3	0.9	0.3

DGP 1 $\mathbf{A} = \mathsf{diag}(0.1, 0.2)$ on the main diagonal, whereas in the five-dimensional model $\mathbf{A} = \mathsf{diag}(0.1, 0.2, 0.1, 0.2, 0.1)$. The five-dimensional conditional correlation matrices are of the form

$$\mathbf{P} = \begin{bmatrix} 1 & \rho & \rho^2 & \rho^3 & \rho^4 \\ \rho & 1 & \rho & \rho^2 & \rho^3 \\ \rho^2 & \rho & 1 & \rho & \rho^2 \\ \rho^3 & \rho^2 & \rho & 1 & \rho \\ \rho^4 & \rho^3 & \rho^2 & \rho & 1 \end{bmatrix}, \tag{36}$$

which is selected simply because it depends on a single parameter. There is no statistical theory behind this choice.

Table 2 summarises the results for $m = 2$. The test has a reasonable size already when $T = 1000$. The only exception is DGP 5 with $T = 1000$ and $r = 4$. Table 3 contains the results for $m = 5$. The test has good size properties even in this case.

Table 2: Empirical size of the LM test for testing the adequacy of the estimated CCC-GARCH model when $m = 2$. and $r = 1, 4$. The nominal significance level is 0.05.

T	DGP 1	DGP 2	DGP 3	DGP 4	DGP 5
			$r = 1$		
1000	0.045	0.050	0.043	0.052	0.047
2500	0.050	0.049	0.048	0.051	0.049
5000	0.052	0.049	0.051	0.052	0.047
10000	0.050	0.047	0.051	0.051	0.051
			$r = 4$		
1000	0.049	0.051	0.052	0.051	0.082
2500	0.048	0.052	0.048	0.051	0.050
5000	0.048	0.050	0.049	0.053	0.047
10000	0.052	0.052	0.053	0.052	0.049

Note: The number of replications equals 10000.

7.2 Power

We begin by considering the power of the test when a CCC-GARCH$(1, 1)$ model is fitted to the data while the data are generated by a CCC-ARCH(2) or a CCC-GARCH$(2, 1)$ process. We continue by studying the situation in which a CCC-GARCH$(1, 1)$ model is fitted to the

Table 3: Empirical size of the LM test for testing the adequacy of the estimated CCC-GARCH model when $m = 5$ and $r = 1, 4$. The nominal significance level is 0.05.

T	DGP 1	DGP 2	DGP 3	DGP 4	DGP 5
			$r = 1$		
1000	0.047	0.055	0.046	0.050	0.047
2500	0.049	0.052	0.040	0.045	0.051
5000	0.051	0.050	0.047	0.053	0.045
10000	0.049	0.052	0.048	0.051	0.043
			$r = 4$		
1000	0.045	0.054	0.052	0.061	0.115
2500	0.052	0.058	0.050	0.050	0.052
5000	0.051	0.058	0.048	0.053	0.043
10000	0.053	0.055	0.055	0.049	0.047

Note: The number of replications equals 5000.

data, but the misspecification is due to the fact that the true process is an MGARCH process with time-varying conditional correlations. We consider cases where the correlations follow the Dynamic Conditional Correlation (DCC) GARCH model of Engle (2002), the Smooth Transition Conditional Correlation (STCC) GARCH model of Silvennoinen & Teräsvirta (2009) and the Baba-Engle-Kraft-Kroner (BEKK) GARCH model, defined in Engle & Kroner (1995).

All simulations are again performed in R. As the empirical size of the our statistic is very close to the nominal 5% size and the simulations require plenty of CPU time we have used the asymptotic null distribution in calculating the power. All estimates of the power of the test statistics are rejection rates under the alternative. We use 5000 replications to simulate the power of the LM_ζ, $Q(r)$ and LMC tests. As $\mathsf{Sup}T_n^*$ takes computationally a long time to calculate we have only been able to simulate it for $m = 2$ and $T = 1000$ and 2500, where we use 500 replications for $T = 1000$ and 200 replications for $T = 2500$. Using the fast method proposed by Davidson & MacKinnon (2006) for simulating the power of bootstrap tests we only draw one wild bootstrap value for each replication.

Three different parameterisation are considered for the CCC-GARCH, one for the DCC- and the STCC-GARCH process and two for the BEKK-GARCH processes. The parameters of these models appear in Table 4.

Table 4: MGARCH-parameters used in the simulations

	DGP 1	DGP 2	DGP 3	DGP 4-5	DGP 6-8	DGP 9	DGP 10
$\mathbf{A}_1$	$0.3 \cdot \mathbf{I}_2$	$\begin{bmatrix} 0.07 & 0 \\ 0 & 0.06 \end{bmatrix}$	$\begin{bmatrix} 0.07 & 0 \\ 0 & 0.06 \end{bmatrix}$	$\begin{bmatrix} 0.1 & 0 \\ 0 & 0.2 \end{bmatrix}$	$\begin{bmatrix} 0.04 & 0 \\ 0 & 0.05 \end{bmatrix}$	$\begin{bmatrix} 0.6 & 0 \\ 0 & 0.3 \end{bmatrix}$	$\sqrt{0.08} \cdot \mathbf{I}_2$
$\mathbf{A}_2$	$0.3 \cdot \mathbf{I}_2$	$\begin{bmatrix} 0.1 & 0 \\ 0 & 0.08 \end{bmatrix}$	$\begin{bmatrix} 0.3 & 0 \\ 0 & 0.2 \end{bmatrix}$				
$\mathbf{B}_1$		$\begin{bmatrix} 0.8 & 0 \\ 0 & 0.85 \end{bmatrix}$	$\begin{bmatrix} 0.25 & 0 \\ 0 & 0.3 \end{bmatrix}$	$\begin{bmatrix} 0.8 & 0 \\ 0 & 0.7 \end{bmatrix}$	$\begin{bmatrix} 0.95 & 0 \\ 0 & 0.9 \end{bmatrix}$	$\begin{bmatrix} 0.3 & 0 \\ 0 & 0.95 \end{bmatrix}$	$\sqrt{0.9} \cdot \mathbf{I}_2$
$\mathbf{CC}'$						$\begin{bmatrix} 0.2 & 0.1 \\ 0.1 & 0.2 \end{bmatrix}$	$\begin{bmatrix} 0.2 & 0.1 \\ 0.1 & 0.2 \end{bmatrix}$

Note: $\mathbf{a}_0 = (0.1, 0.2)'$ in DGPs 1-9.

Table 5 presents the results when $m = 2$. In DGPs 1-3 the constant conditional correlation matrix

$$\mathbf{P} = \begin{bmatrix} 1 & \rho \\ \rho & 1 \end{bmatrix},$$

with $\rho = 0.3$ (DGP 1a-3a) and $\rho = 0.9$ (DGP 1b-3b). The power of LM_ζ is higher than the power of $Q(r)$ or LMC in all six cases. In addition $Q(r)$ outperforms LMC in most cases. $Q(r)$ has good power for DGP 3 and in large samples also for DGPs 1 and 2. LMC has rather low if any power at all sample sizes when $\rho = 0.3$. For LM_ζ and LMC there is an increase in power when the correlation changes from 0.3 to 0.9, while the power of $Q(r)$ in that case slightly decreases. In particular, when the conditional correlation is large, also LMC which is not designed to detect misspecification in GARCH equations, can have considerable power when the time series are sufficiently long. The semiparametric test does not seem useful when the GARCH equations are misspecified but the correlations are constant.

