Journal of Business & Economic Statistics

ISSN: 0735-0015 (Print) 1537-2707 (Online) Journal homepage: https://amstat.tandfonline.com/loi/ubes20

Dynamic Conditional Correlation

A Simple Class of Multivariate Generalized Autoregressive Conditional Heteroskedasticity Models

Robert Engle

To cite this article: Robert Engle (2002) Dynamic Conditional Correlation, Journal of Business & Economic Statistics, 20:3, 339-350, DOI: 10.1198/073500102288618487

To link to this article: https://doi.org/10.1198/073500102288618487

Published online: 01 Jan 2012.

Dynamic Conditional Correlation: A Simple Class of Multivariate Generalized Autoregressive Conditional Heteroskedasticity Models

Robert ENGLE

Department of Finance, New York University Leonard N. Stern School of Business, New York, NY 10012 and Department of Economics, University of California, San Diego (*rengle@stern.nyu.edu*)

Time varying correlations are often estimated with multivariate generalized autoregressive conditional heteroskedasticity (GARCH) models that are linear in squares and cross products of the data. A new class of multivariate models called dynamic conditional correlation models is proposed. These have the flexibility of univariate GARCH models coupled with parsimonious parametric models for the correlations. They are not linear but can often be estimated very simply with univariate or two-step methods based on the likelihood function. It is shown that they perform well in a variety of situations and provide sensible empirical results.

KEY WORDS: ARCH; Correlation; GARCH; Multivariate GARCH.

1. INTRODUCTION

Correlations are critical inputs for many of the common tasks of financial management. Hedges require estimates of the correlation between the returns of the assets in the hedge. If the correlations and volatilities are changing, then the hedge ratio should be adjusted to account for the most recent information. Similarly, structured products such as rainbow options that are designed with more than one underlying asset have prices that are sensitive to the correlation between the underlying returns. A forecast of future correlations and volatilities is the basis of any pricing formula.

Asset allocation and risk assessment also rely on correlations; however, in this case a large number of correlations is often required. Construction of an optimal portfolio with a set of constraints requires a forecast of the covariance matrix of the returns. Similarly, the calculation of the standard deviation of today's portfolio requires a covariance matrix of all the assets in the portfolio. These functions entail estimation and forecasting of large covariance matrices, potentially with thousands of assets.

The quest for reliable estimates of correlations between financial variables has been the motivation for countless academic articles and practitioner conferences and much Wall Street research. Simple methods such as rolling historical correlations and exponential smoothing are widely used. More complex methods, such as varieties of multivariate generalized autoregressive conditional heteroskedasticity (GARCH) or stochastic volatility, have been extensively investigated in the econometric literature and are used by a few sophisticated practitioners. To see some interesting applications, examine the work of Bollerslev, Engle, and Wooldridge (1988), Bollerslev (1990), Kroner and Claessens (1991), Engle and Mezrich (1996), Engle, Ng, and Rothschild (1990), Bollerslev, Chou, and Kroner (1992), Bollerslev, Engle, and Nelson (1994), and Ding and Engle (2001). In very few of these articles are more than five assets considered, despite the apparent need for bigger correlation matrices. In most cases, the number of parameters in large models is too big for easy optimization.

In this article, dynamic conditional correlation (DCC) estimators are proposed that have the flexibility of univariate GARCH but not the complexity of conventional multivariate GARCH. These models, which parameterize the conditional correlations directly, are naturally estimated in two steps—a series of univariate GARCH estimates and the correlation estimate. These methods have clear computational advantages over multivariate GARCH models in that the number of parameters to be estimated in the correlation process is independent of the number of series to be correlated. Thus potentially very large correlation matrices can be estimated. In this article, the accuracy of the correlations estimated by a variety of methods is compared in bivariate settings where many methods are feasible. An analysis of the performance of the DCC methods for large covariance matrices was considered by Engle and Sheppard (2001).

Section 2 gives a brief overview of various models for estimating correlations. Section 3 introduces the new method and compares it with some of the cited approaches. Section 4 investigates some statistical properties of the method. Section 5 describes a Monte Carlo experiment and results are presented in Section 6. Section 7 presents empirical results for several pairs of daily time series, and Section 8 concludes.

2. CORRELATION MODELS

The conditional correlation between two random variables r_1 and r_2 that each have mean zero is defined to be

$$\rho_{12,t} = \frac{E_{t-1}(r_{1,t}r_{2,t})}{\sqrt{E_{t-1}(r_{1,t}^2)E_{t-1}(r_{2,t}^2)}}. \tag{1}$$

Journal of Business & Economic Statistics
July 2002, Vol. 20, No. 3
DOI 10.1198/073500102288618487

In this definition, the conditional correlation is based on information known the previous period; multiperiod forecasts of the correlation can be defined in the same way. By the laws of probability, all correlations defined in this way must lie within the interval $[-1, 1]$. The conditional correlation satisfies this constraint for all possible realizations of the past information and for all linear combinations of the variables.

To clarify the relation between conditional correlations and conditional variances, it is convenient to write the returns as the conditional standard deviation times the standardized disturbance:

$$h_{i,t} = E_{t-1}\left(r_{i,t}^2\right), \qquad r_{i,t} = \sqrt{h_{i,t}}\varepsilon_{i,t}, \qquad i = 1, 2; \tag{2}$$

ε is a standardized disturbance that has mean zero and variance one for each series. Substituting into (4) gives

$$\rho_{12,t} = \frac{E_{t-1}(\varepsilon_{1,t}\varepsilon_{2,t})}{\sqrt{E_{t-1}(\varepsilon_{1,t}^2)E_{t-1}(\varepsilon_{2,t}^2)}} = E_{t-1}(\varepsilon_{1,t}\varepsilon_{2,t}). \tag{3}$$

Thus, the conditional correlation is also the conditional covariance between the standardized disturbances.

Many estimators have been proposed for conditional correlations. The ever-popular rolling correlation estimator is defined for returns with a zero mean as

$$\hat{\rho}_{12,t} = \frac{\sum_{s=t-n-1}^{t-1} r_{1,s}r_{2,s}}{\sqrt{\left(\sum_{s=t-n-1}^{t-1} r_{1,s}^2\right)\left(\sum_{s=t-n-1}^{t-1} r_{2,s}^2\right)}}. \tag{4}$$

Substituting from (4) it is clear that this is an attractive estimator only in very special circumstances. In particular, it gives equal weight to all observations less than n periods in the past and zero weight on older observations. The estimator will always lie in the $[-1, 1]$ interval, but it is unclear under what assumptions it consistently estimates the conditional correlations. A version of this estimator with a 100-day window, called MA100, will be compared with other correlation estimators.

The exponential smoother used by RiskMetrics uses declining weights based on a parameter λ, which emphasizes current data but has no fixed termination point in the past where data becomes uninformative.

$$\hat{\rho}_{12,t} = \frac{\sum_{s=1}^{t-1} \lambda^{t-j-1} r_{1,s}r_{2,s}}{\sqrt{\left(\sum_{s=1}^{t-1} \lambda^{t-s-1} r_{1,s}^2\right)\left(\sum_{s=1}^{t-1} \lambda^{t-s-1} r_{2,s}^2\right)}}. \tag{5}$$

It also will surely lie in $[-1, 1]$; however, there is no guidance from the data for how to choose λ. In a multivariate context, the same λ must be used for all assets to ensure a positive definite correlation matrix. RiskMetrics uses the value of .94 for λ for all assets. In the comparison employed in this article, this estimator is called EX .06.

