Econometric Theory, **24**, 2008, 1554–1583. Printed in the United States of America.
doi:10.1017/S0266466608080614

GENERALIZED AUTOREGRESSIVE CONDITIONAL CORRELATION

MICHAEL MCALEER
University of Western Australia
and
Yokohama National University

FELIX CHAN
Curtin University of Technology

SUHEJLA HOTI
University of Western Australia

OFFER LIEBERMAN
University of Haifa

This paper develops a generalized autoregressive conditional correlation (GARCC) model when the standardized residuals follow a random coefficient vector autoregressive process. As a multivariate generalization of the Tsay (1987, *Journal of the American Statistical Association* 82, 590–604) random coefficient autoregressive (RCA) model, the GARCC model provides a motivation for the conditional correlations to be time varying. GARCC is also more general than the Engle (2002, *Journal of Business & Economic Statistics* 20, 339–350) dynamic conditional correlation (DCC) and the Tse and Tsui (2002, *Journal of Business & Economic Statistics* 20, 351–362) varying conditional correlation (VCC) models and does not impose unduly restrictive conditions on the parameters of the DCC model. The structural properties of the GARCC model, specifically, the analytical forms of the regularity conditions, are derived, and the asymptotic theory is established. The Baba, Engle, Kraft, and Kroner (BEKK) model of Engle and Kroner (1995, *Econometric Theory* 11, 122–150) is demonstrated to be a special case of a multivariate RCA process. A likelihood ratio test is proposed for several special cases

The authors thank the co-editor, Bruce Hansen, and three referees for insightful suggestions and Manabu Asai, Monica Billio, Massimiliano Caporin, Rob Engle, Shiqing Ling, Marcelo Medeiros, Adrian Pagan, Ruey Tsay, Jun Yu, seminar participants at the Institute of Economics, Academia Sinica, Taiwan, Bank of Italy, Chiang Mai University, Chinese University of Hong Kong, Complutense University, Ente Einaudi—Rome, Fondazione Eni Enrico Mattei—Milan, Hong Kong University of Science and Technology, Kyoto University, National University of Singapore, University of Adelaide, University of Auckland, University of Canterbury, University of Hong Kong, University of Milan, University of New South Wales, University of Padua, University Pompeu Fabra, University of Queensland, University of Venice "Ca' Foscari," and University of Vigo, and participants at the New Frontiers in Financial Volatility Modelling Conference, Florence, Italy, May 2003, and the New Frontiers of Statistics Workshop, Chinese Academy of Sciences, Beijing, China, July 2005, for helpful comments and discussion. The authors acknowledge the financial support of the Australian Research Council. Address correspondence to Offer Lieberman, Department of Economics, University of Haifa, Haifa 31905, Israel; e-mail: offerl@econ.haifa.ac.il.

 © 2008 Cambridge University Press 0266-4666/08 $15.00

https://doi.org/10.1017/S0266466608080614 Published online by Cambridge University Press

of GARCC. The empirical usefulness of GARCC and the practicality of the likelihood ratio test are demonstrated for the daily returns of the Standard and Poor's 500, Nikkei, and Hang Seng indexes.

1. INTRODUCTION

The empirical usefulness of the Engle (1982) autoregressive conditional heteroskedasticity (ARCH) model and the Bollerslev (1986) extension of ARCH to the generalized ARCH (GARCH) model have inspired a new generation of models to capture time-varying conditional volatility in financial time series data. Such success in modeling conditional volatility has led to several different research developments. One area of research concerns the flexibility of the models to capture the stylized facts or features that appear in financial time series data. This direction has led to several useful asymmetric extensions of GARCH, including the Glosten, Jagannathan, and Runkle (1992) asymmetric (or threshold) GARCH (GJR) model, the Nelson (1991) exponential GARCH (EGARCH) model, and the Ding, Granger, and Engle (1993) asymmetric power GARCH (APGARCH) model. Although these univariate models can capture the excessive kurtosis and asymmetric behavior that are often found in financial time series, they do not analyze interdependent (or spillover) effects in volatility across different markets or assets.

To accommodate spillovers in conditional volatility, several different multivariate GARCH models have been proposed. These include (a) the diagonal model of Bollerslev, Engle, and Wooldridge (1988); (b) the vech, diagonal vech, and Baba, Engle, Kraft, and Kroner (BEKK) multivariate GARCH models of Engle and Kroner (1995), which model the conditional covariances; (c) the constant conditional correlation (CCC) multivariate GARCH model of Bollerslev (1990), the Ling and McAleer (2003) vector ARMA-GARCH (VARMA-GARCH) model and the VARMA–asymmetric GARCH (VARMA-AGARCH) model of McAleer, Hoti, and Chan (2008), which assume conditional correlations to be constant; and (d) the Engle (2002) dynamic conditional correlation (DCC) and Tse and Tsui (2002) varying conditional correlation (VCC) models, which relaxed the assumption of CCCs and model the dynamics of the conditional correlations and conditional variances jointly. With the exception of the Tse and Tsui (2002) VCC, each of these models has been programmed in standard econometric software packages such as RATS Version 6.

Parts of the multivariate GARCH literature have evolved without establishing the structural and/or statistical properties of the underlying models. Important theoretical properties include sufficient conditions for stationarity and ergodicity, sufficient conditions for the existence of moments, and sufficient conditions for consistency and asymptotic normality of the quasi–maximum likelihood estimators (QMLEs). Satisfying such conditions allows valid inferences to be drawn and facilitates the subsequent testing of restrictions and diagnostic testing of the auxiliary assumptions.

https://doi.org/10.1017/S0266466608080614 Published online by Cambridge University Press

The regularity conditions and asymptotic properties are typically either assumed to hold or sufficient regularity conditions are established to ensure that the asymptotic theory is applicable. Exceptions to the rule are the VARMA-GARCH and VARMA-AGARCH models, for which the structural properties of the models have been developed, the analytical forms of the regularity conditions have been derived, and the asymptotic theory for the QMLE has been established under the second and fourth moments for consistency and asymptotic normality, respectively, and the BEKK model, for which Comte and Lieberman (2003) showed consistency of the QMLE using the conditions established in Jeantheau (1998), and asymptotic normality of the QMLE by assuming the existence of eighth moments.

Multivariate GARCH models have not yet enjoyed the same popularity in practice as have their univariate counterparts. Four major inhibiting factors would seem to be (a) the lack of a theoretical foundation for some of the multivariate models; (b) the interpretations of the multivariate models, which are not always straightforward; (c) the number of parameters in the multivariate models, which can increase significantly with the number of assets and/or markets; and (d) the large number of parameters, which can lead to computational problems (for useful comparisons of alternative multivariate GARCH models, see, e.g., Kroner and Ng, 1998; McAleer et al., 2008). Estimating the large number of parameters jointly can cause numerical problems, and not all multivariate GARCH models are able to accommodate convenient two-step estimation methods. For further details regarding the theoretical and computational properties of univariate and multivariate, conditional and stochastic, volatility models, see McAleer (2005).

Apart from the practical difficulties in estimating the parameters, several of the recently proposed multivariate GARCH models have assumed that the conditional correlations of the standardized shocks remain constant over time. Alternatively, some multivariate GARCH models, such as BEKK, model the conditional covariances, so that the conditional correlations are inherently time varying. Recently, Engle (2002) and Tse and Tsui (2002) proposed closely related multivariate GARCH models with time-varying conditional correlations. Under certain simplifying assumptions, the DCC model of Engle (2002) can be estimated using a two-step procedure, which has resolved some numerical difficulties associated with estimating multivariate GARCH models. Furthermore, both DCC and the Tse and Tsui (2002) VCC multivariate GARCH model lack explicit structural properties and impose strong cross-equation parametric restrictions on the conditional correlations. Although this can be beneficial in terms of reducing the number of parameters to be estimated, the effects of such restrictions when they are false are unknown. Thus, a test for the validity of such restrictions is warranted.

The primary purpose of this paper is to develop a new generalized autoregressive conditional correlation (GARCC) model when the standardized residuals follow a random coefficient vector autoregressive (VAR) process. The GARCC model is a multivariate generalization of the Tsay (1987) random coefficient

https://doi.org/10.1017/S0266466608080614 Published online by Cambridge University Press

autoregressive (RCA) model and provides a motivation for the conditional correlations to be time varying. GARCC is also more general than the DCC and VCC models, as it does not impose unduly restrictive conditions on the parameters of the conditional correlation model. The structural and statistical properties of GARCC will also be derived, which include sufficient conditions for the existence of moments and sufficient conditions for consistency and asymptotic normality of the QMLE. Furthermore, a formal test is proposed of a variety of cross-equation restrictions that can be imposed on GARCC. Empirical results using daily returns of the Standard and Poor's 500 (S&P), Nikkei, and Hang Seng indexes from 1 January 1986 to 11 April 2000 show that these cross-equation restrictions are violated. Therefore, GARCC would seem to be a useful addition to the multivariate GARCH literature for modeling time-varying conditional correlations.

The plan of the remainder of the paper is as follows. Section 2 presents the GARCC model and provides a structural justification of the BEKK model. The structural properties are developed, the analytical forms of the regularity conditions are derived, the asymptotic theory is established, and a likelihood ratio (LR) test is proposed for several special cases of GARCC in Section 3. The empirical usefulness of GARCC and the practicality of the LR test are demonstrated for the daily returns in the S&P, Nikkei, and Hang Seng indexes in Section 4. Section 5 provides some concluding comments. Proofs of the proposition and theorems are given in the Appendix.

2. MULTIVARIATE GARCH MODELS

Consider the following multivariate GARCH model:

$$\begin{aligned} Y_t &= \mathrm{E}_{Y,t-1}(Y_t) + \varepsilon_t, \qquad t = 1, \ldots, T, \\ \varepsilon_t &= D_t^{1/2} \eta_t, \end{aligned} \tag{1}$$

where $Y_t = (y_{1t}, \ldots, y_{kt})'$, $\varepsilon_t = (\varepsilon_{1t}, \ldots, \varepsilon_{kt})'$, $\eta_t = (\eta_{1t}, \ldots, \eta_{kt})'$, $D_t = \mathrm{diag}(h_{1t}, \ldots, h_{kt})$ is a diagonal matrix of conditional variances and the expectation is taken with respect to $\{Y_\tau : \tau \leq t\}$. Let $H_t = (h_{1t}, \ldots, h_{kt})'$ be a $k \times 1$ vector of conditional variances. Ling and McAleer (2003) proposed the VARMA-GARCH model as follows:

$$\left(I - \sum_{i=1}^{p} \Phi_i L^i\right) Y_t = \left(I - \sum_{i=1}^{q} \Theta_i L^i\right) \varepsilon_t, \tag{2}$$

where Φ_i and Θ_i are $k \times k$ matrices, the roots of the characteristic polynomials $|I_k - \sum_{i=1}^{p} \Phi_i L^i|$ and $|I_k - \sum_{i=1}^{q} \Theta_i L^i|$ lie outside the unit circle, and L is the lag operator. Furthermore, the conditional variance vector is given as

$$H_t = W + \sum_{i=1}^{r} A_i \vec{\varepsilon}_{t-i} + \sum_{j=1}^{s} B_j H_{t-j}, \tag{3}$$

https://doi.org/10.1017/S0266466608080614 Published online by Cambridge University Press

where $\vec{\varepsilon}_t = (\varepsilon_{1t}^2, \ldots, \varepsilon_{kt}^2)'$, W is a $k \times 1$ vector, and A_i and B_j are $k \times k$ matrices, $i = 1, \ldots, r$ and $j = 1, \ldots, s$. Ling and McAleer (2003) established the structural and statistical properties, including the sufficient conditions for the existence of moments and the sufficient conditions for consistency and asymptotic normality of the QMLE for the VARMA-GARCH model. Assuming that $\mathrm{E}_{\eta, t-1}(\eta_t \eta_t') = Q_t$, under the strict minimum phase condition the conditional covariance matrix is given by

$$\Sigma_t \equiv \mathrm{E}_{Y, t-1}(\varepsilon_t \varepsilon_t') = D_t^{1/2} Q_t D_t^{1/2}.$$

In other words,

$$Q_t = D_t^{-1/2} \Sigma_t D_t^{-1/2}.$$

If the η_t's are independently and identically distributed (i.i.d.) with $Q_t = \Gamma_t$, then ε_t has a constant conditional correlation. If the A_i's and B_j's are diagonal matrices and $\eta_t \sim iid(0, \Gamma)\ \forall t$, then (1)–(3) reduce to the CCC model of Bollerslev (1990).

