Sequential Conditional Correlations: Inference and Evaluation [*]

Alessandro Palandri[†]

February 2005

FIRST DRAFT

Abstract

This paper presents a new approach to the modeling of the conditional correlation matrix in the multivariate GARCH framework. I will show how to break it into the product of a sequence of matrices with desirable characteristics. This feature will allow for a multi step estimation procedure, thus converting a highly dimensional and intractable optimization problem into a series of simple and feasible estimations. On the wake of this simplicity it is possible to employ richer parameterizations and more complex functional forms for the single components.

I will provide an empirical application studying the conditional second moments of 69 selected stock returns from the NASDAQ100.

Keywords: Multivariate GARCH, Conditional Correlations, High Dimensional GARCH Models, Sequential Estimation, Climber.

JEL classification: C51, C52, C61, G1.

---

[*] I thank my advisor George Tauchen and my committee members Tim Bollerslev, Bjørn Eraker, Ron Gallant, and Barbara Rossi for their guidance, constant moral support, and precious comments. I also thank all the participants to the Duke Financial Econometrics Group for their comments and suggestions.

[†] Department of Economics, Duke University, Box 90097, Durham NC 27708, E-mail:palandri@econ.duke.edu.

1 Introduction

Modeling the temporal dependence in the second order moments and forecasting future volatility has key relevance in many financial-econometric issues such as evaluating risk and derivatives pricing. Because it is relatively simple to estimate and because it seems to track volatility in a satisfactory way, the GARCH(1,1)[1] model has become the most popular specification among theorists and practitioners. Andersen and Bollerslev (1998) showed how these models do *provide strikingly accurate volatility forecasts* . Soon attention turned to the fact that volatilities of assets and markets tend to move together in time. This widely accepted feature of the data has led to the multivariate modeling of volatilities.

In the last years the number of different multivariate GARCH, henceforth M GARCH, models has grown steadily. In each new model there was the effort of the authors to overcome one, and possibly more, undesirable feature of the models present in the literature at the time. However, the trade-off faced by every model is the one between parameters' parsimony and richness in the description of the second order dynamics. In fact, an M-variate model will comprise $M(M+1)/2$ equations each one of which will contain a set of parameters describing the evolution of the dependent variable. The number of parameters of a fairly rich volatility model soon becomes big enough, in some cases even for $M < 6$, to render the estimation infeasible.

The main goal of this paper is to push further the limit determined by the dimensions that can be handled by an MGARCH model without giving up on rich parameterizations of the conditional variances and correlations. This will be achieved by extending the mainstream approach of a multi-step estimation procedure. In this regard, I propose an MGARCH model that allows for the separate estimation of the conditional correlations in an internally consistent manner such that the resulting conditional correlation matrix will always be positive semidefinite.

---

[1] This is true for GARCH models of all flavors such as the GARCH, TS-GARCH, Log-GARCH, IGARCH, AGARCH, EGARCH, NGARCH, A-GARCH, QARCH, STARCH, GJR-GARCH, APARCH, NAGARCH, V-GARCH, ZARCH, SWARCH, H-GARCH, FIGARCH, FIE-GARCH, Aug-GARCH, etc.

2 MGARCH Literature

In general, for an M-dimensional vector stochastic process y_t, the location can be described by:

$$y_t = \mu_t(\theta) + u_t \tag{1}$$

where $\mu_t(\theta)$ is a function, of the vector of parameters θ, describing the evolution of the mean vector conditional to I_{t-1}. The scale will be described as:

$$u_t = H_t^{1/2}(\theta) \cdot \varepsilon_t \tag{2}$$

where $H_t^{1/2}(\theta)$ is a positive definite matrix and ε_t satisfies the following moment conditions:

$$E(\varepsilon_t) = 0 \tag{3}$$

$$VAR(\varepsilon_t) = I_M \tag{4}$$

$H_t^{1/2}$ is such that H_t is the conditional variance-covariance matrix of the process y_t, while its unconditional variance-covariance matrix is given by $\Sigma = E(H)$.

The way in which to specify the functions $\mu_t(\theta)$ and $H_t(\theta)$ depends on the model we are interested in. If a discussion on the most appropriate model for the location is beyond the scope of this paper, the specification of the conditional variance covariance matrix H_t is the core.

The VEC specification of Bollerslev, Engle, and Wooldridge (1988) is one of the least restrictive in the sense that it allows each element of the conditional variance-covariance matrix to be function of every element of the lagged conditional variance-covariance matrices and outer products of lagged realizations. The major drawbacks of this specification are the difficulty to guarantee the positivity of H_t without imposing strong parameter restrictions and the rate at which the number of parameters grows with the dimensionality of the model (M^4).

The R-GARCH of Gallant and Tauchen (2002) solves one of the VEC's problems by guaranteeing the positive definiteness of H_t. This is achieved by modeling its Cholesky decomposition[2] rather than H_t itself. This model solves the positivity problem without sacrificing the richness of the VEC's dynamics which in turn cause it to share the dimensionality problem of the latter (M^4).

Proposed by Engle and Kroner (1995), the BEKK guarantees the positivity of H_t and significantly reduces the number of parameters to be estimated (M^2). It is clear

---

[2] $H_t = R_t R_t'$ and R_t is modeled by a VEC.

that unless some sort of *magic* is introduced the dimensionality of the model cannot be further decreased as the *intercept* itself is high dimensional (M^2). Luckily, the *magic* required comes in the form of targeting (Engle and Mezrich(1996)) which not only locks the model's unconditional variance to the sample counter-part but eliminates $M(M+1)/2$ parameters from the model.

Diebold and Nerlove (1986) and Engle, Ng, and Rothschild (1990) exploit the idea that co-movements of the data are driven by a small number of common underlying *factors*. With variance targeting, the number of parameters of an F-GARCH that need to be estimated drops down considerably (M). This class of models requires the *extraction* of the factors and therefore the estimation of an M-dimensional vector of loadings per factor. Of the same dimensions, when targeting, is the diagonal-BEKK (M), where the matrices of coefficients are not full but diagonal.

