CARF Working Paper

Massimiliano Caporin
Università degli Studi di Padova
Michael McAleer
Erasmus University Rotterdam
Tinbergen Institute
The University of Tokyo

August 2009

Do We Really Need Both BEKK and DCC?
A Tale of Two Covariance Models[1]

Massimiliano Caporin[2]

Dipartimento di Scienze Economiche "Marco Fanno"
Università degli Studi di Padova

Michael McAleer

Econometric Institute
Erasmus School of Economics
Erasmus University Rotterdam
and
Tinbergen Institute
The Netherlands
and
Center for International Research on the Japanese Economy (CIRJE)
Faculty of Economics
University of Tokyo

Revised: August 2009
First draft: July 2008

[2] Corresponding author: Massimiliano Caporin, Dipartimento di Scienze Economiche "Marco Fanno", Università degli Studi di Padova, Facoltà di Scienze Statistiche, Via del Santo, 33, 35123 Padova, Italy – email: Massimiliano.caporin@unipd.it – phone +39-049-827-4258, fax +39-049-827-4211.

Abstract

Large and very large portfolios of financial assets are routine for many individuals and organizations. The two most widely used models of conditional covariances and correlations are BEKK and DCC. BEKK suffers from the archetypal "curse of dimensionality" whereas DCC does not. This is a misleading interpretation of the suitability of the two models to be used in practice. The primary purposes of the paper are to define targeting as an aid in estimating matrices associated with large numbers of financial assets, analyze the similarities and dissimilarities between BEKK and DCC, both with and without targeting, on the basis of structural derivation, the analytical forms of the sufficient conditions for the existence of moments, and the sufficient conditions for consistency and asymptotic normality, and computational tractability for very large (that is, ultra high) numbers of financial assets, to present a consistent two step estimation method for the DCC model, and to determine whether BEKK or DCC should be preferred in practical applications.

1. Introduction

Large and very large portfolios of financial assets are routine for many individuals and organizations. Consequently, a careful analysis, specification, estimation, and forecasting of financial asset returns dynamics, and the construction and evaluation of financial portfolios, are essential in the tool kit of any financial planner and analyst. Correlations are used to determine portfolios, with appropriate attention being given to hedging and asset specialization strategies, whereas variances and covariances are used to forecast Value-at-Risk (VaR) thresholds to satisfy the requirements of the Basel Accords.

There are different models for different purposes, such as correlation models to create and evaluate a portfolio, and covariance models to forecast VaR on a daily basis for a given portfolio (see, for example, McAleer (2005)). The two most widely used models of conditional covariances and correlations are BEKK and DCC, as developed in Engle and Kroner (1995) and Engle (2002), respectively.

There are many similarities between BEKK and DCC. A scalar version of BEKK was compared with DCC, which is inherently scalar in practice, in Caporin and McAleer (2008). It was found empirically that scalar versions of the two models are very similar in forecasting conditional variances, covariances and correlations, which would suggest that they would also be similar in forecasting VaR thresholds and daily capital charges.

Accordingly, there are pertinent aspects regarding alternative versions of the two models that have not yet been addressed and clarified in the literature. These may be summarized as follows:

(i)   BEKK and DCC co-exist, despite one model being able to do virtually everything the other can do, thereby raising the pertinent question: Do we really need both BEKK and DCC?

(ii) DCC is used to forecast conditional correlations only[3], while its structure could easily be applied to forecast conditional covariances;

(iii) BEKK is used to forecast conditional covariances, although it may also be used to forecast conditional correlations indirectly;

(iv) The inherent differences between BEKK and DCC do not seem to be widely known. This situation is particularly relevant as DCC is equivalent to a scalar BEKK model applied to the standardized residuals, and can thereby be interpreted as a conditional correlation matrix only because of the standardization;

(v) The structural and statistical differences and similarities between the two models have not previously been analysed.

Engle and Kroner (1995) is the most widely cited paper in the history of the journal in which it appeared, but most citations would seem to be of a theoretical rather than empirical nature. The model is an archetypical example of an overparameterized model, thereby leading to the "curse of dimensionality". Engle (2002) is also widely cited, but most citations would seem to be of an empirical rather than theoretical nature.

The prevailing empirical wisdom would seem to be that DCC is preferred to BEKK because of the curse of dimensionality associated with the latter model. We argue that this is a misleading interpretation of the suitability of the two models to be used in practice.

The primary purpose of the paper is to define targeting as an aid in estimating matrices associated with large numbers of financial assets, discuss the use of targeting in estimating conditional covariance and correlation matrices in financial econometrics. In addition, we describe the following aspects: the similarities and dissimilarities between BEKK and DCC, both with and without targeting, on the basis of structural derivation; the analytical forms of the sufficient conditions for the existence of moments, and the sufficient conditions for consistency and asymptotic normality; the computational tractability for very large (that is, ultra high) numbers of financial assets; the use of consistent two step estimation methods for

---

[3] note we slightly deviate from the usual naming of DCC as covariance model since we focus on its correlation dynamic only, leaving aside the marginals capturing the conditional variance dynamic

the DCC model to enable it to be used sensibly in practical situations; and, finally, the determination of whether BEKK or DCC is to be preferred in empirical applications.

The plan of the remainder of the paper is as follows. Section 2 introduces the notation and models. The section compares alternative BEKK and DCC specifications, their corresponding mean specifications, the associated curse of dimensionality, the definition of the long run solution of a covariance (correlation) model as the unconditional expectation of the covariance (correlation), and the necessary and sufficient conditions for a covariance (correlation) model to be "targeted". The asymptotic results associated with various forms of the BEKK and DCC models are analyzed and summarized in Section 3. Some concluding remarks are given in Section 4.

2. A Comparison of the BEKK and DCC Specifications

The univariate models underlying DCC can be based on various conditional volatility specifications, such as the asymmetric GJR model of Glosten, Jagannathan and Runkle (1992) or the asymmetric/leverage EGARCH model of Nelson (1991). Similarly, both the BEKK and DCC structures might be generalized in alternative ways to introduce asymmetry, leverage and/or other stylized facts that are observed in financial returns variances and correlations.

However, as the primary focus of the paper is to compare the analytical and statistical performances of directly comparable BEKK and DCC models which are feasible under large cross-sectional dimensions, univariate and multivariate asymmetry and leverage are not considered. Moreover, forecasting comparisons of various versions of BEKK and DCC is left for further research (see also Caporin and McAleer (2008) for some results based on small scale models).