Defining the conditional covariance matrix of returns as

$$E_{t-1}(r_t r_t') \equiv H_t, \tag{6}$$

these estimators can be expressed in matrix notation respectively as

$$H_t = \frac{1}{n}\sum_{j=1}^{n}(r_{t-j}r_{t-j}') \text{ and } H_t = \lambda(r_{t-1}r_{t-1}') + (1-\lambda)H_{t-1}. \tag{7}$$

An alternative simple approach to estimating multivariate models is the Orthogonal GARCH method or principle component GARCH method. This was advocated by Alexander (1998, 2001). The procedure is simply to construct unconditionally uncorrelated linear combinations of the series r. Then univariate GARCH models are estimated for some or all of these, and the full covariance matrix is constructed by assuming the conditional correlations are all zero. More precisely, find A such that $y_t = Ar_t$, $E(y_t y_t') \equiv V$ is diagonal. Univariate GARCH models are estimated for the elements of y and combined into the diagonal matrix V_t. Making the additional strong assumption that $E_{t-1}(y_t y_t') = V_t$, then

$$H_t = A^{'-1}V_t A^{-1}. \tag{8}$$

In the bivariate case, the matrix A can be chosen to be triangular and estimated by least squares where r_1 is one component and the residuals from a regression of r_1 on r_2 are the second. In this simple situation, a slightly better approach is to run this regression as a GARCH regression, thereby obtaining residuals that are orthogonal in a generalized least squares (GLS) metric.

Multivariate GARCH models are natural generalizations of this problem. Many specifications have been considered; however, most have been formulated so that the covariances and variances are linear functions of the squares and cross products of the data. The most general expression of this type is called the vec model and was described by Engle and Kroner (1995). The vec model parameterizes the vector of all covariances and variances expressed as $\text{vec}(H_t)$. In the first-order case this is given by

$$\text{vec}(H_t) = \text{vec}(\Omega) + A\,\text{vec}(r_{t-1}r_{t-1}') + B\,\text{vec}(H_{t-1}), \tag{9}$$

where A and B are $n^2 \times n^2$ matrices with much structure following from the symmetry of H. Without further restrictions, this model will not guarantee positive definiteness of the matrix H.

Useful restrictions are derived from the BEKK representation, introduced by Engle and Kroner (1995), which, in the first-order case, can be written as

$$H_t = \Omega + A(r_{t-1}r_{t-1}')A' + BH_{t-1}B'. \tag{10}$$

Various special cases have been discussed in the literature, starting from models where the A and B matrices are simply a scalar or diagonal rather than a whole matrix and continuing to very complex, highly parameterized models that still ensure positive definiteness. See, for example, the work of Engle and Kroner (1995), Bollerslev et al. (1994), Engle and Mezrich (1996), Kroner and Ng (1998), and Engle and Ding (2001). In this study the scalar BEKK and the diagonal BEKK are estimated.

As discussed by Engle and Mezrich (1996), these models can be estimated subject to the variance targeting constraint by which the long run variance covariance matrix is the sample covariance matrix. This constraint differs from the maximum likelihood estimator (MLE) only in finite samples but reduces the number of parameters and often gives improved performance. In the general vec model of Equation (9), this can be expressed as

$$\mathrm{vec}(\Omega)=(I-A-B)\mathrm{vec}(S), \qquad \text{where } S=\frac{1}{T}\sum_t (r_t r_t'). \quad (11)$$

This expression simplifies in the scalar and diagonal BEKK cases. For example, for the scalar BEKK the intercept is simply

$$\Omega=(1-\alpha-\beta)S. \quad (12)$$

3. DCCs

This article introduces a new class of multivariate GARCH estimators that can best be viewed as a generalization of the Bollerslev (1990) constant conditional correlation (CCC) estimator. In Bollerslev's model,

$$H_t=D_t R D_t, \qquad \text{where } D_t=\mathrm{diag}\left\{\sqrt{h_{i,t}}\right\}, \quad (13)$$

where R is a correlation matrix containing the conditional correlations, as can directly be seen from rewriting this equation as

$$E_{t-1}(\varepsilon_t\varepsilon_t')=D_t^{-1}H_t D_t^{-1}=R \quad \text{since } \varepsilon_t=D_t^{-1}r_t. \quad (14)$$

The expressions for h are typically thought of as univariate GARCH models; however, these models could certainly include functions of the other variables in the system as predetermined variables or exogenous variables. A simple estimate of R is the unconditional correlation matrix of the standardized residuals.

This article proposes the DCC estimator. The dynamic correlation model differs only in allowing R to be time varying:

$$H_t=D_t R_t D_t. \quad (15)$$

Parameterizations of R have the same requirements as H, except that the conditional variances must be unity. The matrix R_t remains the correlation matrix.

Kroner and Ng (1998) proposed an alternative generalization that lacks the computational advantages discussed here. They proposed a covariance matrix that is a matrix weighted average of the Bollerslev CCC model and a diagonal BEKK, both of which are positive definite.

Probably the simplest specification for the correlation matrix is the exponential smoother, which can be expressed as

$$\rho_{i,j,t}=\frac{\sum_{s=1}^{t-1}\lambda^s\varepsilon_{i,t-s}\varepsilon_{j,t-s}}{\sqrt{\left(\sum_{s=1}^{t-1}\lambda^s\varepsilon_{i,t-s}^2\right)\left(\sum_{s=1}^{t-1}\lambda^s\varepsilon_{j,t-s}^2\right)}}=[R_t]_{i,j}, \quad (16)$$

a geometrically weighted average of standardized residuals. Clearly these equations will produce a correlation matrix at each point in time. A simple way to construct this correlation is through exponential smoothing. In this case the process followed by the q's will be integrated,

$$q_{i,j,t}=(1-\lambda)(\varepsilon_{i,t-1}\varepsilon_{j,t-1})+\lambda(q_{i,j,t-1}),$$
$$\rho_{i,j,t}=\frac{q_{i,j,t}}{\sqrt{q_{ii,t}q_{jj,t}}}. \quad (17)$$

A natural alternative is suggested by the GARCH(1, 1) model:

$$q_{i,j,t}=\bar{\rho}_{i,j}+\alpha(\varepsilon_{i,t-1}\varepsilon_{j,t-1}-\bar{\rho}_{i,j})+\beta(q_{i,j,t-1}-\bar{\rho}_{i,j}) \quad (18)$$

where $\bar{\rho}_{i,j}$ is the unconditional correlation between $\varepsilon_{i,t}$ and $\varepsilon_{j,t}$. Rewriting gives

$$q_{i,j,t}=\bar{\rho}_{i,j}\left(\frac{1-\alpha-\beta}{1-\beta}\right)+\alpha\sum_{s=1}^{\infty}\beta^{s-1}\varepsilon_{i,t-s}\varepsilon_{j,t-s}. \quad (19)$$

The average of $q_{i,j,t}$ will be $\bar{\rho}_{i,j}$, and the average variance will be 1.

$$\bar{q}_{i,j}\cong\bar{\rho}_{i,j}. \quad (20)$$

The correlation estimator

$$\rho_{i,j,t}=\frac{q_{i,j,t}}{\sqrt{q_{i,i,t}q_{j,j,t}}} \quad (21)$$

will be positive definite as the covariance matrix, Q_t with typical element $q_{i,j,t}$, is a weighted average of a positive definite and a positive semidefinite matrix. The unconditional expectation of the numerator of (21) is $\bar{\rho}_{i,j}$ and each term in the denominator has expected value 1. This model is mean reverting as long as $\alpha+\beta<1$, and when the sum is equal to 1 it is just the model in (17). Matrix versions of these estimators can be written as

$$Q_t=(1-\lambda)(\varepsilon_{t-1}\varepsilon_{t-1}')+\lambda Q_{t-1} \quad (22)$$

and

$$Q_t=S(1-\alpha-\beta)+\alpha(\varepsilon_{t-1}\varepsilon_{t-1}')+\beta Q_{t-1}, \quad (23)$$

where S is the unconditional correlation matrix of the epsilons.

Clearly more complex positive definite multivariate GARCH models could be used for the correlation parameterization as long as the unconditional moments are set to the sample correlation matrix. For example, the MARCH family of Ding and Engle (2001) can be expressed in first-order form as

$$Q_t=S\circ(\iota\iota'-A-B)+A\circ\varepsilon_{t-1}\varepsilon_{t-1}'+B\circ Q_{t-1}, \quad (24)$$

where ι is a vector of ones and $\circ$ is the Hadamard product of two identically sized matrices, which is computed simply by element-by-element multiplication. They show that if A, B, and $(\iota\iota'-A-B)$ are positive semidefinite, then Q will be positive semidefinite. If any one of the matrices is positive definite, then Q will also be. This family includes both earlier models as well as many generalizations.

