What They Did Not Tell You About Algebraic (Non-)Existence, Mathematical (IR-)Regularity and (Non-)Asymptotic Properties of the Full BEKK Dynamic Conditional Covariance Model*

Michael McAleer**

Department of Finance
Asia University
Taiwan
and
Discipline of Business Analytics
University of Sydney Business School
Australia
and
Econometric Institute
Erasmus School of Economics
Erasmus University Rotterdam
The Netherlands
and
Department of Economic Analysis and ICAE
Complutense University of Madrid
Spain
and
Institute of Advanced Sciences
Yokohama National University
Japan

Abstract

Persistently high negative covariances between risky assets and hedging instruments are intended to mitigate against risk and subsequent financial losses. In the event of having more than one hedging instrument, multivariate covariances need to be calculated. Optimal hedge ratios are unlikely to remain constant using high frequency data, so it is essential to specify dynamic covariance models. These values can either be determined analytically or numerically on the basis of highly advanced computer simulations. Analytical developments are occasionally promulgated for multivariate conditional volatility models. The primary purpose of the paper is to analyse purported analytical developments for the most widely-used multivariate dynamic conditional covariance model to have been developed to date, namely the Full BEKK model of Baba et al. (1985), which was published as Engle and Kroner (1995). Dynamic models are not straightforward (or even possible) to translate in terms of the algebraic existence, underlying stochastic processes, specification, mathematical regularity conditions, and asymptotic properties of consistency and asymptotic normality, or the lack thereof. The paper presents a critical analysis, discussion, evaluation and presentation of caveats relating to the Full BEKK model, and an emphasis on the numerous dos and don'ts in implementing Full BEKK in practice.

Keywords: Hedging, covariances, existence, mathematical regularity, inevitability, likelihood function, statistical asymptotic properties, caveats, practical implementation.
JEL: C22, C32, C51, C52, C58, C62, G32.

1. Introduction

Persistently high negative covariances between risky assets and hedging instruments are intended to mitigate against risk and subsequent financial losses. It is possible to hedge against risky assets using one or more hedging instruments as the benchmark, which requires the calculation of multivariate covariances. As optimal hedge ratios are unlikely to remain constant using high frequency data, it is essential to specify dynamic models of covariances.

Modelling, forecasting and evaluating dynamic covariances between hedging instrument and risky financial assets requires the specification and estimation of multivariate models of covariances. These values can either be determined analytically or numerically on the basis of highly advanced computer simulations. High frequency time periods such daily data can lead to models of conditional volatility, where analytical developments are occasionally promulgated.

The primary purpose of the paper is to analyse purported analytical developments for the most widely-used multivariate dynamic conditional covariance model to have been developed to date, namely the Full BEKK model of Baba et al. (1985), which was published as Engle and Kroner (1995).

Dynamic models are not straightforward (or even possible) to translate in terms of the algebraic existence, underlying stochastic processes, specification, mathematical regularity conditions, and asymptotic properties of consistency and asymptotic normality, or the lack thereof. The paper presents a critical analysis, discussion, evaluation and presentation of caveats relating to the Full BEKK model, and an emphasis on the numerous dos and don'ts in implementing Full BEKK in practice.

For the variety of detailed possible outcomes mentioned above, where problematic issues arise constantly, and sometimes unexpectedly, a companion paper by the author evaluates the recent developments in modelling dynamic conditional correlations on the basis of the Dynamic Conditional correlation (DCC) model (see McAleer (2019)).

The remainder of the paper is as follows. The Full BEKK model is presented in Section 2, which will enable a subsequent critical analysis and emphasis on a discussion, evaluation and presentation of caveats in Section 3 of the numerous dos and don'ts in implementing the Full BEKK model in practice.

2. Model Specification

Some of the results in this section, though not all, are available in the extant literature, but the interpretation of the models and their non-existent underlying stochastic processes, as well as the discussions and caveats in the following section, are not available. Much of the basic material relating to the univariate and multivariate specifications in Sections 2.1 - 2.3 overlap with the presentation in McAleer (2019).