In this section we introduce the most relevant specifications which could be considered when fitting BEKK and DCC conditional covariance (and correlation) models to real data. We present the mean specification, the two conditional covariance and correlation models, and discuss the issues associated with the respective model structures and asymptotic properties. Within this section, we also highlight those issues that have not yet been addressed critically in the literature.

We will make use of the following operators below: ∘ denotes the Hadamard, or element-by-element, product; $dg(a)$ is a diagonal matrix, with scalar a along the main diagonal; $diag(A)$ is a vector formed from the elements of the main diagonal of A; I^{t-1} is the information set to time *t-1*; and **i** is a vector composed of unit elements.

2.1 Mean specification, curse of dimensionality and the concept of "targeting"

In order to make a fair comparison of models for the conditional second-order moments, we assume that the mean dynamics are common across all possible specifications, and are adequately captured by an un-specified conditional model. As a result, the mean innovations (or residuals) will be identically distributed according to a multivariate density with conditional covariance matrix Σ_t, and possibly dependent on a set of parameters θ (including, for instance, degrees of freedom or coefficients driving the distribution asymmetry).

Let x_t denote a *k*-dimensional vector of financial variables (returns), μ_t represent the expected mean of x_t obtained from a conditional mean model, and ε_t the mean innovation vector, as follows:

$$
\begin{aligned}
x_t \mid I^{t-1} &\sim D(\mu_t, \Sigma_t) \\
x_t - \mu_t = \varepsilon_t \mid I^{t-1} &\sim D(0, \Sigma_t)
\end{aligned}
\tag{1}
$$

In the following, we do not consider the effects of different mean specifications. The mean could be fixed at sample values, or could be based on a variety of time series models. The relevant issue is that, for each pair of covariance models we compare, the mean models are identical.

Definition 1: *The long run solution of a conditional covariance (correlation) model is given by the unconditional expectation of the dynamic conditional covariance (correlation).*

We present below two illustrative examples based on the simplest GARCH and BEKK models. These examples will be also used in the following[4].

Example 1: GARCH(1,1)

Consider the simple GARCH(1,1) model for asset i:

$$\sigma_{i,t}^2 = \omega_i + \alpha_i \varepsilon_{i,t-1}^2 + \beta_i \sigma_{i,t-1}^2, \quad i = 1, \ldots, k \tag{2}$$

It can be easily shown that the unconditional variance (the long run solution) of the model is given by:

$$\bar{\sigma}_i^2 = \omega_i \left(1 - \alpha_i - \beta_i\right)^{-1}. \blacksquare \tag{3}$$

Example 2: Scalar BEKK

Consider the Scalar BEKK model of Ding and Engle (2001), which is given as

$$\Sigma_t = CC' + \alpha \varepsilon_{t-1} \varepsilon_{t-1}' + \beta \Sigma_{t-1} \tag{4}$$

The unconditional covariance matrix of the model is

---

[4] These examples are intended to be illustrative of the arguments, and do not purport to be exhaustive.

$$E[\Sigma_t] = E\left[CC' + \alpha\varepsilon_{t-1}\varepsilon_{t-1}' + \beta\Sigma_{t-1}\right] = CC' + \alpha E\left[\varepsilon_{t-1}\varepsilon_{t-1}'\right] + \beta E[\Sigma_{t-1}],$$

$$E\left[\varepsilon_{t-1}\varepsilon_{t-1}'\right] = E[\Sigma_t] \quad E[\Sigma_{t-1}] = E[\Sigma_t],$$

$$E[\Sigma_t] = CC' + \alpha E[\Sigma_t] + \beta E[\Sigma_t],$$

$$E[\Sigma_t] = \bar{\Sigma} = CC'(1-\alpha-\beta)^{-1}. \blacksquare \tag{5}$$

The long-run solution for the BEKK model can be derived easily by mimicking the approach traditionally adopted for the GARCH(1,1) model.

Two topics that are discussed in the financial econometrics literature regarding covariance/correlation model estimation, namely the "curse of dimensionality" and "targeting". The first issue is perceived as the most serious problem in covariance modeling, while the second could be considered as a tool for disentangling the serious problem.

It is known that many fully parameterized conditional covariance models have the number of parameters increasing at an order larger than the number of assets, otherwise known as the "curse of dimensionality". For example, the most general BEKK model of Engle and Kroner (1995) has parameters increasing with order $O(k^4)$, its commonly used specification increases with order $O(k^2)$, the VECH model parameter number has order $O(k^4)$, and the Generalized DCC model of Engle (2002) increases with order $O(k^2)$ (further comments regarding the numbers of parameters and the cited models will be discussed in the following subsections).

In order to control the growth in the number of parameters, several restricted specifications have been proposed in the literature, such as the scalar and diagonal models presented in Ding and Engle (2001), the block structured specifications suggested by Billio, Caporin and Gobbo (2005), Billio and Caporin (2009), Asai, Caporin and McAleer (2009), Bonato, Caporin and Ranaldo (2009), and the parameter restrictions inspired by spatial econometrics concept introduced in Caporin and Paruolo (2009). However, restrictions generally operate on the parameters driving the dynamics, while little can be done on the model intercepts which include $O(k^2)$ parameters in both the conditional covariance and correlation models. This still exposes the models to the curse of dimensionality (to the best of our knowledge, Caporin and Paruolo (2009) is the only paper proposing parameter restrictions on the model intercepts).

The "targeting" constraint then becomes useful because it imposes a structure on the model intercept based on sample information. Within "targeting", the constants in the dynamic equations are structured in order to make explicit the long run target, which is then fixed using a consistent (sample) estimator. As a result, the number of parameters to be estimated by maximizing a conditional log-likelihood function can be reduced substantially. Although targeting can be applied to both BEKK and DCC, in practice it has been used only for DCC. Targeting can be very useful computationally when the number of financial assets is large (say, $k > 20$), but can become essential when the number of assets is very large (such as $k > 100$).