A particular type of F-GARCH is the O-GARCH (orthogonal GARCH) of Alexander and Chibumba (1997) and Alexander (2001) where univariate GARCH processes are fit to a transformation of the data[3]. A criticism to that has been raised to this model is that the use of $m < M$ factors implies a reduced rank H_t. To avoid this problem, one could set the eigenvalues that do not correspond to factors to their unconditional mean[4] rather than to zero. This approach does not suffer from any curse of dimensionality as it only requires the estimation of M univariate GARCH processes. A not very appealing feature of this model is that conditional variances and covariances are modeled as constant linear combinations of univariate GARCH processes. Furthermore, the parameters of these linear combinations are not estimated but determined by the diagonalization of the unconditional variance-covariance matrix.

Patton (2000) and Jondeau and Rockinger (2001) propose copula-MGARCH models in which univariate GARCH processes describe the conditional variances, and a copula function joins the marginal distributions to form a multivariate distribution function. In order to provide a richer parametrization than that of Constant Conditional Correlation model (Bollerslev (1990)), the parameters of the conditional copula must be time-varying. Chen and Fan (2003) introduced a new class of semiparametric copulas in which a parametric copula is evaluated at a nonparametric marginal distribution. Furthermore, they move away from the *traditional* bivariate case and

---

[3] Let y_t be the vector of observables and Σ its unconditional variance. From the diagonalization of the latter: $\Sigma = C\Lambda C'$. Let the transformed data be: $\varepsilon_t = C'y_t$.

[4] In other words: keep the values that have been determined by the diagonalization of the unconditional variance-covariance matrix.

estimate a 10-dimensional copula.

The Flexible-GARCH of Ledoit, Santa-Clara and Wolf (2002) first estimates univariate GARCH models for the volatilities and then moves on to estimate the conditional covariances one at the time using bivariate models. Since the so obtained coefficients are generally incompatible with each other, [5] they will then be *transformed* to guarantee positive definiteness of the conditional variance matrix. This last step is achieved by finding coefficients that guarantee positivity and that are close, with respect to a certain norm, [6] to the estimated coefficients. The appealing feature of this model is that the curse of dimensionality is completely avoided in the estimation phase. Nevertheless, it somehow reappears in the minimization of the norm with respect to the coefficients which has to be done numerically. Furthermore, while the incompatible parameters are estimated free of constraints the minimization of the norm *does* put constraints[7] that are not evident as in the BEKK but rather quite *obscure.*

2.1 Conditional Correlations Models

Introduced by Bollerslev (1990), the Conditional Correlations models, henceforth CC, provide dynamics for the variances and the correlations rather than for the variances and covariances. Furthermore, CC models constrain every conditional variance $h_{i,t}$ to be function of its own lags $h_{i,t-j}$ and realizations $\varepsilon^2_{i,t-j}$, which means that the volatilities can be described by M univariate GARCH processes.

The Constant Conditional Correlation model or CCC, introduced by Bollerslev (1990) defines the conditional variance-covariance matrix in the following way:

$$H_t = D_t R D_t \tag{5}$$

where D_t is a diagonal matrix whose squared elements follow univariate GARCH(p,q) processes[8], and R is the constant correlation matrix with $R_{ii} = 1$, $\forall i$. Empirically, the parsimonious parametrization of the variances[9] leads to less noisy parameters' estimates and improved forecasts. The only critique that has been moved to this

---

[5] In the sense that they do not yield conditional variance-covariance matrices that are positive definite.

[6] The Authors have used the Frobenius norm.

[7] This must be the case as the resulting parameters do guarantee positive definiteness of the conditional variance.

[8] In general the univariate GARCHes can have different number of lags and follow different specifications.

[9] Parsimonious in the sense that there are no interactions among the volatilities.

approach is that the assumption of constant conditional correlations may seem unrealistic in many empirical applications.

Engle (2001) proposed a generalization of the CCC model by providing dynamics for the conditional correlation matrix:

$$H_t = D_t R_t D_t \tag{6}$$

However, in this early setup, the evolution of all the correlations is governed by only 2 parameters not allowing for asset specific news impact nor asymmetries. Engle and Sheppard (2002) eliminate this drawback of the first DCC by using a diagonal BEKK to model the conditional correlations. With *correlation targeting*, the total number of parameters to be estimated is of order M. Furthermore, the DCC allows for a two-stage estimation of H_t: in the first stage univariate GARCHes are estimated and used to construct the matrix D_t; in the second stage the returns, standardized by the GARCH variances, are used to estimate the parameters of the conditional correlation model which produces estimates of R_t.

The most parsimonious model of this class that allows for asset specific news impact and asymmetries requires the estimation of $3M$In a high-dimensional framework this model will be very hard to estimate if not infeasible.

3 Sequential Conditional Correlations

This model fully exploits the idea of separating the various components of the conditional variance covariance matrix put forward in the Conditional Correlations models. While both CCC and DCC only separate the variances from the correlations, therefore allowing only for a *two*-step estimation procedure, it is possible to further separate correlations from partial correlations thus allowing for a truly *multi*-step procedure. The idea is to re-write the conditional variance covariance matrix in the following form:

$$H_t = D_t K_{1,2,t} K_{1,3,t} \ldots K_{M-1,M,t} K'_{M-1,M,t} \ldots K'_{1,3,t} K'_{1,2,t} D_t$$

or in a more compact manner as:

$$H_t = D_t \Big(\prod_{i=1}^{M-1} \prod_{j=i+1}^{M} K_{i,j,t} \Big) \Big(\prod_{i=1}^{M-1} \prod_{j=i+1}^{M} K_{i,j,t} \Big)' D_t$$

The matrices $K_{i,j,t}$ are lower triangular and their generic element (row, col) is given by:

$$\mathbf{K_{i,j,t}}[\mathbf{row}, \mathbf{col}] = \begin{cases} \rho_{i,j,t} & \text{if } row = j \text{ and } col = i \\ (1 - \rho_{i,j,t}^2)^{1/2} & \text{if } row = j \text{ and } col = j \\ \mathbf{I}[row, col] & \text{otherwise} \end{cases}$$

In the Appendix I prove that every correlation matrix allows for such a decomposition. Furthermore, I show that given a sequence of matrices $K_{i,j,t}$, with the only constraint that $|\rho_{i,j,t}| < 1$, their product is a correlation matrix.

