Econometric Theory, **23**, 2007, 464–484. Printed in the United States of America.
DOI: 10.1017/S026646660707020X

ON THE PARAMETRIZATION OF MULTIVARIATE GARCH MODELS

WOLFGANG SCHERRER AND EVA RIBARITS
Institute of Mathematical Methods in Economics, Vienna University of Technology

This paper deals with issues of structure and parametrization of VECH models proposed in Bollerslev, Engle, and Wooldridge (1988) and Baba, Engle, Kraft, and Kroner (BEKK) models. Both general models and also restricted versions such as the widely used *diagonal* VECH (DVECH) and *factor* generalized autoregressive conditional heteroskedastic (F-GARCH) models are discussed. A simple algorithm is presented that checks whether a given VECH model may be cast as a BEKK model. It is shown that in the bivariate case BEKK models are as general as VECH models. In higher dimensional cases however, VECH models allow for more flexibility. In addition, a parametrization for a generic, i.e., open and dense, class of BEKK models is given, and the frequently cited parametrization by Engle and Kroner (1995, *Econometric Theory* 11, 122–150) is analyzed. Two shortcomings of the latter are pointed out. Finally, parametrizations for BEKK(p,q,K) models with $K \leq n$, including DVECH, F-GARCH, and a generalization of the latter, are discussed.

1. INTRODUCTION

The modeling of conditional variance and of conditional covariance matrices is an important issue in financial econometrics. For the scalar case there is a wide range of models, which have been successfully used in practice. In particular, generalized autoregressive conditional heteroskedastic (GARCH) models are standard tools to describe and predict changes in the volatility of financial time series. See, e.g., Bollerslev, Engle, and Nelson (1994), Pagan (1996), and Palm (1996) for a survey.

In this paper the focus is on multivariate GARCH models. In particular, we concentrate on the VECH model class proposed by Bollerslev, Engle, and Wooldridge (1988) and the Baba, Engle, Kraft, and Kroner (BEKK) model class (see Baba et al., 1991, and Engle and Kroner, 1995). The acronym VECH refers to the fact that these models are expressed in terms of vectorized conditional variance matrices. For a recent survey of multivariate GARCH models we refer to Bauwens, Laurent, and Rombout (2006).

This research was supported by the project P17065 "Identification of multivariate dynamic systems with a focus on dimension reduction" of the Austrian Science Foundation (FWF). Address correspondence to Wolfgang Scherrer, Institut für Wirtschaftsmathematik, Forschungsgruppe für Ökonometrie und Systemtheorie (EOS), Technische Universität Wien, Argentinierstr. 8/105-2, A-1040 Wien, Austria; e-mail: wolfgang.scherrer@tuwien.ac.at.

 © 2007 Cambridge University Press 0266-4666/07 $15.00

https://doi.org/10.1017/S026646660707020X Published online by Cambridge University Press

It is shown that there is a one-to-one relation between VECH terms and certain affine sets of symmetric matrices, which are referred to as $\mathcal{Q}$-sets in the following discussion. This link is the main tool used to analyze VECH and BEKK models in this paper. In particular, it is shown that a VECH term allows for an equivalent BEKK representation if and only if the corresponding $\mathcal{Q}$-set contains a positive semidefinite (p.s.d.) element. The corresponding BEKK parameter matrices are obtained from a factorization of such a p.s.d. element. This gives a simple-to-check characterization of VECH models that have an equivalent BEKK representation via a positive semidefinite program. Furthermore, this observation implies that the BEKK parameter matrices are not identifiable without adding further restrictions.

Another conclusion of this link is that in the bivariate case ($n = 2$) any *admissible* VECH model, i.e., any model that ensures positive definite (p.d.) conditional variance matrices for any sample path, has an equivalent BEKK representation. For more than two variables ($n > 2$) however, the class of admissible VECH models is more general than the class of BEKK models.

In addition, we analyze the parametrization of Engle and Kroner (1995) and show that this parametrization is not fully general and identifiable; i.e., the parametrization does not cover a generic set of BEKK models, and two different parameter vectors may represent the same process. Therefore an alternative parametrization is put forward that covers a generic set of BEKK models.

One of the main problems of VECH and BEKK models is the curse of dimensionality, because the number of parameters grows as $O(n^4)$ with the number of observed variables. This implies that full VECH or BEKK models are only feasible for small n. Nevertheless in our opinion, our analysis of general models is fruitful. It gives an understanding of the structure of BEKK models and their relation to VECH models. In particular, it shows the difficulties of deriving suitable parametrizations of BEKK models in general. For moderate n, in particular for the bivariate case, estimation of full BEKK models is an option. Thus the proposed parametrization might be useful because of its advantages over the mostly used Engle–Kroner parametrization. Finally, a good understanding of the structure of BEKK models is indispensable for defining and parametrization of more parsimonious subclasses.

In view of the preceding discussion BEKK model classes where the number K of additive components is restricted to be less than or equal to the number of variables also are considered. In particular, the popular diagonal VECH model and the factor GARCH (F-GARCH) model are discussed in terms of $\mathcal{Q}$-sets. Finally, a generalization for F-GARCH models is proposed.

The outline of the paper is as follows. Section 1.1 summarizes some basics concerning the notation used throughout the paper. In Section 2 the VECH and BEKK models are discussed, and some properties of these model classes are stated. Section 3 gives a parametrization of a generic class of BEKK models. Here "generic" means that this set is an open and dense subset of the set of all BEKK models. Furthermore, the parametrization given by Engle and Kroner

https://doi.org/10.1017/S026646660707020X Published online by Cambridge University Press

(1995) is discussed. In Section 4 a subset of BEKK models with a small number of additive terms, $K \leq n$, is considered. Section 5 concludes. Some technical results are collected in the Appendix.

1.1. Notation

Throughout the paper we will adhere to the following conventions: The symbols I_n, $0_{m\times n}$ denote the $n \times n$ identity matrix and the $m \times n$ zero matrix. If the respective size is clear from the context, the indices are omitted. The transposed, inverse, and transposed inverse of a matrix are denoted with X', X^{-1}, and X^{-T}, respectively. The inequality $H > 0$ means that H is a p.d. matrix. Analogously $H \geq 0$ is used for p.s.d. matrices. The square root of a p.s.d. matrix $H \geq 0$ is denoted by $H^{1/2}$, i.e., $H = H^{1/2}(H^{1/2})'$. The transposed of $H^{1/2}$ is denoted by $H^{T/2}$, and if $H^{1/2}$ is nonsingular its inverse and transposed inverse are denoted by $H^{-1/2}$ and $H^{-T/2}$ respectively.

The Kronecker product of two matrices A and B is denoted by $A \otimes B$. The operator $\odot$ flags the Hadamard product, $A \odot B \in \mathbb{R}^{m\times n}$, of two $m \times n$ matrices A and B, which is formed by the elementwise multiplication of their elements.

The vec operator stacks all elements a_{ij} of a matrix $A \in \mathbb{R}^{m\times n}$ column by column, i.e., $\mathrm{vec}(A) = (a_{11}, a_{21}, \ldots, a_{mn})'$. Given the vector of stacked entries the $\mathrm{mat}(\cdot)$ operator reconstructs the matrix, i.e., $\mathrm{mat}(\mathrm{vec}(A)) = A$. For square matrices $A \in \mathbb{R}^{n\times n}$ the vech operator stacks the diagonal and the lower diagonal entries, i.e., $\mathrm{vech}(A) = (a_{11}, a_{21}, \ldots, a_{n1}, a_{22}, \ldots, a_{nn})' \in \mathbb{R}^{n(n+1)/2}$. The index function

$$I(k,l) = (k - l + 1) + \sum_{s=1}^{l-1} (n + 1 - s) \tag{1.1}$$

gives the position of the (k,l)th element of A in $\mathrm{vech}(A)$. For symmetric matrices, $A = A'$, the math operator is the inverse operator to vech, i.e., $\mathrm{math}(\mathrm{vech}(A)) = A$. The $\underline{\mathrm{vech}}$ operator stacks the elements below the diagonal, i.e., for a square matrix $A \in \mathbb{R}^{n\times n}$, $\underline{\mathrm{vech}}(A) = (a_{21}, a_{31}, \ldots, a_{n1}, a_{32}, \ldots, a_{n,(n-1)})'$. The diagonal elements are obtained with the diag operator, i.e., $\mathrm{diag}(A) = (a_{11}, \ldots, a_{nn})'$. With a slight abuse of notation we use diag also to construct diagonal matrices from some given vector; i.e., $\mathrm{diag}(x) = \mathrm{diag}(x_1, \ldots, x_k) \in \mathbb{R}^{k\times k}$ is a diagonal matrix with elements x_i on its diagonal.

For a square matrix $X \in \mathbb{R}^{m\times m}$ with eigenvalues $\lambda_1, \ldots, \lambda_m$, $\rho(X) := \max_{1\leq i\leq m}(|\lambda_m|)$ denotes the spectral radius of X. The minimum eigenvalue of a symmetric matrix $X = X'$ is denoted as $\lambda_{\min}(X)$.