We define the "targeting" constraint in the following proposition:

Definition 2: *A conditional covariance (correlation) model is "targeted" if and only if the following two conditions are satisfied:*

  i) *the intercept is an explicit function of the long run covariance (correlation);*

  ii) *the long run covariance (correlation) solution is replaced by a consistent estimator of the unconditional sample covariance (correlation) of the observed data.*

Remark 1: Condition i) implicitly requires the long run solution of the covariance (correlation) model to be equal to the long run covariance (correlation), and ensures that the long run solution does not depend on any parameters. Thus targeting should be distinguished from the imposition of parametric restrictions.

Remark 2: Condition ii) implies the use of all the available sample data in constructing a consistent estimator of the observed long run covariance (correlation). The definition of targeting excludes estimating the long run matrices using latent variables. Such exclusion is essential because estimation of latent variables in the conditional volatility literature does not ensure by construction the consistency of the estimator used for the sample covariance (correlation).

9

Remark 3: Further to Remark 2, and contrary to what has been stated to have been proved in the literature, consistency and asymptotic normality of the estimated parameters of any version of DCC has not yet been established (see the Appendix).

The following examples clarify these aspects.

Example 1 (continued)

In the GARCH(1,1) model of equation (2), there are three parameters to estimate for each asset, namely ω_i, α_i and β_i, $i = 1,2, \ldots, k$. We know that the long run solution of the model is $\bar{\sigma}_i^2 = (1 - \alpha_i - \beta_i)^{-1} \omega_i$. This result could be used to make explicit the long run variance in the GARCH equation by replacing the conditional variance constant, ω_i, with an alternative expression:

$$\sigma_{i,t}^2 = (1 - \alpha_i - \beta_i)\bar{\sigma}_i^2 + \alpha_i \varepsilon_{i,t-1}^2 + \beta_i \sigma_{i,t-1}^2, \tag{6}$$

where the long run variance becomes a parameter to be estimated.

Equation (6) is equivalent to the standard GARCH(1,1) model as there are three parameters to be estimated, namely α_i, β_i and $\bar{\sigma}_i^2$. The model is targeted if, in this alternative representation, $\bar{\sigma}_i^2$ is matched with the sample information, in which case there would only be two parameters to be estimated by maximum likelihood, namely α_i and β_i. Therefore, we may estimate the long run variance using the sample variance estimator of $\varepsilon_{i,t}$, and substitute this into the model, thereby reducing the number of parameters to be estimated (namely, 3 parameters in equation (2) but only 2 parameters in equation (6)). Consistency of the other parameters is not influenced as the sample variance can be consistently estimated. ∎

Example 2 (continued)

In a similar manner to Example 1, we can replace the intercept of the BEKK model of equation (4) using the long run solution in (5), thereby obtaining

$$\Sigma_t = \bar{\Sigma}(1 - \alpha - \beta) + \alpha \varepsilon_{t-1} \varepsilon_{t-1}' + \beta \Sigma_{t-1}. \tag{7}$$

Equation (7), without additional constraints, has two parameters associated with the dynamics and $k(k+1)/2$ in the intercept, $\bar{\Sigma}$ (the parameters in the long run covariance). Targeting implies the use of a sample covariance estimator for $\bar{\Sigma}$ and the maximization of the likelihood function with respect to the parameters α and β (maximization is made conditionally on the estimates of the long run covariance). The introduction of targeting reduces the number of intercept parameters, thereby making estimation feasible, even for large cross-sectional dimensions. However, the model will be still computationally complicated for large k because the likelihood evaluation of the model in (7) requires the inversion of a covariance matrix of dimension k. ∎

Although targeting can be computationally useful in terms of reducing the number of parameters to be estimated by maximum likelihood, sometimes dramatically, it requires care in terms of the sample estimator that is used. If targeting were to use an inconsistent estimator to reduce the number of parameters, as is typical in the literature, the resulting estimators will also be inconsistent.

2.2 BEKK models

11

Engle and Kroner (1995) introduced the BEKK class of multivariate GARCH models. The specification they proposed was sufficiently general to allow the inclusion of special factor structures (see Bauwens et al., 2006). In this paper, we consider the simplest BEKK specification with all orders set to 1:

$$\Sigma_t = CC' + A\varepsilon_{t-1}\varepsilon_{t-1}'A' + B\Sigma_{t-1}B' \qquad (8)$$

where A and B are $k \times k$ parameter matrices (not necessarily symmetric) and C is a lower triangular parameter matrix. The fully parameterized model includes $2.5k^2 + 0.5k$ parameters. The conditional covariance matrices are positive definite, by construction, and the conditional variances are positive, regardless of the parameter signs. Covariance stationarity of the BEKK model is discussed in Engle and Kroner (1995), together with a representation that is more general than the one given in (8), but which does not seem to have been used in empirical applications.

Fully parameterized BEKK models are feasible only for small values of k, typically less than 10. In order to make the model feasible for large cross-sectional dimensions, two restricted parameterizations have been proposed in Ding and Engle (2001), namely the diagonal and scalar specifications.

In the scalar BEKK model, the parameter matrices A and B in (8) are replaced by $A = \alpha^{\frac{1}{2}}\mathbf{i}\mathbf{i}'$ and $B = \beta^{\frac{1}{2}}\mathbf{i}\mathbf{i}'$, leading to the following specification:

$$\Sigma_t = CC' + \alpha\varepsilon_{t-1}\varepsilon_{t-1}' + \beta\Sigma_{t-1}. \qquad (9)$$

In the diagonal specification, the parameter matrices A and B are set to be diagonal as $A = dg(a)$ and $B = dg(b)$, so that the model has the following structure:

$$\Sigma_t = CC' + dg(a)\varepsilon_{t-1}\varepsilon_{t-1}'dg(a) + dg(b)\Sigma_{t-1}dg(b) = CC' + (aa') \circ \varepsilon_{t-1}\varepsilon_{t-1}' + (bb') \circ \Sigma_{t-1}. \quad (10)$$

An additional representation of a BEKK-type model may be based on the Hadamard matrix product:

$$\Sigma_t = CC' + A \circ \varepsilon_{t-1}\varepsilon_{t-1}' + B \circ \Sigma_{t-1}. \quad (11)$$

In this case, the parameter matrices A and B must be symmetric and positive definite, and the number of parameters is still O(k^2). Generally, (11) is not estimated directly, but rather by imposing a structure for A and B to ensure positive definiteness (by making A and B equal to the product of triangular matrices, as in the case of the intercept). Positive definiteness of conditional covariance matrices is guaranteed, by construction (see Ding and Engle, 2001). Finally, we note that the diagonal specification in (10) is a restricted parameterization of the BEKK model in equation (8) and also of the Hadamard BEKK model in equation (11). Similarly, the scalar BEKK model in (9) can also be obtained from (11) by setting $A = \alpha \mathbf{ii}'$ and $B = \beta \mathbf{ii}'$. We also highlight that imposing diagonal A and B matrices in (11) is not considered since this restriction induces constant covariances.