4. ESTIMATION

The DCC model can be formulated as the following statistical specification:

$$\begin{aligned}
r_t|\Im_{t-1} &\sim N(0, D_t R_t D_t),\\
D_t^2 &= \operatorname{diag}\{\omega_i\} + \operatorname{diag}\{\kappa_i\} \circ r_{t-1} r_{t-1}' + \operatorname{diag}\{\lambda_i\} \circ D_{t-1}^2,\\
\varepsilon_t &= D_t^{-1} r_t,\\
Q_t &= S \circ (\iota\iota' - A - B) + A \circ \varepsilon_{t-1}\varepsilon_{t-1}' + B \circ Q_{t-1},\\
R_t &= \operatorname{diag}\{Q_t\}^{-1} Q_t \operatorname{diag}\{Q_t\}^{-1}.
\end{aligned} \tag{25}$$

The assumption of normality in the first equation gives rise to a likelihood function. Without this assumption, the estimator will still have the Quasi-Maximum Likelihood (QML) interpretation. The second equation simply expresses the assumption that each asset follows a univariate GARCH process. Nothing would change if this were generalized.

The log likelihood for this estimator can be expressed as

$$\begin{aligned}
r_t|\Im_{t-1} &\sim N(0, H_t),\\
L &= -\frac{1}{2}\sum_{t=1}^{T}\big(n\log(2\pi) + \log|H_t| + r_t' H_t^{-1} r_t\big)\\
&= -\frac{1}{2}\sum_{t=1}^{T}\big(n\log(2\pi) + \log|D_t R_t D_t|\\
&\qquad + r_t' D_t^{-1} R_t^{-1} D_t^{-1} r_t\big)\\
&= -\frac{1}{2}\sum_{t=1}^{T}\big(n\log(2\pi) + 2\log|D_t|\\
&\qquad + \log|R_t| + \varepsilon_t' R_t^{-1}\varepsilon_t\big)\\
&= -\frac{1}{2}\sum_{t=1}^{T}\big(n\log(2\pi) + 2\log|D_t| + r_t' D_t^{-1} D_t^{-1} r_t\\
&\qquad - \varepsilon_t'\varepsilon_t + \log|R_t| + \varepsilon_t' R_t^{-1}\varepsilon_t\big),
\end{aligned} \tag{26}$$

which can simply be maximized over the parameters of the model. However, one of the objectives of this formulation is to allow the model to be estimated more easily even when the covariance matrix is very large. In the next few paragraphs several estimation methods are presented, giving simple consistent but inefficient estimates of the parameters of the model. Sufficient conditions are given for the consistency and asymptotic normality of these estimators following Newey and McFadden (1994). Let the parameters in D be denoted θ and the additional parameters in R be denoted ϕ. The log-likelihood can be written as the sum of a volatility part and a correlation part:

$$L(\theta, \phi) = L_V(\theta) + L_C(\theta, \phi). \tag{27}$$

The volatility term is

$$L_V(\theta) = -\frac{1}{2}\sum_t\big(n\log(2\pi) + \log|D_t|^2 + r_t' D_t^{-2} r_t\big), \tag{28}$$

and the correlation component is

$$L_C(\theta, \phi) = -\frac{1}{2}\sum_t\big(\log|R_t| + \varepsilon_t' R_t^{-1}\varepsilon_t - \varepsilon_t'\varepsilon_t\big). \tag{29}$$

The volatility part of the likelihood is apparently the sum of individual GARCH likelihoods

$$L_V(\theta) = -\frac{1}{2}\sum_t\sum_{i=1}^{n}\left(\log(2\pi) + \log(h_{i,t}) + \frac{r_{i,t}^2}{h_{i,t}}\right), \tag{30}$$

which is jointly maximized by separately maximizing each term.

The second part of the likelihood is used to estimate the correlation parameters. Because the squared residuals are not dependent on these parameters, they do not enter the first-order conditions and can be ignored. The resulting estimator is called DCC LL MR if the mean reverting formula (18) is used and DCC LL INT with the integrated model in (17).

The two-step approach to maximizing the likelihood is to find

$$\hat{\theta} = \arg\max\{L_V(\theta)\} \tag{31}$$

and then take this value as given in the second stage:

$$\max_{\phi}\{L_C(\hat{\theta}, \phi)\}. \tag{32}$$

Under reasonable regularity conditions, consistency of the first step will ensure consistency of the second step. The maximum of the second step will be a function of the first-step parameter estimates, so if the first step is consistent, the second step will be consistent as long as the function is continuous in a neighborhood of the true parameters.

Newey and McFadden (1994), in Theorem 6.1, formulated a two-step Generalized Method of Moments (GMM) problem that can be applied to this model. Consider the moment condition corresponding to the first step as $\nabla_\theta L_V(\theta) = 0$. The moment condition corresponding to the second step is $\nabla_\phi L_C(\hat{\theta}, \phi) = 0$. Under standard regularity conditions, which are given as conditions (i) to (v) in Theorem 3.4 of Newey and McFadden, the parameter estimates will be consistent, and asymptotically normal, with a covariance matrix of familiar form. This matrix is the product of two inverted Hessians around an outer product of scores. In particular, the covariance matrix of the correlation parameters is

$$\begin{aligned}
V(\phi) &= [E(\nabla_{\phi\phi} L_C)]^{-1}\\
&\quad \times E\big(\{\nabla_\phi L_C - E(\nabla_{\phi\theta} L_C)[E(\nabla_{\theta\theta} L_V)]^{-1}\nabla_\theta L_V\}\\
&\quad \times\{\nabla_\phi L_C - E(\nabla_{\phi\theta} L_C)[E(\nabla_{\theta\theta} L_V)]^{-1}\nabla_\theta L_V\}'\big)\\
&\quad \times[E(\nabla_{\phi\phi} L_C)].
\end{aligned} \tag{33}$$

Details of this proof can be found elsewhere (Engle and Sheppard 2001).

Alternative estimation approaches, which are again consistent but inefficient, can easily be devised. Rewrite (18) as

$$\begin{aligned}
e_{i,j,t} &= \bar{\rho}_{i,j}(1 - \alpha - \beta) + (\alpha + \beta)e_{i,j,t-1}\\
&\qquad - \beta(e_{i,j,t-1} - q_{i,j,t-1}) + (e_{i,j,t} - q_{i,j,t}),
\end{aligned} \tag{34}$$

where $e_{i,j,t} = \varepsilon_{i,t}\varepsilon_{j,t}$. This equation is an ARMA(1, 1) because the errors are a Martingale difference by construction. The autoregressive coefficient is slightly bigger if α is a small positive number, which is the empirically relevant case. This equation can therefore be estimated with conventional time series software to recover consistent estimates of the parameters. The drawback to this method is that ARMA with nearly equal roots are numerically unstable and tricky to estimate. A further drawback is that in the multivariate setting, there are many such cross products that can be used for this estimation. The problem is even easier if the model is (17) because then the autoregressive root is assumed to be 1. The model is simply an integrated moving average (IMA) with no intercept,

$$\Delta e_{i,j,t} = -\beta(e_{i,j,t-1} - q_{i,j,t-1}) + (e_{i,j,t} - q_{i,j,t}), \tag{35}$$

which is simply an exponential smoother with parameter $\lambda = \beta$. This estimator is called the DCC IMA.

5. COMPARISON OF ESTIMATORS

In this section, several correlation estimators are compared in a setting where the true correlation structure is known. A bivariate GARCH model is simulated 200 times for 1,000 observations or approximately 4 years of daily data for each correlation process. Alternative correlation estimators are compared in terms of simple goodness-of-fit statistics, multivariate GARCH diagnostic tests, and value-at-risk tests.