The first step in estimating Full BEKK is to estimate the standardized shocks from the univariate conditional mean returns shocks. The most widely used univariate conditional volatility model, namely GARCH, will be presented briefly, followed by Full BEKK. Consider the conditional mean of financial returns, as follows:

$$y_t = E(y_t | I_{t-1}) + \varepsilon_t \tag{1}$$

where the returns, $y_t = \Delta \log P_t$, represents the log-difference in financial asset prices (P_t), I_{t-1} is the information set at time t-1, and ε_t is a conditionally heteroskedastic returns shock that has the same unit of measurement as the returns. In order to derive conditional volatility specifications, it is necessary to specify, wherever possible, the stochastic processes underlying the returns shocks, ε_t.

2.1   Univariate Conditional Volatility Models

Univariate conditional volatilities can be used to standardize the conditional covariances in alternative multivariate conditional volatility models to estimate conditional correlations, which are particularly useful in developing dynamic hedging strategies. The most widely-used

univariate model, GARCH, is presented below as an illustration because the focus of the paper is on estimating and testing Full BEKK.

2.2   Random Coefficient Autoregressive Process and GARCH

Consider the random coefficient autoregressive process of order one:

$$\varepsilon_t = \phi_t \varepsilon_{t-1} + \eta_t \tag{2}$$

where

$\phi_t \sim iid(0, \alpha),$
$\eta_t \sim iid(0, \omega),$
and $\eta_t = \varepsilon_t / \sqrt{h_t}$ is the standardized residual.

The standardized residual is unit-free of measurement, and is a financial fundamental as it represents a riskless asset.

Tsay (1987) derived the ARCH(1) model of Engle (1982) from equation (1) as:

$$h_t = E(\varepsilon_t^2 | I_{t-1}) = \omega + \alpha \varepsilon_{t-1}^2 \tag{3}$$

where h_t is conditional volatility, and I_{t-1} is the information set available at time t-1. The mathematical regularity condition of invertibility is used to relate the conditional variance, h_t , in equation (3) to the returns shocks, ε_t , which has the same measurement as y_t in equation (1), thereby yielding a valid likelihood function of the parameters given the data.

The use of an infinite lag length for the random coefficient autoregressive process in equation (2), with appropriate geometric restrictions (or stability conditions) on the random coefficients, leads to the GARCH model of Bollerslev (1986). From the specification of equation (2), it is clear that both ω and α should be positive as they are the unconditional variances of two independent stochastic processes. The GARCH model is given as:

$$h_t = E(\varepsilon_t^2 | I_{t-1}) = \omega + \alpha\varepsilon_{t-1}^2 + \beta h_{t-1}$$

where α is the short run ARCH effect, and β, which lies in the range (-1,1), is the GARCH contribution to the long run persistence of returns shocks.

2.3   Multivariate Conditional Volatility Models

Multivariate conditional volatility GARCH models are often used to analyze the interaction between the second moments of returns shocks to a portfolio of assets, and can model and the possible risk transmission or spillovers among different assets.

In order to establish volatility spillovers in a multivariate framework, it is useful to define the multivariate extension of the relationship between the returns shocks and the standardized residuals, that is, $\eta_t = \varepsilon_t / \sqrt{h_t}$. The multivariate extension of equation (1), namely:

$$y_t = E(y_t | I_{t-1}) + \varepsilon_t$$

can remain unchanged by assuming that the three components in the above equation are now $m \times 1$ vectors, where m is the number of financial assets.

The multivariate definition of the relationship between ε_t and η_t is given as:

$$\varepsilon_t = D_t^{1/2}\eta_t \tag{4}$$

where $D_t = diag(h_{1t}, h_{2t}, \dots, h_{mt})$ is a diagonal matrix comprising the univariate conditional volatilities. Define the conditional covariance matrix of ε_t as Q_t. As the $m \times 1$ vector, η_t, is assumed to be *iid* for all m elements, the conditional correlation matrix of ε_t, which is equivalent to the conditional correlation matrix of η_t, is given by Γ_t.

Therefore, the conditional expectation of the process in equation (4) is defined as:

$$Q_t = D_t^{1/2}\Gamma_t D_t^{1/2} \tag{5}$$

Equivalently, the conditional correlation matrix, Γ_t , can be defined as:

$$\Gamma_t = D_t^{-1/2} Q_t D_t^{-1/2} \tag{6}$$

Equation (5) is useful if a model of Γ_t is available for purposes of estimating the conditional covariance matrix, Q_t, whereas equation (6) is useful if a model of Q_t is available for purposes of estimating the conditional correlation matrix, Γ_t.

