Generalized Orthogonal GARCH
A multivariate GARCH model

Roy van der Weide*
Department of Economics, CeNDEF, University of Amsterdam
e-mail: rvdweide@fee.uva.nl

February 12, 2002

Abstract

Multivariate GARCH specifications are typically determined by means of practical considerations such as the ease of estimation, which often results in a serious loss of generality. A new type of multivariate GARCH model is proposed, in which potentially large covariance matrices can be parameterized with a fairly large degree of freedom while estimation of the parameters remains feasible. The model can be seen as a natural generalization of the O-GARCH model, while it is nested in the more general BEKK model. Its relation with the latter model is exploited to derive consistency of the maximum likelihood estimator. In order to avoid convergence difficulties of estimation algorithms, we propose to exploit unconditional information first, so that the number of parameters that need to be estimated by means of conditional information is more than halved. Both artificial and empirical examples are included to illustrate the model.

*An earlier version of this paper has been presented at the International Conference on Modelling and Forecasting Financial Volatility, Perth, Western Australia, September 7-9, 2001. Stimulating discussions with participants, in particular Robert Engle and the convenors Philip Hans Franses and Michael McAleer, are gratefully acknowledged. I am especially grateful to Peter Boswijk, Cees Diks and Roald Ramer as they provided helpful comments and suggestions. I also thank Cars Hommes and Carol Alexander for reading preliminary versions of the paper. This research has been supported by the Netherlands Organization for Scientific Research (NWO) under a NWO-MaG Pionier grant. None of the above are responsible for errors in this paper.

1 Introduction

How to model risk remains a key issue for many financial institutions. In order to describe the variances of economic data numerous econometric models have been introduced in recent years. For univariate time series the autoregressive conditional heteroscedasticity (ARCH) model, introduced by Engle (1982), has become a standard for modelling the time-varying variances of both economic and financial data. The ARCH model has been generalized and extended in many ways ever since, of which the generalized ARCH (GARCH) model of Bollerslev (1986) is one of the first and best known. For many financial applications, however, univariate GARCH models are not able to produce reliable measures for risk, so that multivariate versions are desired.

Some of the best known multi-variate GARCH models available include the VECH model of Bollerslev and Wooldridge (1988), the constant correlation model of Bollerslev (1990), the factor ARCH model of Engle et al. (1990), and the BEKK model studied by Engle and Kroner (1995). For an overview of the multivariate GARCH models, as well as tests for misspecification, see the paper by Kroner and Ng (1998). An extensive survey of empirical applications of time-varying covariance models in finance can be found in Bollerslev et al. (1992).

The 'holy grail' in multivariate GARCH modeling is without any doubt a parameterization of the covariance matrix that is feasible in terms of estimation at a minimum loss of generality. The general multivariate GARCH models available parameterize the covariance matrix by a very large number of parameters that are hard to estimate, which often leads to convergence difficulties of estimation algorithms. Therefore, the choice of the multivariate model is often determined by means of practical considerations i.e. the ease of estimation. The strong restrictions are often not believed to reflect the 'truth', but they are imposed to guarantee feasibility.

Except in the case of the constant correlation model of Bollerslev, most of the models have difficulties in coping with high-variate time series, despite their strong restrictions. As a consequence, the constant correlation model has become a very popular choice for modelling large conditional covariance matrices. In many of the empirical studies, the assumption of constant correlation is taken for granted. A test for this assumption is introduced in a recent paper by Tse (2000). The hypothesis is tested against the alternative of a simple scheme that allows for time-varying correlations. It is shown that e.g. for foreign exchange rates the null hypothesis of constant correlations can not be rejected, whereas in other financial markets the assumption of constant correlations seems unrealistic. In an effort to solve this problem, both Engle (2002) and Tsui and Tse (2002) proposed a generalization of the constant correlation model of Bollerslev. Both show that significant improvements in modelling large covariance matrices can be made when allowing the correlations to be time-varying. However, a less attractive feature of these models is that, regardless of how the conditional correlations are modelled, the conditional variances are described by univariate models.

A somewhat new approach is the Orthogonal GARCH (O-GARCH) or principal components GARCH method. The principal components approach has first been applied in a GARCH-type context by Ding (1994). Shortly after, Alexander and Chibumba (1996) introduced the strongly related O-GARCH model. Thereafter, O-GARCH has been a popular choice to model the conditional covariances of financial data (see e.g. Klaassen (1999)), mainly because the model remains feasible for large covariances matrices (see e.g. Alexander (2002)). Recently, the model has been elaborated along with applications by Alexander (1998, 2001).

The O-GARCH model implicitly assumes that the observed data can be linearly transformed into a set of independent components by means of an orthogonal matrix. These unobserved components can be interpreted as a set of independent factors that drive the particular economy or market, similar to that in the Factor (G)ARCH approach of Engle et al. (1990). The estimator for the linear map is given by the orthonormal eigenvectors of the sample covariance matrix. As the transformed data are assumed to be independent, univariate GARCH-type specifications are sufficient to model the conditional variances of the components. The inverse map can be used to obtain the conditional covariances of the observed series. Clearly, one of the (strong) restrictions imposed by the O-GARCH model is that it requires the matrix, that is assumed to link the independent components with the observed variables, to be orthogonal. If a linkage with a set of independent economic components indeed exists, why should the associated matrix be orthogonal? The O-GARCH model is also known to suffer from identification problems, mainly because estimation of the matrix is entirely based on unconditional information (the sample covariance matrix). For example, when the data exhibits weak dependence, the model has substantial difficulties to identify a matrix that is truly orthogonal (see e.g. Alexander (2001)).

The multivariate GARCH model proposed in this paper can best be seen as a natural generalization of the O-GARCH model. Clearly, orthogonal matrices are very special, and they only reflect a very small subset of all possible invertible linear maps. The generalized O-GARCH model (GO-GARCH) allows the linkage to be given by any possible invertible matrix. Estimation of the matrix requires the use of conditional information, which in turn solves possible identification problems[1]. The parameters are relatively easy to estimate, so that a substantial increase in the degrees of freedom is obtained at a very affordable price. In order to avoid possible convergence difficulties of estimation algorithms, we propose to exploit unconditional information first, so that the number of parameters that need to be estimated by means of conditional information is more than halved.

The next section will introduce notation and outline the O-GARCH model. The generalized Orthogonal GARCH model (GO-GARCH) will be introduced in section 3. Estimation is discussed in section 4. The sections 5 and 6 present some simulation results and an empirical example, respectively. Section 7 concludes.

2 Multivariate GARCH

In a multivariate GARCH setting, the conditional covariance matrix of the m-dimensional zero mean random variable depends on elements of the information set up to time $t-1$, denoted by $\Im_{t-1}$. Assume that ε_t is normally distributed and that its conditional covariance matrix V_t is measurable with respect to $\Im_{t-1}$, the multivariate GARCH model is then described by:

$$\varepsilon_t \left| \Im_{t-1} \sim N\left(0, V_t\right), \right. \tag{1}$$

where we have assumed that ε_t is second order stationary so that $V = E(V_t)$ exists. The information set $\Im_t$ contains both lagged values of the squares and cross-products of ε_t and elements of the conditional covariance matrices up to time t, i.e. lagged values of V_t. The challenge in multivariate GARCH modeling is to find a parameterization of V_t as a function of $\Im_{t-1}$ that is fairly general while feasible in terms of estimation. This section we will briefly describe the type of multivariate GARCH model advocated by Alexander and Chibumba (1996) and Alexander (1998), the O-GARCH model. Some disadvantages of the model are given at the end of the section, which provide the incentives to build a more general model.

In the following we will frequently use the terms conditional information and unconditional information. We specify unconditional information as information that can be extracted from the unconditional covariance matrix. By conditional information we mean the information set $\Im_t$ as introduced above.

[1]For example, the data is not required to exhibit strong dependence for the method to wok well.

