A Generalized Dynamic Conditional Correlation Model for Many Asset Returns

Christian M. Hafner[1] Philip Hans Franses[2]

Econometric Institute Report EI 2003–18

June 2003

Abstract

In this paper we put forward a generalization of the Dynamic Conditional Correlation (DCC) Model of Engle (2002). Our model allows for asset-specific correlation sensitivities, which is useful in particular if one aims to summarize a large number of asset returns. The resultant GDCC model is considered for daily data on 18 German stock returns, which are all included in the DAX, and for 25 UK stock returns in the FTSE. We find convincing evidence that the GDCC model improves on the DCC model and also on the CCC model of Bollerslev (1990).

Keywords: Multivariate GARCH, dynamic conditional correlation

JEL Classification: C14, C22.

[1]Econometric Institute, Erasmus University Rotterdam, P.O. Box 1738, NL–3000 DR Rotterdam, The Netherlands, chafner@few.eur.nl, corresponding author

[2]Econometric Institute, Erasmus University Rotterdam, P.O. Box 1738, NL–3000 DR Rotterdam, The Netherlands

1 Introduction

A topic high on the research agenda in financial econometrics is the construction of models that can summarize the dynamic properties of two or more asset returns, with a particular focus on volatility forecasting and portfolio selection. A class of models that addresses this topic is the multivariate GARCH model. By now, there are many variants available, see Bauwens et al. (2003) for a recent survey. The current benchmark models seem to be the Constant Conditional Correlation (CCC) model of Bollerslev (1990) and its extension, the Dynamic Conditional Correlation (DCC) model of Engle (2002). These models impose a useful structure on the many possible model parameters. By doing so, the model parameters can easily be estimated and the model can be evaluated and used in a rather straightforward way.

In this paper we aim to extend on the DCC model by focussing on the notion that one might want to use this model for a large number of asset returns. For example, one might want to summarize 18 important stocks in the DAX for the purpose of portfolio selection, as we will do below. As is shown in Engle and Sheppard (2001), the DCC model leads to sub-optimal portfolio selection in case of many assets (like 20 or 30). This is due to the fact that the DCC model assumes that the asset-specific conditional correlations all follow the same dynamic (ARMA-type) structure. This assumption may be more easily satisfied by a small number of selected asset returns, but it becomes increasingly more unlikely in case of many returns. Hence, intuitively, when one considers many returns, one would want to allow for asset-specific dynamics, and this is precisely what we do in this paper. By allowing one of the ARMA parameters to vary across the assets, and in a sense allowing for a panel structure, we generalize the DCC model towards a GDCC model.

The outline of our paper is as follows. In Section 2, we review the CCC and DCC model, and we introduce our GDCC model. In Section 3, we discuss parameter estimation of the GDCC model, and various ways to compare it with the CCC and DCC model. In Section 4, we consider the three models for daily data on 18 stock returns in the DAX and on 25 stock returns in the FTSE. We document that the GDCC model improves on the other two models in various dimensions. In Section 5, we conclude with some remarks.

2 Dynamic conditional correlation models

Let y_t be an N dimensional time series of length T. Suppose for simplicity that the mean of y_t is zero. For example, y_t could be the returns of the stocks in the DAX index. Our objective is to find a suitable model for the conditional covariance matrix H_t of y_t if both N and T are large.

The main benchmark is the CCC model of Bollerslev (1990), which specifies

$$H_t = D_t R D_t,$$

where D_t is a diagonal matrix with the square root of the estimated univariate GARCH variances on the diagonal, and R is the sample correlation matrix of y_t. Although the model is useful, the assumption of constant conditional correlations can be too restrictive. One may expect higher correlations in extreme market situations like crashes, for example.

Engle (2002) generalizes the CCC model to the Dynamic Conditional Correlation model (DCC). This model is

$$H_t = D_t R_t D_t \tag{1}$$

$$R_t = \text{diag}(Q_t)^{-1/2} Q_t \text{diag}(Q_t)^{-1/2} \tag{2}$$

$$Q_t = S(1-\alpha-\beta) + \alpha\varepsilon_{t-1}\varepsilon'_{t-1} + \beta Q_{t-1} \tag{3}$$

where α and β are parameters and $\varepsilon_t = D_t^{-1} y_t$ are the standardized but correlated residuals. That is, the conditional variances of the components of ε_t are equal to 1, but the conditional correlations are given by R_t. $\text{diag}(Q_t)$ is a diagonal matrix with the same diagonal elements as Q_t. S is the sample correlation matrix of ε_t, which is a consistent estimator of the unconditional correlation matrix. If α and β are zero, one obtains the above CCC model. If they are different from zero one gets a kind of ARMA structure for all correlations. Note however that all correlations would follow the same kind of dynamics, since the ARMA parameters are the same for all correlations.

