

Multivariate GARCH models

To appear in T. G. Andersen, R. A. Davis, J.-P. Kreiss and T. Mikosch, eds.
Handbook of Financial Time Series. New York: Springer.

Annastiina Silvennoinen*

*School of Finance and Economics, University of Technology Sydney
Box 123, Broadway NSW 2007*

and

Timo Teräsvirta†

*CREATES, School of Economics and Management, University of Aarhus
Building 1322, DK-8000 Aarhus C*

and

*Department of Economic Statistics, Stockholm School of Economics,
P. O. Box 6501, SE-113 83 Stockholm, Sweden*

SSE/EFI Working Paper Series in Economics and Finance No. 669

January 2008

Abstract

This article contains a review of multivariate GARCH models. Most common GARCH models are presented and their properties considered. This also includes nonparametric and semiparametric models. Existing specification and misspecification tests are discussed. Finally, there is an empirical example in which several multivariate GARCH models are fitted to the same data set and the results compared.

JEL classification: C32; C51; C52

Key words: Multivariate GARCH; Volatility

*e-mail: annastiina.silvennoinen@uts.edu.au

†e-mail: tterasvirta@econ.au.dk

Acknowledgements. This research has been supported by the Jan Wallander and Tom Hedelius Foundation, Grant No. P2005-0033:1, and the Danish National Research Foundation. We thank Robert Engle for providing the data for the empirical section, Markku Lanne and Pentti Saikkonen for sharing with us their program code for their Generalized Orthogonal Factor GARCH model, and Mika Meitz for programming assistance. Part of the research was done while the first author was visiting CREATES, University of Aarhus, whose kind hospitality is gratefully acknowledged. The responsibility for any errors and shortcomings in this paper remains ours.

1 Introduction

Modelling volatility in financial time series has been the object of much attention ever since the introduction of the Autoregressive Conditional Heteroskedasticity (ARCH) model in the seminal paper of Engle (1982). Subsequently, numerous variants and extensions of ARCH models have been proposed. A large body of this literature has been devoted to univariate models; see for example Bollerslev, Engle, and Nelson (1994), Palm (1996), Shephard (1996), and chapters 1–7 of this Handbook for surveys of this literature.

While modelling volatility of the returns has been the main centre of attention, understanding the comovements of financial returns is of great practical importance. It is therefore important to extend the considerations to multivariate GARCH (MGARCH) models. For example, asset pricing depends on the covariance of the assets in a portfolio, and risk management and asset allocation relate for instance to finding and updating optimal hedging positions. For examples, see Bollerslev, Engle, and Wooldridge (1988), Ng (1991), and Hansson and Hördahl (1998). Multivariate GARCH models have also been used to investigate volatility and correlation transmission and spillover effects in studies of contagion, see Tse and Tsui (2002) and Bae, Karolyi, and Stulz (2003).

What then should the specification of an MGARCH model be like? On one hand, it should be flexible enough to be able to represent the dynamics of the conditional variances and covariances. On the other hand, as the number of parameters in an MGARCH model often increases rapidly with the dimension of the model, the specification should be parsimonious enough to allow for relatively easy estimation of the model and also allow for easy interpretation of the model parameters. However, parsimony often means simplification, and models with only a few parameters may not be able to capture the relevant dynamics in the covariance structure. Another feature that needs to be taken into account in the specification is imposing positive definiteness (as covariance matrices need, by definition, to be positive definite). One possibility is to derive conditions under which the conditional covariance matrices implied by the model are positive definite, but this is often infeasible in practice. An alternative is to formulate the model in a way that positive definiteness is implied by the model structure (in addition to some simple constraints).

Combining these needs has been the difficulty in the MGARCH literature. The first GARCH model for the conditional covariance matrices was the so-called VEC model of Bollerslev, Engle, and Wooldridge (1988), see Engle, Granger, and Kraft (1984) for an ARCH version. This model is a very general one, and a goal of the subsequent literature has been to formulate more parsimonious models. Furthermore, since imposing positive definiteness of the conditional covariance matrix in this model is difficult, formulating models with this feature has been considered important. Furthermore, constructing models in which the estimated parameters have direct interpretation has been viewed as beneficial.

In this paper, we survey the main developments of the MGARCH literature. For another such survey, see Bauwens, Laurent, and Rombouts (2006). This paper is organized as follows. In Section 2, several MGARCH specifications are reviewed. Statistical properties of the models are the topic of Section 3, whereas testing MGARCH models is discussed in Section 4. An empirical comparison of a selection of the models is given in Section 5. Finally, some conclusions and directions for future research are provided in Section 6.

2 Models

Consider a stochastic vector process $\{\mathbf{r}_t\}$ with dimension $N \times 1$ such that $E\mathbf{r}_t = \mathbf{0}$. Let \mathcal{F}_{t-1} denote the information set generated by the observed series $\{\mathbf{r}_t\}$ up to and including time $t - 1$. We assume that \mathbf{r}_t is conditionally heteroskedastic:

$$\mathbf{r}_t = \mathbf{H}_t^{1/2} \boldsymbol{\eta}_t \quad (1)$$

given the information set \mathcal{F}_{t-1} , where the $N \times N$ matrix $\mathbf{H}_t = [h_{ijt}]$ is the conditional covariance matrix of \mathbf{r}_t and $\boldsymbol{\eta}_t$ is an iid vector error process such that $E\boldsymbol{\eta}_t \boldsymbol{\eta}_t' = \mathbf{I}$. This defines the standard multivariate GARCH framework, in which there is no linear dependence structure in $\{\mathbf{r}_t\}$. In financial applications, \mathbf{r}_t is most often viewed as a vector of log-returns of N assets.

What remains to be specified is the matrix process \mathbf{H}_t . Various parametric formulations will be reviewed in the following subsections. We have divided these models into four categories. In the first one, the conditional covariance matrix \mathbf{H}_t is modelled directly. This class includes, in particular, the VEC and BEKK models to be defined in Section 2.1 that were among the first parametric MGARCH models. The models in the second class, the factor models, are motivated by parsimony: the process \mathbf{r}_t is assumed to be generated by a (small) number of unobserved heteroskedastic factors. Models in the third class are built on the idea of modelling the conditional variances and correlations instead of straightforward modelling of the conditional covariance matrix. Members of this class include the Constant Conditional Correlation (CCC) model and its extensions. The appeal of this class lies in the intuitive interpretation of the correlations, and models belonging to it have received plenty of attention in the recent literature. Finally, we consider semi- and nonparametric approaches that can offset the loss of efficiency of the parametric estimators due to misspecified structure of the conditional covariance matrices. Multivariate stochastic volatility models are discussed in a separate chapter of this Handbook, see Chib, Omori, and Asai (2008).

Before turning to the models, we discuss some points that need attention when specifying an MGARCH model. As already mentioned, a problem with MGARCH models is that the number of parameters can increase very rapidly as the dimension of \mathbf{r}_t increases. This creates difficulties in the estimation of the models, and therefore an important goal in constructing new MGARCH models is to make them reasonably parsimonious while maintaining flexibility. Another aspect that has to be imposed is the positive definiteness of the conditional covariance matrices. Ensuring positive definiteness of a matrix, usually through an eigenvalue-eigenvector-decomposition, is a numerically difficult problem, especially in large systems. Yet another difficulty with MGARCH models has to do with the numerical optimization of the likelihood function (in the case of parametric models). The conditional covariance (or correlation) matrix appearing in the likelihood depends on the time index t , and often has to be inverted for all t in every iteration of the numerical optimization. When the dimension of \mathbf{r}_t increases, this is a both time consuming and numerically unstable procedure. Avoiding excessive inversion of matrices is thus a worthy goal in designing MGARCH models. It should be emphasized, however, that practical implementation of all the models to be considered in this chapter is of course feasible, but the problem lies in devising easy to use, automated estimation routines that would make widespread use of these models possible.

2.1 Models of the conditional covariance matrix

The VEC-GARCH model of Bollerslev, Engle, and Wooldridge (1988) is a straightforward generalization of the univariate GARCH model. Every conditional variance and covariance is a

function of all lagged conditional variances and covariances, as well as lagged squared returns and cross-products of returns. The model may be written as follows:

$$\text{vech}(\mathbf{H}_t) = \mathbf{c} + \sum_{j=1}^q \mathbf{A}_j \text{vech}(\mathbf{r}_{t-j} \mathbf{r}'_{t-j}) + \sum_{j=1}^p \mathbf{B}_j \text{vech}(\mathbf{H}_{t-j}) \quad (2)$$

where $\text{vech}(\cdot)$ is an operator that stacks the columns of the lower triangular part of its argument square matrix, \mathbf{c} is an $N(N+1)/2 \times 1$ vector, and \mathbf{A}_j and \mathbf{B}_j are $N(N+1)/2 \times N(N+1)/2$ parameter matrices. In fact, the authors introduced a multivariate GARCH-in-mean model, but in this chapter we only consider its conditional covariance component. The generality of the VEC model is an advantage in the sense that the model is very flexible, but it also brings disadvantages. One is that there exist only sufficient, rather restrictive, conditions for \mathbf{H}_t to be positive definite for all t , see Gouriéroux (1997, Chapter 6). Besides, the number of parameters equals $(p+q)(N(N+1)/2)^2 + N(N+1)/2$, which is large unless N is small. Furthermore, as will be discussed below, estimation of the parameters is computationally demanding.

Bollerslev, Engle, and Wooldridge (1988) presented a simplified version of the model by assuming that \mathbf{A}_j and \mathbf{B}_j in (2) are diagonal matrices. In this case, it is possible to obtain conditions for \mathbf{H}_t to be positive definite for all t , see Bollerslev, Engle, and Nelson (1994). Estimation is less difficult than in the complete VEC model because each equation can be estimated separately. But then, this ‘diagonal VEC’ model that contains $(p+q+1)N(N+1)/2$ parameters seems too restrictive since no interaction is allowed between the different conditional variances and covariances.

A numerical problem is that estimation of parameters of the VEC model is computationally demanding. Assuming that the errors $\boldsymbol{\eta}_t$ follow a multivariate normal distribution, the log-likelihood of the model (1) has the following form:

$$\sum_{t=1}^T \ell_t(\boldsymbol{\theta}) = c - (1/2) \sum_{t=1}^T \ln |\mathbf{H}_t| - (1/2) \sum_{t=1}^T \mathbf{r}'_t \mathbf{H}_t^{-1} \mathbf{r}_t. \quad (3)$$

The parameter vector $\boldsymbol{\theta}$ has to be estimated iteratively. It is seen from (3) that the conditional covariance matrix \mathbf{H}_t has to be inverted for every t in each iteration, which may be tedious when the number of observations is large and when, at the same time, N is not small. Another, an even more difficult problem, is how to ensure positive definiteness of the covariance matrices. In the case of the VEC model there does not seem to exist a general solution to this problem. The problem of finding the necessary starting-values for \mathbf{H}_t is typically solved by using the estimated unconditional covariance matrix as the initial value.