Both equations (5) and (6) are instructive for a discussion of asymptotic properties. As the elements of D_t are consistent and asymptotically normal, the consistency of Q_t in equation (5) depends on consistent estimation of Γ_t , whereas the consistency of Γ_t in equation (6) depends on consistent estimation of Q_t. As both Q_t and Γ_t are products of matrices, and the inverse of the matrix D is not asymptotically normal, even when D is asymptotically normal, neither the QMLE of Q_t nor Γ_t will be asymptotically normal, especially based on the definitions that relate the conditional covariances and conditional correlations given in equations (5) and (6).

2.4   Full BEKK and Diagonal BEKK Models

The vector random coefficient autoregressive process of order one is the multivariate extension of equation (2), and is given as:

$$\varepsilon_t = \Phi_t \varepsilon_{t-1} + \eta_t \tag{7}$$

where

ε_t and η_t are $m \times 1$ vectors,

Φ_t is an $m \times m$ matrix of random coefficients,

$\Phi_t \sim iid(0, A)$,

$\eta_t \sim iid(0, QQ')$.

Technically, a vectorization of a full (that is, non-diagonal) matrix A to *vec A* can have dimension as high as $m^2 \times m^2$, whereas a half-vectorization of a symmetric matrix A to *vech*

7

A can have dimension as low as $m(m-1)/2 \times m(m-1)/2$. The matrix A is crucial in the interpretation of symmetric and asymmetric weights attached to the returns shocks.

As the dimension of the unconditional variance of ε_t in equation (7) is m, if the variance matrix is not restricted parametrically, the dynamic conditional covariance matrix of (7) would depend on the product of the variance of Φ_t, with a dimension that lies between $m(m-1)/2$ and m^2, neither of which would be conformable with the dimension of ε_{t-1}.

In cases where A is either a diagonal matrix, or the special case of a scalar matrix, $A = aI_m$, McAleer et al. (2008) showed that the multivariate extension of GARCH(*1,1*) from equation (7), incorporating an infinite geometric lag in terms of the returns shocks, is given as the diagonal (or scalar) BEKK model, namely:

$$Q_t = QQ' + A\varepsilon_{t-1}\varepsilon'_{t-1}A' + BQ_{t-1}B' \tag{8}$$

where A and B are diagonal (or scalar) matrices.

Although the Full BEKK model is always presented in the form of equation (8), with A and B given as full rather than diagonal matrices, as stated above, the specification is not consistent with equation (8) as the matrices A and B for Full BEKK would have dimensions that lie between $m(m-1)/2$ and m^2. To state the obvious, the dimensions of A and B would not be conformable for multiplication with the $m \times 1$ vector, ε_{t-1}, and the $m \times m$ matrix, Q_{t-1}, respectively.

McAleer et al. (2008) showed that the QMLE of the parameters of the Diagonal BEKK models are consistent and asymptotically normal, so that standard statistical inference on testing hypotheses is valid. Moreover, as Q_t in equation (8) can be estimated consistently, Γ_t in equation (6) can also be estimated consistently. However, as explained above, asymptotic normality cannot be proved given the definitions in equations (5) and (6).

It should be emphasized that the QMLE of the parameters in the conditional means and the conditional variances for univariate GARCH, Diagonal BEKK and Full BEKK will differ as the multivariate models are estimated jointly, whereas the univariate models are estimated

individually. The QMLE of the parameters of the conditional means and the conditional variances of Diagonal BEKK and Full BEKK will differ as Diagonal BEKK imposes parametric restrictions on the off-diagonal terms of the conditional covariance matrix of the Full BEKK model.

3. Discussion and Caveats of Dos and Don'ts Regarding Full BEKK

The results in the previous section allow a clear discussion of the caveats associated with **Full BEKK**. The deficiencies and limitations in virtually all published papers that use the deeply flawed **Full BEKK** model are given below. The discussion and caveats are presented in a clear and entirely straightforward manner that need no further elaboration.

(1) Engle (1982) developed an autoregressive model of conditional correlations, **ARCH**, based on the conditional returns shocks.

(2) Bollerslev (1986) extended **ARCH** by adding a lagged dependent variable to obtain **Generalized ARCH**, **GARCH**.