The data-generating process consists of two Gaussian GARCH models; one is highly persistent and the other is not.

$$\begin{gathered}
h_{1,t} = .01 + .05r_{1,t-1}^2 + .94h_{1,t-1},\\
h_{2,t} = .5 + .2r_{2,t-1}^2 + .5h_{2,t-1},\\
\begin{pmatrix}\varepsilon_{1,t}\\ \varepsilon_{2,t}\end{pmatrix} \sim N\left[0, \begin{pmatrix}1 & \rho_t\\ \rho_t & 1\end{pmatrix}\right],\\
r_{1,t} = \sqrt{h_{1,t}}\varepsilon_{1,t}, \quad r_{2,t} = \sqrt{h_{2,t}}\varepsilon_{2,t}.
\end{gathered} \tag{36}$$

The correlations follow several processes that are labeled as follows:

1. Constant $\rho_t = .9$
2. Sine $\rho_t = .5 + .4\cos(2\pi t/200)$
3. Fast sine $\rho_t = .5 + .4\cos(2\pi t/20)$
4. Step $\rho_t = .9 - .5(t > 500)$
5. Ramp $\rho_t = \operatorname{mod}(t/200)$

These processes were chosen because they exhibit rapid changes, gradual changes, and periods of constancy. Some of the processes appear mean reverting and others have abrupt changes. Various other experiments are done with different error distributions and different data-generating parameters but the results are quite similar.

Eight different methods are used to estimate the correlations—two multivariate GARCH models, orthogonal GARCH, two integrated DCC models, and one mean reverting DCC plus the exponential smoother from Riskmetrics and the familiar 100-day moving average. The methods and their descriptions are as follows:

1. Scalar BEKK—scalar version of (10) with variance targeting as in (12)
2. Diag BEKK—diagonal version of (10) with variance targeting as in (11)
3. DCC IMA—DCC with integrated moving average estimation as in (35)
4. DCC LL INT—DCC by log-likelihood for integrated process
5. DCC LL MR—DCC by log-likelihood with mean reverting model as in (18)
6. MA100—moving average of 100 days
7. EX .06—exponential smoothing with parameter = .06
8. OGARCH—orthogonal GARCH or principle components GARCH as in (8).

Three performance measures are used. The first is simply the comparison of the estimated correlations with the true correlations by mean absolute error. This is defined as

$$\mathrm{MAE} = \frac{1}{T}\sum|\hat{\rho}_t - \rho_t|, \tag{37}$$

and of course the smallest values are the best. A second measure is a test for autocorrelation of the squared standardized residuals. For a multivariate problem, the standardized residuals are defined as

$$\nu_t = H_t^{-1/2} r_t, \tag{38}$$

which in this bivariate case is implemented with a triangular square root defined as

$$\begin{aligned}
\nu_{1,t} &= r_{1,t}/\sqrt{H_{11,t}},\\
\nu_{2,t} &= r_{2,t}\frac{1}{\sqrt{H_{22,t}(1-\hat{\rho}_t^2)}} - r_{1,t}\frac{\hat{\rho}_t}{\sqrt{H_{11,t}(1-\hat{\rho}_t^2)}}.
\end{aligned} \tag{39}$$

The test is computed as an F test from the regression of $\nu_{1,t}^2$ and $\nu_{2,t}^2$ on five lags of the squares and cross products of the standardized residuals plus an intercept. The number of rejections using a 5% critical value is a measure of the performance of the estimator because the more rejections, the more evidence that the standardized residuals have remaining time varying volatilities. This test obviously can be used for real data.

The third performance measure is an evaluation of the estimator for calculating value at risk. For a portfolio with w invested in the first asset and $(1-w)$ in the second, the value at risk, assuming normality, is

$$\mathrm{VaR}_t = 1.65\sqrt{(w^2H_{11,t} + (1-w)^2H_{22,t} + 2^*w(1-w)\hat{\rho}_t\sqrt{H_{11,t}H_{22,t}})}, \tag{40}$$

and a dichotomous variable called hit should be unpredictable based on the past where hit is defined as

$$\mathrm{hit}_t = I(w^* r_{1,t} + (1-w)^* r_{2,t} < -\mathrm{VaR}_t) - .05. \tag{41}$$

The dynamic quantile test introduced by Engle and Manganelli (2001) is an F test of the hypothesis that all coefficients as well as the intercept are zero in a regression of this variable on its past, on current VaR, and any other variables. In this case, five lags and the current VaR are used. The number of rejections using a 5% critical value is a measure of model performance. The reported results are for an equal weighted portfolio with $w = .5$ and a hedge portfolio with weights $1, -1$.

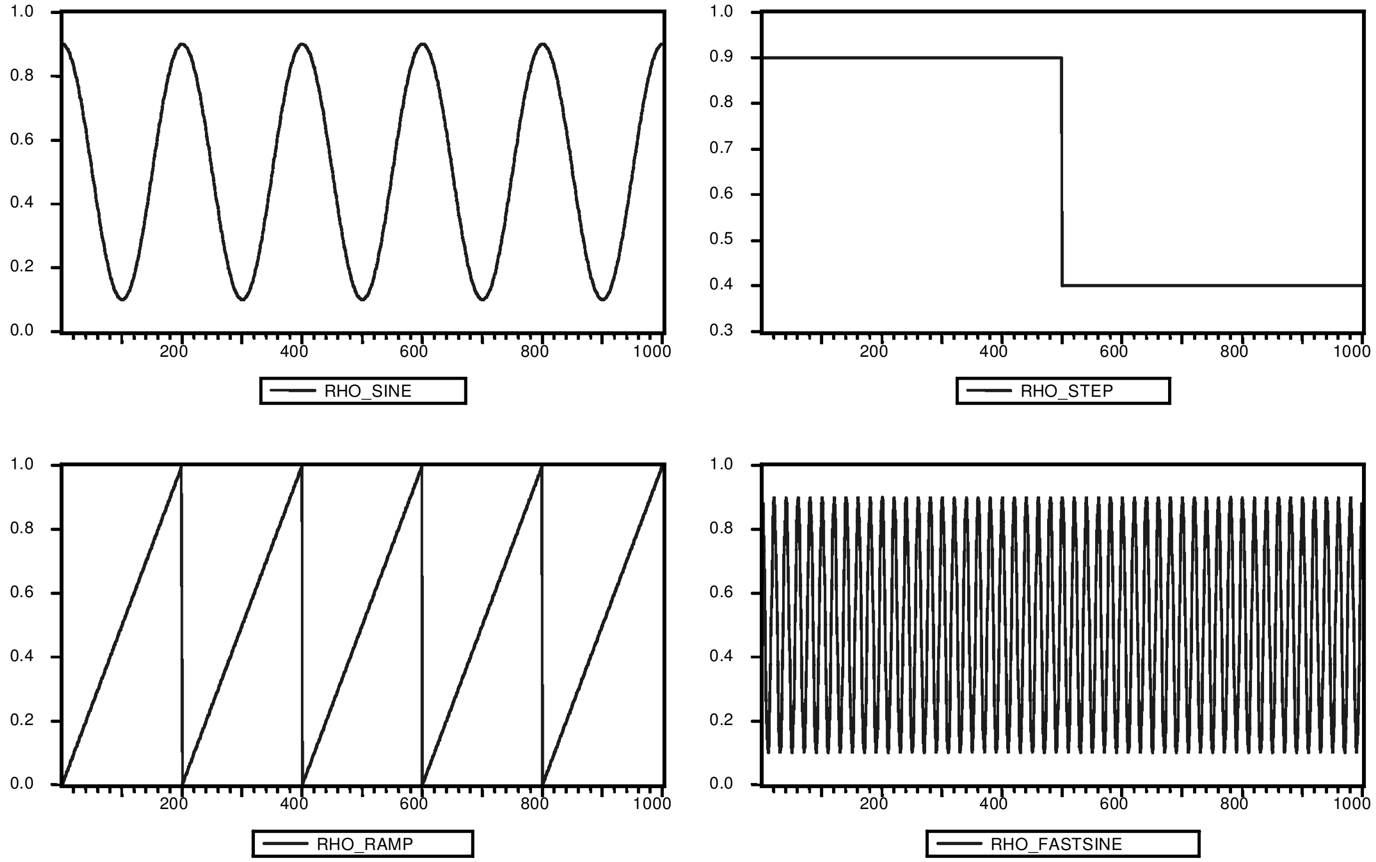

Figure 1. Correlation Experiments.

