Theoretical and Empirical properties of Dynamic Conditional Correlation Multivariate GARCH

Robert F. Engle*
University of California
at San Diego
and NYU Stern School of Business

Kevin Sheppard†
University of California
at San Diego.

First Version: February 2001
This version: December 12, 2001

Abstract

In this paper, we develop the theoretical and empirical properties of a new class of multivariate GARCH models capable of estimating large time-varying covariance matrices, Dynamic Conditional Correlation (DCC) Multivariate GARCH. We show that the problem of multivariate conditional covariance estimation can be simplified by estimating univariate GARCH models for each asset's variance, and then, using transformed residuals resulting from the first stage, estimating a time-varying conditional correlation estimator. The standard errors of the first stage parameters remain consistent, and only the standard errors for the correlation parameters need be modified. We use the model to estimate the conditional covariance of up to 100 assets using S&P 500 Sector Indices and Dow Jones Industrial Average stocks, and conduct specification tests of the estimator using an industry standard benchmark for volatility models. This new estimator demonstrates very strong performance especially considering the ease of implementation of the estimator.

JEL Classification Codes: C32, G0, G1

Keywords: Dynamic Correlation, Multivariate GARCH, Volatility

*Mailing address: 9500 Gilman Drive 0508, Department of Economics, UCSD, La Jolla, CA 92093 email: rengle@weber.ucsd.edu

†Corresponding Author, Mailing address: 9500 Gilman Drive 0508, Department of Economics, UCSD, La Jolla, CA 92093 email: kksheppard@ucsd.edu. The authors would like to thank participants in the Summer Financial Reading Group at UCSD and Andrew Patton for insightful comments. All efforts have been made to ensure the paper is mistake free and any remaining errors are the sole responsibility of the authors. Software used in the estimation in the paper is available at `http://weber.ucsd.edu/~ksheppar/research/frames.htm`

1 Introduction

While univariate GARCH models have met with widespread empirical success, the problems associated with the estimation of multivariate GARCH models with time-varying correlations have constrained researchers to estimating models with either limited scope or considerable restrictions. Large time-varying covariance matrices are needed in portfolio management and optimization, models of the term structure of treasuries or commodities, and large vector autoregressions. In this paper we describe a model which can be used to estimate extremely large time-varying covariance matrices and describe the theoretical properties of the Dynamic Conditional Correlation (DCC) Multivariate GARCH model, first introduced in Engle (2001). This class of MV-GARCH models differs from other specifications in that univariate GARCH models are estimated for each asset series, and then, using the standardized residuals resulting from the first step, a time varying correlation matrix is estimated using a simple specification. This parameterization preserves the simple interpretation of univariate GARCH models with an easy to compute correlation estimator. This multi-stage estimation requires modifying the standard errors of the parameters, however the Bollerslev-Wooldridge standard errors for each univariate GARCH model remain consistent, and only the asymptotic covariance of the parameters of the correlation estimator need be modified.

Bollerslev, Engle and Wooldridge (1988) originally proposed the multivariate GARCH model in the familiar half-vec (*vech*) form which provided a very general framework for multivariate volatility models. The full unrestricted model requires $O(k^4)$ parameters to be estimated by maximum likelihood, where k is the number of time series being modelled. A simpler model, the diagonal *vech* was also proposed which allows for non-zero coefficients only on own lagged effects and cross products, reducing the numbers of parameters needing to be estimated to be $O(k^2)$. The diagonal specification allows for a relatively straightforward interpretation, as each series has a GARCH-like specification. However, deriving the restrictions necessary on the parameters to ensure the positive definiteness of the conditional covariance becomes extremely difficult as k grows to even a moderate size.

Bollerslev (1990) introduced the constant conditional correlation (CCC) multivariate GARCH specification,where univariate GARCH models are estimated for each asset and then the correlation matrix is estimated using the standard closed form MLE correlation estimator by transforming the residuals using their estimated conditional standard deviations. The assumption of constant correlation makes estimating a large model feasible and ensures that the estimator is positive definite, simply requiring each univariate conditional variance to be non-zero and the correlation matrix to be of full rank. However, the constant correlation estimator, as proposed, does not provide a method to construct consistent standard errors using the multi-stage estimation process.

Bollerslev finds the notion of constant correlation plausible, yet recent work by Tsui and Yu (1999) has found that constant correlation can be rejected for ceratin assets. Bera (1996) and Tse (2000) both have developed tests for constant correlation, the former being a bivariate test while the latter is a more general multivariate LM test.

The BEKK formulation, proposed in Engle and Kroner (1995), developed a general quadratic form for the conditional covariance equation which eliminated the problem of assuring the positive definiteness of the conditional covariance estimate of the original *vech* model. In order for the BEKK model to be fully general, the number of parameters needing to be estimated is $O(k^4)$, but a standard BEKK estimation will involve $O(k^2)$ parameters. Other more tractable formulations of the BEKK model include diagonal and scalar which place restrictions on certain parameters to be equal to zero, although these restrictions are typically rejectable. In addition to the large number of parameters needing to be estimated for the general form, the exact interpretation of the individual coefficients is difficult to discern.

Recently, Alexander (2000) has advocated the use of factor GARCH models, as first outlined Engle, Ng, and Rothschild (1990), for estimation of large covariance matrices. Factor or Orthogonal MV-GARCH models provide a method for estimating any dynamic covariance matrix using only univariate GARCH models. Alexander shows how a limited number of factors can explain a significant amount of the volatility in certain cases. However, this approach, while reducing the numbers of parameters estimated to $o(k)$, is limited by both the difficulty in interpreting the coefficients on the univariate GARCH models and the poor performance for less correlated systems such as equities.

Engle (2001) proposed a new class of estimator that both preserves the ease of estimation of Bollerslev's constant correlation model yet allows for non-constant correlations. Dynamic Conditional Correlation MV-GARCH preserves the parsimony of univariate GARCH models of individual assets' volatility with a simple GARCH-like time varying correlation. Further, the number of parameters estimated using maximum likelihood is $O(k)$, a considerable improvement over both the *vech* and the BEKK models. More importantly, the number of parameters requiring simultaneous estimation is $O(1)$. The focus of the current paper is to explore both the theoretical and empirical properties of the DCC MV-GARCH model when estimating large conditional covariance matrices.

Tse and Tsui (1998) have also proposed a dynamic correlation multivariate GARCH model, however no attempt has been made to allow for separate estimation of the univariate GARCH processes and the dynamic correlation estimator. In addition, by not allowing for reversion to the unconditional correlation in the correlation estimator, the number of parameters needing to be simultaneously estimated is $O(k^2)$, and is only slightly less than the typical BEKK formulation.

While this estimator does possess a straight forward interpretation of the coefficients, it still will require simultaneous estimation of 32 parameters in a 5 asset model, and 167 parameters in a 15 asset model.

The paper is organized as follows. The second section outlines the model in detail and discusses the estimation procedure used. Section three establishes the asymptotic consistency and normality of the parameters, discusses the modified mispecification robust standard errors, provides a modified likelihood-ratio test that is valid under either two stage estimation or model mispecification, and discusses using a one-step Newton-Raphson iteration to achieve full efficiency. Section four describes an easy to implement test for constant correlation requiring only the estimation of a restricted VAR. Section five describes the data used and provides empirical results where systems with up to 100 assets are estimated. In section six, specification tests and benchmark comparisons are conducted and section seven describes multi-step forecasting with the DCC model. Finally, section eight concludes and outlines area of future research.

2 Model

The multivariate GARCH model proposed assumes that returns from k assets are conditionally multivariate normal with zero expected value and covariance matrix H_t.[1] The returns can be either mean zero or the residuals from a filtered time series.[2]

$$r_t|\mathcal{F}_{t-1} \sim N(0, H_t)$$

and

$$H_t \equiv D_t R_t D_t$$

where D_t is the $k \times k$ diagonal matrix of time varying standard deviations from univariate GARCH models with $\sqrt{h_{it}}$ on the i^{th} diagonal, and R_t is the time varying correlation matrix. The log-likelihood of this estimator can be written:

$$L = -\frac{1}{2}\sum_{t=1}^{T}(k\log(2\pi) + \log(|H_t|) + r_t' H_t^{-1} r_t)$$

[1]The assumptions of multivariate normality is not required for consistency and asymptotic normality of the estimated parameters. When the returns have non-Gaussian innovations, the DCC estimator can be interpreted as a quasi-maximum likelihood estimator.

[2]The standard errors of the model will not depend on the choice of filtration (ARMA, demeaning), as the cross partial derivative of the log-likelihood with respect to the mean and the variance parameters has expectation zero when using the normal likelihood.

$$
\begin{aligned}
&= -\frac{1}{2}\sum_{t=1}^{T}(k\log(2\pi) + \log(|D_t R_t D_t|) + r_t' D_t^{-1} R_t^{-1} D_t^{-1} r_t) \\
&= -\frac{1}{2}\sum_{t=1}^{T}(k\log(2\pi) + 2\log(|D_t|) + \log(|R_t|) + \epsilon_t' R_t^{-1}\epsilon_t)
\end{aligned}
$$

where $\epsilon_t \sim N(0, R_t)$ are the residuals standardized by their conditional standard deviation. We propose to write the elements of D_t as univariate GARCH models, so that

$$h_{it} = \omega_i + \sum_{p=1}^{P_i}\alpha_{ip} r_{it-p}^2 + \sum_{q=1}^{Q_i}\beta_{iq} h_{it-q} \tag{1}$$

for $i = 1, 2, \ldots, k$ with the usual GARCH restrictions for non-negativity and stationarity being imposed, such as non-negativity of variances and $\sum_{p=1}^{P_i}\alpha_{ip} + \sum_{q=1}^{Q_i}\beta_{iq} < 1$. The subscripts are present on the individual P and Q for each series to indicate that the lag lengths chosen need not be the same. The specification of the univariate GARCH models is not limited to the standard GARCH (p,q), but can include any GARCH process with normally distributed errors that satisfies appropriate stationarity conditions and non-negativity constraints. For instance, one could use TARCH to capture asymmetric effects in volatility or APARCH to allow for long memory volatility processes. The proposed dynamic correlation structure is:

$$Q_t = (1 - \sum_{m=1}^{M}\alpha_m - \sum_{n=1}^{N}\beta_n)\bar{Q} + \sum_{m=1}^{M}\alpha_m(\epsilon_{t-m}\epsilon_{t-m}') + \sum_{n=1}^{N}\beta_n Q_{t-n} \tag{2}$$

$$R_t = Q_t^{*-1} Q_t Q_t^{*-1}$$

where $\bar{Q}$ is the unconditional covariance of the standardized residuals resulting from the first stage estimation, and

$$Q_t^* = \begin{bmatrix} \sqrt{q_{11}} & 0 & 0 & \ldots & 0 \\ 0 & \sqrt{q_{22}} & 0 & \ldots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & \ldots & \sqrt{q_{kk}} \end{bmatrix}$$

so that Q_t^* is a diagonal matrix composed of the square root of the diagonal elements of Q_t. The typical element of R_t will be of the form $\rho_{ijt} = \frac{q_{ijt}}{\sqrt{q_{ii}q_{jj}}}$. The following useful result from linear algebra simplifies finding the necessary conditions for R_t to be positive definite and hence a correlation matrix.[3]

[3] A correlation matrix is defined as a real, symmetric positive semi-definite matrix, with ones on the diagonal.

Proposition 1 (Positive Definiteness) *Let A be a real, symmetric square matrix. Then A is positive definite if and only if $B = A^{*-1}AA^{*-1}$,as defined above, is positive definite.*

Proof: See Appendix

Proposition 1 establishes that for positive definiteness of R_t, we only need to ensure Q_t is positive definite. Applying this proposition, we can describe a set of sufficient conditions for H_t to be uniformly positive definite.

Proposition 2 (Positive Definiteness of DCC) *If the following univariate GARCH parameter restrictions (Equation 1) are satisfied for all asset series $i \in [1, \ldots, k]$:*

a. $\omega_i > 0$

b. $\alpha_{ip}\ \forall p \in [1, \ldots, P_i]$ *and* $\beta_{iq}\ \forall q \in [1, \ldots, Q_i]$ *are such that* h_{it} *will be positive with probability one.*[4]

c. $h_{i0} > 0$

d. The roots of $1 - \sum_{p=1}^{P_i} \alpha_{ip} Z^p + \sum_{q=1}^{Q_i} \beta_{iq} Z^q$ *lie outside the unit circle.*

and the DCC parameters satisfy (Equation 2):

e. $\alpha_m \geq 0\ \forall m \in [1, \ldots, M_i]$

f. $\beta_n \geq 0\ \forall n \in [1, \ldots, N_i]$

g. $\sum_{m=1}^{M} \alpha_m + \sum_{n=1}^{N} \beta_n < 1$

h. The minimum eigenvalue of $\bar{R} > \delta > 0$.

Then H_t will be positive definite for all t.

Proof: Each h_{it} will be strictly positive as each is a sum of ω, a strictly positive parameter with α and β such that $\sum_{p=1}^{P_i} \alpha_{ip} r_{it-p}^2 + \sum_{q=1}^{Q_i} \beta_{iq} h_{it-q}$ is non-negative with probability 1. Q_t will be positive definite for all t as it is a weighted average of a positive definite matrix ($\bar{Q}$), a positive semi-definite matrices ($\epsilon_{t-i}\epsilon'_{t-i}$) and a positive definite matrices Q_{t-i}, with non-negative parameters, and Q_0 is positive definite by Assumption h. By Proposition 1, R_t will be positive definite, and $H_t = D_t R_t D_t$ will be positive definite as it is the product of three positive definite matrices.

Essentially, the requirements for positive definiteness of the conditional covariance are the same for the DCC model as for a univariate GARCH process. The restrictions on parameters in Proposition 2 are not necessary, only sufficient to guarantee positive definiteness for H_t.

[4]Exact conditions are complicated and can be found in Nelson and Cao (1992).

