

Asai, Manabu; Chang, Chia-Lin; McAleer, Michael; Pauwels, Laurent

Article

Asymptotic and finite sample properties for multivariate rotated GARCH models

Econometrics

Provided in Cooperation with:

MDPI – Multidisciplinary Digital Publishing Institute, Basel

Suggested Citation: Asai, Manabu; Chang, Chia-Lin; McAleer, Michael; Pauwels, Laurent (2021) : Asymptotic and finite sample properties for multivariate rotated GARCH models, *Econometrics*, ISSN 2225-1146, MDPI, Basel, Vol. 9, Iss. 2, pp. 1-21, <https://doi.org/10.3390/econometrics9020021>

This Version is available at:

<https://hdl.handle.net/10419/247611>

Standard-Nutzungsbedingungen:

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

Terms of use:

Documents in EconStor may be saved and copied for your personal and scholarly purposes.

You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.

If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.



<https://creativecommons.org/licenses/by/4.0/>

Article

Asymptotic and Finite Sample Properties for Multivariate Rotated GARCH Models

Manabu Asai ^{1,*}, Chia-Lin Chang ^{2,3}, Michael McAleer ^{4,5,6,7,8} and Laurent Pauwels ⁶¹ Faculty of Economics, Soka University, Tokyo 192-8577, Japan² Department of Applied Economics, National Chung Hsing University, Taichung City 402, Taiwan; changchialin@email.nchu.edu.tw³ Department of Finance, National Chung Hsing University, Taichung City 402, Taiwan⁴ Department of Finance, College of Management, Asia University, Taichung City 41354, Taiwan; michael.mcaleer@gmail.com⁵ Department of Bioinformatics and Medical Engineering, College of Information and Electrical Engineering, Asia University, Taichung City 41354, Taiwan⁶ Discipline of Business Analytics, University of Sydney Business School, Darlingtown 2006, Australia; laurent.pauwels@sydney.edu.au⁷ Econometric Institute, Erasmus School of Economics, Erasmus University Rotterdam, 3062 PA Rotterdam, The Netherlands⁸ Department of Economic Analysis and ICAE, Complutense University of Madrid, 28040 Madrid, Spain

* Correspondence: m-asai@soka.ac.jp

† These authors contributed equally to this work.



Citation: Asai, Manabu, Chia-Lin Chang, Michael McAleer, and Laurent Pauwels. 2021. Asymptotic and Finite Sample Properties for Multivariate Rotated GARCH Models. *Econometrics* 9: 21. <https://doi.org/10.3390/econometrics9020021>

Academic Editors: Marc S. Paolella, In Choi and Martin Wagner

Received: 24 December 2020

Accepted: 27 April 2021

Published: 4 May 2021

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Abstract: This paper derives the statistical properties of a two-step approach to estimating multivariate rotated GARCH-BEKK (RBEKK) models. From the definition of RBEKK, the unconditional covariance matrix is estimated in the first step to rotate the observed variables in order to have the identity matrix for its sample covariance matrix. In the second step, the remaining parameters are estimated by maximizing the quasi-log-likelihood function. For this two-step quasi-maximum likelihood (2sQML) estimator, this paper shows consistency and asymptotic normality under weak conditions. While second-order moments are needed for the consistency of the estimated unconditional covariance matrix, the existence of the finite sixth-order moments is required for the convergence of the second-order derivatives of the quasi-log-likelihood function. This paper also shows the relationship between the asymptotic distributions of the 2sQML estimator for the RBEKK model and variance targeting quasi-maximum likelihood estimator for the VT-BEKK model. Monte Carlo experiments show that the bias of the 2sQML estimator is negligible and that the appropriateness of the diagonal specification depends on the closeness to either the diagonal BEKK or the diagonal RBEKK models. An empirical analysis of the returns of stocks listed on the Dow Jones Industrial Average indicates that the choice of the diagonal BEKK or diagonal RBEKK models changes over time, but most of the differences between the two forecasts are negligible.

Keywords: BEKK; rotated BEKK; diagonal BEKK; variance targeting; multivariate GARCH; consistency; asymptotic normality



Copyright: © 2021 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

The BEKK model of [Baba et al. \(1985\)](#) and [Engle and Kroner \(1995\)](#) is widely used for estimating and forecasting time-varying conditional covariance dynamics, especially in the empirical analysis of the multiple asset returns of financial time series ([Bauwens et al. 2006](#); [Laurent et al. 2012](#); [McAleer 2005](#); [Silvennoinen and Teräsvirta 2009](#)). The BEKK model is a natural extension of the ARCH/GARCH models of [Engle \(1982\)](#) and [Bollerslev \(1986\)](#). One of the features of the BEKK model is that it guarantees the positive definiteness of the covariance matrix. However, as it does not satisfy suitable regularity conditions, the corresponding estimators do not possess asymptotic properties, except under restrictive

conditions (Chang and McAleer 2019; Comte and Lieberman 2003; McAleer et al. 2008). To overcome this problem, Hafner and Preminger (2009) showed the asymptotic properties of the quasi-maximum likelihood (QML) estimator under moderate regularity conditions.

Similar to other multivariate GARCH models, a drawback of the BEKK model is that it contains a large number of parameters, even for moderate dimensions. To reduce the number of parameters, the so-called scalar BEKK and diagonal BEKK specifications are used in empirical analyses (Chang and McAleer 2019). Recently, Noureldin et al. (2014) suggested the rotated BEKK (RBEKK) model to handle the high-dimensional BEKK model. They proposed the estimation of the unconditional covariance matrix of the observed variables in the first step to rotate the variables in order to have unit sample variance and zero sample correlation coefficients. In the second step, they considered simplified BEKK models for the QML estimation. We call this procedure the two-step QML (2sQML) estimation. One of the major advantages of the RBEKK model is that it can reduce the number of parameters in the optimization step, while another is that it is more natural to consider simplified specifications after the rotation than to simplify the structure directly without the rotation.

2sQML is closely related to the variance targeting (VT) specification analyzed by Francq et al. (2011) and Pedersen and Rahbek (2014), among others. The VT-QML estimation also uses the estimated unconditional covariance matrix in the first step to reduce the number of parameters in the QML maximization step. Pedersen and Rahbek (2014) showed the consistency and asymptotic normality of the VT-QML estimator under finite sixth-order moments. As Noureldin et al. (2014) discussed the general framework for the asymptotic distribution of the 2sQML estimator for the RBEKK model, it is thus worth examining the detailed moment condition, as in Pedersen and Rahbek (2014).

In this study, we show the consistency and asymptotic normality of the 2sQML estimator for the RBEKK model by extending the approach of Pedersen and Rahbek (2014). For asymptotic normality, we need to impose sixth-order moment restrictions, as in Hafner and Preminger (2009) and Pedersen and Rahbek (2014). We also derive the asymptotic relationship between the VT-QML estimator for the BEKK model and the 2sQML estimator for the RBEKK model. We conduct Monte Carlo experiments to check the finite sample properties of the 2sQML estimator and compare the performance of the estimated diagonal BEKK and diagonal RBEKK models. The proofs of the propositions and corollaries are given in Appendix A. We present an empirical result based on the returns of stocks listed on the Dow Jones Industrial Average.

There are several works related to the idea of Noureldin et al. (2014). First, the transformation suggested by Noureldin et al. (2014) is related to the (generalized) orthogonal GARCH models of Alexander (2001), van der Weide (2002), and Lanne and Saikkonen (2007). While these authors attempt to find orthogonal or unconditionally uncorrelated components in the raw returns, which can then be modeled individually through univariate variance models, Noureldin et al. (2014) suggest fitting flexible multivariate models to the rotated returns using the VT approach. Second, the symmetric square root rotation of Noureldin et al. (2014) is not the most general type of rotation one could use—see, for instance, the hyper-rotation suggested by Asai and McAleer (2020) and the structural multivariate GARCH approach of Hafner et al. (2020). Both use generalized rotations that are not necessarily symmetric. From the structure, we can infer that these works are not based on the VT approach.

We use the following notation throughout the paper. For a matrix A , we define $A^{\otimes 2} = (A \otimes A)$. With ζ_1, \dots, ζ_n , the n eigenvalues of a matrix A , $\rho(A) = \max_{i \in \{1, \dots, n\}} |\zeta_i|$, is the spectral radius of A . The Frobenius norm of the matrix, or vector A , is defined as $\|A\| = \sqrt{\text{tr}(A'A)}$. For a positive matrix A , we define the square root, $A^{1/2}$, by the spectral decomposition of A . For an $m \times n$ matrix A , the $mn \times mn$ commutation matrix C_{mn} has the property $C_{mn} \text{vec}(A) = \text{vec}(A')$.

2. RBEKK-GARCH Model

As in [Hafner and Preminger \(2009\)](#) and [Pedersen and Rahbek \(2014\)](#), we focus on a simple specification of the BEKK model defined by:

$$X_t = H_t^{1/2} Z_t, \quad (1)$$

$$H_t = C^* + A^* X_{t-1} X_{t-1}' A^{*'} + B^* H_{t-1} B^{*'}, \quad (2)$$

where $t = 1, \dots, T$, A^* , and B^* are d -dimensional square matrices, C^* is a d -dimensional positive definite matrix, and Z_t ($d \times 1$) is an *i.i.d.* $(0, I_d)$ sequence of random variables.

We begin with the following assumption.

Assumption 1.

- The distribution of Z_t is absolutely continuous with respect to the Lebesgue measure of \mathbb{R}^d and zero is an interior point of the support of the distribution.
- The matrices A^* and B^* satisfy $\rho((A^* \otimes A^*) + (B^* \otimes B^*)) < 1$.

From Theorem 2.4 of [Boussama et al. \(2011\)](#), Assumption 1 implies the existence of a unique stationary and ergodic solution to the model in (1) and (2). Furthermore, the stationary solution has finite second-order moments, $E\|X_t\|^2 < \infty$, and variance $V(X_t) = E(H_t) = \Omega$, with positive definite Ω , which is the solution to:

$$\Omega = C^* + A^* \Omega A^{*'} + B^* \Omega B^{*'} \quad (3)$$

Lemma 2.4 and Proposition 4.3 of [Boussama et al. \(2011\)](#) indicate that the necessary and sufficient conditions for (3) to have a solution of a positive definite matrix is Assumption 1(b). As in [Pedersen and Rahbek \(2014\)](#), we obtain the VT specification by substituting C^* in (3) into the model (2), giving:

$$H_t = \Omega - A^* \Omega A^{*'} - B^* \Omega B^{*'} + A^* X_{t-1} X_{t-1}' A^{*'} + B^* H_{t-1} B^{*'} \quad (4)$$

Based on this specification, [Noureldin et al. \(2014\)](#) suggested the RBEKK model, which is obtained by setting $A^* = \Omega^{1/2} A \Omega^{-1/2}$ and $B^* = \Omega^{1/2} B \Omega^{-1/2}$ in (2), where A and B are d -dimensional square matrices. The transformation yields:

$$H_t = \Omega^{1/2} \underline{H}_t \Omega^{1/2}, \quad \underline{H}_t = (I_d - AA' - BB') + A \tilde{X}_{t-1} \tilde{X}_{t-1}' A' + B \underline{H}_{t-1} B', \quad (5)$$

with the rotated vector $\tilde{X}_t = \Omega^{-1/2} X_t$, which gives $E(\tilde{X}_t \tilde{X}_t') = I_d$. As discussed by [Noureldin et al. \(2014\)](#), the specification provides a natural interpretation for considering the diagonal matrices A and B to reduce the number of parameters. Rather than the special case with these diagonal matrices, we consider general A and B for asymptotic theory. With respect to the initial values, we consider the estimation conditional on the initial values X_0 and $\underline{H}_0 = h$, where h is a positive definite matrix. Under this structure, it is natural to replace Assumption 1(b) with the following assumption.

Assumption 2. The matrices A and B satisfy $\rho((A \otimes A) + (B \otimes B)) < 1$.

Lemma A2 in Appendix A.2 shows that Assumption 2 is equivalent to Assumption 1(b).