6. RESULTS

Table 1 presents the results for the mean absolute error (MAE) for the eight estimators for six experiments with 200 replications. In four of the six cases the DCC mean reverting model has the smallest MAE. When these errors are summed over all cases, this model is the best. Very close second- and third-place models are DCC integrated with log-likelihood estimation and scalar BEKK.

In Table 2 the second standardized residual is tested for remaining autocorrelation in its square. This is the more revealing test because it depends on the correlations; the test for the first residual does not. Because all models are misspecified, the rejection rates are typically well above 5%. For three of six cases, the DCC mean reverting model is the best. When summed over all cases it is a clear winner.

The test for autocorrelation in the first squared standardized residual is presented in Table 3. These test statistics are typically close to 5%, reflecting the fact that many of these models are correctly specified and the rejection rate should be the size. Overall the best model appears to be the diagonal BEKK with OGARCH and DCC close behind.

The VaR-based dynamic quantile test is presented in Table 4 for a portfolio that is half invested in each asset and in Table 5 for a long-short portfolio with weights plus and minus one. The number of 5% rejections for many of the models is close to the 5% nominal level despite misspecification of the structure. In five of six cases, the minimum is the integrated DCC log-likelihood; overall, it is also the best method, followed by the mean reverting DCC and the IMA DCC.

The value-at-risk test based on the long-short portfolio finds that the diagonal BEKK is best for three of six, whereas the DCC MR is best for two. Overall, the DCC MR is observed to be the best.

From all these performance measures, the DCC methods are either the best or very near the best method. Choosing among these models, the mean reverting model is the general winner, although the integrated versions are close behind and perform best by some criteria. Generally the log-likelihood estimation method is superior to the IMA estimator for the integrated DCC models.

The confidence with which these conclusions can be drawn can also be investigated. One simple approach is to repeat the experiment with different sets of random numbers. The entire Monte Carlo was repeated two more times. The results are very close with only one change in ranking that favors the DCC LL MR over the DCC LL INT.

7. EMPIRICAL RESULTS

Empirical examples of these correlation estimates are presented for several interesting series. First the correlation between the Dow Jones Industrial Average and the NASDAQ composite is examined for 10 years of daily data ending

Table 1. *Mean Absolute Error of Correlation Estimates*

MODEL	SCAL BEKK	DIAG BEKK	DCC LL MR	DCC LL INT	DCC IMA	EX .06	MA 100	O-GARCH
FAST SINE	.2292	.2307	**.2260**	.2555	.2581	.2737	.2599	.2474
SINE	.1422	.1451	**.1381**	.1455	.1678	.1541	.3038	.2245
STEP	.0859	.0931	.0709	.0686	.0672	.0810	**.0652**	.1566
RAMP	.1610	.1631	**.1546**	.1596	.1768	.1601	.2828	.2277
CONST	.0273	.0276	.0070	**.0067**	.0105	.0276	.0185	.0449
T(4) SINE	.1595	.1668	**.1478**	.1583	.2199	.1599	.3016	.2423

Table 2. *Fraction of 5% Tests Finding Autocorrelation in Squared Standardized Second Residual*

MODEL	SCAL BEKK	DIAG BEKK	DCC LL MR	DCC LL INT	DCC IMA	EX .06	MA 100	O-GARCH
FAST SINE	.3100	**.0950**	.1300	.3700	.3700	.7250	.9900	.1100
SINE	.5930	.2677	**.1400**	.1850	.3350	.7600	1.0000	.2650
STEP	.8995	.6700	**.2778**	.3250	.6650	.8550	.9950	.7600
RAMP	.5300	.2600	.2400	.5450	.7500	.7300	1.0000	**.2200**
CONST	.9800	.3600	**.0788**	.0900	.1250	.9700	.9950	.9350
T(4) SINE	.2800	.1900	.2050	.2400	.1650	.3300	.8950	**.1600**

Table 3. *Fraction of 5% Tests Finding Autocorrelation in Squared Standardized First Residual*

MODEL	SCAL BEKK	DIAG BEKK	DCC LL MR	DCC LL INT	DCC IMA	EX .06	MA 100	O-GARCH
FAST SINE	.2250	**.0450**	.0600	.0600	.0650	.0750	.6550	.0600
SINE	.0804	.0657	.0400	**.0300**	.0600	.0400	.6250	.0400
STEP	.0302	.0400	.0505	.0500	.0450	.0300	.6500	**.0250**
RAMP	.0550	.0500	.0500	.0600	.0600	.0650	.7500	**.0400**
CONST	.0200	.0250	.0242	.0250	.0250	.0400	.6350	**.0150**
T(4) SINE	.0950	**.0550**	.0850	.0800	.0950	.0850	.4900	.1050

Table 4. *Fraction of 5% Dynamic Quantile Tests Rejecting Value at Risk, Equal Weighted*

MODEL	SCAL BEKK	DIAG BEKK	DCC LL MR	DCC LL INT	DCC IMA	EX .06	MA 100	O-GARCH
FAST SINE	**.0300**	.0450	.0350	**.0300**	.0450	.2450	.4350	.1200
SINE	.0452	.0556	**.0250**	.0350	.0350	.1600	.4100	.3200
STEP	.1759	.1650	.0758	**.0650**	.0800	.2450	.3950	.6100
RAMP	.0750	.0650	.0500	**.0400**	.0450	.2000	.5300	.2150
CONST	.0600	.0800	.0667	**.0550**	**.0550**	.2600	.4800	.2650
T(4) SINE	.1250	.1150	.1000	**.0850**	.1200	.1950	.3950	.2050

Table 5. *Fraction of 5% Dynamic Quantile Tests Rejecting Value at Risk, Long-Short*

MODEL	SCAL BEKK	DIAG BEKK	DCC LL MR	DCC LL INT	DCC IMA	EX .06	MA 100	O-GARCH
FAST SINE	.1000	.0950	.0900	.2550	.2550	.5800	.4650	**.0850**
SINE	.0553	**.0303**	.0450	.0900	.1850	.2150	.9450	.0650
STEP	.1055	.0850	**.0404**	.0600	.1150	.1700	.4600	.1250
RAMP	.0800	**.0650**	.0800	.1750	.2500	.3050	.9000	.1000
CONST	.1850	.0900	**.0424**	.0550	.0550	.3850	.5500	.1050
T(4) SINE	.1150	**.0900**	.1350	.1300	.2000	.2150	.8050	.1050

in March 2000. Then daily correlations between stocks and bonds, a central feature of asset allocation models, are considered. Finally, the daily correlation between returns on several currencies around major historical events including the launch of the Euro is examined. Each of these datasets has been used to estimate all the models described previously. The DCC parameter estimates for the integrated and mean reverting models are exhibited with consistent standard errors from (33) in Appendix A. In that Appendix, the statistic referred to as likelihood ratio is the difference between the log-likelihood of the second-stage estimates using the integrated model and using the mean reverting model. Because these are not jointly maximized likelihoods, the distribution could be different from its conventional chi-squared asymptotic limit. Furthermore, nonnormality of the returns would also affect this limiting distribution.

7.1 Dow Jones and NASDAQ

The dramatic rise in the NASDAQ over the last part of the 1990s perplexed many portfolio managers and delighted the new internet start-ups and day traders. A plot of the GARCH volatilities of these series in Figure 2 reveals that the NASDAQ has always been more volatile than the Dow but that this gap widens at the end of the sample.

The correlation between the Dow and NASDAQ was estimated with the DCC integrated method, using the volatilities in Figure 2. The results, shown in Figure 3, are quite interesting.

Whereas for most of the decade the correlations were between .6 and .9, there were two notable drops. In 1993 the correlations averaged .5 and dropped below .4, and in March 2000 they again dropped below .4. The episode in 2000 is sometimes attributed to sector rotation between new economy stocks and "brick and mortar" stocks. The drop at the end of the sample period is more pronounced for some estimators than for others. Looking at just the last year in Figure 4, it can be seen that only the orthogonal GARCH correlations fail to decline and that the BEKK correlations are most volatile.