(3) The **GARCH(1,1)** parameters must satisfy the regularity conditions of positivity as they are the unconditional variances from a univariate random coefficient autoregressive process (see Tsay, 1987; McAleer, 2014).

(4) However, the coefficient of the arbitrary lagged conditional variance is a positive or negative fraction (see Bollerslev (1986)).

(5) The **FULL BEKK** model was proposed in Baba, Engle, Kraft and Kroner (1985), after which the model is named.

(6) The BEKK model was published ten years later in Engle and Kroner (1995).

(7) The **Full BEKK** model does not satisfy the definition of a conditional covariance matrix, as the purported conditional covariances do not satisfy the definition of a covariance, except by an untenable assumption.

(8) There is no known underlying stochastic process that leads to the ***Full BEKK*** model, so that there are no regularity conditions relating to its specification.

(9) The regularity conditions include invertibility, which is essential in relating the ***iid*** standardized residuals to the returns data.

(10)   As invertibility does not hold, it follows that there is no likelihood function.

(11)   Consequently, there are no derivatives that would enable the derivation of asymptotic properties for the ***Quasi-Maximum Likelihood Estimates (QMLE)*** of the estimated parameters.

(12)   Therefore, any statements regarding the purported "statistical significance" are meaningless and lack statistical validity (see Chang and McAleer (2019) for a critical analysis).

(13)   It follows that any empirical results based on the ***Full BEKK*** estimates are fatally flawed and lack statistical validity.

(14)   As ***Full BEKK*** does not satisfy appropriate mathematical regularity conditions, the ***QMLE*** do not possess asymptotic properties.

(15)   The only exceptions to the non-existence of asymptotic properties of the QMLE of ***Full BEKK*** are under highly restrictive and untestable assumptions (see Chang and McAleer (2018), Comte and Lieberman (2003), Hafner and Preminger (2009), and McAleer et al. (2008)).

(16)   The results in Tsay (1987) were extended to a vector random coefficient stochastic process to derive ***Diagonal BEKK*** in McAleer et al. (2008):

(17)   McAleer et al. (2008) demonstrate that the ***Diagonal BEKK*** model has a stochastic process that leads to its specification, and hence satisfies the regularity conditions, including invertibility.

(18)　　Consequently, the **QMLE** of the estimated parameters of **Diagonal BEKK** are consistent and asymptotic normal.

(19)　　It should come as little or no surprise that, when the Full BEKK model is estimated using real data, there are always difficulties in terms of computational convergence, especially when $m > 4$, and the fact that the model does not actually exist!

(20)　　Such computational outcomes would almost certainly arise from the addition of between $m(m-1)/2$ and m^2 parameters when $p = 1$, especially when the value of m is high for the large financial portfolios that are observed in practice.

(21)　　In short, **Diagonal BEKK** is mathematically and statistically preferable to the fatally flawed **Full BEKK**.

(22)　　If **Full BEKK** is to be considered at all, except in connection with the algebraic non-existence, absence of an underlying stochastic process, mathematical irregularity, and unknown asymptotic statistical properties, or alternatively, in the presence of problems that should be avoided at all costs, it is advisable that the Full BEKK specification be used with extreme and utter caution in empirical practice.

References

Baba, Y., R. Engle, D. Kraft and K. Kroner (1985), Multivariate simultaneous generalized ARCH, Unpublished Paper, University of California, San Diego. [Published as Engle and Kroner (1995)]

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

Chang, C.-L. and M. McAleer (2019), The fiction of full BEKK: Pricing fossil fuels and carbon emissions, Finance Research Letters, 28, 11-19.

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

Engle, R.F. (1982), Autoregressive conditional heteroskedasticity, with estimates of the variance of United Kingdom inflation, Econornetrica, 50, 987-1007.

Engle, R.F. and K.F. Kroner (1995), Multivariate simultaneous generalized ARCH, Econometric Theory, 11(1), 122-150.

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

McAleer, M. (2014), Asymmetry and leverage in conditional volatility models, Econometrics, 2(3), 145-150.

McAleer, M. (2019), What they did not tell you about algebraic (non-)existence, mathematical (ir-)regularity and (non-)asymptotic properties of the dynamic conditional correlation (DCC) model, unpublished paper, Department of Finance, Asia University, Taiwan.

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

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