A model that can be viewed as a restricted version of the VEC model is the Baba-Engle-Kraft-Kroner (BEKK) defined in Engle and Kroner (1995). It has the attractive property that the conditional covariance matrices are positive definite by construction. The model has the form

$$\mathbf{H}_t = \mathbf{C}\mathbf{C}' + \sum_{j=1}^q \sum_{k=1}^K \mathbf{A}'_{kj} \mathbf{r}_{t-j} \mathbf{r}'_{t-j} \mathbf{A}_{kj} + \sum_{j=1}^p \sum_{k=1}^K \mathbf{B}'_{kj} \mathbf{H}_{t-j} \mathbf{B}_{kj} \quad (4)$$

where \mathbf{A}_{kj} , \mathbf{B}_{kj} , and \mathbf{C} are $N \times N$ parameter matrices, and \mathbf{C} is lower triangular. The decomposition of the constant term into a product of two triangular matrices is to ensure positive definiteness of \mathbf{H}_t . The BEKK model is covariance stationary if and only if the eigenvalues of $\sum_{j=1}^q \sum_{k=1}^K \mathbf{A}_{kj} \otimes \mathbf{A}_{kj} + \sum_{j=1}^p \sum_{k=1}^K \mathbf{B}_{kj} \otimes \mathbf{B}_{kj}$, where \otimes denotes the Kronecker product of two matrices, are less than one in modulus. Whenever $K > 1$ an identification problem arises

because there are several parameterizations that yield the same representation of the model. Engle and Kroner (1995) give conditions for eliminating redundant, observationally equivalent representations.

Interpretation of parameters of (4) is not easy. But then, consider the first order model

$$\mathbf{H}_t = \mathbf{C}\mathbf{C}' + \mathbf{A}'\mathbf{r}_{t-1}\mathbf{r}'_{t-1}\mathbf{A} + \mathbf{B}'\mathbf{H}_{t-1}\mathbf{B}. \quad (5)$$

Setting $\mathbf{B} = \mathbf{A}\mathbf{D}$ where \mathbf{D} is a diagonal matrix, (5) becomes

$$\mathbf{H}_t = \mathbf{C}\mathbf{C}' + \mathbf{A}'\mathbf{r}_{t-1}\mathbf{r}'_{t-1}\mathbf{A} + \mathbf{D}\mathbf{E}[\mathbf{A}'\mathbf{r}_{t-1}\mathbf{r}'_{t-1}\mathbf{A}|\mathcal{F}_{t-2}]\mathbf{D}. \quad (6)$$

It is seen from (6) that what is now modelled are the conditional variances and covariances of certain linear combinations of the vector of asset returns \mathbf{r}_t or ‘portfolios’. Kroner and Ng (1998) restrict $\mathbf{B} = \delta\mathbf{A}$ where $\delta > 0$ is a scalar.

A further simplified version of (5) in which \mathbf{A} and \mathbf{B} are diagonal matrices has sometimes appeared in applications. This ‘diagonal BEKK’ model trivially satisfies the equation $\mathbf{B} = \mathbf{A}\mathbf{D}$. It is a restricted version of the diagonal VEC model such that the parameters of the covariance equations (equations for h_{ijt} , $i \neq j$) are products of the parameters of the variance equations (equations for h_{iit}). In order to obtain a more general model (that is, to relax these restrictions on the coefficients of the covariance terms) one has to allow $K > 1$. The most restricted version of the diagonal BEKK model is the scalar BEKK one with $\mathbf{A} = a\mathbf{I}$ and $\mathbf{B} = b\mathbf{I}$ where a and b are scalars.

Each of the BEKK models implies a unique VEC model, which then generates positive definite conditional covariance matrices. Engle and Kroner (1995) provide sufficient conditions for the two models, BEKK and VEC, to be equivalent. They also give a representation theorem that establishes the equivalence of diagonal VEC models (that have positive definite covariance matrices) and general diagonal BEKK models. When the number of parameters in the BEKK model is less than the corresponding number in the VEC model, the BEKK parameterization imposes restrictions that makes the model different from that of VEC model. Increasing K in (4) eliminates those restrictions and thus increases the generality of the BEKK model towards the one obtained from using pure VEC model. Engle and Kroner (1995) give necessary conditions under which all unnecessary restrictions are eliminated. However, too large a value of K will give rise to the identification problem mentioned earlier.

Estimation of a BEKK model still involves somewhat heavy computations due to several matrix inversions. The number of parameters, $(p+q)KN^2 + N(N+1)/2$ in the full BEKK model, or $(p+q)KN + N(N+1)/2$ in the diagonal one, is still quite large. Obtaining convergence may therefore be difficult because (4) is not linear in parameters. There is the advantage, however, that the structure automatically ensures positive definiteness of \mathbf{H}_t , so this does not need to be imposed separately. Partly because numerical difficulties are so common in the estimation of BEKK models, it is typically assumed $p = q = K = 1$ in applications of (4).

Parameter restrictions to ensure positive definiteness are not needed in the matrix exponential GARCH model proposed by Kawakatsu (2006). It is a generalization of the univariate exponential GARCH model (Nelson, 1991) and is defined as follows:

$$\begin{aligned} \text{vech}(\ln \mathbf{H}_t - \mathbf{C}) &= \sum_{i=1}^q \mathbf{A}_i \boldsymbol{\eta}_{t-i} + \sum_{i=1}^q \mathbf{F}_i (|\boldsymbol{\eta}_{t-i}| - E|\boldsymbol{\eta}_{t-i}|) \\ &+ \sum_{i=1}^p \mathbf{B}_i \text{vech}(\ln \mathbf{H}_{t-i} - \mathbf{C}) \end{aligned} \quad (7)$$

where \mathbf{C} is a symmetric $N \times N$ matrix, and \mathbf{A}_i , \mathbf{B}_i , and \mathbf{F}_i are parameter matrices of sizes $N(N+1)/2 \times N$, $N(N+1)/2 \times N(N+1)/2$, and $N(N+1)/2 \times N$, respectively. There is no need to impose restrictions on the parameters to ensure positive definiteness, because the matrix $\ln \mathbf{H}_t$ need not be positive definite. The positive definiteness of the covariance matrix \mathbf{H}_t follows from the fact that for any symmetric matrix \mathbf{S} , the matrix exponential defined as

$$\exp(\mathbf{S}) = \sum_{i=0}^{\infty} \frac{\mathbf{S}^i}{i!}$$

is positive definite. Since the model contains a large number of parameters, Kawakatsu (2006) discusses a number of more parsimonious specifications. He also considers the estimation of the model, hypothesis testing, the interpretation of the parameters, and provides an application. How popular this model will turn out in practice remains to be seen.

2.2 Factor models

Factor models are motivated by economic theory. For instance, in the arbitrage pricing theory of Ross (1976) returns are generated by a number of common unobserved components, or factors; for further discussion see Engle, Ng, and Rothschild (1990) who introduced the first factor GARCH model. In this model it is assumed that the observations are generated by underlying factors that are conditionally heteroskedastic and possess a GARCH-type structure. This approach has the advantage that it reduces the dimensionality of the problem when the number of factors relative to the dimension of the return vector \mathbf{r}_t is small.

Engle, Ng, and Rothschild (1990) define a factor structure for the conditional covariance matrix as follows. They assume that \mathbf{H}_t is generated by K ($< N$) underlying, not necessarily uncorrelated, factors $f_{k,t}$ as follows:

$$\mathbf{H}_t = \mathbf{\Omega} + \sum_{k=1}^K \mathbf{w}_k \mathbf{w}_k' f_{k,t} \quad (8)$$

where $\mathbf{\Omega}$ is an $N \times N$ positive semi-definite matrix, \mathbf{w}_k , $k = 1, \dots, K$, are linearly independent $N \times 1$ vectors of factor weights, and the $f_{k,t}$'s are the factors. It is assumed that these factors have a first-order GARCH structure:

$$f_{k,t} = \omega_k + \alpha_k (\boldsymbol{\gamma}_k' \mathbf{r}_{t-1})^2 + \beta_k f_{k,t-1}$$

where ω_k , α_k , and β_k are scalars and $\boldsymbol{\gamma}_k$ is an $N \times 1$ vector of weights. The number of factors K is intended to be much smaller than the number of assets N , which makes the model feasible even for a large number of assets. Consistent but not efficient two-step estimation method using maximum likelihood is discussed in Engle, Ng, and Rothschild (1990). In their application, the authors consider two factor-representing portfolios as the underlying factors that drive the volatilities of excess returns of the individual assets. One factor consists of value-weighted stock index returns and the other one of average T-bill returns of different maturities. This choice is motivated by principal component analysis.

Diebold and Nerlove (1989) propose a model similar to the one formulated in Engle, Ng, and Rothschild (1990). However their model is rather a stochastic volatility model than a GARCH one, and hence we do not discuss its properties here; see Sentana (1998) for a comparison of this model with the factor GARCH one.

In the factor ARCH model of Engle, Ng, and Rothschild (1990) the factors are generally correlated. This may be undesirable as it may turn out that several of the factors capture

very similar characteristics of the data. If the factors were uncorrelated, they would represent genuinely different common components driving the returns. Motivated by this consideration, several factor models with uncorrelated factors have been proposed in the literature. In all of them, the original observed series contained in \mathbf{r}_t are assumed to be linked to unobserved, uncorrelated variables, or factors, \mathbf{z}_t through a linear, invertible transformation \mathbf{W} :

$$\mathbf{r}_t = \mathbf{W}\mathbf{z}_t$$

where \mathbf{W} is thus a nonsingular $N \times N$ matrix. Use of uncorrelated factors can potentially reduce their number relative to the approach where the factors can be correlated. The unobservable factors are estimated from the data through \mathbf{W} . The factors \mathbf{z}_t are typically assumed to follow a GARCH process. Differences between the factor models are due to the specification of the transformation \mathbf{W} and, importantly, whether the number of heteroskedastic factors is less than the number of assets or not.

