INSTITUT DE
STATISTIQUE

UNIVERSITÉ CATHOLIQUE DE LOUVAIN

DISCUSSION
PAPER

0904

A GENERALIZED DYNAMIC CONDITIONAL CORRELATION MODEL: SIMULATION AND APPLICATION TO MANY ASSETS

HAFNER, C. M. and P. H. FRANSES

A Generalized Dynamic Conditional Correlation Model: Simulation and Application to Many Assets

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

January 14, 2009

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. We propose two estimation methods, one based on a full likelihood maximization, the other on individual correlation estimates. The resultant GDCC model is considered for daily data on 39 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]Institut de statistique, Université catholique de Louvain, Voie du Roman Pays, 20, 1348 Louvain-la-Neuve, Belgium, `christian.hafner@uclouvain.be`. Financial support from the contract *Projet d'Actions de Recherche Concertées* nr. 07/12/002 of the *Communauté française de Belgique*, granted by the *Académie universitaire Louvain*, is gratefully acknowledged.

[2]Econometric Institute, Erasmus University Rotterdam, P.O. Box 1738, NL–3000 DR Rotterdam, The Netherlands, `franses@few.eur.nl`. We thank the detailed comments made by two anonymous referees.

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, in which volatilities and correlations at a given time are functions of lagged returns. The alternative class of multivariate stochastic volatility models specifies latent stochastic processes, but estimation is typically more involved, see e.g. Asai et al. (2006) for a recent review. By now, there are many variants of multivariate GARCH models available, see Bauwens et al. (2006). The current benchmark models seem to be the Constant Conditional Correlation (CCC) model of Bollerslev (1990) and its extensions, e.g. the vector ARMA-GARCH model of Ling and McAleer (2003) and the Dynamic Conditional Correlation (DCC) model of Engle (2002). A model similar to the DCC model has been proposed by Tse and Tsui (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. However, the imposed restrictions may be too strong in the case of many assets. For example, McAleer, Hoti and Chan (2009) extend the vector ARMA-GARCH model of Ling and McAleer (2003) to allow for asymmetric effects in volatilities. Several extensions of the standard DCC model have been suggested, e.g. Cappiello, Engle and Sheppard (2006) and McAleer et al. (2008). The proposed extensions of the DCC model allow for flexible modelling of correlation dynamics, but it seems that their estimation becomes difficult in very high dimensions.

In this paper we aim to extend the DCC model by focussing on the notion that one would like to use this model for a large number of asset returns. For example, one might want to summarize 39 important stocks in the FTSE 100 stock index 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 (such as 30 or more). 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 the DCC 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 39 stock returns in the FTSE index. We document that the GDCC model improves on the more restrictive models in various dimensions. In Section 5, we provide simulation results which show that certain anomalies of the standard DCC model can be explained by mis-specification if the true model is GDCC, and Section 6 concludes 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 FTSE 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

$$\begin{aligned} H_t &= D_t R_t D_t && (1) \\ R_t &= \mathrm{diag}(Q_t)^{-1/2} Q_t \mathrm{diag}(Q_t)^{-1/2} && (2) \\ Q_t &= S(1-\alpha-\beta) + \alpha \varepsilon_{t-1}\varepsilon'_{t-1} + \beta Q_{t-1} && (3) \end{aligned}$$

where α and β are non-negative 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. The term $\mathrm{diag}(Q_t)$ 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

$$
\begin{aligned}
H_t &= D_t R_t D_t && (4)\\
R_t &= \mathrm{diag}(Q_t)^{-1/2} Q_t \mathrm{diag}(Q_t)^{-1/2} && (5)\\
Q_t &= S\left(1-\bar{\alpha}^2-\bar{\beta}^2\right) + \alpha\alpha' \odot \varepsilon_{t-1}\varepsilon_{t-1}' + \beta\beta' \odot Q_{t-1} && (6)
\end{aligned}
$$

where $\odot$ denotes the Hadamard matrix product operator, i.e., elementwise multiplication. In (6), α and β are $N \times 1$ parameter vectors, $\bar{\alpha} = N^{-1}\sum_{i=1}^{N}\alpha_i$ and $\bar{\beta} = N^{-1}\sum_{i=1}^{N}\beta_i$. In order to ensure stationarity and positivity of Q_t we impose the restrictions $\alpha_i \geq 0$, $\beta_i \geq 0$ and $\alpha_i^2 + \beta_i^2 < 1$, $i = 1, \ldots, N$. If some $\alpha_i = 0$, then β_i is not identified, so that in that case we set $\beta_i = 0$.

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. By defining the diagonal matrices $A = \mathrm{diag}(\mathrm{vech}(\alpha\alpha'))$ and $B = \mathrm{diag}(\mathrm{vech}(\beta\beta'))$, where vech denotes the operator that stacks the lower triangular portion of a symmetric matrix into a vector, we can rewrite (6) as

$$
\mathrm{vech}(Q_t) = \left(1-\bar{\alpha}^2-\bar{\beta}^2\right)\mathrm{vech}(S) + A\mathrm{vech}(\varepsilon_{t-1}\varepsilon_{t-1}') + B\mathrm{vech}(Q_{t-1}),
$$

which can be viewed as a diagonal vec multivariate GARCH model suggested by Bollerslev, Engle and Wooldridge (1988), but applied to the matrix Q_t rather than H_t.