In the next section, we consider the 2sQML estimation for the RBEKK model (1) and (5), as in [Noureldin et al. \(2014\)](#) and [Pedersen and Rahbek \(2014\)](#).

3. 2sQML Estimation

Let $\theta, \theta \in \mathbb{R}^{3d^2}$ denote the parameter vector of the RBEKK model, which is defined by $\theta = (\omega', \lambda')'$, where $\omega = \text{vec}(\Omega)$ and $\lambda = (\alpha', \beta')'$ with $\alpha = \text{vec}(A)$ and $\beta = \text{vec}(B)$. We also define the parameter space $\Theta = \Theta_\omega \times \Theta_\lambda \subset \mathbb{R}^{d^2} \times \mathbb{R}^{2d^2}$. As in [Hafner and Preminger \(2009\)](#) and [Pedersen and Rahbek \(2014\)](#), we emphasize the dependence of H_t and \underline{H}_t

on the parameters ω and λ by writing $H_t(\omega, \lambda)$ and $\underline{H}_t(\omega, \lambda)$, respectively. In addition, we emphasize the initial value of the covariance matrix, h , by denoting $H_{t,h}(\omega, \lambda)$ and $\underline{H}_{t,h}(\omega, \lambda)$. Now, we restate the RBEKK model as:

$$X_t = H_t^{1/2}(\omega, \lambda)Z_t, \quad H_t(\omega, \lambda) = \Omega^{1/2}\underline{H}_t(\omega, \lambda)\Omega^{1/2}, \quad (6)$$

$$\underline{H}_t(\omega, \lambda) = (I_d - AA' - BB') + A\Omega^{-1/2}X_{t-1}X'_{t-1}\Omega^{-1/2}A' + B\underline{H}_{t-1}(\omega, \lambda)B', \quad (7)$$

with the given initial values X_0 and $\underline{H}_{0,h}(\omega, \lambda) = h$.

As mentioned above, we consider the 2sQML estimation, which comprises two steps. In the first step, we estimate ω using the sample covariance matrix, while the second step conducts the QML estimation by optimizing the log-likelihood function for λ conditional on the estimates of ω . For the RBEKK model, the Gaussian log-likelihood function is given by:

$$L_{T,h}(\omega, \lambda) = T^{-1} \sum_{t=1}^T l_{t,h}(\omega, \lambda), \quad (8)$$

with the t th contribution to the log-likelihood given as:

$$l_{t,h}(\omega, \lambda) = -\frac{1}{2} \log(\det(H_{t,h}(\omega, \lambda))) - \frac{1}{2} \text{tr}\left(X_t X'_t H_{t,h}^{-1}(\omega, \lambda)\right), \quad (9)$$

excluding the constant. In the first step, we estimate the unconditional covariance matrix by:

$$\hat{\omega} = \text{vec}(\hat{\Omega}) = \text{vec}\left(T^{-1} \sum_{t=1}^T X_t X'_t\right), \quad (10)$$

to rotate X_t and $H_{t,h}(\omega, \lambda)$ as:

$$\tilde{X}_t = \hat{\Omega}^{-1/2} X_t, \quad \underline{\tilde{H}}_{t,h}(\hat{\omega}, \lambda).$$

Note that the sample covariance matrix is positive definite, since the structure confirms the positive semi-definiteness, and since $T > d$ guarantees that the rank of $[X_1 \cdots X_T]$ is d . By definition, we have $T^{-1} \sum_{t=1}^T \tilde{X}_t \tilde{X}'_t = I_d$. The conditional log-likelihood function is given by:

$$-\frac{1}{2T} \sum_{t=1}^T \left[\log(\det(\underline{\tilde{H}}_{t,h}(\hat{\omega}, \lambda))) + \text{tr}\left(\tilde{X}_t \tilde{X}'_t \underline{\tilde{H}}_{t,h}^{-1}(\hat{\omega}, \lambda)\right) \right],$$

which is equivalent to $L_{T,h}(\hat{\omega}, \lambda) + 0.5T \log(\det(\hat{\Omega}))$. Hence, the second step of the estimator is given by:

$$\hat{\lambda} = \arg \max_{\lambda \in \Theta_\lambda} L_{T,h}(\hat{\omega}, \lambda). \quad (11)$$

We derive the asymptotic theory for the 2sQML estimator, which consists of (10) and (11).

Following Comte and Lieberman (2003), Hafner and Preminger (2009), and Pedersen and Rahbek (2014), we make the following conventional assumptions.

Assumption 3.

- (a) The process $\{X_t\}$ is strictly stationary and ergodic.
- (b) The true parameters $\theta_0 \in \Theta$ and Θ are compact.
- (c) For $\lambda \in \Theta_\lambda$, if $\lambda \neq \lambda_0$, then $\underline{H}_t(\omega_0, \lambda) \neq \underline{H}_t(\omega_0, \lambda_0)$ almost surely, for all $t \geq 1$.

For Assumption 3(a), Assumptions 1(a) and 2 imply the existence of a strictly stationary ergodic solution $\{X_t\}$ in the RBEKK model. Regarding Assumption 3(c), one of the conditions is that the first element in the matrices A and B should be strictly positive, which is a sufficient condition for the parameter identification, as shown by Engle and Kroner (1995).

We now state the following result on the consistency of the 2sQML estimator, $\hat{\theta} = (\hat{\omega}', \hat{\lambda}')'$.

Proposition 1. Under Assumptions 1(a), 2, and 3, as $T \rightarrow \infty$, $\hat{\theta} \xrightarrow{a.s.} \theta_0$.

Assumptions 1(a) and 2 imply the finite second-order moments of X_t , which are necessary for estimating Ω with the sample covariance matrix. As shown by Hafner and Preminger (2009), the consistency of the QML estimator for the BEKK model (1) and (2) does not require the finite second-order moment of X_t .

We make the following assumption about the asymptotic normality of the 2sQML estimator.

Assumption 4.

- (a) $E[||X_t||^6] < \infty$.
- (b) θ_0 is in the interior of Θ .

As in Pedersen and Rahbek (2014), we need to assume finite sixth-order moments to show that the second-order derivatives of the log-likelihood function converge uniformly on the parameter space. This is different than the univariate case, which only requires finite fourth-order moments (Francq et al. 2011).

Proposition 2. Under Assumptions 1(a) and 2–4, as $T \rightarrow \infty$,

$$\sqrt{T}(\hat{\theta} - \theta_0) \xrightarrow{d} N(\mathbf{0}, Q_0 \Gamma_0 Q_0'),$$

where

$$Q_0 = \begin{pmatrix} I_{d^2} & O_{d^2 \times 2d^2} \\ -J_0^{-1} K_0 & -J_0^{-1} \end{pmatrix},$$

with the matrix K_0 and non-singular matrices J_0 and Γ_0 defined by:

$$K_0 = E \left(\frac{\partial^2 l_t(\theta)}{\partial \lambda \partial \omega'} \Big|_{\theta=\theta_0} \right), \quad J_0 = E \left(\frac{\partial^2 l_t(\theta)}{\partial \lambda \partial \lambda'} \Big|_{\theta=\theta_0} \right), \quad \Gamma_0 = E[\gamma_t \gamma_t'],$$

using

$$\gamma_t = \begin{pmatrix} \text{vec}(X_t X_t') - \omega_0 \\ \frac{\partial l_t(\theta)}{\partial \lambda} \Big|_{\theta=\theta_0} \end{pmatrix}.$$

Remark 1. The structure of the asymptotic covariance matrix is similar to that of the VT-QML estimator for Equations (2) and (4), derived by Pedersen and Rahbek (2014). The major difference comes from the model structure, as the RBEKK further assumes that A^* and B^* depend on the non-linear function of Ω .

Remark 2. We can estimate Γ_0 , K_0 , and J_0 using the sample outer product of the gradient and Hessian matrices as follows:

$$\hat{\Gamma} = \frac{1}{T} \sum_{t=1}^T \hat{\gamma}_t \hat{\gamma}_t', \quad \hat{K} = \frac{1}{T} \sum_{t=1}^T \hat{K}_t, \quad \hat{J} = \frac{1}{T} \sum_{t=1}^T \hat{J}_t,$$

where

$$\hat{\gamma}_t = \begin{pmatrix} \text{vec}(X_t X_t') - \hat{\omega} \\ \frac{\partial l_{t,h}(\theta)}{\partial \lambda} \Big|_{\theta=\hat{\theta}} \end{pmatrix}, \quad \hat{K}_t = \frac{\partial^2 l_{t,h}(\theta)}{\partial \lambda \partial \omega'} \Big|_{\theta=\hat{\theta}}, \quad \hat{J}_t = \frac{\partial^2 l_{t,h}(\theta)}{\partial \lambda \partial \lambda'} \Big|_{\theta=\hat{\theta}}.$$

Given the asymptotic distribution of $\hat{\theta}$, we can show the asymptotic distribution of the 2sQML estimator of (Ω, A^*, B^*) in the VT representation of the BEKK. Define $\theta = (\omega', \lambda^*)'$, where $\lambda^* = (\alpha^*, \beta^*)'$.

Proposition 3. Under the assumptions of Proposition 2, as $T \rightarrow \infty$,

$$\sqrt{T}(\hat{\theta}^* - \theta_0^*) \xrightarrow{d} N(\mathbf{0}, R_0 Q_0 \Gamma_0 Q_0' R_0'),$$

while the VT-QML estimator has the asymptotic covariance matrix, $R_0^* Q_0 \Gamma_0 Q_0' R_0^*$, where R_0 and R_0^* are stated in (A5) and (A17), respectively.

Remark 3. The difference in the asymptotic covariance matrix for the 2sQML and VT-QML estimators depends on R_0 and R_0^* . While R_0^* is a symmetric matrix, R_0 is a square matrix in general. The definiteness of $R_0 Q_0 \Gamma_0 Q_0' R_0' - R_0^* Q_0 \Gamma_0 Q_0' R_0^*$ is undetermined.

From Proposition 3 and the delta method, we provide the asymptotic distribution of the 2sQML estimator for (C^*, A^*, B^*) in the original BEKK model. Define $c^* = \text{vec}(C^*)$, $\alpha^* = \text{vec}(A^*)$, and $\beta^* = \text{vec}(B^*)$.

Corollary 1. Under the assumptions of Proposition 2, as $T \rightarrow \infty$,

$$\sqrt{T} \begin{pmatrix} \hat{c}^* - c^* \\ \hat{\alpha}^* - \alpha^* \\ \hat{\beta}^* - \beta^* \end{pmatrix} \xrightarrow{d} N(\mathbf{0}, S_0' R_0 Q_0 \Gamma_0 Q_0' R_0' S_0),$$

where

$$S_0 = \begin{pmatrix} I_{d^2} - (\Omega_0^{1/2} A_0 \Omega_0^{-1/2})^{\otimes 2} - (\Omega_0^{1/2} B_0 \Omega_0^{-1/2})^{\otimes 2} & O_{d^2 \times d^2} & O_{d^2 \times d^2} \\ -(\Omega_0^{-1/2} \otimes \Omega_0^{1/2})(I_{d^2} + C_{dd})((\Omega_0^{1/2} A_0 \Omega_0^{1/2}) \otimes I_d) & I_{d^2} & O_{d^2 \times d^2} \\ -(\Omega_0^{-1/2} \otimes \Omega_0^{1/2})(I_{d^2} + C_{dd})((\Omega_0^{1/2} B_0 \Omega_0^{1/2}) \otimes I_d) & O_{d^2 \times d^2} & I_{d^2} \end{pmatrix}.$$

Remark 4. From Proposition 1, we can estimate S_0 and R_0 using the 2sQML estimate, $\hat{\theta}$.