The BEKK model, among others, is concerned with modeling directly the conditional covariance matrix, Σ_t, for which the (implied) conditional correlation matrix is also dynamic. For this model, the conditional covariance and correlation matrices are given by

$$\Sigma_t = \Sigma + A \varepsilon_{t-1} \varepsilon_{t-1}' A' + B \Sigma_{t-1} B'$$

and

$$\Gamma_t = \mathrm{diag}(\Sigma_t)^{-1/2} \Sigma_t \, \mathrm{diag}(\Sigma_t)^{-1/2},$$

which is typically not a constant matrix.

Using an extension of the Tsay (1987) univariate RCA process, BEKK can be derived from a vector RCA (VRCA) process, as shown in the following proposition.

PROPOSITION 1.

(i) For the process

$$\varepsilon_t = \sum_{i=1}^{p} A_{it} \varepsilon_{t-i} + \eta_t,$$

where ε_t *and* η_t *are* $k \times 1$, $\eta_t \sim iid(0, \Gamma)$, $\forall t$, *the* A_{it}*'s* $(i = 1, \ldots, p)$, $(t = 1, 2, \ldots)$ *satisfy*

$$\begin{aligned}
&\mathrm{E}_{\varepsilon, t-1}(A_{it}) = 0, \quad \forall i, t, \\
&\mathrm{E}_{\varepsilon, t-1}(A_{it, j_1, l_1} A'_{it, j_2, l_2}) = A_{i, j_1, l_1} A'_{i, j_2, l_2} \qquad (j_1, j_2, l_1, l_2 = 1, \ldots, k), \quad \textbf{(4)} \\
&\mathrm{E}_{\varepsilon, t-1}(A_{it, j_1, l_1} A_{js, j_2, l_2}) = 0, \qquad \textit{if} \quad i \neq j \quad \textit{and/or} \quad t \neq s \\
&\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad (j_1, j_2, l_1, l_2 = 1, \ldots, k),
\end{aligned}$$

https://doi.org/10.1017/S0266466608080614 Published online by Cambridge University Press

and η_t *and* A_{it} *are independent* $\forall i$ *and* $\forall t$*, the conditional variance of* ε_t *is*

$$\Sigma_t = \mathrm{E}_{\varepsilon, t-1}(\varepsilon_t \varepsilon_t') = \sum_{i=1}^{p} A_i \varepsilon_{t-i} \varepsilon_{t-i}' A_i' + \Gamma.$$

(ii) The finite-order random coefficient moving average process

$$\varepsilon_t = \eta_t + \sum_{j=1}^{M} D_{jt} \eta_{t-j},$$

with the condition that the roots of $|I - \sum_{j=1}^{M} D_{jt} L^j|$ *lie outside the unit circle with probability 1, is stationary and invertible, and there exists an* equivalent *representation*

$$\varepsilon_t = \sum_{i=1}^{\infty} A_{it} \varepsilon_{t-i} + \eta_t. \tag{5}$$

(iii) The model (5) with the conditions (4) with indexes extending to infinity has a conditional variance given by the BEKK model

$$\Sigma_t = \Sigma + A\varepsilon_{t-1} \varepsilon_{t-1}' A' + B\Sigma_{t-1} B'$$

as long as the roots of $|I - B \otimes B|$ *lie outside the unit circle.*

Proof. See the Appendix.

Remark. In conjunction with the theoretical results in Comte and Lieberman (2003), Proposition 1 provides strong justification for using BEKK to model the conditional covariances directly. (See also Scherrer and Ribarits, 2007.)

An alternative approach to accommodate time-varying conditional correlations is to allow the conditional covariance of the standardized residuals, Q_t, to change over time. However, to maintain H_t as the vector of conditional variances for ε_t, Q_t must satisfy the definition of a conditional correlation matrix. Therefore, we must have

$$\mathrm{E}_{Y, t-1}(\eta_{it}^2) = 1, \quad \forall t = 1, \ldots, T, \quad \forall i = 1, \ldots, k.$$

For this reason, Engle (2002) modeled the time-varying conditional correlation in DCC by defining the process of a $k \times k$ matrix P_t as

$$P_t = (1 - \phi_1 - \phi_2)P + \phi_1 \eta_{t-1} \eta_{t-1}' + \phi_2 P_{t-1}, \tag{6}$$

and with this, the $k \times k$ conditional correlation matrix Q_t is

$$Q_t = \mathrm{diag}(P_t)^{-1/2} P_t \, \mathrm{diag}(P_t)^{-1/2}.$$

https://doi.org/10.1017/S0266466608080614 Published online by Cambridge University Press

Furthermore, using the multivariate GARCH model in Ding and Engle (2001), Engle (2002) suggested a more general specification of P_t as

$$P_t = (i_k i_k' - \Phi_1 - \Phi_2) \circ P + \Phi_1 \circ \eta_{t-1}\eta_{t-1}' + \Phi_2 \circ P_{t-1}, \tag{7}$$

where i_k is the $k \times 1$ unit vector and P is the CCC matrix of η_{it} if Φ_1 and Φ_2 are equal to the null matrix. It is clear that (6) is nested within (7). Although Engle (2002) proposed (7) as an alternative to (6), no formal statistical test has been derived for determining the appropriate specification. Moreover, Engle (2002) did not demonstrate how Q_t could vary over time, and it is difficult to interpret P_t as it is unclear how equation (7) is derived.

In the following sections, GARCC is derived when the standardized residuals are assumed to follow a random coefficient VAR process. The structural properties of the GARCC model are developed, the analytical forms of the regularity conditions are derived, and the asymptotic theory is established.

2.1. The Generalized Autoregressive Conditional Correlation Model

In a model such as (1), the η_t's are typically assumed to be either an i.i.d. random vector or a martingale difference process. Not only is the i.i.d. condition often violated in practice, it also requires the conditional correlation matrix to be static, thereby denying the possibility of dynamic conditional correlations. We propose a specification that addresses this point. Let

$$\eta_{it} = \sum_{l=1}^{L} \phi_{ilt}\nu_t\eta_{it-l} + \xi_{it}\nu_t, \tag{8}$$

where ϕ_{ilt} can variously be a constant, a deterministic function, or a random variable and $\nu_t \sim iid(0,1)$. Furthermore, let $\xi_t = (\xi_{1t}, \ldots, \xi_{kt})'$ and assume that $\xi_t \sim iid(0, \Omega)$. This structure is a multivariate extension of the RCA model in Tsay (1987) (for the structural and statistical properties of finite-order RCA process, see Nicholls and Quinn, 1980, 1981, 1982).

The GARCC model has the following specification:

$$Y_t = \mathrm{E}_{Y,t-1}(Y_t;\theta) + \varepsilon_t, \qquad (t = 1, \ldots, T),$$

$$\varepsilon_t = D_t^{1/2}\eta_t,$$

$$D_t = \mathrm{diag}(h_{1t}, \ldots, h_{kt}),$$

$$\eta_{it} = \left(\sum_{l=1}^{L} \phi_{ilt}\eta_{it-l} + \xi_{it}\right)\nu_t, \qquad (i = 1,..,k), \qquad \nu_t \sim iid(0,1), \quad \forall t,$$

https://doi.org/10.1017/S0266466608080614 Published online by Cambridge University Press

$$H_t = W + \sum_{i=1}^{r} A_i \vec{\varepsilon}_{t-i} + \sum_{j=1}^{s} B_j H_{t-j}, \tag{9}$$

$$\xi_t = (\xi_{1t}, \ldots, \xi_{kt}) \sim iid(0, \Omega),$$

$$\eta_t = (\eta_{1t}, \ldots, \eta_{kt})',$$

$$Q_t = \mathrm{E}_{\eta, t-1}(\eta_t \eta_t'),$$

$$\Gamma_t = \mathrm{diag}\{Q_t\}^{-1/2} Q_t \,\mathrm{diag}\{Q_t\}^{-1/2},$$

where $\theta \in \mathbb{R}^a$ is the parameter vector for the conditional first moment and a is the number of elements in θ. Because ν_t is $iid(0,1)$, $\mathrm{E}_{Y, t-1}(\eta_{it}) = \mathrm{E}_{\eta, t-1}(\eta_{it}) = 0$, and under suitable conditions on ϕ_{ilt}'s (see Cases 1–4, which follow), for all i and t,

$$\mathrm{E}_{Y, t-1}(\eta_{it}\eta_{js}) = \mathrm{E}_{\eta, t-1}(\eta_{it}\eta_{js}) = 0, \quad \text{if } t \neq s$$

and

$$\mathrm{E}_{Y, t-1}(\eta_{it}\eta_{js}) = \mathrm{E}_{\eta, t-1}(\eta_{it}\eta_{js}) = \sum_{l=1}^{L} \phi_{ilt}\phi_{jlt}\eta_{it-l}\eta_{jt-l} + \omega_{ij}, \quad \text{if } t = s,$$

with $\omega_{ij} = [\Omega]_{i,j}$. Note that if the diagonal elements of Q_t are not all equal to one, then Q_t is no longer a conditional correlation matrix and, more importantly, it implies that H_t is no longer a vector of conditional variances. In this case, the conditional variance of ε_t is $\mathrm{diag}(D_t^{1/2} Q_t D_t^{1/2})$, and the conditional correlation matrix, Γ_t, is obtained by standardizing Q_t. The random variable ν_t can be interpreted as a common, or global, shock to all markets, whereas ξ_{it} can be interpreted as a shock that is specific to market i.

This paper considers four specific cases of the GARCC model. Cases 1–3 are specializations of Case 4 and demonstrate that a time-varying conditional correlation model can be derived when the standardized residuals follow a VAR process with random coefficients.

The structural properties of the model are developed, the analytical forms of the regularity conditions are derived, and the asymptotic theory is established for Case 4. As specializations of Case 4, the theoretical results also apply to Cases 1–3.