Our model is related to a similar specification for the BEKK model in Ding and Engle (2001). A model even more general than (6) has been mentioned by Engle (2002) in equations (24) and (25), where the matrices $\alpha\alpha'$ and $\beta\beta'$ are written as general unrestricted parameter matrices. Two problems arise from this general specification: one is that of an exceedingly large number of parameters when the dimension increases, the other is that of keeping Q_t positive definite. Our model (6) can be seen as a compromise between the standard DCC model and the general model proposed by Engle (2002), restricting the parameter matrices to be positive semi-definite of rank one. Furthermore, we note that the basic DCC model is closely related to Case 3 of the generalized autoregressive conditional correlation (GARCC) model of McAleer et al. (2008), while the GDCC model is closely related to Case 4 of the GARCC model.

Note that the exact variance targeting approach as in the standard DCC model is not feasible in the GDCC model (6), 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 $X_{t-1}^2 I(X_{t-1} < \tau)$, where $I(\cdot)$ is the indicator function and X_t is e.g. the return on the market index, because it may be that correlations increase in crash situations where the index return is smaller than a threshold τ. The parameter τ could be either fixed *a priori* or estimated jointly with the other parameters.

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.

Finally, note that extensions of the GDCC model are possible to allow for asymmetries and multiple regimes. For example, similar to Cappiello, Engle and Sheppard (2006), we can define $\eta_t = \varepsilon_t \odot I(\varepsilon < 0)$. Then (6) could be replaced by

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

where δ is an additional $(N \times 1)$ parameter vector. This model would allow for a different impact of negative standardized residuals on correlations than positive ones.

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 regularity conditions these estimators will be consistent and asymptotically normal, see McAleer et al. (2008). 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 used this two-step estimation procedure in the empirical part of Hafner and Herwartz (2003). Inference concerning the correlation parameter vector θ_C has to take the first step into account, as described by Engle and Sheppard (2001).

3.2 Combining individual correlation estimates

Maximization of the likelihood function may be cumbersome if the dimension N is high. 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. This is motivated by the insight that estimation of many small dimensional models may be easier than estimation of one high dimensional model. 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. A similar idea has been followed by Ledoit, Santa Clara and Wolf (2003) for the diagonal vec model. For the standard DCC model, Engle (2008) proposes to use means or medians of individually estimated α and β parameters. In the GDCC model, however, the difficult part is to restrict the univariate models such that the composed multivariate model forms valid covariance matrices. In this section, we discuss one way of solving this problem.

By definition, $r_{ij,t}(\phi_{ij}, \theta_{ij})$ is the conditional correlation of $\varepsilon_{i,t}$ and $\varepsilon_{j,t}$, where $\phi_{ij} = (\alpha_i, \alpha_j)$, $\theta_{ij} = (\beta_i, \beta_j)$. The dependence of $r_{ij,t}$ on S_{ij} is not indicated explicitly because S is treated as given from a first step estimation. Extracting the likelihood for the pair (i, j) from the joint likelihood (10), then we can estimate ϕ_{ij} and θ_{ij} by maximizing

$$L_{ij}(\phi_{ij}, \theta_{ij}) = -\frac{1}{2}\sum_{t=1}^{T}\left(\log(1 - r_{ij,t}^2) + \frac{\varepsilon_{i,t}^2 + \varepsilon_{j,t}^2 - 2r_{ij,t}\varepsilon_{i,t}\varepsilon_{j,t}}{1 - r_{ij,t}^2}\right), \tag{11}$$

which are the same estimators as in Engle (2008). Denote these estimators by $\widehat{\phi}_{ij}$ and $\widehat{\theta}_{ij}$. 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 and asymptotically identical if the GDCC model is correctly specified.

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, where we use an L_2-distance of the elementwise logarithms of these matrices. 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

$$C\log(\alpha) = \log(a) \tag{12}$$

$$C\log(\beta) = \log(b) \tag{13}$$

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 vectors a and b are defined as $a = LT(A)$ and $b = LT(B)$, respectively, where 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

$$\hat{\alpha} = \exp\left\{(C'C)^{-1}C'\log(a)\right\}, \tag{14}$$

$$\hat{\beta} = \exp\left\{(C'C)^{-1}C'\log(b)\right\}. \tag{15}$$

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

$$\hat{\beta}_1 = \sqrt{B_{12}B_{13}/B_{23}}$$

$$\hat{\beta}_2 = \sqrt{B_{12}B_{23}/B_{13}}$$

$$\hat{\beta}_3 = \sqrt{B_{13}B_{23}/B_{12}}$$

Note that in the case $N = 3$ the system is exactly determined, so that an exact solution to the equation system (12) and (13) 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

$$\sum_{i=1}^{N-1}\sum_{j=i+1}^{N} v_{ij}^2 = \sum_{i=1}^{N-1}\sum_{j=i+1}^{N} (\log B_{ij} - \log\beta_i - \log\beta_j)^2. \tag{16}$$

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

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

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 $k < j$. An analogous estimator is obviously defined for α.