Let $\mathcal{S}_m \subset \mathbb{R}^{m\times m}$ be the set of all symmetric $m \times m$ matrices. The set $\mathcal{S}_m^+ = \{Q \in \mathcal{S}_m | Q \geq 0\}$ is the set of all p.s.d. symmetric matrices, and $\mathcal{S}_{m,K}^+ \subseteq \mathcal{S}_m^+$ denotes the set of p.s.d. symmetric matrices of rank K. In particular $\mathcal{S}_{m,m}^+$ is the set of all p.d. matrices of order m. Furthermore, $\mathcal{A}_m \subseteq \mathbb{R}^{m\times m}$ denotes the set of antisymmetric (skew symmetric) matrices.

We will often encounter square matrices, Q, say, of size $n^2 \times n^2$, which will be partitioned into $n \times n$ square subblocks $Q_{ij} \in \mathbb{R}^{n\times n}$, i.e.,

https://doi.org/10.1017/S026646660707020X Published online by Cambridge University Press

$$Q = \begin{pmatrix} Q_{11} & \cdots & Q_{1n} \\ \vdots & & \vdots \\ Q_{n1} & \cdots & Q_{nn} \end{pmatrix}. \tag{1.2}$$

The elements of Q are indexed by $q_{ij,kl}$, where $q_{ij,kl}$ denotes the (k,l)th entry of the (i,j)th subblock Q_{ij}.

Let $\mathcal{S}_{n*n}$ be the set of all such block matrices Q that are symmetric and whose subblocks are symmetric too; i.e., $\mathcal{S}_{n*n} = \{Q \in \mathcal{S}_{n^2} | Q_{ij} = Q_{ij}' \text{ for all } 1 \leq i, j \leq n\}$. In addition, we need the set of $\mathcal{A}_{n*n} = \{\Delta = \Delta' | \Delta_{ij} = -\Delta_{ij}' \text{ for all } 1 \leq i, j \leq n\}$, i.e., the set of all symmetric block matrices whose subblocks are all antisymmetric.

Let $E_{ij} = u_i u_j' - u_j u_i'$, $i > j$, where $u_i = (0, \ldots, 0, 1, 0, \ldots, 0)'$ are the canonical unit vectors of $\mathbb{R}^n$. Then it is immediate that $\mathcal{A}_{n*n}$ is spanned by the matrices $(E_{ij} \otimes E_{kl})$, i.e.,

$$\mathcal{A}_{n*n} = \left\{ \Delta = \sum_{1 \leq i < j \leq n} \sum_{1 \leq k < l \leq n} \omega_{ij,kl} (E_{ij} \otimes E_{kl}) \,\middle|\, \omega_{ij,kl} \in \mathbb{R} \right\}.$$

To simplify notation let $\Delta_1, \ldots, \Delta_{\bar{n}}$, $\bar{n} = (n(n-1)/2)^2$, denote the matrices $(E_{ij} \otimes E_{kl})$ for $i > j$ and $k > l$. Note that $Q = Q'$ is an element of $\mathcal{S}_{n*n}$ if and only if (iff) trace $Q\Delta = 0$ for all $\Delta \in \mathcal{A}_{n*n}$. Furthermore, for $\Delta = \Delta'$, $\Delta \in \mathcal{A}_{n*n}$ holds iff $(\nu \otimes e)'\Delta(\nu \otimes e) = 0$ holds for all $\nu, e \in \mathbb{R}^n$. See the Appendix for more details.

2. THE MODEL CLASSES OF INTEREST

2.1. Basic Framework

Suppose we are given some n-dimensional observable, weakly stationary vector process (y_t). The process (y_t) can be decomposed into the sum of two components, the conditional mean μ_t and some noise component ϵ_t,

$$y_t = \mu_t + \epsilon_t, \tag{2.1}$$

where $\mu_t = \mathbb{E}(y_t | \mathcal{I}_{t-1})$ and $\mathcal{I}_{t-1} = \sigma(y_{t-1}, y_{t-2}, \ldots)$ is the sigma field generated by the past information of (y_t). We assume that the noise process (ϵ_t) is characterized by the relations

$$\epsilon_t = H_t^{1/2} z_t, \tag{2.2}$$

where H_t is some $\mathcal{I}_{t-1}$-measurable $(n \times n)$–dimensional, p.d., symmetric matrix and (z_t) is an n-dimensional vector process independently and identically distributed with mean zero and unit variance, $z_t \sim iid(0, I_n)$. Given all these assumptions, the variance matrix of y_t conditional on $\mathcal{I}_{t-1}$ equals the variance matrix of ϵ_t conditional on $\mathcal{I}_{t-1}$,

$$\text{var}(y_t | \mathcal{I}_{t-1}) = \text{var}(\epsilon_t | \mathcal{I}_{t-1}) = \mathbb{E}(\epsilon_t \epsilon_t' | \mathcal{I}_{t-1}) = H_t^{1/2} \mathbb{E}(z_t z_t' | \mathcal{I}_{t-1}) H_t^{T/2} = H_t.$$

https://doi.org/10.1017/S026646660707020X Published online by Cambridge University Press

In the following discussion, we concentrate on the modeling of the conditional variance H_t. The conditional mean μ_t may, e.g., be modeled by some autoregressive moving average system (ARMA).

For the modeling of the conditional variance matrices here we use VECH and BEKK models, which are well-known multivariate generalizations of the univariate ARCH (autoregressive conditional heteroskedastic) model proposed by Engle (1982) and the univariate GARCH model suggested by Bollerslev (1986).

2.2. The VECH Model

The VECH(p,q) model was proposed in Bollerslev et al. (1988) and is given by

$$\operatorname{vech}(H_t) = c + \sum_{i=1}^{q} A_i \operatorname{vech}(\epsilon_{t-i}\epsilon'_{t-i}) + \sum_{j=1}^{p} A_{q+j} \operatorname{vech}(H_{t-j}), \tag{2.3}$$

where c is an $(n(n+1)/2)$–dimensional vector and the A_i's are square matrices of order $n(n+1)/2$. This is a very general formulation of the conditional second moments of (ϵ_t), because each component of H_t can be formed by a linear combination of its own past and the past of all the other elements in H_t and $\epsilon_t\epsilon'_t$.

However, at least two drawbacks of this model are apparent. First, the high number of parameters to estimate, which increases rapidly with $O(n^4)$, implies that estimation of the model is infeasible for *large n*. Yet a close look at the theory of this very general model class is worthwhile, not only for understanding the model features but also for the construction or definition of reasonable parsimonious subclasses. Second, H_t has to be p.d. for all ϵ_{t-i} and all H_{t-j} in the sample space. This is ensured if we assume that the constant is strictly positive,

$$\text{math}(c) > 0, \tag{2.4}$$

and that the contribution of each ARCH and GARCH term resp. is nonnegative; i.e.,

$$\text{math}(A_i \operatorname{vech}(ee')) \geq 0 \quad \forall e \in \mathbb{R}^n \tag{2.5}$$

holds for $i = 1, \ldots, p+q$. Here we have used the fact that any p.s.d. matrix $H \geq 0$ may be factorized as $H = \sum_{i=1}^{n} e_i e'_i$ and thus the condition

$$\text{math}(A_i \operatorname{vech}(H)) \geq 0 \quad \forall H \in \mathbb{R}^{n\times n}, \qquad H \geq 0$$

is equivalent to (2.5). A VECH model that satisfies conditions (2.4) and (2.5) will be called *admissible*. However, note that admissibility is hard to check for given parameters $c, A_1, \ldots, A_{p+q}$, and thus hard to impose during estimation.

To further analyze the admissibility constraint let us concentrate on one of the preceding terms, $A \operatorname{vech}(ee')$, say. Any matrix $A \in \mathbb{R}^{n(n+1)/2\times n(n+1)/2}$ will be called *VECH term* in the following. A VECH term A is *admissible* iff

https://doi.org/10.1017/S026646660707020X Published online by Cambridge University Press

$$f(\nu, e) = \nu' \operatorname{math}(A \operatorname{vech}(ee'))\nu \geq 0 \quad \forall \nu, e \in \mathbb{R}^n. \tag{2.6}$$

Thus A is admissible iff the optimization problem

$$\begin{cases} \text{minimize} & \nu' \operatorname{math}(A \operatorname{vech}(ee'))\nu \\ \text{subject to} & \|\nu\| = \|e\| = 1 \end{cases} \tag{2.7}$$

has an optimal value, $\gamma^*_{\mathcal{V}}(A)$, say, that is nonnegative. Note that (2.7) is a nonconvex optimization problem that in general is hard to solve.

The function $f(\nu, e)$ has the following interpretation: Suppose that A corresponds to the lag one ARCH term; then $f(\nu, e)$ is the contribution of this term to the conditional variance of the linear combination $(\nu'\epsilon_t)$ for a given $\epsilon_{t-1} = e$. For fixed ν, $f(\nu, e)$ is a quadratic form in e and for fixed e, $f(\nu, e)$ is a quadratic form in ν.

