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

EI 2010-13

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

Revised: January 2010

[1] The first author gratefully acknowledges financial support from Italian MUS Grant Cofin2006-13-1140. The second author wishes to thank the Australian Research Council, National Science Council, Taiwan, Institute of Economics, Academia Sinica, Taiwan, and Center for International Research on the Japanese Economy (CIRJE), Faculty of Economics, University of Tokyo, for financial support. This is an abridged and revised version of a paper entitled "Do we really need both BEKK and DCC? A tale of two covariance models".

[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

The management and monitoring of very large portfolios of financial assets are routine for many individuals and organizations. The two most widely used models of conditional covariances and correlations in the class of multivariate GARCH models are BEKK and DCC. It is well known that BEKK suffers from the archetypal "curse of dimensionality", whereas DCC does not. It is argued in this paper that this is a misleading interpretation of the suitability of the two models for use in practice. The primary purpose of this paper is to analyze the similarities and dissimilarities between BEKK and DCC, both with and without targeting, on the basis of the structural derivation of the models, the availability of analytical forms for the sufficient conditions for existence of moments, sufficient conditions for consistency and asymptotic normality of the appropriate estimators, and computational tractability for ultra large numbers of financial assets. Based on theoretical considerations, the paper sheds light on how to discriminate between BEKK and DCC in practical applications.

Keywords: Conditional correlations, conditional covariances, diagonal models, forecasting, generalized models, Hadamard models, scalar models, targeting.

JEL Classification: C32, G11, G17, G32.

1. Introduction

The management and monitoring of 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 Accord. 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 the consequent 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. First, we note that BEKK and DCC co-exist, despite one model being able to do virtually everything the other can do, thereby raising the pertinent question posed in the title of the paper. Second, we argue that BEKK is used to forecast conditional covariances, although it may also be used to forecast conditional correlations indirectly, while DCC is used to forecast conditional correlations only, while its structure could easily be applied to forecast conditional covariances. Third, the inherent differences between BEKK and DCC do not seem to be widely known. This is particularly relevant as DCC is equivalent to a targeted scalar BEKK model as applied to the variance standardized residuals, and can thereby be interpreted as a conditional correlation matrix only because of the standardization. Fourth, both the structural and statistical differences and similarities between the two models have not previously been analyzed in the literature.

With respect to the first question, we note that Engle and Kroner (1995) is a widely cited paper, but most citations would seem to be of a theoretical rather than empirical nature. The model is an archetypical example of over-parameterization, thereby leading to the moniker “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. It is argued in the paper that this is a misleading interpretation of the suitability of the two models to be used in practice.

A primary purpose of the paper is to shed some light on the similarities and differences between BEKK and DCC. The comparison commences from a theoretical perspective. A comparison of the two models considers several aspects which are generally associated with theoretical econometrics, but which are also fundamental in guaranteeing that the empirical applications, as well as their interpretation, are reliable. With this rationale, we first define targeting as an aid in estimating matrices associated with large numbers of financial assets, and then briefly discuss the use of targeting in estimating conditional covariance and correlation matrices in financial econometrics. We also consider the similarities and dissimilarities between BEKK and DCC, both with and without targeting; the analytical forms of the sufficient conditions for the existence of moments, sufficient conditions for consistency and asymptotic normality, computational tractability for ultra high numbers of financial assets, use of consistent two step estimation methods for the DCC model to enable it to be used sensibly in practical situations, and the determination of whether BEKK or DCC is to be preferred in empirical applications.

The remainder of the paper is organized as follows. Section 2 compares the BEKK and DCC specifications, defines the long run solution of conditional covariances (correlations), and defines the targeting of conditional covariance (correlation) models. Section 3 discusses the asymptotic results for BEKK and DCC. Some concluding comments are given in Section 4.

2. A Comparison of BEKK and DCC

This section evaluates directly comparable BEKK and DCC models which are feasible under large cross-sectional dimensions. Univariate and multivariate asymmetry and leverage, as well as the empirical comparison of the models, are not considered but are left for further research (see Caporin and McAleer (2008) for some results based on small scale models). For the same reason, we restrict the analysis to the original and simplest specifications given in Engle and Kroner (1995), and Engle (2002). Finally, 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 model specifications, and focus on mean innovations whose conditional covariance matrix is denoted by Σ_t. A short description of the models and the specifications considered are given in the Appendix.