Although it is not necessary to do so, BEKK can be specified with targeting. The introduction of this feature may require appropriate constraints to be imposed at the estimation step in order to guarantee that the covariance matrices are positive definite. As argued in Proposition 2, the targeting constraints require two elements: a modification in the model structure, and matching some of the model parameters with appropriate sample estimators.

Define the sample covariance matrix $E\left[\varepsilon_t\varepsilon_t'\right] = \overline{\Sigma}$, which can be consistently estimated by the sample estimator. The BEKK equations may be redefined, as follows:

13

Scalar BEKK with targeting: $\quad \Sigma_t = \bar{\Sigma} + \alpha \left(\varepsilon_{t-1} \varepsilon_{t-1}' - \bar{\Sigma} \right) + \beta \left(\Sigma_{t-1} - \bar{\Sigma} \right)$ $\qquad$ (12)

Diagonal BEKK with targeting: $\quad \Sigma_t = \bar{\Sigma} + \left(aa' \right) \circ \left(\varepsilon_{t-1} \varepsilon_{t-1}' - \bar{\Sigma} \right) + \left(bb' \right) \circ \left(\Sigma_{t-1} - \bar{\Sigma} \right)$ $\qquad$ (13)

BEKK with targeting: $\quad \Sigma_t = \bar{\Sigma} + A \left(\varepsilon_{t-1} \varepsilon_{t-1}' - \bar{\Sigma} \right) A' + B \left(\Sigma_{t-1} - \bar{\Sigma} \right) B'$ $\qquad$ (14)

Hadamard BEKK with targeting: $\quad \Sigma_t = \bar{\Sigma} + A \circ \left(\varepsilon_{t-1} \varepsilon_{t-1}' - \bar{\Sigma} \right) + B \circ \left(\Sigma_{t-1} - \bar{\Sigma} \right)$ $\qquad$ (15)

In all of these specifications, it follows that $E[\Sigma_t] = \bar{\Sigma}$, as $E\left[\varepsilon_{t-1} \varepsilon_{t-1}' \right] = \bar{\Sigma}$ and $E[\Sigma_{t-1}] = \bar{\Sigma}$. However, positive definiteness of the conditional covariance matrices must be imposed at the estimation step by constraining the matrix of intercepts in the model, otherwise the estimates cannot be interpreted as covariance matrices.

For the BEKK, diagonal BEKK and Hadamard BEKK models, we can guarantee positive definiteness of the conditional covariance matrices by imposing positive definiteness of $\bar{\Sigma} - A\bar{\Sigma}A' - B\bar{\Sigma}B'$, $\bar{\Sigma} - dg(a)\bar{\Sigma}dg(a) - dg(b)\bar{\Sigma}dg(b)$ and $\bar{\Sigma} - A \circ \bar{\Sigma} - B \circ \bar{\Sigma}$, respectively. In the scalar case, the inequality constraint $\alpha + \beta < 1$ imposes positive definiteness of the conditional covariances.

Although the constraints may seem to be quite simple, their computational complexity is extremely relevant, in particular, when the cross-sectional dimension is simply moderate rather than high. In fact, imposing positive definiteness of the intercepts of (13), (14) or (15) results in a set of highly non-linear constraints on the parameters. One way of imposing positive definiteness is through the imposition of positivity of the eigenvalues of the intercepts. However, such a constraint is non-linear in the parameters and is extremely complicated, except for the scalar case.

In addition, covariance stationarity constraints should be taken into account. These are generally simple in restricted specifications, as shown by Engle and Kroner (1995). However, in fully parameterized cases, these additional constraints significantly increase the computational complexity of the model.

Example 3: Constraints in the bivariate BEKK model with targeting

Consider the simple case with $k = 2$. We assume that the positive definiteness of the conditional covariance matrices is imposed by constraining the eigenvalues of the intercepts in equations (12) to (15) to be positive. The model intercepts can be represented as

Scalar BEKK with targeting: $\qquad (1 - \alpha - \beta)\bar{\Sigma}$ $\qquad\qquad\qquad$ (16)

Diagonal BEKK with targeting: $\qquad \bar{\Sigma} - (aa') \circ \bar{\Sigma} - (bb') \circ \bar{\Sigma} = \bar{\Sigma} \circ (\mathbf{ii}' - aa' - bb')$ $\qquad$ (17)

BEKK with targeting: $\qquad \bar{\Sigma} - A\bar{\Sigma}A' - B\bar{\Sigma}B'$ $\qquad\qquad\qquad$ (18)

Hadamard BEKK with targeting: $\qquad \bar{\Sigma} - A \circ \bar{\Sigma} - B \circ \bar{\Sigma} = \bar{\Sigma} \circ (\mathbf{ii}' - A - B)$ $\qquad$ (19)

In (16), positive definiteness is imposed by restricting parameters to be positive and their sum to be less than 1. In (17) positive definiteness is achieved by imposing positive definiteness of $(\mathbf{ii}' - aa' - bb')$. In turn, this requires imposing positivity of the eigenvalues of $(\mathbf{ii}' - aa' - bb')$ which translates into the following constraints

$$1 - \frac{1}{2}\alpha_1^2 - \frac{1}{2}\alpha_2^2 - \frac{1}{2}\beta_1^2 - \frac{1}{2}\beta_2^2 - \frac{1}{2}\sqrt{\begin{array}{l} 4 + \alpha_1^4 + \alpha_2^4 + 2\alpha_1^2\alpha_2^2 - 8\alpha_1\alpha_2 + \beta_1^4 + \beta_2^4 + 2\beta_1^2\beta_2^2 - 8\beta_1\beta_2 \\ +2\alpha_1^2\beta_1^2 + 2\alpha_2^2\beta_2^2 - 2\alpha_1^2\beta_2^2 - 2\alpha_2^2\beta_1^2 + 8\alpha_1\alpha_2\beta_1\beta_2 \end{array}} > 0$$

$$1 - \frac{1}{2}\alpha_1^2 - \frac{1}{2}\alpha_2^2 - \frac{1}{2}\beta_1^2 - \frac{1}{2}\beta_2^2 + \frac{1}{2}\sqrt{\begin{array}{l} 4 + \alpha_1^4 + \alpha_2^4 + 2\alpha_1^2\alpha_2^2 - 8\alpha_1\alpha_2 + \beta_1^4 + \beta_2^4 + 2\beta_1^2\beta_2^2 - 8\beta_1\beta_2 \\ +2\alpha_1^2\beta_1^2 + 2\alpha_2^2\beta_2^2 - 2\alpha_1^2\beta_2^2 - 2\alpha_2^2\beta_1^2 + 8\alpha_1\alpha_2\beta_1\beta_2 \end{array}} > 0$$