By rearrangement of terms, $f(\cdot, \cdot)$ may also be written as

$$f(\nu, e) = \nu' \operatorname{math}(A \operatorname{vech}(ee'))\nu = (\nu \otimes e)' Q (\nu \otimes e), \tag{2.8}$$

where the symmetric matrix $Q \in \mathbb{R}^{n^2 \times n^2}$ is a solution of the following relations:

$$\begin{aligned} q_{ij,kl} &= q_{ji,lk} \quad \forall i, j, k, l, \\ q_{ii,kk} &= a_{I(i,i),I(k,k)} \quad \forall i, k, \\ 2q_{ii,kl} &= a_{I(i,i),I(k,l)} \quad \forall i, k > l, \\ q_{ij,kk} &= a_{I(i,j),I(k,k)} \quad \forall i > j, k, \\ q_{ij,kl} + q_{ij,lk} &= a_{I(i,j),I(k,l)} \quad \forall i > j, k > l. \end{aligned} \tag{2.9}$$

Here we have partitioned $Q \in \mathbb{R}^{n^2 \times n^2}$ into $n \times n$ subblocks of size $(n \times n)$ each and used the indexing as described in (1.1) and (1.2). Note that the first set of equations relates to the symmetry of Q and that because of the last set of equations the matrix Q is not uniquely determined for a given matrix A. It follows immediately that the set of Q matrices that satisfy (2.9) for given matrix A is given by

$$\mathcal{Q}(A) = \left\{ Q(\omega) = Q_0 + \sum_{k=1}^{\bar{n}} \omega_k \Delta_k \,\middle|\, \omega_k \in \mathbb{R} \right\} = \{Q_0 + \Delta \,|\, \Delta \in \mathcal{A}_{n*n}\}, \tag{2.10}$$

where Q_0 is an arbitrary matrix for which, given this matrix A, equations (2.9) hold; see also Section 1.1. Note that $\mathcal{Q}(A)$ is an affine subset of $\mathcal{S}_{n^2}$ of dimension $(n(n-1)/2)^2$. These sets will be called $\mathcal{Q}$-sets in what follows.

One can of course always select a unique matrix $\underset{\sim}{Q}$ out of $\mathcal{Q}(A)$, if further normalizing conditions are imposed. For instance, $\underset{\sim}{Q} \in \mathcal{S}_{n*n} \cap \mathcal{Q}(A)$ is the only element in $\mathcal{Q}(A)$ with symmetric subblocks:

https://doi.org/10.1017/S026646660707020X Published online by Cambridge University Press

$$\check{q}_{ij,kl} = \check{q}_{ij,lk} = \frac{1}{2} a_{I(i,j),I(k,l)} \quad \forall i > j, k > l. \tag{2.11}$$

However, as we will see later on, it is advantageous to consider other normalizing conditions.

Note that for a given symmetric $Q \in \mathbb{R}^{n^2 \times n^2}$ the relations (2.9) uniquely define a VECH term $A \in \mathbb{R}^{n(n+1)/2 \times n(n+1)/2}$. As an example the matrices $Q(\omega_1)$, $\omega_1 \in \mathbb{R}$, for the simplest case, $n = 2$, are given here:

$$Q(\omega_1) = \underbrace{\left(\begin{array}{cc|cc} a_{11} & a_{12}/2 & a_{21} & a_{22}/2 \\ a_{12}/2 & a_{13} & a_{22}/2 & a_{23} \\ \hline a_{21} & a_{22}/2 & a_{31} & a_{32}/2 \\ a_{22}/2 & a_{23} & a_{32}/2 & a_{33} \end{array}\right)}_{=\check{Q}} + \omega_1 \underbrace{\left(\begin{array}{cc|cc} 0 & 0 & 0 & 1 \\ 0 & 0 & -1 & 0 \\ \hline 0 & -1 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{array}\right)}_{=\Delta_1}. \tag{2.12}$$

In this paper we will mainly deal with the admissibility constraints, i.e., conditions that ensure the positive definiteness of the conditional variance matrices. For conditions that ensure stationarity of the process (ϵ_t) we refer, e.g., to Engle and Kroner (1995, Prop. 2.7): The VECH model has a (unique) weakly stationary solution (ϵ_t) if and only if

$$\rho\left(\sum_{i=1}^{p+q} A_i\right) < 1 \tag{2.13}$$

holds. This is a highly nonlinear restriction, which is hard to impose on a parametrization. In particular it imposes a "cross restriction," which jointly refers to all terms A_i. In special cases however, the stationarity condition simplifies significantly; see the diagonal VECH (DVECH) model and the F-GARCH model later on.

2.3. The BEKK Model

Baba et al. (1991) proposed the BEKK(p,q,K) model,

$$H_t = CC' + \sum_{i=1}^{q} \sum_{k=1}^{K} B'_{i,k} \epsilon_{t-i} \epsilon'_{t-i} B_{i,k} + \sum_{j=1}^{p} \sum_{k=1}^{K} B'_{q+j,k} H_{t-j} B_{q+j,k}, \tag{2.14}$$

where C is a nonsingular $n \times n$ matrix, $B_{i,k}$ are $n \times n$ parameter matrices, and K in a certain sense determines the generality of the process.

Of course admissibility is not an issue here, because the preceding recursions always give p.d. matrices $H_t > 0$ by construction. For completeness we

https://doi.org/10.1017/S026646660707020X Published online by Cambridge University Press

state the corresponding stationarity condition; see Engle and Kroner (1995): The BEKK model has a weakly stationary solution (ϵ_t) if and only if

$$\rho\left(\sum_{i=1}^{p+q}\sum_{k=1}^{K} B_{i,k}\otimes B_{i,k}\right) < 1 \tag{2.15}$$

holds.

Because p.s.d. matrices $H \geq 0$ may be factorized as $H = \sum e_i e_i'$, it follows that GARCH terms, $\sum_{k=1}^{K} B_{q+j,k}' H_{t-j} B_{q+j,k}$, may again be treated completely analogously to ARCH terms, $\sum_{k=1}^{K} B_{i,k}' \epsilon_{t-i}\epsilon_{t-i}' B_{i,k}$. Now consider one of these terms, $\sum B_k' ee' B_k$, say, and note that

$$\nu'\left(\sum_{k=1}^{K} B_k' ee' B_k\right)\nu = \sum_{k=1}^{K} (\nu' B_k' e)(\nu' B_k' e)'$$

$$= (\nu\otimes e)' \underbrace{\sum_{k=1}^{K} \operatorname{vec}(B_k)\operatorname{vec}(B_k)'}_{=Q} (\nu\otimes e). \tag{2.16}$$

This implies that any BEKK term has an equivalent representation as a VECH term, A, say, where A is computed from $Q = \sum_{k=1}^{K} \operatorname{vec}(B_k)\operatorname{vec}(B_k)' \geq 0$ and (2.9). Furthermore this VECH term is admissible, and the corresponding $\mathcal{Q}$-set $\mathcal{Q}(A)$ contains at least one p.s.d. element. On the other hand, if the $\mathcal{Q}$-set corresponding to a VECH term, A, say, contains a p.s.d. element, $Q \geq 0$, say, then this A is admissible and has an equivalent BEKK representation. To see this, just note that $Q \geq 0$ implies that Q may be factorized as $Q = \sum_{k=1}^{K} U_k U_k'$, where $U_k \in \mathbb{R}^{n^2}$ and $K \leq n^2$ is the rank of Q, and thus $B_k = \operatorname{mat}(U_k)$ is the desired BEKK representation.

As has been noted already it is hard to check whether a given VECH term, A, say, is admissible. However, it is easy to check whether A admits a BEKK representation.

To this end, let some VECH term A be given and consider the following semidefinite program (s.d.p.) for $\lambda \in \mathbb{R}$ and $\omega = (\omega_1, \ldots, \omega_{\bar{n}})' \in \mathbb{R}^{\bar{n}}$:

$$\begin{cases} \text{maximize} & \lambda \\ \text{subject to} & \left(Q_0 + \sum_{k=1}^{\bar{n}} \omega_k \Delta_k\right) - \lambda I_{n^2} \geq 0, \end{cases} \tag{2.17}$$

where Q_0 is an arbitrary element of $\mathcal{Q}(A)$.

One can easily prove that this s.d.p. always has an optimizer, (λ^*, ω^*), say. Furthermore, note that the optimal λ^* is the minimal eigenvalue of $Q(\omega^*) = Q_0 + \sum \omega_k^* \Delta_k$. In other words, the preceding s.d.p. maximizes $\lambda_{\min}(Q)$ over the

https://doi.org/10.1017/S026646660707020X Published online by Cambridge University Press

set $\mathcal{Q}(A)$, where $\lambda_{\min}(Q)$ denotes the minimal eigenvalue of Q. Thus, $\mathcal{Q}(A)$ contains a p.s.d. matrix Q iff the optimal λ is nonnegative. For a detailed description of s.d.p.'s and their properties we refer to Vandenberghe and Boyd (1996). Note that (2.17) is a convex relaxation of problem (2.7).

It is well known that any BEKK model may be cast as a VECH model. However to the best of our knowledge the previously given check whether a VECH model has an equivalent BEKK representation is new. By the preceding considerations one may embed the set of BEKK models/terms into the set of VECH models/terms. Thus we will, for simplicity of notation, from now on call a VECH term that allows for a BEKK representation simply a *BEKK term*.