3 Estimation and Standard Errors

The DCC model was designed to allow for two stage estimation, where in the first stage univariate GARCH models are estimated for each residual series, and in the second stage, residuals, transformed by their standard deviation estimated during the first stage, are used to estimate the parameters of the dynamic correlation. The likelihood used in the first stage involves replacing R_t with I_k, an identity matrix of size k. Let the parameters of the model, $\boldsymbol{\theta}$, be written in two groups $(\phi_1, \phi_2, \ldots, \phi_k, \psi) = (\boldsymbol{\phi}, \boldsymbol{\psi})$, where the elements of ϕ_i correspond to the parameters of the univariate GARCH model for the ith asset series, $\phi_i = (\omega, \alpha_{1i}, \ldots, \alpha_{P_i i}, \beta_{1i}, \ldots, \beta_{Q_i i})$. The resulting first stage quasi-likelihood function is:

$$
\begin{aligned}
QL_1(\phi|r_t) &= -\frac{1}{2}\sum_{t=1}^{T}\left(k\log(2\pi) + \log(|I_k|) + 2\log(|D_t|) + r_t' D_t^{-1} I_k D_t^{-1} r_t\right) && (3)\\
&= -\frac{1}{2}\sum_{t=1}^{T}\left(k\log(2\pi) + 2\log(|D_t|) + r_t' {D_t}^{-2} r_t\right) && (4)\\
&= -\frac{1}{2}\sum_{t=1}^{T}\left(k\log(2\pi) + \sum_{n=1}^{k}\left(\log(h_{it}) + \frac{r_{it}^2}{h_{it}}\right)\right) && (5)\\
&= -\frac{1}{2}\sum_{n=1}^{k}\left(T\log(2\pi) + \sum_{t=1}^{T}\left(\log(h_{it}) + \frac{r_{it}^2}{h_{it}}\right)\right) && (6)
\end{aligned}
$$

which is simply the sum of the log-likelihoods of the individual GARCH models for each of the assets. Once the first stage has been estimated, the second stage is estimated using the correctly specified likelihood, conditioning on the parameters estimated in the first stage likelihood:

$$
\begin{aligned}
QL_2(\psi|\hat{\phi}, r_t) &= -\frac{1}{2}\sum_{t=1}^{T}(k\log(2\pi) + 2\log|D_t| + \log(|R_t|) + r_t' D_t^{-1} R_t^{-1} D_t^{-1} r_t) && (7)\\
&= -\frac{1}{2}\sum_{t=1}^{T}(k\log(2\pi) + 2\log|D_t| + \log(|R_t|) + \epsilon_t' R_t^{-1}\epsilon_t) && (8)
\end{aligned}
$$

Since we are conditioning on $\hat{\phi}$, the only portion of the log-likelihood that will influence the parameter selection is $\log(|R_t|) + \epsilon_t' R_t^{-1}\epsilon_t$, and in estimation of the DCC parameters, it is often easier to exclude the constant terms and simply maximize:

$$
QL_2^*(\psi|\hat{\phi}, r_t) = -\frac{1}{2}\sum_{t=1}^{T}(\log(|R_t|) + \epsilon_t' R_t^{-1}\epsilon_t)
$$

White (1994) has provided a proof for the asymptotic distribution for two-stage QMLE estimators. The proofs for consistency and asymptotic normality of the parameter estimates of the two stage DCC estimator closely follow the results presented in White. The following set of assumptions are sufficient to establish the consistency of the parameters estimated using this two stage procedure, in addition to standard assumptions guaranteeing the completeness of the probability space and measurability of the quasi-likelihood functions.

A 1 *(i) For each* ϕ *in* Φ, $E(\log f_1(r_t, \phi))$ *exists and is finite,* $t = 1, 2, \ldots$,
(ii) $\{log f_1(r_t, \phi)\}$ *obeys the strong uniform law of large numbers.*
(iii) For each $\theta = (\phi, \psi)$ *in* $\Theta = \Phi \times \Psi$, $E(\log f_2(r_t, \theta))$ *exists and is finite,* $t = 1, 2, \ldots$
(iv) $\{log f_2(r_t, \theta)\}$ *obeys the strong ULLN.*

A 2 *(i)* $\theta_0 = (\phi_0, \psi_0)$ *is identifiably unique, interior in* $\Theta = \Phi \times \Psi$ *uniformly in* n, Θ *is compact, and* θ_0 *satisfies the conditions of Proposition 2.*
(ii) $\{\bar{L}_{1T}(\phi) = E(T^{-1} \sum_{t=1}^{T} \log f_1(r_t, \phi))\}$ *is* $O(1)$ *uniformly on* Φ
(iii) $\{\bar{L}_{2T}(\theta) = E(T^{-1} \sum_{t=1}^{T} \log f_2(r_t, \theta))\}$ *is O(1) uniformly on* Θ

A 3 *(i) For all* ϕ *in* Φ, $\nabla \bar{L}_{1T}(\phi) = E(\nabla L_{1T}(r^T, \phi)) < \infty$, *where* $r^T = (r_1, r_2, \ldots, r_T)$, *the T-dimensional vector of observations.*
(ii) For all θ *in* Θ, $\nabla \bar{L}_{2T}(\theta) = E(\nabla L_{2T}(r^T, \theta)) < \infty$

A 4 *(i) For all* ϕ *in* Φ, $\nabla^2 \bar{L}_{1T}(\phi) = E(\nabla^2 L_{1T}(r^T, \phi)) < \infty$
(ii) $E(\nabla^2 L_{1T}(r^T, \cdot))$ *is continuous on* Φ *uniformly in* $T = 1, 2, \ldots$
(iii) $\{\nabla^2 log f_1(r_t, \phi)\}$ *obeys the strong ULLN.*
(iv) For all θ *in* Θ, $\nabla^2 \bar{L}_{2T}(\theta) = E(\nabla^2 L_{2T}(r^T, \theta)) < \infty$
(v) $E(\nabla^2 L_{2T}(r^T, \cdot))$ *is continuous on* Θ *uniformly in* $T = 1, 2, \ldots$
(vi) $\{\nabla^2 log f_2(r_t, \phi)\}$ *obeys the strong ULLN.*

A 5 *(i)* $\{A_{11,T} = \nabla_{\phi\phi} \bar{L}_{1T}(\phi_0)\}$ *is* $O(1)$ *and uniformly negative definite.*
(ii) $\{A_{22,T} = \nabla_{\psi\psi} \bar{L}_{2T}(\theta_0)\}$ *is O(1) and uniformly negative definite.*

Theorem 1 (Consistency) *Under assumptions A1 - A5,* $\hat{\phi}_T \xrightarrow{p} \phi_0$ *and* $(\hat{\phi}_T, \hat{\psi}_T) = \hat{\theta}_T \xrightarrow{p} \theta_0$.

Details of the proof can be found in White. The conditions for consistency are very weak and will be satisfied by numerous data generating processes. We now will add sufficient regularity conditions to allow for asymptotic normality of the estimated parameters. The following additional assumption is needed for this result:

A 6 *(i)* $\{T^{-1/2}\nabla'_{\phi}\ln f_1(r_t,\phi_0), T^{-1/2}\nabla'_{\psi}\ln f_2(r_t,\phi_0,\psi_0)\}$ *obeys the central limit condition with covariance matrix* B_{0T}*, and* B_{0T} *is* $O(1)$ *and uniformly positive definite.*

Using these assumptions the following theorem establishes the asymptotic distribution of the two stage estimation process used in estimating DCC-MVGARCH models.

Theorem 2 (Asymptotic Normality) *Under assumption A1 - A6, for* f_1 *and* f_2*,*

$$\sqrt{T}(\hat{\theta}_T - \hat{\theta}_0) \overset{A}{\sim} N(0, {A_0}^{-1} B_0 {A_0}'^{-1})$$

where

$$A_0 = \begin{bmatrix} \nabla_{\phi\phi}\ln f_1(\phi_0) & 0 \\ \nabla_{\phi\psi}\ln f_2(\theta_0) & \nabla_{\psi\psi}\ln f_2(\theta_0) \end{bmatrix} = \begin{bmatrix} A_{11} & 0 \\ A_{12} & A_{22} \end{bmatrix}$$

and

$$B_0 = var\left[\sum_{t=1}^{T}\{T^{-1/2}\nabla'_{\phi}\ln f_1(r_t,\phi_0), T^{-1/2}\nabla'_{\psi}\ln f_2(r_t,\phi_0,\psi_0)\}\right] = \begin{bmatrix} B_{11} & B_{12} \\ B_{12} & B_{22} \end{bmatrix}$$

Following from the theorem, the asymptotic variance of $\hat{\theta}_n$ is given by ${A_0}^{-1}B_0{A_0}^{-1}$. Applying the partitioned inverse theorems for square matrices, the asymptotic variances of the GARCH parameters for each asset, $\hat{\phi}_n$ are the standard Bollerslev-Wooldridge robust covariance matrix estimators given by ${A_{11}}^{-1}B_{11}{A_{11}}^{-1}$.[5] The asymptotic variance of the second stage DCC parameters is however a much more complicated formula involving all of the covariance terms. Further, the asymptotic variance of the DCC parameters is (weakly) greater in this case than would be the case where A_{12} was a zero matrix, due to the loss of efficiency in estimating the parameters separately.

In addition to having modified standard errors, a likelihood-ratio test with r restrictions will not typically be χ^2_r. Both Foutz and Srivastave (1977) and Liang and Self (1996) have discussed likelihood ratio testing when either the distributional assumption is incorrect or a model is estimated in multiple steps. Both of these results demonstrate when the Information Matrix Equality does not hold, i.e. $A_0 - B_0 \overset{p}{\nrightarrow} 0$, then the asymptotic distribution will be a weighted sum of r independent χ^2_1 variables where the weights will not necessarily be unity. In this case, where the limiting distribution of the parameters is known (Theorem 2), Foutz and Srivastave have shown that to test the null of $H_0 : \theta \in \Theta_r$ against $H_1 : \theta \in \Theta$ that

$$-2\ln\lambda = -2\ln\frac{\max\limits_{\theta\in\Theta}\{\prod_{i-1}^{T} f(r_t,\theta)\}}{\max\limits_{\theta\in\Theta_r}\{\prod_{i-1}^{T} f(r_t,\theta)\}} \xrightarrow{d} C$$

[5] ${A_{11}}^{-1}B_{11}{A_{11}}^{-1}$ will be a block diagonal matrix with the covariance matrix for the i^{th} univariate GARCH model on the i^{th} diagonal block.

where

$$C \sim c(\theta_0)_1\chi_1^2 + c(\theta_0)_2\chi_1^2 + \ldots + c(\theta_0)_r\chi_1^2$$

where $c(\theta_0)_i$ is the i^{th} eigenvalue of $W(\theta_0)M(\theta_0)$, $W(\theta_0) = (-\tilde{A}_1 + \tilde{A}_2\tilde{A}_4^{-1}\tilde{A}_3)$ and $M(\theta_0)$ is the upper $r \times r$ matrix of $\tilde{A}^{-1}\tilde{B}\tilde{A}^{-1}$, where $\tilde{A}$ is A, as defined in theorem 2, with the rows and columns of the restricted parameters interchanged, so that the first r rows and columns of $\tilde{A}$ correspond to the standard errors of the restricted parameters being tested and $\tilde{B}$ is similarly defined. In testing one restriction, we have $(W(\theta_0)M(\theta_0))^{-1}\hat{\lambda} \sim \chi_1^2$. Another useful result in understanding the DCC estimator is that the expectation of the scores of the full loglikelood, evaluated at the estimated parameters from the first stage estimate, is zero, and thus the parameters of the two stage process have the same limit as a jointly estimated model. Proposition three details this claim.

Proposition 3 (Univariate GARCH specification) *If $E_{t-1}(r_{it}^2) = h_{it}$, then the expectation of the partial gradient with respect to ϕ of the second-stage log-likelihood evaluated at the parameters estimated using the first-stage log-likelihood is zero, i.e. $E(\nabla_\phi \ln f_2(r, \hat{\phi}_1(\cdot), \hat{\psi}_2(\cdot)) = 0$.*

Proof: See Appendix

While the parameters are consistent and asymptotically normal with a known covariance, they are not fully efficient as both sets of parameters are estimated using LIML. However, due to a result from Pagan (1986) among others, we know that when an estimator is root-n consistent, that a fully efficient estimate requires only one step using a Newton-Raphson algorithm that involves using the second stage likelihood to obtain consistent estimates of the first and second derivatives.[6] Lemma one gives an exact statement of the one step efficiency for the DCC estimator.

Lemma 1 (One Step Efficiency) *If $\hat{\theta}_1 = \hat{\theta}_{2S} + \widehat{A_{\theta\theta}}^{-1}\nabla_\theta \ln \widehat{f_2(r, \phi_0, \psi_0)}$ where $\widehat{A_{\theta\theta}}$ is a consistent estimate of the second derivative of the QL_2 and $\nabla_\theta \ln \widehat{f_2(r, \phi_0, \psi_0)}$ is a consistent estimate of the gradient of QL_2, then $T^{1/2}(\hat{\theta}_1 - \theta_0)$ has the same limiting distribution as $T^{1/2}(\hat{\theta} - \theta_0)$, where $\hat{\theta}_1$ would is a one step efficient estimator and $\hat{\theta}$ is a standard one-stage estimator.*

Proof: See Pagan (1996)

4 Testing for Constant Correlation

One of the primary motivations for this paper is that the correlations between assets are not constant through time. Testing data for constant correlation has proven to be a difficult problem, as

[6] Estimates of the derivatives can be easily computed from the second stage likelihood function numerically.

testing for dynamic correlation with data that have time-varying volatilities can result in misleading conclusions and can lead to rejecting constant correlation when it is true due to misspecified volatility models. Tse (1998), testing a null of constant conditional correlation against an ARCH in correlation alternative, and Bera (1996), testing a null of constant conditional correlation against a diffuse alternative, have provided tests of a null of constant correlation against an alternative of a dynamic correlation structure. One short coming of both of these tests is that they do not generalize well to higher dimensions. We propose a test that only requires a consistent estimate of the constant conditional correlation, and can be implemented using a vector autoregression.

We are interested in testing the null of constant correlation against an alternative of dynamic conditional correlation. However, there is a significant difficulty in conducting this test as the decay parameters (β_i) in the DCC estimator are unidentified under the standard null hypothesis and must be treated as nuisance parameters. There are two ways to treat this. One is to merely test the null against an alternative with a specific coefficient for beta β. This test can be conducted using standard likelihood ratio test with the usual properties of LR testing holding, however it may lack power if the chosen coefficient for β is far form the truth. However, this is an unnecessarily restrictive test as it should not be necessary to identify β. Andrews and Ploberger (1994) establish a procedure by which tests with unidentified parameters can be conducted and only recently has this framework been extended to cases where the parameter unidentified under the null can be on the boundary of the parameter space. Implementing this type of test is very difficult as it requires many optimizations and monte carlo critical values. The test we propose is

$$H_0 : R_t = \bar{R} \qquad \forall t \in T$$

against

$$H_a : vech^u(R_t) = vech^u(\bar{R}) + \beta_1 vech^u(R_{t-1}) + \beta_2 vech^u(R_{t-2}) + \dots \beta_p vech^u(R_{t-p})$$

where $vech^u$ is a modified $vech$ which only selects elements above the diagonal. The testing procedure is as follows. Estimate the univariate GARCH processes and standardize the residuals for each series. Then estimate the correlation of the standardized residuals, and jointly standardize the vector of univariate standardized residuals by the symmetric square root decomposition of the $\bar{R}$.[7] Under the null of constant correlation, these residuals should be IID with a variance covariance

[7]While the $E(\epsilon_t \epsilon_t')$ is a correlation matrix, in finite samples this is never going to occur. In practice it is more efficient to use $\bar{R}$ as the covariance matrix in place of the correlation matrix. By using a correlation matrix, the test is further weakened as the test is also sensitive to the standardized variance of the univariate GARCH processes not being unity.

matrix given by I_k.[8] The artificial regressions will be a regression of the outer products of the residuals on a constant and lagged outer products. Let $Y_t = vech^u \left[(\bar{R}^{-1/2}D_t^{-1}r_t)(\bar{R}^{-1/2}D_t^{-1}r_t)' - I_k\right]$ where $\bar{R}^{-1/2}D_t^{-1}r_t$ is a k by 1 vector of residuals jointly standardized under the null. The vector autoregression is

$$Y_t = \alpha + \beta_1 Y_{t-1} + \ldots + \beta_s Y_{t-s} + \eta_t$$

Under the null the intercept and all of the lag parameters in the model should be zero. In order to estimate the test statistic, all that is necessary to do is to make the $T \times 1$ vector of outer-products for each univariate regressand and the $T \times s + 1$ matrix of regressors including the constant for each set of regressors. Then the parameters can be estimated by stacking the $k(k-1)/2$ vectors of regressands and regressors and performing a seemingly unrelated regression. The test can then be conducted as $\frac{\hat{\delta}X'X\hat{\delta}'}{\hat{\sigma}^2}$ which is asymptotically $\chi^2_{(s+1)}$, where $\hat{\delta}$ are the estimated regression parameters and X is a matrix consisting of the regressors.[9] In every model considered in this paper, we reject the null of a constant correlation in favor of a dynamic structure.