In the Generalized Orthogonal (GO-) GARCH model of van der Weide (2002), the uncorrelated factors \mathbf{z}_t are standardized to have unit unconditional variances, that is, $E\mathbf{z}_t\mathbf{z}_t' = \mathbf{I}$. This specification extends the Orthogonal (O-) GARCH model of Alexander and Chibumba (1997) in that \mathbf{W} is not required to be orthogonal, only invertible. The factors are conditionally heteroskedastic with GARCH-type dynamics. The $N \times N$ diagonal matrix of conditional variances of \mathbf{z}_t is defined as follows:

$$\mathbf{H}_t^z = (\mathbf{I} - \mathbf{A} - \mathbf{B}) + \mathbf{A} \odot (\mathbf{z}_{t-1}\mathbf{z}_{t-1}') + \mathbf{B}\mathbf{H}_{t-1}^z \quad (9)$$

where \mathbf{A} and \mathbf{B} are diagonal $N \times N$ parameter matrices and \odot denotes the Hadamard (i.e. elementwise) product of two conformable matrices. The form of the constant term imposes the restriction $E\mathbf{z}_t\mathbf{z}_t' = \mathbf{I}$. Covariance stationarity of \mathbf{r}_t in the models with uncorrelated factors is ensured if the diagonal elements of $\mathbf{A} + \mathbf{B}$ are less than one. Therefore the conditional covariance matrix of \mathbf{r}_t can be expressed as

$$\mathbf{H}_t = \mathbf{W}\mathbf{H}_t^z\mathbf{W}' = \sum_{k=1}^N \mathbf{w}_{(k)}\mathbf{w}_{(k)}' h_{k,t}^z \quad (10)$$

where $\mathbf{w}_{(k)}$ are the columns of the matrix \mathbf{W} and $h_{k,t}^z$ are the diagonal elements of the matrix \mathbf{H}_t^z . The difference between equations (8) and (10) is that the factors in (10) are uncorrelated but then, in the GO-GARCH model it is not possible to have fewer factors than there are assets. This is possible in the O-GARCH model but at the cost of obtaining conditional covariance matrices with a reduced rank.

Van der Weide (2002) constructs the linear mapping \mathbf{W} by making use of the singular value decomposition of $E\mathbf{r}_t\mathbf{r}_t' = \mathbf{W}\mathbf{W}'$. That is,

$$\mathbf{W} = \mathbf{U}\mathbf{\Lambda}^{1/2}\mathbf{V}$$

where the columns of \mathbf{U} hold the eigenvectors of $E\mathbf{r}_t\mathbf{r}_t'$ and the diagonal matrix $\mathbf{\Lambda}$ holds its eigenvalues, thus exploiting unconditional information only. Estimation of the orthogonal matrix \mathbf{V} requires use of conditional information; see van der Weide (2002) for details.

Vrontos, Dellaportas, and Politis (2003) have suggested a related model. They state their Full Factor (FF-) GARCH model as above but restrict the mapping \mathbf{W} to be an $N \times N$ invertible triangular parameter matrix with ones on the main diagonal. Furthermore, the parameters in \mathbf{W} are estimated directly using conditional information only. Assuming \mathbf{W} to be triangular

simplifies matters but is restrictive because, depending on the order of the components in the vector \mathbf{r}_t , certain relationships between the factors and the returns are ruled out.

Lanne and Saikkonen (2007) put forth yet another modelling proposal. In their Generalized Orthogonal Factor (GOF-) GARCH model the mapping \mathbf{W} is decomposed using the polar decomposition:

$$\mathbf{W} = \mathbf{C}\mathbf{V}$$

where \mathbf{C} is a symmetric positive definite $N \times N$ matrix and \mathbf{V} an orthogonal $N \times N$ matrix. Since $E\mathbf{r}_t\mathbf{r}'_t = \mathbf{W}\mathbf{W}' = \mathbf{C}\mathbf{C}'$, the matrix \mathbf{C} can be estimated making use of the spectral decomposition $\mathbf{C} = \mathbf{U}\mathbf{\Lambda}^{1/2}\mathbf{U}'$, where the columns of \mathbf{U} are the eigenvectors of $E\mathbf{r}_t\mathbf{r}'_t$ and the diagonal matrix $\mathbf{\Lambda}$ contains its eigenvalues, thus using unconditional information only. Estimation of \mathbf{V} requires the use of conditional information, see Lanne and Saikkonen (2007) for details.

An important aspect of the GOF-GARCH model is that some of the factors can be conditionally homoskedastic. In addition to being parsimonious, this allows the model to include not only systematic but also idiosyncratic components of risk. Suppose K ($\leq N$) of the factors are heteroskedastic, while the remaining $N - K$ factors are homoskedastic. Without loss of generality we can assume that the K first elements of \mathbf{z}_t are the heteroskedastic ones, in which case this restriction is imposed by setting that the $N - K$ last diagonal elements of \mathbf{A} and \mathbf{B} in (9) equal to zero. This results in the conditional covariance matrix of \mathbf{r}_t of the following form (ref. eq. (10)):

$$\begin{aligned} \mathbf{H}_t &= \sum_{k=1}^K \mathbf{w}_{(k)}\mathbf{w}'_{(k)}h_{k,t}^z + \sum_{k=K+1}^N \mathbf{w}_{(k)}\mathbf{w}'_{(k)} \\ &= \sum_{k=1}^K \mathbf{w}_{(k)}\mathbf{w}'_{(k)}h_{k,t}^z + \mathbf{\Omega}. \end{aligned} \tag{11}$$

The expression (11) is very similar to the one in (8), but there are two important differences. In (11) the factors are uncorrelated, whereas in (8), as already pointed out, this is not generally the case. The role of $\mathbf{\Omega}$ in (11) is also different from that of $\mathbf{\Omega}$ in (8). In the factor ARCH model $\mathbf{\Omega}$ is required to be a positive semi-definite matrix and it has no particular interpretation. For comparison, the matrix $\mathbf{\Omega}$ in the GOF-GARCH model has a reduced rank directly related to the number of heteroskedastic factors. Furthermore, it is closely related to the unconditional covariance matrix of \mathbf{r}_t . This results to the model being possibly considerably more parsimonious than the factor ARCH model; for details and a more elaborate discussion, see Lanne and Saikkonen (2007). Therefore, the GOF-GARCH model can be seen as combining the advantages of both the factor models (having a reduced number of heteroskedastic factors) and the orthogonal models (relative ease of estimation due to the orthogonality of factors).

2.3 Models of conditional variances and correlations

Correlation models are based on the decomposition of the conditional covariance matrix into conditional standard deviations and correlations. The simplest multivariate correlation model that is nested in the other conditional correlation models, is the Constant Conditional Correlation (CCC-) GARCH model of Bollerslev (1990). In this model, the conditional correlation matrix is time-invariant, so the conditional covariance matrix can be expressed as follows:

$$\mathbf{H}_t = \mathbf{D}_t\mathbf{P}\mathbf{D}_t \tag{12}$$

where $\mathbf{D}_t = \text{diag}(h_{1t}^{1/2}, \dots, h_{Nt}^{1/2})$ and $\mathbf{P} = [\rho_{ij}]$ is positive definite with $\rho_{ii} = 1$, $i = 1, \dots, N$. This means that the off-diagonal elements of the conditional covariance matrix are defined as follows:

$$[\mathbf{H}_t]_{ij} = h_{it}^{1/2} h_{jt}^{1/2} \rho_{ij}, \quad i \neq j$$

where $1 \leq i, j \leq N$. The models for the processes $\{r_{it}\}$ are members of the class of univariate GARCH models. They are most often modelled as the GARCH(p, q) model, in which case the conditional variances can be written in a vector form

$$\mathbf{h}_t = \boldsymbol{\omega} + \sum_{j=1}^q \mathbf{A}_j \mathbf{r}_{t-j}^{(2)} + \sum_{j=1}^p \mathbf{B}_j \mathbf{h}_{t-j} \quad (13)$$

where $\boldsymbol{\omega}$ is $N \times 1$ vector, \mathbf{A}_j and \mathbf{B}_j are diagonal $N \times N$ matrices, and $\mathbf{r}_t^{(2)} = \mathbf{r}_t \odot \mathbf{r}_t$. When the conditional correlation matrix \mathbf{P} is positive definite and the elements of $\boldsymbol{\omega}$ and the diagonal elements of \mathbf{A}_j and \mathbf{B}_j positive, the conditional covariance matrix \mathbf{H}_t is positive definite. Positivity of the diagonal elements of \mathbf{A}_j and \mathbf{B}_j is not, however, necessary for \mathbf{P} to be positive definite unless $p = q = 1$, see Nelson and Cao (1992) for discussion of positivity conditions for h_{it} in univariate GARCH(p, q) models.

An extension to the CCC-GARCH model was introduced by Jeantheau (1998). In this Extended CCC- (ECCC-) GARCH model the assumption that the matrices \mathbf{A}_j and \mathbf{B}_j in (13) are diagonal is relaxed. This allows the past squared returns and variances of all series to enter the individual conditional variance equations. For instance, in the first-order ECCC-GARCH model, the i th variance equation is

$$h_{it} = \omega_i + a_{11} r_{1,t-1}^2 + \dots + a_{1N} r_{N,t-1}^2 + b_{11} h_{1,t-1} + \dots + b_{1N} h_{N,t-1}, \\ i = 1, \dots, N.$$

An advantage of this extension is that it allows a considerably richer autocorrelation structure for the squared observed returns than the standard CCC-GARCH model. For example, in the univariate GARCH(1,1) model the autocorrelations of the squared observations decrease exponentially from the first lag. In the first-order ECCC-GARCH model, the same autocorrelations need not have a monotonic decline from the first lag. This has been shown by He and Teräsvirta (2004) who considered the fourth-moment structure of first- and second-order ECCC-GARCH models.

The estimation of MGARCH models with constant correlations is computationally attractive. Because of the decomposition (12), the log-likelihood in (3) has the following simple form:

$$\sum_{t=1}^T \ell_t(\boldsymbol{\theta}) = c - (1/2) \sum_{t=1}^T \sum_{i=1}^N \ln |h_{it}| - (1/2) \sum_{t=1}^T \log |\mathbf{P}| \\ - (1/2) \sum_{t=1}^T \mathbf{r}'_t \mathbf{D}_t^{-1} \mathbf{P}^{-1} \mathbf{D}_t^{-1} \mathbf{r}_t. \quad (14)$$

From (14) it is apparent that during estimation, one has to invert the conditional correlation matrix only once per iteration. The number of parameters in the CCC- and ECCC-GARCH models, in addition to the ones in the univariate GARCH equations, equals $N(N-1)/2$ and covariance stationarity is ensured if the roots of $\det(\mathbf{I} - \sum_{j=1}^q \mathbf{A}_j \lambda^j - \sum_{j=1}^p \mathbf{B}_j \lambda^j) = 0$ lie outside the unit circle.