In the DCC-GARCH model (DGPs 4-5) the conditional correlation is generated by the following process:

$$\mathbf{Q}_t = (1 - a - b)\mathbf{P} + a\mathbf{z}_{t-1}\mathbf{z}'_{t-1} + b\mathbf{Q}_{t-1},$$

where a and b are the DCC-parameters and $\mathbf{P}$ is now the unconditional correlation matrix $\mathbf{P} = \{\rho_{ij}\}$. Furthermore, to produce valid correlation matrices $\mathbf{Q}_t$ is rescaled as follows:

$$\mathbf{P}_t = (\mathbf{I} \odot \mathbf{Q}_t)^{-1/2}\mathbf{Q}_t(\mathbf{I} \odot \mathbf{Q}_t)^{-1/2},$$

where $\odot$ is the Hadamard product. The values for the DCC-parameters are

$$\begin{aligned} &\text{DGP 4} \;:\; a = 0.09, b = 0.9 \qquad \text{and} \\ &\text{DGP 5} \;:\; a = 0.05, b = 0.9. \end{aligned}$$

In DGP 4 the persistence in the conditional correlation is very high, i.e. the conditional correlation can deviate substantially from its mean for long periods, whereas in DGP 5 the attraction towards the mean is stronger than in DGP 4. We consider two values for the unconditional correlation: $\rho = 0.3$ (DGP 4a and 5a) and $\rho = 0.9$ (DGP 4b and 5b). From Table 5 we can see that the power of LM_ζ more or less equals its size for all four DGPs at all sample sizes. This is noteworthy as it suggests that the LM test still works as a misspecification test for GARCH equations when the null model is a DCC- and not a CCC-GARCH model. Note, however, that the asymptotic null distribution of LM_ζ is derived under the assumption that the null model is a CCC-GARCH one, so the fact that the null model contains additional parameters is ignored when the test is applied to a DCC-GARCH model.

Interestingly, the power of $Q(r)$ considerably increases when the correlation increases from 0.3 to 0.9. It can be quite high when the persistence of the correlation is high as in DGP 4. As may be expected, LMC is the best performer, displaying strong power against both DGPs at all sample sizes. The semiparametric test does have power when the correlations are persistent (DGP 5), but this is not the case when they are less so (DGP 4). However, the test is less powerful than LMC.

In the STCC-GARCH model (DGPs 6-8) the time-varying correlations are defined as follows:

$$\mathbf{P}_t = (1 - G_t)\mathbf{P}_{(1)} + G_t\mathbf{P}_{(2)},$$

where $\mathbf{P}_t$ fluctuates between two positive definite correlation matrices $\mathbf{P}_{(1)}$ and $\mathbf{P}_{(2)}$ according to a transition function G_t which takes values between 0 and 1 depending on a continuous transition variable s_t. In our simulations G_t is a logistic function:

$$G_t(c, \gamma, s_t) = \left(1 + e^{-\gamma(s_t - c)}\right)^{-1}, \qquad \gamma > 0 \tag{37}$$

where γ is the speed and c the location of transition. In DGPs 6 and 7, $s_t = \varepsilon_{1,t-1}$ in (37) whereas in DGP 8, the transition variable s_t follows a first-order autoregressive process whose innovation is $\varepsilon_{1,t-1}$:

$$s_t = 0.99 s_{t-1} + \varepsilon_{1,t-1}.$$

In this case, the transition variable is quite persistent. The difference between DGP 6 and DGP 7 is that in the former the transition is fairly smooth, $\gamma = 5$, whereas it is rapid in the latter as $\gamma = 100$. In both DGPs, $c = 3$, which means that $\mathbf{P}_t$ on average stays closer to $\mathbf{P}_{(1)}$ than $\mathbf{P}_{(2)}$. In DGP 8, $\gamma = 5$ and $c = 0$, so the transition is smooth and due to persistent $\{s_t\}$ the correlations change slowly over time.

In all these DGPs the two correlation matrices are

$$\mathbf{P}_{(1)} = \begin{bmatrix} 1 & 0.3 \\ 0.3 & 1 \end{bmatrix}, \qquad \mathbf{P}_{(2)} = \begin{bmatrix} 1 & 0.9 \\ 0.9 & 1 \end{bmatrix}.$$

Again, LMC has the highest power of the three tests, but contrary to the DCC-GARCH alternative, LM_ζ also has power against DGPs 6 and 7 where $s_t = \varepsilon_{1,t-1}$. It has very little power against DGP 8. It seems that if the correlation fluctuates sufficiently slowly, LM_ζ does not respond to such time-variation. The performance of $Q(1)$ lies between that of LMC and LM_ζ. This test has power against all three DGPs but the power is clearly weaker than that of LMC, in small samples in particular. Interestingly, $\mathsf{Sup}T_n^*$ is again powerful when the correlation is persistent (DGP 8), while this is not the case in the other situations.

Finally we consider two diagonal BEKK-GARCH alternatives, where the model of the conditional covariances is given by

$$\mathbf{H}_t = \mathbf{C}\mathbf{C}' + \mathbf{A}_1'\boldsymbol{\varepsilon}_{t-1}\boldsymbol{\varepsilon}_{t-1}'\mathbf{A}_1 + \mathbf{B}_1'\mathbf{H}_{t-1}\mathbf{B}_1.$$

Tse & Tsui (1999) found that $Q(1)$ has low power against a diagonal BEKK-GARCH model. The model they use is DGP 10 in our study. The results in Table 4 show that, as in the case of DCC-GARCH, LM_ζ only has trivial power against the BEKK-GARCH models considered. $Q(1)$ has some power against the simplest diagonal BEKK-GARCH alternative (DGP 10) but trivial power against DGP 9. As can be expected, LMC has the highest power of the three tests. The semiparametric test still has its power about equal to the size of the test when $T = 2500$. From the simulations in Long et al. (2011) it appears that the test is generally undersized, so this result may not be surprising. As already mentioned, we do not have the computational resources needed to simulate the test at larger sample sizes.

The power of the tests is also simulated for $m = 5$. The results reported in Table 6 are similar to the ones obtained when $m = 2$. The LM_ζ test has in general higher power when $m = 5$ and the difference in power between the tests in favour of LM_ζ is even larger than in the bivariate case. The portmanteau test has slightly less power when $m = 5$ than when $m = 2$ when the alternative is a CCC-GARCH$(2,1)$ process. When the alternative is an STCC-GARCH model, the power of LM_ζ marginally increases with the dimension of the model.

The test results seem to suggest following guidelines as to what to do in practice after estimating a CCC-GARCH model. First carry out the three tests. If both LM_ζ and $Q(r)$ reject the null hypothesis of no ARCH in GARCH whereas LMC does not or does so only weakly, conclude that at least some of the GARCH equations have to be respecified. If all tests strongly reject, no conclusions can be drawn at this stage. If LMC rejects the null hypothesis of constant correlations whereas LM_ζ does not, tentatively assume that the correlations are not constant and fit a suitable multivariate GARCH model such as DCC-GARCH or BEKK-GARCH to the data. If both tests reject but LMC provides the strongest rejection, consider again giving up the assumption of constant conditional correlations but also consider the STCC-GARCH model as an alternative. If all three tests reject very strongly, reconsidering both the GARCH equations and the CCC-assumption could be useful. Note, however, that these guidelines are based on a rather limited number of simulation designs and are rather tentative. Finally, if one has reason to suspect spillover effects, these tests can be completed by the GARCH misspecification test in Nakatani and Teräsvirta (2009).

Table 5: Simulated power of three test statistics for testing the adequacy of the estimated CCC-GARCH model when $m = 2$ and $r = 1$.