2.1 Orthogonal GARCH

Consider a m-variate time series $x_t = (x_{1,t}...x_{m,t})^T$. Let X be a T by m matrix, where the rows of X are given by $x_1^T, ..., x_T^T$, representing the entire m-variate time series in matrix notation. Let the conditional and unconditional covariance matrix of x_t be denoted by V_t and V, respectively. The data is commonly normalized, so that every series (column of X) has unit sample variance and zero mean. As a result, the unconditional covariance matrix V, estimated by $\hat{V} = \frac{1}{T-1} X^T X$, will have 1's along the diagonal.

Under the null, the observed process is described by an orthogonal linear transformation of a set of independent components $y_t = (y_{1,t}...y_{m,t})^T$:

$$x_t = Zy_t. \tag{2}$$

If we denote the diagonal unconditional covariance matrix of the components by H, it follows that:

$$V = ZHZ^T. \tag{3}$$

The associated orthogonal matrix Z is estimated[2] by the eigenvector matrix P of the sample covariance matrix V. The diagonal eigenvalue matrix Λ of V is considered an estimator for the variances of the components:

$$\hat{V} = P\Lambda P^T. \tag{4}$$

The economic components y_t are assumed independent, so that their conditional covariance matrix H_t is diagonal under the null. The covariances of the components are therefore described by means of univariate models only. Alexander (2001, 2002) and Klaassen (1999) employ the popular GARCH(1,1) model[3] to specify the conditional variances $h_{i,t}$ of the components:

$$y_t = u_t \quad u_t \sim N(0, H_t), \tag{5}$$

where the diagonal elements of H_t are described by:

$$h_{i,t} = c_i + \alpha_i u_{i,t-1}^2 + \beta_i h_{i,t-1}, \quad i = 1, ...m. \tag{6}$$

Note that the O-GARCH model still requires a substantial number of parameters to be estimated, namely $\frac{m(m-1)}{2}$ parameters for the orthogonal matrix

[2] Note that identification problems arise when the some of the eigenvalues of H have multiplicity more than one.

[3] Note that the model can easily be extended to more sophisticated models for the components without seriously complicating estimation.

Z, and $3m$ parameters for the univariate GARCH model specifications. However, none of the parameters are likely to give any estimation problems, such as non-convergence. The matrix Z can be extracted from V, which is consistently estimated by the sample covariance matrix $\hat{V}$. Estimation of the univariate GARCH models is also unlikely to cause any difficulties. High variate time series will mainly require more univariate GARCH models to be estimated, which is feasible in terms of estimation. The O-GARCH model has recently been addressed on its ability to generate large covariance matrices in Alexander (2002).

Having estimated all the parameters, the conditional covariance matrix of the original series is simply obtained by applying the linear map:

$$V_t = E_{t-1}x_tx_t^T = E_{t-1}Zy_ty_t^TZ^T \tag{7}$$
$$= PH_tP^T. \tag{8}$$

It is well known that the O-GARCH estimates are not reliable when the data exhibits weak dependence (see e.g. Alexander (2001)). More generally, the O-GARCH model is not always able to identify the orthogonal matrix Z. These identification problems will be illustrated in the next subsection.

2.2 Identification problem

Let us assume that an orthogonal linear linkage Z indeed exists, so that $x_t = Zy_t$. The unconditional covariance matrix V is then given by: $V = ZHZ^T$, where H is diagonal. Then the orthogonal matrix P, the estimator for Z, is only guaranteed to coincide with Z, when the diagonal elements of H are all distinct. Identification problems thus arise when some of the independent components have similar unconditional variance. To see this, suppose that all components have unit variance, so that $V = ZIZ^T = I$. Clearly, the matrix Z is no longer identified by the eigenvector matrix of V, as for every orthogonal matrix Q, we have $(ZQ)(ZQ)^T = I$. Note that the eigenvalues of V reflect the variances of the components when the model is well identified. The estimations should therefore be interpreted with caution when some of the eigenvalues are almost identical. Problems of this type are known to occur when, for example, the data exhibits weak dependence[4].

Consider a bivariate time series $\{x_t\}$, that is constructed from two univariate GARCH processes, both with zero mean. The unconditional sample

[4]Given that the observed data is normalized to have unit variance, which is common practice.

covariance matrix, estimated by $\frac{1}{T-1}X^TX$ and normalized to have unit variance, will typically be of the form:

$$\widehat{V} = \begin{pmatrix} 1 & \delta \\ \delta & 1 \end{pmatrix}, \tag{9}$$

where δ is small, but non-zero, due to finite sample effects. The orthonormal eigenvectors of $\widehat{V}$ are given by:

$$P = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 1 \\ -1 & 1 \end{pmatrix}. \tag{10}$$

The transpose of P is considered an estimator for Z^{-1}, which is used to transform the observed data to obtain the 'independent' components.

Having computed the components, their diagonal conditional covariances $\{H_t\}$ are estimated by means of univariate GARCH models. The implied conditional covariances $\{V_t\}$ of the observed process are then given by:

$$V_t = PH_tP^T. \tag{11}$$

If we denote the diagonal elements of H_t by h_{it}, we have:

$$V_t = \frac{1}{2} \begin{pmatrix} 1 & 1 \\ -1 & 1 \end{pmatrix} \begin{pmatrix} h_{1t} & 0 \\ 0 & h_{2t} \end{pmatrix} \begin{pmatrix} 1 & -1 \\ 1 & 1 \end{pmatrix} \tag{12}$$

$$= \frac{1}{2} \begin{pmatrix} h_{1t} + h_{2t} & h_{2t} - h_{1t} \\ h_{2t} - h_{1t} & h_{1t} + h_{2t} \end{pmatrix}. \tag{13}$$

It follows that V_t is diagonal, if and only if the estimated conditional variances of the components are identical, i.e. $h_{1t} = h_{2t}$. Clearly, this condition will generally not be satisfied, so that from the estimated conditional covariances one would easily conclude that (strong) correlations are present in the observed time series. However, the bivariate time series is constructed from two univariate models and thus independent by definition.

3 Generalized Orthogonal GARCH

3.1 Representation

The key assumption of the GO-GARCH model is the following:

Assumption 1 *The observed economic process $\{x_t\}$ is governed by a linear combination of independent economic components*[5] $\{y_t\}$*:*

$$x_t = Zy_t. \tag{14}$$

The linear map Z that links the unobserved components with the observed variables is assumed to be constant over time, and invertible.

Without loss of generality[6], we normalize the unobserved components to have unit variance, so that:

$$V = Exx^T = ZZ^T. \tag{15}$$

An explicit example, which we will denote the GO-GARCH(1,1) model, would be:

$$x_t = Zy_t \quad y_t \sim N(0, H_t), \tag{16}$$

where each component is described by a GARCH(1,1) process:

$$H_t = diag(h_{1,t}, ..., h_{m,t}) \tag{17}$$
$$h_{i,t} = (1 - \alpha_i - \beta_i) + \alpha_i y_{i,t-1}^2 + \beta_i h_{i,t-1} \quad i = 1, ..., m, \tag{18}$$

where $H_0 = I$ equals the unconditional covariance matrix of the components[7]. The conditional covariances of $\{x_t\}$ are given by:

$$V_t = ZH_tZ^T. \tag{19}$$

[5] Note that there might be more components than the number of variables observed, so that exposing a set of reliable components could be troublesome. However, as the components are assumed to be described by independent GARCH-type models, a new set of independent components can be constructed by aggregating the 'original' components. Under certain conditions, the (extracted) aggregated components are also described by GARCH-type processes, see for example Drost and Nijman (1993) in which temporal aggregation of GARCH processes is considered. However, it is known that the GARCH-type 'features' typically become weaker under aggregation. As a consequence, the accuracy with which the components are described by GARCH-type models increases as more components can be extracted, which will result in better fits.

[6] Note that the unconditional variances of the components and the matrix Z are directly related. Let $\{y_t\}$ denote the components with original scaling, and let the normalized set of components be denoted by $\{\tilde{y}_t\}$, so that $\{\tilde{y}_t\} = \{Dy_t\}$, where D represents the diagonal normalization matrix. The observed process is then given by $\{x_t\} = \{Zy_t\} = \left\{\tilde{Z}\tilde{y}_t\right\}$, where $\tilde{Z} = ZD^{-1}$.