2.2. Case 1. $\phi_{ilt} = \phi$, $L = 1$

In this case, $\phi_{ilt} = \phi$ with $|\phi| < 1$, so that

$$\eta_{it} = (\phi\eta_{it-1} + \xi_{it})\nu_t \quad \forall\, i = 1, \ldots, k. \tag{10}$$

https://doi.org/10.1017/S0266466608080614 Published online by Cambridge University Press

This is equivalent to a first-order RCA (RCA(1)) process where the random coefficients are the same across all i. It follows from (10) that

$$q_{ijt} = [Q_t]_{i,j} = \mathrm{E}_{\eta,t-1}(\eta_{it}\eta_{jt}) = \omega_{ij} + \phi^2\eta_{it-1}\eta_{jt-1}, \tag{11}$$

or in matrix form,

$$Q_t = \Omega + \phi^2\eta_{t-1}\eta'_{t-1}. \tag{12}$$

The parameter ϕ^2 denotes the short-run persistence of shocks to q_{ijt}, and the conditional variance of ε_t is given by

$$\mathrm{E}_{\varepsilon,t-1}(\varepsilon_{it}^2) = (1 + \phi^2\eta_{it-1}^2)h_{it}, \quad \forall t = 1,\ldots,T, \quad \forall i = 1,\ldots,k.$$

2.3. Case 2. $\phi_{ilt} = \phi_i$, $L = 1$

In this case, $\phi_{ilt} = \phi_i$ with $|\phi_i| < 1$ $\forall i$, so that

$$\eta_{it} = (\phi_i\eta_{it-1} + \xi_{it})\nu_t \quad \forall i = 1,\ldots,k. \tag{13}$$

This is a generalization of Case 1 where the variance of the random coefficients can be different in each $i = 1,\ldots,k$. It follows from (13) that

$$q_{ijt} = \mathrm{E}_{\eta,t-1}(\eta_{it}\eta_{jt}) = \omega_{ij} + \phi_{ij}\eta_{it-1}\eta_{jt-1},$$

where $\phi_{ij} = \phi_i\phi_j$ is a typical element of Φ, so that

$$Q_t = \Omega + \Phi \circ \eta_{t-1}\eta'_{t-1}. \tag{14}$$

The parameter ϕ_{ij} denotes the short-run persistence of shocks to q_{ijt}, and the conditional variance of ε_t is given by

$$\mathrm{E}_{\varepsilon,t-1}(\varepsilon_{it}^2) = (1 + \phi_i^2\eta_{it-1}^2)h_{it}, \quad \forall t = 1,\ldots,T, \quad \forall i = 1,\ldots,k.$$

2.4. Case 3. $\phi_{ilt} = \phi_1^{1/2}\delta_l$, $L = \infty$

Let $\delta_1 = 1$ and assume that $\delta_l \sim iid(0, \phi_2^{l-1})$ $\forall l \geq 2$, with $|\phi_1| < 1$, $|\phi_2| < 1$. We have

$$\eta_{it} = \left(\phi_1^{1/2}\sum_{l=1}^{\infty}\delta_l\eta_{it-l} + \xi_{it}\right)\nu_t, \quad \forall i = 1,\ldots,k. \tag{15}$$

As $\mathrm{E}(\delta_l\delta_m) = \mathrm{E}(\delta_l)\mathrm{E}(\delta_m) = 0$,

$$q_{ijt} = \mathrm{E}_{\eta,t-1}(\eta_{it}\eta_{jt}) = \omega_{ij} + \phi_1\sum_{l=1}^{\infty}\phi_2^{l-1}\eta_{it-l}\eta_{jt-l} \quad \forall i = 1,\ldots,k, \tag{16}$$

https://doi.org/10.1017/S0266466608080614 Published online by Cambridge University Press

so that

$$Q_t = \Omega + \phi_1 \sum_{l=1}^{\infty} \phi_2^{l-1} \eta_{t-l} \eta_{t-l}'. \tag{17}$$

The parameter $\phi_1 \phi_2^{l-1}$ denotes the short-run persistence of shocks to q_{ijt}, and the conditional variance of ε_t is given by

$$\mathrm{E}_{\varepsilon, t-1}(\varepsilon_{it}^2) = \left(1 + \phi_1 \sum_{l=1}^{\infty} \phi_2^{l-1} \eta_{it-l}^2\right) h_{it}, \quad \forall t = 1, \ldots, T, \quad \forall i = 1, \ldots, k.$$

2.5. Case 4. $\phi_{ilt} = \phi_{1i}\, \delta_{il}$, $L = \infty$

Let $\delta_{i1} = 1$, $\forall i = 1, \ldots, k$, and assume that for each $i = 1, \ldots, k$ and $l \geq 2$, $\delta_{il} \sim (0, \phi_{2i}^{l-1})$, with $|\phi_{1i}| < 1$ and $|\phi_{2i}| < 1$, $\forall i = 1, \ldots, k$. Furthermore, assume that for any $l \neq m$, δ_{il} is independent of δ_{im} and that $E(\delta_{il} \delta_{jl}) = (\phi_{2i} \phi_{2j})^{(l-1)/2}$, $\forall i, j = 1, \ldots, k$. We have

$$\eta_{it} = \left(\phi_{1i} \sum_{l=1}^{\infty} \delta_{il} \eta_{it-l} + \xi_{it}\right) \nu_t, \quad \forall i = 1, \ldots, k,$$

and

$$q_{ijt} = \mathrm{E}_{\eta, t-1}(\eta_{it} \eta_{jt}) = \omega_{ij} + \phi_{1ij} \left(\eta_{it-1} \eta_{jt-1} + \sum_{l=2}^{\infty} \phi_{2ij}^{l-1} \eta_{it-l} \eta_{jt-l}\right),$$

$$\forall i, j = 1, \ldots, k, \tag{18}$$

where $\phi_{1ij} = \phi_{1i} \phi_{1j}$ and $\phi_{2ij} = (\phi_{2i} \phi_{2j})^{1/2}$ are typical elements of Φ_1 and Φ_2, respectively, so that

$$Q_t = \Omega + \Phi_1 \circ \left(\eta_{t-1} \eta_{t-1}' + \sum_{l=2}^{\infty} \Phi_2^{l-1} \circ \eta_{t-l} \eta_{t-l}'\right). \tag{19}$$

Hence, $\phi_{1ij} \phi_{2ij}^{l-1}$ is the short-run persistence of shocks to q_{ijt}, and the conditional variance of ε_t is given by

$$\mathrm{E}_{\varepsilon, t-1}(\varepsilon_{it}^2) = \left[1 + \phi_{1i}^2 \left(\eta_{1t-1}^2 + \sum_{l=2}^{\infty} \phi_{2i}^{l-1} \eta_{it-l}^2\right)\right] h_{it},$$

$$\forall t = 1, \ldots, T, \quad \forall i = 1, \ldots, k.$$

Cases 3 and 4 demonstrate that a time-varying conditional correlation model can be derived from an infinite vector RCA process in the standardized residuals.

https://doi.org/10.1017/S0266466608080614 Published online by Cambridge University Press

3. STRUCTURAL AND STATISTICAL PROPERTIES

This section develops the structural properties of Case 4 of GARCC, derives the analytical forms of the sufficient conditions for the existence of moments, and establishes the sufficient conditions for consistency and asymptotic normality of the QMLE. As Cases 1–3 are specializations of Case 4, the structural and statistical properties of these special cases also follow. Finally, a LR test is proposed to test the parametric restrictions for various special cases of GARCC.

3.1. Theory

Denote the parameter vector

$\lambda = (\theta', \tau', \rho')'$,

$\tau = \text{vech}(W, A_1, \ldots, A_r, B_1, \ldots, B_s)$, $\rho = \text{vech}(\Omega, \Phi_1, \Phi_2)$, and the true parameter vector as λ_0. It is assumed that the parameter space Λ is a compact subspace of euclidean space such that λ_0 is an interior point in Λ. We do not consider in this paper the situation in which the parameter is on the boundary of the parameter space. In the case where more than one parameter is on the boundary, the asymptotic distribution of the maximum likelihood estimator is generally nonstandard (see Chant, 1974, eqn. (8); Andrews, 2001, eqns. (3.14) and (4.16)).

For each $\lambda \in \Lambda$, we make the following assumptions.

Assumption 1. $\text{E}_{Y,t-1}(Y_t)$ is stationary and possesses an $\Im_t$-measurable second-order stationary solution.

Assumption 2. All the elements of A_i and B_j are nonnegative, $i = 1, \ldots, r$, $j = 1, \ldots, s$; each element of W has positive lower and upper bounds over λ; and all the roots of $|I_k - \sum_{i=1}^{r} A_i L^i - \sum_{i=1}^{s} B_i L^i| = 0$ are outside the unit circle. Moreover, $I_k - \sum_{i=1}^{r} A_i L^i$ and $\sum_{i=1}^{s} B_i L^i$ are left coprime and satisfy other identifiability conditions given in Ling and McAleer (2003).

Assumption 3. The functions D and Q are such that $\forall \lambda \in \Lambda$ and $\forall \lambda_0 \in \Lambda$, $D_{t,\lambda} = D_{t,\lambda_0}$ and $Q_{t,\lambda} = Q_{t,\lambda_0}$, almost surely (a.s.), if and only if $\lambda = \lambda_0$.

Assumption 4. Φ_1 and Φ_2 are positive semidefinite matrices, and $(1 - \phi_{2i})\, \sigma_\nu^4 \sigma_{\xi i}^4 < 1 - (\phi_{1i} + \phi_{2i})$ $\forall i = 1, \ldots, k$ where $E(\nu_t^4) = \sigma_\nu^4$.

Assumption 5. $\forall i = 1, \ldots, k$, $\exists c_i \in (0,1)$ such that δ_{il} are independent random variables with support $[-c_i^l, c_i^l]$ $\forall l > 1$.

Note that Assumption 3 is an identifiability condition, analogous to Assumption A4 of Jeantheau (1998). Assumption 5 provides sufficient conditions to ensure the existence of a stationary solution to (19). The structural properties of the model are developed and the analytical forms of the regularity conditions are derived in Theorems 1 and 2, respectively.

https://doi.org/10.1017/S0266466608080614 Published online by Cambridge University Press

THEOREM 1. *Under Assumptions 1, 2, 4, and 5, Case 4 possesses an* $\Im_{Y,t}$*-measurable second-order stationary solution* $\{Y_t, \varepsilon_t, H_t\}$ *that is unique, given the* Y_t*, where* $\Im_{Y,t}$ *is a* σ*-field generated by* $\{Y_k : k \leq t\}$*. The solution* $\{H_t\}$ *has the following causal representation:*

$$H_t = W + \sum_{j=1}^{\infty} c' \left(\prod_{i=1}^{j} \tilde{A}_{t+1-i} \right) \pi_{t-j}, \quad a.s., \tag{20}$$

where $\pi_t = [(\tilde{\eta}_t W)', 0, \ldots, 0, W', 0, \ldots, 0]'_{(r+s)k \times 1}$*, that is, the subvector consisting of the first* k *components is* $\tilde{\eta}_t W$ *and the subvector consisting of the* $(rk+1)$*th to* $(r+1)$*mth components is* W*;* $\tilde{\eta}_t = \nu_t^2 \operatorname{diag}(\eta_{1t}^2, \ldots, \eta_{kt}^2)$*,* $c' = (0, \ldots, 0, I_k, 0, \ldots, 0)_{m \times (r+s)k}$*, with the subvector consisting of the* $(rk+1)$*th to* $(r+1)$*kth columns being* I_k*, and*

$$\tilde{A}_t = \left(\begin{array}{ccc|ccc} \tilde{\eta}_t A_1 & \cdots & \tilde{\eta}_t A_r & \tilde{\eta}_t B_1 & \cdots & \tilde{\eta}_t B_s \\ & I_{k(r-1)} & O_{k(r-1)\times k} & & O_{k(r-1)\times ks} & \\ \hline A_1 & \cdots & A_r & B_1 & \cdots & B_s \\ & O_{k(s-1)\times kr} & & & I_{k(s-1)} & O_{k(s-1)\times k} \end{array} \right).$$

Proof. See the Appendix.

THEOREM 2. *Let the assumptions of Theorem 1 hold. If* $\rho[\mathrm{E}(\tilde{A}_t^{\otimes m})] < 1$*, with* m *being a strictly positive integer, then the* $2m$*th moments of* $\{Y_t, \varepsilon_t\}$ *are finite, where* $\tilde{A}_t$ *is defined as in Theorem 1 and* $A^{\otimes m}$ *is the Kronecker product of the* m *matrices* A*.*

Proof. See the Appendix.

Given these structural properties, the statistical properties of the model are established in Theorems 3–5, with sufficient multivariate log-moment conditions for consistency in Theorem 3, sufficient second-order moment conditions for consistency in Theorem 4, and sufficient conditions for asymptotic normality in Theorem 5.