	CCC-GARCH						DCC-GARCH				STCC-GARCH			BEKK-GARCH	
DGP	1a	2a	3a	1b	2b	3b	4a	5a	4b	5b	6	7	8	9	10
ρ		0.3			0.9		0.3		0.9						
$T = 1000$															
LM_ζ	0.752	0.660	1.000	0.816	0.805	1.000	0.048	0.043	0.043	0.040	0.142	0.145	0.063	0.047	0.044
$Q(r)$	0.260	0.338	0.764	0.283	0.314	0.740	0.056	0.051	0.274	0.112	0.176	0.180	0.439	0.048	0.129
LMC	0.060	0.063	0.074	0.197	0.253	0.616	0.883	0.433	0.868	0.502	0.559	0.543	0.662	0.989	0.944
$\mathsf{Sup}T_n^*$	0.028	0.024	0.022	0.000	0.002	0.000	0.072	0.024	0.070	0.000	0.000	0.000	0.334	0.070	0.028
$T = 2500$															
LM_ζ	0.988	0.973	1.000	0.995	0.995	1.000	0.049	0.049	0.049	0.046	0.265	0.280	0.063	0.051	0.055
$Q(r)$	0.535	0.709	0.983	0.587	0.687	0.980	0.062	0.053	0.559	0.173	0.350	0.365	0.713	0.059	0.265
LMC	0.059	0.072	0.103	0.285	0.386	0.899	0.998	0.762	0.993	0.794	0.869	0.862	0.937	1.000	1.000
$\mathsf{Sup}T_n^*$	0.030	0.037	0.040	0.000	0.005	0.000	0.155	0.005	0.200	0.000	0.000	0.000	0.890	0.049	0.050
$T = 5000$															
LM_ζ	1.000	1.000	1.000	1.000	1.000	1.000	0.048	0.049	0.052	0.047	0.511	0.516	0.068	0.049	0.052
$Q(r)$	0.825	0.947	1.000	0.869	0.931	1.000	0.085	0.058	0.833	0.279	0.617	0.603	0.917	0.052	0.445
LMC	0.070	0.079	0.151	0.415	0.547	0.989	1.000	0.957	1.000	0.963	0.985	0.988	0.995	1.000	1.000
$T = 10000$															
LM_ζ	1.000	1.000	1.000	1.000	1.000	1.000	0.053	0.054	0.064	0.051	0.799	0.804	0.067	0.047	0.047
$Q(r)$	0.984	0.999	1.000	0.992	0.998	1.000	0.111	0.056	0.980	0.507	0.881	0.884	0.995	0.046	0.735
LMC	0.087	0.104	0.244	0.584	0.715	0.999	1.000	1.000	1.000	0.999	1.000	1.000	1.000	1.000	1.000

Note: The nominal significance level is 5%.

Table 6: Simulated power of three test statistics for testing the adequacy of the estimated CCC-GARCH model when $m = 5$ and $r = 1$.

	CCC-GARCH						DCC-GARCH				STCC-GARCH			BEKK-GARCH	
DGP	1	2	3	1	2	3	4	5	4	5	6	7	8	9	10
ρ		0.3			0.9		0.3		0.9						
$T = 1000$															
LM_ζ	0.969	0.934	1.000	0.994	0.994	1.000	0.042	0.043	0.053	0.047	0.154	0.162	0.088	0.052	0.064
$Q(r)$	0.234	0.334	0.672	0.387	0.304	0.791	0.060	0.059	0.513	0.159	0.375	0.371	0.959	0.059	0.214
LMC	0.083	0.083	0.104	0.361	0.466	0.907	1.000	0.949	1.000	0.896	0.300	0.290	0.858	1.000	1.000
$T = 2500$															
LM_ζ	1.000	1.000	1.000	1.000	1.000	1.000	0.048	0.050	0.059	0.052	0.329	0.340	0.098	0.053	0.068
$Q(r)$	0.474	0.674	0.959	0.714	0.641	0.987	0.076	0.052	0.881	0.330	0.736	0.738	0.999	0.065	0.485
LMC	0.083	0.076	0.115	0.512	0.663	0.998	1.000	1.000	1.000	0.998	0.522	0.521	0.991	1.000	1.000
$T = 5000$															
LM_ζ	1.000	1.000	1.000	1.000	1.000	1.000	0.056	0.048	0.071	0.057	0.602	0.632	0.102	0.047	0.073
$Q(r)$	0.740	0.927	0.999	0.944	0.909	1.000	0.109	0.067	0.994	0.563	0.961	0.950	1.000	0.063	0.796
LMC	0.080	0.089	0.173	0.686	0.837	1.000	1.000	1.000	1.000	1.000	0.815	0.806	1.000	1.000	1.000
$T = 10000$															
LM_ζ	1.000	1.000	1.000	1.000	1.000	1.000	0.047	0.048	0.096	0.063	0.906	0.911	0.104	0.058	0.073
$Q(r)$	0.957	0.998	1.000	0.999	0.998	1.000	0.161	0.075	1.000	0.852	1.000	0.999	1.000	0.070	0.975
LMC	0.083	0.121	0.316	0.890	0.955	1.000	1.000	1.000	1.000	1.000	0.984	0.985	1.000	1.000	1.000

Note: The nominal significance level is 5%.

8 Conclusion

We derive an LM test for testing the adequacy of a fitted CCC-GARCH model. Monte Carlo simulations show that the test has good size properties. The test has reasonable power when the GARCH equations are misspecified, and the power of the test increases with the dimension of the model. In comparison with other tests, our test has higher power than the portmanteau test of Ling & Li (1997) when the GARCH equations are misspecified. On the other hand, the test is not greatly affected by misspecification in the conditional correlations in the sense that its power remains close to its size. A special case of an STCC-GARCH alternative constitutes a sole exception. Therefore the test is well suited for considering misspecification of GARCH equations. At the same time we find that the LMC test for time-varying correlations of Tse (2000), while having very low power when the misspecification is in the conditional covariances, performs remarkably well when the conditional correlation structure is misspecified. The portmanteau test of Ling & Li (1997) has some power against misspecification in both the GARCH equations and in the conditional correlations structure, but is in both cases outperformed by either our test or the test of Tse (2000). It therefore seems a good idea to carry out the last two tests or perhaps all three and, based on the outcomes, decide how to proceed from there. The test of Long et al. (2011) may be performed as well, but in our simulations it appears inferior to its parametric counterparts. It would be interesting to find cases where it is superior to the other test, but considering this problem is beyond the scope of this paper.

Appendix

The matrix derivations are based on results in Lütkepohl (1996), see also Nakatani & Teräsvirta (2009).

Proof of Lemma 1

First consider

$$\frac{\partial l_t(\boldsymbol{\theta})}{\partial \boldsymbol{\zeta}} = -\frac{\partial \ln |\mathbf{G}_t|}{\partial \boldsymbol{\zeta}} - \frac{1}{2}\frac{\partial \boldsymbol{\varepsilon}_t'\mathbf{H}_t^{-1}\boldsymbol{\varepsilon}_t}{\partial \boldsymbol{\zeta}} \tag{38}$$

where

$$-\frac{\partial \ln |\mathbf{G}_t|}{\partial \boldsymbol{\zeta}} = -\frac{\partial \mathsf{vec}(\mathbf{G}_t)'}{\partial \boldsymbol{\zeta}}\mathrm{vec}(\mathbf{G}_t^{-1}). \tag{39}$$

The second term of (38) becomes

$$-\frac{1}{2}\frac{\partial \boldsymbol{\varepsilon}_t'\mathbf{H}_t^{-1}\boldsymbol{\varepsilon}_t}{\partial \boldsymbol{\zeta}} = \frac{1}{2}\frac{\partial \mathsf{vec}(\mathbf{G}_t)'}{\partial \boldsymbol{\zeta}}\mathsf{vec}(\mathbf{H}_t^{-1}\boldsymbol{\varepsilon}_t\boldsymbol{\varepsilon}_t'\mathbf{G}_t^{-1} + \mathbf{G}_t^{-1}\boldsymbol{\varepsilon}_t\boldsymbol{\varepsilon}_t'\mathbf{H}_t^{-1}). \tag{40}$$