For exposition purposes let's assume that the elements in D_t are known:

$$\varepsilon_t = D_t^{-1/2} u_t$$

If $\rho_{1,2,t}$ measured the conditional correlation between $\varepsilon_{1,t}$ and $\varepsilon_{2,t}$, then:

$$\varepsilon_{1,2,t} = K_{1,2,t}^{-1/2} \varepsilon_t$$

Now, the conditional correlation between the first and second element of $\varepsilon_{1,2,t}$ is zero. In the same way we can think of $\rho_{1,3,t}$ as being the conditional correlation between $\varepsilon_{1,2,t,}[1]$ and $\varepsilon_{1,2,t,}[3]$:

$$\varepsilon_{1,3,t} = K_{1,3,t}^{-1/2} \varepsilon_{1,2,t}$$

Proceeding in the same fashion for all the couples $(1,4),(1,5),\dots(1,M-1),(1,M)$ we will obtain a vector $\varepsilon_{1,M,t}$ where the conditional correlation between the first element and all the others is zero:

$$E_{t-1}[\varepsilon_{1,M,t}\varepsilon'_{1,M,t}] = \begin{pmatrix} 1 & 0 & 0 & 0 & \dots & 0 \\ 0 & 1 & \# & \# & \dots & \# \\ 0 & \# & 1 & \# & \dots & \# \\ 0 & \# & \# & 1 & \dots & \# \\ \vdots & \vdots & \vdots & & \ddots & \\ 0 & \# & \# & \# & \dots & 1 \end{pmatrix}$$

where the signs # indicate the conditional correlations among the *standardized* series.

This particular decomposition allows for a multi-step estimation of H_t that guarantees its positive definiteness therefore translating a very complex estimation into a sequence of simple estimates. The first step of the procedure will be analogous to that of CCC and DCC: estimation of univariate GARCH models for each of the M series and their standardization:

$$\varepsilon_t = \widehat{D}_t^{-1/2} u_t$$

The second step involves the estimation of the conditional correlation matrix. This is achieved by estimating the conditional correlations $\widehat{\rho}_{i,j,t}$ for a pair of returns $y_{i,t}, y_{j,t}$ followed by their transformation in $y_{i,t}, (1-\widehat{\rho}^2_{i,j,t})^{-1/2}(y_{j,t} - \widehat{\rho}_{i,j,t} y_{i,t})$. Iterating until all pairwise correlations have been estimated will yield the sequence of values of ρ necessary to *re-construct* the whole conditional correlation matrix R_t.

4 A Model for Conditional Correlations

The proposed univariate model, to be used in the SCC steps, is a model for correlations rather than a model for the corresponding variance-covariance matrix. instead of using the product $\varepsilon_{i,t-1}\varepsilon_{j,t-1}$ as a measure of the realized correlation at time $t-1$, I will follow the of idea Tse and Tsui (2002) and instead use an average of past values. In their paper they use the following measure $q_{i,j,t-1}$:

$$q_{i,j,t-1} = \frac{Q_{i,j,t-1}}{Q^{1/2}_{i,i,t-1} Q^{1/2}_{j,j,t-1}}$$

where

$$Q_{t-1} = \sum_{l=1}^{L} u_{t-L} u'_{t-L}$$

and $u_{t-k} = (\varepsilon_{i,t-k}, \varepsilon_{j,t-k})'$
While fixing the number of lags L does not seem an attractive alternative, its estimation may nevertheless be problematic as it is a discontinuous parameter. Both these issues can be solved by substituting the arithmetic average with a simple exponential smoothing:

$$Q_{t-1} = \alpha Q_{t-2} + (1-\alpha) u_{t-1} u'_{t-1}$$

where the parameter α, being continuous on the interval $(0,1)$, will be estimated together with the rest of the parameters of the model. Since correlations are bound between plus and minus one it is convenient not to model them directly but rather model their Fisher transform. Therefore, instead of using $q_{i,j,t-1}$ the model will use the variable $\psi_{i,j,t-1}$:

$$\begin{aligned} q_{i,j,t-1} &= \frac{Q_{t-1}[1,2]}{Q_{t-1}[1,1] \cdot Q_{t-1}[2,2]} \\ \psi_{i,j,t-1} &= \frac{1}{2} \ln \left(\frac{1 + q_{i,j,t-1}}{1 - q_{i,j,t-1}} \right) \end{aligned}$$

The Fisher transform of the conditional correlation χ at time t will then be modeled as a linear combination of χ_{t-1} and $\psi_{i,j,t-1}$:

$$\chi_t = \omega + \delta \chi_{t-1} + (\theta + \beta \cdot d_{t-1}) \psi_{i,j,t-1}$$

where d_{t-1} is a dummy variable that allows for an asymmetric response to the realizations.
Targeting the parameter ω so that the long-run mean of the process matches the sample unconditional correlation:

$$\chi_t = (1 - \delta - \theta)\overline{\chi} - \beta \cdot \overline{(d\psi)} + \delta \chi_{t-1} + (\theta + \beta \cdot d_{t-1}) \psi_{i,j,t-1}$$

where $\overline{d}$ is the sample average of d_t and $\overline{\chi} = \frac{1}{2} \ln \left(\frac{1+\overline{\rho}_{i,j}}{1-\overline{\rho}_{i,j}} \right)$. Notice that $\overline{(d\psi)}$ is the sample analog of $E(d\psi)$ which can be rewritten as $Cov(d,\psi) + E(d)E(\psi)$. The dummy variable d_{t-1} is a function of u_{t-1} which enters Q_{t-1} with a weight of $(1-\alpha)$. As the parameter α is usually close to one, the covariance between d_{t-1} and ψ_{t-1} is negligible and therefore $\overline{d\psi} \approx \overline{d} \cdot \overline{\chi}$.