Two definitions are given below in order to emphasize the approach taken in the paper:

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

For the Scalar BEKK model of Ding and Engle (2001), which is described in the Appendix, the unconditional covariance matrix is $E\left[\Sigma_t\right]=\bar{\Sigma}=CC'\left(1-\alpha-\beta\right)^{-1}$.

Two topics that are frequently discussed in the financial econometrics literature regarding covariance/correlation model estimation are the “curse of dimensionality” and “targeting”. The first issue is perceived as the most serious problem in covariance modelling, 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 that increase at an order greater 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 $O(k^2)$, the VECH model parameter number is of order $O(k^4)$, and the Generalized DCC model of Engle (2002) increases with order $O(k^2)$.

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 (2006), 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 regarding 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.

The "targeting" constraint is 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.

We define the "targeting" constraint as follows:

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.

Note that 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. Furthermore, 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).

Referring again to the Scalar BEKK model that is given in the Appendix, targeting leads to a specification where the intercept is given as $\bar{\Sigma}(1-\alpha-\beta)$. The model 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 conditional 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 still be computationally complicated for large k because the likelihood evaluation of the model in (2) requires the inversion of a covariance matrix of dimension k.

Although targeting can be computationally useful in terms of reducing, sometimes dramatically, the number of parameters to be estimated by maximum likelihood, 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 dynamic correlation literature, the resulting estimators will also be inconsistent.

Consider the BEKK model of Engle and Kroner (1995), with model orders set to 1. This model is exposed to the curse of dimensionality and is feasible for small cross-sectional dimensions, typically with fewer than 10 assets.

Although it is not necessary to do so, BEKK can be specified with targeting. The introduction of this feature requires some constraints to be imposed at the estimation step in order to guarantee that the covariance matrices are positive definite (for further details, see the Appendix). Fortunately, these constraints are extremely simple in the scalar case.

Focusing on the DCC model, we note it has been proposed directly with a targeting constraint, expressing the intercept as a function of the long run correlation. However, the most general specification of DCC without targeting is exposed to the curse of dimensionality, and has parameters with order $O(k^2)$, as in the BEKK model. Without targeting, DCC has the same problems as BEKK. Put differently, if targeting were to be included, the constraints required by DCC to ensure positive definiteness of the correlation matrix are identical to the constraints required by the Scalar BEK model.

In summary, with respect to computational complexity, when targeting is included, BEKK and DCC are equivalent; DCC has a structure equivalent to that of BEKK, and is a correlation model only because it includes a standardization. Finally, it should be noted that DCC is more flexible than BEKK because it models the conditional variances separately as a first step.

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 and VECH, 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.

Both DCC and BEKK require the imposition of parameter constraints to ensure covariance stationarity. The constraints are discussed in Engle and Kroner (1995), and are valid for the Generalized DCC model of Engle (2002). Constraints for the scalar representations have a very simple structure, are identical for targeted BEKK and DCC, and are closely related to the constraints needed to achieve a positive variance for BEKK and positive definiteness of the conditional covariance (correlation) matrices in the two models.

For BEKK, Jeantheau (1998) proved consistency 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. 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. Finally, Hafner and Preminger (2009) proved asymptotic normality of the VECH model (which nests BEKK) of Engle and Kroner (1995) under the existence of sixth-order moments.

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 (see the appendix) can be derived similarly by noting that Hadamard BEKK has a companion VECH representation with diagonal parameter matrices. The proofs in Jeantheau (1998), Comte and Lieberman (2003), and Hafner and Preminger (2009) can be generalized to include the BEKK representations where the long run solution of the model enters the intercept explicitly. In such cases, appropriate modifications of the regularity conditions are required. Therefore, the asymptotic theory for BEKK models has been established, albeit under untestable conditions.