Inserting (39) and (40) into (38) yields

$$\frac{\partial l(\boldsymbol{\theta})}{\partial \boldsymbol{\zeta}} = -\frac{\partial \mathsf{vec}(\mathbf{G}_t)'}{\partial \boldsymbol{\zeta}}\mathsf{vec}(\mathbf{G}_t^{-1} - \frac{1}{2}\mathbf{H}_t^{-1}\boldsymbol{\varepsilon}_t\boldsymbol{\varepsilon}_t'\mathbf{G}_t^{-1} - \frac{1}{2}\mathbf{G}_t^{-1}\boldsymbol{\varepsilon}_t\boldsymbol{\varepsilon}_t'\mathbf{H}_t^{-1}) \tag{41}$$

where $\mathbf{H}_t = \mathbf{G}_t\mathbf{D}_t\mathbf{P}\mathbf{D}_t\mathbf{G}_t$. Evaluated under H_0, (41) has the form

$$\left.\frac{\partial l(\boldsymbol{\theta})}{\partial \boldsymbol{\zeta}}\right|_{H_0} = -\frac{\partial \mathsf{vec}(\mathbf{G}_t)'}{\partial \boldsymbol{\zeta}}\mathsf{vec}(\mathbf{I} - \frac{1}{2}\mathbf{M}_t^{-1}\boldsymbol{\varepsilon}_t\boldsymbol{\varepsilon}_t' - \frac{1}{2}\boldsymbol{\varepsilon}_t\boldsymbol{\varepsilon}_t'\mathbf{M}_t^{-1})$$

where $\boldsymbol{\theta} = (\boldsymbol{\omega}', \boldsymbol{\rho}', \boldsymbol{\zeta}')'$, $\boldsymbol{\omega} = (\boldsymbol{\omega}_1', ..., \boldsymbol{\omega}_m')'$, $\boldsymbol{\omega}_i = (\alpha_{i0}, \alpha_{i1}, \beta_{i1})'$, $\boldsymbol{\rho} = \mathsf{vecl}(\mathbf{P})$, $\boldsymbol{\zeta} = (\boldsymbol{\zeta}_1', ..., \boldsymbol{\zeta}_m')'$, $\boldsymbol{\zeta}_i' = (\zeta_{i1}, ..., \zeta_{ir})'$. ■

Proof of Lemma 2

The second partial derivatives of the log-likelihood function w.r.t. $\boldsymbol{\theta}$, the Hessian, are given by:

$$\mathbf{J}_t(\boldsymbol{\theta}) = \begin{bmatrix} \frac{\partial^2 l(\boldsymbol{\theta})}{\partial\boldsymbol{\omega}\partial\boldsymbol{\omega}'} & \frac{\partial^2 l(\boldsymbol{\theta})}{\partial\boldsymbol{\rho}\partial\boldsymbol{\omega}'} & \frac{\partial^2 l(\boldsymbol{\theta})}{\partial\boldsymbol{\zeta}\partial\boldsymbol{\omega}'} \\ \frac{\partial^2 l(\boldsymbol{\theta})}{\partial\boldsymbol{\omega}\partial\boldsymbol{\rho}'} & \frac{\partial^2 l(\boldsymbol{\theta})}{\partial\boldsymbol{\rho}\partial\boldsymbol{\rho}'} & \frac{\partial^2 l(\boldsymbol{\theta})}{\partial\boldsymbol{\zeta}\partial\boldsymbol{\rho}'} \\ \frac{\partial^2 l(\boldsymbol{\theta})}{\partial\boldsymbol{\omega}\partial\boldsymbol{\zeta}'} & \frac{\partial^2 l(\boldsymbol{\theta})}{\partial\boldsymbol{\rho}\partial\boldsymbol{\zeta}'} & \frac{\partial^2 l(\boldsymbol{\theta})}{\partial\boldsymbol{\zeta}\partial\boldsymbol{\zeta}'} \end{bmatrix}.$$

Begin by

$$\begin{aligned} \frac{\partial^2 l_t(\boldsymbol{\theta})}{\partial\boldsymbol{\omega}'\partial\boldsymbol{\zeta}} &= -\frac{\partial}{\partial\boldsymbol{\omega}'}\left(\frac{\partial \mathsf{vec}(\mathbf{G}_t)'}{\partial\boldsymbol{\zeta}}\mathsf{vec}(\mathbf{G}_t^{-1})\right) \\ &\quad +\frac{1}{2}\frac{\partial}{\partial\boldsymbol{\omega}'}\left(\frac{\partial \mathsf{vec}(\mathbf{G}_t)'}{\partial\boldsymbol{\zeta}}\mathsf{vec}(\mathbf{G}_t^{-1}\mathbf{D}_t^{-1}\mathbf{P}^{-1}\mathbf{D}_t^{-1}\mathbf{G}_t^{-1}\boldsymbol{\varepsilon}_t\boldsymbol{\varepsilon}_t'\mathbf{G}_t^{-1})\right) \\ &\quad +\frac{1}{2}\frac{\partial}{\partial\boldsymbol{\omega}'}\left(\frac{\partial \mathsf{vec}(\mathbf{G}_t)'}{\partial\boldsymbol{\zeta}}\mathsf{vec}(\mathbf{G}_t^{-1}\boldsymbol{\varepsilon}_t\boldsymbol{\varepsilon}_t'\mathbf{G}_t^{-1}\mathbf{D}_t^{-1}\mathbf{P}^{-1}\mathbf{D}_t^{-1}\mathbf{G}_t^{-1})\right) \\ &= \mathbf{A}_1 + \mathbf{A}_2 + \mathbf{A}_{3.}. \end{aligned} \tag{42}$$

First we see that $\mathbf{A}_1 = \mathbf{0}$. Second,

$$\begin{aligned} \mathbf{A}_2 &= \frac{1}{2}\frac{\partial \mathsf{vec}(\mathbf{G}_t)}{\partial\boldsymbol{\zeta}'}\frac{\partial \mathsf{vec}(\mathbf{G}_t^{-1}\mathbf{D}_t^{-1}\mathbf{P}^{-1}\mathbf{D}_t^{-1}\mathbf{G}_t^{-1}\boldsymbol{\varepsilon}_t\boldsymbol{\varepsilon}_t'\mathbf{G}_t^{-1})}{\partial \mathsf{vec}(\mathbf{D}_t^{-1})'}\frac{\partial \mathsf{vec}(\mathbf{D}_t^{-1})'}{\partial \mathsf{vec}(\mathbf{D}_t)}\frac{\partial \mathsf{vec}(\mathbf{D}_t)}{\partial\boldsymbol{\omega}} \\ &= -\frac{1}{2}\frac{\partial \mathsf{vec}(\mathbf{G}_t)}{\partial\boldsymbol{\zeta}'}(\mathbf{G}_t^{-1}\boldsymbol{\varepsilon}_t\boldsymbol{\varepsilon}_t'\mathbf{G}_t^{-1}\mathbf{D}_t^{-1}\mathbf{P}^{-1} \otimes \mathbf{G}_t^{-1} \\ &\quad +\mathbf{G}_t^{-1}\boldsymbol{\varepsilon}_t\boldsymbol{\varepsilon}_t'\mathbf{G}_t^{-1} \otimes \mathbf{G}_t^{-1}\mathbf{D}_t^{-1}\mathbf{P}^{-1})(\mathbf{D}_t^{-1} \otimes \mathbf{D}_t^{-1})\frac{\partial \mathsf{vec}(\mathbf{D}_t)}{\partial\boldsymbol{\omega}} \\ &= -\frac{1}{2}\frac{\partial \mathsf{vec}(\mathbf{G}_t)}{\partial\boldsymbol{\zeta}'}(\mathbf{G}_t^{-1}\boldsymbol{\varepsilon}_t\boldsymbol{\varepsilon}_t'\mathbf{G}_t^{-1}\mathbf{D}_t^{-1}\mathbf{P}^{-1}\mathbf{D}_t^{-1} \otimes \mathbf{G}_t^{-1}\mathbf{D}_t^{-1} \\ &\quad +\mathbf{G}_t^{-1}\boldsymbol{\varepsilon}_t\boldsymbol{\varepsilon}_t'\mathbf{G}_t^{-1}\mathbf{D}_t^{-1} \otimes \mathbf{G}_t^{-1}\mathbf{D}_t^{-1}\mathbf{P}^{-1}\mathbf{D}_t^{-1})\frac{\partial \mathsf{vec}(\mathbf{D}_t)}{\partial\boldsymbol{\omega}}. \end{aligned} \tag{43}$$