[7] Ling and McAleer (2002) provide a method for treating the initial value when it comes to asymptotic theory for multivariate GARCH.

3.2 Identification

Let P and Λ denote the matrices with, respectively, the orthonormal eigenvectors and the eigenvalues of the unconditional covariance matrix $V = ZZ^T$.

Lemma 2 *Let Z be the map that links the independent components $\{y_t\}$ with the observed process $\{x_t\}$. Then there exists an orthogonal matrix U_0 such that:*

$$P\Lambda^{\frac{1}{2}}U_0 = Z. \tag{20}$$

Proof. The result follows directly from Singular Value Decomposition, see e.g. Horn and Johnson (1999). ■

Let the estimator for U_0 be denoted by U. Without loss of generality, we restrict the determinant of U to be 1[8].

It can be verified that the orthogonal matrices P and Λ have $\frac{m(m-1)}{2}$ and m degrees of freedom, respectively. Together with the $\frac{m(m-1)}{2}$ degrees of freedom for U, we have $m + m(m-1) = m^2$ degrees of freedom for the invertible matrix Z. The matrices P and Λ will be estimated by means of unconditional information, as they will be extracted from the sample covariance matrix V. Conditional information is required to estimate U_0.

Note that there is a continuum of matrices Q for which a set of linearly independent components $u_t = Qx_t$ can be obtained. For every choice of orthogonal matrix U, the linear transformation $Q = U^T\Lambda^{-\frac{1}{2}}P^T$ induces an uncorrelated series with unit variance: $Eu_tu_t^T = QVQ^T = U^TU = I$. Clearly, these components rarely exhibit complete independence. Therefore, linear independence can be very deceiving, as it might give the impression that the linkage between the observed variables and the independent components is uncovered, when more often it is not. The original independent components can only[9] be restored by means of the inverse of Z.

According to lemma 2, the model is well identified as there exists a U_0 that is associated with the original Z. Indeed, the additional $\frac{m(m-1)}{2}$ degrees of freedom induced by the extra term U extends the representation to full generality, in the sense that any invertible linkage Z can in principle be estimated from the data, instead of orthogonal matrices only.

[8] More precisely, U is considered an element of $SO(m)$, which denotes the set of all m-dimensional orthogonal matrices with positive determinant.

[9] Equivalent matrices, in the sense that they only exchange variables for example, are included.

One way to parameterize the estimator for the orthogonal matrix U_0 would be by means of rotation matrices[10]:

Lemma 3 *Every m-dimensional orthogonal matrix U with* $\det(U) = 1$ *can be represented as a product of* $\binom{m}{2} = \frac{m(m-1)}{2}$ *rotation matrices:*

$$U = \prod_{i<j} R_{ij}(\theta_{ij}) \quad -\pi \leq \theta_{ij} \leq \pi, \tag{21}$$

where $R_{ij}(\theta_{ij})$ performs a rotation in the plane spanned by e_i and e_j over an angle θ_{ij}.

Proof. See Vilenkin (1968). ■

The rotation angles[11] $\{\theta_{ij}\}$ are commonly referred to as the Euler angles, which can be estimated by means of maximum likelihood.

In the previous section it was demonstrated that the O-GARCH model suffers from identification problems, for example when the data exhibits weak dependence. These problems should not arise when conditional information is exploited, as proposed in the GO-GARCH model. For example, when the independent components appear to be observed directly, we expect the estimator for U_0 to be close to P^T, since $\hat{Z} = P\Lambda^{\frac{1}{2}}P^T = V^{\frac{1}{2}}$ is approximately diagonal when the data is virtually independent.

3.3 Time-varying correlations

The implied conditional correlations $\{R_t\}$ of the observed process $\{x_t\}$ can be computed as:

$$R_t = D_t^{-1} V_t D_t^{-1}, \quad D_t = (V_t \circ I), \tag{22}$$

where $\{V_t\} = \{Z H_t Z^T\}$ denotes the conditional covariances of $\{x_t\}$.

This theoretical example illustrates how possible lower and upper bounds for the correlation depend on the type of linear map Z_θ. Let Z_θ be the following two dimensional map:

$$Z_\theta = \begin{pmatrix} 1 & 0 \\ \cos\theta & \sin\theta \end{pmatrix}, \tag{23}$$

[10] An alternative parameterization of the orthogonal matrix can be found in van der Weide (2002).

[11] Note that the values for the angles will depend on the ordering of the rotation matrices. The ordering should not affect the estimation results.

where θ measures the extent to which the independent components are mapped in the same direction. For $\theta = 0$ the map is not invertible yielding perfect correlation between the observed variables, whereas for $\theta = \frac{1}{2}\pi$ we have the identity map, so that the observed variables are completely uncorrelated. Let the conditional variances of the observed variables and the independent components be denoted by $(\sigma_{1t}^2, \sigma_{2t}^2)$, and (h_{1t}, h_{2t}), respectively. A straightforward computation shows that the conditional variances of the observed process are given by:

$$\sigma_{1t}^2 = h_{1t} \tag{24}$$

$$\sigma_{2t}^2 = h_{1t}\cos^2\theta + h_{2t}\sin^2\theta. \tag{25}$$

For the conditional covariance we have $\sigma_{12,t} = h_{1t}\cos\theta$. It follows that the conditional correlation, denoted by ρ_t, is given by:

$$\rho_t = \frac{h_{1t}\cos\theta}{\sqrt{h_{1t}}\sqrt{h_{1t}\cos^2\theta + h_{2t}\sin^2\theta}} \tag{26}$$

$$= \frac{\sqrt{h_{1t}}}{\sqrt{h_{1t} + h_{2t}\tan^2\theta}}. \tag{27}$$

If we assume that $h_{it} > 0$, we can define $z_t = \frac{h_{2t}}{h_{1t}}$, so that ρ_t can be expressed as:

$$\rho_t = \frac{1}{\sqrt{1 + z_t\tan^2\theta}}. \tag{28}$$

For finite samples, the variable z_t will have finite lower and upper bounds, denoted by $z_{\min}$ and $z_{\max}$. As a consequence, the conditional correlation ρ_t is also bounded:

$$\frac{1}{\sqrt{1 + z_{\max}\tan^2\theta}} \leq \rho_t \leq \frac{1}{\sqrt{1 + z_{\min}\tan^2\theta}}. \tag{29}$$

For the example presented in section 5.2 with an artificially generated data set of 3000 observations, the bounds $z_{\min}$ and $z_{\max}$ are given by 0.03 and 28.21. The lower and upper bounds for the correlation, when these components are linked by Z_θ, are shown in figure 1. The bounds are plotted as a function of θ, where θ ranges from 0 to $\frac{\pi}{2}$. For $\theta = 0$, we always have perfect correlation, i.e. $\rho_t = 1$, whereas for $\theta = \frac{\pi}{2}$ the observed variables exhibit no correlation at all, i.e. $\rho_t = 0$ for all t.

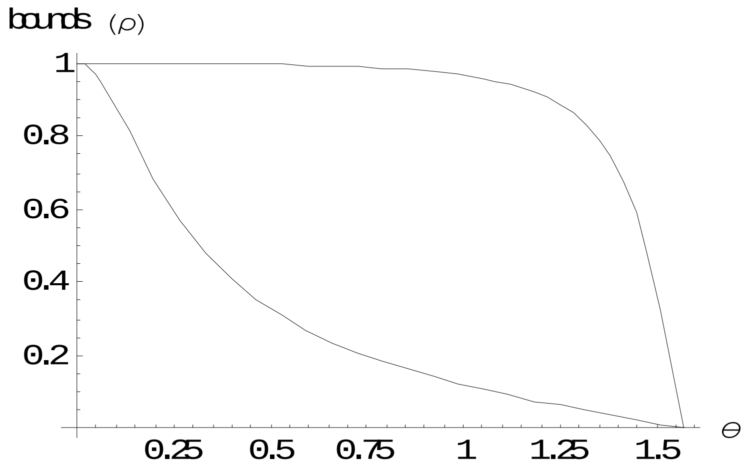

Figure 1: The lower and upper bounds for the correlations

Note that a constant linkage Z gives rise to time-varying correlations between the observed variables. These correlations rise on average when the components are mapped more in the same direction. We can not exclude the possibility that the 'economic mechanism' Z evolves over time. If so, endogenizing Z and make it time-varying, might improve the fit of the time-varying correlations. Extending the GO-GARCH model to allow for a non-constant Z, however, is left for further research. A first step would be to test for a constant linkage, for example by means of test on structural change such as the Chow test.