The theorem that follows summarizes some basic facts on VECH and BEKK models. For simplicity of notation we restrict ourselves to single terms; however the generalization to complete models is straightforward. We define the following sets.

Let $\mathcal{V} \subseteq \mathbb{R}^{n(n+1)/2 \times n(n+1)/2}$ denote the set of all admissible VECH terms A. Note that $A \in \mathcal{V}$ iff $\gamma^*_{\mathcal{V}}(A) \geq 0$ holds; see (2.7). Let $\mathcal{B} \subseteq \mathcal{V}$ denote the set of all VECH terms A that have a BEKK representation; i.e., using the previous convention $\mathcal{B}$ is the set of all BEKK terms. Let $\lambda^*_{\mathcal{B}}(A) = \lambda^*$ denote the optimum value attained for problem (2.17) and note that $A \in \mathcal{B}$ iff $\mathcal{Q}(A)$ contains a p.s.d. element $Q \geq 0$, i.e., iff $\lambda^*_{\mathcal{B}}(A) \geq 0$ holds. Finally, let $\mathcal{B}^+ \subset \mathcal{B} \subseteq \mathcal{V}$ denote the set of all BEKK terms A where $\mathcal{Q}(A)$ contains a p.d. element $Q > 0$, i.e., $\lambda^*_{\mathcal{B}}(A) > 0$.

THEOREM 2.1.

(i) $\mathcal{V}$ *contains an open subset of* $\mathbb{R}^{n(n+1)/2 \times n(n+1)/2}$.
(ii) $\mathcal{B}$ *contains an open subset of* $\mathbb{R}^{n(n+1)/2 \times n(n+1)/2}$.
(iii) *The set* $\mathcal{B}^+ \subset \mathcal{B}$ *is open in* $\mathbb{R}^{n(n+1)/2 \times n(n+1)/2}$ *and dense in* $\mathcal{B}$.
(iv) $\mathcal{B} = \mathcal{V}$ *holds for* $n = 2$, *whereas for* $n > 2$ *the set* $(\mathcal{V} \backslash \mathcal{B}) = \{A \in \mathcal{V} | A \notin \mathcal{B}\}$ *contains an open subset of* $\mathbb{R}^{n(n+1)/2 \times n(n+1)/2}$.

Theorem 2.1(i) makes precise that a parametrization of the set of admissible VECH terms $\mathcal{V}$—and of any general subset—needs $(n(n+1)/2)^2 = O(n^4)$ free parameters. Here "general" means that this subclass contains an open set of VECH terms.

Theorem 2.1(ii) implies that this applies in particular to the set of all BEKK terms $\mathcal{B}$. In the next section we will consider the subclass $\mathcal{B}^+ \subseteq \mathcal{B}$ of BEKK terms and give a parametrization for this class. From Theorem 2.1(iii) it follows that this subset covers "almost all" BEKK terms. Finally, the preceding theorem shows that in the bivariate case, $n = 2$, BEKK models are as general as (admissible) VECH models. Thus in this case there is no loss of generality if one considers BEKK models only. However for $n > 2$ there is a "thick" set of admissible VECH models that have no equivalent BEKK representation.

https://doi.org/10.1017/S026646660707020X Published online by Cambridge University Press

3. PARAMETRIZATION OF BEKK MODELS

By the preceding considerations it is clear that the BEKK parameter matrices $(B_1, \ldots, B_K)$ are by no means identified; even the number K is not unique because the set $\mathcal{Q}(A) \cap \mathcal{S}^+_{n^2}$, in general, contains elements Q of different rank. One way to get unique parameter matrices $(B_1, \ldots, B_K)$ corresponding to a certain $\mathcal{Q}$-set is to choose a unique p.s.d. representative element $\bar{Q} \in \mathcal{Q}(A) \cap \mathcal{S}^+_{n^2}$ and then use a uniquely defined factorization of this $\bar{Q} \geq 0$.

3.1. A Parametrization for $\mathcal{B}^+$

Here we apply the preceding scheme to derive a parametrization to the generic set $\mathcal{B}^+$. By definition the $\mathcal{Q}$-sets for all $A \in \mathcal{B}^+$ contain at least one p.d. element $Q > 0$. We then choose a representative $\bar{Q} > 0$ that is, in a certain sense to be defined subsequently, the center of all positive elements in $\mathcal{Q}(A)$. In particular this implies that $K = n^2$.

Let a $\mathcal{Q}$-set $\mathcal{Q}(A) = \{Q(\omega) = Q_0 + \sum_{k=1}^{\bar{n}} \omega_k \Delta_k\}$ be given where $\mathcal{Q}^+(A) = \mathcal{Q}(A) \cap \mathcal{S}^+_{n^2,n^2} = \{Q \in \mathcal{Q}(A) | Q > 0\}$ is nonvoid. We then define a function $\phi(\cdot)$ on $\mathbb{R}^{\bar{n}}$ by

$$\phi(\omega) = \begin{cases} \log\det Q(\omega)^{-1} & \text{if } Q(\omega) > 0 \\ +\infty & \text{otherwise.} \end{cases} \tag{3.1}$$

Note that the set $\mathcal{W} = \{\omega | Q(\omega) \geq 0\}$ is a compact, convex subset of $\mathbb{R}^{\bar{n}}$ and that, by assumption, the interior of this set, i.e., the set $\mathcal{W}^+ = \{\omega | Q(\omega) > 0\}$, is nonvoid. The function $\phi(\cdot)$, when restricted to $\mathcal{W}^+$, is convex and analytic and thus has a unique minimizer, $\bar{\omega}$, say. The corresponding $\bar{Q} = Q(\bar{\omega}) = Q_0 + \sum_k \bar{\omega}_k \Delta_k$ is called the *analytic center* of $\mathcal{Q}^+(A)$; see Vandenberghe and Boyd (1996). The gradient, $\nabla\phi(\omega)$, of ϕ for $\omega \in \mathcal{W}^+$ is given by

$$(\nabla\, \phi(\omega))_k = -\text{trace}(Q(\omega)^{-1}\Delta_k). \tag{3.2}$$

Hence, in the analytic center, $\bar{Q} = Q(\bar{\omega})$, we have $\text{trace}(\bar{Q}^{-1}\Delta_k) = 0$ for all $k = 1, \ldots, \bar{n}$. This is guaranteed if and only if each subblock matrix of $\bar{Q}^{-1}$ is symmetric; see Section 1.1.

Next, let $\bar{Q}^{-1} = VV'$ be the Cholesky decomposition of $\bar{Q}^{-1}$; i.e., $V \in \mathbb{R}^{n^2 \times n^2}$ is a lower triangular matrix where without loss of generality all diagonal elements are assumed to be positive. The parameter vector θ is defined as

$$\theta' = (\text{diag}(V_{11})', \ldots, \text{diag}(V_{nn})', \underline{\text{vech}}(V_{11})', \ldots, \underline{\text{vech}}(V_{nn})',$$
$$\text{vech}(V_{21})', \ldots, \text{vech}(V_{n1}), \text{vech}(V_{32}), \ldots, \text{vech}(V_{n(n-1)})'); \tag{3.3}$$

https://doi.org/10.1017/S026646660707020X Published online by Cambridge University Press

i.e., we stack the lower triangular entries of all lower triangular blocks of V. Note that we have put all the diagonal elements of V to the first n^2 positions and thus the positivity constraint implies that

$$\theta \in \Theta^+ := \mathbb{R}_+^{n^2} \times \mathbb{R}^{(n(n+1)/2)^2 - n^2} \subseteq \mathbb{R}^{(n(n+1)/2)^2}.$$

Because of the block symmetric structure of $\bar{Q}^{-1}$ these entries are sufficient to reconstruct $\bar{Q}^{-1}$.

Let us recapitulate the previous procedure. Suppose we have given a VECH term, $A \in \mathcal{B}^+$, say. Then a unique set of BEKK parameter matrices $(B_1, \ldots, B_{n^2})$ corresponding to this VECH term may be constructed as follows. First, compute the analytic center $\bar{Q}$ of $\mathcal{Q}^+(A)$ as described before. Next, compute a Cholesky factor V of $\bar{Q}^{-1}$ with strictly positive diagonal elements. A parameter vector θ is given by (3.3), and the BEKK parameter matrices are given by $B_k = \mathrm{mat}(U_k)$, where U_k is the kth column of $U = V^{-T}$.