Rather than minimizing (16), one could minimize directly a distance measure between B and $\beta\beta'$, for example the Frobenius norm of $B - \beta\beta'$. In general, however, there is no analytic solution to this minimization problem and numerical algorithms would have to be used.

The advantage of the individual estimation approach is its computational feasibility. Since the logarithm of the pooled estimator of α_i is a weighted average of individual estimates, its properties can be derived similar to Engle (2008) and Engle, Shephard and Sheppard (2008). Of course it is less efficient than the joint estimator, which is however infeasible in high dimensions.

In order to show the asymptotic result we need to introduce some more notation. First, we stack all likelihood functions L_{ij} in (11) with $i < j$ into a column vector l of length $K = N(N-1)/2$ with components $l_k = \sum_{t=1}^{T} l_{kt}$. Furthermore, we define $\lambda_k = (\eta_{1k}', \eta_{2k}', s_k)'$, where η_{ik} is the parameter vector of the score ζ_{ikt} of the volatility part L_V, $i = 1, 2$, and where s_k is the k-th element of $LT(S)$. Then, let

$$F_{k,T} = \sum_{t=1}^{T} \frac{\partial^2 l_{kt}}{\partial\alpha\partial\lambda_k'} \left(\sum_{t=1}^{T} \frac{\partial g_{kt}}{\partial\lambda_k'}\right)^{-1}$$

and $g_{kt} = (\zeta_{1kt}', \zeta_{2jt}', T^{-1}\sum_{t=1}^{T} e_{kt}' - s_k')'$, where e_{kt} is the k-th element of $LT(\varepsilon_t\varepsilon_t')$. Finally, denote

$$D_{k,T} = \frac{1}{T}\sum_{t=1}^{T} \frac{\partial l_{kt}^2}{\partial\alpha\partial\alpha'} - F_{k,T}\frac{1}{T}\sum_{t=1}^{T} \frac{\partial g_{kt}}{\partial\alpha'}.$$

Consider for example the asymptotic distribution of $\hat{\alpha}$. Let $\gamma = \log \alpha$ and $\hat{\gamma} = \log \hat{\alpha}$. From (14), we have $\hat{\gamma} = W \log a$, where $W = (C'C)^{-1}C'$ is a fixed weight matrix. Then we can show that under the conditions given by Engle, Shephard and Sheppard (2008),

$$\sqrt{T}(\hat{\gamma} - \gamma) \to_d N(0, WD^{-1}VD^{-1}W'),$$

where $D = \text{plim} D_T$ and D_T is a block-diagonal matrix with block elements given by $D_{k,T}$, $k = 1, \ldots, K$. Furthermore, $V = \lim_{T\to\infty} \text{Var}(T^{-1/2} \sum_{t=1}^{T} Z_t)$, where $Z_t = (Z'_{1t}, \ldots, Z'_{Kt})'$ and $Z_{kt} = \partial l_{kt}/\partial\alpha - F_{k,T} g_{kt}$ is the projected score function. A HAC estimator can be used to estimate V. Note however that Engle et al. (2008) do not prove that V can be estimated consistently, but it is one of the assumptions. Finally, the asymptotic distribution of $\sqrt{T}(\hat{\alpha} - \alpha)$ follows by applying the delta method. We will use this estimation and inference method in the following empirical section.

4 Empirical application

In the following some results are given for 39 selected daily stock returns of the London FTSE 100 index. The series is adjusted for dividends and stock splits. The sample period is from 1/1/1973 to 12/8/2008 for the FTSE returns ($T = 9375$ 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.

The joint maximum likelihood estimates of α and β of the standard DCC model are 0.0015 and 0.9968, respectively. Note the very small numerical value of the estimate of α, which hints at a negative bias that was noted already by Engle and Sheppard (2001). We will see in our simulation study that one of the reasons for this bias may be neglected heterogeneity of α_i parameters if the GDCC model was correct. The estimates of α and β using the 'MacGyver' estimator of Engle (2008) that does not suffer from this bias are 0.0086 and 0.9815, respectively. We use these estimates for the comparisons of the alternative models.

Estimation results for the GDCC model are reported in Table 1. The smallest estimated α is 0.0769 (Land), the largest 0.1331 (Diageo). The average standard error of α is 0.0047, that of β 0.0009, such that all parameters are highly significant. For the two stocks with smallest α (Land Securities and WPP), Figure 1 shows the estimated

conditional correlation series using the DCC model and Figure 2 using the GDCC model. Obviously, the DCC model (with a larger α estimate) implies a more volatile correlation series, whereas the GDCC model delivers a smoother, less erratic correlation series. The converse holds for the stocks with highest GDCC α estimates, Diageo and GSK. Figures 3 and 4 show the corresponding estimated correlation series, where now the DCC model yields a smoother series than the GDCC model. Thus, while for the DCC model the degree of sensitivity of correlation to news is the same for all stocks, the GDCC model allows to obtain estimated correlations that differ in terms of this sensitivity.