$\qquad$ (20)

Similar restrictions, but clearly more complex (6 parameters involved rather than 4), are required for (19). Finally, in the case of (18), positive definiteness is imposed by requiring the eigenvalues of the entire intercept to be positive, which induces two non-linear constraints involving the 8 model parameters and the long run covariance elements. ∎

Other than for the scalar case, it is clear that imposing positive definiteness and covariance stationarity for various versions of BEKK is extremely complicated when there is more than one asset. The different degrees of complexity of several BEKK models can also be drawn from Table 1, where columns 3 to 5 report the parameter numbers for each specification in a general representation, and also for the cases $k=10$ and $k=100$.

2.3 DCC models

The Dynamic Conditional Correlation (DCC) model was introduced by Engle (2002) as a generalization of the Constant Conditional Correlation (CCC) model of Bollerslev (1990). In this case, the focus is on the separate modeling of the conditional variances and conditional correlations.

The covariance matrix is decomposed as follows:

$$\Sigma_t = D_t R_t D_t \tag{21}$$

$$D_t = diag\left(\sigma_{1,t}, \sigma_{2,t}, \ldots \sigma_{k,t}\right) \tag{22}$$

$$R_t = \bar{Q}_t^{-1} Q_t \bar{Q}_t^{-} \ , \qquad \bar{Q}_t = dg\left(Q_t\right) \tag{23}$$

where D_t includes the conditional volatilities which are modeled by a set of univariate GARCH equations (see Bollerslev (1990) and Engle (2002)). The dynamic correlation matrix, R_t, is not explicitly driven by a dynamic equation, but is derived from a standardization of a different matrix Q_t which has a dynamic structure. The form of Q_t determines the model complexity and the feasibility in large cross-sectional dimensions.

16

Several specifications have been suggested for Q_t. We present here the most simple specifications, which can be matched with the BEKK cases illustrated in equations (8)-(11). The DCC model (or Hadamard DCC) is given in Ding and Engle (2001) and Engle (2002) as:

$$Q_t = S + A \circ \left(D_{t-1}^{\boxempty} \varepsilon_{t-1} \varepsilon_{t-1}' D_{t-1} - S \right) + B \circ \left(Q_{t-1} - S \right), \tag{24}$$

where A and B are symmetric parameter matrices and S is a long run correlation matrix. As distinct from standard practice, we maintain explicitly in the model the dependence on the conditional variances. This model has parameter numbers of order $O(k^2)$, meaning that the model is affected by the curse of dimensionality. Notably, the model has been proposed in the literature directly with a targeting constraint, thereby highlighting the long run component. However, we note that imposing targeting in (24) is counterintuitive since the Q_t are then standardized to obtain dynamic conditional correlations. Targeting was included as a tool for the reduction of the numbers of parameters, given that the S matrix could be estimated by the sample correlation matrix, so that A and B can be estimated by maximum likelihood, conditional on the value assigned to S.

Aielli (2008) shows that the sample correlation is an inconsistent estimator of S, thereby eliminating the advantage of targeting as a tool for controlling the curse of dimensionality for DCC models. A deeper discussion of the asymptotics is given in the following section.

The Hadamard DCC (HDCC) model without targeting has the following structure:

$$Q_t = CC' + A \circ D_{t-1}^{\boxempty} \varepsilon_{t-1} \varepsilon_{t-1}' D_{t-1} + B \circ Q_{t-1}, \tag{25}$$

where C is a lower triangular matrix whose diagonal elements are constrained in order to ensure CC' is a correlation matrix. Considering now the alternative specifications for Q_t, we

17

first highlight that diagonal specifications cannot be used for the GDCC model, as for the Hadamard BEKK.

An alternative fully parameterized model, the Generalized DCC (GDCC) specification, is given by Cappiello, Engle and Sheppard (2006). The dynamic equation driving the conditional correlation matrix are as follows for the cases with and without targeting, respectively:

$$Q_t = S + A\left(D_{t-1}^{蝸}\varepsilon_{t-1}\varepsilon_{t-1}'D_{t-1} - S\right)A' + B\left(Q_{t-1} - S\right)B', \tag{26}$$

$$Q_t = CC' + AD_{t-1}^{蝸}\varepsilon_{t-1}\varepsilon_{t-1}'D_{t-1}A' + BQ_{t-1}B', \tag{27}$$

where *A* and *B* are parameter matrices (not necessarily symmetric), while *S* and *C* are as in equations (24) and (25). The GDCC model has parameter numbers increasing with order $O(k^2)$ as the Hadamard DCC. However, despite the introduction of correlation targeting, the two models, Hadamard DCC and Generalized DCC, are infeasible with large cross sectional dimensions because the parameter matrices *A* and *B* in both models include $O(k^2)$ parameters.

Two major restricted specifications may be considered, namely the diagonal and scalar models. Notably, as in the BEKK model, the scalar representation is a special case of both the HDCC and GDCC models, while the diagonal specification of GDCC may be associated with a restricted HDCC model.