Although the CCC-GARCH model is in many respects an attractive parameterization, empirical studies have suggested that the assumption of constant conditional correlations may be

too restrictive. The model may therefore be generalized by retaining the previous decomposition but making the conditional correlation matrix in (12) time-varying. Thus,

$$\mathbf{H}_t = \mathbf{D}_t \mathbf{P}_t \mathbf{D}_t. \quad (15)$$

In conditional correlation models defined through (15), positive definiteness of \mathbf{H}_t follows if, in addition to the conditional variances h_{it} , $i = 1, \dots, N$, being well-defined, the conditional correlation matrix \mathbf{P}_t is positive definite at each point in time. Compared to the CCC–GARCH models, the advantage of numerically simple estimation is lost, as the correlation matrix has to be inverted for each t during every iteration.

Due to the intuitive interpretation of correlations, there exist a vast number of proposals for specifying \mathbf{P}_t . Tse and Tsui (2002) imposed GARCH type of dynamics on the conditional correlations. The conditional correlations in their Varying Correlation (VC–) GARCH model are functions of the conditional correlations of the previous period and a set of estimated correlations. More specifically,

$$\mathbf{P}_t = (1 - a - b)\mathbf{S} + a\mathbf{S}_{t-1} + b\mathbf{P}_{t-1}$$

where \mathbf{S} is a constant, positive definite parameter matrix with ones on the diagonal, a and b are non-negative scalar parameters such that $a + b \leq 1$, and \mathbf{S}_{t-1} is a sample correlation matrix of the past M standardized residuals $\hat{\boldsymbol{\varepsilon}}_{t-1}, \dots, \hat{\boldsymbol{\varepsilon}}_{t-M}$ where $\hat{\boldsymbol{\varepsilon}}_{t-j} = \widehat{\mathbf{D}}_{t-j}^{-1} \mathbf{r}_{t-j}$, $j = 1, \dots, M$. The positive definiteness of \mathbf{P}_t is ensured by construction if \mathbf{P}_0 and \mathbf{S}_{t-1} are positive definite. A necessary condition for the latter to hold is $M \geq N$. The definition of the ‘intercept’ $1 - a - b$ corresponds to the idea of ‘variance targeting’ in Engle and Mezrich (1996).

Kwan, Li, and Ng (in press) proposed a threshold extension to the VC–GARCH model. Within each regime, indicated by the value of an indicator or threshold variable, the model has a VC–GARCH specification. Specifically, the authors partition the real line into R subintervals, $r_0 = -\infty < l_1 < \dots < l_{R-1} < l_R = \infty$, and define an indicator variable $s_t \in \mathcal{F}_{t-1}^*$, the extended information set. The r th regime is defined by $l_{r-1} < s_t \leq l_r$, and both the univariate GARCH models and the dynamic correlations have regime-specific parameters. Kwan, Li, and Ng (in press) also apply the same idea to the BEKK model and discuss estimation of the number of regimes. In order to estimate the model consistently, one has to make sure that each regime contains a sufficient number of observations.

Engle (2002) introduced a Dynamic Conditional Correlation (DCC–) GARCH model whose dynamic conditional correlation structure is similar to that of the VC–GARCH model. Engle considered a dynamic matrix process

$$\mathbf{Q}_t = (1 - a - b)\mathbf{S} + a\boldsymbol{\varepsilon}_{t-1}\boldsymbol{\varepsilon}'_{t-1} + b\mathbf{Q}_{t-1}$$

where a is a positive and b a non-negative scalar parameter such that $a + b < 1$, \mathbf{S} is the unconditional correlation matrix of the standardized errors $\boldsymbol{\varepsilon}_t$, and \mathbf{Q}_0 is positive definite. This process ensures positive definiteness but does not generally produce valid correlation matrices. They are obtained by rescaling \mathbf{Q}_t as follows:

$$\mathbf{P}_t = (\mathbf{I} \odot \mathbf{Q}_t)^{-1/2} \mathbf{Q}_t (\mathbf{I} \odot \mathbf{Q}_t)^{-1/2}.$$

Both the VC– and the DCC–GARCH model extend the CCC–GARCH model, but do it with few extra parameters. In each correlation equation, the number of parameters is $N(N-1)/2 + 2$ for the VC–GARCH model and two for in the DCC–GARCH one. This is a strength of these models but may also be seen as a weakness when N is large, because all $N(N-1)/2$ correlation processes are restricted to have the same dynamic structure.

To avoid this limitation, various generalizations of the DCC–GARCH model have been proposed. Billio and Caporin (2006) suggested a model that imposes a BEKK structure on the conditional correlations. Their Quadratic Flexible DCC (GFDDCC) GARCH model where the matrix process Q_t is defined as

$$Q_t = C'SC + A'\varepsilon_{t-1}\varepsilon'_{t-1}A + B'Q_{t-1}B \quad (16)$$

where the matrices A , B , and C are symmetric, and S is the unconditional covariance matrix of the standardized errors ε_t . To obtain stationarity, $C'SC$ has to be positive definite and the eigenvalues of $A + B$ are less than one in modulus. The number of parameters governing the correlations in the GFDDCC–GARCH model in its fully general form is $3N(N + 1)/2$ which is unfeasible in large systems. The authors therefore suggested several special cases: One is to group the assets according to their properties, sector, or industry and restricting the coefficient matrices to be block diagonal following the partition. Another is to restrict the coefficient matrices to be diagonal with possibly suitable partition.

Cappiello, Engle, and Sheppard (2006) generalized the DCC–GARCH model in a similar manner, but also including asymmetric effects. In their Asymmetric Generalized DCC (AG–DCC) GARCH model the dynamics of Q_t is the following:

$$Q_t = (S - A'SA - B'SB - G'S^-G) + A'\varepsilon_{t-1}\varepsilon'_{t-1}A + B'Q_{t-1}B + G'\varepsilon_{t-1}^-\varepsilon_{t-1}^-G \quad (17)$$

where A , B , and G are $N \times N$ parameter matrices, $\varepsilon^- = \mathbb{I}_{\{\varepsilon_t < 0\}} \odot \varepsilon_t$, where \mathbb{I} is an indicator function, and S and S^- are the unconditional covariance matrices of ε_t and ε_t^- , respectively. Again, the number of parameters increases rapidly with the dimension of the model, and restricted versions, such as diagonal, scalar, and symmetric, were suggested.

In the VC–GARCH as well as the DCC–GARCH model, the dynamic structure of the time-varying correlations is a function of past returns. There is another class of models that allows the dynamic structure of the correlations to be controlled by an exogenous variable. This variable may be either an observable variable, a combination of observable variables, or a latent variable that represents factors that are difficult to quantify. One may argue that these models are not pure vector GARCH models because the conditioning set in them can be larger than in VC–GARCH or DCC–GARCH models. The first one of these models to be considered here is the Smooth Transition Conditional Correlation (STCC–) GARCH model.

In the STCC–GARCH model of Silvennoinen and Teräsvirta (2005), the conditional correlation matrix varies smoothly between two extreme states according to a transition variable. The following dynamic structure is imposed on the conditional correlations:

$$P_t = (1 - G(s_t))P_{(1)} + G(s_t)P_{(2)}$$

where $P_{(1)}$ and $P_{(2)}$, $P_{(1)} \neq P_{(2)}$, are positive definite correlation matrices that describe the two extreme states of correlations, and $G(\cdot) : \mathbb{R} \rightarrow (0, 1)$, is a monotonic function of an observable transition variable $s_t \in \mathcal{F}_{t-1}^*$. The authors define $G(\cdot)$ as the logistic function

$$G(s_t) = \left(1 + e^{-\gamma(st-c)}\right)^{-1}, \quad \gamma > 0 \quad (18)$$

where the parameter γ determines the velocity and c the location of the transition. In addition to the univariate variance equations, the STCC–GARCH model has $N(N - 1) + 2$ parameters. The sequence $\{P_t\}$ is a sequence of positive definite matrices because each P_t is a convex

combination of two positive definite correlation matrices. The transition variable s_t is chosen by the modeller to suit the application at hand. If there is uncertainty about an appropriate choice of s_t , testing the CCC–GARCH model can be used as tool for judging the relevance of a given transition variable to the dynamic conditional correlations. A special case of the STCC–GARCH model is obtained when the transition variable is calendar time. The Time Varying Conditional Correlation (TVCC–) GARCH model was in its bivariate form introduced by Berben and Jansen (2005).

A recent extension of the STCC–GARCH model, the Double Smooth Transition Conditional Correlation (DSTCC–) GARCH model by Silvennoinen and Teräsvirta (2007) allows for another transition around the first one:

$$\begin{aligned} \mathbf{P}_t = & (1 - G_2(s_{2t})) \{ (1 - G_1(s_{1t})) \mathbf{P}_{(11)} + G_1(s_{1t}) \mathbf{P}_{(21)} \} \\ & + G_2(s_{2t}) \{ (1 - G_1(s_{1t})) \mathbf{P}_{(12)} + G_1(s_{1t}) \mathbf{P}_{(22)} \}. \end{aligned} \quad (19)$$

For instance, one of the transition variables can simply be calendar time. If this is the case, one has the Time Varying Smooth Transition Conditional Correlation (TVSTCC–) GARCH model that nests the STCC–GARCH as well as the TVCC–GARCH model. The interpretation of the extreme states is the following: At the beginning of the sample, $\mathbf{P}_{(11)}$ and $\mathbf{P}_{(21)}$ are the two extreme states between which the correlations vary according to the transition variable s_{1t} and similarly, $\mathbf{P}_{(12)}$ and $\mathbf{P}_{(22)}$ are the corresponding states at the end of the sample. The TVSTCC–GARCH model allows the extreme states, constant in the STCC–GARCH framework, to be time-varying, which introduces extra flexibility when modelling long time series. The number of parameters, excluding the univariate GARCH equations, is $2N(N - 1) + 4$ which restricts the use of the model in very large systems.

The Regime Switching Dynamic Correlation (RSDC–) GARCH model introduced by Pelletier (2006) falls somewhere between the models with constant correlations and the ones with correlations changing continuously at every period. The model imposes constancy of correlations within a regime while the dynamics enter through switching regimes. Specifically,

$$\mathbf{P}_t = \sum_{r=1}^R \mathbb{I}_{\{\Delta_t=r\}} \mathbf{P}_{(r)}$$

where Δ_t is a (usually first-order) Markov chain independent of $\boldsymbol{\eta}_t$ that can take R possible values and is governed by a transition probability matrix $\boldsymbol{\Pi}$, \mathbb{I} is the indicator function, and $\mathbf{P}_{(r)}$, $r = 1, \dots, R$, are positive definite regime-specific correlation matrices. Correlation component of the model has $RN(N - 1)/2 - R(R - 1)$ parameters. A version that involves fewer parameters is obtained by restricting the R possible states of correlations to be linear combinations of a state of zero correlations and that of possibly high correlations. Thus,

$$\mathbf{P}_t = (1 - \lambda(\Delta_t)) \mathbf{I} + \lambda(\Delta_t) \mathbf{P}$$

where \mathbf{I} is the identity matrix (‘no correlations’), \mathbf{P} is a correlation matrix representing the state of possibly high correlations, and $\lambda(\cdot) : \{1, \dots, R\} \rightarrow [0, 1]$ is a monotonic function of Δ_t . The number of regimes R is not a parameter to be estimated. The conditional correlation matrices are positive definite at each point in time by construction both in the unrestricted and restricted version of the model. If N is not very small, Pelletier (2006) recommends two-step estimation. First estimate the parameters of the GARCH equations and, second, conditionally on these estimates, estimate the correlations and the switching probabilities using the EM algorithm of Dempster, Laird, and Rubin (1977).