We propose to extend the DCC model to a generalized DCC (GDCC) model in the following way, that is

$$H_t = D_t R_t D_t \tag{4}$$

$$R_t = \text{diag}(Q_t)^{-1/2} Q_t \text{diag}(Q_t)^{-1/2} \tag{5}$$

$$Q_t = S\left(1-\bar{\alpha}^2-\bar{\beta}^2\right) + \alpha\alpha' \odot \varepsilon_{t-1}\varepsilon'_{t-1} + \beta\beta' \odot Q_{t-1} \tag{6}$$

where $\odot$ denotes the Hadamard matrix product operator, i.e., elementwise multiplication. In (6), α and β are $N \times 1$ parameter vectors, $\bar{\alpha} = 1/N \sum_{i=1}^{N} \alpha_i$ and $\bar{\beta} = 1/N \sum_{i=1}^{N} \beta_i$. Clearly, the DCC model results as a special case if $\alpha_1 = \cdots = \alpha_N$ and $\beta_1 = \cdots = \beta_N$. The GDCC model guarantees to deliver positive definite H_t, because Q_t is a sum of positive (semi-)definite matrices, provided that a suitable starting value for Q_0 is used, for example the sample correlation matrix S.

Note that the exact variance targeting approach as in the DCC model does not work here, as the matrix $S \odot (\iota\iota' - \alpha\alpha' - \beta\beta')$ is not positive definite in general. Thus, replacing the first term in (6) by this matrix would not guarantee a positive definite Q_t. The GDCC specification (6) leads to a bias in the unconditional correlations in the sense that they do no longer correspond necessarily to the sample correlations. However, this should be weighted against the flexibility gain for the dynamics of the correlations. As the DCC model is nested in the GDCC model, the null hypothesis of DCC can be tested using standard Wald or Likelihood ratio statistics. An exact variance targeting would be possible if the residuals ε_t were orthogonalized such that $S = I_N$, because the matrix $I_N \odot (\iota\iota' - \alpha\alpha' - \beta\beta')$ is positive semi-definite if $\alpha_i^2 + \beta_i^2 < 1$ for all i. We tried an orthogonalization in one empirical application but did not find any substantial improvement.

The GDCC model (6) contains $2N$ parameters for the conditional correlations. This may still be problematic for estimation if N is very large. A compromise between the models (3) and (6) could be found by noting that often the parameters associated with the innovations, α, are more varying over the panel than the parameters associated with the autoregression, β. In that case, we can specify

$$Q_t = S\left(1 - \bar{\alpha}^2 - \beta\right) + \alpha\alpha' \odot \varepsilon_{t-1}\varepsilon_{t-1}' + \beta Q_{t-1} \tag{7}$$

with only $N + 1$ parameters to estimate. One can still reduce the number of parameters by pooling variables with similar values α_i into meaningful clusters.

On the other hand, one may still add flexibility and introduce exogenous variables or factors in the equation for Q_t. For example, we could include a factor $DAX_{t-1}^2 1_{DAX_{t-1}<\tau}$, because, for example, it may be that correlations increase in crash situations where the DAX return is smaller than a threshold τ.

Note that the ijth element of Q_t can be written as

$$q_{ij,t} = S_{ij}(1 - \bar{\alpha}^2 - \beta) + \alpha_i\alpha_j e_{ij,t-1} + \beta q_{ij,t-1}$$

where $e_{ij,t} = \varepsilon_{i,t}\varepsilon_{j,t}$ can be called the correlation innovation. The ijth element of R_t, the conditional correlation matrix is given by

$$r_{ij,t} = \frac{q_{ij,t}}{\sqrt{q_{ii,t}q_{jj,t}}} \tag{8}$$

The α parameters could be given the following interpretation: If an α_i is large (small), then the correlation of the corresponding asset with other assets tends to be (in)sensitive to correlation innovations. In the extreme case that $\alpha_i = 0$, we can write $r_{ij,t}$ as

$$r_{ij,t} = \frac{S_{ij}\sqrt{1 - \bar{\alpha}^2}}{\sqrt{S_{ii}q_{jj,t}}}$$

Thus, if $\alpha_i = 0$, then all variation of $r_{ij,t}$ originates from variation of $q_{jj,t}$, which does not depend on correlation innovations $e_{ij,t}$. In other words, we can characterize the α's as the individual asset's sensitivity with respect to correlation innovations.