3.1. Do Asymptotic Results Exist for DCC?

The primary appeal of the DCC specification, at least in its scalar incarnation, 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 the following three conditions hold: (i) the model can be targeted; (ii) the two step estimators are consistent; and (iii) the number of parameters increases as a power function of the cross-sectional dimension, with an exponent less 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 can be derived from the estimated parameters and the likelihood function. Furthermore, it ensures that the forecasts will not be influenced by parameter distortions. Finally, point (iii) controls for the parameters in the model dynamics. 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)) and 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) proved that the two step estimation of DCC models with targeting is inconsistent (see also Aielli (2009)). In fact, Aielli (2008) showed that the sample correlation estimator is an *inconsistent* estimator of the long run correlation appearing in the DCC intercept. As a result, the parameters driving the dynamics cannot be consistently estimated by Quasi Maximum Likelihood (QML), conditional on an inconsistent estimator of S. Therefore, 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 the consistency of the QML estimates of the other parameters, as well as their asymptotic distribution, thereby eliminating point (ii). Consequently, all the purported proofs for models with targeting, as presented 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.

Aielli (2008) suggested a correction to the DCC model to resolve the previous inconsistency between the 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 regularity conditions without actually stating them). It is worth mentioning that Aielli's (2008) model was used in Engle, Sheppard and Shephard (2008), under the assumption that it included targeting, which is not possible.

Aielli's (2008) results preclude the estimation of DCC with targeting, but this does not affect the DCC specifications without targeting. Hence, the asymptotic properties are still unknown. Clearly, despite the possibility of estimating DCC models in a single step, the curse of dimensionality will always be present as the intercept includes $0.5k(k-1)$ 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).

3.2 Consistent Estimation of Correlations for 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) showed that a theoretical relation could be derived to compare scalar DCC and BEKK with and without targeting. They suggested the derivation of conditional correlations from alternative BEKK representations, and referred to the derived model as Indirect DCC (which, despite its name, is not a different model but rather a bi-product of BEKK).

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

Theorem 1: *The indirect DCC conditional correlations derived from BEKK representations are consistent for the true conditional correlations.*

Proof: The conditional covariance matrix, Q_t, satisfies 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 BEKKs, 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 hence consistent.■

The theorem shows how BEKK may be used to obtain consistent estimates of the conditional correlation matrix. The BEKK model may also be 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 parameters may be calibrated at the corresponding parameters in a given BEKK model.

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

4. Concluding Remarks

The efficient management and monitoring of 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 in 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 Value-at-Risk (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, in the multivariate GARCH class. Although the two models are similar in many respects, the literature has not yet 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, namely: 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 both structural and statistical differences and similarities between the two models.

The primary purpose of the paper has been to examine these issues. For this purpose, we highlighted that BEKK possessed asymptotic properties under untestable moment conditions, whereas the asymptotic properties of DCC have simply been stated under a set of 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 BEKK could be used to obtain consistent estimates of dynamic conditional correlations, with a direct link to the Indirect DCC model suggested in Caporin and McAleer (2008).

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

References

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

Aielli, G.P., 2009, Dynamic conditional correlations: On properties and estimation, Available at SSRN: http://ssrn.com/abstract=1507743.

Billio, M., M. Caporin and M. Gobbo, 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.

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

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

Caporin, M. and M. McAleer, 2009, Do we really need both BEKK and DCC? A tale of two covariance models, Available at SSRN: http://ssrn.com/abstract=1338190.

Caporin, M., and P. Paruolo, 2009, Structured multivariate volatility models, Available at SSRN: http://ssrn.com/abstract=1318639.

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

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

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

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

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

Hafner, C.M., and A. Preminger, 2009, On asymptotic theory for multivariate GARCH models, Journal of Multivariate Analysis, 100, 2044-2054.

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

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

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

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 B.G. Quinn, 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.1. BEKK models

Engle and Kroner (1995) introduced the BEKK class of multivariate GARCH models. We will consider the simplest BEKK specification, which is standard, with all lag orders set to 1:

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

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. Engle and Kroner (1995) propose a more general representation than is given in (A.1), but it does not seem to have been used in empirical applications.

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 (A.1) are replaced by scalar coefficients ($A = \alpha I$ and $B = \beta I$, where I is an identity matrix of dimension k), whereas in the Diagonal BEKK version, A and B are diagonal matrices. A further representation of a BEKK-type model may be based on the Hadamard matrix product, as follows (we name it Hadamard BEKK):

$$\Sigma_t = CC' + A \circ \varepsilon_{t-1}\varepsilon_{t-1}{}' + B \circ \Sigma_{t-1}. \qquad \text{(A.2)}$$

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, (A.2) 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 the conditional covariance matrices is guaranteed, by construction (see Ding and Engle, 2001). Finally, we note that the diagonal specification is a restricted parameterization of the Hadamard BEKK model in equation (A.2).