The conditional correlation ρ_t is given by the inverse-Fisher transform of χ_t:

$$\rho_t = \frac{\exp(2\chi_t) - 1}{\exp(2\chi_t) + 1}$$

In the SCC framework, ρ_t is used to construct the matrices $K_{i,j,t}$ and their inverses. The latter are necessary to decompose the conditional correlation matrix while the first will allow the reconstruction of the whole conditional variance-covariance matrix.

5 Empirical Results

The data employed in this paper consists of 69 series from the NASDAQ100. The sample is made of 10 years of daily observations (from 9/1/1994 to 8/31/2004) for a total of 2517 returns. 31 series have not been included in this study because there were not enough observations available: from Apollo Group (2454 obs.) to Kmart Holding (340 obs.). A list of symbols of the series included in this work can be found in Table 1.

Returns have been calculated as log differences of the closing prices. Unconditional means were subtracted from each series of returns before proceeding to the estimation of the conditional second moments. Both the EGARCH(1,1) and the ZARCH(1,1) have been fit to the 69 returns. In order to select the *best* model for the whole data set two measures of goodness of fit have been used: the BIC and the MAD of a 20 days in-sample prediction. While according to the BIC the two models are equivalent (ZARCH is preferred in 52% of the cases against 48% of the EGARCH) the MAD based on the predictions clearly favored the ZARCH (83% of the times) against the EGARCH (17% of the times). Therefore, the selected model was ZARCH for each stock. Furthermore, in a large-scale framework where the univariate GARCH estimation is part of an automated procedure, the ZARCH proved to be the most reliable as it is the easiest to implement and converge.

To provide a reference for the performance of the proposed SCC, the CCC and scalar DCC with asymmetry will also be estimated.

As a first measure of the goodness of fit for the models I computed the Value-at-Risk of level α or the α left and right quantiles of the conditional distribution of returns. This in-sample exercise has been carried out constructing portfolios with random weights for every t in the sample. For each portfolio the conditional one-period-ahead variance is easily computed using the estimates for the conditional variance-covariance matrix of the multivariate-GARCH-models. In order to evaluate the performance of the models w.r.t. the VaR it is enough to check the frequency with which the portfolio returns will lie outside given confidence intervals.

The results in Table 2 show the frequency with which the random portfolios exceed the 95% and 99% confidence intervals under the assumption of Gaussian distribution of the errors. While the models' performances do not seem to differ too much, in this simulation the most accurate measures of risk are provided by the CCC, followed by SCC(3), DCC, SCC(4), SCC(1), and SCC(2). These results would suggest that for this data set the assumption of constant conditional correlations is not unrealistic. Under this circumstance the more flexibles DCC and SCC do not have much space left for improvement. Nevertheless, these results can be considered as somewhat in favor of the SCC. Considering that in this application the conditional correlation matrix was modeled by the SCC using a total of 2384 estimated parameters it is striking how, even in the presence of no information or constant correlations, the model's predictions do not *drown* in the noise.

Another measure of goodness of CCC, DCC, and SCC is the magnitude of the unconditional daily variance of minimum variance portfolios. For each model the one step-ahead prediction of the conditional variance-covariance matrix was used to construct the minimum variance portfolio with short-selling constraint. The unconditional variance of each portfolio was then computed and reported in Table 3. The gain, in terms of variance, of switching from a Balanced (equally weighted) portfolio to a minimum variance portfolio based on H_t from the CCC is striking. The scalar DCC with asymmetries can produce a marginally less volatile portfolios than the CCC. A noticeable improvement, however, comes from the SCC: portfolios based on its conditional variance-covariance matrix exhibit the smallest variances. Again, while the comparison between CCC and DCC would suggest that conditional correlations are indeed constant for this data set, SCC(1)-SCC(4) is still capable of extracting that extra bit of information from the data and produce better results than CCC and DCC without suffering from any over parameterization problem. The particular decomposition of the correlation matrix of the SCC is based on a certain ordering of the series in the data set. While a permutation of such order does not effect the decomposition when the values of the correlations are known, it will have an effect when they need to be estimated. In order to measure the impact of the ordering of the series on the conditional variance-covariance matrix a couple of *shuffles* will be considered. In SCC(1) the stocks are in alphabetical order and in SCC(2) they are in reversed alphabetical order.

The variance of the estimates of the correlations is a function of the fourth moments of the series. Since the higher the kurtosis the bigger the variance of the estimates it seems reasonable to order the series in terms of their fourth moments, starting with the less kurtic. Because of the transformation, every pair of series undergoes after the conditional correlations have been estimated it is clear how the uncertainty at any given point of the estimation will potentially affect all those that follow. Hence, it seems to be a good strategy to estimate R_t beginning with the pair stocks that are least kurtic. The results of SCC(3) seem to confirm this intuition while starting from the most kurtic series as in SCC(4) does not appear to be an attractive alternative.

6 Conclusion

The main purpose of this paper was to present a new approach at the modeling of the conditional correlation matrix. The structure of the SCC allows for a sequential estimation of the various components of such matrix. This way, the traditional intractability of multivariate GARCH models of non-trivial dimensions is eliminated at once. The multi step procedure translates the high dimensionality of the MGARCH into a series of simple univariate estimations. Furthermore, because of the particular decomposition of the variance-covariance matrix in the SCC the latter is guaranteed to be symmetric and positive definite without the imposition of any parametric restriction.

Another strong feature of the SCC is the flexibility it allows in the modeling of the various correlations: not only there are no constraints on the choice of the model but different univariate models can be used at different steps. In this paper I have proposed a new univariate model for correlations that: lets the data decide what is the best proxy for the lagged realizations, allows for an immediate interpretation of its parameters, guarantees that correlations are bound between plus and minus one without the need for parameters' restrictions.