3 Estimation

This section discusses estimation methods for the GDCC model. We first review the simultaneous estimation of all parameters, before discussing possible ways to combine estimation of the individual correlations.

3.1 Simultaneous estimation

Estimation of the GDCC model parameters can be performed by quasi maximum likelihood (QML) by maximizing the criterion function

$$L(\theta) = -\frac{1}{2}\sum_{t=1}^{T}\left(\log|H_t(\theta)| + y_t' H_t^{-1}(\theta) y_t\right)$$

with respect to the parameter vector θ. Under quite general conditions, listed by Engle and Sheppard (2001), these estimators will be consistent and asymptotically normal. If the estimation for the variances (contained in D_t) and the correlations (contained in R_t) is performed simultaneously, the QML estimation will be efficient provided that innovations are indeed Gaussian. If estimation is split up in two parts, where first the variances are estimated, and then the correlations, then estimators will no longer be efficient but still consistent. Following Engle (2002), the likelihood can be split in two parts,

$$L(\theta) = L_V(\theta_V) + L_C(\theta_C)$$

where

$$L_V(\theta_V) = -\frac{1}{2}\sum_{t=1}^{T}\left(\log|D_t(\theta_V)|^2 + y_t' D_t(\theta_V)^{-2} y_t\right) \tag{9}$$

is the volatility part of the likelihood, and

$$L_C(\theta_C) = -\frac{1}{2}\sum_{t=1}^{T}\left(\log|R_t(\theta_C)| + \varepsilon_t' R_t(\theta_C)^{-1} \varepsilon_t\right) \tag{10}$$

is the correlation part, with $\theta = (\theta_V', \theta_C')'$. At the first step, (9) is maximized with respect to θ_V by estimating the univariate GARCH models for y_{it}, $i = 1, \ldots, N$. Define the estimate of θ_V by $\widehat{\theta}_V = \arg\max L_V(\theta_V)$. Conditional on the first step, standardized residuals

$\varepsilon_t = D_t(\widehat{\theta}_V)^{-1}y_t$ can be calculated. At the second step, (10) is maximized with respect to θ_C, giving the estimate $\widehat{\theta}_C = \arg\max L_C(\theta_C)$. We use this two-step estimation procedure in the empirical part of the paper. Inference concerning the correlation parameter vector θ_C has to take the first step into account, as described by Engle and Sheppard (2002). We also use their results to compute standard errors.

3.2 Combining individual correlation estimates

Maximization of the likelihood function may be cumbersome if the dimension N is high, as in every step, $N \times N$ covariance matrices must be inverted. The inversions are numerically difficult because the covariance matrices are typically ill-conditioned. It might therefore be preferable to look for estimation routines of the individual correlations that still restrict the composed covariance matrix to be positive definite. For example, estimating univariate ARMA-type models for each component of the covariance matrix can be achieved so quickly that the task of estimating $N(N-1)/2$ such univariate models can still be much faster than estimating the multivariate model. The difficult part is to restrict the univariate models such that the composed multivariate model forms valid covariance matrices.

In the following, we discuss one way of achieving this for the GDCC model. In the standard DCC(1,1) model, Engle (2002) suggests to rewrite the ij-th cross product as an ARMA(1,1) process,

$$e_{ij,t} = S_{ij}(1-\alpha-\beta) + (\alpha+\beta)e_{ij,t-1} - \beta u_{ij,t-1} + u_{ij,t}, \tag{11}$$

where $e_{ij,t} = \varepsilon_{i,t}\varepsilon_{j,t}$ and $u_{ij,t} = e_{ij,t} - q_{ij,t}$ has mean zero and can be treated as an error term. Note however that it is not a martingale difference as the conditional expectation of $e_{ij,t}$ is $r_{ij,t}$ and not $q_{ij,t}$, and it is not obvious to show that they are serially uncorrelated. We found in a small simulation exercise that for typical parameter values, there is some autocorrelation in the $u_{ij,t}$, which implies that (11) is not an ARMA process and parameter estimates are biased.