4. Monte Carlo Experiments

4.1. Experimental Framework

In this section, we illustrate the theoretical results presented in the previous section using Monte Carlo experiments for d -dimensional rotated diagonal GARCH models. We use the diagonal GARCH model for the model comparison because the number of parameters to be estimated is the same. It is difficult to estimate the fully parametrized BEKK model when d takes a higher value such as $d = 30$, which gives the number of parameters for A^* and B^* as $2d^2 = 1800$. On the contrary, rotated and unrotated diagonal GARCH models use $2d = 60$ parameters. We consider four experiments to examine the (i) finite sample property of the 2sQML estimator, (ii) difference between the true and estimated covariance matrices, (iii) approximation of the fully parametrized BEKK model via the diagonal models, and (iv) finite sample property of the 2sQML estimator when fourth moments of X_t do not exist. We generate Z_t from $N(0, I_d)$ for all the experiments.

4.2. Performance of the 2sQML Estimator

In the first Monte Carlo experiment, we consider the bivariate case ($d = 2$) for the data-generating processes (DGPs) with the following structure in (5):

$$\begin{aligned} \Omega_0 &= \begin{pmatrix} s_{01} & 0 \\ 0 & s_{02} \end{pmatrix} \begin{pmatrix} 1 & \rho_0 \\ \rho_0 & 1 \end{pmatrix} \begin{pmatrix} s_{01} & 0 \\ 0 & s_{02} \end{pmatrix}, \\ A_0 &= \begin{pmatrix} A_{11,0} & 0 \\ 0 & A_{22,0} \end{pmatrix}, \quad B_0 = \begin{pmatrix} B_{11,0} & 0 \\ 0 & B_{22,0} \end{pmatrix}. \end{aligned} \tag{12}$$

We consider two types of parameter sets:

$$\text{DGP1: } (s_{01}, s_{02}) = (1, 0.9), \quad \rho_0 = 0.5, \quad (A_{11,0}, A_{22,0}) = (0.6, 0.4), \\ (B_{11,0}, B_{22,0}) = (0.7, 0.9),$$

$$\text{DGP2: } (s_{01}, s_{02}) = (0.8, 1.1), \quad \rho_0 = -0.3, \quad (A_{11,0}, A_{22,0}) = (0.6, -0.3), \\ (B_{11,0}, B_{22,0}) = (0.7, -0.9),$$

which are used to obtain (C_0^*, A_0^*, B_0^*) for the DGPs from (1) and (2). We use $\underline{H}_1 = I_2$ for the initial value to generate $T = 500$ observations. We set the number of replications to 2000. Tables 1 and 2 provide the values of (Ω_0, A_0, B_0) and corresponding values of (C_0^*, A_0^*, B_0^*) , respectively. While DGP1 describes the positive unconditional correlation, DGP2 uses the negative correlation. Using this specification, we can verify that $\rho((A_0^* \otimes A_0^*) + (B_0^* \otimes B_0^*)) = 0.9$. From this setting, we examine the finite sample property of the 2sQML estimator for (Ω, A, B) .

Table 1. Finite sample properties of the 2sQML estimator for the diagonal RBEEK model.

| Parameters | DGP1 | | | | DGP2 | | | |
|---------------|------|--------|-----------|--------|--------|---------|-----------|--------|
| | True | Mean | Std. Dev. | RMSE | True | Mean | Std. Dev. | RMSE |
| Ω_{11} | 1.00 | 1.0150 | 0.5831 | 0.5832 | 0.640 | 0.6375 | 0.2031 | 0.2031 |
| Ω_{21} | 0.54 | 0.5492 | 0.3654 | 0.3654 | -0.264 | -0.2635 | 0.0552 | 0.0552 |
| Ω_{22} | 0.81 | 0.8250 | 0.7143 | 0.7143 | 1.210 | 1.2067 | 0.1474 | 0.1474 |
| A_{11} | 0.60 | 0.5853 | 0.0531 | 0.0551 | 0.600 | 0.5855 | 0.0567 | 0.0586 |
| A_{22} | 0.40 | 0.3921 | 0.0424 | 0.0431 | -0.300 | -0.3032 | 0.0523 | 0.0524 |
| B_{11} | 0.70 | 0.6939 | 0.0593 | 0.0596 | 0.700 | 0.6920 | 0.0710 | 0.0714 |
| B_{22} | 0.90 | 0.8921 | 0.0463 | 0.0470 | -0.900 | -0.8666 | 0.1025 | 0.1078 |

Table 2. Finite sample properties of the 2sQML estimator for the full BEKK model.

| Parameters | DGP1 | | | | DGP2 | | | |
|------------|---------|---------|-----------|--------|---------|---------|-----------|--------|
| | True | Mean | Std. Dev. | RMSE | True | Mean | Std. Dev. | RMSE |
| C_{11}^* | 0.1392 | 0.1469 | 0.0346 | 0.0354 | 0.0950 | 0.1007 | 0.0262 | 0.0268 |
| C_{21}^* | 0.0505 | 0.0559 | 0.0175 | 0.0183 | -0.0319 | -0.0396 | 0.0174 | 0.0190 |
| C_{22}^* | 0.0351 | 0.0433 | 0.0165 | 0.0184 | 0.1220 | 0.1707 | 0.1185 | 0.1281 |
| A_{11}^* | 0.6249 | 0.6113 | 0.0613 | 0.0628 | 0.6212 | 0.6072 | 0.0582 | 0.0599 |
| A_{21}^* | 0.0706 | 0.0685 | 0.0260 | 0.0261 | -0.1644 | -0.1656 | 0.0281 | 0.0281 |
| A_{12}^* | -0.0794 | -0.0817 | 0.0375 | 0.0375 | 0.1187 | 0.1181 | 0.0230 | 0.0230 |
| A_{22}^* | 0.3751 | 0.3661 | 0.0487 | 0.0495 | -0.3212 | -0.3250 | 0.0542 | 0.0543 |
| B_{11}^* | 0.6751 | 0.6678 | 0.0597 | 0.0601 | 0.7376 | 0.7313 | 0.0613 | 0.0616 |
| B_{21}^* | -0.0706 | -0.0714 | 0.0266 | 0.0266 | -0.2922 | -0.2912 | 0.0521 | 0.0521 |
| B_{12}^* | 0.0794 | 0.0824 | 0.0303 | 0.0304 | 0.2110 | 0.2067 | 0.0360 | 0.0363 |
| B_{22}^* | 0.9249 | 0.9195 | 0.0311 | 0.0315 | -0.9376 | -0.9045 | 0.1073 | 0.1123 |

Table 1 shows the sample mean, standard deviations, and root mean squared error of the 2sQML estimates. The table indicates that the bias of the estimators is negligible, even for $T = 500$. The standard deviation for the sample covariance matrix is relatively large for DGP1, which is expected to decrease as the sample size increases. Figure 1 shows the histograms and QQ Plots of 2sQML estimates for representative parameters. As the estimate of Ω_{11} is obtained from the sample covariance matrix, its distribution is skewed to the right. By contrast, the distributions of the 2sQML estimates of A_{11} and B_{11} are close to the normal distribution.

We also check the effects of the transformation from $(\hat{\Omega}, \hat{A}, \hat{B})$ to $(\hat{C}^*, \hat{A}^*, \hat{B}^*)$, as shown in Corollary 1. Table 2 shows the sample mean, standard error, and root mean squared error of the transformed estimator. As in Table 1, the bias of the estimators is negligible.

Figure 2 indicates that the transformed estimators for the full BEKK specification can be approximated using the normal distribution.

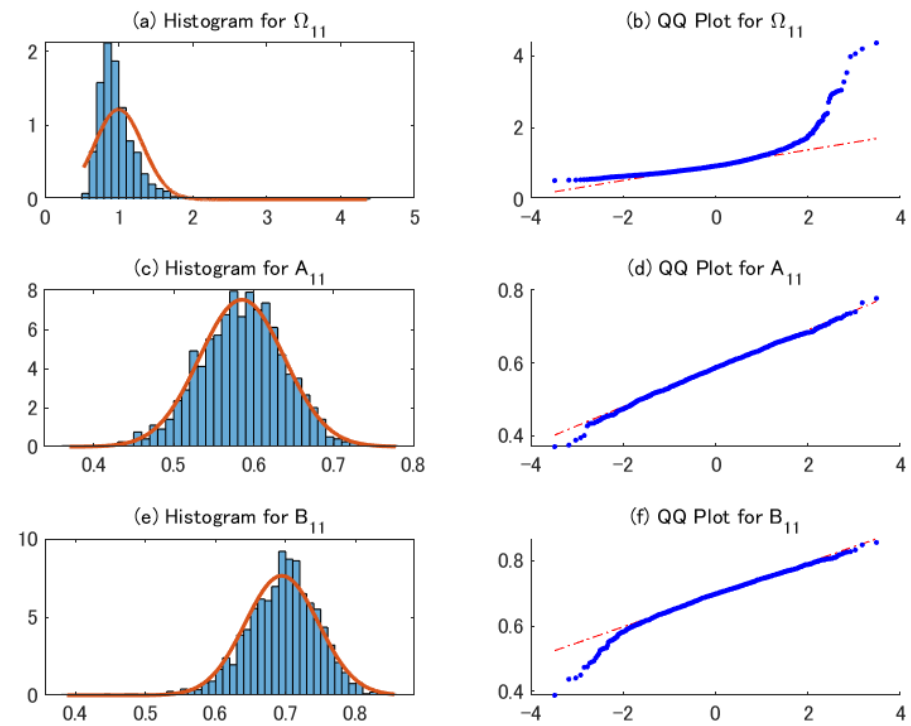


Figure 1. Histograms and QQ plots of 2sQML estimates for DGP1.

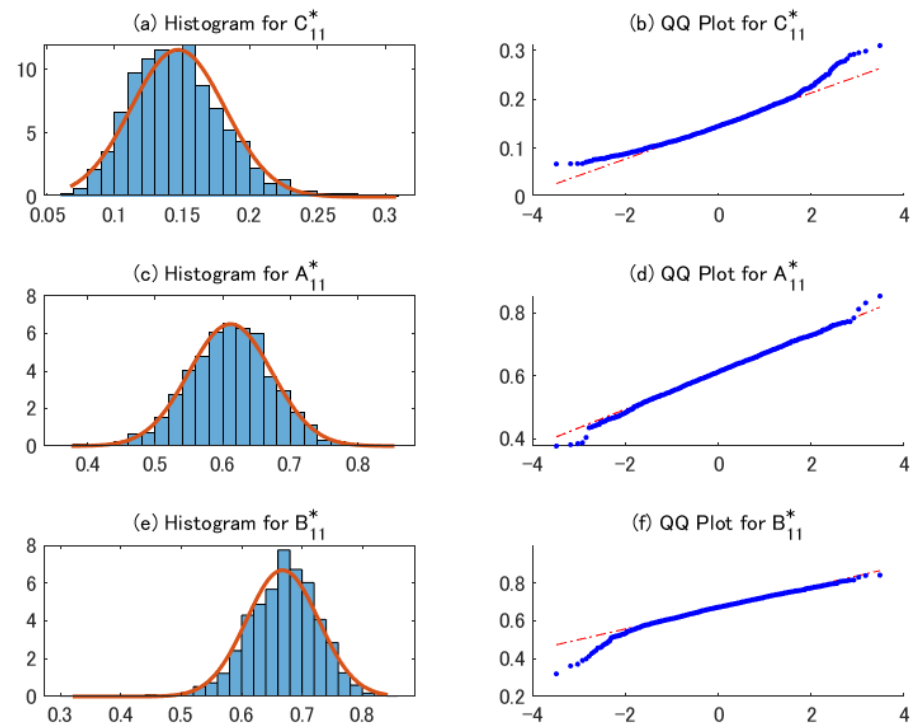


Figure 2. Histograms and QQ plots of 2sQML estimates for DGP1 with full BEKK parameters.