2.4 Nonparametric and semiparametric approaches

Non- and semiparametric models form an alternative to parametric estimation of the conditional covariance structure. These approaches have the advantage of not imposing a particular (possibly misspecified) structure on the data. One advantage of at least a few fully parametric multivariate GARCH models is, however, that they offer an interpretation of the dynamic structure of the conditional covariance or correlation matrices. Another is that the quasi-maximum likelihood estimator is consistent when the errors are assumed multivariate normal. However, there may be considerable efficiency losses in finite samples if the returns are not normally distributed. Semiparametric models combine the advantages of a parametric model in that they reach consistency and sustain the interpretability, and those of a nonparametric model which is robust against distributional misspecification. Nonparametric models, however, suffer from the ‘curse of dimensionality’: due to the lack of data in all directions of the multidimensional space, the performance of the local smoothing estimator deteriorates quickly as the dimension of the conditioning variable increases, see Stone (1980). For this reason, it has been of interest to study methods for dimension-reduction or to use a single, one-dimensional conditioning variable. Developments in semi- and nonparametric modelling are discussed in detail in a separate chapter of this Handbook, see Linton (2008).

One alternative is to specify a parametric model for the conditional covariance structure but estimate the error distribution nonparametrically, thereby attempting to offset the efficiency loss of the quasi-maximum likelihood estimator compared to the maximum likelihood estimator of the correctly specified model. In the semiparametric model of Hafner and Rombouts (2007) the data are generated by any particular parametric MGARCH model and the error distribution is unspecified but estimated nonparametrically. Their approach leads to the log-likelihood

$$\sum_{t=1}^T \ell_t(\boldsymbol{\theta}) = c - (1/2) \sum_{t=1}^T \ln |\mathbf{H}_t| + \sum_{t=1}^T \ln g(\mathbf{H}_t^{-1/2} \mathbf{r}_t) \quad (20)$$

where $g(\cdot)$ is an unspecified density function of the standardized residuals $\boldsymbol{\eta}_t$ such that $E[\boldsymbol{\eta}_t] = \mathbf{0}$ and $E[\boldsymbol{\eta}_t \boldsymbol{\eta}_t'] = \mathbf{I}$. This model may be seen as a multivariate extension of the semiparametric GARCH model by Engle and González-Rivera (1991). A flexible error distribution blurs the line between the parametric structure and the distribution of the errors. For example, if the correlation structure of a semiparametric GARCH model is misspecified, a nonparametric error distribution may absorb some of the misspecification. The nonparametric method for estimating the density g is discussed in detail in Hafner and Rombouts (2007). They assume that g belongs to the class of spherical distributions. Even with this restriction their semiparametric estimator remains more efficient than the maximum likelihood estimator if the errors z_t are non-normal.

Long and Ullah (2005) introduce an approach similar to the previous one in that the model is based on any parametric MGARCH model. After estimating a parametric model, the estimated standardized residuals $\hat{\boldsymbol{\eta}}_t$ are extracted. When the model is not correctly specified, these residuals may have some structure in them, and Long and Ullah (2005) use nonparametric estimation to extract this information. This is done by estimating the conditional covariance matrix using the Nadaraya-Watson estimator

$$\mathbf{H}_t = \widehat{\mathbf{H}}_t^{1/2} \frac{\sum_{\tau=1}^T \hat{\boldsymbol{\eta}}_{\tau} \hat{\boldsymbol{\eta}}_{\tau}' K_h(s_{\tau} - s_t)}{\sum_{\tau=1}^T K_h(s_{\tau} - s_t)} \widehat{\mathbf{H}}_t^{1/2} \quad (21)$$

where $\widehat{\mathbf{H}}_t$ is the conditional covariance matrix estimated parametrically from an MGARCH model, $s_t \in \mathcal{F}_{t-1}^*$ is an observable variable that the model is conditioned on, $\hat{\boldsymbol{\epsilon}}_t = \widehat{\mathbf{D}}_t^{-1} \mathbf{r}_t$,

$K_h(\cdot) = K(\cdot/h)/h$, $K(\cdot)$ is a kernel function, and h is the bandwidth parameter. Positive definiteness of $\widehat{\mathbf{H}}_t$ ensures positive definiteness of the semiparametric estimator \mathbf{H}_t .

In the Semi-Parametric Conditional Correlation (SPCC-) GARCH model of Hafner, van Dijk, and Franses (2005), the conditional variances are modelled parametrically by any choice of univariate GARCH model, where $\hat{\boldsymbol{\varepsilon}}_t = \widehat{\mathbf{D}}_t^{-1} \mathbf{r}_t$ is the vector consisting of the standardized residuals. The conditional correlations \mathbf{P}_t are then estimated using a transformed Nadaraya-Watson estimator:

$$\mathbf{P}_t = (\mathbf{I} \odot \mathbf{Q}_t)^{-1/2} \mathbf{Q}_t (\mathbf{I} \odot \mathbf{Q}_t)^{-1/2}$$

where

$$\mathbf{Q}_t = \frac{\sum_{\tau=1}^T \hat{\boldsymbol{\varepsilon}}_\tau \hat{\boldsymbol{\varepsilon}}_\tau' K_h(s_\tau - s_t)}{\sum_{\tau=1}^T K_h(s_\tau - s_t)}. \quad (22)$$

In (22), $s_t \in \mathcal{F}_{t-1}^*$ is a conditioning variable, $K_h(\cdot) = K(\cdot/h)/h$, $K(\cdot)$ is a kernel function, and h is the bandwidth parameter.

Long and Ullah (2005) also suggest estimating the covariance structure in a fully nonparametric fashion so that the model is not an MGARCH model, but merely a parameter-free multivariate volatility model. The estimator of the conditional covariance matrix is

$$\mathbf{H}_t = \frac{\sum_{\tau=1}^T \mathbf{r}_\tau \mathbf{r}_\tau' K_h(s_\tau - s_t)}{\sum_{\tau=1}^T K_h(s_\tau - s_t)}$$

where $s_t \in \mathcal{F}_{t-1}^*$ is a conditioning variable, $K_h(\cdot) = K(\cdot/h)/h$, $K(\cdot)$ is a kernel function, and h is the bandwidth parameter. This approach ensures positive definiteness of \mathbf{H}_t .

The choice of the kernel function is not important and it could be any probability density function, whereas the choice of the bandwidth parameter h is crucial, see for instance Pagan and Ullah (1999, Sections 2.4.2 and 2.7). Long and Ullah (2005) consider the choice of an optimal fixed bandwidth, whereas Hafner, van Dijk, and Franses (2005) discuss a way of choosing a dynamic bandwidth parameter such that the bandwidth is larger in the tails of the marginal distribution of the conditioning variable s_t than it is in the mid-region of the distribution.

3 Statistical Properties

Statistical properties of multivariate GARCH models are only partially known. For the development of statistical estimation and testing theory, it would be desirable to have conditions for strict stationarity and ergodicity of a multivariate GARCH process, as well as conditions for consistency and asymptotic normality of the quasi-maximum likelihood estimator. The results that are available establish these properties in special cases and sometimes under strong conditions.

Jeantheau (1998) considers the statistical properties and estimation theory of the ECCC-GARCH model he proposes. He provides sufficient conditions for the existence of a weakly stationary and ergodic solution, which is also strictly stationary. This is done by assuming $E\mathbf{r}_t \mathbf{r}_t' < \infty$. It would be useful to have both a necessary and a sufficient condition for the existence of a strictly stationary solution, but this question remains open. Jeantheau (1998) also proves the strong consistency of the QML estimator for the ECCC-GARCH model. Ling and McAleer (2003) complement Jeantheau's results and also prove the asymptotic normality of the QMLE in the case of the ECCC-GARCH model. For the global asymptotic normality result, the existence of the sixth moment of \mathbf{r}_t is required. The statistical properties of the second-order model are also investigated in He and Teräsvirta (2004), who provide sufficient conditions for

the existence of fourth moments, and, furthermore, give expressions for the fourth moment as well as the autocorrelation function of squared observations as functions of the parameters.

Comte and Lieberman (2003) study the statistical properties of the BEKK model. Relying on a result in Boussama (1998), they give sufficient, but not necessary conditions for strict stationarity and ergodicity. Applying Jeantheau's results, they provide conditions for the strong consistency of the QMLE. Furthermore, they also prove the asymptotic normality of the QMLE, for which they assume the existence of the eighth moment of \mathbf{r}_t . The fourth-moment structure of the BEKK and VEC models is investigated by Hafner (2003), who gives necessary and sufficient conditions for the existence of the fourth moments and provides expressions for them. These expressions are not functions of the parameters of the model. As the factor models listed in Section 2.2 are special cases of the BEKK model, the results of Comte and Lieberman (2003) and Hafner (2003) also apply to them.

4 Hypothesis testing in multivariate GARCH models

Testing the adequacy of estimated models is an important part of model building. Existing tests of multivariate GARCH models may be divided into two broad categories: general misspecification tests and specification tests. The purpose of the tests belonging to the former category is to check the adequacy of an estimated model. Specification tests are different in the sense that they are designed to test the model against a parametric extension. Such tests have been constructed for the CCC-GARCH model, but obviously not for other models. We first review general misspecification tests.