Another way to obtain estimates for individual correlations is described in the following. As by definition $r_{ij,t}$ is the conditional expectation of $e_{ij,t}$, we can write

$$e_{ij,t} = r_{ij,t}(\phi_{ij}, \theta_{ij}) + \eta_{ij,t}, \tag{12}$$

where $\phi_{ij} = (\alpha_i, \alpha_j)$, $\theta_{ij} = (\beta_i, \beta_j)$, and $\eta_{ij,t}$ is an error term with variance σ_{ij}^2, say, such that $\mathrm{E}[\eta_{ij,t} \mid \mathcal{F}_{t-1}] = 0$. If the conditional distribution of $\eta_{ij,t}$ can be approximated by a normal distribution, then we can estimate ϕ_{ij} and θ_{ij} by maximizing

$$L_{ij}(\phi_{ij}, \theta_{ij}) = -\frac{1}{2}\sum_{t=1}^{n}\left(\log\sigma_{ij}^2 + \frac{(e_{ij,t} - r_{ij,t}(\phi_{ij}, \theta_{ij}))^2}{\sigma_{ij}^2}\right).$$

Denote by $\widehat{\phi}_{ij}$ and $\widehat{\theta}_{ij}$ the corresponding estimates. Note however that for another pair, say (i, j'), $j' \neq j$, one obtains estimates for α_i and $\alpha_{j'}$, where α_i is not necessarily equal to the estimate of α_i using the pair (i, j). Ideally they should be close if the GDCC model is correctly specified, and this can be viewed already as a first specification test.

To obtain the composed estimates of α and β, define the symmetric matrices A and B with entries

$$A_{ij} = \widehat{\phi}_{ij,1}\widehat{\phi}_{ij,2} \quad B_{ij} = \widehat{\theta}_{ij,1}\widehat{\theta}_{ij,2}$$

In words, A_{ij} is just the product of the estimates of α_i and α_j using the pair (i, j). The objective now is to find α and β such that $\alpha\alpha'$ is close to A and $\beta\beta'$ close to B. Ideally, one would like to solve the system of equations

$$A_{ij} = \alpha_i\alpha_j \quad \text{and} \quad B_{ij} = \beta_i\beta_j$$

for all $i \neq j$. By taking logarithms this can be written as a linear equation system, that is

$$C\log(\alpha) = \log(LT(A)) \tag{13}$$

$$C\log(\beta) = \log(LT(B)) \tag{14}$$

where C is an $(N(N-1)/2 \times N)$ matrix with a 1 at positions (k_{ij}, i) and (k_{ij}, j), where $k_{ij} = i - j + (j-1)(N - j/2)$, $i > j$, and zeros elsewhere. The operator LT stacks the lower triangular part of a symmetric matrix, excluding the diagonal, into a vector. By convention, we define the logarithm of a vector as the vector of the componentwise logarithms. It can be shown that the matrix C is of full column rank. Thus, we can define estimators of α and β as

$$\hat{\alpha} = \exp\left\{(C'C)^{-1}C'\log(LT(A))\right\},$$

$$\hat{\beta} = \exp\left\{(C'C)^{-1}C'\log(LT(B))\right\}.$$

For example, consider the case with $N = 3$. Then the system for β can be written as

$$\begin{bmatrix} 1 & 1 & 0 \\ 1 & 0 & 1 \\ 0 & 1 & 1 \end{bmatrix}\begin{bmatrix} \log\beta_1 \\ \log\beta_2 \\ \log\beta_3 \end{bmatrix} = \begin{bmatrix} \log B_{12} \\ \log B_{13} \\ \log B_{23} \end{bmatrix}$$

After matrix inversion one finds the exact solution

$$\begin{aligned}\hat{\beta}_1 &= \sqrt{B_{12}B_{13}/B_{23}}\\ \hat{\beta}_2 &= \sqrt{B_{12}B_{23}/B_{13}}\\ \hat{\beta}_3 &= \sqrt{B_{13}B_{23}/B_{12}}\end{aligned}$$

Note that in the case $N = 3$ the system is exactly determined, so that an exact solution to the equation system (13) and (14) can be found. For larger N, the system is overdetermined so that one would have to add an 'error term' v_{ij}, say, to each equation. The least squares estimates then minimize the sum of squared errors. For example, the estimator for β minimizes

$$\sum_{i<j} v_{ij}^2 = \sum_{i<j} (\log B_{ij} - \log \beta_i - \log \beta_j)^2. \tag{15}$$

A general expression for the least squares estimator for β_i for $N > 2$, can be shown to be

$$\hat{\beta}_i = \frac{\left(\prod_{j\neq i} B_{ij}\right)^{\frac{1}{N-1}}}{\left(\prod_{j,k\neq i} B_{kj}\right)^{\frac{1}{(N-1)(N-2)}}},$$

so that $\hat{\beta}_i$ is just the geometrical mean of all $B_{ij}, j \neq i$, divided by the square root of the geometrical mean of all B_{kj} with $j, k \neq i$ and $k < j$.