A Monte Carlo study was conducted to study size and power of the test for constant conditional correlation. In addition to studying the size under correctly specified models for the conditional variance, we felt that studying the size with misspecified conditional variance models was also warranted, as in practice one can never be completely sure of correct specification. To begin studying the size under the simplest case, correct specification of the conditional variance models, we used 4 different levels of correlation for the constant conditional correlation: The first three simply set the constant conditional correlations (between all pairs of assets) to .3, .5 and .8, while the fourth drew, at random, unconditional correlations from the correlation matrix of the 100 S&P 500 indices, by selecting the appropriate number of rows and columns. All univariate GARCH models were GARCH(1,1) with $\omega = 0.1$, $\alpha = 0.1.$, and $\beta = 0.8$, resulting in an unconditional variance of unity. The tests were conducted with the number of assets ranging over 2, 3, 4, 5 and 10; 5000 thousand replications were performed for each choice of unconditional correlation and number of assets. Table 1 contains the results for the size of the test with 5 lags when the critical value is chosen at the 5% level form a χ^2_6 distribution.

The test was under sized for all correlations and number of assets. This under size of the test is most likely due to parameter estimation uncertainty in the conditional variance estimation.

[8] A limitation of this test is that it cannot necesarily differentiate between a dynamic correlation structure and misspecified conditional heteroscedasticity of the univariate series. However, as demonstrated in this section, the test has od size properties even when the univariate models are misspecified.

[9] If the covariance of the standardized residuals is used in place of the correlation, the intercept will necessarily be zero, and the test could be conducted as $(k)(k-1)TR^2$ from the regression which will be distributed $\chi^2_{(s)}$.

However, the size distortion is quite manageable, typically falling in a very narrow range, about 2% too low. The second experiment conducted was to examine the power of the test when the conditional variance models are misspecified. To examine this, we used the different GARCH specifications for the univariate volatility generation, and estimated a GARCH(1,1) in lieu of the correct model. The three misspecifications considered where a GARCH(2,1), a GARCH(2,2) and a GARCH(3,2). The parameters of the three models were [0.1, 0.05, 0.05, 0.85], [0.1, 0.05, 0.05, 0.5, 0.35], and [0.1, 0.05, 0.05, 0.05, 0.5, 035] respectively. As the unconditional correlation does not appear to affect the size of the test, the unconditional correlation was set to be .5 between all assets. Table 2 contains the size of the tests when the conditional volatility models are misspecified. The conditional volatility model misspecidifcation does not appear to affect the size of the test in the least. The test is still under sized with a size of approximately 3% when using a 5% critical value form a χ^2_6. The robustness to this type of misspecification makes this a very attractive test, allowing of fitting a GARCH(1,1) model for the conditional volatility when a specification search could be costly.[10]

Finally, we were interested in the power of this test in detecting DCC. The power experiment was conducted using 2, 5 and 10 assets with the unconditional correlation set to 0.5. The experiment allows the $\alpha \times \beta$ to range over $[0.005, 0.01, 0.02, \ldots, .1] \times [.9, .91, \ldots, .99]$ subject to $\alpha + \beta < 1$. The test appears to have good power properties, even against near alternatives. When considering 2 asset models, the power ranged form a low of .423 to one for models with large α. The 5 asset power test was similar to the 2 asset, having power of about $\frac{1}{2}$ for near alternatives, and increasing to have unity power for distant alternatives. For the 10 asset models, the test had unity power for all choices of $\alpha > .01$, and had a minimum power of .657 when $\alpha = 0.005, \beta = 0.9$. Naturally, the power is lowest for small values of α, and is increasing in β. Tables 3, 4, and 5 contain the calculated power of the test for the varying combinations of α and β. The proposed test seems to have consistent size properties and good power, ranging from nearly one-half for local alternatives to unity for distant alternatives.

5 Empirical Results

The data used in this paper consist of 100 S&P 500 Sector Indices including the S&P 500 Composite and the 30 Dow Jones Industrial Average stocks plus the average. Both data series were from January 1, 1994 until December 31, 1999, and were provided by Datastream. All days that the

[10] While the test appears robust to misspecification, it is necessary that the volatility models to have correct dynamics for the standard errors presented in Section 3 to be correct. Thus, we would recommend using an information criteria in selecting the univariate GARCH models.

market was closed were removed, with the number of days removed being either eight or nine, depending on the year.[11] After removing these days there were 1509 observations for the sample. We feel these two sets of data allow for meaningful analysis of high dimensional systems without experiencing problems due to thinly traded assets or nonsynchronous issues.

The model used in the empirical section was a simple DCC (1,1)-MVGARCH where each of the univariate GARCH models estimated for the conditional variances was selected by finding the minimum of the AIC allowing for $P < 4$ and $Q < 3$.[12] In addition, an integrated form of this model was estimated where $\lambda = 1 - \beta = \alpha$ was imposed. The models were built in an expanding fashion so that the three asset model included the assets of the two asset model plus an additional asset, the four asset model nested the three asset model, and so forth.[13]

Table 6 summarizes the estimated $\hat{\alpha}$ and $\hat{\beta}$ of the mean reverting model for the different portfolios of assets using the S&P indices, as well as the value of the statistic derived from a corrected likelihood ratio test that tests $H_0 : \alpha = 1 - \beta$ and thus the correlations have an integrated structure. T-stats are reported for each coefficient using the mispecification robust standard errors in parenthesis. Table 6 also reports the estimated $\hat{\lambda}$ and the modified LR statistic of the test of the null of an integrated model against an alternative of a mean reverting model.[14] For every model with more than six assets, the estimated $\hat{\lambda}$ was on the boundary, and the test statistic is most likely not distributed χ^2_1.[15] Also, the mean reverting model was preferred to the integrated model for all data sets, with the rejection of the integrated model occurring at the 0.1% level for all models. The estimated parameters imply a highly persistent correlation, with a half-life of innovation of 21 days for the 10 asset model.

Table 7 presents the estimated parameters for the Dow Jones Industrial Average stocks. All of the coefficients of the mean reverting models were significant at the 5% level, with most significant at the 1% level. The estimated integrated parameter $\hat{\lambda}$ was in all instances estimated at the boundary of 0. The typical estimated set of parameters had slow decay ($\hat{\beta} > .97$) with a small news parameter ($\hat{\alpha} < .01$). While the distribution of the LR is not know, we would solidly reject

[11] There were at least 8 days removed each year:New Year's Day, President's Day, Good Friday, Memorial Day, Independence Day, Labor Day, Thanksgiving Day, and Christmas Day. In addition in 1994, the markets were closed for President Nixon's funeral, and in 1998 and 1999, the markets were closed on Martin Luther King's Day

[12] The data were not filtered other than simple subtraction of the mean.

[13] The S&P 500 was included as the first asset for the models estimated S&P 500 data with the remaining assets entering in alphabetical order, while the Dow Jones Industrial Average was included as the 31^{st} asset in the DJIA models. There was no perfectly redundant asset in the 31 asset Dow Jones model as the DJIA uses variable weights.

[14] The nuisance parameter is only present for the test of the mean reverting model against an alternative of a constant correlation, and is not present for either the null of an integrated model against the alternative of constant correlation, or a null of mean reverting against an alternative of an integrated model.

[15] When $\hat{\lambda} = 0$, the first test statistic can be interpreted as a test of $H_0 : \alpha = 0, \beta = 0$ against a DCC.

the null of constant correlation in favor of dynamic conditional correlation at the 1% level based on the likelihood ratio test statistic if it were between a χ^2_1 and a χ^2_2.

To ensure that the expanding data sets used in estimation were not driving the results that the parameters seem to settle down and are contained in a fairly narrow range, we estimated 10 models where the data series were chosen at random from the 100 data series of the S&P 500 indices used in the paper. Figure 1 contains the estimated $\hat{\alpha}$ and $\hat{\beta}$ for the ten models. The $\hat{\alpha}$ range from 0.004 to 0.013 while the $\hat{\beta}$ range from 0.86 to 0.98. None of the models produced parameters which were constrained on the boundary of $\alpha + \beta \leq 1$. Comparing these values to the estimated coefficients of 0.0255 and 0.9410 for the 10 asset model estimated using the S&P 500 indices, these typically have a slightly smaller news impact parameter and a slightly larger decay parameter, most likely due to the correlation structure of the different sector indices and the S&P 500 index. We also estimated the integrated model on these same ten pairs of assets, finding the estimated parameter to be on the boundary for all but one of the models. For the one set of assets where the integrated model did not estimate the parameter on the boundary, the null of an integrated moving average was rejected in favor of an alternative mean reverting model at 0.001%. We feel that the evidence does not provide support for the integrated version as all models considered preferred the mean reverting model using the LR test. In writing this paper, we considered using a longer sample of 10 years. When estimating models with ten years of data, we found that the integrated model typically had a coefficient that was significantly different from zero, although the mean reverting model was preferred to the integrated model for this data length as well. One possible explanation for this phenomena is that there are breaks in the unconditional correlation, which allow the integrated model to fit better than a constant correlation model over longer horizons. This line of research is beyond the scope of this paper, and will be saved for future work.

Finally, we estimated a variety of specifications for the DCC model, allowing for more lags of both the news term (α) and the decay term (β). Table 8 contains the corrected likelihood ratio results for both the S&P and the Dow assets. The models estimated were DCC (2,1), DCC (2,2), and DCC (3,2), where the first number represents the number of α lags included. The models of the Dow assets never preferred a longer lag length, with most parameters estimated as zero. The likelihoods improved when using the (3,2) specification, although none were significant. The original specification was always preferred to the (2,1) specification for all S&P 500 index models. The DCC (2,2) was preferred to the DCC (1,1) for the 25 asset model, and 10 of the 12 models estimated with the DCC (3,2) rejected the null of a DCC (1,1) process. However, upon inspection of the fitted correlations, the larger models typically generated correlation which were much noisier than the original model, although the basic dynamics of the correlations remained the same.

Figure 2 contains a plot of the cumulative returns of the 4 S&P 500 indices used in estimation, while figure 3 plots the time-varying correlations of the S&P 500 composite and 3 sector indices. Also in figure 3 are the CCC estimated correlations and 95% confidence intervals. The correlations range form 0.2 to 0.8 over the course of the sample, and are often outside of the confidence intervals from the CCC. Figure 4 contains the estimated variances and correlations for these stocks over the sample period. The variances for each series are simply the result of the univariate GARCH specifications. Figure 5, contains the minimum variance portfolio weights for these same four assets.[16] For the first 3 years of the sample, the weights are relatively constant with a high loading in the S&P 500 index and slight shorts in most of the others. However, in the latter periods of the sample, the weights become extremely volatile and include a short position in the S&P 500 index and a weight greater than 1 in the S&P 500 Auto Parts index. Figure 6 contains the dynamic correlations of the first four Dow Jones Industrial Average stocks, as well as the CCC estimated correlations and confidence intervals. Consistent with the smaller news parameter estimated for this model, the conditional correlations are less volatile than the four asset S&P 500 model. However, there do appear to be periods where the correlation is above or below the unconditional correlation for hundreds of days, and outside of the CCC confidence bands.

6 Specification Testing

In order to test the specification of the DCC model, we propose to examine the model's performance using three methods: the standard deviation of portfolios of returns standardized by the portfolio standard deviation implied by the conditional covariance estimate, Value-at-Risk performance, and relative performance to the industry standard RiskMetrics exponential smoother. The first test we conducted involved testing the variance of returns of portfolios against the predicted variance. We use three different portfolio weighting methods: equally weighted, value weighted, and minimum variance portfolio weighted. The minimum variance portfolio is of particular interest as the weights on the assets are determined by the estimated variance covariance matrix. We feel that if a particular estimated conditional covariance estimate is misspecified, that the minimum variance portfolio should exacerbate the short coming. The time-varying weights on the minimum variance portfolio were calculated using

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

where

$$C_t \equiv \iota' H_t^{-1}\iota$$

[16]The procedure used to construct minimum variance weights is discussed in section 6.

and where H_t is the one-step ahead forecast of the conditional covariance constructed at time $t-1$, and ι is a k by 1 vector of ones. There was no limit on short selling, however no portfolio ever took an extreme position. The portfolio weights of the value weighted portfolio were calculated using the following formula:

$$w_{t+1} = \frac{w_t \odot (1 + r_t)}{(1 + r_t)'\iota}$$

where r_t is a k by 1 vector of time t asset returns with $w_0 = k^{-1}\iota$. Finally a simple equally weighted portfolio was used where the weights were $w_t = k^{-1}\iota$. If the model was correctly specified for the conditional covariance, then we would expect that the variance of any portfolio with weights w_t would be $w_t' H_t w_t$. To test this hypothesis we used a symmetric confidence interval with $\alpha/2$ probability in each tail.[17] Portfolio variances which are too small relative to the predicted variance are indication of excess correlation while variances which are too big indicate underestimation of the correlation. When standardized by their estimated standard deviations, all assets had standardized (by univariate GARCH processes) residual variance that were in the 95% confidence interval centered at 1 for both the S&P 500 indices and the DJIA stocks. The RiskMetrics EWMA covariance filter is

$$H_t = .06 e_t e_t' + .94 H_{t-1}$$

where H_0 can be taken to be the sample covariance matrix or a presample data selection to begin the smoother. The primary advantage of the RiskMetrics model is that it is extremely easy to estimate, given that it has no parameters to estimate. According to RiskMetrics, the choice of the smoothing parameter has been calibrated using an extensive model search. The obvious drawback to the model is that it has no estimated parameters, and that it forces all assets have the same smoothing coefficient (.94) irrespective of the type of asset of the asset's volatility dynamics. The RiskMetrics model is widely used in industry, especially for portfolio Value-at-Risk.

For the S&P 500 indices, the minimum variance portfolio proved to be troublesome for both the mean reverting DCC estimator and the industry standard RiskMetrics EWMA. The standardized variances of the minimum variance portfolios only fall with in the confidence interval for portfolios under 5 assets using the DCC estimator, while the predicted variance of the RiskMetrics model is uniformly too small, resulting in a portfolio standard deviation 35 times larger than what would have been expected. For the equally weighted and value weighted portfolios, the DCC estimator produced portfolio standard deviations insignificantly different from one for all portfolios. Further,

[17] Confidence intervals were constructed to find a and b such that $\int_0^a f(u)du = \int_b^\infty f(u)du$, $\sigma : \{\frac{(t-1)s^2}{b} \leq \sigma \leq \frac{(t-1)s^2}{a}\}$ where $f(u)$ is the probability density function of a χ^2_{t-1}.

we reject the null that the variance is equal to one for all of the RiskMetrics generated portfolios using equally weighting or value weighting at the 5% level. Table 9 summarizes the results for models estimated on the S&P 500 indices.