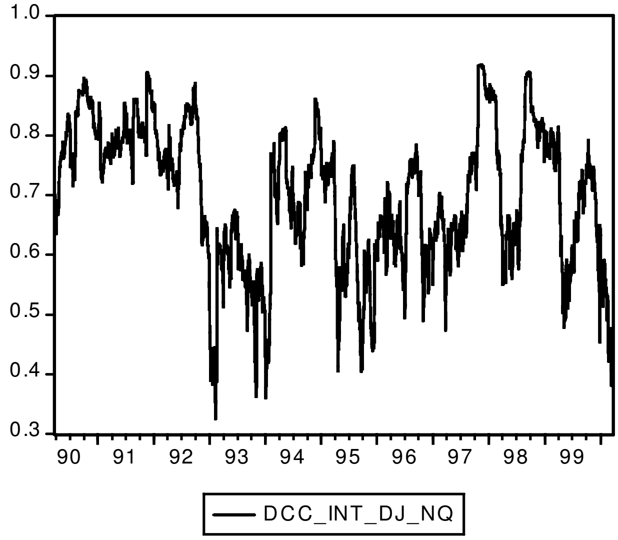

Figure 3. Ten Years of Dow Jones–NASDAQ Correlations.

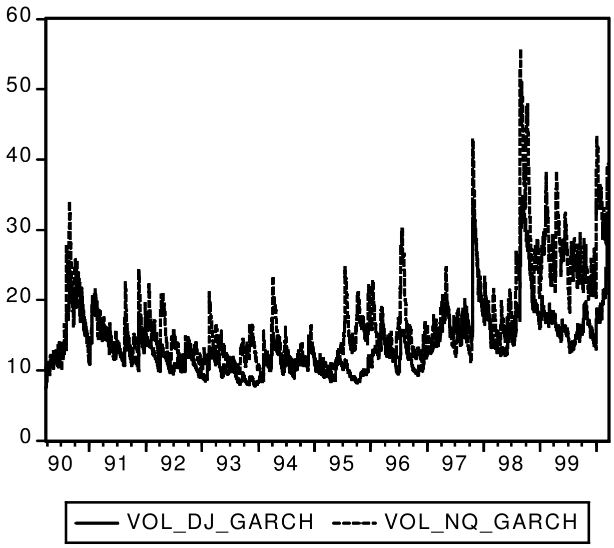

Figure 2. Ten Years of Volatilities.

7.2 Stocks and Bonds

The second empirical example is the correlation between domestic stocks and bonds. Taking bond returns to be minus the change in the 30-year benchmark yield to maturity, the correlation between bond yields and the Dow and NASDAQ are shown in Figure 5 for the integrated DCC for the last 10 years. The correlations are generally positive in the range of .4 except for the summer of 1998, when they become highly negative, and the end of the sample, when they are about 0. Although it is widely reported in the press that the NASDAQ does not seem to be sensitive to interest rates, the data suggests that this is true only for some limited periods, including the first quarter of 2000, and that this is also true for the Dow. Throughout the decade it appears that the Dow is slightly more correlated with bond prices than is the NASDAQ.

7.3 Exchange Rates

Currency correlations show dramatic evidence of nonstationarity. That is, there are very pronounced apparent structural changes in the correlation process. In Figure 6, the breakdown of the correlations between the Deutschmark and the pound and lira in August of 1992 is very apparent. For the pound this was a return to a more normal correlation, while for the lira it was a dramatic uncoupling.

Figure 7 shows currency correlations leading up to the launch of the Euro in January 1999. The lira has lower correlations with the Franc and Deutschmark from 93 to 96, but then they gradually approach one. As the Euro is launched, the estimated correlation moves essentially to 1. In the last year it drops below .95 only once for the Franc and lira and not at all for the other two pairs.

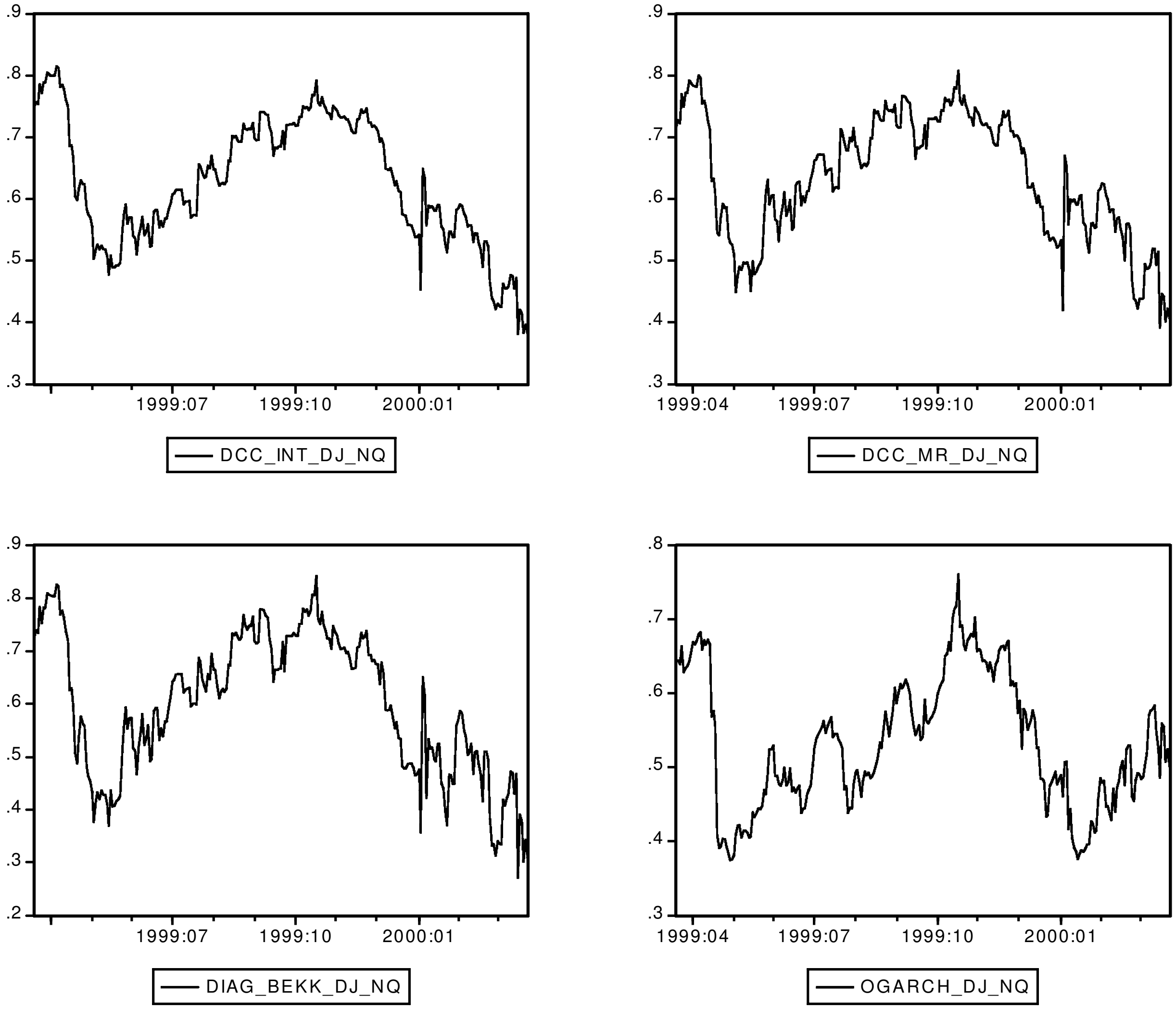

Figure 4. Correlations from March 1999 to March 2000.