4.1 General misspecification tests

Ling and Li (1997) derived a rather general misspecification test for multivariate GARCH models. It is applicable for many families of GARCH models. The test statistic has the following form:

$$Q(k) = T\boldsymbol{\gamma}'_k\widehat{\boldsymbol{\Omega}}_k^{-1}\boldsymbol{\gamma}_k \quad (23)$$

where $\boldsymbol{\gamma}_k = (\gamma_1, \dots, \gamma_k)'$ with

$$\gamma_j = \frac{\sum_{t=j+1}^T (\mathbf{r}'_t\widehat{\mathbf{H}}_t^{-1}\mathbf{r}_t - N)(\mathbf{r}'_{t-j}\widehat{\mathbf{H}}_{t-j}^{-1}\mathbf{r}_{t-j} - N)}{\sum_{t=1}^T (\mathbf{r}'_t\widehat{\mathbf{H}}_t^{-1}\mathbf{r}_t - N)^2} \quad (24)$$

$j = 1, \dots, k$, $\widehat{\mathbf{H}}_t$ is an estimator of \mathbf{H}_t , and $\widehat{\boldsymbol{\Omega}}_k$ is the estimated covariance matrix of $\boldsymbol{\gamma}_k$, see Ling and Li (1997) for details. Under the null hypothesis that the GARCH model is correctly specified, that is, $\boldsymbol{\eta}_t \sim \text{IID}(\mathbf{0}, \mathbf{I})$, statistic (23) has an asymptotic χ^2 distribution with k degrees of freedom. Under H_0 , $E\mathbf{r}'_t\widehat{\mathbf{H}}_t^{-1}\mathbf{r}_t = N$, and therefore the expression (24) is the j th-order sample autocorrelation between $\mathbf{r}'_t\widehat{\mathbf{H}}_t^{-1}\mathbf{r}_t = \boldsymbol{\eta}'_t\boldsymbol{\eta}_t$ and $\mathbf{r}'_{t-j}\widehat{\mathbf{H}}_{t-j}^{-1}\mathbf{r}_{t-j} = \boldsymbol{\eta}'_{t-j}\boldsymbol{\eta}_{t-j}$. The test may thus be viewed as a generalization of the portmanteau test of Li and Mak (1994) for testing the adequacy of a univariate GARCH model. In fact, when $N = 1$, (23) collapses into the Li and Mak statistic. The McLeod and Li (1983) statistic (Ljung-Box statistic applied to squared residuals), frequently used for evaluating GARCH models, is valid neither in the univariate nor in the multivariate case, see Li and Mak (1994) for the univariate case.

A simulation study by Tse and Tsui (1999) indicates that the Ling and Li portmanteau statistic (24) often has low power. The authors show examples of situations in which a portmanteau test based on autocorrelations of pairs of individual standardized residuals performs

better. The drawback of this statistic is, however, that its asymptotic null distribution is unknown, and the statistic tends to be undersized. Each test is based only on a single pair of residuals.

Duchesne (2004) introduced the test which is a direct generalization of the portmanteau test of Li and Mak (1994) to the VEC–GARCH model (2). Let $\hat{\boldsymbol{\eta}}_t$ denote the maximum likelihood estimator of the error vector $\boldsymbol{\eta}_t$ in the VEC–GARCH model. The idea is to derive the asymptotic distribution of $\hat{\boldsymbol{c}}_j = \text{vech}(\hat{\boldsymbol{\eta}}_t \hat{\boldsymbol{\eta}}'_{t-j})$ where $j = 1, \dots, k$, under the null hypothesis that $\{\boldsymbol{\eta}_t\} \sim \text{NID}(\mathbf{0}, \mathbf{I})$. Once this has been done, one can combine the results and obtain the asymptotic null distribution of $\hat{\boldsymbol{c}}_{(k)} = (\hat{\boldsymbol{c}}'_1, \dots, \hat{\boldsymbol{c}}'_k)'$, where the vectors $\hat{\boldsymbol{c}}_j$, $j = 1, \dots, k$, are asymptotically uncorrelated when the null holds. This distribution is normal since the asymptotic distribution of each $\hat{\boldsymbol{c}}_k$ is normal. It follows that under the null hypothesis,

$$Q_D(k) = T \hat{\boldsymbol{c}}'_{(k)} \widehat{\boldsymbol{\Omega}}_k^{-1} \hat{\boldsymbol{c}}_{(k)} \xrightarrow{d} \chi^2(kN(N+1)/2) \quad (25)$$

where $\widehat{\boldsymbol{\Omega}}_k$ is a consistent estimator of the covariance matrix of $\hat{\boldsymbol{c}}_{(k)}$, defined in Duchesne (2004). This portmanteau test statistic collapses into the statistic of Li and Mak (1994) when $N = 1$. When $\{\boldsymbol{\eta}_t\} = \{\boldsymbol{\varepsilon}_t\}$, that is, when $\mathbf{H}_t \equiv \sigma^2 \mathbf{I}$, the test (25) is a test of no multivariate ARCH. For $N = 1$, it is then identical to the well known portmanteau test of McLeod and Li (1983).

Yet another generalization of univariate tests can be found in Kroner and Ng (1998). Their misspecification tests are suitable for any multivariate GARCH model. Let

$$\mathbf{G}_t = \mathbf{r}_t \mathbf{r}'_t - \widehat{\mathbf{H}}_t$$

where $\widehat{\mathbf{H}}_t$ has been estimated from a GARCH model. The elements of $\mathbf{G}_t = [g_{ijt}]$ are ‘generalized residuals’. When the model is correctly specified, they form a matrix of martingale difference sequences with respect to the information set \mathcal{F}_{t-1} that contains the past information until $t-1$. Thus any variable $x_s \in \mathcal{F}_{t-1}$ is uncorrelated with the elements of \mathbf{G}_t . Tests based on these misspecification indicators may now be constructed. This is done for each g_{ijt} separately. The suggested tests are generalizations of the sign-bias and size-bias tests of Engle and Ng (1993). The test statistics have an asymptotic χ^2 distribution with one degree of freedom when the null hypothesis is valid. If the dimension of the model is large and there are several misspecification indicators, the number of available tests may be very large.

Testing the adequacy of the CCC–GARCH model has been an object of interest since it was found that the assumption of constant correlations may sometimes be too restrictive in practice. Tse (2000) constructed a Lagrange multiplier (LM) test of the CCC–GARCH model against the following alternative, \mathbf{P}_t , to constant correlations:

$$\mathbf{P}_t = \mathbf{P} + \boldsymbol{\Delta} \odot \mathbf{r}_{t-1} \mathbf{r}'_{t-1} \quad (26)$$

where $\boldsymbol{\Delta}$ is a symmetric parameter matrix with the main diagonal elements equal to zero. This means that the correlations are changing as functions of the previous observations. The null hypothesis is $H_0 : \boldsymbol{\Delta} = \mathbf{0}$ or, expressed as a vector equation, $\text{vecl}(\boldsymbol{\Delta}) = 0$.¹ Equation (26) does not define a particular alternative to conditional correlations as \mathbf{P}_t is not necessarily a positive definite matrix for every t . For this reason we interpret the test as a general misspecification test.

Bera and Kim (2002) present a test of a bivariate CCC–GARCH model against the alternative that the correlation coefficient is stochastic. The test is an Information Matrix test and as such an LM or score test. It is designed for a bivariate model, which restricts its usefulness in applications.

¹The operator $\text{vecl}(\cdot)$ stacks the columns of the strictly lower triangular part (excluding main diagonal elements) of its argument matrix.

4.2 Tests for extensions of the CCC–GARCH model

The most popular extension of the CCC–GARCH model to-date is the DCC–GARCH model of Engle (2002). However, there does not seem to be any published work on developing tests of constancy of correlations directly against this model.

As discussed in Section 2.3, Silvennoinen and Teräsvirta (2005) extend the CCC–GARCH into a STCC–GARCH model in which the correlations fluctuate according to a transition variable. They construct an LM test for testing the constant correlation hypothesis against the smoothly changing correlations. Since the STCC–GARCH model is only identified when the correlations are changing, standard asymptotic theory is not valid. A good discussion of this problem can be found in Hansen (1996). The authors apply the technique in Luukkonen, Saikkonen, and Teräsvirta (1988) in order to circumvent the identification problem. The null hypothesis is $\gamma = 0$ in (18), and a linearization of the correlation matrix \mathbf{P}_t by the first-order Taylor expansion of (18) yields

$$\mathbf{P}_t^* = \mathbf{P}_{(1)} - s_t \mathbf{P}_{(2)}^*.$$

Under H_0 , $\mathbf{P}_{(2)}^* = \mathbf{0}$ and the correlations are thus constant. The authors use this fact to build their LM-type test on the transformed null hypothesis H'_0 : $\text{vecl}(\mathbf{P}_{(2)}^*) = \mathbf{0}$ (the diagonal elements of $\mathbf{P}_{(2)}^*$ equal zero by definition). When H'_0 holds, the test statistic has an asymptotic χ^2 distribution with $N(N - 1)/2$ degrees of freedom. The authors also derive tests for the constancy hypothesis under the assumption that some of the correlations remain constant also under the alternative. Silvennoinen and Teräsvirta (2007) extend the Taylor expansion based test to the situation where the STCC–GARCH model is the null model and the alternative is the DSTCC–GARCH model. This test collapses into the test of the CCC–GARCH model against STCC–GARCH model when $G_1(s_{1t}) \equiv 1/2$ in (19).

5 An application

In this section we compare some of the multivariate GARCH models considered in previous sections by fitting them to the same data set. In order to keep the comparison transparent, we only consider bivariate models. Our observations are the daily returns of S&P 500 index futures and 10-year bond futures from January 1990 to August 2003. This data set has been analyzed by Engle and Colacito (2006).² There is no consensus in the literature about how stock and long term bond returns are related. Historically, the long-run correlations have been assumed constant, an assumption that has led to contradicting conclusions because evidence for both positive and negative correlation has been found over the years (short-run correlations have been found to be affected, among other things, by news announcements). From a theoretical point of view, the long-run correlation between the two should be state-dependent, driven by macroeconomic factors such as growth, inflation, and interest rates. The way the correlations respond to these factors may, however, change over time.

For this reason it is interesting to see what the correlations between the two asset returns obtained from the models are and how they fluctuate over time. The focus of reporting results will therefore be on conditional correlations implied by the estimated models, that is, the BEKK, GOF–, DCC–, DSTCC–, and SPCC–GARCH ones. In the last three models, the individual GARCH equations are simply symmetric first-order ones. The BEKK model is also of order one

²The data set in Engle and Colacito (2006) begins in August 1988, but our sample starts from January 1990 because we also use the time series for a volatility index that is available only from that date onwards.

with $K = 1$. All computations have been performed using Ox, version 4.02, see Doornik (2002), and our own source code.

Estimation of the BEKK model turned out to be cumbersome. Convergence problems were encountered in numerical algorithms, but the iterations seemed to suggest diagonality of the coefficient matrices \mathbf{A} and \mathbf{B} . A diagonal BEKK model was eventually estimated without difficulty.

In the estimation of the GOF–GARCH model it is essential to obtain good initial estimates of the parameters; for details, see Lanne and Saikkonen (2007). Having done that, we experienced no difficulties in the estimation of this model with a single factor. Similarly, no convergence problems were encountered in the estimation of the DCC model of Engle (2002).