The performance of the DCC estimator on the DJIA stocks was similar. We reject the null that the portfolio variance is one for all of the RiskMetrics portfolios, and for the larger DCC generated minimum variance portfolio. We are also not able to reject the hypothesis that the portfolio standard deviations are unity for any of the value or equally weighted portfolio variances estimated by the DCC model. The slight better performance of DCC and slightly worse performance of the RiskMetrics model is most likely due to the larger news parameters of the RiskMetrics model ($\lambda = .06$). These tests on the DCC estimator show for many portfolio strategies it performs well at fitting the portfolio variance, although the structure of the dynamics of the correlation may need to allow for more than one news parameter as the size of the portfolio grows. Table 10 summarizes the results of the portfolio variance diagnostic test for the DJIA stocks.

The second measure of performance used to test the empirical validity of the models was the HIT test (Engle and Manganelli (2000)). The test is designed to test the performance of a model's prediction of Value-at-Risk. A series of HITs is defined as a binary variable, $\mathbf{1}_{[r_t<VaR(q)]}$, where a hit is a return below the forecasted Value-at-Risk and q is the VaR quantile. Under the null of a correctly specified model for Value-at-Risk, the HITs should have mean q and should be independent of everything in the conditioning information set, including lagged HITs. An artificial regression can be constructed to test both the mean and the independence of the HITs using OLS:

$$HIT_t - q = \delta_0 + \delta_1 HIT_{t-1} + \delta_2 HIT_{t-2} + \ldots + \delta_r HIT_{t-r} + \delta_{r+1} VaR_t + \nu_t$$

The artificial regression tests the independence of a HIT from past hits and from the predicted Value-at-Risk (which is in the time $t-1$ information set).[18] We constructed one step ahead forecasts of the variance of a portfolio using the mean reverting DCC model. The 5% VaR was defined as $-1.65\hat{\sigma}_p$ where $\hat{\sigma}_p$ is the forecasted portfolio standard deviation. Under the assumption that returns are conditionally multivariate normal, this level would be appropriate. However, all of the standardized residuals from either the S&P 500 indices or the Dow stocks reject the null of normality using a Jarque-Bera test at the 5% level. This leptokurtosis of the univariate returns would invalidate using the $-1.65\hat{\sigma}_p$ as a test for the number of HITs. We found that when including the constant term in the artificial regression, we were able to reject the null of a correct VaR model for all cases. Thus, in the test conducted, we tested for independence of the HITs

[18] This is the out of sample version of the HIT test as presented in Engle and Manganelli. As noted in their paper, the distribution of the HITs is not known when the same data are used for fitting the model and evaluating the HIT test, although this test is still a useful diagnostic.

without simultaneously testing if the percentage of HITs was correct by excluding the constant and replacing q with $\hat{q}$,the mean number of HITs at $-1.65\hat{\sigma}_p$[19] The DCC model results in a consistently lower percentage of HITs when estimated on equally weighted portfolios of S&P 500 indices data (Table 11). Further, even with the mean of the HITs subtracted, there was still serial correlation in the HITs when modelled on the past 5 HITs and the forecasted VaR. The VaR performance on the minimum variance portfolio varies from slightly better to considerably worse, with nearly 5% for smaller portfolios to 10% when estimated on all 100 indices. This is consistent with the estimated variance performance of the DCC on the minimum variance portfolio significantly underestimating the variances as the number of assets increased.

The performance of the DCC model using the HIT test was better for the Dow Jones stocks than it was for the S&P 500 indices. However, the DCC again consistently underestimated the number of violations of Value-at-Risk for the equally weighted portfolio at the 5% level. Table 12 contains the results of the HIT testing for the DJIA 30 stocks. The performance relative to the 1% level is much more accurate, usually falling within 20% of the anticipated level.[20] The probability the HITs are independent was also better for all portfolios, however for the largest portfolios the HITs seem to be correlated with past hits and/or contemporaneous VaR. The performance of the minimum variance portfolio was generally worse than the equally weighted portfolio with respect to the percentage of HITs, but better with regards to the probability of independence. We fail to reject the null of independence of the HITs for none of the 14 portfolios at the 5% level using the minimum variance portfolio.

The final specification test we used was to compare the results of the DCC estimator against the estimator used widely by practitioners, the RiskMetrics exponential smoother. The comparison was conducted using 4 criteria. The first criterion examines the percentage of HITs predicted with each model using the equally weighted portfolio. The RiskMetrics estimator consistently outperformed the DCC estimator in terms of the correct percentage of hits. The RiskMetrics filter using S&P 500 indices data also performs slightly better on the HIT test, however the performance of both estimators is inadequate by this metric, with both series having at least 11 out of 15 series failing independence at the 5% level. A second criteria is the percentage of multivariate variance standardized residuals which have variance in a confidence interval of one. The test relies on the assumption that assets are multivariate normal, so

$$r_t|\mathcal{F}_{t-1} \sim N(0, H_t) \Rightarrow H_t^{-\frac{1}{2}} r_t|\mathcal{F}_{t-1} \sim N(0, I)$$

[19] An alternative aproach would be to treat the DCC estimator as a QMLE and use the empirical distribution when choosing the scaling on the variance in a VaR setting.

[20] However this difference was still significant at the 5% level.

The DCC estimator produced standardized residuals in the range for all models with less than 10 assets, and performed well in this metric for larger models. The RiskMetrics model produced no standardized series in the 95% confidence interval for any of the models. This is consistent with the standardized portfolio variance of Table 13 where the RiskMetrics estimator consistently underestimated true variance. The final test was a Ljung-Box Q test at lag 15 of the residual to determine if there was excess serial correlation in the squares and cross products of the covariance standardized residuals using a 5% significance level. We found that the percentage of cross products failing (there were $(k \times k + 1)/2$ cross products for a given number of assets k) was relatively small for the DCC estimator for assets with models with fewer than 10 assets , with 15% typically failing the test for serial correlation for larger models. The RiskMetrics model standardized residual consistently performed worse, having over 40% of the residuals failing this test for the largest models. We also found that the percentage failing was always greater than the 5% which would have been expected with the test at the 5% level.

The Value-at-Risk calculation using the DJIA stocks performed similarly to the S&P Indices, with the percentage of HITs being consistently under estimated for the equally weighted portfolios and with the RiskMetrics estimator was always closer to the expected percentage of HITs for all but one portfolio. In addition, the HITs generated by the RiskMetrics model only fail the independence test for one of the 15 portfolios at the 1% level, while 7 out of the 15 DCC generated portfolios fail the independence test at that level. However, for all assets and sizes of the models, with the exception of one asset in the 31 asset model, the standard deviations of the returns standardized by the square root of the estimated covariance were always in the confidence interval for the DCC estimator. In addition, the DCC estimator significantly outperformed the RiskMetrics estimator in terms of the Ljung-Box Q-statistics in the outer products, having fewer series fail for dependance at lag 15 than the RiskMetrics model, usually by at least a factor of two. This provides strong evidence in favor of the DCC model over the RiskMetrics model for assets with a less dynamic correlation structure.

Overall, the DCC estimator performs very well. While it is not possible to directly compare the DCC and the constant conditional correlation multivariate models using LR statistics due to an unidentified parameter, even allowing 2 degrees of freedom, the corrected likelihood ratio statistics using S&P index data are always greater than 80 which corresponds to a p-value of less than .001%. The larger models estimated with the Dow Jones data also have very large likelihood ratio statistics, typically greater then 50. The standard deviations of portfolios were also fit much better using the DCC than the RiskMetrics model, resulting in no rejections of the null of a portfolio standard deviation of 1 for the value or equally weighted portfolios using the S&P index data or the Dow

Jones stocks. The RiskMetrics estimator performed especially poor using the minimum variance portfolios, possibly due to the correlation dynamics imposed by the choice of smoothing parameter. Figure 7 contains a graph of the correlations estimated using the DCC model and the RiskMetrics smoother for the 4 asset S&P 500 model. The RiskMetrics produced correlations are much more volatile and imply that the correlation matrix was nearly singular at certain points in time. Further, the DCC estimator produced multivariate standardized residuals in a confidence interval of one for all models with less than 10 assets, and never had a rejection rate greater than 14%, while the RiskMetrics model never produced multivariate standardized residuals in the confidence interval. Finally, the DCC estimator typically produces residuals with less serial dependence at lag 15 then the RiskMetrics estimator using a Ljung-Box Q statistics on the standardized outer-products of the residuals.

7 Multi-Step Ahead Forecasting

Forecasting covariances is a requirement of GARCH models. Most GARCH models provide an easy to implement method to generate r-step ahead forecasts. For instance, the r-step ahead forecast of a standard GARCH (1,1) is given by

$$h_{t+r} = \sum_{i=0}^{r-2} \omega(\alpha+\beta)^i + (\alpha+\beta)^{r-1} h_{t+1}$$

However, the DCC evolution process is a non-linear process, where

$$Q_{t+r} = (1-\alpha-\beta)\bar{Q} + \alpha(\epsilon_{t+r-1}\epsilon_{t+r-1}') + \beta Q_{t+r-1}$$

where $E_t[\epsilon_{t+r-1}\epsilon_{t+r-1}'] = E_t[R_{t+r-1}]$ and $R_{t+r} = {Q_{t+r}^*}^{-1} Q_{t+r} {Q_{t+r}^*}^{-1}$. Thus, the r-step ahead forecast of the correlation cannot be directly solved forward to provide a convenient method for forecasting. In examining methods to overcome this difficulty, two forecasts seem to be the most natural, each requiring a different set of approximations. The first technique proposed would be to generate the r-step ahead forecast of Q by making the approximation that $E_t[\epsilon_{t+1}\epsilon_{t+1}'] \approx Q_{t+1}$ for $i \in [1, \ldots, r]$. Using this approximation, we then have the r-step ahead forecast of Q_t is

$$E_t[Q_{t+r}] = \sum_{i=0}^{r-2} (1-\alpha-\beta)\bar{Q}(\alpha+\beta)^i + (\alpha+\beta)^{r-1} Q_{t+1}$$

and $R_{t+r} = {Q_{t+r}^*}^{-1} Q_{t+r} {Q_{t+r}^*}^{-1}$. An alternative approximation would be that $\bar{Q} \approx \bar{R}$ and that $E_t[Q_{t+1}] \approx E_t[R_{t+1}]$. Using this approximation, we can forecast R_{t+r} directly using the relationship

$$E_t[R_{t+r}] = \sum_{i=0}^{r-2}(1-\alpha-\beta)\bar{R}(\alpha+\beta)^i + (\alpha+\beta)^{r-1}R_{t+1}$$

In order to test which of these approximations performs better (if either), a Monte Carlo experiment was conducted. In the Monte Carlo, 1000 days of bivariate normal data were simulated using $\alpha = .01, \beta = .98$, and varying the unconditional correlation over the set $[-0.8, -0.5, 0, 0.35, 0.5, 0.65, 0.8, 0.9]$. At the 750^{th} observation, Q_{750}, r_{750}, and R_{750} were saved. Forecasts were then constructed for days 751 through 1000 using the two methods described above. In order to study these two forecasting techniques, we treated $\hat{\alpha}$ $\hat{\beta}$, and $\bar{R}$ as known, setting each parameter to the value used to simulate the data, to avoid the effects of parameter estimation uncertainty. It is also worth noting that this experiment would be exactly the same had we allowed for time varying volatility as the volatility errors would be exactly the same for either correlation forecasting technique. With these parameters, the half life of an innovation would be 68 days. Figure 8 contains a plot of the bias of the two forecasting methods. The first observation is that both of the forecasting techniques produce forecasts with a very small bias. This bias is generally toward 1 (or -1 for the series with negative unconditional correlation). The forecast produced by solving forward for Q_{t+r} (the dash-dotted line) was always closer to 1 (or -1) and consequently more biased. The method for solving R_{t+r} forward had better bias properties for almost all correlations and horizons. Also of interest is that both forecasts appear to be unbiased when the unconditional correlation is zero, and that they make essentially the same forecast when the unconditional correlation is zero.

Figure 9 contains a plot of the ratio of the MSFE of the forecast produced by solving R_t forward directly divided by the forecast produced by solving Q_t forward, then transforming the final Q_{t+r} to a correlation matrix. Neither forecast seems to produce uniformly smaller MSFE in the first fifty days. For certain unconditional correlations, the forecast which solves Q_t forward produces better MSFE (represented by a value greater than 1), yet this performance is hardly uniform. The ratio of the MSFEs is extremely close to 1, despite the lower bias produced by the R_t forecast as the square of the bias was typically at least three orders of magnitude smaller than the variance. Finally, figure 10 contains a fan plot of the density of the forecast errors. Two observations can be made from this figure. First, as the unconditional correlation approaches 1 (or -1), the variance of the forecast errors decreases. Second, the amount of skewness in the distribution of errors increases as the unconditional correlation increases. This is not surprising given that correlations are bounded between 1 and -1. While neither of these two techniques significantly outperformed the other, it would seem that a logical choice for forecasting would be the method that directly solves forward R_t. Not only is this method easier to implement, it appears that it also suffers less bias without

any detectable increase in variance.

8 Conclusion

This paper presents a class of estimators which join the simplicity and empirical success of univariate GARCH models with an easy to estimate and interpret dynamic correlation estimator. The two-step estimator is shown to be consistent and asymptotically normal and a consistent estimator of standard errors is provided. The standard errors for the correlation parameters depend on the cross partial derivatives of the second stage likelihood with respect to the first and second stage parameters in addition to the typical Bollerslev-Wooldridge robust standard errors. Ensuring the positive definiteness of the estimator is also shown to be easy to achieve as it simply requires using the same restrictions as univariate GARCH processes, and in the DCC(1,1) case could be treated as a scalar BEKK model for the correlation to allow estimation without lower bounds on the parameters. In addition to the two stage process, a fully efficient estimation procedure is outlined which involves a single Newton-Raphson step from the original consistent estimates.

A simple test is presented to test the null of constant correlation against an alternative of a dynamic conditional correlation-like structure. This test involves running a simple restricted VAR and can be easily estimated by OLS. A Monte Carlo study of this test has shown it to have good size and power against local alternatives. Specification testing provided a solid picture that this estimator exhibits better performance than the industry standard benchmark RiskMetrics for both large and small models. The real strength of the DCC estimation process is the flexibility provided in modelling the dynamics of the univariate volatility process. For instance, it would be easy to allow for asymmetric affects in volatility or to consider long memory volatility models, such as APARCH. In addition, the theory of DCC specification can be easy extended to include exogenous factors in the correlation model or alternative parameterizations which allow for heterogeneous dynamics in the correlations. In particular, it would be of interest to investigate whether increased correlation between assets is caused by either large volatilities or the correlations of other assets.