Rather than minimizing (15), one can also minimize directly the distance between B and $\beta\beta'$. For example, this distance can be measured by the Frobenius norm $\|\cdot\|_F$, i.e., the sum of the squared elements of $B - \beta\beta'$,

$$\begin{aligned}\|B - \beta\beta'\|_F &= \text{vec}(B - \beta\beta')'\text{vec}(B - \beta\beta')\\ &= \text{Tr}\left((B - \beta\beta')(B - \beta\beta')\right)\\ &= \text{Tr}(BB) + (\beta'\beta)^2 - 2\beta' B\beta \end{aligned} \tag{16}$$

As the first term of (16) does not depend on β, minimizing the Frobenius norm is equivalent to minimizing

$$Q(\beta) = (\beta'\beta)^2 - 2\beta' B\beta$$

As there is no analytic solution to this minimization problem, numerical algorithms have to be used.

Finally, analogous estimators can be found for α,

$$\hat{\alpha}_i = \frac{\left(\prod_{j\neq i} A_{ij}\right)^{\frac{1}{N-1}}}{\left(\prod_{j,k\neq i} A_{kj}\right)^{\frac{1}{(N-1)(N-2)}}} \tag{17}$$

and

$$\tilde{\alpha} = \arg\min_{\alpha} \left[(\alpha'\alpha)^2 - 2\alpha' A \alpha \right]. \tag{18}$$

The advantage of this individual estimation approach is the computational feasibility. A drawback, however, is that the theoretical properties of these estimators are far from clear. Even consistency is doubtful, as the regression equations (12) are linked through the denominator of $r_{ij,t}$ in (8), but these links are neglected in the estimation. In the next section, we will therefore use the estimation method described in Section 3.1.

4 Empirical results

In the following some results are given for 18 selected daily stock returns of the Frankfurt DAX30 index and 25 selected daily stock returns of the London FTSE 100 index. Both series are adjusted for dividends and stock splits. The sample period is from 1/1/1973 to 3/1/2003 for the DAX returns ($T = 7876$ observations), and from 1/1/1973 to 5/13/2003 for the FTSE returns ($T = 7921$ observations). The series were selected such that they are available over the entire sample period. In most of the FTSE returns we found significant first order autocorrelation, so that we first estimated a linear AR(1) model and continue to work with the residuals of that model in the following. The finding of first order autocorrelation is not unusual, see for instance Chapter 2 of Campbell, Lo, and MacKinlay (1997) and Hafner and Herwartz (2000) for empirical evidence.

Table 1 reports parameter estimates of the DCC model, as well as likelihood ratio statistics for testing CCC against DCC and DCC against GDCC. In both cases, the simpler model is clearly rejected. Table 2 summarizes the estimation results of the GDCC model for the DAX data. The largest estimated α_i is 0.0724, the smallest 0.0489, so the range is quite narrow. To see the difference to the DCC model, we show in Figure 1 for the stock with smallest α the estimated conditional correlation series. It is obvious that the DCC model (with a larger α estimate) implies a more volatile correlation series, whereas the GDCC model permits a correlation series that is closer to a constant, as in the CCC case.

For the FTSE data, results are reported in Table 3. The smallest estimated α is 0.0328, the largest 0.0629. As reported in Table 1, the likelihood value is significantly improved also for the FTSE data. The likelihoods for the CCC, DCC and GDCC models are, respectively, -23.0395, -22.8028 and -22.7915, yielding in both cases significant likelihood ratio statistics in favor of the more general model.

Table 3 also contains an indication of the sectors. To see if the α_i values are perhaps sector-specific, we run a regression of the estimated values in this table on an intercept and five sector dummies. No parameter for these dummies is significant, except for chemical stocks, with a t-ratio of 2.801. Deleting redundant dummies leads to the conclusion that the average value of α_i is 0.044 for all sectors, while it is 0.057 for chemicals. Hence, chemical stocks seem to have more volatile correlations with other stocks. For the DAX data, we do not have enough observations to perform a similar regression, but, interestingly, the BASF and Bayer α_i values are also higher than the average value of 0.062.