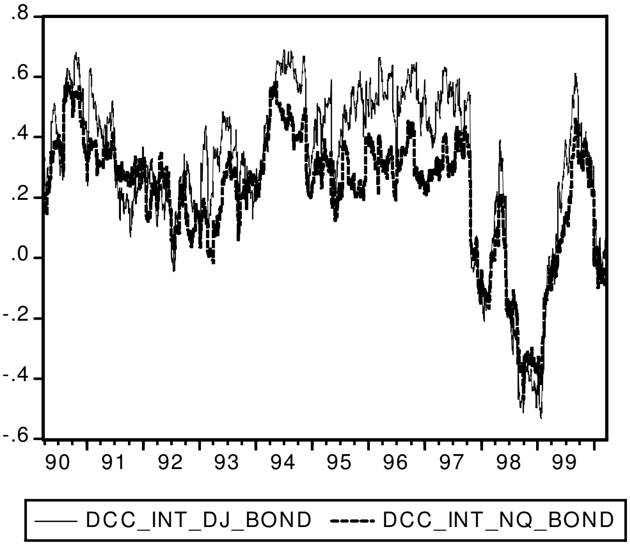

Figure 5. Ten Years of Bond Correlations.

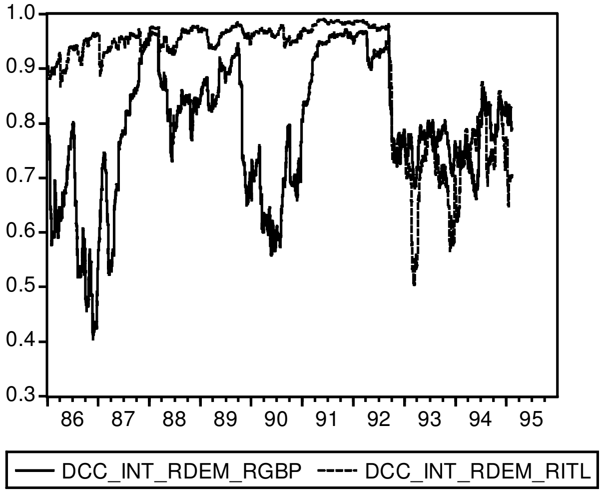

Figure 6. Ten Years of Currency Correlations.

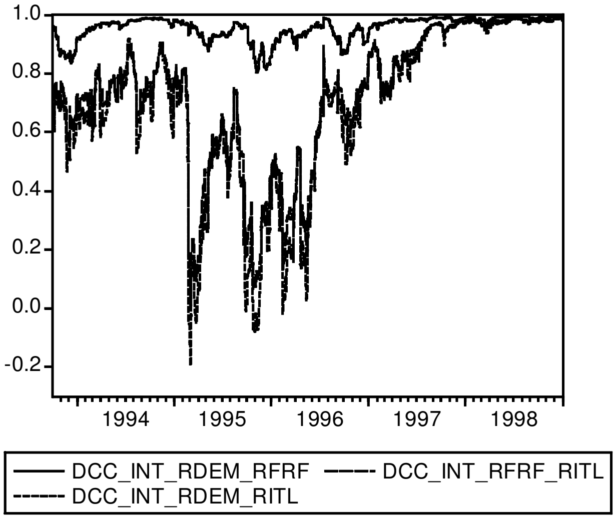

Figure 7. Currency Correlations.

From the results in Appendix A, it is seen that this is the only dataset for which the integrated DCC cannot be rejected against the mean reverting DCC. The nonstationarity in these correlations presumably is responsible. It is somewhat surprising that a similar result is not found for the prior currency pairs.

7.4 Testing the Empirical Models

For each of these datasets, the same set of tests that were used in the Monte Carlo experiment can be constructed. In this case of course, the mean absolute errors cannot be observed, but the tests for residual ARCH can be computed and the tests for value at risk can be computed. In the latter case, the results are subject to various interpretations because the assumption of normality is a potential source of rejection. In each case the number of observations is larger than in the Monte Carlo experiment, ranging from 1,400 to 2,600.

The p-statistics for each of four tests are given in Appendix B. The tests are the tests for residual autocorrelation in squares and for accuracy of value at risk for two portfolios. The two portfolios are an equally weighted portfolio and a long-short portfolio. They presumably are sensitive to rather different failures of correlation estimates. From the four tables, it is immediately clear that most of the models are misspecified for most of the data sets. If a 5% test is done for all the datasets on each of the criteria, then the expected number of rejections for each model would be just over 1 of 28 possibilities. Across the models there are from 10 to 21 rejections at the 5% level!

Without exception, the worst performer on all of the tests and datasets is the moving average model with 100 lags. From counting the total number of rejections, the best model is the diagonal BEKK with 10 rejections. The DCC LL MR, scalar BEKK, O_GARCH, and EX .06 all have 12 rejections, and the DCC LL INT has 14. Probably, these differences are not large enough to be convincing.

If a 1% test is used reflecting the larger sample size, then the number of rejections ranges from 7 to 21. Again the MA 100 is the worst but now the EX .06 is the winner. The DCC LL MR, DCC LL INT, and diagonal BEKK are all tied for second with 9 rejections each.

The implications of this comparison are mainly that a bigger and more systematic comparison is required. These results suggest first that real data are more complicated than any of these models. Second, it appears that the DCC models are competitive with the other methods, some of which are difficult to generalize to large systems.

8. CONCLUSIONS

In this article a new family of multivariate GARCH models was proposed that can be simply estimated in two steps from univariate GARCH estimates of each equation. A maximum likelihood estimator was proposed and several different specifications were suggested. The goal for this proposal is to find specifications that potentially can estimate large covariance matrices. In this article, only bivariate systems were estimated to establish the accuracy of this model for simpler structures. However, the procedure was carefully defined and should also work for large systems. A desirable practical feature of the DCC models is that multivariate and univariate volatility forecasts are consistent with each other. When new variables are added to the system, the volatility forecasts of the original assets will be unchanged and correlations may even remain unchanged, depending on how the model is revised.

The main finding is that the bivariate version of this model provides a very good approximation to a variety of time-varying correlation processes. The comparison of DCC with simple multivariate GARCH and several other estimators shows that the DCC is often the most accurate. This is true whether the criterion is mean absolute error, diagnostic tests, or tests based on value at risk calculations.

Empirical examples from typical financial applications are quite encouraging because they reveal important time-varying features that might otherwise be difficult to quantify. Statistical tests on real data indicate that all these models are misspecified but that the DCC models are competitive with the multivariate GARCH specifications and are superior to moving average methods.

ACKNOWLEDGMENTS

This research was supported by NSF grant SBR-9730062 and NBER AP group. The author thanks Kevin Sheppard for research assistance and Pat Burns and John Geweke for insightful comments. Thanks also to seminar participants at New York University; University of California at San Diego; Academica Sinica; Taiwan; CIRANO at Montreal; University of Iowa; Journal of Applied Econometrics Lectures, Cambridge, England; CNRS Aussois; Brown University; Fields Institute University of Toronto; and Riskmetrics.

APPENDIX A

	Mean Reverting Model				*Integrated Model*		
	Asset 1 NQ	*Asset 2 DJ*			*Asset 1 NQ*	*Asset 2 DJ*	
	Parameter	*T-stat*	*Log-likelihood*		*Parameter*	*T-stat*	*Log-likelihood*
alphaDCC	.039029	6.916839405		lambdaDCC	.030255569	4.66248	18062.79651
betaDCC	.941958	92.72739572	18079.5857	LR TEST			33.57836423
	Asset 1 RATE	*Asset 2 DJ*			*Asset 1 RATE*	*Asset 2 DJ*	
	Parameter	*T-stat*	*Log-likelihood*		*Parameter*	*T-stat*	*Log-likelihood*
alphaDCC	.037372	2.745870787		lambdaDCC	.02851073	3.675969	13188.63653
betaDCC	.950269	44.42479805	13197.82499	LR TEST			18.37690833
	Asset 1 NQ	*Asset 2 RATE*			*Asset 1 NQ*	*Asset 2 RATE*	
	Parameter	*T-stat*	*Log-likelihood*		*Parameter*	*T-stat*	*Log-likelihood*
alphaDCC	.029972	2.652315309		lambdaDCC	.019359061	2.127002	12578.06669
betaDCC	.953244	46.61344925	12587.26244	LR TEST			18.39149373
	Asset 1 DM	*Asset 2 ITL*			*Asset 1 DM*	*Asset 2 ITL*	
	Parameter	*T-stat*	*Log-likelihood*		*Parameter*	*T-stat*	*Log-likelihood*
alphaDCC	.0991	3.953696951		lambdaDCC	.052484321	4.243317	20976.5062
betaDCC	.863885	21.32994852	21041.71874	LR TEST			13.4250734
	Asset 1 DM	*Asset 2 GBP*			*Asset 1 DM*	*Asset 2 GBP*	
	Parameter	*T-stat*	*Log-likelihood*		*Parameter*	*T-stat*	*Log-likelihood*
alphaDCC	.03264	1.315852908		lambdaDCC	.024731692	1.932782	19480.21203
betaDCC	.963504	37.57905053	19508.6083	LR TEST			56.79255661
	Asset 1 rdem90	*Asset 2 rfrf90*			*Asset 1 rdem90*	*Asset 2 rfrf90*	
	Parameter	*T-stat*	*Log-likelihood*		*Parameter*	*T-stat*	*Log-likelihood*
alphaDCC	.059413	4.154987386		lambdaDCC	.047704833	2.880988	12416.84873
betaDCC	.934458	59.19216459	12426.89065	LR TEST			20.08382828
	Asset 1 rdem90	*Asset 2 ritl90*			*Asset 1 rdem90*	*Asset 2 ritl90*	
	Parameter	*T-stat*	*Log-likelihood*		*Parameter*	*T-stat*	*Log-likelihood*
alphaDCC	.056675	3.091462338		lambdaDCC	.053523717	2.971859	11442.50983
betaDCC	.943001	5.77614662	11443.23811	LR TEST			1.456541924