The QMLE of the model is obtained by maximizing, conditional on the initial values of $(Y_t, \varepsilon_t, Q_t)$, the following log-likelihood function:

$$L_T(\lambda) = \frac{1}{T} \sum_{t=1}^{T} l_t(\lambda),$$

$$l_t(\lambda) = -\frac{1}{2} \left(\ln |D_t^{1/2} Q_t D_t^{1/2}| + \varepsilon_t' (D_t^{1/2} Q_t D_t^{1/2})^{-1} \varepsilon_t \right), \tag{21}$$

https://doi.org/10.1017/S0266466608080614 Published online by Cambridge University Press

where $l_t(\lambda)$ takes the form of the Gaussian log-likelihood function, so that the QMLE is given as

$$\hat{\lambda} = \underset{\lambda \in \Lambda}{\operatorname{argmax}}\, L_T(\lambda). \tag{22}$$

Maximization of (21) leads to the following consistency result.

THEOREM 3. *Denote $\hat{\lambda}$ as the QMLE of λ. Under Conditions C1–C6 in the Appendix, $\hat{\lambda} \rightarrow_p \lambda$.*

Proof. See the Appendix.

An alternative proof of consistency of the QMLE based on second moments is to verify the sufficient conditions of Theorem 4.1.1 in Amemiya (1985), as demonstrated for the VARMA-GARCH model in Ling and McAleer (2003).

THEOREM 4. *Denote $\hat{\lambda}$ as the QMLE of λ_0. Under Conditions D1–D3 in the Appendix, $\hat{\lambda} \rightarrow_p \lambda_0$.*

Proof. See the Appendix.

Given the consistency of $\hat{\lambda}$, the following theorem provides sufficient conditions for asymptotic normality.

THEOREM 5. *Let Y_t be generated by model (9), with ϕ_{ilt} as defined in Case 4. Given the consistency of $\hat{\lambda}$ for λ_0, under Conditions E1–E3 in the Appendix, $\sqrt{T}(\hat{\lambda} - \lambda_0) \xrightarrow{d} N(0, \Sigma_0^{-1} \Omega_\lambda \Sigma_0^{-1})$.*

Proof. See the Appendix.

3.2. Tests of Restrictions

In view of (16) and (18), Case 3 is a restricted version of Case 4 with $\phi_{1i} = \phi_1$ and $\phi_{2i} = \phi_2$, $\forall i = 1, \ldots, k$. It is important to test whether these restrictions hold in practice.

Formally, the null hypothesis of Case 3 against GARCC in Case 4 is given by

$$H_0 : \phi_{1ij} = \phi_1, \qquad \phi_{2ij} = \phi_2, \quad \forall i, j = 1, \ldots, k. \tag{23}$$

As Φ_1 and Φ_2 are $k \times k$ symmetric matrices, when ε_t is normally distributed, the LR statistic is given by

$$2T(L_T(\hat{\lambda}^U) - L_T(\hat{\lambda}^R)) \overset{a}{\sim} \chi^2(k(k-1) - 2), \tag{24}$$

where $\hat{\lambda}^U$ and $\hat{\lambda}^R$ denote the unrestricted and restricted estimates, respectively, and $k(k-1) - 2$ is the number of parametric restrictions in (23).

https://doi.org/10.1017/S0266466608080614 Published online by Cambridge University Press

The usefulness of the LR test in (24) is demonstrated through an empirical example in the next section using Cases 3 and 4 of GARCC.

4. EMPIRICAL EXAMPLE

4.1. Estimation of DCC and GARCC

This section estimates DCC and various special cases of GARCC using the S&P, Nikkei, and Hang Seng daily returns. The daily data were obtained through DataStream DataBase Services, with the sample from 1 January 1986 to 11 April 2000, giving 3,726 observations. The sample volatilities for each of the returns are calculated as

$$v_t = (y_t - \bar{y})^2,$$

where y_t denotes the stock return (or the log-difference of the stock index) and $\bar{y}$ is the sample mean of y_t.

The sample standard deviations for the returns of S&P, Nikkei, and Hang Seng are 0.0102, 0.0137, and 0.0176, respectively. Thus, S&P appears to be less volatile than either Nikkei or Hang Seng, especially during the early to late 1990s before the Asian economic and financial crises. Although Nikkei seems to be more volatile than S&P, the volatility is relatively low compared with Hang Seng, which seems to be the most volatile. Nikkei seems to have more positive shocks than S&P and Hang Seng. There are some obvious outliers and extreme observations for all three indexes, with Hang Seng appearing to have the largest number of outliers. S&P appears to have fewer extreme observations and outliers than Nikkei and Hang Seng.

An obvious similarity among the three indexes is the enormous decrease in returns at observation 474, which corresponds to the share market crash in October 1987. This is also the most obvious outlier in the three indexes. The second largest decrease in returns is observation 894 for Hang Seng, which corresponds with the Tianenman Square incident in Beijing on 4 June 1989. It is evident that the three markets can be affected by some common factors or shocks (ν_t), such as the share market crash in October. However, outliers such as 4 June 1989 affected only Hong Kong and hence are market specific (ξ_{it}). The distinction between such outliers can then be characterized by the distributions associated with ν_t and ξ_{it}.

In this example, the conditional mean of the GARCC model is assumed to be an AR(1) process for each return, that is,

$$y_{it} = \theta_{i0} + \theta_{i1} y_{it-1} + \varepsilon_{it}.$$

Furthermore, $r = s = 1$, with both A_1 and B_1 being diagonal matrices. This is equivalent to having the returns follow a univariate GARCH(1,1) process.

https://doi.org/10.1017/S0266466608080614 Published online by Cambridge University Press

Estimates for the DCC model can be found in Table 1. All three returns exhibit strong long-run persistence as $\alpha + \beta$ is close to one in all cases. Furthermore, the log-moment condition, $\mathrm{E}[\log(\alpha\eta_{it}^2 + \beta)] < 0$, $i = 1, \ldots, k$, is satisfied in all cases, which suggests that the QMLE are consistent and asymptotically normal.

Moreover, $\hat{\phi}_1$ and $\hat{\phi}_2$ in DCC are statistically significant, as shown in Table 1. This indicates that the conditional correlations between returns are not constant. Furthermore, the short-run persistence of shocks to the time-varying conditional correlations is statistically significant at the 0.01 level, albeit at a small value, with long-run persistence of shocks to the conditional correlations being statistically significant at $0.993 = 0.01 + 0.983$. The corresponding estimates for Cases 3 and 4 of GARCC are given in Tables 2 and 3, respectively.

As shown in Table 2, the estimates of the conditional mean and the GARCH component for GARCC Case 3 are very similar to those of DCC, with the exception of the $\hat{\beta}$ estimates. In the case of GARCC, the $\hat{\beta}$ estimates are typically much lower than their DCC counterparts, which implies that GARCC Case 3 suggests a lower long-run persistence in the conditional variance contributed from the GARCH component.

TABLE 1. DCC

Parameter estimates	S&P	Nikkei	Hang Seng
$\hat{\theta}_0$	8.683	10.867	15.824
	(6.991)	(11.953)	(8.883)
$\hat{\theta}_1$	0.022	−0.007	0.08
	(1.356)	(−0.428)	(4.879)
$\hat{\omega}$	1.09E-6	2.823E-6	6.815E-6
	(3.881)	(5.017)	(7.614)
$\hat{\alpha}$	0.074	0.133	0.145
	(17.99)	(21.654)	(26.16)
$\hat{\beta}$	0.919	0.866	0.844
	(195.653)	(153.237)	(171.804)
Log-moment	−0.007	−0.012	−0.019
Second moment	0.993	0.999	0.988
	Estimates of the P_t equation		
$\hat{\phi}_1$	0.010		
	(10.615)		
$\hat{\phi}_2$	0.983		
	(533.51)		
Mean log-likelihood	−9.121		

Note: The two entries for each parameter are their respective estimate and the Bollerslev and Wooldridge (1982) robust *t*-ratio.

https://doi.org/10.1017/S0266466608080614 Published online by Cambridge University Press

TABLE 2. GARCC: Case 3 with $\phi_{ilt} = \phi_1^{1/2}\delta_l$

Parameter estimates	S&P	Nikkei	Hang Seng
$\hat{\theta}_0$	8.494	10.402	14.998
	(7.518)	(13.626)	(8.912)
$\hat{\theta}_1$	0.023	−0.006	0.079
	(1.422)	(−0.352)	(4.753)
$\hat{\omega}$	1.068E-5	1.225E-5	1.601E-5
	(4.771)	(4.811)	(7.117)
$\hat{\alpha}$	0.070	0.146	0.135
	(15.776)	(19.029)	(18.833)
$\hat{\beta}$	0.790	0.795	0.823
	(53.039)	(77.784)	(152.197)
Log-moment	−0.069	−0.043	−0.037
Second moment	0.860	0.940	0.958
	Estimates of the Q_t equation		
$\hat{\phi}_1$	0.014		
	(11.341)		
$\hat{\phi}_2$	0.980		
	(617.119)		
Mean log-likelihood	−9.031		

Note: The two entries for each parameter are their respective estimate and the Bollerslev and Wooldridge (1982) robust *t*-ratio.

The estimates $\hat{\phi}_1$ and $\hat{\phi}_2$ are both statistically significant, with short-run persistence of 0.014 and long-run persistence of 0.994 = 0.014 + 0.980. This also indicates that the conditional correlations between the returns vary over time.

As shown in Table 3, the estimates in the conditional mean for GARCC Case 4 are similar to the other two models. The α (β) estimates in the GARCH component are generally lower (higher) than for the other two cases. However, both the log-moment and second moment conditions are satisfied for GARCC Case 4, indicating that the QMLE are consistent and asymptotically normal.

The short-run persistence of shocks to the time-varying conditional correlations is generally statistically significant, namely, at 0.074 between S&P and Nikkei (SP, NK), 0.007 between S&P and Hang Seng (SP, HS), and 0.017 between Nikkei and Hang Seng (NK, HS), whereas the corresponding long-run persistence of shocks to the time-varying conditional correlations is 0.400 = 0.074 + 0.326 for (SP, NK), −0.05 = 0.007 − 0.057 for (SP, HS), and 0.169 = 0.017 + 0.152 for (NK, HS). It is interesting to note that neither $\hat{\phi}_{1,(NK,HS)}$ nor $\hat{\phi}_{2,(NK,HS)}$ is statistically significant, indicating that the conditional correlation between Nikkei and Hang Seng is likely to be constant over time, whereas the conditional correlations between S&P and Nikkei, and between S&P and Hang Seng, are time varying. This demonstrates the advantage of GARCC Case 4

https://doi.org/10.1017/S0266466608080614 Published online by Cambridge University Press

TABLE 3. GARCC: Case 4 with $\phi_{ilt} = \phi_{1i}\delta_{il}$

Parameter estimates	S&P	Nikkei	Hang Seng
$\hat{\theta}_0$	8.510	10.067	14.960
	(8.459)	(30.685)	(7.694)
$\hat{\theta}_1$	0.023	−0.011	0.093
	(1.409)	(−0.684)	(5.568)
$\hat{\omega}$	5.200E-6	1.831E-6	6.65E-6
	(3.658)	(4.520)	(6.645)
$\hat{\alpha}$	0.023	0.087	0.073
	(8.482)	(15.836)	(11.991)
$\hat{\beta}$	0.971	0.911	0.901
	(304.957)	(169.476)	(135.149)
Log-moment	−0.026	−0.077	−0.088
Second moment	0.943	0.837	0.817
	Estimates of the Q_t equation		
$\hat{\phi}_{1,(SP,SP)}$	0.056	$\hat{\phi}_{2,(SP,SP)}$	0.680
	(2.959)		(2.466)
$\hat{\phi}_{1,(SP,NK)}$	0.074	$\hat{\phi}_{2,(SP,NK)}$	0.326
	(6.297)		(2.494)
$\hat{\phi}_{1,(SP,HS)}$	0.007	$\hat{\phi}_{2,(SP,HS)}$	−0.057
	(1.724)		(−5.818)
$\hat{\phi}_{1,(NK,NK)}$	0.100	$\hat{\phi}_{2,(NK,NK)}$	0.198
	(6.561)		(0.999)
$\hat{\phi}_{1,(NK,HS)}$	0.017	$\hat{\phi}_{2,(NK,HS)}$	0.152
	(1.439)		(1.132)
$\hat{\phi}_{1,(HS,HS)}$	0.061	$\hat{\phi}_{2,(HS,HS)}$	0.772
	(7.809)		(9.243)
Mean log-likelihood	−9.020		

Note: The two entries for each parameter are their respective estimate and the Bollerslev and Wooldridge (1982) robust *t*-ratio.

over both GARCC Case 3 and DCC, as the latter two models do not permit the conditional correlations to vary between different pairs of stock indexes.