To obtain an idea of the distribution of estimated correlations, Figures 5 and 6 show the median and the 5 and 95 percent quantiles of the 741 correlations for each day. We see a very similar shape but the GDCC correlations are slightly more dispersed with the 5 and 95 percent quantiles lying further in the tails. This corresponds to the finding that there are some stocks with relatively large correlation sensitivity α whose estimated correlations attain more extreme values than using a uniform α parameter.

Table 1 also contains an indication of the sectors. To see if the α_i values are sector-specific, we run a regression of the estimated values in this table on an intercept and six sector dummies. We selected the sectors oil and gas, food and retail, finance, media, mining and chemicals. No parameter for these dummies is significant except for stocks in the retail sector with a t-statistic of 2.18 and corresponding p-value of 0.036. Hence, retail stocks have more volatile correlations than other sectors. For a reduced sample until 2003, Hafner and Franses (2003) report significant sector dummy for chemicals. Even though GSK still has a high α estimate, the chemicals dummy becomes insignificant for our new sample.

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

$$P_h = T^2 \sum_{i=1}^{h} (T-i)^{-1} \mathrm{Tr}(\hat{C}_i' \hat{C}_0^{-1} \hat{C}_i \hat{C}_0^{-1}), \quad \hat{C}_i = \frac{1}{T} \sum_{t=i+1}^{T} \hat{\xi}_t \hat{\xi}_{t-i}', \tag{17}$$

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 5,839,992.4, that for the DCC model is 5,568,228.0 and that of the GDCC model is 5,245,814.5. All are higher than the 5% critical value of a χ^2_{15210} distribution. This may indicate remaining residual autocorrelation, but it also shows that the GDCC model provides a better fit to the data.

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 one important criterion to evaluate the performance of models for the covariance of stock returns, see also Chan, Karceski, and Lakonishok (1999), as it does not depend on a correct specification of the conditional mean vector. 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 our data, the variance of the portfolio that uses H_t estimated by the standard DCC model is about 5% higher than the one that uses the GDCC model.

For the same minimum variance criterion, Engle and Sheppard (2001) report that the DCC model performs well for a small number of assets, but that the model fails to find the optimal portfolio as N increases. Another interesting phenomenon 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, we will consider a small simulation study in the next section.

Other criteria have been used, for example minimizing tracking error variance as in Chan, Karceski, and Lakonishok (1999) and calculating fees that an investor would pay to switch from one model to another as in Fleming et al. (2001) and Fleming et al. (2003). However, we think that in our context the minimum variance analysis clearly detects empirical differences between the DCC and GDCC model. A more detailed study including the above mentioned approaches is left for future research.

5 A 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 the restricted GDCC model (7) with multinormal innovations. Before each simulation, a realization of an $N \times 1$ parameter vector α is drawn from the distribution

$$\alpha_i \sim Beta(10, 90) \tag{18}$$

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 4. 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, but we will give some comments on this issue below. 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 2 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 3 reports estimates of a DCC model applied to 50 generated GDCC models, where the GDCC parameters are again generated by (18). 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 phe-

nomena 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.

6 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. One could also develop specifications to test for correlation spillover, similar to volatility spillover or causality as proposed by Caporin (2007). This is also left for future research.

References

Asai, M., McAleer, M., Yu, J. (2006). Multivariate stochastic volatility: A review. *Econometric Reviews* 25: 145-175.

Bauwens L., Laurent, S., Rombouts, J.V.K. (2006). Multivariate GARCH models: a survey. *Journal of Applied Econometrics* 21: 79-109.

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.

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

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

Cappiello, L., Engle, R.F., Sheppard, K. (2006). Asymmetric dynamics in the correlations of global equity and bond returns, Journal of Financial Econometrics 4: 537-572.

Caporin, M. (2007). Variance (non)causality in multivariate GARCH, *Econometric Reviews* 26: 124.

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

Ding, Z., Engle, R.F. (2001). Large scale conditional covariance modeling, estimation and testing, *Academia Economic Papers* 29: 157-184.

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

Engle, R.F. (2008). High dimensional dynamic correlations. In J.L. Castle and N. Shephard (Eds.), The Methodology and Practice of Econometrics: Papers in Honour of David F. Hendry. Oxford University Press, forthcoming.

Engle, R.F., Shephard, N., Sheppard, K. (2008). Fitting vast dimensional time-varying covariance models, discussion paper.

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

Fleming, J., Kirby, C., Ostdiek, B. (2001). The economic value of volatility timing, *Journal of Finance* 56: 329-352.

Fleming, J., Kirby, C., Ostdiek, B. (2003). The economic value of volatility timing using realized volatility, *Journal of Financial Economics* 67: 473-509.

Hafner, C.M., Franses, Ph.H. (2003). A generalized dynamic conditional correlation model with applications to many asset returns, Econometric Institute Report 18 (2003), Erasmus Universiteit Rotterdam.

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

Ledoit, O., Santa Clara, P., Wolf, M. (2003), Flexible multivariate GARCH modeling with an application to international stock markets, *Review of Economics and Statistics* 85: 735–747.