Similarly,

$$\begin{aligned} \mathbf{A}_3 &= -\frac{1}{2}\frac{\partial \mathsf{vec}(\mathbf{G}_t)'}{\partial\boldsymbol{\zeta}}(\mathbf{G}_t^{-1}\mathbf{D}_t^{-1}\mathbf{P}^{-1}\mathbf{D}_t^{-1} \otimes \mathbf{G}_t^{-1}\boldsymbol{\varepsilon}_t\boldsymbol{\varepsilon}_t'\mathbf{G}_t^{-1}\mathbf{D}_t^{-1} \\ &\quad +\mathbf{G}_t^{-1}\mathbf{D}_t^{-1} \otimes \mathbf{G}_t^{-1}\boldsymbol{\varepsilon}_t\boldsymbol{\varepsilon}_t'\mathbf{G}_t^{-1}\mathbf{D}_t^{-1}\mathbf{P}^{-1}\mathbf{D}_t^{-1})\frac{\partial \mathsf{vec}(\mathbf{D}_t)}{\partial\boldsymbol{\omega}'}. \end{aligned} \tag{44}$$

Inserting (43) and (44) into (42), setting $\mathbf{A}_1 = \mathbf{0}$ and using $\boldsymbol{\varepsilon}_t = \mathbf{G}_t\mathbf{D}_t\mathbf{u}_t$ yields

$$\begin{aligned} \frac{\partial^2 l_t(\boldsymbol{\theta})}{\partial\boldsymbol{\omega}'\partial\boldsymbol{\zeta}} &= -\frac{1}{2}\frac{\partial \mathsf{vec}(\mathbf{G}_t)'}{\partial\boldsymbol{\zeta}}(\mathbf{D}_t\mathbf{u}_t\mathbf{u}_t'\mathbf{P}^{-1}\mathbf{D}_t^{-1} \otimes \mathbf{G}_t^{-1}\mathbf{D}_t^{-1} + \mathbf{D}_t\mathbf{u}_t\mathbf{u}_t' \otimes \mathbf{G}_t^{-1}\mathbf{M}_t^{-1} \\ &\quad +\mathbf{G}_t^{-1}\mathbf{M}_t^{-1} \otimes \mathbf{D}_t\mathbf{u}_t\mathbf{u}_t' + \mathbf{G}_t^{-1}\mathbf{D}_t^{-1} \otimes \mathbf{D}_t\mathbf{u}_t\mathbf{u}_t'\mathbf{P}^{-1}\mathbf{D}_t^{-1})\frac{\partial \mathsf{vec}(\mathbf{D}_t)}{\partial\boldsymbol{\omega}'}. \end{aligned}$$

Next consider

$$\begin{aligned}
\frac{\partial^2 l_t(\boldsymbol{\theta})}{\partial \boldsymbol{\rho}' \partial \boldsymbol{\zeta}} &= -\frac{\partial}{\partial \boldsymbol{\rho}'}\left(\frac{\partial \mathsf{vec}(\mathbf{G}_t)'}{\partial \boldsymbol{\zeta}}\mathsf{vec}(\mathbf{G}_t^{-1})\right) \\
&\quad +\frac{1}{2}\frac{\partial}{\partial \boldsymbol{\rho}'}\left(\frac{\partial \mathsf{vec}(\mathbf{G}_t)'}{\partial \boldsymbol{\zeta}}\mathsf{vec}(\mathbf{G}_t^{-1}\mathbf{D}_t^{-1}\mathbf{P}^{-1}\mathbf{D}_t^{-1}\mathbf{G}_t^{-1}\boldsymbol{\varepsilon}_t\boldsymbol{\varepsilon}_t\mathbf{G}_t^{-1})\right) \\
&\quad +\frac{1}{2}\frac{\partial}{\partial \boldsymbol{\rho}'}\left(\frac{\partial \mathsf{vec}(\mathbf{G}_t)'}{\partial \boldsymbol{\zeta}}\mathsf{vec}(\mathbf{G}_t^{-1}\boldsymbol{\varepsilon}_t\boldsymbol{\varepsilon}_t\mathbf{G}_t^{-1}\mathbf{D}_t^{-1}\mathbf{P}^{-1}\mathbf{D}_t^{-1}\mathbf{G}_t^{-1})\right) \\
&= \mathbf{B}_1 + \mathbf{B}_2 + \mathbf{B}_3. \qquad (45)
\end{aligned}$$

First, $\mathbf{B}_1 = \mathbf{0}$. Second,

$$\begin{aligned}
\mathbf{B}_2 &= \frac{1}{2}\left(\frac{\partial \mathsf{vec}(\mathbf{G}_t)'}{\partial \boldsymbol{\zeta}}\frac{\partial \mathsf{vec}(\mathbf{G}_t^{-1}\mathbf{D}_t^{-1}\mathbf{P}^{-1}\mathbf{D}_t^{-1}\mathbf{G}_t^{-1}\boldsymbol{\varepsilon}_t\boldsymbol{\varepsilon}_t'\mathbf{G}_t^{-1})}{\partial \mathsf{vec}(\mathbf{P}^{-1})'}\frac{\partial \mathsf{vec}(\mathbf{P}^{-1})}{\partial \mathsf{vec}(\mathbf{P})'}\frac{\partial \mathsf{vec}(\mathbf{P})}{\partial \boldsymbol{\rho}'}\right) \\
&= -\frac{1}{2}\frac{\partial \mathsf{vec}(\mathbf{G}_t)'}{\partial \boldsymbol{\zeta}}\left(\mathbf{G}_t^{-1}\boldsymbol{\varepsilon}_t\boldsymbol{\varepsilon}_t'\mathbf{G}_t^{-1}\mathbf{D}_t^{-1} \otimes \mathbf{G}_t^{-1}\mathbf{D}_t^{-1}\right)\left(\mathbf{P}^{-1} \otimes \mathbf{P}^{-1}\right)\frac{\partial \mathsf{vec}(\mathbf{P})}{\partial \boldsymbol{\rho}'} \\
&= -\frac{1}{2}\frac{\partial \mathsf{vec}(\mathbf{G}_t)'}{\partial \boldsymbol{\zeta}}\left(\mathbf{G}_t^{-1}\boldsymbol{\varepsilon}_t\boldsymbol{\varepsilon}_t'\mathbf{G}_t^{-1}\mathbf{D}_t^{-1}\mathbf{P}^{-1} \otimes \mathbf{G}_t^{-1}\mathbf{D}_t^{-1}\mathbf{P}^{-1}\right)\frac{\partial \mathsf{vec}(\mathbf{P})}{\partial \boldsymbol{\rho}'} \qquad (46)
\end{aligned}$$

and, finally,

$$\mathbf{B}_3 = -\frac{1}{2}\frac{\partial \mathsf{vec}(\mathbf{G}_t)'}{\partial \boldsymbol{\zeta}}\left(\mathbf{G}_t^{-1}\mathbf{D}_t^{-1}\mathbf{P}^{-1} \otimes \mathbf{G}_t^{-1}\boldsymbol{\varepsilon}_t\boldsymbol{\varepsilon}_t'\mathbf{G}_t^{-1}\mathbf{D}_t^{-1}\mathbf{P}^{-1}\right)\frac{\partial \mathsf{vec}(\mathbf{P})}{\partial \boldsymbol{\rho}'}. \qquad (47)$$