As the LR statistic in (24) is 81.97, it is clear that the null hypothesis in (23) is strongly rejected at conventional levels. Thus, short- and long-run persistence are not the same for shocks to the conditional correlations for the three indexes, so that Case 4 is superior to Case 3 and hence is also superior to Cases 1 and 2. Furthermore, Theorem 5 shows that the QMLE for Case 4 is asymptotically normal. Therefore, the results in Table 3 show that the conditional correlation between Nikkei and Hang Seng is constant over time, as $\hat{\phi}_{1,(NK,HS)}$ and $\hat{\phi}_{2,(NK,HS)}$ are not statistically significant. However, the conditional correlation between S&P and Hang Seng is time varying, with both $\hat{\phi}_{1,(SP,HS)}$ and $\hat{\phi}_{2,(SP,HS)}$ being statistically significant.

https://doi.org/10.1017/S0266466608080614 Published online by Cambridge University Press

TABLE 4. Summary statistics for conditional correlations between S&P and Nikkei

Model	Mean	Variance	Skewness	Kurtosis	Minimum	Maximum
DCC	0.268	0.008	0.885	6.037	0.003	0.688
GARCC Case 3	0.267	0.011	0.729	5.563	−0.052	0.742
GARCC Case 4	0.291	0.004	0.359	12.413	−0.266	0.934

The plot of the estimated conditional correlations between each return from each model can be found in Figures 1–3, with the summary statistics in Tables 4–6. The conditional correlations from DCC and GARCC Case 3 are very similar in all three cases. However, the conditional correlations from DCC Case 4 are different from the other two models, as both DCC and GARCC Case 3 have imposed unjustified restrictions on the parameters in the conditional correlation model.

The means of the conditional correlations between S&P and Nikkei are 0.268, 0.267, and 0.291 for DCC, GARCC Case 3, and GARCC Case 4, respectively. Although the means of the correlations are similar for the three models, the ranges of the conditional correlations for the three models are quite different. In the case of DCC, the range of conditional correlations between S&P and Nikkei is [0.003, 0.688], whereas the ranges of the conditional correlations for GARCC Case 3 and Case 4 are [−0.052, 0.742] and [−0.266, 0.934], respectively. It is interesting to note that some conditional correlations from GARCC are negative, whereas the conditional correlations estimated using DCC are all positive. Thus, hedging strategies derived under DCC and GARCC could be quite different.

Figure 2 contains the dynamic paths of the conditional correlations between S&P and Hang Seng for the three models. As in the previous case, the conditional correlations between S&P and Hang Seng from DCC and GARCC Case 3 are very similar. The mean correlations are 0.319, 0.318, and 0.319 for DCC, GARCC Case 3, and GARCC Case 4, respectively. The conditional correlations between S&P and Hang Seng for the three models are positive throughout

TABLE 5. Summary statistics for conditional correlations between S&P and Hang Seng

Model	Mean	Variance	Skewness	Kurtosis	Minimum	Maximum
DCC	0.319	0.008	−0.164	2.690	0.085	0.537
GARCC Case 3	0.318	0.011	−0.268	2.714	0.034	0.576
GARCC Case 4	0.319	0.001	−2.264	12.709	0.032	0.379

https://doi.org/10.1017/S0266466608080614 Published online by Cambridge University Press

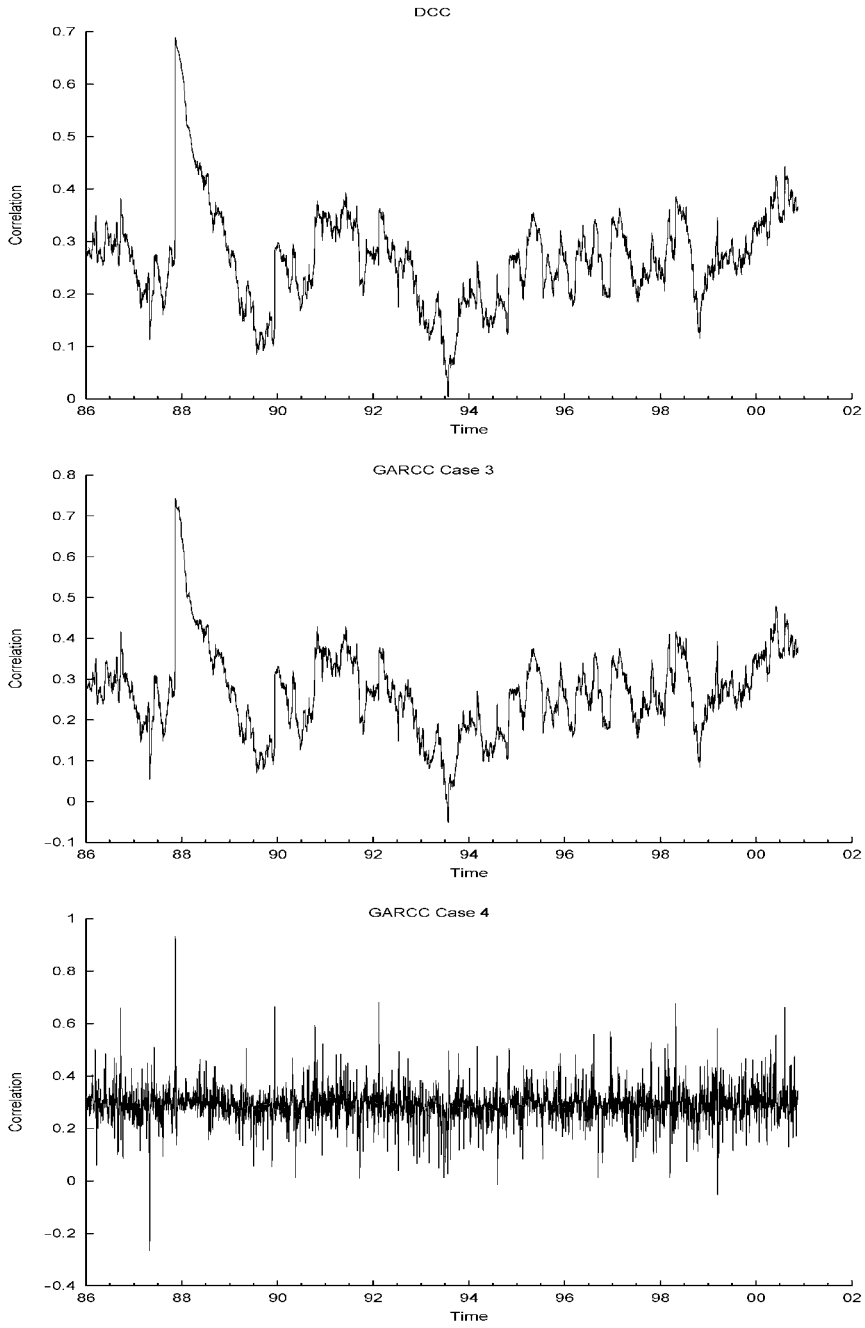


FIGURE 1. Conditional correlations between S&P and Nikkei.

https://doi.org/10.1017/S0266466608080614 Published online by Cambridge University Press

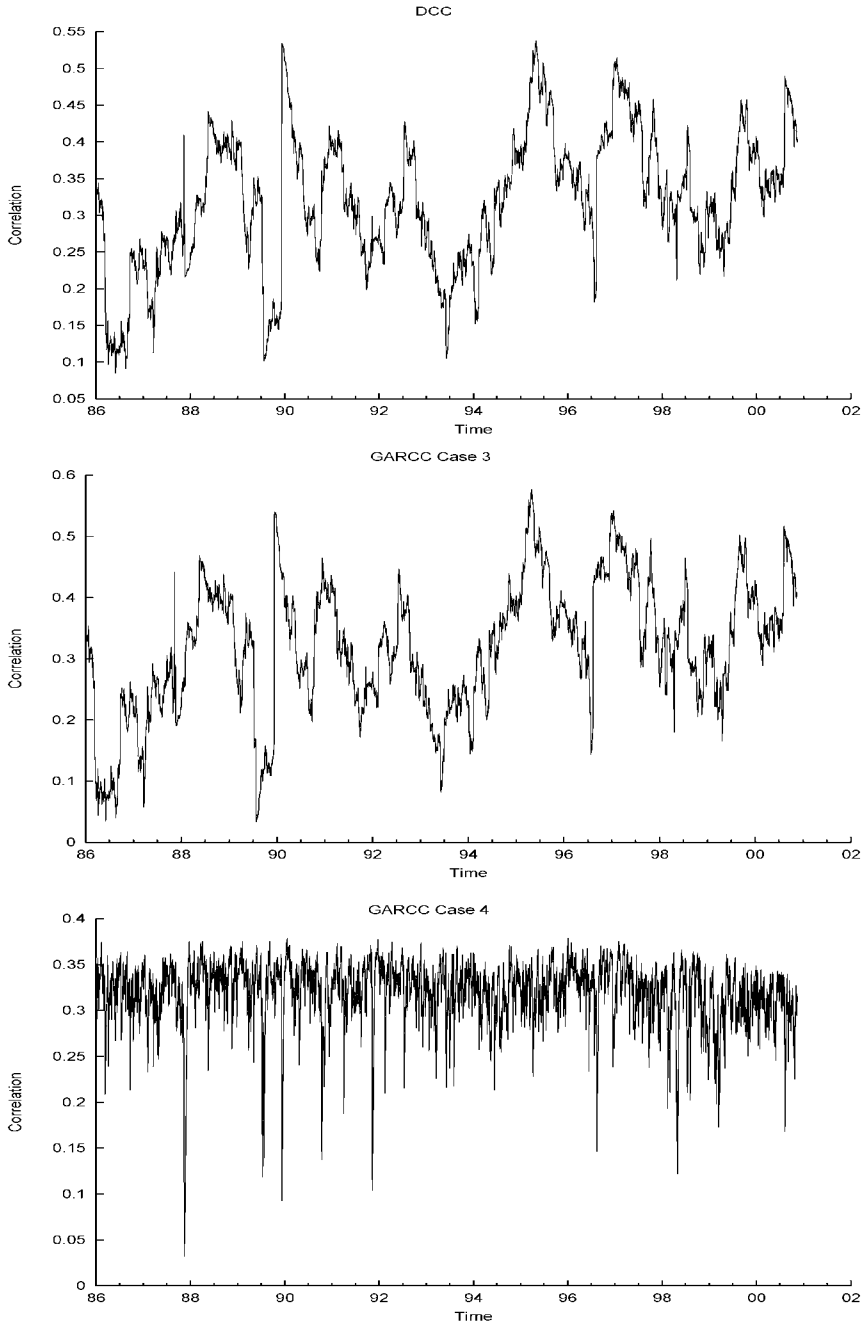


FIGURE 2. Conditional correlations between S&P and Hang Seng.