3.4 Relation to alternative specifications

It can be verified that for U equal to identity, we have the O-GARCH model nested as a special case. This allows one to test the O-GARCH setting with GO-GARCH as the alternative, for example by means of the Likelihood ratio test. The relation with Factor GARCH and the more general BEKK model will be addressed hereupon.

3.4.1 The relation with Factor GARCH

The GO-GARCH model is strongly related to the Factor ARCH model, which is initially proposed in Engle (1987) and implemented by Engle, Ng and Rothschild (1990). Suppose that the observed m-dimensional vector x_t is governed by K factors denoted by a K-dimensional vector ξ_t, where $K \leq m$. According to Factor ARCH, the observed variables and the unobserved factors are related by means of some m by K matrix B with full rank, i.e. $\text{rank}(B) = K$:

$$x_t = B\xi_t + \varepsilon_t, \tag{30}$$

where $\{\varepsilon_t\}$ represent idiosyncratic shocks that are either independent of the factors $\{\xi_t\}$, or they exhibit constant correlation. It is assumed that the shocks have constant covariances Ψ, whereas the factors have conditional covariance matrix denoted by Λ_t. Moreover, assume that both the factors and the shocks have conditional mean zero, i.e. $E_{t-1}\xi_t = E_{t-1}\varepsilon_t = 0$. The conditional covariance matrix of the observed process, denoted by Ω_t, is then given by:

$$\Omega_t = E_{t-1}x_t x_t^T = B\Lambda_t B^T + \Psi. \tag{31}$$

If we neglect the idiosyncratic shocks, i.e. assume $\Psi = 0$, it follows that GO-GARCH is a Factor ARCH model with the maximum number of factors, i.e. $K = m$, where conditional covariances of the factors $\{\Lambda_t\}$ are restricted to be diagonal. In the case of a maximum number of factors, the squared matrix B can be associated with the linkage Z in the GO-GARCH framework.

3.4.2 A special kind of BEKK

Throughout the paper, we specify univariate GARCH(1,1) models for the unobserved components. The GO-GARCH model, however, can easily be adapted to allow for arbitrary model specifications for the components. For example, if one wants to take into account possible asymmetries such as the 'leverage'-effect, an E-GARCH specification might be preferable. As the economic components are assumed independent, more sophisticated specifications for the associated univariate series are not expected to complicate estimation of the GO-GARCH model.

As we will show, a GO-GARCH model with GARCH(1,1)[12] components has the advantage that it is nested in a more general and well-known multivariate GARCH model, namely the BEKK model. As a consequence, most of the theory of maximum-likelihood estimation available for the BEKK model can be applied in our GO-GARCH framework. Consider the following version of Engle and Kroner's BEKK representation for the multivariate GARCH(1,1) model:

$$V_t = C + \sum_{i=1}^{m} A_i x_{t-1} x_{t-1}^T A_i^T + BV_{t-1}B^T, \tag{32}$$

[12]The relation also holds for GARCH(p,q) components.

where C, $\{A_i\}$ and B are all m by m matrices. Moreover, C is assumed to be positive definite, in order to guarantee positive definiteness of V_t for all t.

Lemma 4 *Let the matrices $\{A_i\}_{i=1}^m$ and B be restricted to have identical eigenvector matrix Z, where the eigenvalues of A_i are all zero except for the i'th one. Moreover, assume that C can be decomposed as ZD_CZ^T, where D_C is some positive definite diagonal matrix. Then the associated BEKK model, as in (32), is a GO-GARCH model with GARCH(1,1) components where the Z reflects the linkage between the independent components and the observed process.*

Proof. See Appendix A. ■

4 Estimation

The parameters that need to be estimated by means of conditional information, include the vector θ of rotation coefficients that will identify the invertible matrix Z (see lemma 3), and the parameters (α, β) for the m uni-variate GARCH(1,1) specifications. The log likelihood $L_{\theta,\alpha,\beta}$ for the GO-GARCH model can be represented as:

$$L_{\theta,\alpha,\beta} = -\frac{1}{2}\sum_t m\log(2\pi) + \log|V_t| + x_t^T V_t^{-1} x_t \quad (33)$$

$$L_{\theta,\alpha,\beta} = -\frac{1}{2}\sum_t m\log(2\pi) + \log\left|Z_\theta H_t Z_\theta^T\right| + y_t^T Z_\theta^T (Z_\theta H_t Z_\theta^T)^{-1} Z_\theta y_t \quad (34)$$

$$L_{\theta,\alpha,\beta} = -\frac{1}{2}\sum_t m\log(2\pi) + \log\left|Z_\theta Z_\theta^T\right| + \log|H_t| + y_t^T H_t^{-1} y_t, \quad (35)$$

where $Z_\theta Z_\theta^T = P\Lambda P^T$ is independent of θ. Even in high-variate cases, when the covariance matrices are very large, it should not be a problem to maximize the log likelihood over the $\frac{m(m-1)}{2} + 2m$ parameters. Note that in order to avoid convergence difficulties of estimation algorithms, we propose a kind two-step estimation. We exploit unconditional information first, so that the number of parameters for Z that are estimated through maximum likelihood is $\frac{m(m-1)}{2}$ instead of m^2 (see lemma 2).

4.1 Consistency

Conditions for strong consistency of the maximum likelihood estimator for general multivariate GARCH are derived by Jeantheau (1998). Let Θ be the parameter space, so that $\theta \in \Theta \subset \Re^d$. The true parameter value will be denoted by $\theta_0 \in \Theta$. Jeantheau's conditions are:

A0 Θ is compact.

A1 $\forall \theta_0 \in \Theta$, the model admits a unique strictly stationary and ergodic solution, following a stationary law P_{θ_0}.

A2 There exists a deterministic constant $c > 0$ such that $\forall t$, $\forall \theta \in \Theta$, $det(V_{t,\theta}) \geq c$.

A3 $\forall \theta_0 \in \Theta$, $E_{\theta_0}(|ln(det(V_{t,\theta_0}))|) < \infty$.

A4 The model is identifiable.

A5 $V_{t,\theta}$ is a continuous function of θ.

These conditions are verified by Comte and Lieberman (2000) for the general BEKK model, in which a result of Boussama (1998), concerning the existence of a stationary and ergodic solution to the multivariate GARCH(p,q) process, is used. They summarize their findings in the following theorem.

Theorem 5 *For the GARCH(p,q) process defined by (32), assume that*

B0 Θ *is compact,*

B1 The model is identifiable in the sense of Engle and Kroner (1995),

B2 The standardized errors are i.i.d. and admit a density absolutely continuous w.r.t. the Lebesque measure and positive in a neighbourhood of the origin,

B3 $\rho\left(\sum_{i=1}^{q} A_i + \sum_{i=1}^{p} B_i\right) < 1$, *where* ρ *returns the largest modules of the eigenvalues.*

Then the MLE $\hat{\theta}$ *is strongly consistent for any initial condition.*

We have shown already that the more general BEKK model has the GO-GARCH model nested as a special case, see lemma 4. Strong consistency of the quasi MLE for GO-GARCH can therefore be established by appealing to Jeantheau's conditions, following Comte and Lieberman. To keep it simple, we focus on the GO-GARCH(1,1) model as in (16), but it can be verified that the results also hold for the more general GO-GARCH(p,q) model. For the initial value we choose the unconditional covariance matrix. A method for treating the initial value we refer to Ling and McAleer (2002).