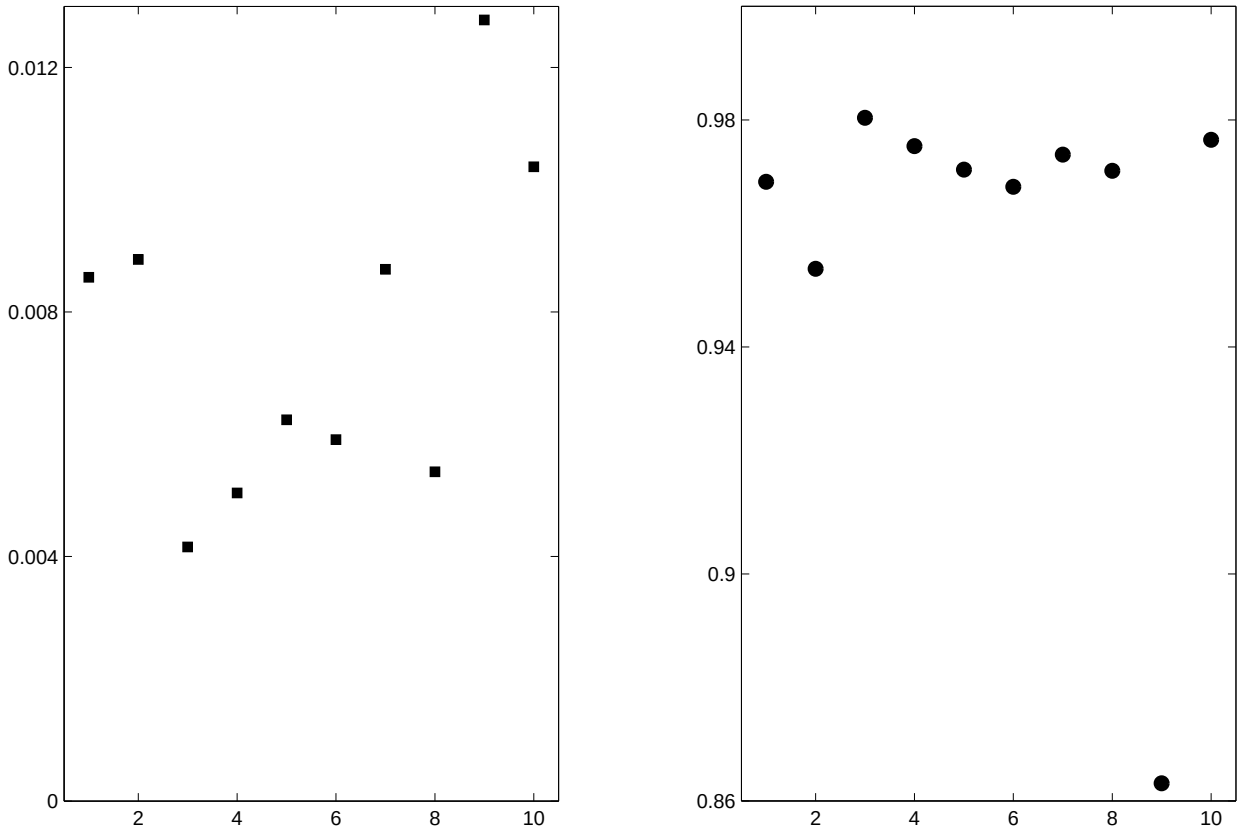

Figure 1: Graph of $\hat{\alpha}$ (left) and $\hat{\beta}$ for 10 portfolios randomly selected from the 100 S&P Indices. The models were strictly mean reverting in the sense that $\alpha + \beta < 1$ for all models.

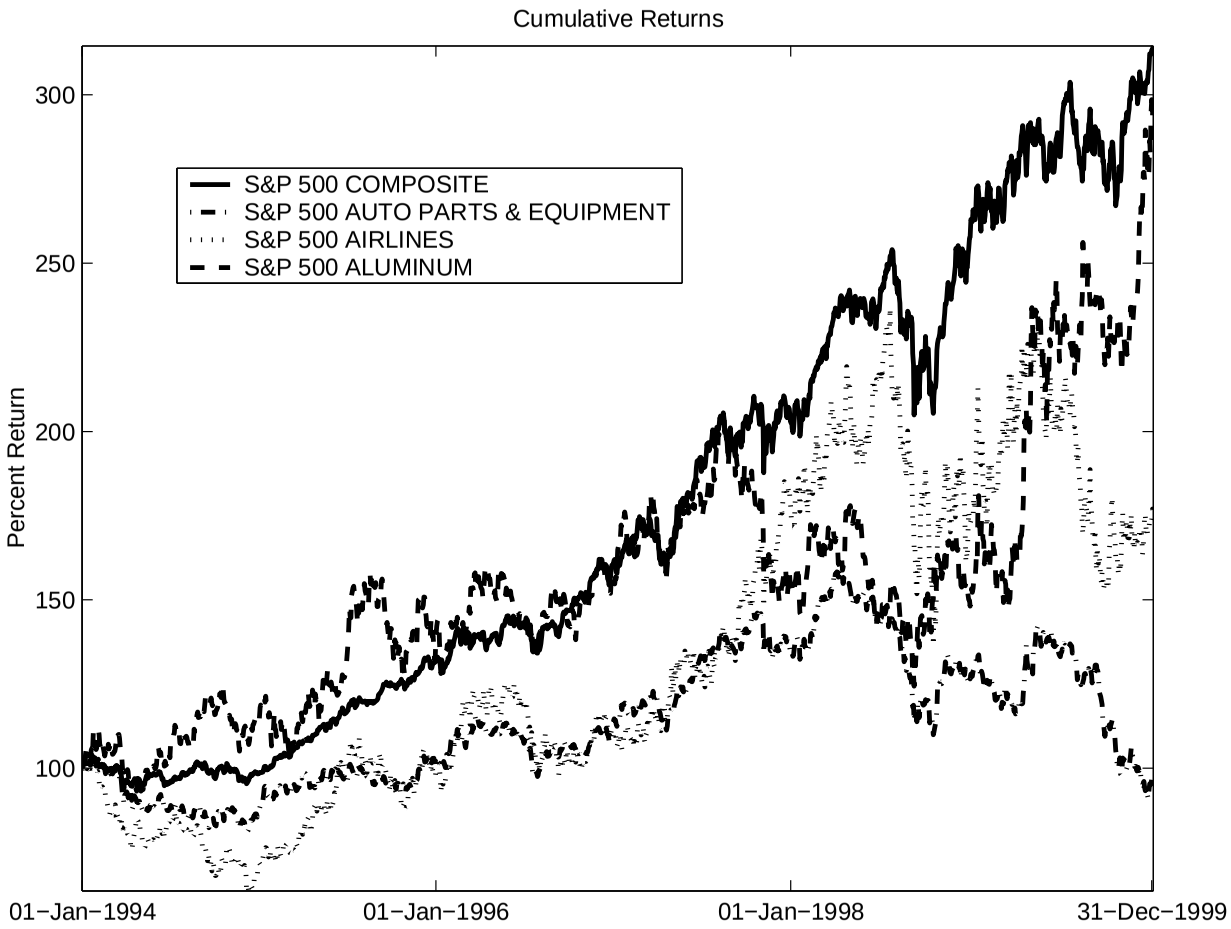

Figure 2: Graph showing the daily cumulative return of the 4 assets used in the S&P 500 indices estimation.

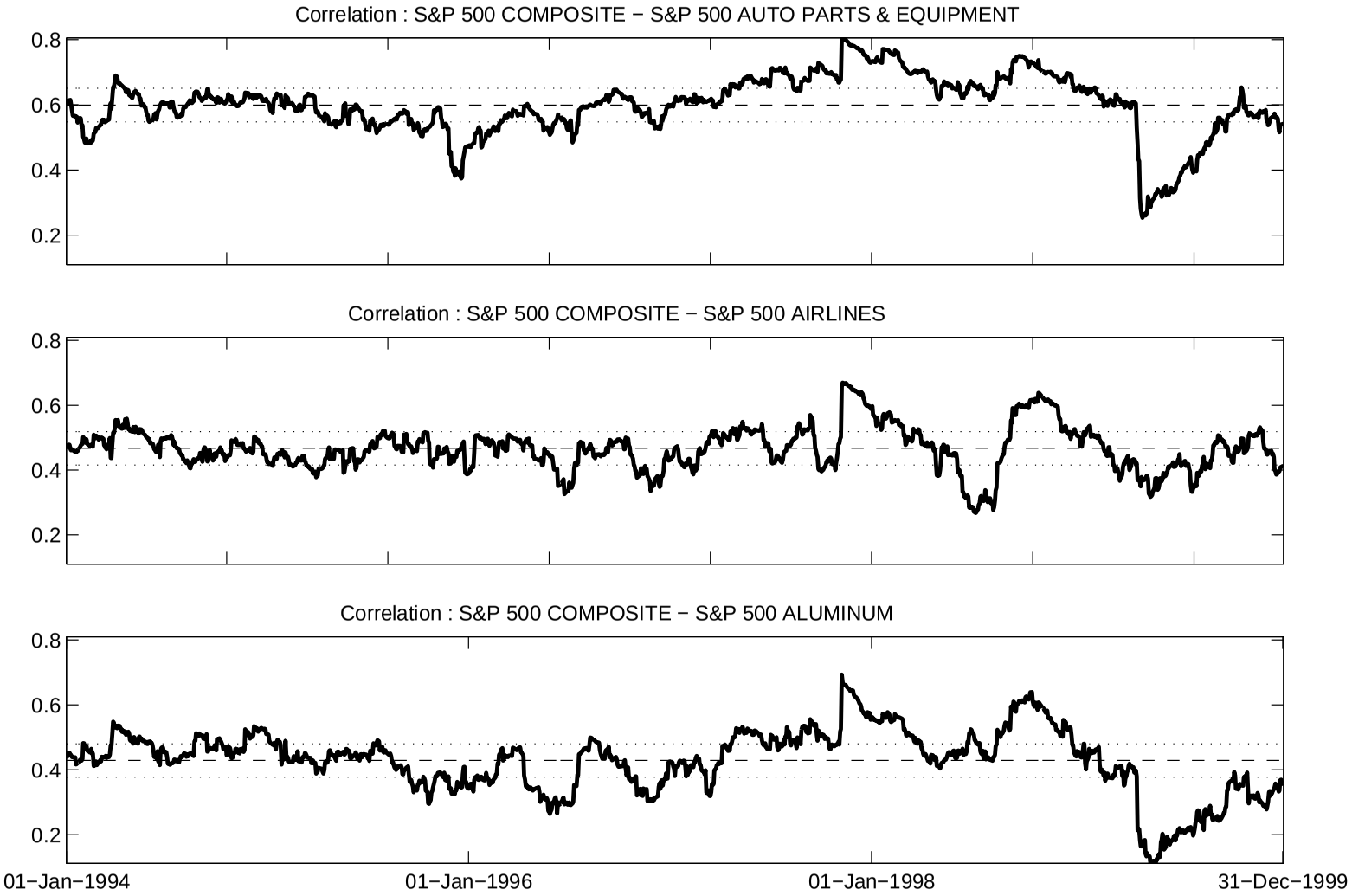

Figure 3: Graph is the dynamic conditional correlations between the S&P 500 Composite and 3 S&P Sector indices from the 4 asset model. The dashed line is the estimated constant conditional correlation (with confidence bands). The DCC correlation has a very wide range, and is often outside of the confidence bands resulting from the CCC MV-GARCH .

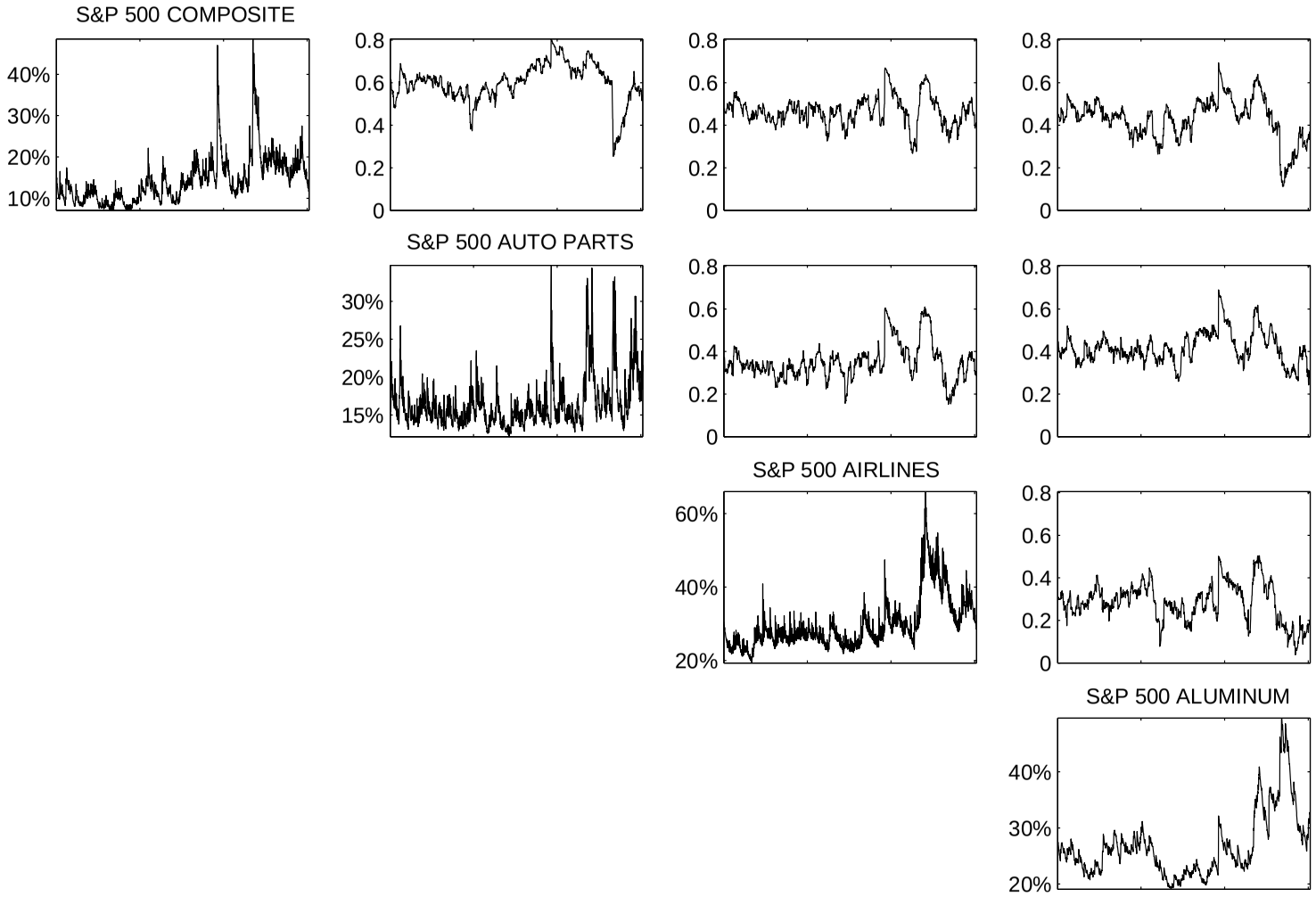

Figure 4: Graph of the time-varying annualized standard deviations (diagonal elements) and the estimated conditional correlation between the S&P sector indices in the four asset model. DCC-MVGARCH presents a natural method to examine the relationship between volatility and correlation.