While the chosen empirical application did not allow for a full appreciation of the model because of the *constant conditional correlations*, at the same time it showed that the highly parametrized SCC was still capable of providing a good fit without being overwhelmed by noise.

Future work will include more methods of evaluating the SCC and the competing models in the literature, simulations to grasp a better understanding of the performance of the SCC when the data generating process is one of the competing models and to better asses the importance of the ordering. Given the lack of dynamics in the conditional correlations of this sample, the future work will include a more heterogeneous data set for which dynamic conditional correlation models do outperform the more simple constant conditional correlations model.

7 Appendix

Definition 7.1. *A correlation matrix is a symmetric positive semi-definite matrix giving the correlations between all pairs of data sets.*

Corollary 7.1. *All the elements on the main diagonal of a correlation matrix must be equal to 1 (a random variable co-varies perfectly with itself) and all the off-diagonal elements must be less than 1 in modulus.*

Definition 7.2. $R_{p,q}$*, with* $p < q$*, is an* $(M \times M)$ *correlation matrix of the following form:*

$$\mathbf{R_{p,q}} = \left(\begin{array}{ccccccccc} & & & 0 & & & & & \\ & \mathbf{I}_{(p-1)} & & 0 & & & \mathbf{0} & & \\ & & & \vdots & & & & & \\ 0 & 0 & \dots & 1 & 0 & \dots & \rho & \# & \# \\ & & & 0 & & & & & \\ & & & \vdots & & & & & \\ & \mathbf{0} & & \rho & & & \mathbf{F}_{(M-p)} & & \\ & & & \# & & & & & \\ & & & \# & & & & & \end{array} \right)$$

where ρ *is the* $[p,q]$ *element,* $\mathbf{I}$ *is the identity matrix,* $\mathbf{F}$ *is a full correlation matrix, and* $\#$ *indicate (generally) non-zero elements.*

Definition 7.3. $K^{-1}_{p,q}$ *is a lower triangular matrix of the form:*

$$\mathbf{K^{-1}_{p,q}[i,j]} = \begin{cases} -\rho(1-\rho^2)^{-1/2} & \textit{if } i = q \textit{ and } j = p \\ (1-\rho^2)^{-1/2} & \textit{if } i = q \textit{ and } j = q \\ \mathbf{I}[i,j] & \textit{otherwise} \end{cases}$$

Definition 7.4. $S_{p,q} = K^{-1}_{p,q} R_{p,q} (K^{-1}_{p,q})'$

Corollary 7.2. $S_{p,q}$ *is a correlation matrix such that:*

$$\mathbf{S_{p,q}[i,j]} = 0 \qquad \textit{if } i \leq p,\ j \leq q \textit{ and } i \neq j$$

Proof:

$$\mathbf{S_{p,q}} = \left(\begin{array}{ccccccccc}
 & & & 0 & \cdots & & s_1 & \cdots & 0 \\
 & \mathbf{I}_{(p-1)} & & 0 & \cdots & & s_2 & \cdots & 0 \\
 & & & \vdots & & & \vdots & & \vdots \\
0 & 0 & \cdots & 1 & 0 & \cdots & s_p & \# & \# \\
\vdots & \vdots & & 0 & F_{1,1} & \cdots & s_{p+1} & \cdots & F_{1,M-p} \\
 & & & \vdots & \vdots & & \vdots & & \vdots \\
s_1 & s_2 & \cdots & s_p & s_{p+1} & \cdots & s_q & \cdots & s_M \\
\vdots & \vdots & & \# & \vdots & & \vdots & & \vdots \\
0 & 0 & \cdots & \# & F_{1,M-p} & \cdots & s_M & \cdots & F_{M-p,M-p}
\end{array}\right)$$

where:

$$s_i = \begin{cases} -\frac{\rho}{(1-\rho^2)^{1/2}} R_{pi} + \frac{1}{(1-\rho^2)^{1/2}} R_{qi} & if\ i \neq q \\ \frac{\rho^2}{(1-\rho^2)} R_{pp} - \frac{2\rho}{(1-\rho^2)} R_{pq} + \frac{1}{(1-\rho^2)} R_{qq} & if\ i = q \end{cases}$$

The latter allows to easily compute the following values of interest:

$$\begin{array}{cccccc}
i < p & \Rightarrow & R_{pi} = 0, R_{qi} = 0 & \Rightarrow & s_i = 0 \\
i = p & \Rightarrow & R_{pp} = 1, R_{qp} = \rho & \Rightarrow & s_p = 0 \\
i = q & \Rightarrow & R_{pp} = 1, R_{qq} = 1, R_{pq} = \rho & \Rightarrow & s_q = 1
\end{array}$$

which can be substituted back in $S_{p,q}$:

$$\mathbf{S_{p,q}} = \left(\begin{array}{ccccccccc}
 & & & 0 & \cdots & & 0 & \cdots & 0 \\
 & \mathbf{I}_{(p-1)} & & 0 & \cdots & & 0 & \cdots & 0 \\
 & & & \vdots & & & \vdots & & \vdots \\
0 & 0 & \cdots & 1 & 0 & \cdots & 0 & \# & \# \\
\vdots & \vdots & & 0 & F_{1,1} & \cdots & s_{p+1} & \cdots & F_{1,M-p} \\
 & & & \vdots & \vdots & & \vdots & & \vdots \\
0 & 0 & \cdots & 0 & s_{p+1} & \cdots & 1 & \cdots & s_M \\
\vdots & \vdots & & \# & \vdots & & \vdots & & \vdots \\
0 & 0 & \cdots & \# & F_{1,M-p} & \cdots & s_M & \cdots & F_{M-p,M-p}
\end{array}\right)$$