Proposition 6 *Consider the GO-GARCH(1,1) model, where α_i and β_i denote the GARCH(1,1) parameters of the independent components. Assume that B0-B2 holds, and that the components are stationary, i.e.*

$$\alpha_i + \beta_i < 1 \quad fori = 1, ..., m. \tag{36}$$

Then the MLE is consistent.

Proof. See appendix B. ∎

How to conduct inference is beyond the scope of this paper. However, as Comte and Lieberman (2000) have proven asymptotic normality of the quasi-MLE for the BEKK model, having GO-GARCH nested as a special case, we conjecture that this property is also inherited by GO-GARCH. Some caution will be in place though, since we proposed a kind of two-step estimation which will affect the distribution of the estimator. The standard errors might be underestimated by the Fisher Information matrix. We conjecture that everything is fine asymptotically, but we leave the proof for further research. For tests on possible misspecification of the multivariate GARCH model see Kroner and Ng (1998), and the more recent paper by Tse (2002).

5 Simulation results

This section aims to illustrate the behavior of the GO-GARCH model by experimenting with artificial data.

5.1 Orthogonal linkage

We constructed the independent components by generating from 4 uni-variate GARCH(1,1) models to build a 4-variate time series, so that the conditional variance of each component is described by:

$$h_{i,t} = c_i + \alpha_i u_{i,t-1}^2 + \beta_i h_{i,t-1} \quad i = 1, ...4. \tag{37}$$

The values that are assigned to the parameters (c, α, β) are summarized in Table 1.

The parameters are chosen so that variances are nearly integrated, which is commonly observed in financial data. Also note that the parameters are chosen in such a way that some of the unconditional variances are identical. The length of the artificial data set is 3000 observations, which is equivalent to approximately 12 years of daily data.

component	c	α	β
1	0.08	0.10	0.88
2	0.03	0.08	0.90
3	0.05	0.15	0.80
4	0.10	0.20	0.70

Table 1: GARCH parameters

The first orthogonal matrix considered is the identity matrix. As it preserves independence, the components will be observed directly. In the second part, we simulate with an orthogonal matrix that induces dependency among the observed variables.

5.1.1 Independent multivariate GARCH

In this part, we test whether the models are able to detect the independent nature of the observed data. It is known, and demonstrated by a theoretical example in subsection 2.2, that the O-GARCH model can not deal properly with virtually independent data. In contrast, GO-GARCH should be able to estimate a linear representation that induces weak dependence or even independence. The results are presented in Table 2.

	O-GARCH				GO-GARCH			
$\widehat{Z^{-1}}$	0.39	−0.25	−0.64	−0.58	1.00	0.02	0.01	0.01
	−0.27	0.83	−0.06	−0.47	−0.01	1.00	0.01	0.01
	−0.88	−0.36	−0.32	−0.08	0.00	0.02	−1.00	0.00
	−0.06	−0.34	0.70	−0.66	0.00	0.00	0.04	−1.00
c	0.09	0.04	0.03	0.09	0.03	0.02	0.05	0.10
α	0.09	0.06	0.08	0.10	0.11	0.06	0.15	0.20
β	0.81	0.90	0.89	0.80	0.86	0.92	0.80	0.70

Table 2: Estimates for the linkage and GARCH parameters

As expected, O-GARCH was not able to detect the independence of the process. The estimated matrix is far from being diagonal, so that conditional dependence is 'brought into' the residuals. The substantial errors in the GARCH parameters estimates also indicate that O-GARCH did not extract the independent components, but some dependent variables instead. The GO-GARCH model, however, performs very well in this example. The estimated linkage correctly reflects the independent nature of the data. Also the GARCH parameters are estimated properly[13].

5.1.2 Dependent multivariate GARCH

The independent components are described by exactly the same process as in the first part. The key difference is that in this example they are not observed directly. The observed process will be an orthogonal representation of the components that exhibits strong dependence. In principle, the O-GARCH model could also perform well in this example, as the observed variables are no longer independent, while the associated matrix is orthogonal. However, note that some of the components have a similar scaling (unconditional variance). As a consequence, O-GARCH might still suffer from identification problems, see subsection 2.2.

The orthogonal matrix, denoted by Z, is constructed as a product of four rotation matrices, and is shown in table 3. Table 4 summarizes the results.

map	Z
matrix	$\begin{pmatrix} \frac{\sqrt{3}}{2\sqrt{2}} & -\frac{1}{2} & \frac{1}{2} & -\frac{1}{2\sqrt{2}} \\ \frac{\sqrt{3}}{2\sqrt{2}} & \frac{1}{2} & -\frac{1}{2} & -\frac{1}{2\sqrt{2}} \\ -\frac{\sqrt{3}}{4} & \frac{1}{2\sqrt{2}} & \frac{1}{2\sqrt{2}} & -\frac{3}{4} \\ \frac{1}{4} & \frac{\sqrt{3}}{2\sqrt{2}} & \frac{\sqrt{3}}{2\sqrt{2}} & \frac{\sqrt{3}}{4} \end{pmatrix}$

Table 3: Orthogonal linkage

	O-GARCH				GO-GARCH			
c	0.05	0.08	0.05	0.07	0.10	0.02	0.05	0.03
α	0.06	0.07	0.07	0.06	0.20	0.06	0.15	0.11
β	0.89	0.84	0.89	0.87	0.70	0.92	0.80	0.86

Table 4: Estimates for the GARCH parameters

In the case of O-GARCH, the estimates for the GARCH parameters are clearly different from the true parameters suggesting that the model was not able to identify the independent components. In the previous subsection we have seen that the trivial orthogonal matrix, namely identity, could also not be identified by O-GARCH. Thus even when the linkage is truly orthogonal, there is no guarantee that O-GARCH is able to identify it. The model additionally requires that all the components have a different scaling, which might often not be the case (see subsection 2.2).

[13] Note that some components might have been switched.

When we look at the estimates of the GO-GARCH model, we find that the GARCH parameters of the components are estimated with reasonable accuracy[14]. From this we conclude that the linkage estimated by GO-GARCH can not be far from the 'truth' as we build it.

5.2 Non-orthogonal invertible linkage

In this subsection, non-orthogonal invertible matrices are chosen to link the independent components with the observed process. This will be an important example, as we generalized the O-GARCH model to be able to expose linkages that are not orthogonal. In section 3.3 it was illustrated with a theoretical example that lower bounds for the conditional correlations can be observed when the matrices approach singularity.

In order to have a more controlled experiment, we confine ourselves to the 2-dimensional case. Similar to the first examples, we construct two independent components in order to build up a bivariate time series. The conditional variances of both components are specified by means of the same GARCH model, as in (37). Also the sample size is chosen to be identical, namely 3000 observations. Table 5 lists the values at which the GARCH parameters were set to simulate the independent components.

component	c	α	β
1	0.05	0.15	0.80
2	0.05	0.25	0.70

Table 5: The GARCH parameters

It can easily be verified that both components have unit unconditional variance, so that their unconditional covariance matrix equals the identity matrix.

We will consider four different invertible linear maps, linking the independent components with the observed bivariate process. The associated matrices, denoted by Z_1 till Z_4, are shown in Table 6.

map	Z_1	Z_2	Z_3	Z_4
matrix	$\begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$	$\begin{pmatrix} \frac{1}{2} & 1 \\ 0 & 2 \end{pmatrix}$	$\begin{pmatrix} 2 & -2 \\ 1 & 1 \end{pmatrix}$	$\begin{pmatrix} 1 & 2 \\ 2 & 1 \end{pmatrix}$

Table 6: The invertible linkages

[14]Note that some components have been switched.