Ling S., McAleer, M. (2003). Asymptotic theory for a vector ARMA-GARCH model. *Econometric Theory* 19: 280-310.

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

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

McAleer, M., Hoti, S., Chan, F. (2009). Structure and asymptotic theory for multivariate asymmetric conditional volatility, *Econometric Reviews*, forthcoming.

Tse, Y.K., Tsui, A.K.C. (2002). A multivariate GARCH model with time-varying correlations, *Journal of Business and Economic Statistics* 20: 351–362.

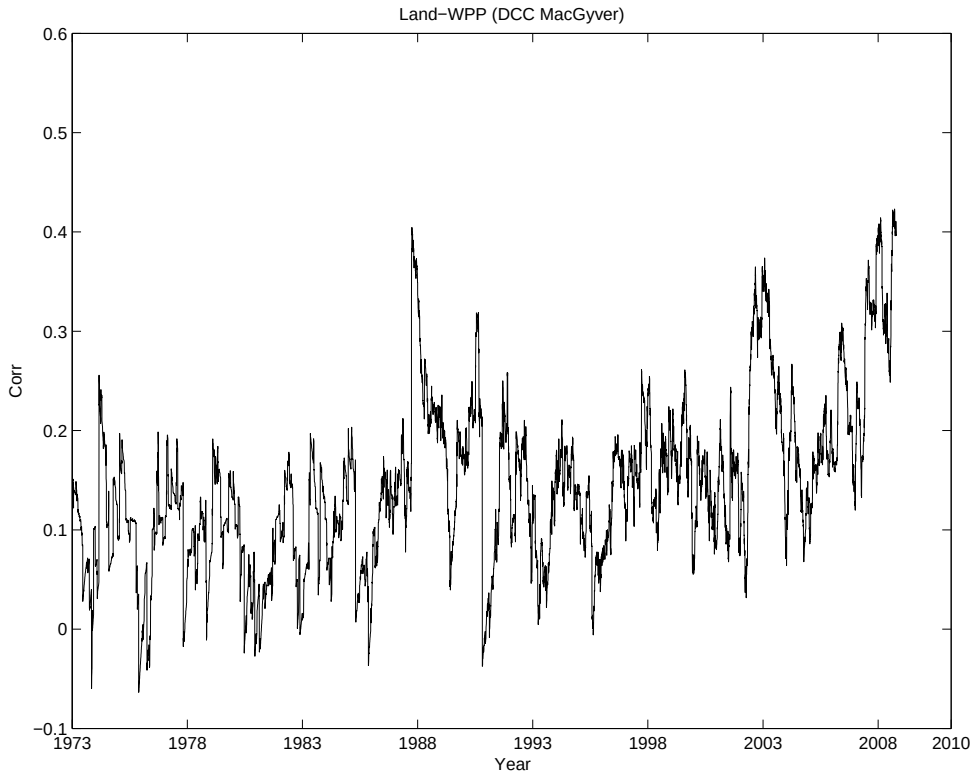

Figure 1: *Estimated conditional correlation for Land Securities and WPP using the DCC model. The MacGyver estimators of* α *and* β *are 0.0086 and 0.9815, respectively.*

		AR(1)	GARCH		GDCC	
Stock	Sector	ϕ	α_G	β_G	α	β
ALLIANCE TRUST	F	0.0454	0.1287	0.8502	0.0981	0.9913
ASSOCIATED BRIT.FOODS	R	0.0064	0.0701	0.9167	0.0938	0.9894
AVIVA	F	0.0167	0.0673	0.9233	0.0812	0.9936
BARCLAYS	F	0.0743	0.0789	0.9086	0.1144	0.9880
BP	G	0.0392	0.0515	0.9398	0.0852	0.9897
BRITISH AMERICAN TOBACCO	O	0.0299	0.0631	0.9184	0.0924	0.9889
BRITISH LAND	F	-0.0208	0.0744	0.9197	0.1155	0.9872
BUNZL	O	0.0926	0.0446	0.9428	0.0793	0.9941
CADBURY	R	0.0488	0.0818	0.8944	0.1044	0.9899
COBHAM	O	0.0385	0.0678	0.9169	0.1171	0.9814
DIAGEO	R	0.0320	0.0549	0.9357	0.1331	0.9808
GLAXOSMITHKLINE	C	0.0431	0.0835	0.8907	0.1196	0.9869
HAMMERSON	F	0.1090	0.0575	0.9293	0.1046	0.9887
INVENSYS	O	0.0646	0.0276	0.9696	0.0909	0.9929
JOHNSON MATTHEY	C	0.0276	0.0166	0.9821	0.0786	0.9932
LAND SECURITIES GROUP	F	0.0745	0.0596	0.9316	0.0769	0.9938
LEGAL & GENERAL	F	0.0080	0.0655	0.9239	0.0941	0.9903
LONMIN	Mi	0.0403	0.0958	0.8807	0.0854	0.9906
MARKS & SPENCER GROUP	R	-0.0060	0.0784	0.9068	0.0933	0.9881
MORRISON(WM)SPMKTS	R	0.0295	0.1051	0.8316	0.1002	0.9899
NEXT	R	0.0349	0.0553	0.9355	0.1037	0.9893
PEARSON	M	0.0786	0.0609	0.9283	0.1032	0.9820
PRUDENTIAL	F	0.0343	0.0991	0.8799	0.0913	0.9918
RECKITT BENCKISER	R	0.0624	0.0653	0.9189	0.1040	0.9885
REED ELSEVIER	M	0.0435	0.0692	0.9162	0.0897	0.9922
REXAM	O	0.0401	0.0422	0.9515	0.0822	0.9941
RIO TINTO	Mi	0.0544	0.0612	0.9317	0.0810	0.9923
ROYAL BANK OF SCTL.GP	F	0.0819	0.1306	0.8451	0.0801	0.9946
ROYAL DUTCH SHELL	G	0.0353	0.0555	0.9368	0.0825	0.9929
RSA INSURANCE GROUP	F	0.0546	0.0519	0.9406	0.0784	0.9932
SCHRODERS	F	-0.0161	0.0733	0.9165	0.0929	0.9905
SMITH & NEPHEW	O	0.0259	0.0786	0.9005	0.0938	0.9912
SMITHS GROUP	O	0.0862	0.0692	0.9095	0.0859	0.9926
STANDARD CHARTERED	F	0.0717	0.0961	0.8847	0.0846	0.9930
TESCO	R	0.0356	0.0694	0.9132	0.0994	0.9844
UNILEVER (UK)	R	0.0517	0.0517	0.9391	0.1114	0.9812
WHITBREAD	R	0.0659	0.0775	0.9036	0.1032	0.9831
WOLSELEY	O	0.0654	0.0789	0.9003	0.0799	0.9962
WPP	M	0.1158	0.0455	0.9455	0.0792	0.9947