The alternative DCC specifications are reported in the following equations:

- Scalar DCC with targeting: $\qquad Q_t = S + \alpha\left(D_{t-1}^{蝸}\varepsilon_{t-1}\varepsilon_{t-1}'D_{t-1} - S\right) + \beta\left(Q_{t-1} - S\right)$ $\qquad$ (28)

- Scalar DCC without targeting: $\qquad Q_t = CC' + \alpha D_{t-1}^{蝸}\varepsilon_{t-1}\varepsilon_{t-1}'D_{t-1} + \beta Q_{t-1}$ $\qquad$ (29)

- Diagonal DCC with targeting: $\qquad Q_t = S + \left(aa'\right)\circ\left(D_{t-1}^{蝸}\varepsilon_{t-1}\varepsilon_{t-1}'D_{t-1} - S\right) + \left(bb\right)\circ\left(Q_{t-1} - S\right)$ (30)

- Diagonal DCC without targeting: $\quad Q_t = CC' + (aa') \circ \left(D_{t-1}^{擶} \varepsilon_{t-1} \varepsilon_{t-1}' D_{t-1} \right) + (bb) \circ Q_{t-1},$ (31)

where α and β are scalars, a and b are vectors, while C and S are as in equations (24) and (25), respectively.

The most frequently estimated version of DCC (put simply, the DCC model) is what we will call the scalar DCC model, for purposes of strict comparability with its scalar BEKK counterpart. Note that the models without targeting require the joint estimation of all the parameters, including the long run correlations. If targeting is excluded, all models are affected by the curse of dimensionality. Such a result can be observed in Table 1, where the parameter dimension for DCC without targeting is comparable to that of the standard BEKK models.

3. Asymptotic Theory

Several papers have purported to establish the consistency and asymptotic normality of the Quasi Maximum Likelihood Estimation (QMLE) of BEKK and DCC. Apart from two papers that have proved consistency and asymptotic normality of BEKK, albeit under high-order stated but untestable assumptions, the proofs for DCC have typically being based on unstated regularity conditions. When the regularity conditions have been stated, they are untestable or irrelevant for the stated purposes. These comments will become clearer in the remainder of this section.

3.1 BEKK

Jeantheau (1998) proved consistency of BEKK under the multivariate log-moment condition. However, the derivation of the log-moment condition requires the assumption of the existence of sixth order moments, which cannot be tested.

19

Using the consistency result proved in Jeantheau (1998), Comte and Lieberman (2003) established the asymptotic normality of the QMLE of BEKK under eighth order moments which, though stated explicitly, cannot be tested.

The consistency and asymptotic normality results for Scalar and Diagonal BEKK follow as special cases of the results given above, while those of Hadamard BEKK can be derived similarly by noting that the Hadamard BEKK has a companion VECH representation with diagonal parameter matrices.

The proofs of Jeantheau (1998) and Comte and Lieberman (2003) can also be generalized to include the BEKK representations where the long run solution of the model enters explicitly in the intercept. In such a case, appropriate modifications of the regularity conditions are required.

Therefore, the asymptotic theory for BEKK models is established, albeit under untestable conditions.

3.2 DCC

The primary appeal of the DCC specification, at least in its scalar representation, is supposed to be its computational tractability for very large numbers of financial assets, with two step estimation reducing the computational complexity relative to systems maximum likelihood estimation.

This presumption is appropriate if:

i) the model can be targeted;

ii) the two step estimators are consistent;

iii) the parameter number increase as a power function of the cross-sectional dimension with an exponent smaller than or equal to 1.

Point i), targeting, reduces by $0.5k(k-1)$ the number of parameters to be estimated by QMLE, given that it fixes part of the intercept. Differently, point ii) ensures that correct inferential procedures could be derived from the estimated parameters and the likelihood (including

parameter testing, model restrictions and LM tests against more general representations). Furthermore, it ensures that the forecasts will not be influenced by parameter distortions. Finally, point iii) controls for the parameters in the model dynamic. Conditions i) and iii) avoid the curse of dimensionality, while the inclusion of just one of the two previous points (either i) or iii)) makes the model feasible only for small dimensional systems (the full model parameters will increase at least with power $O(k^2)$).

Engle (2002) suggests the introduction of targeting (point i)), the use of scalar representations (point iii)), and assumes that the standard regularity conditions yielding consistent and asymptotically normal QML two step estimators are satisfied (point ii)).

However, Aielli (2008) has proved that two step estimation of DCC models with targeting is inconsistent. In fact, Aielli shows that the sample correlation estimator is an *inconsistent* estimator of the matrix S in the dynamic equations (24), (26), (28) and (30). As a result, the parameters driving the dynamics cannot be consistently estimated by QML, conditional on an inconsistent estimator of S. His result is based on the observation that the unconditional expectations of Q_t may differ from the unconditional expectation of $D_t^{-1}\varepsilon_t\varepsilon_t'D_t$, the first being a covariance while the second is a correlation, by construction. In fact, the Q_t matrices are never referred to as correlations, sometimes as non-standardised correlations, which, presuming positive definiteness and symmetry, are covariance matrices. As a result, the sample correlation estimator of $D_t^{-1}\varepsilon_t\varepsilon_t'D_t$ is not a consistent estimator of the S matrix. The long run solution cannot be estimated with a sample estimator which, in turn, eliminates the targeting constraint in point i) and, as a consequence, makes the parameter number at least of order $O(k^2)$. In turn, this affects consistency of the QML estimates of the other parameters as well as their asymptotic distribution, also eliminating point ii). Therefore, all the purported proofs for models with targeting in Engle (2002) and Engle and Sheppard (2001) must be reconsidered.

The need to introduce the long run solution matrix S into the estimation step of QML makes DCC (even in the scalar case) inconsistent with its primary intended purpose, namely the computational tractability for large cross sections of assets (see Table 1).

Aielli (2008) suggests a correction to the DCC model to resolve the previous inconsistency between unconditional expectations. However, the new model proposed does not allow targeting, as given in Definition 2. Furthermore, the asymptotic results are not fully reported (the author presumes a number of regularity conditions without stating them). It is worth mentioning that Aielli's model was used in Engle, Sheppard and Shephard (2008), under the assumption that it included targeting, which is not possible.

Aielli's results preclude the estimation of the DCC model with targeting, but this does not affect the DCC specifications without targeting. Hence, the asymptotic properties are still not known. Clearly, despite the possibility of estimating DCC models in a single step, the curse of dimensionality will always be present as the intercept include *0.5k(k-1)* parameters in the long run correlation matrix.

In summary, the purported asymptotic theory for DCC models has simply been stated without formal proofs of the conditions required for the results to hold, and without checking any of the assumptions underlying the general results in Newey and McFadden (1994).