For the reverse direction, suppose we have given a parameter vector $\theta \in \Theta^+$; then we have to construct the BEKK matrices $B_1, \ldots, B_{n^2}$ and the corresponding VECH parameter matrix A. First, note that the matrix V is a lower triangular matrix, i.e., $v_{ij,kl} = 0$ for $i < j$ and for $i = j$ and $k < l$. The lower triangular entries $v_{ij,kl}$, $i \geq j$, and $k \geq l$ are stored in the parameter vector θ; see (3.3). Thus, it remains to construct the upper triangular elements $v_{ij,kl}$, $k < l$ for $i > j$ from the condition that the product $P = \bar{Q}^{-1} = VV'$ has symmetric subblocks $P_{ij} = P_{ij}'$. For $i \geq j$ such a subblock is given by $P_{ij} = V_{i1}V_{j1}' + \cdots + V_{ij}V_{jj}'$. Because V_{jj} is a lower triangular matrix the (k,l)th entry of subblock P_{ij} is of the form

$$p_{ij,kl} = \underbrace{\sum_{r=1}^{j-1} \sum_{s=1}^{n} v_{ir,ks} v_{jr,ls} + \sum_{s=1}^{l-1} v_{ij,ks} v_{jj,ls}}_{:=\bar{p}_{ij,kl}} + v_{ij,kl} v_{jj,ll}.$$

From the symmetry condition $p_{ij,kl} = p_{ij,lk}$ and by the constraint $v_{jj,ll} > 0$ we get

$$v_{ij,kl} = \frac{p_{ij,lk} - \bar{p}_{ij,kl}}{v_{jj,ll}}. \tag{3.4}$$

For $i > j$ and $k < l$ this expression only depends on V_{jj}, V_{ir}, and V_{jr} with $r < j$, and on $v_{ij,kr}$, $v_{ij,lr}$ with $r < l$. Thus, V may be recursively reconstructed from θ. The lower diagonal blocks of V are reconstructed in the sequence $V_{21}, \ldots, V_{n1}, V_{32}, \ldots, V_{n,n-1}$, and in each block V_{ij} the upper diagonal elements are computed in the sequence $v_{ij,12}, \ldots, v_{ij,1n}, v_{ij,23}, \ldots, v_{ij,(n-1),n}$. Finally, we get $\bar{Q} = UU'$ from $U = V^{-T}$, the BEKK parameter matrices from $B_k = \mathrm{mat}(U_k)$, and the VECH matrix A from the relations (2.9).

https://doi.org/10.1017/S026646660707020X Published online by Cambridge University Press

As an illustration the structure of V for the simplest case $n = 2$ is given here:

$$V = \left(\begin{array}{cc|cc} \theta_1 & 0 & 0 & 0 \\ \theta_5 & \theta_2 & 0 & 0 \\ \hline \theta_7 & \dfrac{\theta_1\theta_8 - \theta_5\theta_7}{\theta_2} & \theta_3 & 0 \\ \theta_8 & \theta_9 & \theta_6 & \theta_4 \end{array}\right).$$

THEOREM 3.1. *The previously defined mapping*

$$\pi^+ : \Theta^+ \to \mathcal{B}^+$$
$$\theta \mapsto A$$

is a diffeomorphism, and the previously defined mapping

$$\pi_B^+ : \Theta^+ \to \mathbb{R}^{n^4}$$
$$\theta \mapsto \mathrm{vec}(B_1, \ldots, B_{n^2})$$

is differentiable and injective.

We want to emphasize that this parametrization covers a generic set of BEKK models; i.e., the corresponding set is an open and dense subset of all BEKK models. We here considered only the case of one term, i.e., the case of a BEKK$(0,1,K)$ model, but the generalization to the case of BEKK(p,q,K) models is straightforward.

3.2. BEKK Parametrization of Engle and Kroner

In Proposition 2.3 of Engle and Kroner (1995) a parametrization for BEKK models is presented. The authors claim that this parametrization is *fully general* and *identifiable*. Here "fully general" means that the parametrization covers as many VECH models as possible and "identifiable" means that two different parameter vectors do not represent the same BEKK model. However, both statements are not correct, as we will show in this section.

In their representation $K = n^2$ and the parameter matrices B_k are chosen such that the corresponding matrix $U = [\mathrm{vec}(B_1), \ldots, \mathrm{vec}(B_K)]$ is a lower block triangular matrix where all $(n \times n)$ subblocks are again lower triangular; i.e., U is a Cholesky factor of $Q = UU'$ with the additional zero restrictions given by $u_{ij,kl} = 0$ for $i > j$ and $k < l$. By these additional zero restrictions implicitly a unique representative, $\hat{Q}$, say, in the $\mathcal{Q}$-set $\mathcal{Q}(A)$ is chosen. In addition, the entries of the last row of U are assumed to be positive, $u_{nj,nl} > 0$, for all j, l to get rid of the nonuniqueness of the Cholesky factorization of $\hat{Q}$.

https://doi.org/10.1017/S026646660707020X Published online by Cambridge University Press

Using the relations between VECH models, BEKK models, and $\mathcal{Q}$-sets, respectively, it is easy to see that two problems occur with this parametrization:

1. The positivity constraints do not guarantee uniqueness in all cases.
2. There is a "thick" set of BEKK models that is not covered by this parametrization, because the corresponding $\mathcal{Q}$-sets do not contain a p.s.d. representative $\hat{Q}$ of this structure.

To see this consider the case $n = 2$. The Engle–Kroner parametrization leads to a U-matrix of the form

$$U = \left(\begin{array}{cc|cc} \theta_1 & 0 & 0 & 0 \\ \theta_2 & \theta_3 & 0 & 0 \\ \hline \theta_4 & 0 & \theta_7 & 0 \\ \theta_5 & \theta_6 & \theta_8 & \theta_9 \end{array}\right),$$

where $\theta_5 > 0$, $\theta_6 > 0$, $\theta_8 > 0$, $\theta_9 > 0$ and thus $\hat{Q}$ is given by

$$\hat{Q} = UU' = \left(\begin{array}{cc|cc} \theta_1^2 & \theta_1\theta_2 & \theta_1\theta_4 & \theta_1\theta_5 \\ \theta_1\theta_2 & \theta_2^2+\theta_3^2 & \theta_2\theta_4 & \theta_2\theta_5+\theta_3\theta_6 \\ \hline \theta_1\theta_4 & \theta_2\theta_4 & \theta_4^2+\theta_7^2 & \theta_4\theta_5+\theta_7\theta_8 \\ \theta_1\theta_5 & \theta_2\theta_5+\theta_3\theta_6 & \theta_4\theta_5+\theta_7\theta_8 & \theta_5^2+\theta_6^2+\theta_8^2+\theta_9^2 \end{array}\right).$$

Now, consider a BEKK term—in Engle–Kroner form—given by

$$U = \left(\begin{array}{cc|cc} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ \hline 1 & 0 & 0 & 0 \\ 1 & 1 & \theta_8 & \theta_9 \end{array}\right) \quad \text{and} \quad \hat{Q} = UU' = \left(\begin{array}{cc|cc} 1 & 0 & 1 & 1 \\ 0 & 1 & 0 & 1 \\ \hline 1 & 0 & 1 & 1 \\ 1 & 1 & 1 & 2+\theta_8^2+\theta_9^2 \end{array}\right).$$

In this case we lose identifiability, because $\theta_7 = 0$, and thus only the sum $\theta_8^2 + \theta_9^2$ is identified from Q. (In other words, the positivity constraint, in this case, is not sufficient to select a unique Cholesky factor.) However, this problem may be circumvented by replacing the restrictions $u_{nj,nl} > 0$ by the restriction that the diagonal elements of U are positive, i.e., $u_{jj,ll} > 0$ for all j, l.

Next, consider a BEKK term of the form

$$U = \left(\begin{array}{cc|cc} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ \hline 0 & x & 1 & 0 \\ 1 & 1 & 1 & 1 \end{array}\right) \quad \text{and} \quad Q = UU' = \left(\begin{array}{cc|cc} 1 & 0 & 0 & 1 \\ 0 & 1 & x & 1 \\ \hline 0 & x & 1+x^2 & 1+x \\ 1 & 1 & 1+x & 4 \end{array}\right).$$

https://doi.org/10.1017/S026646660707020X Published online by Cambridge University Press

The Engle–Kroner parameters θ_k and thus $\hat{Q} \in \mathcal{Q}$ may be computed from Q as follows:

$$\theta_1 = \sqrt{q_{11,11}} = 1,$$

$$\theta_2 = q_{11,21}/\theta_1 = 0,$$

$$\theta_3 = \sqrt{q_{11,22} - \theta_2^2} = 1,$$

$$\theta_4 = q_{21,11}/\theta_1 = 0,$$

$$\theta_5 = (q_{21,21} + q_{21,12} - \theta_2\theta_4)/\theta_1 = 1 + x,$$

$$\theta_6 = (q_{21,22} - \theta_2\theta_5)/\theta_3 = 1,$$

$$\theta_7 = \sqrt{q_{22,11} - \theta_4^2} = \sqrt{1 + x^2},$$

$$\theta_8 = (q_{22,21} - \theta_4\theta_5)/\theta_7 = (1 + x)/\sqrt{1 + x^2},$$

$$\theta_9 = \sqrt{q_{22,22} - \theta_5^2 - \theta_6^2 - \theta_8^2} = \sqrt{4 - (1 + x)^2 - 1 - (1 + x)^2/(1 + x^2)}.$$

Note that θ_5 is computed from the sum $q_{21,21} + q_{21,12}$ because only the sum is uniquely defined. However, this procedure breaks down for "large" x, because $4 - (1 + x)^2 - 1 - (1 + x)^2/(1 + x^2) < 0$ for sufficiently large x. In this case the BEKK model defined as before has no representation in the form suggested by Engle and Kroner because the set $\mathcal{Q}(A)$ does not contain a p.s.d. element $\hat{Q} \geq 0$ conforming to the restrictions imposed by the Engle–Kroner parametrization. It is clear that this problem occurs on a "thick" set of BEKK terms, i.e., on a set that contains an open subset of all BEKK terms. In this sense, using the Engle–Kroner parametrization we lose a nonnegligible set of BEKK models.