NOTE: Empirical results for bivariate DCC models. T-statistics are robust and consistent using (33). The estimates in the left column are DCC LL MR and the estimates in the right column are DCC LL INT. The LR statistic is twice the difference between the log likelihoods of the second stage. The data are all logarithmic differences: NQ=Nasdaq composite, DJ=Dow Jones Industrial Average, RATE=return on 30 year US Treasury, all daily from 3/23/90 to 3/22/00. Furthermore: DM=Deutschmarks per dollar, ITL=Italian Lira per dollar, GBP=British pounds per dollar, all from 1/1/85 to 2/13/95. Finally rdem90=Deutschmarks per dollar, rfrf90=French Francs per dollar, and ritl90=Italian Lira per dollar, all from 1/1/93 to 1/15/99.

APPENDIX B

P-Statistics From Tests of Empirical Models

ARCH in Squared RESID1

MODEL	*SCAL BEKK*	*DIAG BEKK*	*DCC LL MR*	*DCC LL INT*	*EX .06*	*MA100*	*O-GARCH*
NASD&DJ	.0047	.0281	.3541	.3498	.3752	.0000	.2748
DJ&RATE	.0000	.0002	.0003	.0020	.0167	.0000	.0001
NQ&RATE	.0000	.0044	.0100	.0224	.0053	.0000	.0090
DM&ITL	.4071	.3593	.2397	.1204	.5503	.0000	.4534
DM&GBP	.4437	.4303	.4601	.3872	.4141	.0000	.4213
FF&DM90	.2364	.2196	.1219	.1980	.3637	.0000	.0225
DM&IT90	.1188	.3579	.0075	.0001	.0119	.0000	.0010

ARCH in Squared RESID2

MODEL	*SCAL BEKK*	*DIAG BEKK*	*DCC LL MR*	*DCC LL INT*	*EX .06*	*MA100*	*O-GARCH*
NASD&DJ	.0723	.0656	.0315	.0276	.0604	.0000	.0201
DJ&RATE	.7090	.7975	.8251	.6197	.8224	.0007	.1570
NQ&RATE	.0052	.0093	.0075	.0053	.0023	.0000	.1249
DM&ITL	.0001	.0000	.0000	.0000	.0000	.0000	.0000
DM&GBP	.0000	.0000	.0000	.0000	.1366	.0000	.4650
FF&DM90	.0002	.0010	.0000	.0000	.0000	.0000	.0018
DM&IT90	.0964	.1033	.0769	.1871	.0431	.0000	.5384

Dynamic Quantile Test VaR1

MODEL	*SCAL BEKK*	*DIAG BEKK*	*DCC LL MR*	*DCC LL INT*	*EX .06*	*MA100*	*O-GARCH*
NASD&DJ	.0001	.0000	.0000	.0000	.0002	.0000	.0018
DJ&RATE	.7245	.4493	.3353	.4521	.5977	.4643	.2085
NQ&RATE	.5923	.5237	.4248	.3203	.2980	.4918	.8407
DM&ITL	.1605	.2426	.1245	.0001	.3892	.0036	.0665
DM&GBP	.4335	.4348	.4260	.3093	.1468	.0026	.1125
FF&DM90	.1972	.2269	.1377	.1375	.0652	.1972	.2704
DM&IT90	.1867	.0852	.5154	.7406	.1048	.4724	.0038

Dynamic Quantile Test VaR2

MODEL	*SCAL BEKK*	*DIAG BEKK*	*DCC LL MR*	*DCC LL INT*	*EX .06*	*MA100*	*O-GARCH*
NASD&DJ	.0765	.1262	.0457	.0193	.0448	.0000	.0005
DJ&RATE	.0119	.6219	.6835	.4423	.0000	.1298	.3560
NQ&RATE	.0432	.4324	.4009	.6229	.0004	.4967	.3610
DM&ITL	.0000	.0000	.0000	.0000	.0209	.0081	.0000
DM&GBP	.0006	.0043	.0002	.0000	.1385	.0000	.0003
FF&DM90	.4638	.6087	.7098	.0917	.4870	.1433	.5990
DM&IT90	.2130	.4589	.2651	.0371	.3248	.0000	.1454

NOTE: Data are the same as in the previous table and tests are based on the results in the previous table.

[Received January 2002. Revised January 2002.]

REFERENCES

Alexander, C. O. (1998), "Volatility and Correlation: Methods, Models and Applications," in *Risk Manangement and Analysis: Measuring and Modelling Financial Risk*, ed. C. O. Alexander, New York: John Wiley.

——— (2001), "Orthogonal GARCH," *Mastering Risk* (Vol. 2), ed. C. O. Alexander, New York: Prentice–Hall, pp. 21–38.

Bollerslev, T. (1990), "Modeling the Coherence in Short-Run Nominal Exchange Rates: A Multivariate Generalized ARCH Model," *Review of Economics and Statistics*, 72, 498–505.

Bollerslev, T., Chao, R. Y., and Kroner, K. F. (1992), "ARCH Modeling in Finance: A Review of the Theory and Empirical Evidence," *Journal of Econometrics*, 52, 115–128.

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

Bollerslev, T., Engle, R., and Nelson, D. (1994), "ARCH Models," *Handbook of Econometrics* (Vol. 4), eds. R. Engle and D. McFadden, Amsterdam: North–Holland, pp. 2959–3038.

Ding, Z., and Engle, R. (2001), "Large Scale Conditional Covariance Matrix Modeling, Estimation and Testing," *Academia Economic Papers*, 29, 157–184.

Engle, R., and Kroner, K. (1995), "Multivariate Simultaneous GARCH," *Econometric Theory*, 11, 122–150.

Engle, R., and Manganelli, S. (1999), "CAViaR: Conditional Value at Risk by Regression Quantiles," NBER Working Paper 7341.

Engle, R., and Mezrich, J. (1996), "GARCH for Groups," *Risk*, 9, 36–40.

Engle, R., and Sheppard, K. (2001), "Theoretical and Empirical Properties of Dynamic Conditional Correlation Multivariate GARCH," NBER Working Paper 8554, National Bureau of Economic Research.

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

Kroner, K. F., and Claessens, S. (1991), "Optimal Dynamic Hedging Portfolios and the Currency Composition of External Debt," *Journal of International Money and Finance*, 10, 131–148.

Kroner, K. F., and Ng, V. K. (1998), "Modeling Asymmetric Comovements of Asset Returns," *Review of Financial Studies*, pp. 817–844.

Newey, W., and McFadden, D. (1994), "Large Sample Estimation and Hypothesis Testing," in *Handbook of Econometrics* (Vol. 4), eds. R. Engle and D. McFadden, New York: Elsevier Science, pp. 2113–2245.