3.3 Consistent estimation of correlation matrices from BEKK

McAleer et al. (2008) showed that scalar BEKK and diagonal BEKK could be derived as a multivariate extension of the vector random coefficient autoregression (RCA) model of Tsay (1987) (see Nicholls and Quinn (1982) for a statistical analysis of random coefficient models). However, BEKK and Hadamard BEKK cannot be derived using the RCA approach.

Caporin and McAleer (2008) show that a theoretical relation can be derived comparing scalar DCC and BEKK models with and without targeting. They suggest the derivation of conditional correlations from BEKK representations, and refer to these derived models as Indirect DCC (which, despite its name, is not a different model but rather a bi-product of the BEKK estimates).

As there is presently no consistency result for DCC parameters estimated by QML, the theorem below will represent a first contribution in the area. Its advantage will be clarified in the following:

Theorem 1: *The indirect DCC Conditional Correlation matrices derived from BEKK representations are consistent estimates of the underlying true conditional correlations.*

Proof: The conditional covariance matrix Q_t satisfy the decomposition $Q_t = D_t \Gamma_t D_t$. If the dynamic covariances have been estimated by a BEKK model with or without targeting, they are consistent. The matrices D_t contain the conditional volatilities along the main diagonal. In turn, these may be obtained as part of the conditional variance matrix Q_t or from a different univariate or multivariate GARCH model. In all cases, they will include consistent estimates of the conditional volatilities, as given by the results for BEKK models, or in Bougerol and Picard (1992) (for univariate models), and Ling and McAleer (2003) (for VARMA-GARCH specifications). Therefore, the indirect conditional correlations, $\Gamma_t = D_t^{-1} Q_t D_t^{-1}$, are given by the product of consistent estimators of the conditional covariance matrices and conditional standard deviations, and are therefore consistent.∎

This shows how the BEKK model may be used for obtaining consistent estimates of the conditional correlation matrix. The BEKK model may be also used to derive starting values for a full system estimation of DCC models by QML. In this case, the intercept may be calibrated as the sample mean of indirect conditional correlations while the DCC parameter may be calibrated at the corresponding parameters in a given BEKK model.

An example showing the indirect derivation of dynamic conditional correlations from scalar BEKK estimates is given in Caporin and McAleer (2008).

4. Conclusions

The efficient management and monitoring of large and very large portfolios of financial assets are routine for many individuals and organizations. Quantitative tools are then used to analyze

financial asset returns for the purposes of generating forecasts, and then constructing, managing and evaluating financial portfolios. There are different models for different purposes, such as correlation models to create and evaluate a portfolio, and covariance models to forecast VaR on a daily basis for a given portfolio.

BEKK and DCC are the two most widely used models of conditional covariances and correlations, as developed in Engle and Kroner (1995) and Engle (2002), respectively. The two models are similar in many respects but, despite this situation, the literature has not addressed some critical issues pertaining to these models, namely: clarification of the reasons for BEKK and DCC to co-exist when one model can do virtually everything the other model can do; determination as to why DCC is used to forecast conditional correlations rather than conditional covariances, and why BEKK is used to forecast conditional covariances rather than conditional correlations; examination of the inherent differences between BEKK and DCC, especially when DCC is equivalent to a scalar BEKK model applied to the standardized residuals; and comparisons of the structural and statistical differences and similarities between the two models.

The paper tackled some of these issues. In particular, we organized and rationalized the theoretical literature on the BEKK and DCC models, highlighting that the latter model has asymptotic properties under untestable moment conditions. In short, the purported asymptotic properties of DCC are simply stated under a set of untested and untestable regularity conditions.

In addition, we clarified the concept of targeting as a tool for reducing the curse of dimensionality associated with multivariate conditional covariance models.

Finally, we provided a result which demonstrated that the BEKK model can be used to obtain consistent estimates of dynamic conditional correlations, with a direct link to the Indirect DCC given in Caporin and McAleer (2008).

In summary, the paper demonstrated that, from a theoretical perspecrtive, the optimal model for estimating conditional covariances (and conditional correlations) is the scalar BEKK model, regardless of whether targeting was used.

References

Aielli, 2008, Consistent estimation of large scale dynamic conditional correlations, University of Messina, Department of Economics, Statistics, Mathematics and Sociology, Working paper n. 47

Asai, M., Caporin, M., and McAleer, M., 2009, Block structure multivariate stochastic volatility models, unpublished manuscript

Bauwens L., Laurent, S. and Rombouts, J.K.V., 2006, Multivariate GARCH models: a survey, *Journal of Applied Econometrics*, 21, 79-109.

Billio, M. and Caporin M., 2009, A generalised dynamic conditional correlation model for portfolio risk evaluation, *Mathematics and Computers in Simulation*, 79(8), 2566-2578.

Billio, M., Caporin, M. and Gobbo, M., 2006, Flexible dynamic conditional correlation multivariate GARCH for asset allocation, *Applied Financial Economics Letters*, 2, 123-130.

Bollerslev T., 1990, Modelling the coherence in short-run nominal exchange rates: a multivariate generalized ARCH approach, *Review of Economic and Statistics*, 72, 498-505.

Bonato, M., Caporin, M., and Ranaldo, A., 2009, Forecasting realized (co)variances with a block structure Wishart autoregressive model, Swiss National Bank Working Paper

Bougerol, P. And Picard, N., 1992, Stationarity of GARCH processes, *Journal of Econometrics*, 52, 115-127.

Caporin, M. And McAleer, M., 2008, Scalar BEKK and Indirect DCC, *Journal of Forecasting*, 27-6, 537-549.

Caporin, M., and Paruolo, P., 2009, Structured multivariate volatility models, Department of Economics "Marco Fanno", University of Padova, Working Paper 90

Cappiello L., Engle, R.F. And Sheppard, K., 2006, Asymmetric dynamics in the correlations of global equity and bond returns, *Journal of Financial Econometrics*, 4, 537-572.

Comte, F. and Lieberman, O., 2003, Asymptotic theory for multivariate GARCH processes, *Journal of Multivariate Analysis*, 84, 61-84

Ding, Z. and Engle, R., 2001, Large scale conditional covariance modelling, estimation and testing. *Academia Economic Papers*, 29, 157-184

Engle, R.F., 2002, Dynamic conditional correlation: a simple class of multivariate generalized autoregressive conditional heteroskedasticity models, *Journal of Business and Economic Statistics*, 20, 339-350.