Inserting (46) and (47) into (45) and setting $\mathbf{B}_1 = \mathbf{0}$ gives

$$\begin{aligned}
\frac{\partial^2 l(\boldsymbol{\theta})}{\partial \boldsymbol{\rho}' \partial \boldsymbol{\zeta}} &= -\frac{1}{2}\frac{\partial \mathsf{vec}(\mathbf{G}_t)'}{\partial \boldsymbol{\zeta}}\left\{\mathbf{D}_t\mathbf{u}_t\mathbf{u}_t'\mathbf{P}^{-1} \otimes \mathbf{G}_t^{-1}\mathbf{D}_t^{-1}\mathbf{P}^{-1}\right. \\
&\quad \left. +\mathbf{G}_t^{-1}\mathbf{D}_t^{-1}\mathbf{P}^{-1} \otimes \mathbf{D}_t\mathbf{u}_t\mathbf{u}_t'\mathbf{P}^{-1}\right\}\frac{\partial \mathsf{vec}(\mathbf{P})}{\partial \boldsymbol{\rho}'}.
\end{aligned}$$

Finally, look at

$$\begin{aligned}
\frac{\partial^2 l_t(\boldsymbol{\theta})}{\partial \boldsymbol{\zeta}' \partial \boldsymbol{\zeta}} &= -\frac{\partial}{\partial \boldsymbol{\zeta}'}\left(\frac{\partial \mathsf{vec}(\mathbf{G}_t)'}{\partial \boldsymbol{\zeta}}\mathsf{vec}(\mathbf{G}_t^{-1})\right) \\
&\quad +\frac{1}{2}\frac{\partial}{\partial \boldsymbol{\zeta}'}\left(\frac{\partial \mathsf{vec}(\mathbf{G}_t)'}{\partial \boldsymbol{\zeta}}\mathsf{vec}(\mathbf{G}_t^{-1}\mathbf{D}_t^{-1}\mathbf{P}^{-1}\mathbf{D}_t^{-1}\mathbf{G}_t^{-1}\boldsymbol{\varepsilon}_t\boldsymbol{\varepsilon}_t'\mathbf{G}_t^{-1})\right) \\
&\quad +\frac{1}{2}\frac{\partial}{\partial \boldsymbol{\zeta}'}\left(\frac{\partial \mathsf{vec}(\mathbf{G}_t)'}{\partial \boldsymbol{\zeta}}\mathsf{vec}(\mathbf{G}_t^{-1}\boldsymbol{\varepsilon}_t\boldsymbol{\varepsilon}_t'\mathbf{G}_t^{-1}\mathbf{D}_t^{-1}\mathbf{P}^{-1}\mathbf{D}_t^{-1}\mathbf{G}_t^{-1})\right) \\
&= \mathbf{C}_1 + \mathbf{C}_2 + \mathbf{C}_3. \qquad (48)
\end{aligned}$$

First,

$$\begin{aligned}
\mathbf{C}_1 &= \frac{\partial \mathsf{vec}(\mathbf{G}_t)'}{\partial \boldsymbol{\zeta}}(\mathbf{G}_t^{-1} \otimes \mathbf{G}_t^{-1})\frac{\partial \mathsf{vec}(\mathbf{G}_t)}{\partial \boldsymbol{\zeta}'} \\
&\quad - \left[\mathsf{vec}(\mathbf{G}_t^{-1})' \otimes \mathbf{I}\right]\frac{\partial^2 \mathsf{vec}(\mathbf{G}_t)'}{\partial \boldsymbol{\zeta}' \partial \boldsymbol{\zeta}}. \qquad (49)
\end{aligned}$$

Next,

$$
\begin{aligned}
\mathbf{C}_2 &= \frac{1}{2}\left[\mathsf{vec}(\mathbf{G}_t^{-1}\mathbf{M}_t^{-1}\mathbf{G}_t^{-1}\boldsymbol{\varepsilon}_t\boldsymbol{\varepsilon}_t'\mathbf{G}_t^{-1})'\otimes\mathbf{I}\right]\frac{\partial^2\mathsf{vec}(\mathbf{G}_t)'}{\partial\boldsymbol{\zeta}'\partial\boldsymbol{\zeta}} \\
&\quad +\frac{1}{2}\frac{\partial\mathsf{vec}(\mathbf{G}_t)'}{\partial\boldsymbol{\zeta}}\frac{\partial\mathsf{vec}(\mathbf{G}_t^{-1}\mathbf{M}_t^{-1}\mathbf{G}_t^{-1}\boldsymbol{\varepsilon}_t\boldsymbol{\varepsilon}_t'\mathbf{G}_t^{-1})}{\partial\mathsf{vec}(\mathbf{G}_t^{-1})'}\frac{\partial\mathsf{vec}(\mathbf{G}_t^{-1})}{\partial\mathsf{vec}(\mathbf{G}_t)'}\frac{\partial\mathsf{vec}(\mathbf{G}_t)}{\partial\boldsymbol{\zeta}'} \\
&= \frac{1}{2}\left[\mathsf{vec}(\mathbf{G}_t^{-1}\mathbf{M}_t^{-1}\mathbf{G}_t^{-1}\boldsymbol{\varepsilon}_t\boldsymbol{\varepsilon}_t'\mathbf{G}_t^{-1})'\otimes\mathbf{I}\right]\frac{\partial^2\mathsf{vec}(\mathbf{G}_t)'}{\partial\boldsymbol{\zeta}'\partial\boldsymbol{\zeta}} \\
&\quad -\frac{1}{2}\frac{\partial\mathsf{vec}(\mathbf{G}_t)'}{\partial\boldsymbol{\zeta}}(\mathbf{G}_t^{-1}\boldsymbol{\varepsilon}_t\boldsymbol{\varepsilon}_t'\mathbf{G}_t^{-1}\otimes\mathbf{G}_t^{-1}\mathbf{M}_t^{-1}\mathbf{G}_t^{-1}+\mathbf{G}_t^{-1}\otimes\mathbf{G}_t^{-1}\mathbf{M}_t^{-1}\mathbf{G}_t^{-1}\boldsymbol{\varepsilon}_t\boldsymbol{\varepsilon}_t'\mathbf{G}_t^{-1} \\
&\quad +\mathbf{G}_t^{-1}\boldsymbol{\varepsilon}_t\boldsymbol{\varepsilon}_t'\mathbf{G}_t^{-1}\mathbf{M}_t^{-1}\mathbf{G}_t^{-1}\otimes\mathbf{G}_t^{-1})\frac{\partial\mathsf{vec}(\mathbf{G}_t)}{\partial\boldsymbol{\zeta}'} \\
&= \frac{1}{2}\left[\mathsf{vec}(\mathbf{G}_t^{-1}\mathbf{D}_t^{-1}\mathbf{P}^{-1}\mathbf{u}_t\mathbf{u}_t'\mathbf{D}_t)'\otimes\mathbf{I}\right]\frac{\partial^2\mathsf{vec}(\mathbf{G}_t)'}{\partial\boldsymbol{\zeta}'\partial\boldsymbol{\zeta}} \\
&\quad -\frac{1}{2}\frac{\partial\mathsf{vec}(\mathbf{G}_t)'}{\partial\boldsymbol{\zeta}}(\mathbf{D}_t\mathbf{u}_t\mathbf{u}_t'\mathbf{D}_t\otimes\mathbf{H}_t^{-1}+\mathbf{G}_t^{-1}\otimes\mathbf{G}_t^{-1}\mathbf{M}_t^{-1}\mathbf{D}_t\mathbf{u}_t\mathbf{u}_t'\mathbf{D}_t \\
&\quad +\mathbf{D}_t\mathbf{u}_t\mathbf{u}_t'\mathbf{D}_t\mathbf{M}_t^{-1}\mathbf{G}_t^{-1}\otimes\mathbf{G}_t^{-1})\frac{\partial\mathsf{vec}(\mathbf{G}_t)}{\partial\boldsymbol{\zeta}'} \qquad (50)
\end{aligned}
$$