The $F-$block of $S_{p,q}$ has every element on the main diagonal equal to 1. Furthermore, since $S_{p,q}$ is a quadratic form of a positive semi-definite matrix it is itself a positive *semi-definite matrix and therefore the $F-$block must also be positive semi-definite. Hence, it is a correlation matrix:*

$$\mathbf{S_{p,q}} = \left(\begin{array}{ccccccccc} & & & 0 & \dots & & 0 & \dots & 0 \\ & \mathbf{I}_{(p-1)} & & 0 & \dots & & 0 & \dots & 0 \\ & & & \vdots & & & \vdots & & \vdots \\ 0 & 0 & \dots & 1 & 0 & \dots & 0 & \# & \# \\ \vdots & \vdots & & 0 & & & & & \\ & & & \vdots & & & & & \\ 0 & 0 & \dots & 0 & & & \widehat{\mathbf{F}}_{(M-p)} & & \\ \vdots & \vdots & & \# & & & & & \\ 0 & 0 & \dots & \# & & & & & \end{array} \right)$$

Definition 7.5. $R_{p,q+1} = S_{p,q}$ *if* $q+1 \leq M$ *or* $R_{p+1,M} = S_{p,q}$ *if* $q+1 > M$.

Corollary 7.3. *Given a correlation matrix $R_{p,q}$, successive pre and post-multiplication by the matrices $K^{-1}_{p,q}, K^{-1}_{p,q+1}, ..., K^{-1}_{M-1,M}$ will yield the identity matrix.*
Proof: The application of the matrix $K^{-1}_{r,s}$ and its inverse has the property of changing the correlation in the r-row and s-column (and therefore the s-row and r-column) into a zero without affecting the correlations that have been set to zero by the matrices $K^{-1}_{i,j}$ with $i \leq r$ or $j \geq s$:

$$I = K^{-1}_{M-1,M} \cdot K_{M-2,M} \cdot \ldots \cdot K^{-1}_{p,q} \cdot R_{p,q} \cdot (K^{-1}_{p,q})' \cdot \ldots (K_{M-2,M})' \cdot (K^{-1}_{M-1,M})'$$

Corollary 7.4. *Given a full correlation matrix R, there exists a sequence of matrices $K^{-1}_{1,2}, K^{-1}_{1,3}, ..., K^{-1}_{M-1,M}$ such that:*

$$I = K^{-1}_{M-1,M} \cdot \ldots \cdot K^{-1}_{1,3} \cdot K^{-1}_{1,2} \cdot R \cdot (K^{-1}_{1,2})' \cdot (K^{-1}_{1,3})' \cdot \ldots \cdot (K^{-1}_{M-1,M})'$$

Corollary 7.5. *Given a full correlation matrix R, it can be expressed as the product of a sequence of matrices $K^{-1}_{1,2}, K^{-1}_{1,3}, ..., K^{-1}_{M-1,M}$:*

$$R = K_{1,2} \cdot K_{1,3} \cdot \ldots \cdot K_{M-1,M} \cdot I \cdot K'_{M-1,M} \cdot \ldots \cdot K'_{1,3} \cdot K'_{1,2}$$

Definition 7.6. $K_{p,q} = (K^{-1}_{p,q})^{-1}$ *is a lower triangular matrix of the form:*

$$\mathbf{K_{p,q}[i,j]} = \begin{cases} \rho & \textit{if } i = q \textit{ and } j = p \\ (1-\rho^2)^{1/2} & \textit{if } i = q \textit{ and } j = q \\ \mathbf{I}[i,j] & \textit{otherwise} \end{cases}$$

Definition 7.7. *Let $R_{p,q}$ be the correlation matrix of Definition 3.2 and without loss of generality let $q-1>p$.*

Definition 7.8. *Let $K_{p,q-1}$ be a lower triangular matrix of the form:*

$$\mathbf{K_{p,q-1}[i,j]} = \begin{cases} \rho & \textit{if } i=q-1 \textit{ and } j=p \\ (1-\rho^2)^{1/2} & \textit{if } i=q-1 \textit{ and } j=q-1 \\ \mathbf{I}[i,j] & \textit{otherwise} \end{cases}$$

where ρ can take any value between -1 and 1: $|\rho|<1$.

Definition 7.9. $R_{p,q-1} = K_{p,q-1} R_{p,q} K'_{p,q-1}$

Corollary 7.6. *$R_{p,q-1}$ is a correlation matrix.*
Proof:

$$\mathbf{R_{p,q-1}} = \left(\begin{array}{ccccccccc} & & & 0 & \dots & & x_1 & \dots & 0 \\ & \mathbf{I}_{(p-1)} & & 0 & \dots & & x_2 & \dots & 0 \\ & & & \vdots & & & \vdots & & \vdots \\ 0 & 0 & \dots & 1 & 0 & \dots & x_p & \# & \# \\ \vdots & \vdots & & 0 & F_{1,1} & \dots & x_{p+1} & \dots & F_{1,M-p} \\ & & & \vdots & \vdots & & \vdots & & \vdots \\ x_1 & x_2 & \dots & x_p & x_{p+1} & \dots & x_{q-1} & \dots & x_M \\ \vdots & \vdots & & \# & \vdots & & \vdots & & \vdots \\ 0 & 0 & \dots & \# & F_{1,M-p} & \dots & x_M & \dots & F_{M-p,M-p} \end{array} \right)$$

where:

$$\mathbf{x_i} = \begin{cases} \rho R_{pi} + (1-\rho^2)^{1/2} R_{(q-1)i} & \textit{if } i \neq q-1 \\ \rho^2 R_{pp} + 2\rho(1-\rho^2)^{1/2} R_{(q-1)p} + (1-\rho^2) R_{(q-1)(q-1)} & \textit{if } i = q-1 \end{cases}$$

The latter allows to easily compute the following values of interest:

$$\begin{array}{llccl} i<p & \Rightarrow & R_{pi}=0, R_{(q-1)i}=0 & \Rightarrow & x_i=0 \\ i=p & \Rightarrow & R_{pp}=1, R_{(q-1)p}=0 & \Rightarrow & x_p=\rho \\ i=q-1 & \Rightarrow & R_{pp}=1, R_{(q-1)p}=0, R_{(q-1)(q-1)}=1 & \Rightarrow & x_{q-1}=1 \end{array}$$