The preceding examples were given for the case $n = 2$; however, the reasoning may be easily extended to the case $n > 2$.

4. PARAMETRIZATION OF BEKK MODELS WITH $K \leq n$

The main advantage of BEKK models over VECH models is that BEKK models are admissible by construction. However, the number of free parameters to estimate for a general BEKK$(p,q,K = n^2)$ model is equal to the number of parameters for a VECH(p,q) model, which is of the order $O(n^4)$. For small vector dimensions, $n \leq 3$, estimation of a BEKK(p,q,n^2) model might still be reasonable. In the case of higher dimensional vector processes however, it becomes inevitable to impose further structure and restrictions on the parameters to make estimation of multivariate GARCH models applicable in practice. There have been many suggestions to further restrict this model class. Here we will first discuss two widely used model classes, the DVECH model and

https://doi.org/10.1017/S026646660707020X Published online by Cambridge University Press

the F-GARCH model. Thereafter a generalization of the F-GARCH model is proposed.

4.1. The DVECH Model

The DVECH model has been proposed by Bollerslev et al. (1988). This model class is very popular in practice because of its parsimony and the fact that the admissibility and stationarity conditions are very simple. Here we give a short review of this model using the previously derived methodology.

A VECH model is called a DVECH model if all parameter matrices A_i in (2.3) are *diagonal* matrices. Thus, the elements of H_t depend only on their own past and the respective elements of $\epsilon_{t-i}\epsilon_{t-i}'$. The number of parameters reduces to $(p+q+1)n(n+1)/2 = O(n^2)$.

In the following discussion we will again concentrate on one (diagonal) VECH term, A, say. Because $A \in \mathbb{R}^{n(n+1)/2 \times n(n+1)/2}$ is diagonal we may write $A = \operatorname{diag}(\operatorname{vech}(\Omega))$ for some matrix $\Omega \in \mathcal{S}_n$. The term A is admissible if $f(\nu, e) = \nu' \operatorname{math}(A \operatorname{vech}(ee'))\nu \geq 0$ holds for all vectors $\nu, e \in \mathbb{R}^n$. By some easy algebra it follows that

$$f(\nu,e) = \nu' \operatorname{math}(A \operatorname{vech}(ee'))\nu = \nu'(\Omega \odot (ee'))\nu = (\nu \odot e)'\Omega(\nu \odot e), \quad \textbf{(4.1)}$$

$$= (\nu \otimes e)'\underbrace{(S'\Omega S)}_{\bar{Q}}(\nu \otimes e), \quad \textbf{(4.2)}$$

where $S \in \mathbb{R}^{n \times n^2}$ is a selection matrix that satisfies $S(\nu \otimes e) = (\nu \odot e)$ for all $\nu, e \in \mathbb{R}^n$. From (4.1) it follows that $A = \operatorname{diag}(\operatorname{vech}(\Omega))$ is admissible if and only if $\Omega \geq 0$ holds. Furthermore (4.2) implies $\bar{Q} \geq 0$, and thus any admissible DVECH term A has a BEKK representation. Note that $\bar{Q}$ is the only p.s.d. element in the corresponding $\mathcal{Q}$-set. To see this just note that the diagonal blocks of $\bar{Q}$ only have one nonzero diagonal element: $\bar{q}_{ii,kk} = 0$ for all $k \neq i$. Thus $Q = \bar{Q} + \Delta \geq 0$, $\Delta \in \mathcal{A}_{n*n}$ implies $\Delta = 0$.

To construct the BEKK parameter matrices consider a factorization of $\Omega \geq 0$ as $\Omega = \sum_{k=1}^{K} L_k L_k'$ where L_k are vectors in $\mathbb{R}^n$ and $K = \operatorname{rank}\Omega \leq n$. This gives a factorization of $\bar{Q}$, and thus $B_k = \operatorname{mat}(SL_k) = \operatorname{diag}(L_k)$. Therefore the corresponding BEKK matrices are diagonal too, and we obtain a *diagonal BEKK model*. Note that the number of components K is in any case less than or equal to n.

To get a unique BEKK representation for A it suffices to define a unique factorization of Ω. For example, a Cholesky decomposition leads to diagonal matrices B_k, for which the kth diagonal element is positive and the first $k-1$ diagonal elements are zero. (Strictly speaking this holds true for "generic" Ω, which have a regular $K \times K$ left upper block.)

It is obvious that a diagonal BEKK model gives a diagonal (and admissible) VECH model. Therefore there is a one to one correspondence of admissible DVECH models and diagonal BEKK models.

https://doi.org/10.1017/S026646660707020X Published online by Cambridge University Press

The stationarity conditions in this case are especially simple. Because of the diagonal structure (2.13) simplifies to

$$\sum_{i=1}^{p+q} a_{i,jj} < 1; \quad \text{for } 1 \le j \le n(n+1)/2,$$

where $a_{i,jj}$ is the jth diagonal element of the DVECH term A_i.

4.2. The Factor GARCH Model

This model class has been proposed in Engle, Ng, and Rothschild (1990). An F-GARCH model is a BEKK(p,q,K) model (see equation (2.14)), where the following restrictions are imposed:

1. $K \le n$.
2. The BEKK matrices $B_{i,k}$ are rank one matrices of the form

$$B_{i,k} = \mu_{i,k}(\Gamma_k \Lambda_k'), \qquad \Gamma_k, \Lambda_k \in \mathbb{R}^n, \qquad \|\Gamma_k\| = 1, \qquad \mu_{i,k} > 0. \tag{4.3}$$

3. Let $\Gamma = (\Gamma_1, \ldots, \Gamma_K) \in \mathbb{R}^{n \times K}$ and $\Lambda = (\Lambda_1, \ldots, \Lambda_K) \in \mathbb{R}^{n \times K}$. Then

$$\Gamma'\Lambda = \Lambda'\Gamma = I_K. \tag{4.4}$$

Let us first consider the case $0 < K < n$. By (4.4) there exist two matrices $\bar{\Gamma}$, $\bar{\Lambda} \in \mathbb{R}^{n \times (n-K)}$ such that

$$(\Gamma, \bar{\Gamma})'(\Lambda, \bar{\Lambda}) = (\Lambda, \bar{\Lambda})(\Gamma, \bar{\Gamma})' = I_n.$$

Therefore (ϵ_t) has a factor model representation of the form

$$\epsilon_t = \sum_{k=1}^{K} \Lambda_k f_{k,t} + v_t,$$

where $(f_{k,t} = \Gamma_k' \epsilon_t)$ are the factors and where $v_t = \bar{\Lambda}\bar{\Gamma}'\epsilon_t$ is the "noise." The following relations give the conditional variances and covariances, respectively:

$$\begin{aligned}
\operatorname{var}(f_{k,t}|\mathcal{I}_{t-1}) &= \Gamma_k' H_t \Gamma_k = \Gamma_k' CC' \Gamma_k \\
&\quad + \sum_{i=1}^{q} \mu_{i,k}^2 f_{k,t}^2 + \sum_{j=1}^{p} \mu_{q+j,k}^2 \operatorname{var}(f_{k,t-j}|\mathcal{I}_{t-1-j}), \\
\operatorname{cov}(f_{k,t}, f_{l,t}|\mathcal{I}_{t-1}) &= \Gamma_k' H_t \Gamma_l = \Gamma_k' CC' \Gamma_l = \operatorname{cov}(f_{k,t}, f_{l,t}) \quad \text{for } k \ne l, \\
\operatorname{var}(v_t|\mathcal{I}_{t-1}) &= \bar{\Lambda}\bar{\Gamma}' H_t \bar{\Gamma}\bar{\Lambda}' = \bar{\Lambda}\bar{\Gamma}' CC' \bar{\Gamma}\bar{\Lambda}' = \operatorname{var}(v_t), \\
\operatorname{cov}(v_t, f_{k,t}|\mathcal{I}_{t-1}) &= \bar{\Lambda}\bar{\Gamma}' H_t \Gamma_k = \bar{\Lambda}\bar{\Gamma}' CC' \Gamma_k = \operatorname{cov}(v_t, f_{k,t}).
\end{aligned}$$

Thus, the factors $(f_{k,t})$ are univariate GARCH processes. The conditional variance matrix of the noise (v_t) and the conditional covariance matrices are con-

https://doi.org/10.1017/S026646660707020X Published online by Cambridge University Press

stant. For $K = n$ there is no noise term (v_t), and hence ϵ_t is a linear combination of scalar GARCH processes only.

Note that the preceding model imposes strong cross-restrictions on the parameters for the respective ARCH and GARCH terms. The number of parameters reduces to $n(n+1)/2 + (p+q)K + 2nK - K - K^2 = O(n^2)$.