https://doi.org/10.1017/S0266466608080614 Published online by Cambridge University Press

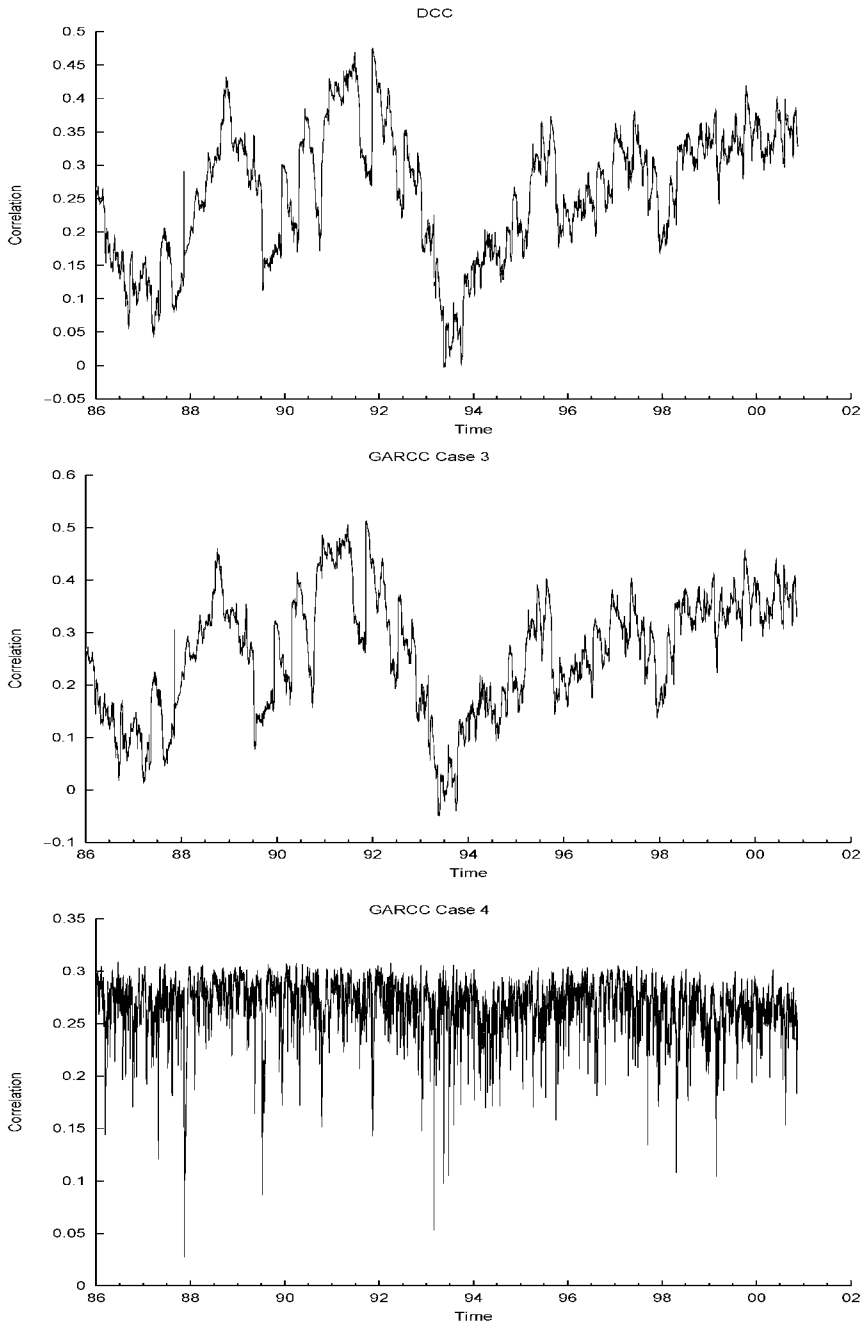


FIGURE 3. Conditional correlations between Nikkei and Hang Seng.

https://doi.org/10.1017/S0266466608080614 Published online by Cambridge University Press

TABLE 6. Summary statistics for conditional correlations between Nikkei and Hang Seng

Model	Mean	Variance	Skewness	Kurtosis	Minimum	Maximum
DCC	0.260	0.009	−0.308	2.434	−0.003	0.474
GARCC Case 3	0.261	0.013	−0.260	2.448	−0.049	0.512
GARCC Case 4	0.267	0.001	−2.172	12.031	0.027	0.308

the entire sample. However, GARCC Case 3 has the highest variance, with the range of the conditional correlations being [0.034, 0.576], followed by DCC with variance and range being 0.008 and [0.085, 0.537], respectively. GARCC Case 4 has the lowest variance at 0.001, with the range of conditional correlations being [0.032, 0.379].

The dynamic paths of the conditional correlations between Nikkei and Hang Seng can be found in Figure 3. As in the previous cases, the dynamic paths are very similar between DCC and GARCC Case 3. The means of the conditional correlations are 0.260, 0.261, and 0.267 for DCC, GARCC Case 3, and GARCC Case 4, respectively. However, some of the conditional correlations from DCC and GARCC Case 3 are negative, whereas the conditional correlations from GARCC Case 4 are all positive. The variance of the conditional correlations from GARCC Case 4 is still the lowest at 0.001, with the range given by [0.027, 0.308]. This is followed by DCC, with the variance and range of the conditional correlations being 0.009 and [−0.003, 0.474], respectively. The variance and range of the conditional correlations from GARCC Case 3 are 0.013 and [−0.049, 0.512], respectively.

5. CONCLUDING REMARKS

This paper presented and motivated a Generalized Autoregressive Conditional Correlation (GARCC) model when the standardized residuals follow a random coefficient VAR process. GARCC was developed as a multivariate generalization of the Tsay (1987) random coefficient autoregressive (RCA) model, which provided a motivation for the conditional correlations to be time varying. The GARCC model was shown to be more general than the Engle (2002) dynamic conditional correlation (DCC) and the Tse and Tsui (2002) varying conditional correlation (VCC) models, and GARCC does not impose strong parameteric restrictions on the conditional correlation model. The structural and statistical properties of the GARCC model were developed, which included the analytical forms of the regularity conditions and the asymptotic properties of the QMLE. This was particularly important as many of the multivariate GARCH models in the literature lack structural and statistical properties.

https://doi.org/10.1017/S0266466608080614 Published online by Cambridge University Press

A LR test was proposed for several special cases of GARCC. The empirical usefulness of GARCC and the practicality of the LR test were demonstrated for the daily returns of the S&P, Nikkei, and Hang Seng indexes. It was shown that standard cross-equation restrictions were violated. Thus, the test suggested in this paper should serve as a useful diagnostic check of various special cases of GARCC and also of the DCC and VCC models.

There are some useful extensions of GARCC for future research. Outliers could usefully be accommodated by choosing appropriate vector RCA process for the standardized residuals and by investigating the effects of extreme observations and outliers on the QMLE. Filtering common outliers from market-specific outliers would seem to raise a serious challenge. Another extension would be to incorporate time-varying nonlinear coefficients in the VAR process for the standardized residuals, which would be useful if regime switching behavior were to be present in the conditional correlations across assets or markets.

REFERENCES

Amemiya, T. (1985) *Advanced Econometrics*. Harvard University Press.

Andrews, D.W.K. (2001) Estimation when a parameter is on a boundary. *Econometrica* 67, 1341–1383.

Bollerslev, T. (1986) Generalized autoregressive conditional heteroskedasticity. *Journal of Econometrics* 31, 307–327.

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

Bollerslev, T., R. Engle, & J. Wooldridge (1988) A capital asset pricing model with time varying covariance. *Journal of Political Economy* 96, 116–131.

Bollerslev, T. & J. Wooldridge (1982) Quasi-maximum likelihood estimation and inference in dynamic models with time-varying covariances. *Econometric Reviews* 11, 143–173.

Chant, D. (1974) On asymptotic tests of composite hypothesis in nonstandard conditions. *Biometrika* 61, 291–298.

Comte, F. & O. Lieberman (2003) Asymptotic theory for multivariate GARCH processes. *Journal of Multivariate Analysis* 84, 61–84.

Ding, Z. & R. Engle (2001) Large scale conditional covariance matrix modeling, estimation and testing. *Academia Economic Papers* 29, 157–184.

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

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

Engle, R. (2002) Dynamic conditional correlation: A new simple class of multivariate GARCH models. *Journal of Business & Economic Statistics* 20, 339–350.

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

Glosten, L., R. Jagannathan, & D. Runkle (1992) On the relation between the expected value and volatility of nominal excess return on stocks. *Journal of Finance* 46, 1779–1801.

Jeantheau, T. (1998) Strong consistency of estimators for multivariate ARCH models. *Econometric Theory* 14, 70–86.

Kroner, K. & V. Ng (1998) Modeling asymmetric comovements of asset returns. *Review of Financial Studies* 11, 817–844.

https://doi.org/10.1017/S0266466608080614 Published online by Cambridge University Press

Ling, S. & M. McAleer (2003) Asymptotic theory for a vector ARMA-GARCH model. *Econometric Theory* 19, 278–308.
McAleer, M. (2005) Automated inference and learning in modeling financial volatility. *Econometric Theory* 21, 232–261.
McAleer, M., S. Hoti, & F. Chan (2008) Structure and asymptotic theory for multivariate asymmetric conditional volatility. *Econometric Reviews*, forthcoming.
Nelson, D. (1991) Conditional heteroscedasticity in asset returns: A new approach. *Econometrica* 59, 347–370.
Nicholls, D.F. & B.G. Quinn (1980) The estimation of random coefficient autoregressive models. I. *Journal of Time Series Analysis* 1, 37–46.
Nicholls, D.F. & B.G. Quinn (1981) The estimation of random coefficient autoregressive models. II. *Journal of Time Series Analysis* 2, 185–203.
Nicholls, D.F. & B.G. Quinn (1982) *Random Coefficient Autoregressive Models: An Introduction*. Lecture Notes in Statistics 11. Springer-Verlag.
Scherrer, W. & E. Ribarits (2007) On the parametrization of multivariate GARCH models. *Econometric Theory* 23, 464–484.
Stout, W. (1974) *Almost Sure Convergence*. Academic Press.
Tsay, R. (1987) Conditional heteroscedastic time series models. *Journal of the American Statistical Association* 82, 590–604.
Tse, Y. & A. Tsui (2002) A multivariate generalized autoregressive conditional heteroscedasticity model with time-varying correlations. *Journal of Business & Economic Statistics* 20, 351–362.
Tweedie, R. (1988) Invariant measure for Markov chains with no irreducibility assumptions. *Journal of Applied Probability* 25A, 275–285.