Therefore:

$$\mathbf{R_{p,q-1}} = \begin{pmatrix} & & & 0 & \dots & & 0 & \dots & 0 \\ & \mathbf{I}_{(p-1)} & & 0 & \dots & & 0 & \dots & 0 \\ & & & \vdots & & & \vdots & & \vdots \\ 0 & 0 & \dots & 1 & 0 & \dots & \rho & \# & \# \\ \vdots & \vdots & & 0 & F_{1,1} & \dots & x_{p+1} & \dots & F_{1,M-p} \\ & & & \vdots & \vdots & & \vdots & & \vdots \\ 0 & 0 & \dots & \rho & x_{p+1} & \dots & 1 & \dots & x_M \\ \vdots & \vdots & & \# & \vdots & & \vdots & & \vdots \\ 0 & 0 & \dots & \# & F_{1,M-p} & \dots & x_M & \dots & F_{M-p,M-p} \end{pmatrix}$$

$R_{p,q-1}$ *is positive semi-definite because it is the quadratic form of the positive semi-definite matrix* $R_{p,q}$*. Every element on its main diagonal is equal to 1. Hence,* $R_{p,q-1}$ *is a correlation matrix.*

Notice that while $R_{p,q}[p, q-1] = 0$ *for the new correlation matrix we have that* $R_{p,q-1}[p, q-1] = \rho, \quad \forall |\rho| < 1$.

8 References

Alexander, C. (2001), "A Primer on the Orthogonal GARCH Model.", ISMA Center, Mimeo.

Alexander, C. and A. Chibumba (1997), "Multivariate Orthogonal Factor GARCH.", University of Sussex, Mimeo.

Andersen, T.G. and T. Bollerslev (1998), "Answering the Skeptics: Yes, Standard Volatility Models do Provide Accurate Forecasts." *International Economic Review* 39, 885-905.

Baillie, R.T., T. Bollerslev and O. Mikkelson (1996), "Fractionally integrated generalized autoregressive conditional heteroscedasticity" *Journal of Econometrics* 74, 3-30.

Bauwens, L., S. Laurent and J.V.K. Rombouts (2003), "Multivariate GARCH Models: A survey.", CORE Discussion Paper.

Bera, A.K., M.L. Higgins and S. Lee (1992), "Interaction between Autocorrelation and Conditional Heteroscedasticity: A Random-Coefficient Approach" *Journal of Business and Economic Statistics* 10, 133-42.

Bollerslev, T. (1990), "Modeling the Coherence in Short-run Nominal Exchange Rates: A Multivariate Generalized ARCH model." *Review of Economics and Statistics* 72, 498-595.

Bollerslev, T. (1987), "A conditional Heteroscedastic Time Series Model for Speculative Prices and Rates of Return." *Review of Economics and Statistics* 69, 542-547.

Bollerslev, T. (1986), "Generalized Autoregressive Conditional Heteroscedasticity." *Journal of Econometrics* 31, 307-327.

Bollerslev, T. and R. Engle (1986), "Modeling the Persistence of Conditional Variances" *Econometric Reviews* 5, 1-50.

Bollerslev, T., R. Engle, and J. Wooldridge (1988), "A Capital Asset Pricing Model with Time Varying Covariances" *Journal of Political Economy* 96, 116-131.

Bollerslev T. and O. Mikkelsen (1996), “Modeling and pricing long memory in stock market volatility” *Journal of Econometrics* 73, 151-184.

Brooks, C., S.P. Burke, and G. Persand (2003), “Multivariate GARCH Models: Software Choice and Estimation Issues.” *Journal of Applied Econometrics* 18, 725-734.

Cappiello, L., R. Engle, and K. Sheppard (2003), “Asymmetric Dynamics in the Correlations of Global Equity and Bond Returns.”, Working Paper 204, Working Paper Series of the European Central Bank.

Chen, X. and Y. Fan (2004), “Estimation and Model Selection of Semiparametric Copula-Based Multivariate Dynamic Models Under Copula Misspecification.”

Cumby, R., S. Figlewski, and J. Hasbrouck (1993), “Forecasting Volatility and Correlations with EGARCH Models.” *Journal of Derivatives*, Winter, 51-63.

Diebold, F.X. and M. Nerlove (1989), “The Dynamics of Exchange Rate Volatility: A Multivariate Latent Factor Arch Model.” *Journal of Applied Econometrics* 1, 1-21.

Ding, Z., R. Engle, and C.W.J. Granger (1993), “A Long Memory Property of Stock Market Returns and a New Model.” *Journal of Empirical Finance* 1, 83-106.

Duan, J. (1997), “Augmented GARCH(p,q) process and its diffusion limit” *Journal of Econometrics* 79(1), 97-127.

Engle, R. (2001), “Dynamic Conditional Correlation - a Simple Class of Multivariate GARCH Models”, forthcoming in *Journal of Business and Economic Statistics*.

Engle, R. (1990), “Stock Volatility and the Crash of ’87.” *The Review of Financial Studies* 3, 103-106.

Engle, R. and F. Kroner (1995), “Multivariate Simultaneous Generalized ARCH.” *Econometric Theory* 11, 122-150.

Engle, R. and J. Mezrich (1996), “GARCH for groups” *RISK* 9, 36-40.

Engle, R. and V. Ng (1993), “Measuring and Testing the Impact of News On Volatility.” *Journal of Finance* 48, 1749-1778.

Engle, R. and G.G.J. Lee (1993), "A permanent and transitory component model of stock return volatility" *Economics Working Paper Series from Department of Economics, UC San Diego.*

Engle, R., V. Ng, and M. Rothschild (1990), "Asset Pricing with a Factor-ARCH Covariance Structure: Empirical Estimates for Treasury Bills." *Journal of Econometrics* 45, 213-238.

Figlewski, S. (1997), "Forecasting Volatility." *Financial Markets, Institutions and Instruments* 6, 1-88.