As a diagnostic test, we use the multivariate Portmanteau statistic given by (see e.g. Lütkepohl, 1993)

$$P_h = T^2 \sum_{i=1}^{h} (T-i)^{-1} \mathrm{Tr}(\hat{C}_i' \hat{C}_0^{-1} \hat{C}_i \hat{C}_0^{-1}), \quad \hat{C}_i = \frac{1}{T} \sum_{t=i+1}^{T} \hat{\xi}_t \hat{\xi}_{t-i}', \tag{19}$$

where $\hat{\xi}_t = \hat{R}_t^{-1/2} \varepsilon_t$. The statistic P_h is conjectured to have an asymptotic χ^2 distribution with hN^2 degrees of freedom. We use P_h as a measure for residual autocorrelation rather than as a formal test statistic, as to our knowledge the asymptotic theory for the present model framework has not been worked out. The value of P_{10} for the CCC model applied to the FTSE data is 1,314,287.6, that for the DCC model is 1,296,375.8 and that of the GDCC model is 1,245,106. All are higher than the 5% critical value of a χ^2_{6250} distribution. This may indicate remaining residual autocorrelation, but it also shows that the GDCC model provides a better fit to the data.

For the sake of completeness, we also report the results for the combined individual estimates described in Section 3.2, applied to the FTSE data. The mean of $\hat{\alpha}$ in (17) is 0.1241 with a standard deviation of 0.0306. The mean of $\tilde{\alpha}$ in (18) is 0.1370 with a standard deviation of 0.0349. All $\hat{\alpha}_i$'s are close to the $\tilde{\alpha}_i$'s but tend to be slightly smaller. However, both are substantially larger than the estimates using simultaneous estimation, and this may indicate the inconsistency of the approach. The likelihood of the individual estimates is -23.4656, much smaller even than the likelihood of the CCC model. Surprisingly, however, the individual estimates have smaller Portmanteau statistics of 912,583.9 and 876847.8, respectively.

As another specification test of the models, we can apply the estimated models to the problem of finding the minimum variance portfolio. This has become a standard criterion to evaluate the performance of models for the covariance of stock returns, see also Chan, Karceski, and Lakonishok (1999). It is well known since Markowitz that the optimal weight vector at time t is given by

$$w_t = \frac{H_t^{-1} \iota}{\iota' H_t^{-1} \iota},$$

where ι is an $(N \times 1)$ vector of ones. If the model for H_t is correctly specified, then this weight vector should provide the minimum variance portfolio. For the DAX (FTSE) data, the variance of the portfolio that uses H_t estimated by the standard DCC model is about 4% (1%) higher than the one that uses the GDCC model.

For the same criterion, Engle and Sheppard (2001) report that the DCC model performs well for small number of assets, but that the model fails to find the optimal portfolio for N increasing. Another interesting phenomenon they find is that the estimated α parameter of the DCC model decreases when the number of assets is increased.

To see whether these features can be explained by different correlation sensitivities α, consider a small simulation study. Let us assume that the volatility part of the model does not play a role here, so to simplify we set $D_t = I_N$, such that $H_t = R_t$. The following results were checked for robustness with respect to this assumption, and no counter-evidence was found. We generate time series y_t following a GDCC model with multinormal innovations. Before each simulation, a realization of an $N \times 1$ parameter vector α is drawn from the distribution

$$\alpha_i \sim Beta(10, 90) \tag{20}$$

which implies a population mean of 0.1 and a standard deviation of 0.0298. This is close to the empirical moments of the reported estimates in Section 3. The autoregressive part of Q_t is fixed at $\beta = 0.999 - \max(\alpha)^2$, so that the maximum persistence, measured by $\alpha_i^2 + \beta$, is given by 0.999. The unconditional correlation matrix is computed by drawing a random $N \times N$ matrix Z of a uniform distribution letting $Z^* = Z \odot Z$ and $S = \text{diag}(Z^*)^{-1/2} Z^* \text{diag}(Z^*)^{-1/2}$. This gives unconditional correlations similar to what is observed in the stock data.

We generate 500 time series, each of length 1000, calculate for each series the minimum variance portfolio using either CCC, DCC or GDCC. The CCC model implies $R_t = S$, so for every simulated series the generated S matrix is used to compute w_t. For the DCC model we use for every simulated series the mean of the generated α_i parameters. That is, the approximating DCC model reads $Q_t = S(1 - \bar{\alpha}^2 - \beta) + \bar{\alpha}^2 \varepsilon_{t-1}\varepsilon_{t-1}' + \beta Q_{t-1}$. This should provide a reasonable approximation to the true GDCC model, however, see below for some comments on this issue. For the GDCC model, we use the generated α_i parameters.

To assess the relative performances, we then calculate the ratios of the CCC and DCC portfolio variances with respect to the optimal GDCC one. Table 4 reports the means and standard errors of these ratios. As can be seen, the ratios tend to increase with the number of assets N. For small N it does not seem to make a difference whether to use

DCC or GDCC, but for large N the difference becomes more and more important. This holds true even though the distribution of the α is kept fixed. The interpretation of this result is that, as N increases, it is more likely to have one asset that has a correlation sensitivity α_i in the tails of the distribution, so that the assumption that all αs are the same becomes too restrictive and yields sub-optimal portfolios. In sum, this could be the explanation of the failure of the DCC model to correctly identify the minimum variance portfolio in high dimensions, as reported by Engle and Sheppard (2001).