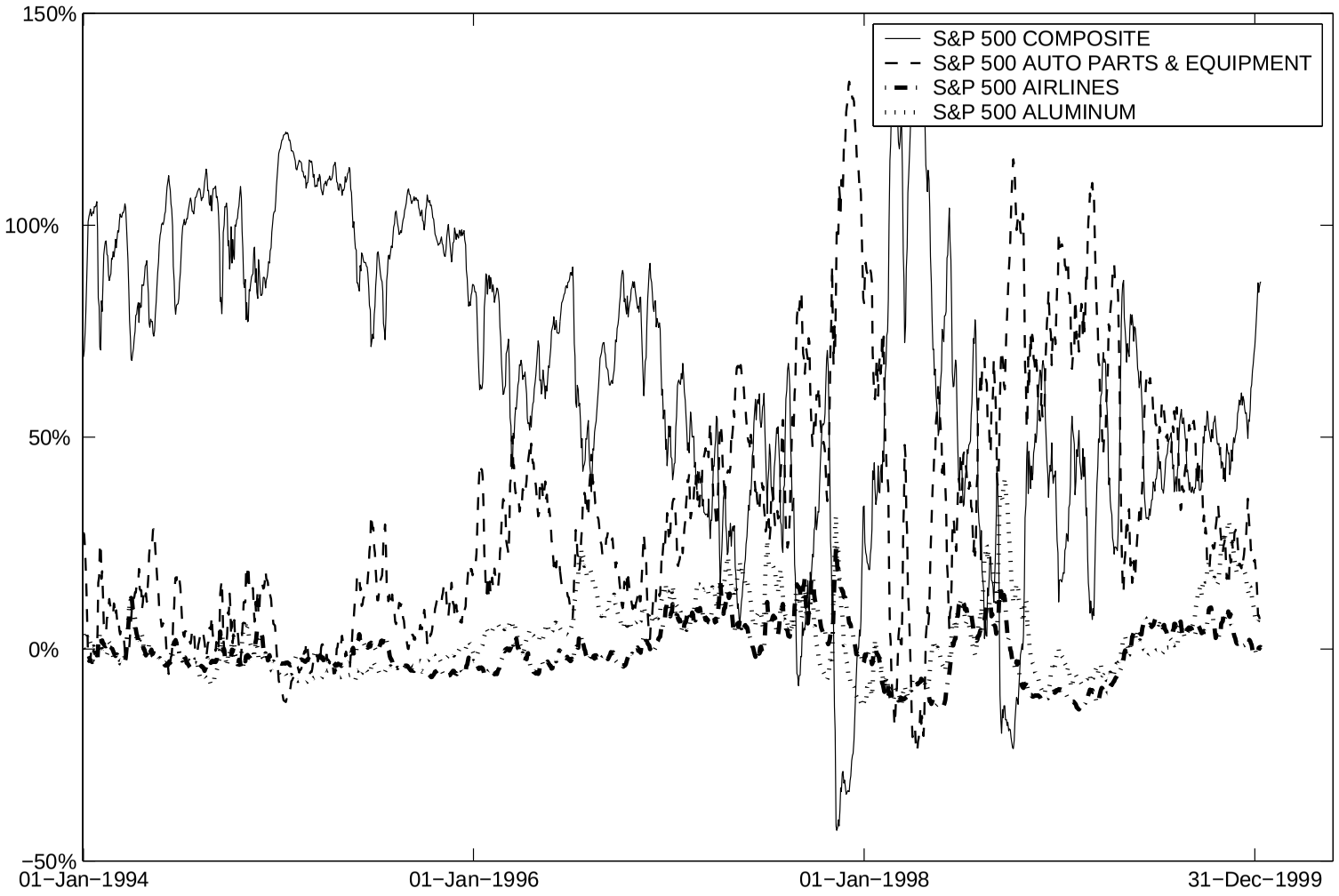

Figure 5: Graph of the minimum variance portfolio weights generates using the conditional variance covariance estimates from the 4 asset mean reverting DCC MV-GARCH model.

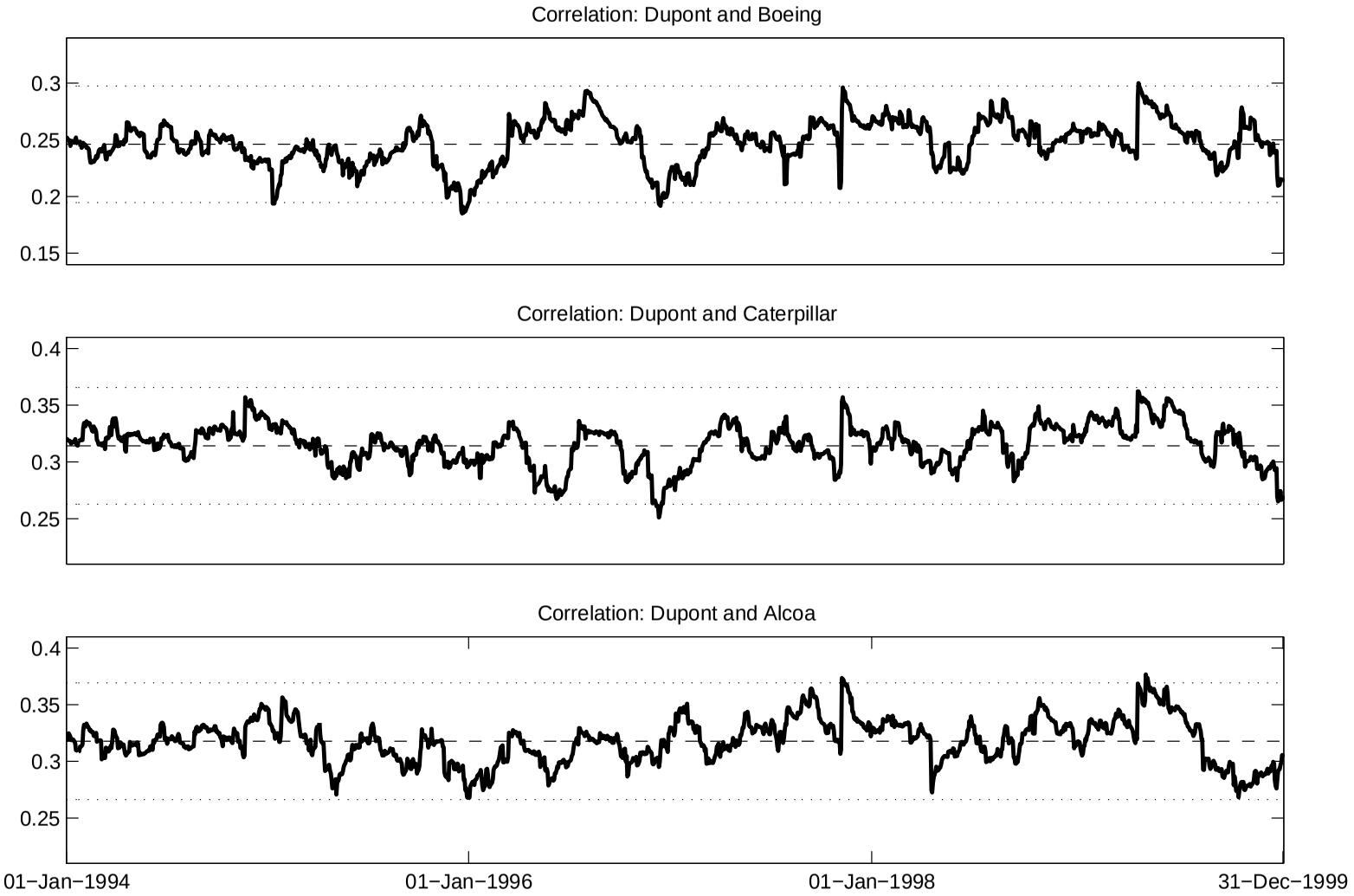

Figure 6: Graph is the dynamic conditional correlations between the Dupont and 3 DJIA indices from the 4 asset model. The dashed line is the estimated constant conditional correlation (with confidence bands). The DCC correlation, while less volatile than the estimated dynamic correlation of the 4 asset S&P model, spans then range of the CCC confidence interval and is outside this interval for each series at one point in time.

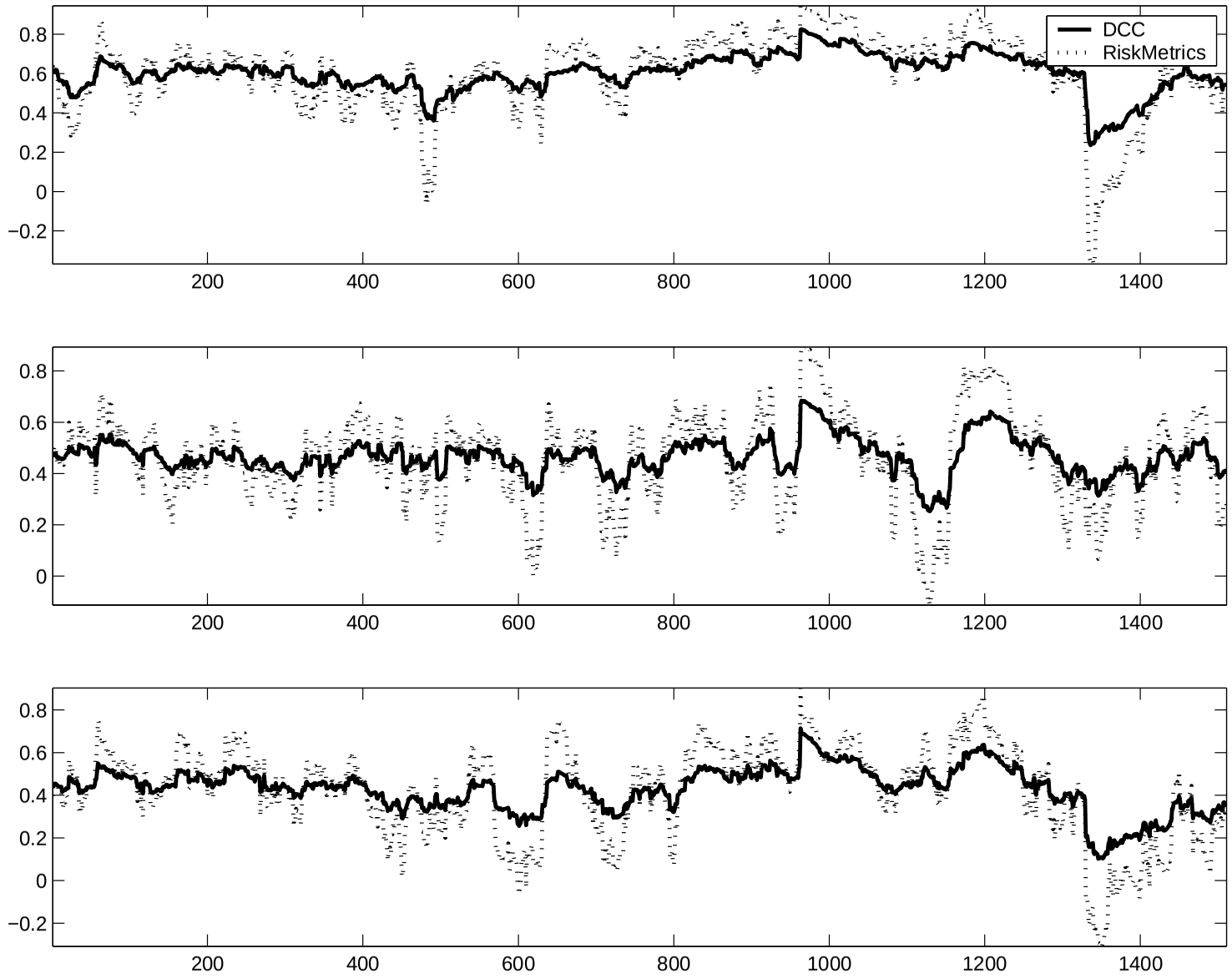

Figure 7: Graph of the conditional correlations between the S&P 500 Composite and 3 sector indices produced using the DCC model and the RiskMetrics smoother. The dynamics are similar, however the RiskMetrics smoother produces much more volatile conditional correlations with a number of near non-singularities($\hat{\rho}_t > .9$)

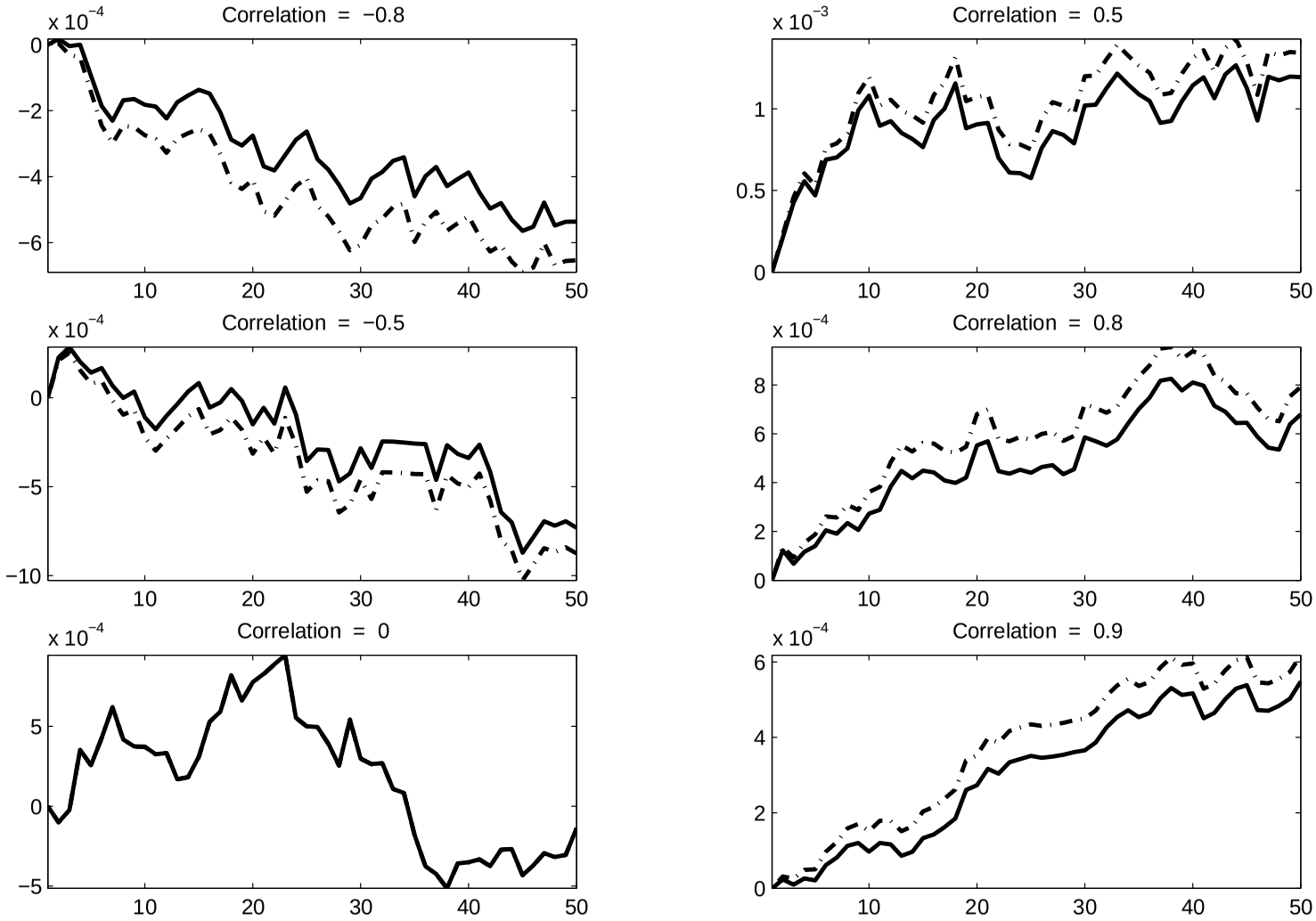

Figure 8: Graph of the bias of forecast of the conditional correlations generated by solving forward R_t (solid line) and Q_t (dot dashed line). While both are biased towards one (or -1 for negative unconditional correlation), the forecast generated by solving forward R_t is always closer to 0.