Gallant, A.R. and G. Tauchen (2002), "SNP: A Program for Nonparametric Time Series Analysis", User's Guide Version 8.8

Gallant, A.R. and G. Tauchen (1989), "Seminonparametric Estimation of Conditionally Constrained Heterogeneous Processes: Asset Pricing Applications." *Econometrica* 57, 1091-1120.

Geweke, J. (1986), "Modeling the persistence in conditional variances: A comment" *Econometric Review* 5, 57-61.

Gilks, W.R., S. Richardson, and D.J. Spiegelhalter (1996), *Markov Chain Monte Carlo in Practice*, London, Chapman & Hall.

Glosten, L., R. Jagannathan, and D. Runkle (1993), "On the Relationship Between the Expected Value and the Volatility of the Nominal Excess Return on Stocks." *Journal of Finance*, 48, 1779-1801.

Gouriéroux, C. and A. Monfort (1996), *Simulation-Based Econometric Methods*, New York, Oxford University Press.

Gouriéroux, C., A. Monfort and A. Trognon (1984), "Pseudo-Maximum Likelihood Methods: Theory." *Econometrica* 52, 681-700.

Hamilton, J.D. (1994), *Time Series Analysis*, Princeton, Princeton University Press.

Hamilton J. and R. Susmel (1994), "Autoregressive conditional heteroscedasticity and changes in regime" *Journal of Econometrics* 64, 307-333.

Harvey, A., E. Ruiz and E. Sentana (1992), "Unobserved component time series models with ARCH disturbances" *Journal of Econometrics* 52, 129-157.

Hausman, J. (1975), "An Instrumental Variable Approach to Full-Information Estimators for Linear and Certain Nonlinear Models." *Econometrica* 43, 727-738.

Hentschel, L. (1995), "All in the family: Nesting symmetric and asymmetric GARCH models" *Journal of Financial Economics*, 39, 71-104.

Higgins, M.L. and A.K. Bera (1992), "A Class of Nonlinear ARCH Models." *International Economic Review* 33, 137-158.

Ingber, L. (1989), "Very Fast Simulated Re-Annealing." *Mathematical Computer Modeling* 12, 967-973.

Jondeau, E. and M. Rockinger (2001), "The copula-GARCH model of conditional dependencies: An international stock-market application." *Forthcoming in Journal of International Money and Finance*

Jorion, P. (1995), "Predicting Volatility in the Foreign Exchange Market." *Journal of Finance* 50, 507-528.

Jorion, P. (1996), "Risk and Turnover in the Foreign Exchange Market." in J.A. Frankel, G. Galli, and A. Giovannini, eds., *The Microstructure of Foreign Exchange Markets* (Chicago: The University of Chicago Press, pp. 19-37).

Kawakatsu, H. (2003), "Matrix Exponential GARCH"

Ledoit, O., P. Santa-Clara, and M. Wolf (2002), "Flexible Multivariate GARCH Modeling With an Application to International Stock Markets.", forthcoming in *The Review of Economics and Statistics*

Lin, W. (1992), "Alternative Estimators for Factor GARCH Models - A Monte Carlo Comparison." *Journal of Applied Econometrics* 7, 259-279.

Milhøj, A. (1987), "A Conditional Variance Model for Daily Deviations of Exchange Rate" *Journal of Business and Economic Statistics* 5(1), 99-103.

Nelson, D.B. (1991), "Conditional Heteroscedasticity in Asset Returns: A New Approach." *Econometrica* 59, 347-370.

Pantula, S.G. (1986), "Modeling the Persistence of Conditional Variances: A Comment" *Econometric Reviews* 5, 71-74.

Patton, A. (2000), "Modeling time-varying exchange rate dependence using the conditional copula." *University of California, San Diego, Discussion Paper* 01-09.

Sentana, E. (1995), "Quadratic ARCH models" *Review of Economic Studies* 62(4), 639-661.

Schwert, G. (1989), "Why Does Stock Market Volatility Change over Time?" *Journal of Finance* 44, 1115-1153.

Sheppard, K. (2002), "Understanding the Dynamics of Equity Covariance.", Manuscript, UCSD.

Taylor, S.J. (1986), *Modeling Financial Time Series*, John Wiley and Sons Ltd.

Tse, Y. and A. Tsui (2002), "A Multivariate GARCH Model with Time-Varying Correlations." *Journal of Business and Economic Statistics* 20, 351-362.

Zakonian, J. (1994), "Threshold Heteroscedastic Models." *Journal of Economic Dynamics and Control* 18, 931-955.

Table 1: Symbols of Included Stocks

ADBE	ALTR	APCC	AMGN	AAPL	AMAT
BBBY	BIIB	BMET	CDWC	CEPH	CHIR
CTAS	CSCO	CMCSA	CPWR	CMVT	COST
DELL	XRAY	ERTS	EXPD	ESRX	FAST
FHCC	FISV	FLEX	GNTX	GENZ	GILD
IACI	INTC	INTU	JDSU	KLAC	LRCX
LNCR	LLTC	MXIM	MEDI	MERQ	MCHP
MSFT	MOLX	NXTL	NVLS	ORCL	PCAR
PDCO	PTEN	PAYX	PSFT	PETM	QLGC
QCOM	ROST	SANM	SIAL	SSCC	SPLS
SBUX	SUNW	SYMC	SNPS	TLAB	TEVA
VRTS	WFMI	XLNX			

Table 2: Sample Probabilities of Lying Outside the Confidence Intervals

	95%	99%
CCC	4.56	0.91
DCC	4.33	0.95
SCC(1)	4.21	1.11
SCC(2)	4.13	1.07
SCC(3)	4.48	1.11
SCC(4)	4.29	1.03

Table 3: Unconditional Daily Variances for Minimum Variance Portfolios with Short Selling Constraints

Balanced	3.5422
CCC	1.2459
DCC	1.2303
SCC(1)	1.1497
SCC(2)	1.1426
SCC(3)	1.1566
SCC(4)	1.1470