The stationarity condition (2.15) simplifies to the condition that the K univariate factor processes $(f_{k,t})$ are stationary:

$$\sum_{i=1}^{p+q} \mu_{i,k}^2 < 1; \quad \text{for } k = 1, \ldots, K.$$

4.3. A Generalization of F-GARCH Models

One of the main difficulties in defining a suitable parametrization of BEKK models is to define a rule for how to pick a unique p.s.d. element $Q \in \mathcal{Q}(A)$, $Q \geq 0$. A natural candidate would be $\breve{Q} \in \mathcal{Q}(A)$, which is the only element with symmetric subblocks; see (2.11). However, this element does not have to be p.s.d.; see, e.g., the DVECH model. Therefore here we simply *assume* that $\breve{Q} \geq 0$ holds; i.e., we only consider the corresponding subset of BEKK terms. In addition we demand that the rank of $\breve{Q}$ is equal to K, for some $K \leq n$, to reduce the number of parameters needed.

Let $\mathcal{B}_K^s \subset \mathcal{B}$ denote the set of BEKK terms where the block symmetric representative $\breve{Q} \in \mathcal{Q}(A)$ is a p.d. rank K matrix. To avoid some degenerated cases we also impose the restriction that there exists a transformation of the form $(S \otimes T)$ with full rank matrices $S, T \in \mathbb{R}^{n \times n}$ such that the transformed matrix

$$(S \otimes T)\breve{Q}(S \otimes T)'$$

has a full rank $K \times K$ left upper block.

The following theorem states that this subclass of BEKK terms is closely related to the F-GARCH model.

THEOREM 4.1. *For $K \leq n$ the set $\mathcal{B}_K^s$ is the set of all BEKK terms that have a representation of the form*

$$B_k = \mu_k \Gamma_k \Lambda_k', \qquad k = 1, \ldots, K,$$

where $\|\Lambda_k\| = \|\Gamma_k\| = 1$ and $\mu_k > 0$. Here Λ and Γ are two $n \times K$ matrices where Γ has full rank K and where Λ_k and Γ_k denote the respective kth columns. This representation is unique up to sign changes and permutations if and only if $\Lambda_k \neq \Lambda_s$ holds for all $k \neq s$.

Hence, the preceding restrictions essentially state that the BEKK matrices B_k are rank one matrices as in the F-GARCH model; see (4.3). In this sense we get a generalization of F-GARCH models for which the following restrictions hold:

https://doi.org/10.1017/S026646660707020X Published online by Cambridge University Press

1. $K \leq n$.
2. The BEKK matrices $B_{i,k}$ are rank one matrices:

$$B_{i,k} = \mu_{i,k}(\Gamma_{i,k}\Lambda_{i,k}'); \qquad \Gamma_{i,k}, \Lambda_{i,k} \in \mathbb{R}^n, \\ \|\Gamma_{i,k}\| = \|\Lambda_{i,k}\| = 1, \qquad \mu_{i,k} > 0. \tag{4.5}$$

Of course one has to impose some more technical conditions to get rid of trivial nonidentifiabilities, such as sign changes and permutations. (The same holds true for the F-GARCH model.)

This model is more flexible than the F-GARCH model, of course at the cost of the number of parameters, which is now equal to $n(n+1)/2 + (p+q)(2nK - K)$. We think this might be a valuable alternative in cases where the F-GARCH model is too restrictive. However, we have to admit that the stationarity condition here is much more involved than for the F-GARCH model. In addition, the "factor interpretation" gets lost, because the parameters $\Gamma_{i,k}$, $\Lambda_{i,k}$ may depend on the lag i and because the condition (4.4) is not imposed.

Let us conclude this section with some final remarks: A natural question is whether the rank constraint $K \leq n$ is sufficient to define unique BEKK parameter matrices. For $K = 1$ the question has been answered by Engle and Kroner (1995): If a VECH term A has a BEKK representation with $K = 1$ components, then the corresponding BEKK matrix B_1 is unique up to sign changes.

The case $1 < K \leq n$ has been discussed in Ribarits (2006). The main findings are as follows: Let $\mathcal{B}_K \subset \mathcal{B}$ denote the set of *BEKK* terms that have a BEKK representation with K components. There is a generic (i.e., open and dense) subset of $\mathcal{B}_K$ for which $\mathcal{Q}(A) \cap \mathcal{S}^+_{n^2,K}$ is a singleton. This means that for BEKK terms within this set the BEKK parameter matrices $(B_1, B_2, \ldots, B_K)$ are unique up to orthogonal transformations. Using, e.g., a Cholesky decomposition one may define unique BEKK matrices.

However there are also "nasty" BEKK terms in $\mathcal{B}_K$ for which $\mathcal{Q}(A) \cap \mathcal{S}^+_{n^2,K}$ is not a singleton and thus the BEKK matrices are not uniquely defined without adding further structure.

5. CONCLUSIONS

In this paper some issues of structure and parametrization of VECH and BEKK models are considered. A close look at the theory of this very general model class is worthwhile, not only for understanding the model features but also for the construction and parametrization of reasonable parsimonious subclasses.

Connected to each VECH term there is an affine subset of symmetric matrices, called $\mathcal{Q}(A)$. Here this fact is the main tool used to analyze VECH and BEKK terms. In particular, a VECH term has an equivalent BEKK representation iff this set $\mathcal{Q}(A)$ contains a p.s.d. matrix $Q \geq 0$. The corresponding BEKK

https://doi.org/10.1017/S026646660707020X Published online by Cambridge University Press

representation essentially corresponds to a factorization of such a $Q \geq 0$, and thus the number K of additive terms is related to the rank of Q. Therefore, one may check if a BEKK representation exists via an s.d.p. We also show that for the bivariate case admissible VECH models and BEKK models are equivalent. For $n > 2$ there is a "thick" class of admissible VECH models that have no equivalent BEKK representation. Thus, in this case VECH models are more general.

The problem of finding a unique BEKK representation may be decomposed into two steps. First, define a unique p.s.d. element in $\mathcal{Q}(A)$ and then use a unique factorization of this element. Based on this idea a parametrization of a generic set of BEKK models is presented. In addition, it is also shown that the Engle–Kroner parametrization does not cover a generic set.

The general BEKK model uses $O(n^4)$ free parameters, which makes estimation infeasible for large n. Anyway, for small n the parametrization of the general BEKK model is useful, and estimation can be performed.

In the literature there are a vast number of papers that define (parsimonious) subclasses of the general VECH and BEKK models. Here in particular, we discuss the popular DVECH model and the F-GARCH model using the previously described methodology. Finally, we consider a model class that relaxes some of the restrictions imposed by the F-GARCH model. An interpretation of this class in terms of restriction on the $\mathcal{Q}$-sets is given.

REFERENCES

Baba, Y., R.F. Engle, D.F. Kraft, & K.F. Kroner (1991) Multivariate simultaneous generalised ARCH. Discussion paper 89-57, University of California, San Diego, Department of Economics.

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

Bollerslev, T. (1986) Generalised autoregressive heteroscedasticity. *Journal of Econometrics* 31, 307–327.

Bollerslev, T., R. Engle, & D. Nelson (1994) ARCH models. In R. Engle & D. McFadden (eds.), *Handbook of Econometrics*, vol. 4, pp. 2959–3038. North-Holland.

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

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

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

Engle, R., V. Ng, & M. Rothschild (1990) Asset pricing with a factor-ARCH covariance structure, empirical estimates for treasury bills. *Journal of Econometrics* 45, 213–237.

Harville, D.A. (1997) *Matrix Algebra from a Statistician's Perspective*. Springer-Verlag.

Pagan, A. (1996) The econometrics of financial markets. *Journal of Empirical Finance* 3, 15–102.

Palm, F. (1996) GARCH models of volatility. In G. Maddala & C. Rao (eds.), *Handbook of Statistics*, vol. 14, pp. 209–239. Elsevier.

Ribarits, E. (2006) Multivariate modeling of financial time series. Ph.D. dissertation, Vienna University of Technology.

Vandenberghe, L. & S. Boyd (1996) Semidefinite programming. *SIAM Review* 38, 49–95.

https://doi.org/10.1017/S026646660707020X Published online by Cambridge University Press

APPENDIX: Technical Results and Proofs

Proof of Theorem 2.1. (i) and (ii): The minimum eigenvalue $\lambda_{\min}(Q)$ is a continuous function of $Q \in \mathcal{S}_m$. Using this property it is straightforward to show that $\gamma^*_{\mathcal{V}}(A)$ and $\lambda^*_{\mathcal{B}}(A)$ are continuous functions too; see, e.g., Ribarits (2006) for a detailed proof.

Let A_I be the VECH term corresponding to $Q = I_{n^2}$. Obviously, $1 = \gamma^*_{\mathcal{V}}(A_I) \geq \lambda^*_{\mathcal{B}}(A_I) \geq 1$ holds, and thus an open neighborhood of A_I is contained in $\mathcal{V}$ and in $\mathcal{B}$.

Concerning (iii): The set $\mathcal{B}^+$ is open in $\mathbb{R}^{n(n+1)/2 \times n(n+1)/2}$ because $\lambda^*_{\mathcal{B}}(\cdot)$ is continuous and because $\mathcal{B}^+$ is nonvoid. The set $\mathcal{B}^+$ is dense in $\mathcal{B}$ because of the fact that the set $\mathcal{S}^+_{n^2,n^2}$ is dense in $\mathcal{S}^+_{n^2}$.