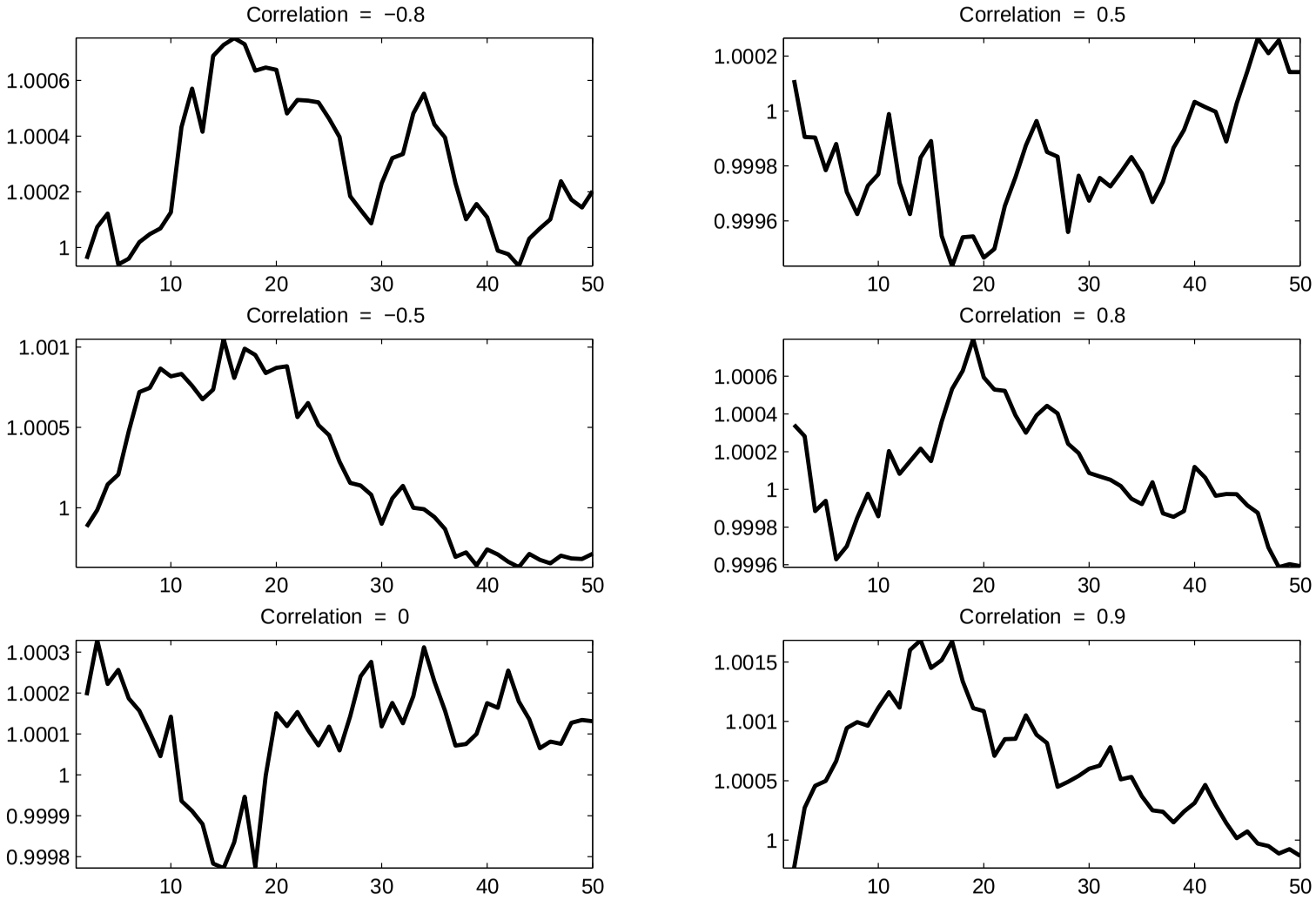

Figure 9: Graph of the ratio of the mean squared forecast error of the forecast generated by solving R_t forward to the mean squared forecast error of the forecast generated by solving Q_t forward. Values less than one indicate relatively good performance of the R_t generated forecast.

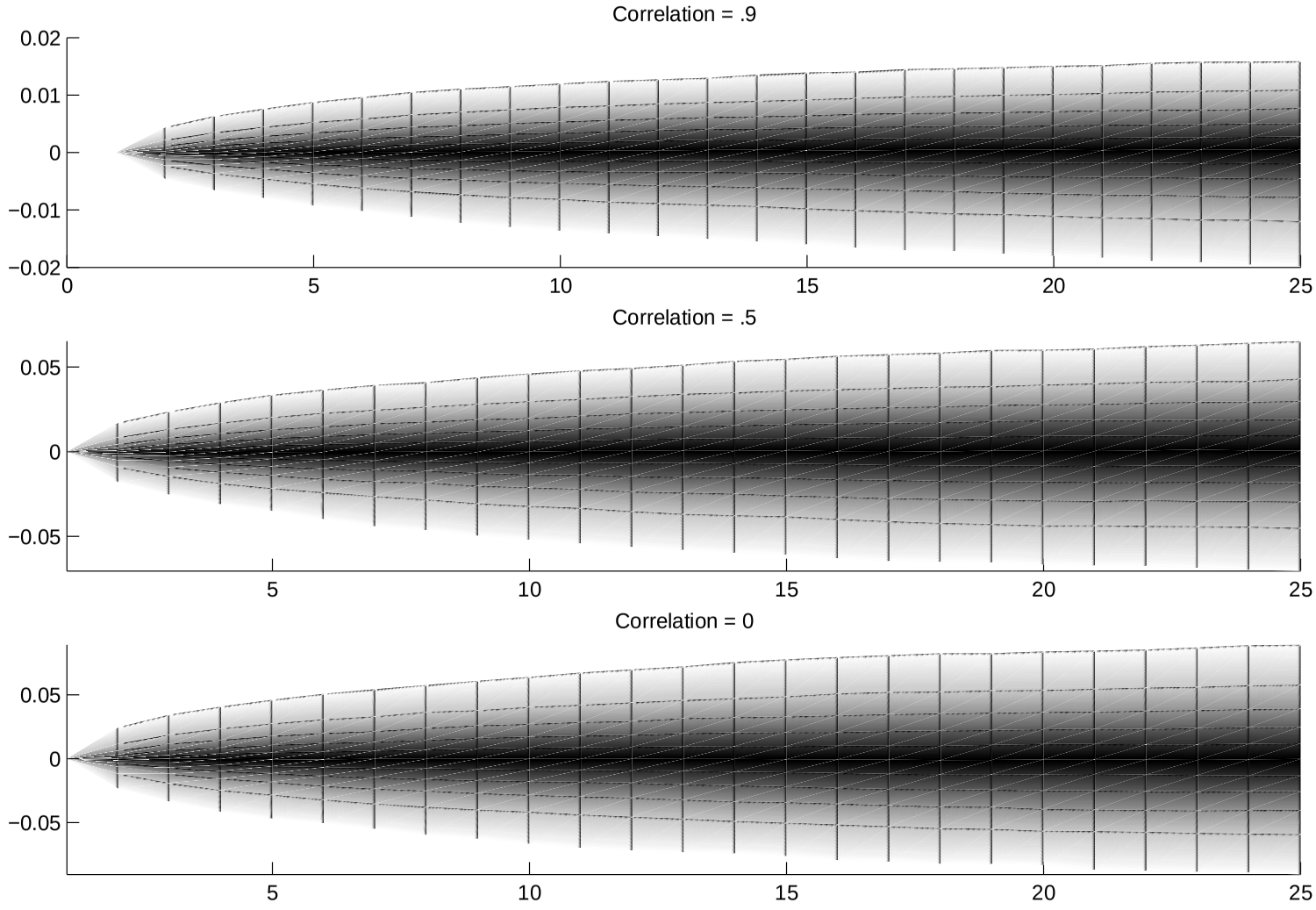

Figure 10: Fan plots of the density of forecast errors for up to 50 days. .

No. of Assets	$\rho = 0.3$	$\rho = 0.5$	$\rho = 0.8$	ρ from S&P indices
2	0.0312	0.0274	0.0290	0.0274
3	0.0268	0.0286	0.0212	0.0274
4	0.0290	0.0290	0.0298	0.0292
5	0.0280	0.0276	0.0278	0.0296
10	0.0308	0.0292	0.0444	0.0308

Table 1: Size under correct specification

DGP Model	2 assets	3 assets	4 assets	5 assets	10 assets
0.0266	0.0310	0.0262	0.0242	0.0284	
0.0230	0.0290	0.0250	0.0258	0.0268	
0.0264	0.0308	0.0266	0.0218	0.0310	

Table 2: Size under conditional volatility misspecification

	0.90	0.91	0.92	0.93	0.94	0.95	0.96	0.97	0.98	0.99
0.005	0.4230	0.4033	0.4144	0.4424	0.4073	0.4127	0.4094	0.4000	0.4197	0.4725
0.01	0.4879	0.4525	0.4627	0.4657	0.4547	0.4789	0.4523	0.4781	0.4925	
0.02	0.5619	0.6096	0.5690	0.6356	0.5951	0.6150	0.6553	0.7298		
0.03	0.7367	0.7571	0.7710	0.7957	0.8474	0.8794	0.9282			
0.04	0.8952	0.8854	0.9216	0.9437	0.9705	0.9843				
0.05	0.9686	0.9858	0.9819	0.9925	0.9998					
0.06	0.9972	0.9992	0.9982	0.9999						
0.07	0.9971	0.9999	0.9997							
0.08	1.0000	1.0000								
0.09	1.0000									

Table 3: Power of the test with 2 asset portfolios as α (across) and β (down) vary.

	0.90	0.91	0.92	0.93	0.94	0.95	0.96	0.97	0.98	0.99
0.005	0.4876	0.4991	0.4717	0.4675	0.4738	0.4894	0.5195	0.5213	0.4908	0.5144
0.01	0.6832	0.7103	0.7139	0.7498	0.7383	0.7489	0.7515	0.7591	0.8714	
0.02	0.9727	0.9862	0.9866	0.9930	0.9953	0.9973	0.9982	0.9999		
0.03	0.9998	0.9999	1.0000	1.0000	1.0000	1.0000	1.0000			
0.04	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000				
0.05	1.0000	1.0000	1.0000	1.0000	1.0000					
0.06	1.0000	1.0000	1.0000	1.0000						
0.07	1.0000	1.0000	1.0000							
0.08	1.0000	1.0000								
0.09	1.0000									

Table 4: Power of the test with 5 asset portfolios as α (across) and β (down) vary.

	0.90	0.91	0.92	0.93	0.94	0.95	0.96	0.97	0.98	0.99
0.005	0.6570	0.6413	0.6944	0.6862	0.7065	0.7440	0.7568	0.7690	0.7464	0.7753
0.01	0.9761	0.9736	0.9848	0.9810	0.9865	0.9869	0.9972	0.9946	0.9988	
0.02	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000		
0.03	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000			
0.04	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000				
0.05	1.0000	1.0000	1.0000	1.0000	1.0000					
0.06	1.0000	1.0000	1.0000	1.0000						
0.07	1.0000	1.0000	1.0000							
0.08	1.0000	1.0000								
0.09	1.0000									

Table 5: Power of the test with 10 asset portfolios as α (across) and β (down) vary.

			$H_0 : \alpha = 1 - \beta$			$H_0 : \lambda = 0$	
No.of Assets	$\hat{\alpha}$	$\hat{\beta}$	χ^2 value	P-value	$\hat{\lambda}$	χ^2 value	P-value
2	0.0335	0.9559	10.7490	0.0010	0.0305	90.1908	0.0000
	(2.96)	(60.20)			(3.5079)		
3	0.0266	0.9531	30.9750	0.0000	0.0198	61.9264	0.0000
	(2.34)	(38.13)			(4.4168)		
4	0.0163	0.9688	50.5591	0.0000	0.0144	43.3558	0.0000
	(4.33)	(112.83)			(5.5370)		
5	0.0123	0.9719	81.6736	0.0000	0.0108	3.7207	0.0537
	(6.09)	(168.13)			(5.0889)		
6	0.0103	0.9732	83.8247	0.0000	0.0000	0.0000	1.0000
	(6.34)	(172.13)			(0.0000)		
7	0.0101	0.9746	125.6139	0.0000	0.0000	0.0000	1.0000
	(7.74)	(224.44)			(0.0000)		
8	0.0097	0.9718	122.1910	0.0000	0.0000	0.0000	1.0000
	(6.95)	(164.23)			(0.0000)		
9	0.0090	0.9688	115.1191	0.0000	0.0000	0.0000	1.0000
	(5.78)	(116.99)			(0.0000)		
10	0.0255	0.9410	832.5525	0.0000	0.0000	0.0000	1.0000
	(3.26)	(28.85)			(0.0000)		
15	0.0176	0.9539	1012.6965	0.0000	0.0000	0.0000	1.0000
	(3.70)	(43.10)			(0.0000)		
20	0.0133	0.9662	1246.8991	0.0000	0.0000	0.0000	1.0000
	(3.51)	(57.48)			(0.0000)		
25	0.0097	0.9696	1099.0391	0.0000	0.0000	0.0000	1.0000
	(3.27)	(60.69)			(0.0000)		
50	0.0072	0.9643	1644.6180	0.0000	0.0000	0.0000	1.0000
	(5.56)	(85.47)			(0.0000)		
75	0.0052	0.9597	1571.7479	0.0000	0.0000	0.0000	1.0000
	(9.85)	(98.65)			(0.0000)		
100	0.0049	0.9497	2100.8853	0.0000	0.0000	0.0000	1.0000
	(14.58)	(154.53)			(0.0000)		

Table 6: Parameters Estimated on the S&P 500 Indices. The numbers in parentheses are robust T-statistics. The leftmost χ^2 value is for the null of integrated DCC against an alternative of dynamic conditional correlation, while the rightmost is for the null of constant conditional correlation against an alternative of a integrated dynamic conditional correlation.