4.3. Performance of Conditional Covariance Matrix Estimator

In the second experiment, we consider higher-dimensional cases with $d = \{5, 10, 30\}$. Denoting by $U(a, b)$, we simulate the uniform distribution on $[a, b]$:

$$\text{DGP3: } \Omega_0 = \Xi_0 \Xi_0', \quad \Xi_0 = \begin{pmatrix} \Xi_{11,0} & 0 & \cdots & 0 \\ \Xi_{21,0} & \Xi_{22,0} & \ddots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ \Xi_{d1,0} & \Xi_{d2,0} & \cdots & \Xi_{dd,0} \end{pmatrix},$$

$$A_0 = \begin{pmatrix} A_{11,0} & 0 & \cdots & 0 \\ 0 & A_{22,0} & \ddots & 0 \\ \vdots & \ddots & \ddots & \vdots \\ 0 & 0 & \cdots & A_{dd,0} \end{pmatrix}, \quad B_0 = \begin{pmatrix} B_{11,0} & 0 & \cdots & 0 \\ 0 & B_{22,0} & \ddots & 0 \\ \vdots & \ddots & \ddots & \vdots \\ 0 & 0 & \cdots & B_{dd,0} \end{pmatrix},$$

where $\Xi_{ij,0} \sim U(-0.5, 0.5)$ ($i \geq j$), $A_{ii,0} \sim U(0.1, 0.4)$, and $B_{ii,0} \sim U(0.85, 0.9)$ ($i, j = 1, 2, \dots, d$). We discard the parameters which do not satisfy the stationarity condition. Based on the simulated parameters, we generate X_t for the sample size $T = \{500, 1000\}$, and estimate the diagonal RBEKK and diagonal BEKK models to calculate the average of the Frobenius norm of the difference in the conditional covariance matrices:

$$\frac{1}{T} \sum_{t=1}^T \|\hat{H}_t - H_{t,0}\|,$$

where \hat{H}_t is the estimated covariance matrix and $H_{t,0}$ is the true covariance matrix. While $\hat{H}_t = H_{t,h}(\hat{\theta})$ for the diagonal BEKK model, \hat{H}_t is similarly defined for the diagonal RBEKK model. We use the common random parameters for the two models and repeat the procedure 100 times with different random parameters.

Table 3 shows the sample means of the average distances. The values of the measure for $d = 10$ are expected to take values approximately four times larger than those for $d = 5$, as implied by the dimension of \hat{H}_t ($d \times d$). Table 3 supports this result. As the sample size increases, the average distance decreases. Compared with the diagonal BEKK model, the diagonal RBEKK model has a smaller distance measure.

Table 3. Average distance between the true and estimated covariance matrices.

| Sample Size | Diagonal RBEKK | | | Diagonal BEKK | | |
|-------------|----------------|----------|----------|---------------|----------|----------|
| | $d = 5$ | $d = 10$ | $d = 30$ | $d = 5$ | $d = 10$ | $d = 30$ |
| $T = 500$ | 1.1377 | 3.9358 | 41.047 | 1.5176 | 6.3447 | 64.909 |
| $T = 1000$ | 0.7883 | 2.7948 | 24.570 | 1.3800 | 5.9794 | 57.433 |

4.4. Effects of Diagonal Specification

In the third experiment, we examine the effects of the diagonal specification for the BEKK and RBEKK models when the true model is full BEKK. For this purpose, we consider several measures to check the distance from the diagonal BEKK and RBEKK models to the full BEKK model. The non-diagonal indices are defined as:

$$\gamma = \|A^* - \text{diag}(A^*)\| + \|B^* - \text{diag}(B^*)\| \quad (\text{Diagonal BEKK}),$$

$$\gamma^r = \left\| A^* - \Omega^{1/2} \text{diag}(\Omega^{-1/2} A^* \Omega^{1/2}) \Omega^{-1/2} \right\| + \left\| B^* - \Omega^{1/2} \text{diag}(\Omega^{-1/2} B^* \Omega^{1/2}) \Omega^{-1/2} \right\| \quad (13)$$

(Diagonal RBEKK),

where $\text{diag}(Y)$ creates a diagonal matrix from the square matrix Y . Using the non-diagonal indices, we can calculate the theoretical distance of the diagonal BEKK and RBEKK models.

For the remaining measures, we use the estimated values of the parameters of these models. The maximized log-likelihood $L_{T,h}(\hat{\theta})$ is used, as is the average of the Frobenius norm of the difference of the conditional covariance matrices, as explained above.

Using these measures, the following Monte Carlo simulations investigate the effects of the diagonal specification of the BEKK and RBEKK models when the true model is bivariate full BEKK. For this purpose, we consider the specification for (4):

$$A_0^* = wD_{a1} + (1-w)\Omega_0^{1/2}D_{a2}\Omega_0^{-1/2}, \quad B_0^* = wD_{b1} + (1-w)\Omega_0^{1/2}D_{b2}\Omega_0^{-1/2}, \quad (14)$$

for $0 \leq w \leq 1$, where D_{a1} , D_{a2} , D_{b1} , and D_{b2} are diagonal matrices. When $w = 1$, the specification reduces to the diagonal BEKK model, while it becomes the diagonal RBEKK model for $w = 0$. Except for these endpoints, the full BEKK specification provides a non-diagonal structure for A_0^* and B_0^* in (4) and A_0 and B_0 in (5). For the specification in (4), the non-diagonal indices provide the linear functions of w :

$$\gamma_w = \xi(1-w), \quad \gamma_w^r = \xi_r w,$$

where

$$\xi = \left\| \Omega_0^{1/2}D_{a2}\Omega_0^{-1/2} - \text{diag}(\Omega_0^{1/2}D_{a2}\Omega_0^{-1/2}) \right\| + \left\| \Omega_0^{1/2}D_{b2}\Omega_0^{-1/2} - \text{diag}(\Omega_0^{1/2}D_{b2}\Omega_0^{-1/2}) \right\|,$$

$$\xi_r = \left\| D_{a1} - \Omega_0^{1/2}\text{diag}(\Omega_0^{-1/2}D_{a1}\Omega_0^{1/2})\Omega_0^{-1/2} \right\| + \left\| D_{b1} - \Omega_0^{1/2}\text{diag}(\Omega_0^{-1/2}D_{b1}\Omega_0^{1/2})\Omega_0^{-1/2} \right\|,$$

to calculate the theoretical distances. Consider the parameter settings for the DGPs as:

DGP4_w : (Ω_0, A_0, B_0) in DGP1, with $D_{a1} = D_{a2} = A_0$ and $D_{b1} = D_{b2} = B_0$ in (14),

DGP5_w : (Ω_0, A_0, B_0) in DGP2, with $D_{a1} = D_{a2} = A_0$ and $D_{b1} = D_{b2} = B_0$ in (14).

Set $w = 0, 0.1, \dots, 1$ to examine 11 cases, with $T = 500$ and the number of replications set to 100. We estimate the diagonal RBEKK model using the 2sQML method, while VT-QML is used for the diagonal BEKK model.

Figures 3 and 4 show the sample means of the average bias for the conditional covariance matrices and sample means of the maximized log-likelihood function for DGP3 and DGP4, respectively. As expected from the structure, the superiority of the diagonal models depends on the structure of the true BEKK model. If w is closer to zero, the diagonal RBEKK model is preferred. The non-diagonality indices are

$$\text{DGP4}_w : \gamma_w = 0.0212(1-w), \quad \gamma_w^r = 0.0309w, \quad \text{crossing at } w^\dagger = 0.407,$$

$$\text{DGP5}_w : \gamma_w = 0.0563(1-w), \quad \gamma_w^r = 0.0614w, \quad \text{crossing at } w^\dagger = 0.479,$$

and these theoretical values of w^\dagger correspond to the intersections shown in Figures 3b and 4b, respectively. The Akaike information criterion and Bayesian information criterion lead to the same conclusion, as the numbers of parameters in these two models are the same.

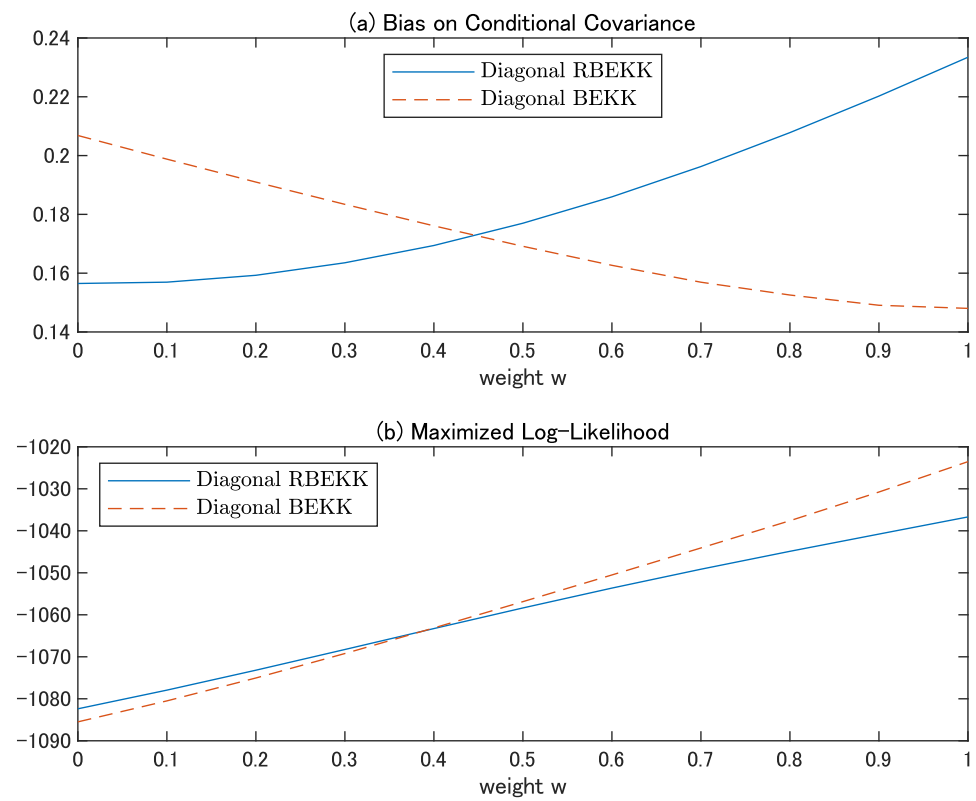


Figure 3. Comparison of the diagonal specifications for the BEKK and RBEKK models: DGP4.

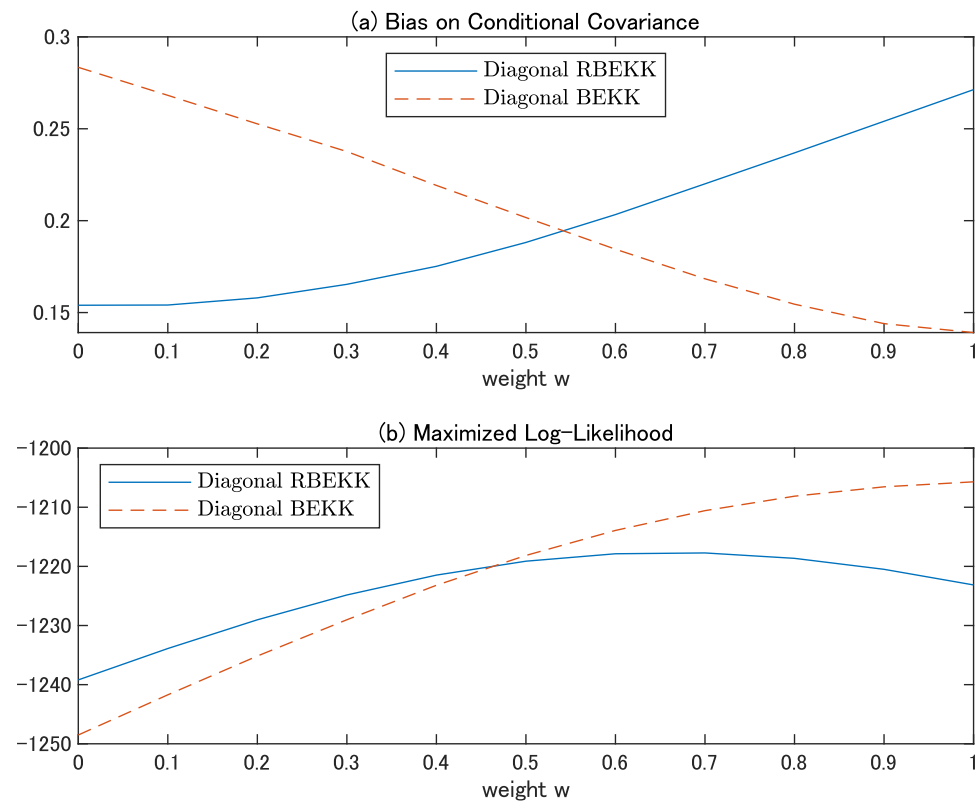


Figure 4. Comparison of the diagonal specifications for the BEKK and RBEKK models: DGP5.