APPENDIX

Proof of Proposition 1. Under the assumptions of Proposition 1, the process

$$\varepsilon_t = \sum_{i=1}^{p} A_{it}\,\varepsilon_{t-i} + \eta_t \tag{A.1}$$

gives

$$\begin{aligned}(\mathrm{E}_{\varepsilon,t-1}(\varepsilon_t\varepsilon_t'))_{j_1,j_2} &= \left(\mathrm{E}_{\varepsilon,t-1}\left(\sum_{i=1}^{p}\sum_{m=1}^{p} A_{it}\,\varepsilon_{t-i}\,\varepsilon_{t-m}'A_{mt}'\right)\right)_{j_1,j_2} + \Gamma_{j_1,j_2}\\ &= \sum_{i=1}^{p}\sum_{m=1}^{p}\sum_{l_1=1}^{k}\sum_{l_2=1}^{k}(\varepsilon_{t-i}\,\varepsilon_{t-m}')_{l_1,l_2}\mathrm{E}_{\varepsilon,t-1}A_{it,j_1,l_1}A_{mt,l_2,j_2}' + \Gamma_{j_1,j_2}\\ &= \sum_{i=1}^{p}\sum_{l_1=1}^{k}\sum_{l_2=1}^{k}(\varepsilon_{t-i}\,\varepsilon_{t-i}')_{l_1,l_2}A_{i,j_1,l_1}A_{i,l_2,j_2}' + \Gamma_{j_1,j_2},\end{aligned}$$

or, in matrix form,

$$\mathrm{E}_{\varepsilon,t-1}(\varepsilon_t\varepsilon_t') = \sum_{i=1}^{p} A_i\,\varepsilon_{t-i}\,\varepsilon_{t-i}'A_i' + \Gamma.$$

https://doi.org/10.1017/S0266466608080614 Published online by Cambridge University Press

Part (ii) of the proposition is obvious. As for part (iii), for model (5), under the conditions (4), with the relevant indexes extending to infinity,

$$\Sigma_t \equiv \mathrm{E}_{\varepsilon, t-1}(\varepsilon_t \varepsilon_t') = \sum_{i=1}^{\infty} A_i \varepsilon_{t-i} \varepsilon_{t-i}' A_i' + \Gamma.$$

Therefore

$$\mathrm{vec}(\Sigma_t) = \sum_{i=1}^{\infty} (A_i \otimes A_i)\mathrm{vec}(\varepsilon_{t-i} \varepsilon_{t-i}') + \mathrm{vec}(\Gamma).$$

Now consider the BEKK model

$$\Sigma_t = \Sigma + A\varepsilon_{t-1} \varepsilon_{t-1}' A' + B\Sigma_{t-1} B'.$$

This is equivalent to

$$\mathrm{vec}(\Sigma_t) = \mathrm{vec}(\Sigma) + (A \otimes A)\mathrm{vec}(\varepsilon_{t-1} \varepsilon_{t-1}') + (B \otimes B)\mathrm{vec}(\Sigma_{t-1}). \tag{A.2}$$

Under the condition that the roots of $|I - (B \otimes B)|$ lie outside the unit circle, (A.2) is equivalent to

$$\begin{aligned}\mathrm{vec}(\Sigma_t) &= (I - (B \otimes B)L)^{-1}\mathrm{vec}(\Sigma) \\ &\quad + (I - (B \otimes B)L)^{-1}(A \otimes A)\mathrm{vec}(\varepsilon_{t-1} \varepsilon_{t-1}') \\ &= \mathrm{vec}(\Gamma) + \sum_{i=1}^{\infty} (A_i \otimes A_i)\mathrm{vec}(\varepsilon_{t-i} \varepsilon_{t-i}'),\end{aligned}$$

say. Therefore, we have established equivalence between the BEKK model and the infinite-order RCA. This completes the proof. ■

Proof of Theorem 1. Let $X_t = (\vec{\varepsilon}_t', \ldots, \vec{\varepsilon}_{t-r+1}', H_t', \ldots, H_{t-s+1}')'$. It is straightforward to show that

$$X_t = \tilde{A}_t X_{t-1} + \pi_t, \tag{A.3}$$

where π_t is defined as in (20). By recursive substitution of (A.3), it follows that

$$X_t = \pi_t + \sum_{j=1}^{\infty} \left(\prod_{i=1}^{j} \tilde{A}_{t+1-i} \right) \pi_{t-j}$$

and $H_t = c' X_t$, where $c' = (0, \ldots, 0, I_k, 0, \ldots, 0)_{k \times (r+s)k}$. Let

$$S_{T,t} = \pi_t + \sum_{j=1}^{T} \left(\prod_{i=1}^{j} \tilde{A}_{t+1-i} \right) \pi_{t-j}$$

and denote the mth element of $(\prod_{i=1}^{j} \tilde{A}_{t+1-i})\pi_{t-j}$ by $s_{T,t}$. Then

$$\mathrm{E}|s_{T,t}| = e_m' \mathrm{E}\left[\left(\prod_{i=1}^{j} \tilde{A}_{t+1-i} \right) \pi_{t-j} \right],$$

https://doi.org/10.1017/S0266466608080614 Published online by Cambridge University Press

where $e_m = (0,\ldots,0,1,0,\ldots,0)'_{k(r+s)\times 1}$ and 1 appears in the mth position. As η_t is not i.i.d., it follows that

$$\mathrm{E}|s_{T,t}| \neq e_m' \left(\prod_{i=1}^{j} \mathrm{E}(\tilde{A}_{t+1-i}) \right) \mathrm{E}(\pi_{t-j}).$$

Let F_t and Q be $k(r+s)$ matrices such that

$$F_t = \begin{pmatrix} \tilde{\eta}_t & 0_{k\times k(r+s-1)} \\ 0_{k(r+s-1)\times k} & I_{k(r+s-1)\times k(r+s-1)} \end{pmatrix}$$

and

$$Q = \left(\begin{array}{ccc|ccc} A_1 & \cdots & A_r & B_1 & \cdots & B_s \\ & I_{k(r-1)} & O_{k(r-1)\times k} & & O_{k(r-1)\times ks} & \\ \hline A_1 & \cdots & A_r & B_1 & \cdots & B_s \\ & O_{k(s-1)\times kr} & & & I_{k(s-1)} & O_{k(s-1)\times k} \end{array} \right).$$

Then $\tilde{A}_t$ can be written as

$$\tilde{A}_t = F_t Q.$$

Hence,

$$e_m' \mathrm{E}\left[\left(\prod_{i=1}^{j} \tilde{A}_{t+1-i} \right) \pi_{t-j} \right] = e_m' \mathrm{E}\left[\left(\prod_{i=1}^{j} F_{t+1-i} Q \right) \pi_{t-j} \right],$$

and using Assumption 2,

$$e_m' \mathrm{E}\left[\left(\prod_{i=1}^{j} F_{t+1-i} Q \right) \pi_{t-j} \right] \leq e_m' \mathrm{E}\left[\left(\prod_{i=1}^{j} F_{t+1-i} \right) \pi_{t-j} \right].$$

So,

$$\left(\prod_{i=1}^{j} F_{t+1-i} \right) \pi_{t-j} = \begin{pmatrix} \prod_{i=1}^{j} \tilde{\eta}_{t+1-i} & 0_{k\times k(r+s-1)} \\ 0_{k(r+s-1)\times k} & I_{k(r+s-1)\times k(r+s-1)} \end{pmatrix} \begin{pmatrix} (\tilde{\eta}_{t-j} W) \\ 0 \\ \cdots \\ 0 \\ W \\ 0 \\ \cdots \\ 0 \end{pmatrix}.$$

https://doi.org/10.1017/S0266466608080614 Published online by Cambridge University Press

Therefore, it is sufficient to focus only on the submatrix,

$$\prod_{i=1}^{j} \tilde{\eta}_{t+1-i}(\tilde{\eta}_{t-j} W) = \mathrm{E}\left[\prod_{i=1}^{j} \tilde{\eta}_{t-i}\right] W.$$

Notice that the mth row of the matrix $\mathrm{E}[\prod_{i=1}^{j} \tilde{\eta}_{t-i}]$ is

$$\mathrm{E}\left[\prod_{i=1}^{j} \eta_{mt-i}^2\right] \leq \prod_{i=1}^{j} \mathrm{E}[\eta_{mt-i}^4]^{1/2},$$

where the inequality can be achieved by Hölder's inequality. Under the definition of Case 4 and Assumption 5,

$$\mathrm{E}_{\eta,t-1}(\eta_{mt}^2) = \sum_{l=1}^{\infty} \phi_{1m} \phi_{2m}^{l-1} \eta_{mt-l}^2 + \sigma_{\xi m}^2,$$

$$\mathrm{E}_{\eta,t-1}(\eta_{mt}^4) = \sigma_{\nu}^4 \left(\sum_{l=1}^{\infty} \mathrm{E}(\delta_l^4) \eta_{mt-l}^4 + \phi_{1m} \phi_{2m}^{l-1} \eta_{mt-l}^2 + \sigma_{\xi m}^4 \right), \tag{A.4a}$$

$$\leq \sum_{l=1}^{\infty} \phi_{1m} \phi_{2m}^{l-1} \sigma_{\nu}^4 \eta_{mt-l}^4 + \phi_{1m} \phi_{2m}^{l-1} \sigma_{\nu}^4 \eta_{mt-l}^2 + \sigma_{\nu}^4 \sigma_{\xi m}^4. \tag{A.4b}$$

Now, consider the following sequence:

$$\kappa_{mt} = \sum_{l=1}^{\infty} \phi_{1m} \phi_{2m}^{l-1} \kappa_{mt-l} + \sigma_{\nu}^4 \sigma_{\xi m}^4, \tag{A.5}$$

where $0 \leq \kappa_{mt-l} < \infty$, $\forall l \in \mathbb{N}^+$. Under the assumption that $\phi_{1m} + \phi_{2m} < 1$, the sequence (A.5) converges to a finite number, namely,

$$\kappa_{mt} \to \frac{(1 - \phi_{2m}) \sigma_{\nu}^4 \sigma_{\xi m}^4}{1 - (\phi_{1m} + \phi_{2m})} = c_{mt}, \qquad t \to \infty.$$

Under Assumptions 4 and 5, $c_{mt} < 1$, and there exists a sequence $\{\kappa_{mt-l}\}_{l=1}^{\infty}$ such that $\kappa_{mt-l} \geq \sigma_{\nu}^4(\eta_{mt-l}^2 + \eta_{mt-l}^4)$, where $\eta_{mt-l}^2 \in \Im_{\eta,t-1}$ and $\eta_{mt-l}^4 \in \Im_{\eta,t-1}$, $\forall l \in \mathbb{N}^+$. Therefore

$$\mathrm{E}_{\varepsilon,t-1}(\eta_{mt}^4) \leq c_{mt}.$$

The law of iterated expectations then implies that each component of $e_m' \mathrm{E}[(\prod_{i=1}^{j} \tilde{A}_{t+1-i}) \pi_{t-j}]$ *is* $O(\bar{\rho}^j)$. Note that the characteristic polynomial of $\mathrm{E}(\tilde{A}_t)$ has all roots outside the unit circle by Assumption 2, and the moments of δ_{il} follow a geometric progression with finite limits by Assumption 5, $\forall i = 1, \ldots, k$. Furthermore, each element in $\tilde{A}_t$ and π_t is nonnegative. Therefore, each component of $S_{T,t}$ converges a.s. as $T \to \infty$, as does $S_{T,t}$. Hence, there exists an $\Im_t$-measurable second-order solution ε_t to (3), where η_t follows a VAR process with random coefficients, as defined in Case 4.

To show uniqueness, let ε_t^* be another $\Im_t$-measurable second-order stationary solution of (3). Thus, $X_t^* = \tilde{A}_t X_{t-1}^* + \pi_t^*$, where $X_t^* = (\vec{\varepsilon}_t^*, \ldots, \vec{\varepsilon}_{t-r+1}^*, H_t^*, \ldots, H_{t-s+1}^*)'$, with $\vec{\varepsilon}_t^* = ((\varepsilon_{1t}^*)', \ldots, (\varepsilon_{kt}^*)')'$, and

https://doi.org/10.1017/S0266466608080614 Published online by Cambridge University Press

$$H_t^* = W + \sum_{i=1}^{r} A_i \vec{\varepsilon}_{t-i}^* + \sum_{j=1}^{s} B_j H_{t-j}^*.$$

Let $U_t = X_t - X_t^*$, which is first-order stationary by Assumption 1. Furthermore, $U_t = (\prod_{i=0}^{T} \tilde{A}_{t+1-i})U_{t-j}$. *As each element in* $\tilde{A}_t$ is nonnegative, by Assumption 5 it follows that

$$\mathrm{E}|u_{mt}| \leq \mathrm{E}|e_k' \left(\prod_{i=0}^{T} \tilde{A}_{t+1-i} \right) U_{t-j}| \to 0 \quad \text{as } T \to \infty,$$

where u_{mt} is the mth component of U_t. Therefore, $H_t = H_t^*$ and $\vec{\varepsilon}_t = \vec{\varepsilon}_t^*$. Hence, the solution is unique. As $H_t = c' X_t$, it follows that the unique causal representation is given by

$$H_t = W + \sum_{j=1}^{\infty} c' \left(\prod_{i=1}^{j} \tilde{A}_{t-i} \right) \pi_{t-j-1}, \quad \text{a.s.} \qquad \blacksquare$$

The following results will be used in the proof of Theorem 2.