It remains to prove (iv). First, consider the case $n = 2$: Let an admissible VECH term A be given and let $\mathcal{Q}(A) = \{Q_0 + \omega_1 \Delta_1\}$; see (2.12). Let (λ^*, ω_1^*) denote the optimizer of the s.d.p. (2.17) where without loss of generality we assume that $\omega_1^* = 0$ holds. Furthermore let $W \in \mathbb{R}^{4\times k}$, $1 \leq k \leq 4$ be an orthonormal basis of the eigenspace of Q_0 corresponding to its minimal eigenvalue λ^*. If $W'\Delta_1 W > 0$ then $(Q_0 - \lambda^* I_4) + \omega_1 \Delta_1 > 0$ holds for all sufficiently small $\omega_1 > 0$, which is a contradiction to the optimality of $\omega_1^* = 0$. Analogously we can rule out the case $W'\Delta_1 W < 0$. Hence, there must exist a vector $o \in \mathbb{R}^k$, $\|o\| = 1$ such that $(Wo)'\Delta_1(Wo) = \text{trace}(Wo)(Wo)'\Delta_1 = 0$. By Lemma A.1 in this Appendix it follows that (Wo) has a representation as $(Wo) = (\nu \otimes e)$ for suitably chosen $\nu, e \in \mathbb{R}^2$. Because A is admissible

$$\lambda^* = (Wo)'Q_0(Wo) = (\nu \otimes e)'Q_0(\nu \otimes e) \geq 0$$

holds, and thus Q_0 corresponds to a BEKK term.

For the case $n \geq 3$ it suffices to find an example, A, say, such that $\lambda^*_{\mathcal{B}}(A) < 0 < \gamma^*_{\mathcal{V}}(A)$ holds. See Ribarits (2006) for such an example. ■

Proof of Theorem 3.1. By what has been said before, in particular by the restriction $\theta_i > 0$ for all indices corresponding to the diagonal elements of V, it is immediate to see that both mappings are injective and differentiable. In addition, $\pi^+(\cdot)$ is surjective by construction.

It remains to prove that the inverse mapping $(\pi^+)^{-1}(\cdot)$ is differentiable. To this end, we prove that the derivative of $\pi^+(\cdot)$ has full rank for all $\theta \in \Theta^+$. Note that $\pi^+(\cdot)$ is a concatenation of the following mappings:

$$\theta \mapsto V \mapsto P = VV' \mapsto \bar{Q} = P^{-1} \mapsto A.$$

Let ∂V, ∂P, $\partial \bar{Q}$, and ∂A denote the respective derivatives of V, P, $\bar{Q}$, and A along a direction $\partial\theta \in \mathbb{R}^{(n(n+1)/2)^2}$. We have to prove that $\partial A = 0$ implies $\partial\theta = 0$. First, note that $\partial A = 0$ holds iff $\partial\bar{Q} = \partial\bar{Q}' = \sum_{k=1}^{\bar{n}} \Delta_k \omega_k$ for some $\omega_k \in \mathbb{R}$. Because $\partial\bar{Q} = -P^{-1}\partial P P^{-1}$ we get $\partial P = -\sum_k P\Delta_k P\omega_k$. The matrix P has symmetric subblocks by construction, and thus the same holds for the derivative ∂P; i.e., we have $\text{trace}\,\Delta_l \partial P = 0$ for $l = 1, \ldots, \bar{n}$. This implies $0 = \sum_k \text{trace}(\Delta_l P \Delta_k P)\omega_k$ for all $l = 1, \ldots, \bar{n}$, and thus $0 = \text{trace}(\partial\bar{Q}P\partial\bar{Q}P) = \text{trace}(P^{T/2}\partial\bar{Q}'P\partial\bar{Q}P^{1/2})$. Because $P > 0$ one obtains $\partial\bar{Q} = 0$ and $\partial P = 0$. Next, we use $v_{ij,ll} > 0$ to show that $\partial V = 0$ as $0 = \partial P = V\partial V' + \partial V V'$. Now, because θ is composed from the entries of V we have $\partial\theta = 0$ as desired. ■

https://doi.org/10.1017/S026646660707020X Published online by Cambridge University Press

The sets $\mathcal{S}_{n*n}$ and $\mathcal{A}_{n*n}$ have been defined in Section 1.1. These sets are invariant under transformations of the form $X \to (S \otimes T)X(S \otimes T)'$. This means that if $S, T \in \mathbb{R}^{n\times n}$ are two nonsingular matrices, then $Q \in \mathcal{S}_{n*n}$ holds iff $(S \otimes T)Q(S \otimes T)' \in \mathcal{S}_{n*n}$, and the same statement holds for matrices $\Delta \in \mathcal{A}_{n*n}$.

Suppose we have given two matrices $\Lambda, \Gamma \in \mathbb{R}^{n\times K}$ and let $U = (\Lambda_1 \otimes \Gamma_1, \ldots, \Lambda_K \otimes \Gamma_K) \in \mathbb{R}^{n^2\times K}$ where Λ_k, Γ_k denote the kth column of Λ, Γ, respectively. It is trivial to see that $Q = UU' \in \mathcal{S}_{n*n} \cap \overline{\mathcal{S}^+_{n^2,K}}$ holds. The next lemma states that for $K \le n$ the reverse statement also holds, given some regularity conditions.

LEMMA A.1. *Let $Q \in \mathcal{S}_{n*n} \cap \mathcal{S}^+_{n^2,K}$ be given. If $K \le n$ and if there exist nonsingular matrices $S, T \in \mathbb{R}^{n\times n}$ such that the first K rows and columns of $\bar{Q} = (S \otimes T)Q(S \otimes T)'$ form a p.d. $K \times K$ matrix, then there exist two matrices $\Lambda, \Gamma \in \mathbb{R}^{n\times K}$ such that Q has a factorization as $Q = UU'$, where $U = (\Lambda_1 \otimes \Gamma_1, \ldots, \Lambda_K \otimes \Gamma_K) \in \mathbb{R}^{n^2\times K}$ and where Λ_k, Γ_k denote the kth column of Λ, Γ, respectively.*

Proof. By $Q \ge 0$ and $\operatorname{rank} Q = K$ there exists a factorization of Q as $Q = \tilde{U}\tilde{U}'$, $\tilde{U} \in \mathbb{R}^{n^2\times K}$. The matrix $\tilde{U}$ is partitioned as

$$\tilde{U} = (\tilde{U}_1', \ldots, \tilde{U}_n')',$$

where $\tilde{U}_i = (\tilde{U}_{i1}', \tilde{U}_{i2}')'$ with $\tilde{U}_{i1} \in \mathbb{R}^{K\times K}$ and $\tilde{U}_{i2} \in \mathbb{R}^{(n-K)\times K}$.

First, note that following the preceding assumptions there exists a transformation $\tilde{U} \to \bar{U} = (S \otimes T)\tilde{U}$ such that $\bar{U}_{11} = I$ and $\bar{U}_{12} = 0$ holds. The block symmetry conditions imply $\bar{U}_i\bar{U}_j' = \bar{U}_j\bar{U}_i'$. For $i = 1$ these relations give $\bar{U}_{j2} = 0$ and $\bar{U}_{j1} = \bar{U}_{j1}'$ for all j. Furthermore we have $\bar{U}_{i1}\bar{U}_{j1}' = \bar{U}_{j1}\bar{U}_{i1}'$ for all i, j. From Harville (1997, Thm. 21.13.1) we know that in this specific case there exists an orthogonal matrix $O \in \mathbb{R}^{K\times K}$ such that $O'\bar{U}_{i1}O = \bar{D}_i$, with $\bar{D}_i = \operatorname{diag}(\bar{d}_{i1}, \ldots, \bar{d}_{iK})$ diagonal, holds for all $i = 2, \ldots, n$. Hence,

$$\begin{aligned}
&\left(I_n \otimes \begin{pmatrix} O' & 0 \\ 0 & I_{n-K} \end{pmatrix}\right)(S \otimes T)\tilde{U}O \\
&\quad = (I_K \quad 0_{n-K\times K} \mid \bar{D}_2 \quad 0_{n-K\times K} \mid \ldots \mid \bar{D}_n \quad 0_{n-K\times K})' \\
&\quad = (L_1 \otimes G_1, \ldots, L_K \otimes G_K),
\end{aligned}$$

where for $j = 1, \ldots, K$, $L_j = (1, d_{2j}, \ldots, d_{nj})'$ and $G_j = u_j$, the jth unit vector of dimension n. Therefore,

$$\begin{aligned}
U := \tilde{U}O &= (S^{-1} \otimes T^{-1})\left(I_n \otimes \begin{pmatrix} O & 0 \\ 0 & I_{n-K} \end{pmatrix}\right) \\
&\quad \times (L_1 \otimes G_1, \ldots, L_K \otimes G_K) = (\Lambda_1 \otimes \Gamma_1, \ldots, \Lambda_K \otimes \Gamma_K)
\end{aligned}$$

has the desired structure. ■

Proof of Theorem 4.1. This follows immediately from Lemma A.1. ■

https://doi.org/10.1017/S026646660707020X Published online by Cambridge University Press