The unconditional covariance matrix of the observed process is simply given by: $V_i = Z_i Z_i^T$. The covariances V_1 till V_4 are listed in Table 7.

map	Z_1	Z_2	Z_3	Z_4
V_i	$\begin{pmatrix} 2 & 1 \\ 1 & 1 \end{pmatrix}$	$\begin{pmatrix} \frac{5}{4} & 2 \\ 2 & 4 \end{pmatrix}$	$\begin{pmatrix} 2 & 1 \\ 1 & 5 \end{pmatrix}$	$\begin{pmatrix} 5 & 4 \\ 4 & 5 \end{pmatrix}$

Table 7: The unconditional covariances

The observed data is commonly normalized to have unit variance by a diagonal matrix D, so that the covariance matrices of the normalized series $\tilde{V}_i = D_i Z_i Z_i^T D_i^T$ has 1's along the diagonal. In our example, the diagonal elements of D_1till D_4 are easily seen to be $\left\{\frac{1}{\sqrt{2}}, 1\right\}$, $\left\{\frac{2}{\sqrt{5}}, \frac{1}{2}\right\}$, $\left\{\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{5}}\right\}$, and $\left\{\frac{1}{\sqrt{5}}, \frac{1}{\sqrt{5}}\right\}$, respectively. It follows that the true matrix that links the normalized observed variables with its independent components, is given by $(D_i Z_i)^{-1}$. These matrices, denoted by W_i, are shown in Table 8.

map	Z_1	Z_2	Z_3	Z_4
W_i	$\begin{pmatrix} 1.41 & -1 \\ 0 & 1 \end{pmatrix}$	$\begin{pmatrix} 2.24 & -2 \\ 0 & 1 \end{pmatrix}$	$\begin{pmatrix} 0.47 & 0.75 \\ -0.94 & 0.75 \end{pmatrix}$	$\begin{pmatrix} -0.75 & 1.49 \\ 1.49 & -0.75 \end{pmatrix}$

Table 8: The true linear representations

Note that in all cases the orthonormal eigenvectors of the unconditional covariance matrix are given by $P = \frac{1}{\sqrt{2}} \begin{pmatrix} -1 & 1 \\ 1 & 1 \end{pmatrix}$, so that O-GARCH is expected to estimate scaled versions of P for the linkages.

The results for both the O-GARCH as the GO-GARCH model are presented in Table 9 and 10, respectively.

This example illustrates that the GO-GARCH model is able to deal with decompositions that are not of the orthogonal type. The estimated linkages are in all cases very close[15] to the 'truth', the matrices from Table 8. The estimates for the GARCH parameters are also accurate.

A priori we know that the O-GARCH model can not uncover the non-orthogonal linkages, as it restricts the matrix to be orthogonal. As a consequence, it extracts components that are not independent, which is reflected by the biased estimates for the GARCH parameters. Particularly in example 4, the O-GARCH estimates for the GARCH parameters show substantial

[15]Neglect signs, as they do not yield a different representation. Also note that some components might have been switched.

O-GARCH

map	Z_1		Z_2		Z_3		Z_4	
$\hat{W}_i$	0.54	0.54	0.51	0.51	-0.61	-0.61	0.53	0.53
	1.31	-1.31	2.17	-2.17	0.86	-0.86	1.59	-1.59
$\hat{c}_i$	0.06	0.05	0.05	0.05	0.05	0.05	0.05	0.06
$\hat{\alpha}_i$	0.18	0.12	0.21	0.13	0.14	0.23	0.11	0.12
$\hat{\beta}_i$	0.76	0.83	0.73	0.82	0.82	0.72	0.84	0.81

Table 9: The estimates for the linkages and the GARCH parameters

GO-GARCH

map	Z_1		Z_2		Z_3		Z_4	
$\hat{W}_i$	0.01	0.99	0.02	0.98	-0.48	-0.73	1.49	-0.74
	1.41	-1.01	2.23	-2.00	0.94	-0.76	0.77	-1.51
$\hat{c}_i$	0.05	0.05	0.05	0.05	0.05	0.05	0.05	0.05
$\hat{\alpha}_i$	0.25	0.14	0.25	0.14	0.14	0.25	0.25	0.14
$\hat{\beta}_i$	0.71	0.81	0.71	0.81	0.81	0.71	0.71	0.81

Table 10: The estimates for the linkages and the GARCH parameters

error. The difference between the estimated correlations is therefore most notable in example 4, which can be seen in figure 2.

In example 4, the GO-GARCH estimates for the correlations never fall below 0.8, say, whereas the correlations estimated by O-GARCH show much stronger declines and sometimes even drop till below 0.4. The reason for this effect is that the matrix from example 4 shows the strongest ‘deviation’ from an orthogonal matrix compared with the matrices from the other examples. The linkage from example 4 maps both independent components in almost the same direction which induces a strong correlation between the observed variables. Exactly the same feature is observed in the empirical example described in the next section. Indeed, it seems reasonable that observed variables that are strongly related exhibit high correlation at all times. As the linkage with the components that induces the high correlation between the observed variables is assumed constant over time, it will be surprising to observe periods in time where the variables suddenly appear almost uncorrelated. This feature is illustrated by a theoretical bivariate example, where the lower bound and upper bounds of the correlation are derived as a function of a characteristic parameter θ of the linkage Z_θ, in subsection 3.3.

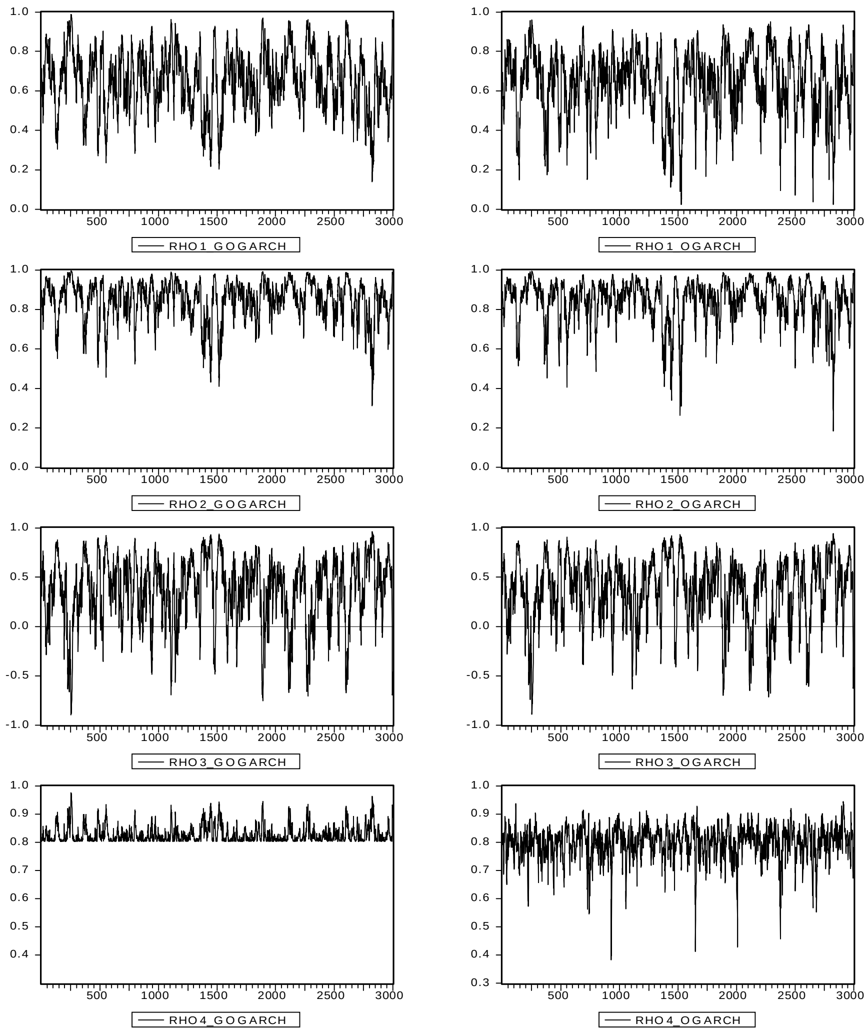

Figure 2: The estimated time evolutions of the (conditional) correlations for all four examples.