LEMMA 1. *Suppose that* $\mathrm{E}(\|\xi_t\|)^{2m} < \infty$ *and* $\rho[\mathrm{E}(\tilde{A}_t^{\otimes m})] < 1$. *Then there exists a vector* $M > 0$ *such that* $[I_k - \mathrm{E}(\tilde{A}_t^{\otimes m})']M > 0$, *where a vector* $B > 0$ *means that each element of* B *is positive.*

Proof. Note that $\mathrm{E}(\|\xi_t\|)^{2m} < \infty \Rightarrow \mathrm{E}(\|\eta_t\|)^{2m} < \infty$ by Assumption 5, so that $[I_k - \mathrm{E}(\tilde{A}_t^{\otimes m})']M > 0$ is invertible. As every element in $\mathrm{E}(\tilde{A}_t^{\otimes m})$ is nonnegative, there is a vector $L_1 > 0$ such that

$$M = [I_k - \mathrm{E}(\tilde{A}_t^{\otimes m})']^{-1} L_1 = L_1 + \sum_{i=0}^{\infty} [\mathrm{E}(\tilde{A}_{0t}^{\otimes m})']^i L_1 > 0.$$

Thus, $[I_k - \mathrm{E}(\tilde{A}_t^{\otimes m})']M = L_1 > 0$. $\blacksquare$

Remark 1. Lemma 1 extends the results of Lemma A.2 in Ling and McAleer (2003), which assumes that the standardized shocks are i.i.d.

Proof of Theorem 2. Using the results on finite moments in Tweedie (1988), Lemma A.3 in Ling and McAleer (2003), and Lemma 1 given previously, Hölder's inequality implies that $\mathrm{E}_{\pi_1}\|\varepsilon_t\|^2 < (\mathrm{E}_{\pi_1}\|\varepsilon_t\|^{2k})^{1/k} < \infty$, where π_1 are the stationary distributions of $\{\varepsilon_t\}$. Furthermore, $\mathrm{E}_{\pi_2}\|Y_t\|^2 < \infty$ by Assumption 1. Thus, $\{Y_t, \varepsilon_t\}$ is a secondary stationary solution of (9). Moreover, the solution $\{Y_t, \varepsilon_t\}$ is unique and ergodic by Theorem 1. Therefore, $\{Y_t, \varepsilon_t\}$ satisfying model (9) has finite $2m$th moment. $\blacksquare$

Proof of Theorem 3. It is sufficient to verify the following conditions for consistency in Jeantheau (1998).

C1. Λ is compact.

C2. $\forall \lambda \in \Lambda$, the model admits a unique strictly stationary and ergodic solution Y_t.

C3. There exists a deterministic constant $c > 0$ such that, $\forall t$ and $\forall \lambda \in \Lambda$, $|H_t| > c$.

C4. Assumption 3.

C5. Y_t, H_t, and Q_t are continuous functions of the parameter λ.

C6. $\mathrm{E}_{\lambda_0}|\log(|H_t|)| < \infty \ \forall \lambda_0 \in \Lambda$.

https://doi.org/10.1017/S0266466608080614 Published online by Cambridge University Press

Under Theorem 1, (9) admits a unique strictly stationary and ergodic solution of Y_t (C2). Furthermore, the model is identifiable under Assumption 3 (C4). Note that the determinant of the conditional covariance matrix is given as

$$|\mathrm{E}_{\varepsilon,t-1}(\varepsilon_t \varepsilon_t')| = |D_t^{1/2} Q_t D_t^{1/2}| = |D_t|\,|Q_t|.$$

As each element in W, A_i, and B_j is strictly positive $\forall i = 1,\ldots,r$ and $\forall j = 1,\ldots,s$, there exists a constant $c_1 > 0$ so that $|D_t| > c_1$ $\forall t$. Moreover, because each element in Ω, Φ_1, and Φ_2 is positive, and Q_t is positive definite, by construction, there exists a constant $c_2 > 0$ such that $|Q_t| > c_2$. Let $c = c_1 c_2$; then there exists a constant $c > 0$ such that $|\mathrm{E}_{\varepsilon,t-1}(\varepsilon_t \varepsilon_t')| > c$ $\forall t$ and $\forall \lambda \in \Lambda$, where Λ is a compact subspace of euclidean space (C1 and C3). By the square integrability of X_t, $\mathrm{E}_{\lambda_0}(\mathrm{vech}(H_{t,\lambda})) < \infty$, which establishes C6 (for details, see Comte and Lieberman, 2003, p. 67). Under Assumption 1, C6, and the representation (7), $\mathrm{E}(Y_t|\Im_{t-1})$, H_t, and Q_t are continuous functions of the parameter λ (C5). ■

Proof of Theorem 4. It is sufficient to verify the following conditions in Theorem 4.1.1 in Amemiya (1985).

D1. Λ is compact.

D2. $L_T(\lambda)$ is continuous in $\lambda \in \Lambda$ for Y_t and is a measurable function of Y_t $\forall \lambda \in \Lambda$.

D3. $T^{-1}L_T(\lambda)$ converges to a nonstochastic function $L(\lambda)$ in probability uniformly in $\lambda \in \Lambda$ as $T \to \infty$, and $L(\lambda)$ attains a unique global maximum at λ_0.

Condition D1 is equivalent to C1, and D2 follows from C5, and so D1 and D2 are satisfied under Theorem 3. To verify D3, it is convenient to introduce the unobserved process $\{(\varepsilon_t^*, H_t^*, Q_t^*) : t = 0, \pm 1, \pm 2, \ldots\}$. Define the unobserved log-likelihood function conditional on the infinite past observations:

$$L_T^u(\lambda) = \frac{1}{T}\sum_{t=1}^{T} l_t^*(\lambda),$$

$$l_t^*(\lambda) = -\frac{1}{2}\left(\ln|D_t^{*1/2} Q_t^* D_t^{*1/2}| + \varepsilon_t^{*\prime}(D_t^{*1/2} Q_t^* D_t^{*1/2})^{-1}\varepsilon_t^{*\prime}\right),$$

where $D_t^* = \mathrm{diag}(h_{1t}^*, \ldots, h_{kt}^*)$. Furthermore,

$$\frac{\partial Q_t^*}{\partial \rho'} = (I_k - \Phi_2 L)^{-1}(I_k, Q_{t-1}^* \otimes I_k),$$

with the following recursive equation:

$$U_{1t} = \Phi_2 U_{1t-1} + V_{1t},$$

where $U_{1t} = \partial H_t^*/\partial \rho'$ and $V_{1t} = (I_k, Q_{t-1}^* \otimes I_k)$. Hence, if $U_{1t}c = 0$, then $V_{1t}c = 0$ a.s. By Assumptions 3–5, it is straightforward to show that, if $V_{1t}c = 0$, then $c = 0$. This result and Lemma 4.2 in Ling and McAleer (2003) imply that $L(\lambda)$ exists $\forall \lambda \in \Lambda$ and $\sup_{\lambda\in\Lambda}|L_T^*(\lambda) - L(\lambda)| = o_p(1)$. Moreover,

https://doi.org/10.1017/S0266466608080614 Published online by Cambridge University Press

$$(Q_t^* - Q_t)|_{\rho=\rho_0} = \left.\frac{\partial Q_t^*}{\partial \rho'}\right|_{\rho_0} (\rho - \rho_0) = 0$$

if and only if $\rho = \rho_0$. This result and Lemma 4.4 in Ling and McAleer (2003) show that $L(\lambda)$ has a unique maximum at λ_0. Furthermore, C3 implies that

$$T^{-1} \sum_{i=1}^{T} \mathrm{E} \sup_{\lambda \in \Lambda} |\ln|D_t^{*1/2} Q_t D_t^{*1/2}| - \ln|D_t^{1/2} Q_t D_t^{1/2}|| \leq o(1).$$

Lemma 4.6 in Ling and McAleer (2003) implies that $|L_T^*(\lambda) - L_T(\lambda)| = o_p(1)$. Thus,

$$\sup_{\lambda \in \Lambda} |L_T(\lambda) - L(\lambda)| \leq \sup_{\lambda \in \Lambda} |L_T^*(\lambda) - L(\lambda)| + \sup_{\lambda \in \Lambda} |L_T^*(\lambda) - L_T(\lambda)| = o_p(1).$$

Therefore, $L_T^*(\lambda) \to_p L(\lambda)$ uniformly in Λ (D3). ■

Proof of Theorem 5. Given the consistency of $\hat{\lambda}$ for λ_0 in Theorems 3 and 4, it is sufficient to verify the following conditions of Theorem 4.1.3 in Amemiya (1985).

E1. $\partial^2 L_T / \partial \lambda \partial \lambda'$ exists and is continuous in an open, convex neighborhood of λ_0.

E2. $T^{-1}(\partial^2 L_T / \partial \lambda \partial \lambda')|_{\lambda_T}$ converges to a finite nonsingular matrix $\Sigma_0 = \mathrm{E} T^{-1}(\partial^2 L_T / \partial \lambda \partial \lambda')|_{\lambda_0}$ in probability for any sequence λ_T, such that $\hat{\lambda} \to_p \lambda_0$.

E3. $T^{-1/2}(\partial L_T / \partial \lambda)|_{\lambda_0} \xrightarrow{d} N(0, \Omega_\lambda)$, where $\Omega_\lambda = \lim \mathrm{E} T^{-1}(\partial L_t / \partial \lambda)|_{\lambda_0} \times (\partial L_t / \partial \lambda')|_{\lambda_0}$.

By Theorems 3 and 4, $\hat{\lambda}$ is consistent for λ_0. It follows from the conditions in Theorem 3 that $\partial^2 L_T / \partial \lambda \partial \lambda'$ exists and is continuous in Λ. As Q_t is a function of ρ, it follows that

$$\frac{\partial l_t^*}{\partial \rho} = -\frac{1}{2} \frac{\partial \mathrm{vec}'(Q_t)}{\partial \rho} \mathrm{vec}(Q_t^{-1} - Q_t^{-1} D_t^{*-1} \varepsilon_t^* \varepsilon_t^{*'} D_t^{*-1} Q_t^{-1}).$$

Lemma 5.4 in Ling and McAleer (2003) can be used to verify that

$$\mathrm{E} \sup_{\lambda \in \Lambda} \|\partial^2 L_T / \partial \tau \partial \rho'\| < \infty$$

and

$$\mathrm{E} \sup_{\lambda \in \Lambda} \|\partial^2 L_T / \partial \rho \partial \rho'\| < \infty$$

(E1 and E2). Under the assumption that $\mathrm{E}(\xi_t^4) < \infty$, which implies $\mathrm{E}(\eta_t^4) < \infty$, using the central limit theorem of Stout (1974) and the Cramér–Wold device, it follows that

$$T^{-1/2} \sum_{t=1}^{T} \frac{\partial l_t}{\partial \lambda} \xrightarrow{d} N(0, \Omega_\lambda),$$

where $\Omega_\lambda < \infty$ (E3). ■

https://doi.org/10.1017/S0266466608080614 Published online by Cambridge University Press