and, finally,

$$
\begin{aligned}
\mathbf{C}_3 &= \frac{1}{2}\left[\mathsf{vec}(\mathbf{G}_t^{-1}\boldsymbol{\varepsilon}_t\boldsymbol{\varepsilon}_t'\mathbf{G}_t^{-1}\mathbf{M}_t^{-1}\mathbf{G}_t^{-1})'\otimes\mathbf{I}\right]\frac{\partial^2\mathsf{vec}(\mathbf{G}_t)'}{\partial\boldsymbol{\zeta}'\partial\boldsymbol{\zeta}} \\
&\quad +\frac{1}{2}\frac{\partial\mathsf{vec}(\mathbf{G}_t)'}{\partial\boldsymbol{\zeta}}\frac{\partial\mathsf{vec}(\mathbf{G}_t^{-1}\boldsymbol{\varepsilon}_t\boldsymbol{\varepsilon}_t'\mathbf{G}_t^{-1}\mathbf{M}_t^{-1}\mathbf{G}_t^{-1})}{\partial\mathsf{vec}(\mathbf{G}_t^{-1})'}\frac{\partial\mathsf{vec}(\mathbf{G}_t^{-1})}{\partial\mathsf{vec}(\mathbf{G}_t)'}\frac{\partial\mathsf{vec}(\mathbf{G}_t)}{\partial\boldsymbol{\zeta}'} \\
&= \frac{1}{2}\left[\mathsf{vec}(\mathbf{D}_t\mathbf{u}_t\mathbf{u}_t'\mathbf{P}^{-1}\mathbf{D}_t^{-1}\mathbf{G}_t^{-1})'\otimes\mathbf{I}\right]\frac{\partial^2\mathsf{vec}(\mathbf{G}_t)'}{\partial\boldsymbol{\zeta}'\partial\boldsymbol{\zeta}} \\
&\quad -\frac{1}{2}\frac{\partial\mathsf{vec}(\mathbf{G}_t)'}{\partial\boldsymbol{\zeta}}(\mathbf{G}_t^{-1}\otimes\mathbf{D}_t\mathbf{u}_t\mathbf{u}_t'\mathbf{D}_t\mathbf{M}_t^{-1}\mathbf{G}_t^{-1}+\mathbf{G}_t^{-1}\mathbf{M}_t^{-1}\mathbf{D}_t\mathbf{u}_t\mathbf{u}_t'\mathbf{D}_t\otimes\mathbf{G}_t^{-1} \\
&\quad +\mathbf{H}_t^{-1}\otimes\mathbf{D}_t\mathbf{u}_t\mathbf{u}_t'\mathbf{D}_t)\frac{\partial\mathsf{vec}(\mathbf{G}_t)}{\partial\boldsymbol{\zeta}'}. \qquad (51)
\end{aligned}
$$

Inserting (49), (50) and (51) into (48) results in

$$
\begin{aligned}
\frac{\partial^2 l(\boldsymbol{\theta})}{\partial\boldsymbol{\zeta}'\partial\boldsymbol{\zeta}} &= -\Big\{\left[(\mathsf{vec}(\mathbf{G}_t^{-1})'\otimes\mathbf{I})\right]-\frac{1}{2}\left[\mathsf{vec}(\mathbf{G}_t^{-1}\mathbf{M}_t^{-1}\mathbf{D}_t\mathbf{u}_t\mathbf{u}_t'\mathbf{D}_t)'\otimes\mathbf{I}\right] \\
&\quad -\frac{1}{2}\left[\mathsf{vec}(\mathbf{G}_t^{-1}\mathbf{M}_t^{-1}\mathbf{D}_t\mathbf{u}_t\mathbf{u}_t'\mathbf{D}_t)'\otimes\mathbf{I}\right]\Big\}\frac{\partial^2\mathsf{vec}(\mathbf{G}_t)'}{\partial\boldsymbol{\zeta}'\partial\boldsymbol{\zeta}} \\
&\quad +\frac{1}{2}\frac{\partial\mathsf{vec}(\mathbf{G}_t)'}{\partial\boldsymbol{\zeta}}\left\{2(\mathbf{G}_t^{-1}\otimes\mathbf{G}_t^{-1})-\mathbf{D}_t\mathbf{u}_t\mathbf{u}_t'\mathbf{D}_t\otimes\mathbf{H}_t^{-1}-\mathbf{G}_t^{-1}\otimes\mathbf{G}_t^{-1}\mathbf{M}_t^{-1}\mathbf{D}_t\mathbf{u}_t\mathbf{u}_t'\mathbf{D}_t\right. \\
&\quad -\mathbf{D}_t\mathbf{u}_t\mathbf{u}_t'\mathbf{D}_t\mathbf{M}_t^{-1}\mathbf{G}_t^{-1}\otimes\mathbf{G}_t^{-1}-\mathbf{G}_t^{-1}\otimes\mathbf{D}_t\mathbf{u}_t\mathbf{u}_t'\mathbf{D}_t\mathbf{M}_t^{-1}\mathbf{G}_t^{-1} \\
&\quad \left.-\mathbf{G}_t^{-1}\mathbf{M}_t^{-1}\mathbf{D}_t\mathbf{u}_t\mathbf{u}_t'\mathbf{D}_t\otimes\mathbf{G}_t^{-1}-\mathbf{H}_t^{-1}\otimes\mathbf{D}_t\mathbf{u}_t\mathbf{u}_t'\mathbf{D}_t\right\}\frac{\partial\mathsf{vec}(\mathbf{G}_t)}{\partial\boldsymbol{\zeta}'}.
\end{aligned}
$$

This completes the proof. ■

Proof of Corollary

Let $\widetilde{\boldsymbol{\theta}}_T = (\widetilde{\boldsymbol{\omega}}'_T, \widetilde{\boldsymbol{\rho}}'_T, \widetilde{\boldsymbol{\zeta}}'_T)'$ be the maximum likelihood estimator of the true parameter $\boldsymbol{\theta}_0 = (\boldsymbol{\omega}'_0, \boldsymbol{\rho}'_0, \boldsymbol{\zeta}'_0)'$ under H_0. The first-order Taylor expansion of $\mathbf{q}(\widetilde{\boldsymbol{\theta}}_T)$ around $\boldsymbol{\theta}_0$ equals

$$\mathbf{0} = \mathbf{q}(\widetilde{\boldsymbol{\theta}}_T) = \mathbf{q}(\boldsymbol{\theta}_0) + \frac{\partial \mathbf{q}(\overline{\boldsymbol{\theta}})}{\partial \boldsymbol{\theta}}(\widetilde{\boldsymbol{\theta}}_T - \boldsymbol{\theta}_0) \tag{52}$$

where $||\overline{\boldsymbol{\theta}}|| \leq ||\widetilde{\boldsymbol{\theta}}_T||$. Assume that for each T there is a model with the true parameter value $\boldsymbol{\theta}_{0T} = (\boldsymbol{\omega}'_0, \boldsymbol{\rho}'_0, \boldsymbol{\zeta}'_{0T})'$ where $\boldsymbol{\zeta}_{0T} = \boldsymbol{\delta}_\zeta / T^{1/2}$ ($\boldsymbol{\delta}_\zeta \neq \mathbf{0}$ is a fixed vector). This implies that there is a sequence of models with the parameter vector $\boldsymbol{\theta}_{0T} = (\boldsymbol{\omega}'_0, \boldsymbol{\rho}'_0, \boldsymbol{\delta}'_\zeta / T^{1/2})'$ that form a sequence of (Pitman) local alternatives to H_0. Multiplying both sides of (52) by $T^{1/2}(\frac{\partial \mathbf{q}(\overline{\boldsymbol{\theta}})}{\partial \boldsymbol{\theta}})^{-1}$ and reshuffling yields