Finally, we considered the issue of approximating a GDCC process by a DCC model more closely. Table 5 reports estimates of a DCC model applied to 50 generated GDCC models, where the GDCC parameters are again generated by (20). The striking result is that the estimated α parameter tends to decrease with the dimension N. Moreover, for $N \geq 30$ the estimated α is significantly smaller than the mean of the true parameter distribution. This could explain yet another empirical phenomenon of the DCC model, namely the decreasing parameter estimates when the number of assets is increased, see for example Tables 1 and 2 of Engle and Sheppard (2001) who use S&P500 and DJIA stocks. We also tried the estimated DCC α parameter instead of the means of the GDCC parameters in the minimum variance portfolio simulations, but did not find substantial differences.

To summarize these simulation experiments, we find evidence that two empirical phenomena of DCC models could be explained by the imposed restriction when applied to a process that has a diversity of correlation sensitivities, such as the GDCC model. These phenomena are the failure of the DCC model to identify the minimum variance portfolio in high dimensions, and the decreasing α parameter estimates when the dimension is increased.

5 Conclusion

In this paper we proposed an extended DCC model that allows for asset-specific heterogeneity in the correlation structure. The model was successfully fitted to DAX and FTSE series, and it significantly improved on the DCC model in various dimensions.

A next topic of research in this area amounts to the interpretation of this heterogeneity. In this paper we simply ran a regression of estimated parameters on sector dummies, but more elegant approaches exist. One of them is to assume that the α_i's also are the outcomes of a model with explanatory variables and an error term. This multi-level model allows then for a further reduction of the number of parameters.

	DAX	FTSE
N	18	25
T	7876	7921
sample period	1/1/1973–3/11/2003	1/1/1973–5/13/2003
Parameter	estimate (std.err)	estimate (std.err)
α	0.0038 (0.0003)	0.0021 (0.0001)
β	0.9944 (0.0005)	0.9957 (0.0004)
L_{CCC}	5776	3750.8
L_{DCC}	122.87	178.7

Table 1: *Estimation results of the DCC model for the DAX and FTSE returns.* L_{CCC} *is the value of the likelihood ratio statistic that tests the CCC model against the DCC model, and* L_{DCC} *the statistic that tests the DCC model against our GDCC model. Both are larger than the 1% critical values of the asymptotic distribution.*

Stock	α_i	standard error	Stock	α_i	standard error
Allianz	0.0680	0.0065	Lufthansa	0.0489	0.0043
BASF	0.0636	0.0046	MAN	0.0596	0.0049
Hypo-Bank	0.0593	0.0044	Münchner Rück	0.0575	0.0042
BMW	0.0619	0.0055	RWE	0.0614	0.0038
Bayer	0.0675	0.0054	Schering	0.0568	0.0058
Commerzbank	0.0666	0.0058	Siemens	0.0696	0.0047
Deutsche Bank	0.0714	0.0049	Thyssen	0.0632	0.0068
E.ON	0.0724	0.0042	TUI	0.0571	0.0049
Linde	0.0513	0.0037	VW	0.0663	0.0057

Table 2: *Estimation results of the GDCC model for the 18 DAX returns, 1973–2003. The estimate of β is 0.9942 with a standard error of 0.0008. The mean of the α_is is 0.062 with a standard deviation of 0.0065.*

Stock	α_i	std. err.	Stock	α_i	std. err.
Allied Domecq (R)	0.0472	0.0024	Diageo (R)	0.0449	0.0029
Amvescap (F)	0.0426	0.0041	Dixons (R)	0.0407	0.0036
Assd. Brit. Foods (R)	0.0342	0.0030	EMAP (M)	0.0453	0.0040
Aviva (F)	0.0465	0.0030	EXEL (T)	0.0384	0.0033
Barclays (F)	0.0458	0.0027	Foreign & Colonial (F)	0.0482	0.0024
BOC (C)	0.0494	0.0025	GKN (T)	0.0461	0.0027
Boots (R)	0.0541	0.0025	Glaxosmithkline (C)	0.0629	0.0028
BP (O)	0.0418	0.0034	Granada (M)	0.0452	0.0042
Brit. Ame. Tobacco (O)	0.0522	0.0029	GUS (R)	0.0469	0.0032
British Land (F)	0.0396	0.0033	Hanson (O)	0.0502	0.0025
BUNZL (O)	0.0328	0.0041	Hilton (O)	0.0454	0.0032
Cadbury Schweppes (R)	0.0419	0.0024	IMP (C)	0.0591	0.0024
Daily Mail (M)	0.0340	0.0039			