Table 1: *Estimation results of the GDCC model for the 39 FTSE returns, 1973–2008. Associated sectors are given in parentheses: R: Food, Beverages, Retail, F: Banks, Insurance, Real Estate, C: Chemicals and Pharmaceuticals, M: Media, Mi: Mining, G: Gas and Oil, O: Other.*

	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 2: *Ratios of variances of the minimum variance portfolios. The series were generated by GDCC using (18), 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 to 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 3: *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 (18), which implies a mean of 0.10 and a standard deviation of 0.0298.*

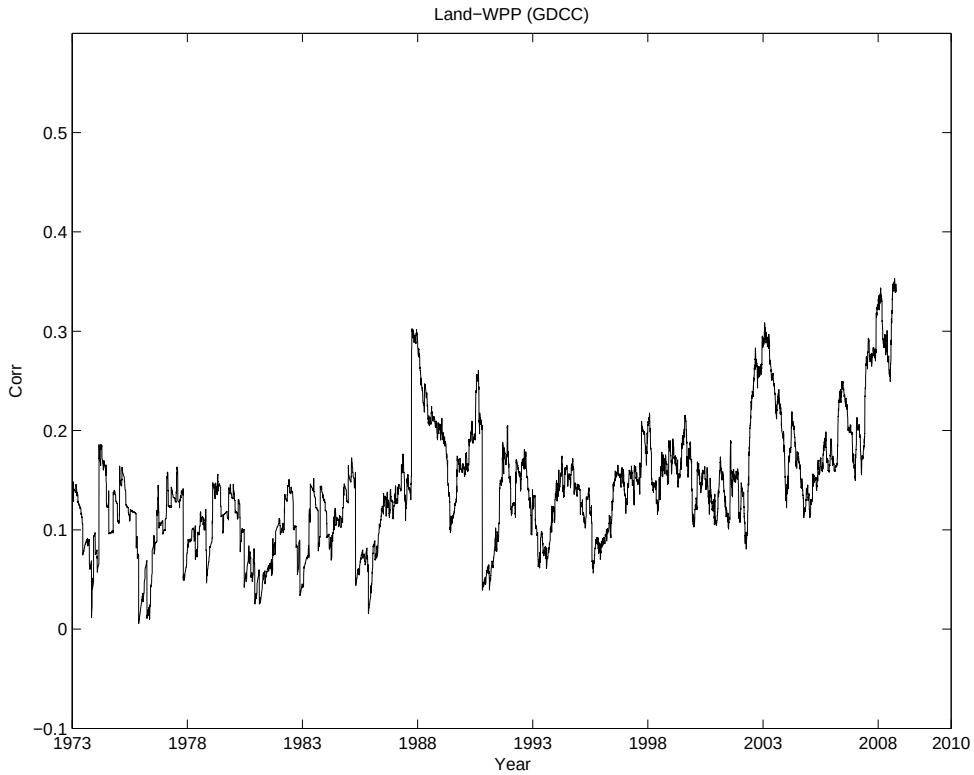

Figure 2: *Estimated conditional correlation for Land Securities and WPP using the GDCC model. The estimators of* α_1 *(Land) and* α_2 *(WPP) are 0.0769 and 0.0792, respectively, giving a product of 0.0061, smaller than the DCC MacGyver estimator 0.0086.*