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 equation may be redefined as follows:

$$\Sigma_t = \overline{\Sigma} + A\left(\varepsilon_{t-1}\varepsilon_{t-1}{}' - \overline{\Sigma}\right)A' + B\left(\Sigma_{t-1} - \overline{\Sigma}\right)B' \qquad \text{(A.3)}$$

Similar representations could be obtained for restricted BEKK models and for the Hadamard BEKK version. We can easily check that the model in (A.3) gives $E\left[\Sigma_t\right]=\overline{\Sigma}$, as $E\left[\varepsilon_{t-1}\varepsilon_{t-1}{}'\right]=\overline{\Sigma}$ and $E\left[\Sigma_{t-1}\right]=\overline{\Sigma}$. Note that (A.3) allows the introduction of targeting by replacing $\overline{\Sigma}$ with a sample estimator. Positive definiteness of the conditional covariance matrices must be imposed at the estimation step by constraining the model intercept; otherwise the estimates cannot be interpreted as covariance matrices. In (A.3) positive definiteness of the conditional covariance matrices is guaranteed by imposing positive definiteness of $\overline{\Sigma} - A\overline{\Sigma}A' - B\overline{\Sigma}B'$.

Although the constraints may seem to be quite simple for heavily restricted models, their computational complexity is entirely relevant, in particular, when the cross-sectional dimension is simply moderate rather than high. In fact, imposing positive definiteness of the intercepts results in a set of highly non-linear constraints on the parameters. In addition, it should be stressed that covariance stationarity constraints need to be taken into account. These constraints are extremely simple in the Scalar BEKK case, and collapse to $\alpha^2+\beta^2<1$.

A.2.. 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). The covariance matrix is decomposed as follows:

$$\Sigma_t = D_t R_t D_t \tag{A.4}$$

$$D_t = diag\left(\sigma_{1,t}, \sigma_{2,t}, ...\sigma_{k,t}\right) \tag{A.5}$$

$$R_t = \bar{Q}_t^{-½} Q_t \bar{Q}_t^{-½}, \qquad \bar{Q}_t = dg\left(Q_t\right) \tag{A.6}$$

where D_t includes the conditional volatilities, which are modelled as 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 feasibility in large cross-sectional dimensions.

Several specifications have been suggested for Q_t. The DCC model (or Hadamard DCC) is given in Engle (2002) as:

$$Q_t = S + A \circ \left(D_{t-1}^{½} \varepsilon_{t-1} \varepsilon_{t-1}' D_{t-1}^{½} - S\right) + B \circ \left(Q_{t-1} - S\right), \tag{A.7}$$

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. The number of parameters in this model is of order $O(k^2)$, such that it 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 (A.7) is counterintuitive since Q_t is 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 methods, conditional on the value assigned to S. Note that the model requires appropriate constraints for covariance stationarity and positive definiteness of Q_t. 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.

An alternative fully parameterized model, the Generalized DCC (GDCC) specification, is given in Cappiello, Engle and Sheppard (2006). The dynamic equation driving the conditional correlation matrix is:

$$Q_t = S + A\left(D_{t-1}^{½}\varepsilon_{t-1}\varepsilon_{t-1}{}'D_{t-1}^{½} - S\right)A' + B\left(Q_{t-1} - S\right)B', \qquad \text{(A.8)}$$

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

Two major restricted specifications may be considered, namely the diagonal and scalar models. The most frequently estimated version of the DCC model is what we will call the scalar DCC model, where A=a**ii'**, B=b**ii'**, and **i** is a vector of ones. Note that the DCC models can be represented without targeting, but this will require the joint estimation of all the parameters, including the long run correlations. In the scalar DCC model, the constraints for covariance stationarity and positive definiteness collapse to $a>0$, $b>0$, and $a+b<1$. These are observationally equivalent to the constraints for the Scalar BEKK model as the DCC model can be considered as a Scalar BEKK model.