The DSTCC–GARCH model makes use of two transition variables. Because the DSTCC framework allows one to test for relevance of a variable, or variables, to the description of the dynamic structure of the correlations, we relied on the tests in Silvennoinen and Teräsvirta (2005,2007), described in Section 4.2, to select relevant transition variables. Out of a multitude of variables, including both exogenous ones and variables constructed from the past observations, prices or returns, the Chicago Board Options Exchange volatility index (VIX) that represents the market expectations of 30-day volatility turned out to lead to the strongest rejection of the null hypothesis, measured by the p -value. Calendar time seemed to be another obvious transition variable. As a result, the first-order TVSTCC–GARCH model was fitted to the bivariate data.

The semiparametric model of Hafner, van Dijk, and Franses (2005) also requires a choice of an indicator variable. Because the previous test results indicated that VIX is informative about the dynamics of the correlations, we chose VIX as the indicator variable. The SPCC–GARCH model was estimated using a standard kernel smoother with an optimal fixed bandwidth, see Pagan and Ullah (1999, Sections 2.4.2 and 2.7) for discussion on the choice of constant bandwidth.

The estimated conditional correlations are presented in Figure 1, whereas Table 1 shows the sample correlation matrix of the estimated time-varying correlations. The correlations from the diagonal BEKK model and the DCC–GARCH model are very strongly positively correlated, which is also obvious from Figure 1. The second-highest correlation of correlations is the one between the SPCC–GARCH and the GOF–GARCH model. The time-varying correlations are mostly positive during the 1990’s and negative after the turn of the century. In most models, correlations seem to fluctuate quite randomly, but the TVSTCC–GARCH model constitutes an exception. This is due to the fact that one of the transition variables is calendar time. Interestingly, in the beginning of the period the correlation between the S&P 500 and bond futures is only mildly affected by the expected volatility (VIX) and remains positive. Towards the end, not only does the correlation gradually turn negative, but expected volatility seems to affect it very strongly. Rapid fluctuations are a consequence of the fact that the transition function with VIX as the transition variable has quite a steep slope. After the turn of the century, high values of VIX generate strongly negative correlations.

Although the estimated models do not display fully identical correlations, the general message in them remains more or less the same. It is up to the user to select the model he wants to use in portfolio management and forecasting. A way of comparing the models consists of inserting the estimated covariance matrices \mathbf{H}_t , $t = 1, \dots, T$, into the Gaussian log-likelihood function (3) and calculate the maximum value of log-likelihood. These values for the estimated models appear in Table 1.

The models that are relatively easy to estimate seem to fit the data less well than the other models. The ones with a more complicated structure and, consequently, an estimation procedure that requires care, seem to attain higher likelihood values. However, the models do not make use of the same information set and, besides, they do not contain the same number of parameters.

	diag BEKK	GOF	DCC	TVSTCC	SPCC
diag BEKK	1.0000				
GOF	0.7713	1.0000			
DCC	0.9875	0.7295	1.0000		
TVSTCC	0.7577	0.7381	0.7690	1.0000	
SPCC	0.6010	0.8318	0.5811	0.7374	1.0000
log-likelihood	-6130	-6091	-6166	-6006	-6054
AIC	12275	12198	12347	12041	12120
BIC	12286	12211	12359	12062	12130

Table 1: Sample correlations of the estimated conditional correlations. The lower part of the table shows the log-likelihood values and the values of the corresponding model selection criteria.

Taking this into account suggests the use of model selection criteria for assessing the performance of the models. Nevertheless, rankings by Akaike’s information criterion (AIC) and the Bayesian information criterion (BIC) are the same as the likelihood-based ranking; see Table 1. Note that in theory, rankings based on a model selection criterion favour the SPCC model. This is because no penalty is imposed on the nonparametric correlation estimates that improve the fit compared to constant correlations.

Nonnested testing as a means of comparison is hardly a realistic option here since the computational effort would be quite substantial. Out-of-sample forecasting would be another way of comparing models. However, the models involved would be multivariate and the quantities to be forecast would be measures of (unobserved) volatilities and cross-volatilities. This would give rise to a number of problems, beginning from defining the quantities to be forecast and appropriate loss functions, and from comparing forecast vectors instead of scalar forecasts. It appears that plenty of work remains to be done in that area.

6 Final remarks

In this review, we have considered a number of multivariate GARCH models and highlighted their features. It is obvious that the original VEC model contains too many parameters to be easily applicable, and research has been concentrated on finding parsimonious alternatives to it. Two lines of development are visible. First, there are the attempts to impose restrictions on the parameters of the VEC model. The BEKK model and the factor models are examples of this. Second, there is the idea of modelling conditional covariances through conditional variances and correlations. It has led to a number of new models, and this family of conditional correlation models appears to be quite popular right now. The conditional correlation models are easier to estimate than many of their counterparts and their parameters (correlations) have a natural interpretation.

As previously discussed, there is no statistical theory covering all MGARCH models. This may be expected, since models in the two main categories differ substantially from each other. Progress has been made in some special occasions, and these cases have been considered in previous sections.

Estimation of multivariate GARCH models is not always easy. BEKK models appear more

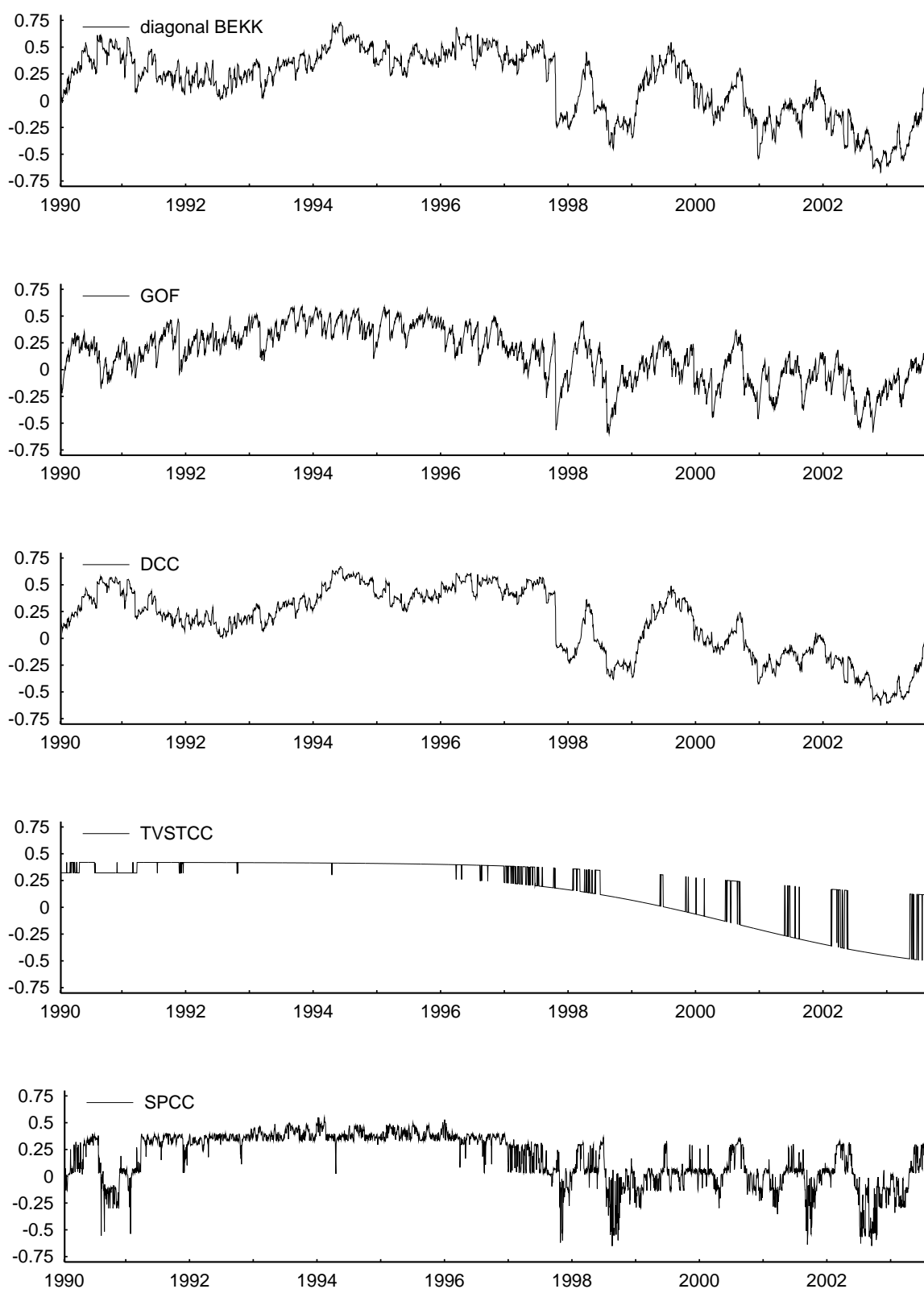


Figure 1: Conditional correlations implied by the estimated models: Diagonal BEKK, GOF-GARCH, DCC-GARCH, TVSTCC-GARCH, and SPCC-GARCH.

difficult to estimate than the CCC-GARCH model and its generalizations. While it has not been the objective of this review to cover algorithms for performing the necessary iterations, Brooks, Burke, and Persaud (2003) compared four software packages for estimating MGARCH models. They used a single bivariate dataset and only fitted a first-order VEC-GARCH model to the data. A remarkable thing is that already the parameter estimates resulting from these packages are quite different, not to mention standard deviation estimates. The estimates give rather different ideas of the persistence of conditional volatility. These differences do not necessarily say very much about properties of the numerical algorithms used in the packages. It is more likely that they reflect the estimation difficulties. The log-likelihood function may contain a large number of local maxima, and different starting-values may thus lead to different outcomes. See Silvennoinen (2008) for more discussion. The practitioner who may wish to use these models in portfolio management should be aware of these problems.

Not much has been done as yet to construct tests for evaluating MGARCH models. A few tests do exist, and a number of them have been considered in this review.

It may be that VEC and BEKK models, with the possible exception of factor models, have already matured and there is not much that can be improved. The situation may be different for conditional correlation models. The focus has hitherto been on modelling the possibly time-varying correlations. Less emphasis has been put on the GARCH equations that typically have been GARCH(1,1) specifications. Designing diagnostic tools for testing and improving GARCH equations may be one of the challenges for the future.