$$\begin{aligned} -(\frac{\partial \mathbf{q}_\varsigma(\overline{\boldsymbol{\theta}})}{\partial \boldsymbol{\theta}})^{-1} T^{1/2} \mathbf{q}(\boldsymbol{\theta}_0) &= T^{1/2}\{(\widetilde{\boldsymbol{\theta}}_{0T} - \boldsymbol{\theta}_{0T}) + (\boldsymbol{\theta}_{0T} - \boldsymbol{\theta}_0)\} \\ &= T^{1/2}(\widetilde{\boldsymbol{\theta}}_T - \boldsymbol{\theta}_{0T}) + \boldsymbol{\delta} \end{aligned}$$

where $\boldsymbol{\delta} = (\mathbf{0}'_{3m}, \mathbf{0}'_{m(m-1)/2}, \boldsymbol{\delta}'_\zeta / T^{1/2})'$. Let $\mathbf{J}(\boldsymbol{\theta}_0) = \mathrm{plim}_{T\to\infty} \partial \mathbf{q}(\boldsymbol{\theta}_0)/\partial \boldsymbol{\theta}$. As $T \to \infty$, $\boldsymbol{\theta}_{0T} \to \boldsymbol{\theta}_0$, $\widetilde{\boldsymbol{\theta}}_T \xrightarrow{p} \boldsymbol{\theta}_0$, $\overline{\boldsymbol{\theta}} \to \boldsymbol{\theta}_0$ and therefore, under regularity conditions and using the continuous mapping theorem, $T^{1/2}(\widetilde{\boldsymbol{\theta}}_T - \boldsymbol{\theta}_{0T}) \xrightarrow{D} \mathcal{N}(\mathbf{0}, \mathbf{J}^{-1}(\boldsymbol{\theta}_0)\mathbf{I}(\boldsymbol{\theta}_0)\mathbf{J}^{-1}(\boldsymbol{\theta}_0)))$. Since $T^{1/2}(\widetilde{\boldsymbol{\theta}}_T - \boldsymbol{\theta}_{0T})$ is asymptotically normal with mean zero, it follows that $T^{1/2}(\widetilde{\boldsymbol{\theta}}_T - \boldsymbol{\theta}_{0T}) + \boldsymbol{\delta} \xrightarrow{D} \mathcal{N}(\boldsymbol{\delta}, \mathbf{J}^{-1}(\boldsymbol{\theta}_0)\mathbf{I}(\boldsymbol{\theta}_0)\mathbf{J}^{-1}(\boldsymbol{\theta}_0)))$. Thus, under H_1, the LM statistic (30) has an asymptotic noncentral χ^2-distribution with noncentrality parameter

$$\begin{aligned} & \boldsymbol{\delta}'\mathbf{J}^{-1}(\boldsymbol{\theta}_0)(\mathbf{J}^{-1}(\boldsymbol{\theta}_0)\mathbf{I}(\boldsymbol{\theta}_0)\mathbf{J}^{-1}(\boldsymbol{\theta}_0))^{-1}\mathbf{J}^{-1}(\boldsymbol{\theta}_0)\boldsymbol{\delta} \\ = \; & \boldsymbol{\delta}'\mathbf{I}^{-1}(\boldsymbol{\theta}_0)\boldsymbol{\delta} = \boldsymbol{\delta}'_\zeta \mathbf{I}^{\zeta\zeta}(\boldsymbol{\theta}_0)\boldsymbol{\delta}'_\zeta. \end{aligned}$$

This completes the proof. ■

References

Amado, C. & Teräsvirta, T. (2013), 'Modelling volatility by variance decomposition', *Journal of Econometrics* **175**, 142–153.

Bauwens, L., Laurent, S. & Rombouts, J. V. K. (2006), 'Multivariate GARCH models: a survey', *Journal of Applied Econometrics* **21**, 79–109.

Bollerslev, T. (1990), 'Modelling the coherence in short-run nominal exchange rates: A multivariate generalized ARCH model', *Review of Economics and Statistics* **72**, 498–505.

Bollerslev, T., Engle, R. F. & Woodridge, J. M. (1988), 'A capital asset pricing model with time-varying covariances', *Journal of Political Economy* **96**, 116–131.

Davidson, R. & MacKinnon, J. G. (2006), 'The power of bootstrap and asymptotic tests', *Journal of Econometrics* **133**, 421 – 441.

Eklund, B. & Teräsvirta, T. (2007), 'Testing constancy of the error covariance matrix in vector models', *Journal of Econometrics* **140**, 753–780.

Engle, R. F. (1982), 'Autoregressive conditional heteroscedasticity with estimates of the variance of United Kingdom inflation', *Econometrica* **50**, 987–1007.

Engle, R. F. (2002), 'Dynamic conditional correlation: A simple class of multivariate generalized autoregressive conditional heteroskedasticity models', *Journal of Business on Economic Statistics* **20**, 339–350.

Engle, R. F. & Kelly, B. (2012), 'Dynamic equicorrelation', *Journal of Business and Economic Statistics* **30**, 212–228.

Engle, R. F. & Kroner, K. F. (1995), 'Multivariate simulataneous generalized ARCH', *Econometric Theory* **11**, 122–150.

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

Jeantheau, T. (1998), 'Strong consistency of estimators for multivariate ARCH models', *Econometric Theory* **14**, 70–86.

Li, W. K. & Mak, T. K. (1994), 'On the squared residual autocorrelations in non-linear time series with conditional heteroskedasticity', *Journal of Time Series Analysis* **15**, 627–636.

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

Ling, S. & McAleer, M. (2003), 'Asymptotic theory for a vector ARMA-GARCH model', *Econometric Theory* **19**, 280–310.

Long, X., Su, L. & Ullah, A. (2011), 'Estimation and forecasting of dynamic conditional covariance: a semiparametric multivariate model', *Journal of Business and Economic Statistics* **29**, 109–125.

Lundbergh, S. & Teräsvirta, T. (2002), 'Evaluating GARCH models', *Journal of Econometrics* **110**, 417–435.

Lütkepohl, H. (1996), *Handbook of Matrices*, Wiley, New York.

McLeod, A. I. & Li, W. K. (1983), 'Diagnostic checking ARMA time series models using squared residual autocorrelations', *Journal of Time Series Analysis* **4**, 269–273.

Nakatani, T. (2013), *ccgarch: An R Package for Modelling Multivariate GARCH Models with Conditional Correlations.* R package version 0.2.0-2, URL <http://CRAN.r-project/package=ccgarch>.

Nakatani, T. & Teräsvirta, T. (2009), 'Testing for volatility interactions in the constant conditional correlation GARCH model', *Econometrics Journal* **12**, 147–163.

Pelletier, D. (2006), 'Regime switching for dynamic correlations', *Journal of Econometrics* **131**, 445–473.

R Core Team (2013), *R: A Language and Environment for Statistical Computing*, R Foundation for Statistical Computing, Vienna, Austria.
URL: *http://www.R-project.org*

Silvennoinen, A. & Teräsvirta, T. (2009), Multivariate GARCH models, *in* T. G. Andersen, R. A. Davis, J.-P. Kreiss & T. V. Mikosch, eds, 'Handbook of Financial Time Series', Springer, New York, pp. 201–229.

Silvennoinen, A. & Teräsvirta, T. (2015), 'Modelling conditional correlations of asset returns: A smooth transition approach', *Econometric Reviews* **34**, 174–197.

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

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