Engle, R.F. and Kroner, K.F., 1995, Multivariate simultaneous generalized ARCH, *Econometric Theory*, 11, 122-150.

Engle, R.F., and Sheppard, K., 2001, Theoretical and Empirical Properties of Dynamic Conditional Correlation Multivariate GARCH, Working Paper 2001-15, University of California at San Diego.

Engle, R.F., Shephard, N., and Sheppard, K., 2008, Fitting vast dimensional time-varying covariance models, Oxford Financial Research Centre, Financial Economics Working Paper n. 30

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

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

Ling, S. and McAleer, M., 2003, Asymptotic theory for a vector ARMA-GARCH model, *Econometric Theory*, 19, 278-308

McAleer, M., 2005, Automated inference and learning in modeling financial volatility, *Econometric Theory*, 21, 232-261.

McAleer, M., Chan, F., Hoti, S., and Lieberman, O., 2008, Generalized autoregressive conditional correlation, *Econometric Theory*, 24-6, 1554-1583

Nelson, D.B., 1991, Conditional heteroskedasticity in asset pricing: a new approach, *Econometrica*, 59, 347-370.

Newey, W.K. and D. McFadden, 1994, Large Sample Estimation and Hypothesis Testing, in *Handbook of Econometrics*, Vol. 4, Elsevier North-Holland.

Nicholls, D.F. and Quinn, B.G., 1982, *Random Coefficient Autoregressive Models: An Introduction*, Springer-Verlag, New York.

Tsay, R.S., 1987, Conditional heteroscedastic time series models, *Journal of the American Statistical Association*, 82, 590-604.

APPENDIX: A review of the purported proofs of consistency and asymptotic normality of various DCC models.

(1) Engle and Sheppard (2001)

The authors assume, but do not verify, that the standard regularity conditions required for two step GMM to yield consistent and asymptotically normal estimators, as given, for example, in Newey and McFadden (1994), are satisfied for the DCC model. This ignores the fact that temporal dependence of correlations was not considered in Newey and McFadden (1994).

(2) Ding and Engle (2001)

The authors discuss estimation of various multivariate conditional covariance (correlation) models, without discussing their statistical properties.

Two diagnostic checks are presented, without establishing their statistical properties, and evaluates the tests using Monte Carlo simulations.

(3) Engle (2002)

The author assumes "reasonable regularity conditions" and "standard regularity conditions" (p. 342), without stating them, and refers to the theoretical results in Engle and Sheppard (2001) (see point 1) above).

(4) Cappiello, Engle and Sheppard (2006)

The authors develop an extension of the DCC model to incorporate asymmetries, but do not establish the asymptotic properties of the estimators.

(5) Aielli (2008)

The authors makes the following statements:

"We assume that QL regularities are satisfied. Basically, this requires correct specification and identification of the first two conditional moments …" (p.10)

"a vector of QML estimators, then, it is consistent by assumption of QL regularities." (p. 11)

Thus, the author assumes that typical regularity conditions are satisfied, without stating them.

(6) Engle, Shephard and Sheppard (2008)

The authors refer to Aielli's (2008) model and estimation method, but not to his purported proofs of consistency and asymptotic normality. Moreover, they do not refer to the main result in Aielli (2008), which is the inconsistency of QMLE for Engle's (2002) scalar DCC parameters. In addition, it seems they have used Aielli's model with targeting, which is impossible, by construction.

The authors purport to prove consistency and asymptotic normality. However, Theorem 1 for consistency is not a proof of consistency of the estimator of the appropriate parameter, while Theorem 2 for asymptotic normality assumes consistency of the estimator of the appropriate parameter (which was not proved in Theorem 1).

(7) McAleer et al (2008)

In comparison with the purported proofs of consistency and asymptotic normality of the QML estimators of the DCC parameters in the literature, McAleer et al. (2008) develop a generalized autoregressive conditional correlation (GARCC) model when the standardized residuals for each asset follow a multivariate random coefficient autoregressive (RCA) process. The scalar and diagonal versions of BEKK are also shown to be special cases of a multivariate RCA process.

As a multivariate generalization of Tsay's (1987) RCA model, GARCC provides a motivation for the conditional correlations to be time varying. Although GARCC is non-nested with respect to DCC model, the non-nestedness arising from different parametric restrictions that are imposed in the two models, special cases of GARCC are virtually identical to the scalar and Hadamard versions of DCC. The analytical forms of the sufficient conditions for the existence of moments are derived, and the sufficient conditions for the asymptotic properties of the QML estimators are established.

28

Table 1: Model dimension and asymptotic properties.

Model	Two step	Number of parameters			Consistency	Asymptotic Normality
		K assets	**K=10**	**K=100**		
Scalar BEKK without targeting	No	$0.5k(k+1)+2$	57	5052	Yes	Yes
Diagonal BEKK without targeting	No	$0.5k(k+1)+2k$	75	5250	Yes	Yes
BEKK without targeting	No	$0.5k(k+1)+2k^2$	255	25050	Yes	Yes
Hadamard BEKK without targeting	No	$1.5k(k+1)$	165	15150	Yes	Yes
Scalar BEKK with targeting	No	2	2	2	Yes	Yes
Diagonal BEKK with targeting	No	$2k$	20	200	Yes	Yes
BEKK with targeting	No	$2k^2$	200	20000	Yes	Yes
Hadamard BEKK with targeting	No	$k(k+1)$	110	10100	Yes	Yes
Scalar DCC with targeting	Yes	2	2	2	No	No
Diagonal DCC with targeting	Yes	$2k$	20	200	No	No
Generalized DCC with targeting	Yes	$2k^2$	200	20000	No	No
Hadamard DCC with targeting	Yes	$k(k+1)$	110	10100	No	No
Scalar DCC without targeting	Yes	$0.5k(k-1)+2$	47	4952	?	?
Diagonal DCC without targeting	Yes	$0.5k(k-1)+2k$	65	5150	?	?
Generalized DCC without targeting	Yes	$0.5k(k-1)+2k^2$	245	24950	?	?
Hadamard DCC without targeting	Yes	$0.5k(k-1)+k(k+1)$	155	15050	?	?

Note. For DCC models, the number of parameters does not include the univariate GARCH parameters, namely *3K* at a minimum, that are estimated for each asset. The number of parameters for DCC is higher if various asymmetric and leverage-based univariate models were to be used in conjunction with DCC.