4.5. Heavy Tails and Moment Conditions

The last experiment uses DGPs 1 and 2 with the multivariate standardized t distribution and the degree-of-freedom parameter ν (denoted by $St(\nu)$), instead of the multivariate standard normal distribution. We consider three cases: (i) a heavy-tailed distribution ($\nu = 7$), which satisfies Assumption 4(a); (ii) DGPs in which the sixth moments are not finite ($\nu = 5$); and (iii) DGPs in which the fourth moments are not finite ($\nu = 3$). The latter two cases imply that the 2sQML estimator is consistent, but its asymptotic normality is not guaranteed. As an alternative approach, we may use the parameter condition derived from Theorem 3 of Hafner (2003). To save space, we omitted the tables for the sample mean, standard deviations, and root mean squared error of the 2sQML estimates, which are available in the Supplementary Materials.

Figure 5 shows the histograms and QQ plots of 2sQML estimates for DGP1 with $St(7)$. The Monte Carlo results indicate that the bias is negligible, and the standard deviations are larger than in the case of standard normal distribution. The distributions of the estimates of Ω_{11} , A_{11} , and B_{11} are similar to those of Figure 1. This result supports Proposition 2.

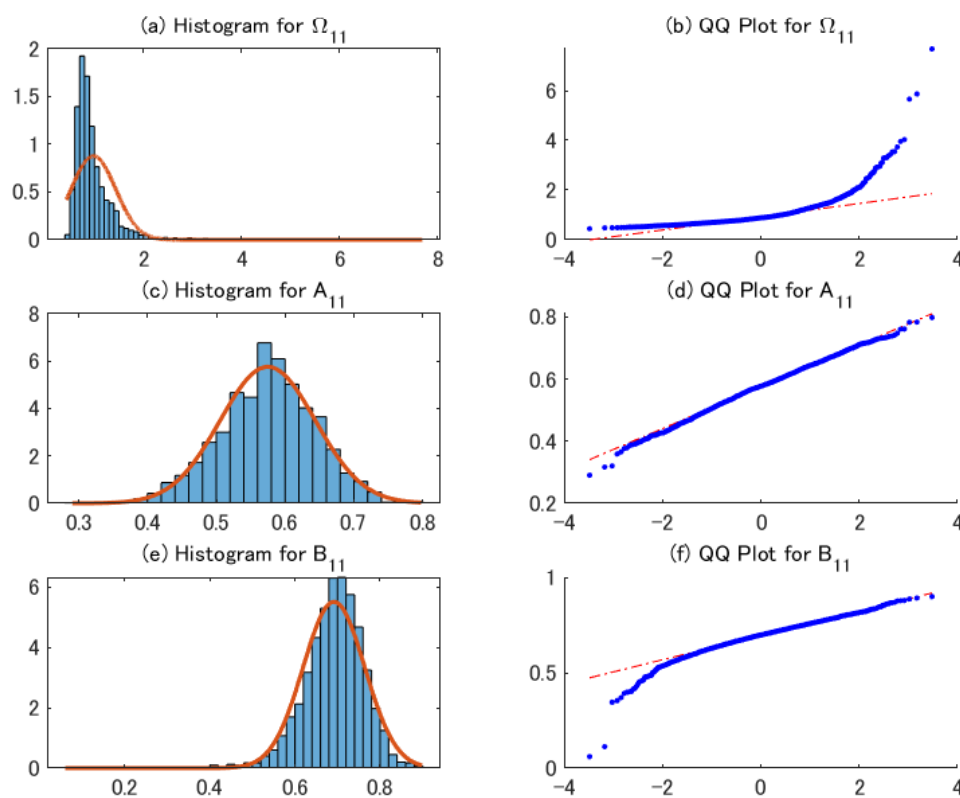


Figure 5. Histograms and QQ plots of 2sQML estimates for DGP1 with $St(7)$.

Figure 6 shows the histograms and QQ plots of 2sQML estimates for DGP1 with $St(5)$. Although the sixth moments are not finite, the distributions of the estimates of Ω_{11} , A_{11} , and B_{11} are close to those of Figure 5. This result implies that we can relax Assumption 4 to guarantee asymptotic normality for the second-step estimator.

For the DGPs with $St(3)$, the Monte Carlo result shown in the Supplementary Materials indicates that the bias of the estimators for Ω and A are relatively small, compared with those of B_{11} and B_{22} . Figure 7 shows the histograms and QQ plots of 2sQML estimates for DGP1 with $St(5)$. The QQ plot of Figure 7a shows the instability of the first-step estimator. The instability affects the standard deviations of the second-step estimator. Figure 7e implies that the effects are more serious on the estimates of B_{11} , with a pressure towards zero. The result indicates that a larger sample size is required in order to improve the estimates for the parameters of B .

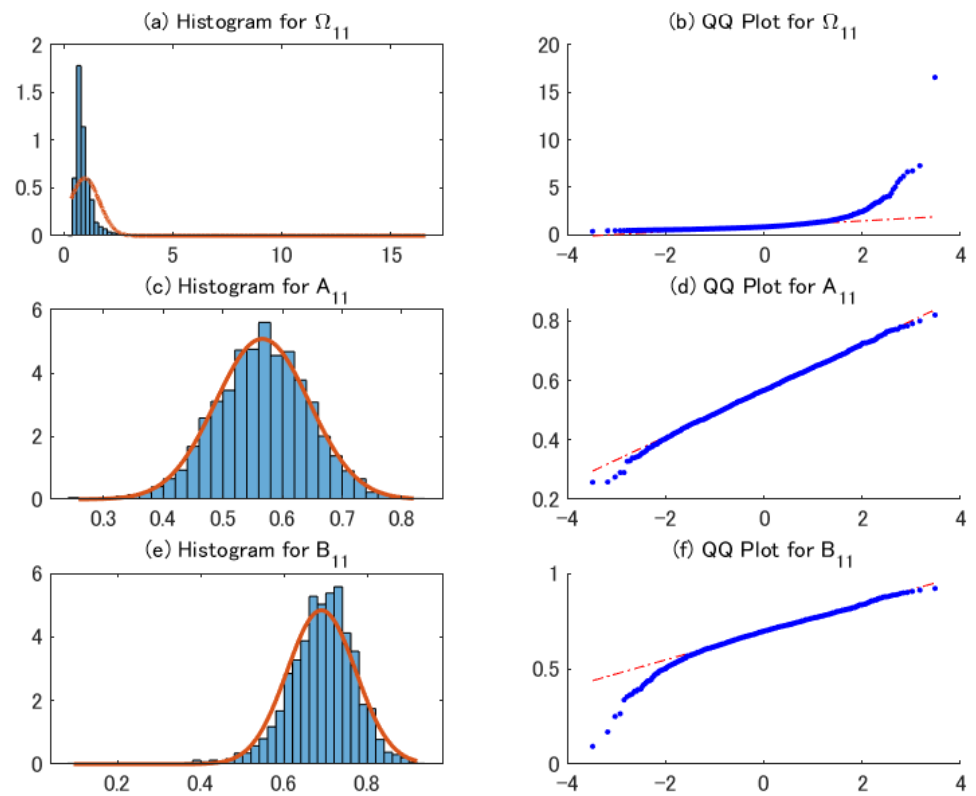


Figure 6. Histograms and QQ plots of 2sQML estimates for DGP1 with $St(5)$.

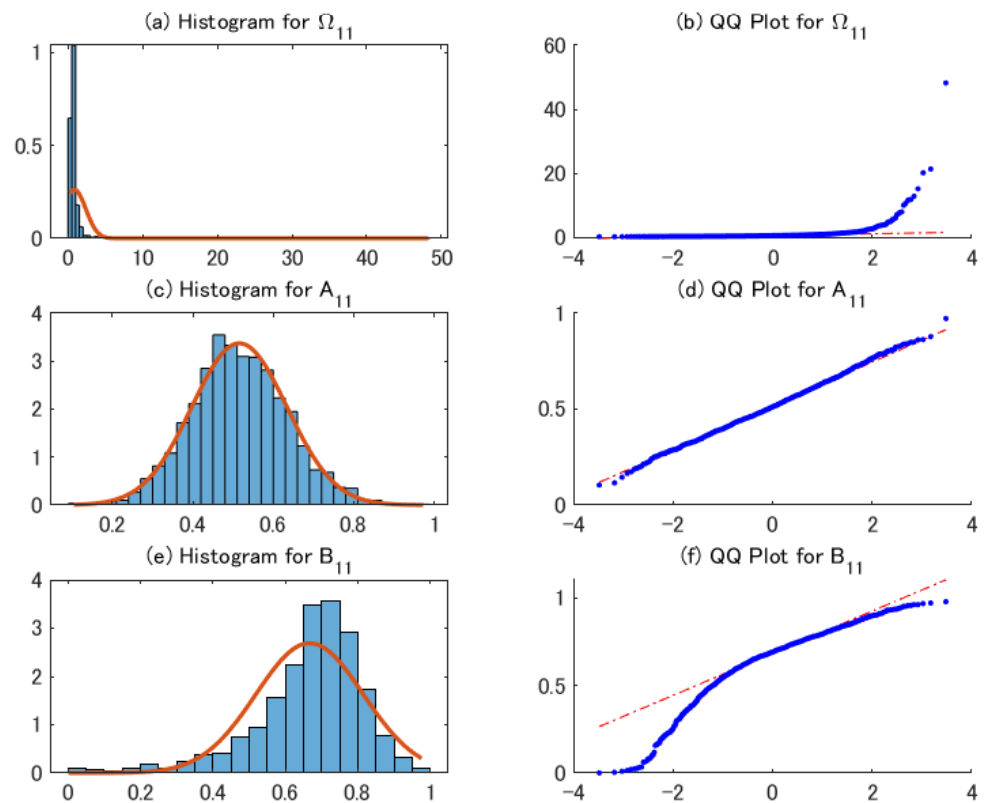


Figure 7. Histograms and QQ plots of 2sQML estimates for DGP1 with $St(3)$.

The Monte Carlo experiments show that the finite sample properties of the 2sQML estimator are satisfactory and that the average distance between the true and estimated

covariance matrices indicates that the difference between the two diagonal BEKK models is not negligible. Because it is difficult to estimate the fully parametrized BEKK model for higher d , it is necessary to examine the model specifications, as in [Noureldin et al. \(2014\)](#), using parsimonious specifications. The Monte Carlo experiments indicate that we may relax the assumption for the sixth moments for the second step estimator.

5. Empirical Analysis

In this section, we assess the diagonal specification of the RBEKK model compared with the diagonal BEKK model. For this purpose, we focus on the out-of-sample forecasts evaluation, adopting the approach of [Engle and Colacito \(2006\)](#), for the mean-variance portfolio. We calculate the returns of stocks listed on the Dow Jones Industrial Average, except for Dow Inc. ($d = 29$), for the period starting from 18 February 2010 to 23 January 2020, yielding 2500 observations. We exclude Dow Inc. since it went public on April 1, 2019. Fixing the sample size as $T = 2000$, we use rolling windows to obtain one-step-ahead forecasts for the last 500 observations for model i ($i = 1, 2$), denoted by $\hat{H}_t^{(i)}$ ($t = 2001, \dots, 2500$). While Model 1 is the diagonal RBEKK model, Model 2 is the diagonal BEKK.