References

- ALEXANDER, C. O., AND A. M. CHIBUMBA (1997): "Multivariate orthogonal factor GARCH," University of Sussex Discussion Papers in Mathematics.
- BAE, K.-H., G. A. KAROLYI, AND R. M. STULZ (2003): "A new approach to measuring financial contagion," *The Review of Financial Studies*, 16, 717–763.
- BAUWENS, L., S. LAURENT, AND J. V. K. ROMBOUTS (2006): "Multivariate GARCH models: A survey," *Journal of Applied Econometrics*, 21, 79–109.
- BERA, A. K., AND S. KIM (2002): "Testing constancy of correlation and other specifications of the BGARCH model with an application to international equity returns," *Journal of Empirical Finance*, 9, 171–195.
- BERBEN, R.-P., AND W. J. JANSEN (2005): "Comovement in international equity markets: A sectoral view," *Journal of International Money and Finance*, 24, 832–857.
- BILLIO, M., AND M. CAPORIN (2006): "A generalized dynamic conditional correlation model for portfolio risk evaluation," unpublished manuscript, Ca' Foscari University of Venice, Department of Economics.
- BOLLERSLEV, T. (1990): "Modelling the coherence in short-run nominal exchange rates: A multivariate generalized ARCH model," *Review of Economics and Statistics*, 72, 498–505.
- BOLLERSLEV, T., R. F. ENGLE, AND D. B. NELSON (1994): "ARCH models," in *Handbook of Econometrics*, ed. by R. F. Engle, and D. L. McFadden, vol. 4, pp. 2959–3038. Elsevier Science, Amsterdam.

- BOLLERSLEV, T., R. F. ENGLE, AND J. M. WOOLDRIDGE (1988): “A capital asset pricing model with time-varying covariances,” *The Journal of Political Economy*, 96, 116–131.
- BOUSSAMA, F. (1998): “Ergodicité, mélange et estimation dans les modèles GARCH,” Thèse de l’Université Paris 7.
- 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. F. ENGLE, AND K. SHEPPARD (2006): “Asymmetric dynamics in the correlations of global equity and bond returns,” *Journal of Financial Econometrics*, 4, 537–572.
- CHIB, S., Y. OMORI, AND M. ASAI (2008): “Multivariate stochastic volatility,” in *Handbook of Financial Time Series*, ed. by T. G. Andersen, R. A. Davis, J.-P. Kreiss, and T. Mikosch. Springer, New York.
- COMTE, F., AND O. LIEBERMAN (2003): “Asymptotic theory for multivariate GARCH processes,” *Journal of Multivariate Analysis*, 84, 61–84.
- DEMPSTER, A. P., N. M. LAIRD, AND D. B. RUBIN (1977): “Maximum likelihood from incomplete data via the EM algorithm,” *Journal of the Royal Statistical Society*, 39, 1–38.
- DIEBOLD, F. X., AND M. NERLOVE (1989): “The dynamics of exchange rate volatility: a multivariate latent factor ARCH model,” *Journal of Applied Econometrics*, 4, 1–21.
- DOORNIK, J. A. (2002): *Object-Oriented Matrix Programming Using Ox*. Timberlake Consultants Press, 3rd edn., see also www.doornik.com.
- DUCHESNE, P. (2004): “On matricial measures of dependence in vector ARCH models with applications to diagnostic checking,” *Statistics and Probability Letters*, 68, 149–160.
- ENGLE, R. F. (1982): “Autoregressive conditional heteroscedasticity with estimates of the variance of United Kingdom inflation,” *Econometrica*, 50, 987–1006.
- (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 R. COLACITO (2006): “Testing and valuing dynamic correlations for asset allocation,” *Journal of Business and Economic Statistics*, 24, 238–253.
- ENGLE, R. F., AND G. GONZÁLEZ-RIVERA (1991): “Semiparametric ARCH models,” *Journal of Business and Economic Statistics*, 9, 345–359.
- ENGLE, R. F., C. W. J. GRANGER, AND D. KRAFT (1984): “Combining competing forecasts of inflation using a bivariate ARCH model,” *Journal of Economic Dynamics and Control*, 8, 151–165.
- ENGLE, R. F., AND K. F. KRONER (1995): “Multivariate simultaneous generalized ARCH,” *Econometric Theory*, 11, 122–150.
- ENGLE, R. F., AND J. MEZRICH (1996): “GARCH for groups,” *Risk*, 9, 36–40.

- ENGLE, R. F., AND V. K. NG (1993): “Measuring and testing the impact of news on volatility,” *Journal of Finance*, 48, 1749–78.
- ENGLE, R. F., V. K. NG, AND M. ROTHSCHILD (1990): “Asset pricing with a factor ARCH covariance structure: empirical estimates for treasury bills,” *Journal of Econometrics*, 45, 213–238.
- GOURIÉROUX, C. (1997): *ARCH Models and Financial Applications*. Springer-Verlag, New York.
- HAFNER, C. M. (2003): “Fourth moment structure of multivariate GARCH models,” *Journal of Financial Econometrics*, 1, 26–54.
- HAFNER, C. M., AND J. V. K. ROMBOUTS (2007): “Semiparametric multivariate volatility models,” *Econometric Theory*, 23, 251–280.
- HAFNER, C. M., D. VAN DIJK, AND P. H. FRANSES (2005): “Semi-parametric modelling of correlation dynamics,” in *Advances in Econometrics*, ed. by T. Fomby, C. Hill, and D. Terrell, vol. 20/A, pp. 59–103. Amsterdam: Elsevier Sciences.
- HANSEN, B. E. (1996): “Inference when a nuisance parameter is not identified under the null hypothesis,” *Econometrica*, 64, 413–430.
- HANSSON, B., AND P. HÖRDAHL (1998): “Testing the conditional CAPM using multivariate GARCH–M,” *Applied Financial Economics*, 8, 377–388.
- HE, C., AND T. TERÄSVIRTA (2004): “An extended constant conditional correlation GARCH model and its fourth-moment structure,” *Econometric Theory*, 20, 904–926.
- JEANTHEAU, T. (1998): “Strong consistency of estimators for multivariate ARCH models,” *Econometric Theory*, 14, 70–86.
- KAWAKATSU, H. (2006): “Matrix exponential GARCH,” *Journal of Econometrics*, 134, 95–128.
- KRONER, K. F., AND V. K. NG (1998): “Modeling asymmetric comovements of asset returns,” *The Review of Financial Studies*, 11, 817–844.
- KWAN, C. K., W. K. LI, AND K. NG (in press): “A multivariate threshold GARCH model with time-varying correlations,” *Econometric Reviews*.
- LANNE, M., AND P. SAIKKONEN (2007): “A multivariate generalized orthogonal factor GARCH model,” *Journal of Business and Economic Statistics*, 25, 61–75.
- LI, W. K., AND T. K. MAK (1994): “On the squared residual autocorrelations in non-linear time series with conditional heteroskedasticity,” *Journal of Time Series Analysis*, 15, 627–636.
- LING, S., AND W. K. LI (1997): “Diagnostic checking of nonlinear multivariate time series with multivariate ARCH errors,” *Journal of Time Series Analysis*, 18, 447–464.
- LING, S., AND M. MCALEER (2003): “Asymptotic theory for a vector ARMA–GARCH model,” *Econometric Theory*, 19, 280–310.
- LINTON, O. B. (2008): “Semiparametric and nonparametric ARCH modelling,” in *Handbook of Financial Time Series*, ed. by T. G. Andersen, R. A. Davis, J.-P. Kreiss, and T. Mikosch. Springer, New York.

- LONG, X., AND A. ULLAH (2005): “Nonparametric and semiparametric multivariate GARCH model,” Unpublished manuscript.
- LUUKKONEN, R., P. SAIKKONEN, AND T. TERÄSVIRTA (1988): “Testing linearity against smooth transition autoregressive models,” *Biometrika*, 75, 491–499.
- MCLEOD, A. I., AND W. K. LI (1983): “Diagnostic checking ARMA time series models using squared-residual autocorrelations,” *Journal of Time Series Analysis*, 4, 269–273.
- NELSON, D. B. (1991): “Conditional heteroskedasticity in asset returns: A new approach,” *Econometrica*, 59, 347–70.
- NELSON, D. B., AND C. Q. CAO (1992): “Inequality constraints in the univariate GARCH model,” *Journal of Business and Economic Statistics*, 10, 229–235.
- NG, L. (1991): “Tests of the CAPM with time-varying covariances: a multivariate GARCH approach,” *The Journal of Finance*, 46, 1507–1521.
- PAGAN, A., AND A. ULLAH (1999): *Nonparametric Econometrics*. Cambridge University Press.
- PALM, F. C. (1996): “GARCH models of volatility,” in *Handbook of Statistics*, ed. by G. S. Maddala, and C. R. Rao, vol. 14, pp. 209–240. Amsterdam: Elsevier Sciences.
- PELLETIER, D. (2006): “Regime switching for dynamic correlations,” *Journal of Econometrics*, 131, 445–473.
- ROSS, S. A. (1976): “The arbitrage theory of capital asset pricing,” *Journal of Economic Theory*, 13, 341–360.
- SENTANA, E. (1998): “The relation between conditionally heteroskedastic factor models and factor GARCH models,” *Econometrics Journal*, 1, 1–9.
- SHEPHARD, N. (1996): “Statistical aspects of ARCH and stochastic volatility,” in *Time Series Models in Econometrics, Finance and Other Fields*, ed. by D. R. Cox, D. V. Hinkley, and O. E. Barndorff-Nielsen, pp. 1–67. Chapman and Hall, London.
- SILVENNOINEN, A. (2008): “Numerical aspects of the estimation of multivariate GARCH models,” QFRC Research Paper, University of Technology, Sydney.
- SILVENNOINEN, A., AND T. TERÄSVIRTA (2005): “Multivariate autoregressive conditional heteroskedasticity with smooth transitions in conditional correlations,” SSE/EFI Working Paper Series in Economics and Finance No. 577.
- (2007): “Modelling multivariate autoregressive conditional heteroskedasticity with the double smooth transition conditional correlation GARCH model,” SSE/EFI Working Paper Series in Economics and Finance No. 652.
- STONE, C. J. (1980): “Optimal rates of convergence for nonparametric estimators,” *Annals of Statistics*, 8, 1348–1360.
- TSE, Y. K. (2000): “A test for constant correlations in a multivariate GARCH model,” *Journal of Econometrics*, 98, 107–127.
- TSE, Y. K., AND K. C. TSUI (1999): “A note on diagnosing multivariate conditional heteroscedasticity models,” *Journal of Time Series Analysis*, 20, 679–691.

- (2002): “A multivariate generalized autoregressive conditional heteroscedasticity model with time-varying correlations,” *Journal of Business and Economic Statistics*, 20, 351–362.
- VAN DER WEIDE, R. (2002): “GO-GARCH: A multivariate generalized orthogonal GARCH model,” *Journal of Applied Econometrics*, 17, 549–564.
- VRONTOS, I. D., P. DELLAPORTAS, AND D. N. POLITIS (2003): “A full-factor multivariate GARCH model,” *Econometrics Journal*, 6, 312–334.