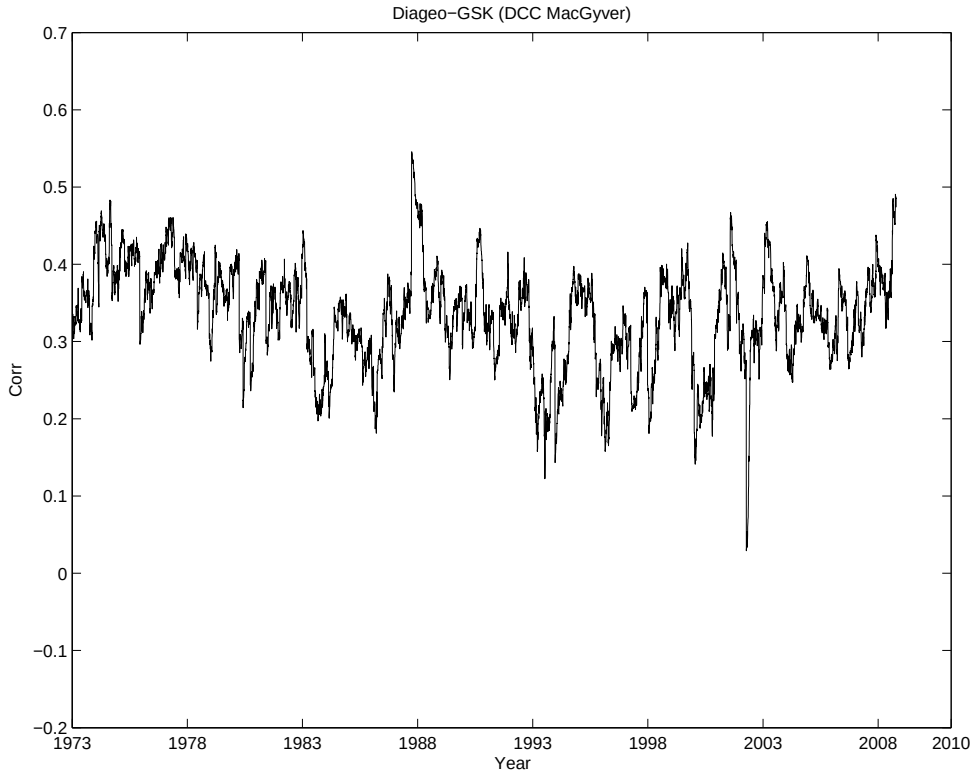

Figure 3: *Estimated conditional correlation for Diageo and GSK using the DCC model.*

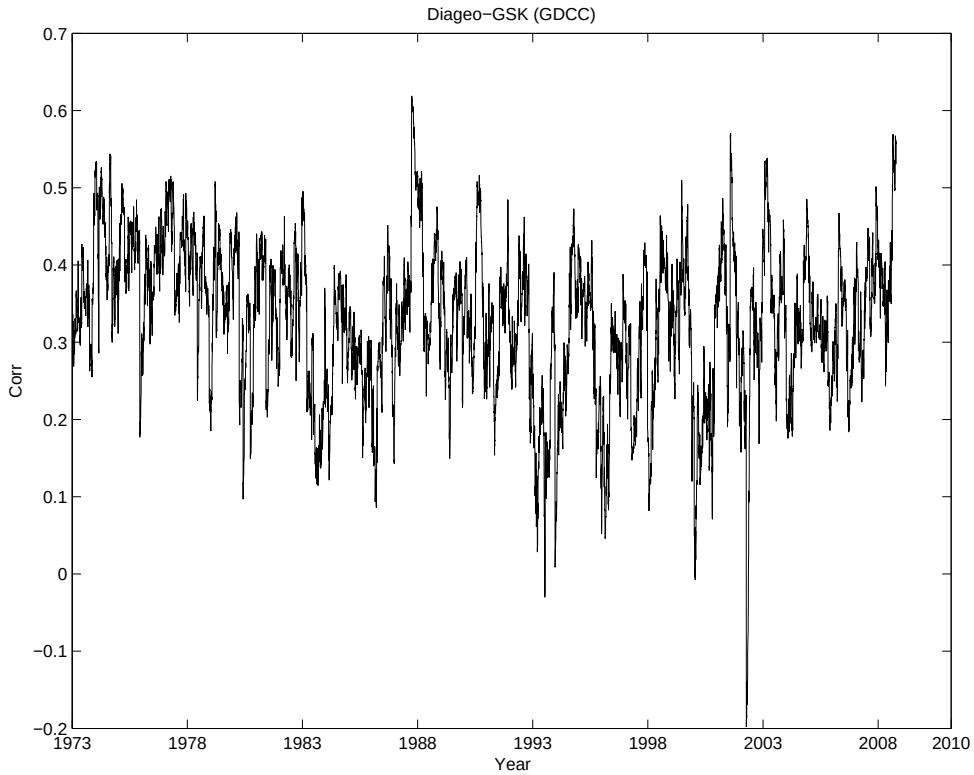

Figure 4: *Estimated conditional correlation for Diageo and GSK using the GDCC model. The estimators of* α_1 *(Diageo) and* α_2 *(GSK) are 0.1331 and 0.1196, respectively, giving a product of 0.0159, larger than the DCC MacGyver estimator 0.0086.*