For our portfolio analysis, we consider the minimum variance portfolio, which gives the weight vector for model i chosen at the end of time $t - 1$, $w_{it} = \hat{H}_t^{(i)-1} \iota / (\iota' \hat{H}_t^{(i)-1} \iota)$, where ι is the $d \times 1$ vector of ones. [Engle and Colacito \(2006\)](#) show that the realized portfolio volatility is the smallest one when the variance-covariance matrices are correctly specified. We define the distance based on the difference of the squared returns of the two portfolios as:

$$e_t = \{w'_{1t} X_t\}^2 - \{w'_{2t} X_t\}^2. \quad (15)$$

Since the portfolio variances are the same if the forecasts of the covariance matrices are the same, we examine the null hypothesis $H_0 : E(e_t) = 0$ using the [Diebold and Mariano \(1995\)](#) test. In this case, the test can be constructed in the following manner. Consider the linear regression model given by $e_t = \mu_e + u_{et}$ with $E(u_{et}) = 0$, and test $H_0 : \mu_e = 0$ using the heteroskedasticity- and autocorrelation-consistent standard errors. If the mean of e_t is negative (positive), the diagonal RBEKK (the diagonal BEKK) model is preferred.

Table 4 indicates that the [Engle and Colacito \(2006\)](#) test fails to reject the null hypothesis that the two forecasts are equivalent. Figure 8 shows the difference of squared portfolio returns, defined by Equation (15), accompanied by the 95% confidence interval. With a few exceptions, there is no significant difference between the two portfolio weights calculated by the forecasts of the two different models. For instance, at time $t = 2178$, the forecast by the diagonal RBEKK is preferred. As discussed in Section 4.4, the choice of the diagonal RBEKK or the diagonal BEKK model depends on the true structure of the full BEKK. Figure 8 supports the result of the [Engle and Colacito \(2006\)](#) test in Table 4. To improve the forecasting performance, we may consider a more general rotation matrix, as in [Asai and McAleer \(2020\)](#) and [Hafner et al. \(2020\)](#).

Table 4. Engle and Colacito (2006) test.

| Parameter | Estimate | HAC S.E. | t-Value |
|-----------|----------|----------|---------|
| μ_e | 0.0010 | 0.0043 | 0.2192 |

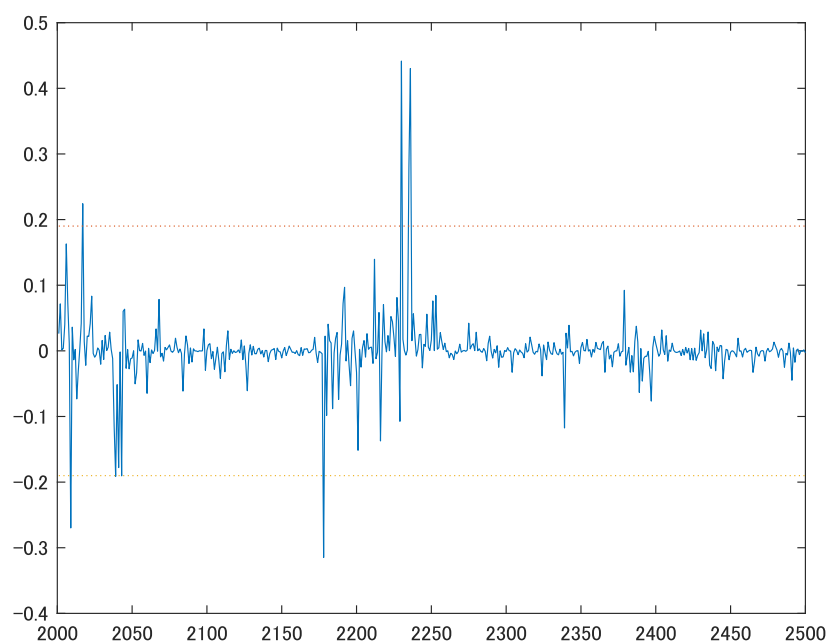


Figure 8. Difference of squared portfolio returns. Note: The solid line shows the value of e_t , defined by Equation (15), while the dotted lines indicate the confidence interval.

6. Conclusions

For the RBEKK-GARCH model, we show the consistency and asymptotic normality of the 2sQML estimator under weak conditions. The 2sQML estimation uses the unconditional covariance matrix for the first step and rotates the observed vector to have the identity matrix for its sample covariance matrix. The second step conducts the QML estimation for the remaining parameters. While we require second-order moments for consistency because of the estimation of the covariance matrix, we need finite sixth-order moments for asymptotic normality, as in Pedersen and Rahbek (2014). We also show the asymptotic relation of the 2sQML estimator for the RBEKK model and the VT-QML estimator for the VT-BEKK model. The Monte Carlo results show that the finite sample properties of the 2sQML estimator are satisfactory, and that the adequacy of the diagonal RBEKK depends on the structure of the true parameters. The empirical result for the returns of stocks listed on the DOW30 indicates that the diagonal RBEKK and diagonal BEKK models are competitive, with the superiority of each model changing over time.

As an extension of the dynamic conditional correlation (DCC) model of Engle (2002), Noureldin et al. (2014) suggested rotated DCC models (for a caveat about the regularity conditions underlying DCC, see McAleer (2018)). We can apply the rotation to the different kinds of correlation models suggested by McAleer et al. (2008) and Tse and Tsui (2002). Together with such extensions, the derivation of asymptotic theory for the rotated DCC models is an important direction for future research.

Supplementary Materials: The following are available at <https://www.mdpi.com/article/10.3390/econometrics9020021/s1>.

Author Contributions: All authors contributed equally to this work. All authors have read and agreed to the published version of the manuscript.

Funding: Japan Society for the Promotion of Science (19K01594).

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Not applicable.

Acknowledgments: The authors are most grateful to the editor, four anonymous reviewers, and Yoshihisa Baba for very helpful comments and suggestions. The first author acknowledges the financial support of the Japan Ministry of Education, Culture, Sports, Science and Technology, Japan Society for the Promotion of Science, and the Australian Academy of Science. The second author thanks the Ministry of Science and Technology (MOST) for financial support. The third author is most grateful for the financial support of the Australian Research Council, Ministry of Science and Technology (MOST), Taiwan, and the Japan Society for the Promotion of Science.

Conflicts of Interest: The authors declare no conflict of interest. The funders had no role in the design of the study; in the collection, analysis, or interpretation of data; in the writing of the manuscript; or in the decision to publish the results.

Abbreviations

The following abbreviations are used in this manuscript:

| | |
|-------|---|
| 2sQML | Two-step quasi-maximum likelihood |
| ARCH | Autoregressive conditional heteroskedasticity |
| BEKK | Baba, Engle, Kraft, and Kroner |
| DCC | Dynamic conditional correlation |
| GARCH | Generalized autoregressive conditional heteroskedasticity |
| QML | Quasi-maximum likelihood |
| RBEKK | Rotated BEKK |
| VT | Variance targeting |

Appendix A

Appendix A.1. Derivatives of the Log-Likelihood Function

Although Pedersen and Rahbek (2014) demonstrated the derivatives with respect to Ω , A^* , and B^* , they are not applicable, as A^* and B^* in (2) depend on $\Omega^{1/2}$ and $\Omega^{-1/2}$ in the RBEKK models (6) and (7), respectively. Related to this issue, we need the following lemma to show the derivatives of the log-likelihood function.

Lemma A1.

$$\frac{\partial \text{vec}(\Omega^{1/2})}{\partial \omega'} = \left[(\Omega^{1/2} \otimes I_d) + (I_d \otimes \Omega^{1/2}) \right]^{-1},$$

$$\frac{\partial \text{vec}(\Omega^{-1/2})}{\partial \omega'} = - \left[(\Omega^{-1/2} \otimes I_d) + (I_d \otimes \Omega^{-1/2}) \right]^{-1} (\Omega^{-1})^{\otimes 2}.$$

Proof. According to the product rule, it is straightforward to obtain

$$\frac{\partial \omega}{\partial \omega'} = \frac{\partial \text{vec}(\Omega^{1/2} \Omega^{1/2})}{\partial \omega'} = \left[(\Omega^{1/2} \otimes I_d) + (I_d \otimes \Omega^{1/2}) \right] \frac{\partial \text{vec}(\Omega^{1/2})}{\partial \omega'}.$$

Because $\Omega^{1/2}$ is positive definite, we obtain the result. A similar application produces the following:

$$\frac{\partial \text{vec}(\Omega^{-1})}{\partial \omega'} = \frac{\partial \text{vec}(\Omega^{-1/2} \Omega^{-1/2})}{\partial \omega'} = \left[(\Omega^{-1/2} \otimes I_d) + (I_d \otimes \Omega^{-1/2}) \right] \frac{\partial \text{vec}(\Omega^{-1/2})}{\partial \omega'}.$$

From the derivative of the inverse of the symmetric matrix shown by 10.6.1(1) of Lütkepohl (1996), we obtain the second result. \square

The gradient and Hessian matrices of the log-likelihood function are given by

$$\frac{\partial L_T}{\partial \theta} = \frac{1}{T} \sum_{t=1}^T \frac{\partial l_t}{\partial \theta'}, \quad \frac{\partial^2 L_T}{\partial \theta \partial \theta'} = \frac{1}{T} \sum_{t=1}^T \frac{\partial^2 l_t}{\partial \theta \partial \theta'}.$$

Applying the chain rule and product rule, we obtain

$$\begin{aligned} \frac{\partial l_t}{\partial \theta} &= \frac{\partial \text{vec}(H_t)'}{\partial \theta} \frac{\partial l_t}{\partial \text{vec}(H_t)'}, \\ \frac{\partial^2 l_t}{\partial \theta_i \partial \theta_j} &= \frac{\partial^2 \text{vec}(H_t)'}{\partial \theta_i \partial \theta_j} \frac{\partial l_t}{\partial \text{vec}(H_t)'} + \frac{\partial \text{vec}(H_t)'}{\partial \theta_i} \frac{\partial^2 l_t}{\partial \text{vec}(H_t) \partial \text{vec}(H_t)} \frac{\partial \text{vec}(H_t)}{\partial \theta_j}, \end{aligned} \tag{A1}$$

where θ_i ($i = 1, \dots, 3d^2$) is the i th element of θ :

$$\begin{aligned} \frac{\partial l_t}{\partial H_t} &= -\frac{1}{2} H_t^{-1} + \frac{1}{2} H_t^{-1} X_t X_t' H_t^{-1}, \\ \frac{\partial^2 l_t}{\partial \text{vec}(H_t) \partial \text{vec}(H_t)} &= \frac{1}{2} \left[I_{d^2} - \{ (H_t^{-1} X_t X_t) \otimes I_d \} - \{ I_d \otimes (H_t^{-1} X_t X_t) \} \right] (H_t^{-1})^{\otimes 2}. \end{aligned} \tag{A2}$$

The first equation of (A2) uses 10.3.2(23) and 10.3.3(10) of Lütkepohl (1996), while we apply 10.6.1(1) for the second equation.