6 Empirical example

We include an example from real life, as an attempt to gain insight in the relation between observed economic and financial variables and the independent factors that are assumed to drive the market. Our example considers the Dow Jones Industrial Average (DJIA) versus the NASDAQ composite. The sample contains more than ten years of daily observations, starting at the first of January in 1990 and ending in October 2001. First, we estimate a first order[16] vector autoregressive (VAR) model to account for the linear structure present in the data. Subsequently, we use the residuals to estimate the conditional covariances from which the (conditional) correlations between the DJIA and the NASDAQ can be computed. Questions that arise naturally include: (i) Are non-orthogonal linkages common in real life examples? (ii) Will allowing for a more general linkage (all invertible matrices) typically induce a better description of the time-varying correlations between economic and financial variables?

The results for the first order VAR model, which removes the linear structure from the data, are summarized in Table 11.

	DJIA	NASDAQ
DJIA(-1)	0.05^{*}	0.04
NASDAQ(-1)	-0.02	0.03
Constant	$3.87\ 10^{-4*}$	$4.07\ 10^{-4}$

Table 11: The estimates for the first order VAR model

In order to examine how the correlations are affected by the generality of the linkage, we estimate both the O-GARCH and GO-GARCH model, and compare their results. The estimates are summarized in Table 12.

model	O-GARCH	GO-GARCH
$\widehat{Z^{-1}}$	$\begin{pmatrix} -0.55 & -0.55 \\ 1.19 & -1.19 \end{pmatrix}$	$\begin{pmatrix} -1.18 & 0.32 \\ 0.58 & -1.27 \end{pmatrix}$
c	0.009 0.003	0.009 0.003
α	0.070 0.039	0.054 0.079
β	0.922 0.957	0.939 0.915

Table 12: The estimates for the linkages and GARCH parameters

To address the question whether non-orthogonal linkages can be found in financial data, we first verify whether the unrestricted matrix estimated by the GO-GARCH model is approximately orthogonal. Let $\widehat{Z_{go}^{-1}}$ denote the unrestricted representation, then $(\widehat{Z_{go}^{-1}}(\widehat{Z_{go}^{-1}})^T = \begin{pmatrix} 1.49 & -1.09 \\ -1.09 & 1.95 \end{pmatrix}$ should be identity for $\widehat{Z_{go}^{-1}}$ to be orthogonal. As this is clearly not the case, we have found support for our conjecture that orthogonal matrices are not able to reflect the possible linkages hidden in financial markets, so that O-GARCH is too restrictive. In order to quantify the impact of the restrictions imposed

[16]Higher order specifications do not significantly contribute to a better linear fit.

by O-GARCH, we compute the Likelihood Ratio statistic LR to test O-GARCH against GO-GARCH for several lengths of the time series. The results are listed in Table 13. For all lengths of the time-series considered, the hypothesis of an orthogonal linkage is rejected at a 1% level.

Length	250	500	1000	3082
LR	23.7	13.8	42.6	731.4

The critical value of χ^2_1 at a 1% level is 6.63.

Table 13: Likelihood-Ratio Test of O-GARCH against GO-GARCH

To examine to what extend the restrictions on the linkage affect the (conditional) correlations, we compare the implied correlations of both models. The time evolution of the correlations is shown in figure 3.

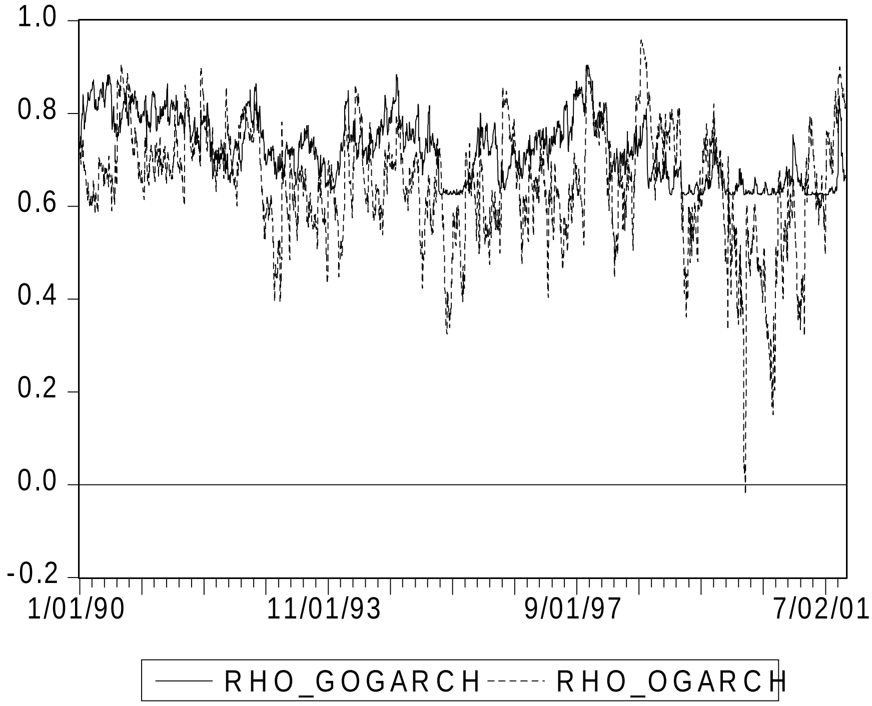

Figure 3: The estimated correlations between the DJIA and the NASDAQ

Figure 3 reveals several interesting features. Perhaps most striking is that the correlations estimated by GO-GARCH are much less volatile. Furthermore, the GO-GARCH correlations never seem to fall below 0.6, say, whereas the more volatile correlations estimated by O-GARCH show several substantial drops where the last one is undoubtedly most notable. In the beginning of 2000, the correlation according to O-GARCH is even just below zero where the GO-GARCH correlations show no decline at all. Since the matrix estimated by GO-GARCH is explicitly not orthogonal, where the independent factors are mapped in a similar direction, we have reason to believe that O-GARCH often underestimates the correlations (see section 3.3). Indeed, it seems plausible that the DJIA and the NASDAQ, which are strongly related, exhibit high correlation at all times. As the linkage with its components is assumed constant over time, it will be surprising to observe periods in time where the DJIA and the NASDAQ appear to be almost unrelated. However, the differences observed for the year 2000 are kind of extreme. This suggests perhaps that the linkage is not constant over time, and that it exhibits a structural change in the year 2000. Indeed, at the beginning of 2000 we experienced a technological boom which might explain such a structural break. To test for a constant linkage will be left for further research.

7 Concluding remarks

A new type of multivariate GARCH model is proposed that can best be seen as a generalization of the O-GARCH model. The O-GARCH model aims to address the demand for multivariate GARCH models that are able to cope with large covariance matrices. In order to do so, the model is based on strong restrictions, maybe too strong. The GO-GARCH model generalizes the O-GARCH model in an effort to shorten the distance from the 'truth' while preserving feasibility. It supports the assumption that the observed variables are driven by some unobserved independent components, linked by a linear map. In order to identify these components, invertibility of the associated matrix is all we need. Under the null of O-GARCH, however, the matrix is assumed orthogonal which only covers a very small subset of all possible invertible matrices. Moreover, even when the matrix is truly orthogonal, the estimator proposed by O-GARCH is not always able to identify it. The GO-GARCH model considers every invertible matrix as a possible linkage, which will be parametrized in such a way that it is not expected to complicate estimation while excluding any idenification problems. In order to avoid possible convergence difficulties of estimation algorithms, we propose a kind two-step estimation. First unconditional information is exploited, so that the number of parameters that are estimated through maximum likelihood is about halved. The parameters are relatively easy to estimate, so that a substantial gain in generality is obtained at a very affordable price.

It is shown that GO-GARCH is related to existing multivariate GARCH models, such as the Factor GARCH model and the more general BEKK model. The asymptotic theory available for BEKK assisted to derive consistency of the maximum likelihood estimator.