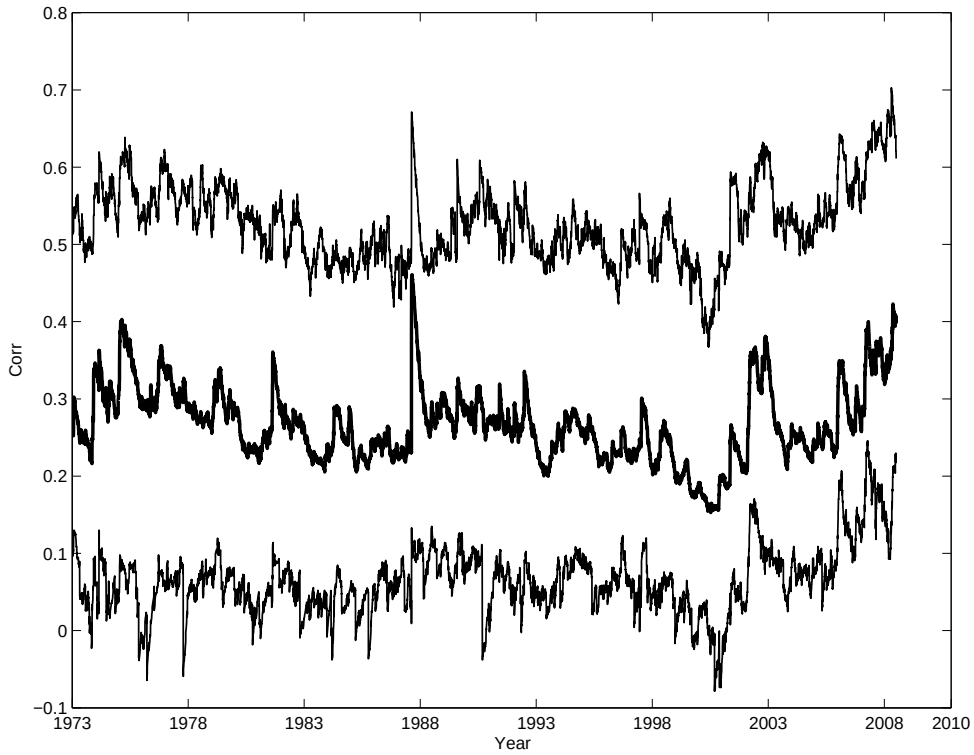

Figure 5: *The pointwise median, 5 and 95 percent quantiles of estimated correlations using the DCC model.*

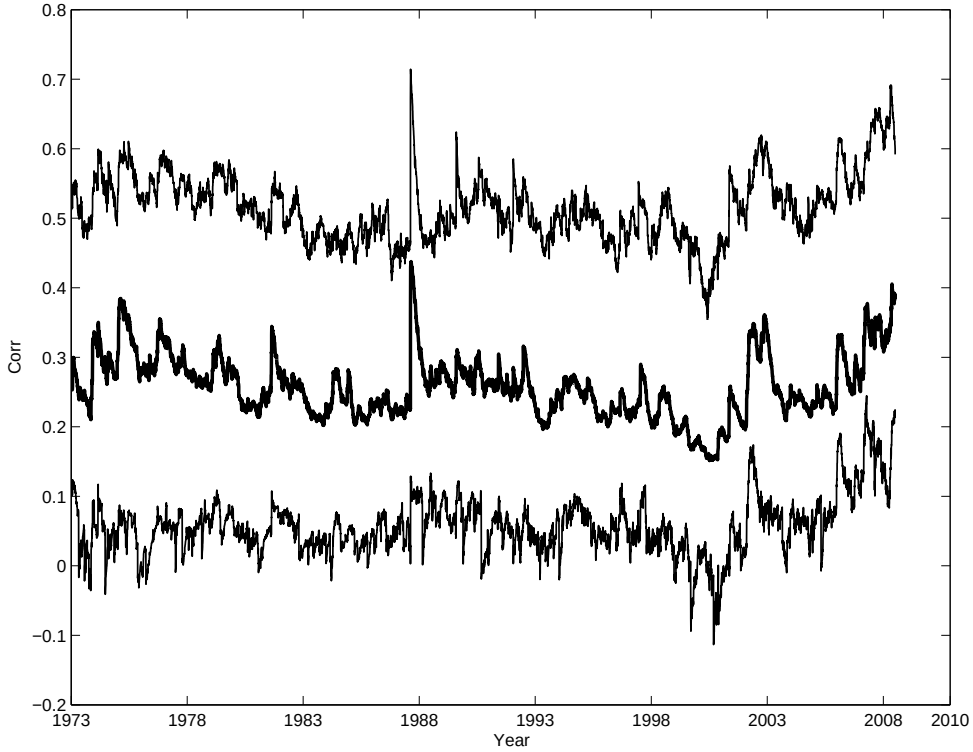

Figure 6: *The pointwise median, 5 and 95 percent quantiles of estimated correlations using the GDCC model.*