From Lemma A1, the product rule, and the chain rule, we obtain the first derivatives:

$$\begin{aligned} \frac{\partial \text{vec}(H_t)}{\partial \omega'} &= \left[(\Omega^{1/2} \underline{H}_t \otimes I_d) + (I_d \otimes \Omega^{1/2} \underline{H}_t) \right] \left[(\Omega^{1/2} \otimes I_d) + (I_d \otimes \Omega^{1/2}) \right]^{-1} \\ &\quad + (\Omega^{1/2})^{\otimes 2} \frac{\partial \text{vec}(\underline{H}_t)}{\partial \omega'}, \\ \frac{\partial \text{vec}(H_t)}{\partial \lambda'} &= (\Omega^{1/2})^{\otimes 2} \frac{\partial \text{vec}(\underline{H}_t)}{\partial \lambda'}, \end{aligned} \tag{A3}$$

and

$$\begin{aligned} \frac{\partial \text{vec}(\underline{H}_t)}{\partial \omega'} &= B^{\otimes 2} \frac{\partial \text{vec}(\underline{H}_{t-1})}{\partial \omega'} - A^{\otimes 2} \left[(I_d \otimes \Omega^{-1/2} X_{t-1} X_{t-1}') + (\Omega^{-1/2} X_{t-1} X_{t-1}' \otimes I_d) \right] \\ &\quad \times \left[(\Omega^{-1/2} \otimes I_d) + (I_d \otimes \Omega^{-1/2}) \right]^{-1} (\Omega^{-1})^{\otimes 2}, \\ \frac{\partial \text{vec}(\underline{H}_t)}{\partial \alpha'} &= B^{\otimes 2} \frac{\partial \text{vec}(\underline{H}_{t-1})}{\partial \alpha'} + \left(A \{ \Omega^{-1/2} X_{t-1} X_{t-1}' \Omega^{-1/2} - I_d \} \otimes I_d \right) \\ &\quad + \left(I_d \otimes A \{ \Omega^{-1/2} X_{t-1} X_{t-1}' \Omega^{-1/2} - I_d \} \right) C_{dd}, \\ \frac{\partial \text{vec}(\underline{H}_t)}{\partial \beta'} &= B^{\otimes 2} \frac{\partial \text{vec}(\underline{H}_{t-1})}{\partial \beta'} + (B \{ \underline{H}_{t-1} - I_d \} \otimes I_d) + (I_d \otimes B \{ \underline{H}_{t-1} - I_d \}) C_{dd}, \end{aligned} \tag{A4}$$

where C_{dd} is the commutation matrix, which consists of one and zero satisfying $\text{vec}(A') = C_{dd} \text{vec}(A)$.

Similarly, the second derivatives of H_t are given by

$$\begin{aligned} \frac{\partial^2 \text{vec}(H_t)}{\partial \omega_i \partial \omega_j} &= \left[\left(\Omega^{1/2} \frac{\partial \underline{H}_t}{\partial \omega_i} \otimes I_d \right) + \left(I_d \otimes \Omega^{1/2} \frac{\partial \underline{H}_t}{\partial \omega_i} \right) \right] \left[(\Omega^{1/2} \otimes I_d) + (I_d \otimes \Omega^{1/2}) \right]^{-1} e^{(j)} \\ &\quad + (\Omega^{1/2})^{\otimes 2} \frac{\partial^2 \text{vec}(\underline{H}_t)}{\partial \omega_i \partial \omega_j} \quad (i, j = 1, \dots, d^2), \\ \frac{\partial \text{vec}(H_t)}{\partial \lambda_i \partial \lambda_j} &= (\Omega^{1/2})^{\otimes 2} \frac{\partial \text{vec}(\underline{H}_t)}{\partial \lambda_i \partial \lambda_j} \quad (i, j = 1, \dots, 2d^2), \\ \frac{\partial^2 \text{vec}(H_t)}{\partial \lambda_i \partial \omega_j} &= \left[\left(\Omega^{1/2} \frac{\partial \underline{H}_t}{\partial \lambda_i} \otimes I_d \right) + \left(I_d \otimes \Omega^{1/2} \frac{\partial \underline{H}_t}{\partial \lambda_i} \right) \right] \left[(\Omega^{1/2} \otimes I_d) + (I_d \otimes \Omega^{1/2}) \right]^{-1} e^{(j)}, \\ &\quad + (\Omega^{1/2})^{\otimes 2} \frac{\partial^2 \text{vec}(\underline{H}_t)}{\partial \lambda_i \partial \omega_j} \quad (i = 1, \dots, 2d^2, j = 1, \dots, d^2), \end{aligned}$$

where $e^{(j)}$ is a $d^2 \times 1$ vector of zeros except for the j th element, which takes one. We omit the derivatives of H_t .

Before we proceed, we show the equivalence of Assumptions 1(b) and 2.

Lemma A2. For the RBEKK model defined by (4) and (5), it can be shown that

$$\rho((A^* \otimes A^*) + (B^* \otimes B^*)) = \rho((A \otimes A) + (B \otimes B)).$$

Proof. Noting that

$$(A^* \otimes A^*) + (B^* \otimes B^*) = (\Omega^{1/2} \otimes \Omega^{1/2})\{(A \otimes A) + (B \otimes B)\}(\Omega^{-1/2} \otimes \Omega^{-1/2}),$$

5.2.1(8) of Lütkepohl (1996) indicates that the eigenvalues of $(A^* \otimes A^*) + (B^* \otimes B^*)$ are the same as those of $(A \otimes A) + (B \otimes B)$, which proves the lemma. \square

Appendix A.2. Proofs of Propositions 1–3

For notational convenience, let $H_{0t} = H_t(\omega_0, \lambda_0)$.

Proof of Propositions 1 and 2. Using the log-likelihood function as well as the gradient and Hessian matrices in Appendix A.1, we can show strong consistency and asymptotic normality using similar arguments as in the proofs of Theorems 4.1 and 4.2 of Pedersen and Rahbek (2014), respectively. See the Supplementary Materials for the details. \square

Lemma A3. Under the assumptions of Proposition 2, as $T \rightarrow \infty$,

$$\sqrt{T}(\hat{\theta}^* - \theta_0^*) \xrightarrow{d} N(\mathbf{0}, R_0 Q_0 \Gamma_0 Q_0' R_0'),$$

where

$$R_0 = \begin{pmatrix} I_{d^2} & O_{d^2 \times d^2} & O_{d^2 \times d^2} \\ R_{21,0} & (\Omega_0^{-1/2} \otimes \Omega_0^{1/2}) & O_{d^2 \times d^2} \\ R_{31,0} & O_{d^2 \times d^2} & (\Omega_0^{-1/2} \otimes \Omega_0^{1/2}) \end{pmatrix}, \tag{A5}$$

$$R_{21,0} = (A_0 \Omega_0^{-1/2} \otimes I_d) \left[(\Omega_0^{1/2} \otimes I_d) + (\Omega_0^{1/2} \otimes I_d) \right]^{-1} \\ - (I_d \otimes \Omega_0^{1/2} A_0) \left[(\Omega_0^{-1/2} \otimes I_d) + (\Omega_0^{-1/2} \otimes I_d) \right]^{-1} (\Omega_0^{-1})^{\otimes 2},$$

$$R_{31,0} = (B_0 \Omega_0^{-1/2} \otimes I_d) \left[(\Omega_0^{1/2} \otimes I_d) + (\Omega_0^{1/2} \otimes I_d) \right]^{-1} \\ - (I_d \otimes \Omega_0^{1/2} B_0) \left[(\Omega_0^{-1/2} \otimes I_d) + (\Omega_0^{-1/2} \otimes I_d) \right]^{-1} (\Omega_0^{-1})^{\otimes 2},$$

with Q_0 and Γ_0 as stated in Proposition 2.

Proof of Lemma A3. From the definition of A^* and B^* and rule of vectorization, $\alpha^* = (\Omega^{-1/2} \otimes \Omega^{1/2})\alpha$ and $\beta^* = (\Omega^{-1/2} \otimes \Omega^{1/2})\beta$. Define $R_0 = \frac{\partial \theta^*}{\partial \theta'} \Big|_{\theta=\theta_0}$. From Proposition 2 and the delta method, $\sqrt{T}(\hat{\theta}^* - \theta_0^*) \xrightarrow{d} N(\mathbf{0}, R_0 Q_0 \Gamma_0 Q_0' R)$. Since

$$R_{21} = \frac{\partial \alpha^*}{\partial \omega'} = (A \Omega^{-1/2} \otimes I_d) \frac{\partial \text{vec}(\Omega^{1/2})}{\partial \omega'} + (I_d \otimes \Omega^{1/2} A) \frac{\partial \text{vec}(\Omega^{-1/2})}{\partial \omega'},$$

$$R_{31} = \frac{\partial \beta^*}{\partial \omega'} = (B \Omega^{-1/2} \otimes I_d) \frac{\partial \text{vec}(\Omega^{1/2})}{\partial \omega'} + (I_d \otimes \Omega^{1/2} B) \frac{\partial \text{vec}(\Omega^{-1/2})}{\partial \omega'},$$

applying Lemma A1 under the true vector yields $R_{21,0}$ and $R_{31,0}$ in Lemma A3. \square

Lemma A4. Under the assumptions of Proposition 2, the Γ_0 stated in Proposition 2 can be written as

$$\Gamma_0 = E \left[Y_t(\omega_0, \lambda_0) \text{vec}(Z_t Z_t' - I_d) (\text{vec}(Z_t Z_t' - I_d))' Y_t'(\omega_0, \lambda_0) \right], \tag{A6}$$

where

$$Y_t(\omega_0, \lambda_0) = \begin{pmatrix} Y_{\omega t}(\omega_0, \lambda_0) \\ Y_{\alpha t}(\omega_0, \lambda_0) \\ Y_{\beta t}(\omega_0, \lambda_0) \end{pmatrix} = \begin{pmatrix} (\Omega_0^{1/2})^{\otimes 2} (I_{d^2} - A_0^{\otimes 2} - B_0^{\otimes 2})^{-1} (I_{d^2} - B_0^{\otimes 2}) (\Omega_0^{-1/2} H_{0t}^{1/2})^{\otimes 2} \\ \frac{1}{2} \left[\sum_{i=0}^{\infty} (B_0^{\otimes 2})^i N_{t-1-i}(\omega_0, \lambda_0) \right]' (\Omega_0^{1/2} H_{0t}^{-1/2})^{\otimes 2} \\ \frac{1}{2} \left[\sum_{i=0}^{\infty} (B_0^{\otimes 2})^i \tilde{N}_{t-1-i}(\omega_0, \lambda_0) \right]' (\Omega_0^{1/2} H_{0t}^{-1/2})^{\otimes 2} \end{pmatrix} \tag{A7}$$

with

$$N_t(\omega_0, \lambda_0) = [A_0(\Omega_0^{-1/2} X_t X_t' \Omega_0^{-1/2} - I_d) \otimes I_d] + [I_d \otimes A_0(\Omega_0^{-1/2} X_t X_t' \Omega_0^{-1/2} - I_d)] C_{dd}, \tag{A8}$$

$$\tilde{N}_t(\omega_0, \lambda_0) = [B_0(H_{0t} - I_d) \otimes I_d] + [I_d \otimes B_0(H_{0t} - I_d)] C_{dd}.$$

Proof of Lemma A4. Using an argument similar to the proof of Lemma B.8 in Pedersen and Rahbek (2014), we can show

$$\sqrt{T} \begin{pmatrix} \hat{\omega} - \omega_0 \\ \partial L_T(\omega_0, \lambda_0) / \partial \lambda \end{pmatrix} = \frac{1}{\sqrt{T}} \sum_{t=1}^T Y_t(\omega_0, \lambda_0) \text{vec}(Z_t Z_t' - I_d) + o_p(1), \tag{A9}$$

which implies the equivalence of the asymptotic covariance matrices on both sides. \square