The model is tested on both artificial and financial data. The simulation results show that the model correctly estimates both orthogonal and non-orthogonal invertible linkages. The results are not affected by the scaling of the independent factors or a possible weak dependence among the observed variables. The latter is known to be responsible for the identification problems of O-GARCH, which is confirmed by our experimental results. The nature of the linkage, for example whether it is orthogonal or not, is strongly related with the implied correlations between the observed variables. This relation is made explicit by a theoretical example, and illustrated by some of the simulation results. The effect of the linkage on the correlations is also observed in the empirical example, the Dow Jones Industrial Average versus the NASDAQ. A likelihood ratio test rejects the hypothesis that the associated matrix is orthogonal, which supports the conjecture that the GO-GARCH model provides a better description of the process, and the conditional correlations in particular. We argue that by restricting the matrix to be orthogonal, O-GARCH will often underestimate the correlations. The differences are kind of extreme during the year 2000, which coincides with the technology boom that initiated early that year. It could be that the linkage is not constant over time and that it exhibited a structural change during the technological bust in 2000. A test for such a structural break, and perhaps even extending the model to allow for a time-varying linkage, is left for further research. Also the aymptotic properties of the two-step estimator still needs to be addressed. Probably the most important question that remains, is to what extent is GO-GARCH able to improve the modelling of large covariance matrices? Indeed, it would be very interesting to compare the model with other recently developed multivariate GARCH models, such as the Dynamic Conditional Correlation model of Engle (2002). The misspecification tests recently proposed by Tse (2002) can be used to measure performance.

References

Alexander, C. (1998). Volatility and correlation: Methods, models and applications. *Risk Management and Analysis: Measuring and Modelling Financial Risk, Wileys*, 125–172.

Alexander, C. (2001). Orthogonal garch. *Mastering Risk, Financial Times - Prentice Hall*, **2,** 21–38.

Alexander, C. (2002). Principal component models for generating large garch covariance matrices. *Review of Banking, Finance and Monetary Economics, forthcoming.*

Alexander, C. and Chibumba, A. (1996). Multivariate orthogonal factor garch. *University of Sussex Discussion Papers in Mathematics.*

Bollerslev, Chou R. T. and Kroner, K. (1992). Arch modeling in finance: A selective review of the theory and empirical evidence. *Journal of Econometrics*, **52,** 5–59.

Bollerslev, Engle R. T. and Wooldridge, J. (1988). A capital asset pricing model with time varying covariances. *Journal of Political Economy*, **96,** 116–131.

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

Bollerslev, T. (1990). Modelling the coherence in short-run nominal exchange rates: A multivariate generalized arch approach. *Review of Economics and Statistics*, **72,** 498–505.

Boussama, F. (1998). Ergodicity, mixing and estmation in garch models. *Ph.D. dissertation, University Paris* 7.

Comte, F. and Lieberman, O. (2000). Asymptotic theory for multivariate garch processes. *Working Paper, CREST-ENSAE, Paris.*

Ding, Z. (1994). Time series analysis of speculative returns. *Ph.D. dissertation, University of California, San Diego.*

Drost, F. and Nijman, T. (1993). Temporal aggregation of garch processes. *Econometrica*, **61,** 909–927.

Engle, Ng V. R. and Rothschild, M. (1990). Asset pricing with a factor arch covariance structure: Empirical estimates for treasury bills. *Journal of Econometrics*, **45,** 213–238.

Engle, R. (1982). Autoregressive conditional heteroskedasticity with estimates of the variance of u.k. inflation. *Econometrica*, **50,** 987–1008.

Engle, R. (2002). Dynamic conditional correlation - a simple class of multivariate garch models. *Journal of Business and Economic Statistics, forthcoming.*

Engle, R. and Kroner, K. (1995). Multivariate simultaneous generalized arch. *Econometric Theory*, **11,** 122–150.

Horn, R. and Johnson, C. (1999). *Matrix Analysis.* Cambridge: Cambridge University Press.

Jeantheau, T. (1998). Strong consistency of estimators for multivariate arch models. *Econometric Theory*, **14,** 70–86.

Klaassen, F. (1999). Have exchange rates become more closely tied? evidence from a new multivariate garch model. *CentER Working Paper, Tilburg Univesrity.*

Kroner, K. and Ng, V. (1998). Modeling asymmetric comovements of asset returns. *The Review of Financial Studies*, **11,** 817–844.

Ling, S. and McAleer, M. (2002). Asymptotic theory for a vector arma-garch model. *Econometric Theory, forthcoming.*

Tse, Y. (2000). A test for constant correlations in a multivariate garch model. *Journal of Econometrics*, **98,** 107–127.

Tse, Y. (2002). Residual-based diagnostics for conditional heteroscedasticity models. *Econometrics Journal, forthcoming.*

Tsui, A. and Tse, Y. (2002). A multivariate garch model with time-varying correlations. *Journal of Business and Economic Statistic, forthcoming.*

Vilenkin, N. (1968). Special functions and the theory of group representation, translations of mathematical monographs. *American Math. Soc., Providence, Rhode Island, U.S.A.*, **22**.

A Proof of lemma 4

As the matrices $\{A_i\}_{i=1}^q$ and B are assumed to have identical eigenvector matrix Z, they can be diagonalized as:

$$A_i = ZD_{A_i}Z^{-1} \text{ and } B = ZD_BZ^{-1}, \tag{38}$$

where $\{D_{A_i}\}$ and D_B denote diagonal eigenvalue matrices. All element of D_{A_i} are zero except for its i'th diagonal element, which represents the only non-zero eigenvalue of A_i and will be denoted by a_i. Substituting these decompositions , along with $C = ZD_CZ^T$, we have:

$$V_t = Z\left(D_C + \sum_{i=1}^{m} D_{A_i}Z^{-1}x_{t-1}x_{t-1}^T\left(Z^{-1}\right)^T D_{A_i} + D_B Z^{-1}V_{t-1}\left(Z^{-1}\right)^T D_B\right)Z^T. \tag{39}$$

Let $\{y_t\} = \{Z^{-1}x_t\} = \{Wx_t\}$ represent the unobserved economic components in the GO-GARCH framework, and let $\{H_t\} = \left\{WV_tW^T\right\}$ denote the conditional covariance of $\{y_t\}$. Rearranging terms we find:

$$H_t = D_C + \sum_{i=1}^{m} D_{A_i}y_{t-1}y_{t-1}^T D_{A_i} + D_B H_{t-1} D_B. \tag{40}$$

By construction of the matrices $\{D_{A_i}\}$ it follows that:

$$\sum_{i=1}^{m} D_{A_i}y_{t-1}y_{t-1}^T D_{A_i} = D_A \circ y_{t-1}y_{t-1}^T, \tag{41}$$

where $D_A = diag\{a_1, ..., a_m\}$, and where $\circ$ denotes the Hadamard product. As the matrices D_C, D_B and $D_A \circ y_{t-1}y_{t-1}^T$ are all diagonal, the conditional covariance matrix of y_t, denoted by H_t, is also diagonal. Therefore, equation (40) implies univariate GARCH(1,1) specifications for the components $\{y_t\}$, as is assumed by the GO-GARCH model.

B Proof of proposition 6

In order to prove consistency we appeal to the conditions A0-A5 of Jeantheau (1998). Comte and Lieberman (2000) show that by assumptions B0-B2, the conditions A0 and A2-A5 are satisfied for the more general BEKK model as in (32). To prove that A1 is also fulfilled, they use a theorem from Boussama (1998), which requires that $\rho(A + B) < 1$, where ρ returns the largest modulus of the eigenvalues (see B3). By lemma 4 we know that GO-GARCH is nested in BEKK, and that $A = ZD_AZ^{-1}$ and $B = ZD_BZ^{-1}$ where $D_A = diag(\alpha_1, ..., \alpha_m)$ and $D_B = diag(\beta_1, ..., \beta_m)$. Thus A1 is satisfied when $\rho(Z(D_A + D_B)Z^{-}) = \rho(D_A + D_B) = \max_i(\alpha_i + \beta_i) < 1$, which is ensured by the stationarity assumption: $\alpha_i + \beta_i < 1 \quad$ for i=1,...,m.