Table 3: *Estimation results of the GDCC model for the 25 FTSE returns, 1973–2003. Associated sectors are given in parentheses: R: Food, Beverages, Retail, F: Banks, Insurance, Real Estate, C: Chemicals, M: Media, T: Transport, O: Other. The estimate of β is 0.996 with a standard error of 0.0001. The mean of the α_is is 0.045 with a standard deviation of 0.007.*

	DCC		CCC	
N	mean	std err	mean	std err
3	1.0228	0.0021	3.8495	0.8243
4	1.0357	0.0025	4.2428	0.7041
5	1.0560	0.0032	6.1709	1.1595
10	1.1180	0.0054	11.4363	2.0441
15	1.1814	0.0056	18.0791	3.6982
20	1.2267	0.0061	35.3159	7.6578
25	1.2761	0.0063	53.6910	9.5502
30	1.3245	0.0108	115.0078	40.3950

Table 4: *Ratios of variances of the minimum variance portfolios. The series were generated by GDCC using (20), then the variance of the minimum variance portfolio using the best DCC and CCC approximation is divided by the GDCC variance. If the ratio is close one, the restricted model (DCC or CCC) does not differ in determining the minimum variance portfolio.*

N	mean	std err
2	0.1080	0.0063
3	0.1088	0.0034
4	0.1109	0.0039
5	0.1091	0.0026
10	0.1076	0.0018
20	0.0993	0.0009
30	0.0945	0.0012
40	0.0929	0.0008
50	0.0907	0.0009
100	0.0811	0.0005

Table 5: *Means and standard errors of estimated α parameters in the DCC model, where 50 processes of length 1000 were generated by a GDCC model. The parameters of the GDCC model are generated according to (20), which implies a mean of 0.10 and a standard deviation of 0.0298.*

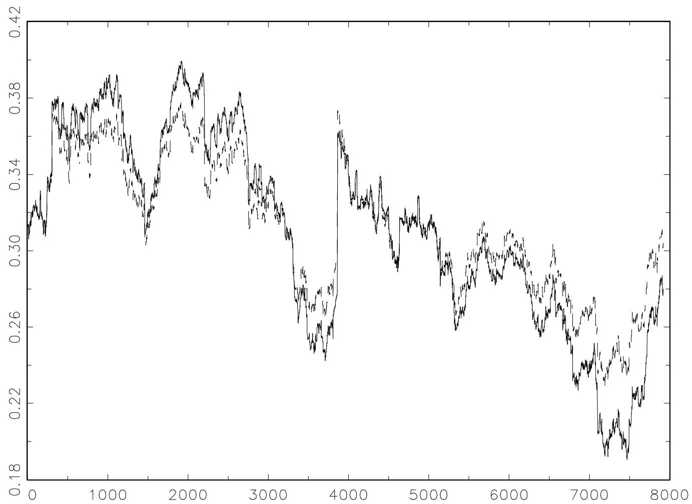

Figure 1: *Estimated conditional correlation for Lufthansa with BASF. Solid line: DCC estimate, dashed line: GDCC estimate.*

References

Bauwens, L., Laurent, S., Rombouts, J.V.K. (2003) Multivariate GARCH models: A survey, CORE discussion paper 2003/31, Université catholique de Louvain.

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.

Campbell, J., A. Lo, A. MacKinlay (1997). *The Econometrics of Financial Markets*, Princeton University Press, Princeton, New Jersey.

Chan, L.K.C., Karceski, J. and Lakonishok, J. (1999). On portfolio optimization: Forecasting covariances and choosing the risk model, *The Review of Financial Studies*, 12, 937–674.

Engle, R.F. (2002). Dynamic conditional correlation - A simple class of multivariate GARCH models, *Journal of Business and Economic Statistics* 20, 339–350.

Engle, R.F. and Sheppard, K. (2001). Theoretical and empirical properties of dynamic conditional correlation multivariate GARCH, UCSD DP 2001-15.

Hafner, C.M. and H. Herwartz (2000). Testing for linear autoregressive dynamics under heteroskedasticity,*The Econometrics Journal* 3, 177–197.

Lütkepohl, H. (1993). *Introduction to Multiple Time Series Analysis*, 2nd ed., Berlin, New York: Springer Verlag.