Lemma A5. Under the assumptions of Proposition 2, the asymptotic covariance matrix for the VT-QML estimator of Pedersen and Rahbek (2014) is given by $Q_0^* \Gamma_0^* Q_0^*$, where

$$\Gamma_0^* = E \left[Y_t^*(\omega_0, \lambda_0^*) \text{vec}(Z_t Z_t' - I_d) (\text{vec}(Z_t Z_t' - I_d))' Y_t^{*'}(\omega_0, \lambda_0^*) \right], \tag{A10}$$

$$Q_0^* = \begin{pmatrix} I_{d^2} & O_{d^2 \times 2d^2} \\ -J_0^{*-1} K_0^* & -J_0^{*-1} \end{pmatrix}, \tag{A11}$$

$$K_0^* = E \left(\frac{\partial^2 l_t}{\partial \lambda^* \partial \omega'} \right), \quad J_0^* = E \left(\frac{\partial^2 l_t}{\partial \lambda^* \partial \lambda^{*'}} \right), \tag{A12}$$

with

$$Y_t^*(\omega_0, \lambda_0^*) = \begin{pmatrix} Y_{\omega t}^*(\omega_0, \lambda_0^*) \\ Y_{\alpha t}^*(\omega_0, \lambda_0^*) \\ Y_{\beta t}^*(\omega_0, \lambda_0^*) \end{pmatrix} = \begin{pmatrix} (I_{d^2} - (A_0^*)^{\otimes 2} - (B_0^*)^{\otimes 2})^{-1} (I_{d^2} - (B_0^*)^{\otimes 2}) (H_{0t}^{1/2})^{\otimes 2} \\ \frac{1}{2} \left[\sum_{i=0}^{\infty} ((B_0^*)^{\otimes 2})^i M_{t-1-i}(\omega_0, \lambda_0^*) \right]' (H_{0t}^{-1/2})^{\otimes 2} \\ \frac{1}{2} \left[\sum_{i=0}^{\infty} ((B_0^*)^{\otimes 2})^i \tilde{M}_{t-1-i}(\omega_0, \lambda_0^*) \right]' (H_{0t}^{-1/2})^{\otimes 2} \end{pmatrix}, \tag{A13}$$

and

$$M_t(\omega_0, \lambda_0^*) = [A_0^*(X_t X_t' - \Omega_0) \otimes I_d] + [I_d \otimes A_0^*(X_t X_t' - \Omega_0)] C_{dd}, \tag{A14}$$

$$\tilde{M}_t(\omega_0, \lambda_0^*) = [B_0^*(H_{0t} - \Omega_0) \otimes I_d] + [I_d \otimes B_0^*(H_{0t} - \Omega_0)] C_{dd}.$$

Proof of Lemma A5. We can verify the result using Theorem 4.2 of Pedersen and Rahbek (2014). \square

Proof of Proposition 3. Lemma A3 shows the asymptotic distribution of $\hat{\theta}$ and corresponding asymptotic covariance matrix. Lemma A5 provides the asymptotic covariance matrix of the VT estimator. In the following, we show an alternative representation of the latter asymptotic covariance matrix.

By definition, we obtain $\lambda^* = P\lambda$, where

$$P = \begin{pmatrix} (\Omega^{-1/2} \otimes \Omega^{1/2}) & O_{d^2 \times d^2} \\ O_{d^2 \times d^2} & (\Omega^{-1/2} \otimes \Omega^{1/2}) \end{pmatrix}. \tag{A15}$$

First, consider the second derivatives of the t th contribution to the likelihood function to obtain

$$\frac{\partial^2 l_t}{\partial \lambda^* \partial \omega'} = P^{-1} \frac{\partial^2 l_t}{\partial \lambda \partial \omega'}, \quad \frac{\partial^2 l_t}{\partial \lambda^* \partial \lambda'} = P^{-1} \frac{\partial^2 l_t}{\partial \lambda \partial \lambda'} P^{-1}.$$

Then, we obtain $K_0^* = P_0^{-1} K_0$ and $J_0^* = P_0^{-1} J_0 P_0^{-1}$. For Q_0^* defined by Lemma A5,

$$Q_0^* = R_0^* Q_0 R_0^*, \tag{A16}$$

where

$$R_0^* = \begin{pmatrix} I_{d^2} & O_{d^2 \times 2d^2} \\ O_{2d^2 \times d^2} & P_0 \end{pmatrix}, \tag{A17}$$

with P_0 , which is evaluated at the true value Ω_0 .

Second, we show that

$$Y_t^*(\omega_0, \lambda_0^*) = R_0^{*-1} Y_t(\omega_0, \lambda_0). \tag{A18}$$

Noting that

$$\begin{aligned} I_{d^2} - (A_0^*)^{\otimes 2} - (B_0^*)^{\otimes 2} &= (\Omega_0^{1/2})^{\otimes 2} [I_{d^2} - A_0^{\otimes 2} - B_0^{\otimes 2}] (\Omega_0^{-1/2})^{\otimes 2}, \\ I_{d^2} - (B_0^*)^{\otimes 2} &= (\Omega_0^{1/2})^{\otimes 2} [I_{d^2} - B_0^{\otimes 2}] (\Omega_0^{-1/2})^{\otimes 2}, \end{aligned}$$

we can verify that $Y_{\omega t}^*(\omega_0, \lambda_0^*) = Y_{\omega t}(\omega_0, \lambda_0)$. For $Y_{\alpha t}^*(\omega_0, \lambda_0^*)$ and $Y_{\beta t}^*(\omega_0, \lambda_0^*)$, we obtain

$$\begin{aligned} [(B_0^*)^{\otimes 2}]^i &= [(\Omega_0^{1/2} B_0 \Omega_0^{-1/2})^{\otimes 2}]^i = [(\Omega_0^{1/2})^{\otimes 2} (B_0^{\otimes 2}) (\Omega_0^{-1/2})^{\otimes 2}]^i \\ &= (\Omega_0^{1/2})^{\otimes 2} (B_0^{\otimes 2})^i (\Omega_0^{-1/2})^{\otimes 2}. \end{aligned}$$

From 9.3.2(5)(a) of Lütkepohl (1996), $(\Omega_0^{-1/2} \otimes \Omega_0^{1/2}) C_{dd} = C_{dd} (\Omega_0^{1/2} \otimes \Omega_0^{-1/2})$. Hence,

$$\begin{aligned} M_t(\omega_0, \lambda_0^*) &= (\Omega_0^{1/2})^{\otimes 2} N_t(\omega_0, \lambda_0) (\Omega_0^{1/2} \otimes \Omega_0^{-1/2}), \\ \tilde{M}_t(\omega_0, \lambda_0^*) &= (\Omega_0^{1/2})^{\otimes 2} \tilde{N}_t(\omega_0, \lambda_0) (\Omega_0^{1/2} \otimes \Omega_0^{-1/2}). \end{aligned}$$

Combining these two results, we show that $Y_{\alpha t}^*(\omega_0, \lambda_0^*) = (\Omega_0^{1/2} \otimes \Omega_0^{-1/2}) Y_{\alpha t}(\omega_0, \lambda_0)$ and $Y_{\beta t}^*(\omega_0, \lambda_0^*) = (\Omega_0^{1/2} \otimes \Omega_0^{-1/2}) Y_{\beta t}(\omega_0, \lambda_0)$. Hence, (A18) holds.

Third, we can verify that $Q_0^* \Gamma_0^* Q_0^{*'} = R_0^* Q_0 \Gamma_0 Q_0' R_0^*$ from (A16), (A18), and Lemma A4. \square

References

Alexander, Carol, ed. 2001. Orthogonal GARCH. In *Mastering Risk*. Hoboken: Financial Times-Prentice Hall, pp. 21–28.

Asai, Manabu, and Michael McAleer. 2020. *Multivariate Hyper-Rotated GARCH-BEKK: Asymptotic Theory and Practice*. Tokyo: Soka University. Unpublished Paper.

Baba, Yoshihisa, Robert F. Engle, Dennis Kraft, and Kenneth F. Kroner. 1985. *Multivariate Simultaneous Generalized ARCH*. San Diego: University of California. Unpublished Paper.

- Bauwens, Luc, Sébastien Laurent, and Jeroen K. V. Rombouts. 2006. Multivariate GARCH models: A survey. *Journal of Applied Econometrics* 21: 79–109. [\[CrossRef\]](#)
- Bollerslev, Tim. 1986. Generalized autoregressive conditional heteroskedasticity. *Journal of Econometrics* 31: 307–27. [\[CrossRef\]](#)
- Boussama, Farid, Florian Fuchs, and Robert Stelzer. 2011. Stationarity and geometric ergodicity of BEKK multivariate GARCH models. *Stochastic Processes and Their Applications* 121: 2331–60. [\[CrossRef\]](#)
- Chang, Chia-Lin, and Michael McAleer. 2019. The fiction of full BEKK: Pricing fossil fuels and carbon emissions. *Finance Research Letters* 28: 11–19. [\[CrossRef\]](#)
- Comte, Fabienne, and Offer Lieberman. 2003. Asymptotic theory for multivariate GARCH processes. *Journal of Multivariate Analysis* 84: 61–84. [\[CrossRef\]](#)
- Diebold, Francis X., and Roberto S. Mariano. 1995. Comparing predictive accuracy. *Journal of Business & Economic Statistics* 13: 253–63.
- Engle, Robert F. 1982. Autoregressive conditional heteroskedasticity with estimates of the variance of United Kingdom inflation. *Econometrica* 50: 987–1007. [\[CrossRef\]](#)
- Engle, Robert F. 2002. Dynamic conditional correlation: A simple class of multivariate generalized autoregressive conditional heteroskedasticity models. *Journal of Business & Economic Statistics* 20: 339–50.
- Engle, Robert F., and Kenneth F. Kroner. 1995. Multivariate simultaneous generalized ARCH. *Econometric Theory* 11: 122–50. [\[CrossRef\]](#)
- Engle, Robert F., and Riccardo Colacito. 2006. Testing and valuing dynamic correlations for asset allocation. *Journal of Business & Economic Statistics* 24: 238–53.
- Franco, Christian, Lajos Horváth, and Jean-Michel Zakoïan. 2011. Merits and drawbacks of variance targeting in GARCH models. *Journal of Financial Econometrics* 9: 619–56. [\[CrossRef\]](#)
- Hafner, Christian M. 2003. Fourth moment structure of multivariate GARCH models. *Journal of Financial Econometrics* 1: 26–54. [\[CrossRef\]](#)
- Hafner, Christian M., Helmut Herwartz, and Simone Maxand. 2020. Identification of structural multivariate GARCH models. *Journal of Econometrics* doi:10.1016/j.jeconom.2020.07.019. [\[CrossRef\]](#)
- Hafner, Christian M., and Arie Preminger. 2009. On asymptotic theory for multivariate GARCH models. *Journal of Multivariate Analysis* 100: 2044–54. [\[CrossRef\]](#)
- Lanne, Markku, and Pentti Saikkonen. 2007. A multivariate generalized orthogonal factor GARCH model. *Journal of Business & Economic Statistics* 25: 61–75.
- Laurent, Sébastien, Jeroen V. K. Rombouts, and Francesco Violante. 2012. On the forecasting accuracy of multivariate GARCH models. *Journal of Applied Econometrics* 27: 934–55. [\[CrossRef\]](#)
- Lütkepohl, Helmut. 1996. *Handbook of Matrices*. New York: John Wiley & Sons.
- McAleer, Michael. 2005. Automated inference and learning in modeling financial volatility. *Econometric Theory* 21: 232–61. [\[CrossRef\]](#)
- McAleer, Michael. 2018. Stationarity and invertibility of a dynamic correlation matrix. *Kybernetika* 54: 363–74. [\[CrossRef\]](#)
- McAleer, Michael, Felix Chan, Suhejla Hoti, and Offer Lieberman. 2008. Generalized autoregressive conditional correlation. *Econometric Theory* 24: 1554–83. [\[CrossRef\]](#)
- Noureldin, Diaa, Neil Shephard, and Kevin Sheppard. 2014. Multivariate rotated ARCH models. *Journal of Econometrics* 179: 16–30. [\[CrossRef\]](#)
- Pedersen, Rasmus S., and Anders Rahbek. 2014. Multivariate variance targeting in the BEKK-GARCH model. *Econometrics Journal* 17: 24–55. [\[CrossRef\]](#)
- Silvennoinen, Annastiina, and Timo Teräsvirta. 2009. Multivariate GARCH models. In *Handbook of Financial Time Series*. Edited by Andersen, Toben G., Richard A. Davis, Jens-Peter Kreiss and Thomas Mikosch. New York: Springer, pp. 201–29.
- Tse, Y. K., and Albert K. C. Tsui. 2002. A multivariate generalized autoregressive conditional heteroscedasticity model with time-varying correlations. *Journal of Business & Economic Statistics* 20: 351–61.
- van der Weide, Roy. 2020. GO-GARCH: A multivariate generalized orthogonal GARCH model. *Journal of Applied Econometrics* 17: 549–64. [\[CrossRef\]](#)