			$H_0 : \alpha = 1 - \beta$			$H_0 : \lambda = 0$	
No.of Assets	$\hat{\alpha}$	$\hat{\beta}$	χ^2 value	P-value	$\hat{\lambda}$	χ^2 value	P-value
2	0.0104	0.9642	2.9885	0.0839	0.0000	0.0000	1.0000
	(1.27)	(43.58)			(0.00)		
3	0.0038	0.9679	1.2988	0.2544	0.0000	0.0000	1.0000
	(0.90)	(44.96)			(0.00)		
4	0.0048	0.9627	3.6213	0.0570	0.0000	0.0000	1.0000
	(1.70)	(81.11)			(0.00)		
5	0.0042	0.9672	5.5486	0.0185	0.0000	0.0000	1.0000
	(2.14)	(95.01)			(0.00)		
6	0.0043	0.9639	7.7226	0.0055	0.0000	0.0000	1.0000
	(2.55)	(65.33)			(0.00)		
7	0.0061	0.9690	26.7969	0.0000	0.0000	0.0000	1.0000
	(3.69)	(68.47)			(0.00)		
8	0.0061	0.9778	51.5646	0.0000	0.0000	0.0000	1.0000
	(4.34)	(61.85)			(0.00)		
9	0.0061	0.9768	63.2891	0.0000	0.0000	0.0000	1.0000
	(5.15)	(58.43)			(0.00)		
10	0.0066	0.9770	93.3452	0.0000	0.0000	0.0000	1.0000
	(5.44)	(55.87)			(0.00)		
15	0.0048	0.9690	70.3416	0.0000	0.0000	0.0000	1.0000
	(2.67)	(54.66)			(0.00)		
20	0.0044	0.9626	84.9840	0.0000	0.0000	0.0000	1.0000
	(4.45)	(76.16)			(0.00)		
25	0.0047	0.9384	92.5593	0.0000	0.0000	0.0000	1.0000
	(5.12)	(60.12)			(0.00)		
30	0.0043	0.9118	81.6511	0.0000	0.0000	0.0000	1.0000
	(4.74)	(26.97)			(0.00)		
31	0.0059	0.9441	252.2591	0.0000	0.0000	0.0000	1.0000
	(8.62)	(88.50)			(0.00)		

Table 7: Parameters Estimated on the Dow Jones Industrial Average Stocks. The numbers in parentheses are robust T-statistics. The leftmost χ^2 value is for the null of integrated DCC against an alternative of dynamic conditional correlation, while the rightmost is for the null of constant conditional correlation against an alternative of a integrated dynamic conditional correlation.

	S & P assets			Dow assets		
No.of Assets	DCC (2,1)	DCC (2,2)	DCC (3,2)	DCC (2,1)	DCC (2,2)	DCC (3,2)
2	1.4654	4.7682	11.5884*	0.0073	2.7891	6.5251
3	0.5825	5.3769	13.4353*	0.0073	0.4521	1.1840
4	1.5316	4.8220	7.8573	0.4276	0.3018	2.5090
5	0.4124	5.0408	17.0898*	0.0000	0.8375	4.2364
6	0.1304	3.6116	9.9360*	0.0000	0.3382	3.1993
7	0.0936	2.9785	15.6030*	0.0000	2.4545	5.9331
8	0.0000	0.1521	9.3299*	0.0000	1.5234	3.4858
9	0.0000	0.2382	4.8142	0.0000	3.0485	5.3958
10	0.0000	0.0169	18.8947*	0.0000	1.1618	5.0438
15	0.0000	2.0201	11.9072*	0.0000	0.5200	6.7282
20	0.0000	8.2936	19.5614*	0.0000	0.4654	3.3290
25	0.0000	10.6189*	21.2871*	0.0086	0.0000	5.2352
30	-	-	-	0.0298	0.4725	9.6378
31	-	-	-	0.0000	1.3408	9.7959

Table 8: Table of likelihood ratio statistics for alternative forms of the DCC estimator.

No. of Assets	DCC MVP	RM MVP	DCC Equal	RM Equal	DCC Value	RM Value
2	1.0271*	1.0917**	0.9967	1.0509*	0.9930	1.0478*
3	1.0342*	1.1238**	0.9908	1.0421*	0.9888	1.0409*
4	1.0386*	1.1600**	0.9899	1.0431*	0.9881	1.0415*
5	1.0424*	1.1949**	0.9881	1.0419*	0.9883	1.0408*
6	1.0504**	1.2311**	0.9855	1.0395*	0.9856	1.0385*
7	1.0663**	1.2757**	0.9856	1.0397*	0.9867	1.0397*
8	1.0726**	1.3188**	0.9852	1.0398*	0.9869	1.0401*
9	1.0753**	1.3573**	0.9851	1.0419*	0.9860	1.0415*
10	1.1299**	1.4261**	0.9868	1.0423*	0.9883	1.0422*
15	1.2707**	1.6540**	0.9842	1.0433*	0.9858	1.0435*
20	1.3380**	1.9362**	0.9830	1.0431*	0.9871	1.0423*
25	1.3627**	2.2994**	0.9767	1.0435*	0.9814	1.0425*
50	1.5459**	6.1302**	0.9778	1.0491**	0.9810	1.0476*
75	1.6471**	11.2570**	0.9781	1.0507**	0.9830	1.0494**
100	1.7794**	35.8968**	0.9777	1.0514**	0.9831	1.0497**

Table 9: Standard Deviation of different portfolios (VP, Equally weighted, and value weighted) using the S&P 500 Indices. (* indicates significantly different from 1 at the 5% level, ** indicates significantly different from 1 at the 1% level)

No. of Assets	DCC MVP	RM MVP	DCC Equal	RM Equal	DCC Value	RM Value
2	1.0113	1.1120**	0.9881	1.0531**	0.9876	1.0519**
3	1.0090	1.1502**	0.9901	1.0521**	0.9908	1.0496**
4	1.0121	1.1825**	0.9901	1.0549**	0.9921	1.0539**
5	1.0160	1.2139**	0.9885	1.0483*	0.9912	1.0454*
6	1.0234	1.2451**	0.9845	1.0476*	0.9884	1.0457*
7	1.0252	1.2891**	0.9846	1.0475*	0.9905	1.0466*
8	1.0281	1.3447**	0.9830	1.0459*	0.9883	1.0444*
9	1.0285	1.3790**	0.9837	1.0453*	0.9905	1.0437*
10	1.0355	1.4142**	0.9826	1.0435*	0.9896	1.0419*
15	1.0454*	1.6613**	0.9814	1.0459*	0.9901	1.0452*
20	1.0715**	1.9195**	0.9780	1.0454*	0.9879	1.0441*
25	1.0783**	2.2456**	0.9795	1.0464*	0.9932	1.0450*
30	1.0906**	2.6087**	0.9799	1.0448*	0.9945	1.0445*
31	1.2689**	2.7048**	0.9778	1.0449*	0.9917	1.0445*

Table 10: Standard Deviation of the minimum variance portfolio created from the estimated VCV of the DJIA 30 stocks (* indicates significantly different from 1 at the 5% level, ** indicates significantly different from 1 at the 1% level)

	% of HITs(Equal)		Probability	% of HITs (MVP)		Probability
No. of Assets	5%	1%	Independent.	5%	1%	Independent
2	4.83	1.52	0.0048	5.29	1.92	0.0360
3	4.57	1.26	0.0030	5.36	2.05	0.0005
4	4.04	1.26	0.0309	5.29	1.99	0.0032
5	4.04	1.39	0.0004	5.16	1.85	0.0595
6	4.17	1.19	0.0177	5.36	2.05	0.1359
7	3.90	1.32	0.0468	5.69	1.85	0.0040
8	4.10	1.32	0.1013	5.76	1.85	0.0179
9	4.37	1.39	0.0445	5.69	2.12	0.0242
10	4.57	1.39	0.0222	6.02	2.38	0.0238
15	4.63	1.26	0.0212	7.81	3.64	0.0304
20	4.30	1.26	0.0014	8.07	3.77	0.0004
25	4.43	1.26	0.0132	8.93	4.24	0.0003
50	4.17	1.52	0.0002	9.23	5.76	0.0035
75	4.30	1.59	0.0002	10.71	6.42	0.0004
100	4.04	1.65	0.0009	13.70	7.08	0.0030

Table 11: Value-at-Risk violations and HIT regression results for S&P 500 Indices for an equally weighted and a minimum variance portfolio

	% of HITs(Equal)		Probability	% of HITs (MVP)		Probability
No. of Assets	5%	1%	Independent.	5%	1%	Independent
2	3.97	0.79	0.5241	3.84	0.99	0.2213
3	4.30	1.26	0.2116	4.17	0.99	0.9035
4	4.43	0.93	0.1374	4.10	0.99	0.1078
5	3.90	1.19	0.0162	4.17	1.39	0.1658
6	3.84	1.13	0.6381	4.30	1.46	0.5429
7	3.90	1.26	0.1130	4.30	1.32	0.5460
8	3.84	1.32	0.4277	3.90	1.52	0.2341
9	3.97	1.26	0.3687	3.90	1.59	0.5338
10	4.10	1.26	0.7277	3.97	1.59	0.4796
15	4.24	1.19	0.0684	4.37	1.79	0.1703
20	4.10	1.26	0.0757	4.83	1.99	0.5754
25	4.70	1.32	0.0737	5.43	2.05	0.8767
30	4.63	1.19	0.0128	5.43	2.25	0.1494
31	4.63	1.26	0.0047	7.41	3.51	0.3155

Table 12: Value-at-Risk violations and HIT regression results for DJIA stocks for an equally weighted and a minimum variance portfolio

No. of Assets	% HIT at 5%		Pr. Independent		% std. resid. in CI		% failing LJ-Q	
	DCC	RiskMetrics	DCC	RiskMetrics	DCC	RiskMetrics	DCC	RiskMetrics
2	4.83	5.43	0.0048	0.0098	100	0	0.0000	0.0000
3	4.57	4.70	0.0030	0.0034	100	0	0.1667	0.3333
4	4.04	4.43	0.0309	0.0391	100	0	0.2000	0.3000
5	4.04	4.63	0.0004	0.2611	100	0	0.0667	0.2000
6	4.17	5.36	0.0177	0.0059	100	0	0.0476	0.0476
7	3.90	5.16	0.0468	0.0160	100	0	0.0714	0.2143
8	4.10	5.43	0.1013	0.0210	100	0	0.1111	0.1389
9	4.37	5.43	0.0445	0.0042	100	0	0.1333	0.2444
10	4.57	5.29	0.0222	0.0727	90	0	0.1818	0.3273
15	4.63	5.10	0.0212	0.0315	93	0	0.1750	0.2500
20	4.30	4.96	0.0014	0.0064	90	0	0.1714	0.2048
25	4.43	5.10	0.0132	0.0816	88	0	0.1508	0.2523
50	4.17	5.29	0.0002	0.0457	86	0	0.1490	0.3137
75	4.30	5.29	0.0002	0.0688	90	0	0.1312	0.4277
100	4.04	5.36	0.0009	0.0283	90	0	0.1628	0.4157

Table 13: Comparison of the DCC GARCH estimated covariances and RiskMetrics estimated covariances using the S&P 500 Sector Indices

No. of Assets	% HIT at 5%		Pr. Independent		% std. resid. in CI		% failing LJ-Q	
	DCC	RiskMetrics	DCC	RiskMetrics	DCC	RiskMetrics	DCC	RiskMetrics
2	3.97	4.63	0.5241	0.5788	100	0	0.0000	0.0000
3	4.30	5.03	0.2116	0.8779	100	0	0.1666	0.5000
4	4.43	5.03	0.1374	0.0525	100	0	0.2000	0.5000
5	3.90	4.43	0.0162	0.0305	100	0	0.2667	0.4000
6	3.84	4.63	0.6381	0.4369	100	0	0.1429	0.3333
7	3.90	4.43	0.1130	0.1628	100	0	0.1786	0.3214
8	3.84	4.37	0.4277	0.4865	100	0	0.1111	0.2500
9	3.97	4.43	0.3687	0.6460	100	0	0.1111	0.2222
10	4.10	4.70	0.7277	0.7494	100	0	0.0909	0.1818
15	4.24	4.57	0.0684	0.0847	100	0	0.0667	0.1500
20	4.10	4.77	0.0757	0.0252	100	0	0.0667	0.1429
25	4.70	4.70	0.0737	0.1274	100	0	0.0646	0.1600
30	4.63	4.43	0.0128	0.4364	100	0	0.0495	0.1441
31	4.63	4.43	0.0047	0.4367	96	0	0.0665	0.1472

Table 14: Comparison of the DCC GARCH estimated covariances and RiskMetrics estimated covariances using the Dow Jones Industrial Average Stocks

9 Appendix

Proof of Proposition 1:

Assume A is positive definite. Further, $B = A^{*-1\prime}AA^{*-1}$ where A^* is as defined in section 2. Since A is real, symmetric and positive definite, we know there exists a cholesky factorization $A = P'P$ where P is upper triangular. Rewriting

$$B = A^{*-1\prime}P'PA^{*-1} = (PA^{*-1})'(PA^{*-1})$$

Finally, we know that (PA^{*-1}) has full rank as it is the product of a diagonal matrix and a triangular matrix,both with non-zero diagonal elements, and will necessarily be a triangular matrix with nonzero diagonal elements. Using a familiar result form Linear Algebra, that $X'X$ is positive definite if and only if the X has full rank, thus B is positive definite. The proof assuming B is real, symmetric and positive definite follows directly replacing A with B and A^{*-1} with A^*.

□

Proof of Proposition 3:

$$QL_2(r_t) = -\frac{1}{2}\sum_{t=1}^{T}(k\log(2\pi) + 2\log|D_t| + \log(|R_t|) + (r_t'D_t^{-1}R_t^{-1}D_t^{-1}r_t)$$

Differentiating QL_2 with respect to the parameter vector $(\phi_1, \phi_2, \ldots, \phi_k)'$, we have

$$\frac{\partial QL_2}{\partial \phi} = \frac{\partial QL_2}{\partial H_t}\frac{\partial H_t}{\partial \phi}$$

where H_t is a k by 1 vector $(h_{1t}\ h_{2t}\ \ldots\ h_{kt})'$. Rewriting,

$$QL_2 = -\frac{1}{2}Tk\log(2\pi) - \frac{1}{2}\sum_{t=1}^{T}\log(h_{1t}) - \frac{1}{2}\sum_{t=1}^{T}\log(h_{2t}) + \ldots$$

$$-\frac{1}{2}\sum_{t=1}^{T}\log(h_{kt}) - \frac{1}{2}\sum_{t=1}^{T}(\log(|R_t|) + (r_t'D_t^{-1}R_t^{-1}D_t^{-1}r_t))$$

Making a change of variables $u_{it} = h_{it}^{-\frac{1}{2}}$, with U_t a column vector of u_{it}, the likelihood becomes

$$QL_2 = -\frac{1}{2}Tk\log(2\pi) - \frac{1}{2}\sum_{t=1}^{T}\log(u_{1t}^{-2}) - \frac{1}{2}\sum_{t=1}^{T}\log(u_{2t}^{-2}) + \ldots$$

$$-\frac{1}{2}\sum_{t=1}^{T}\log(u_{kt}^{-2})+\sum_{t=1}^{T}(U_t\tilde{r_t})R_t^{-1}(U_t\tilde{r_t})$$

where $\tilde{r_t}$ is a diagonal matrix with i^{th} element r_{it}. Differentialting this with respect to U_t, we have

$$\frac{\partial QL_2}{\partial U_t}=2U_t^{-1}-2\tilde{r_t}R_t^{-1}\tilde{r_t}U_t$$

where U_t^{-1} denotes an element by element inverse of U_t. And differentiating with respect to H_t

$$\frac{\partial QL_2}{\partial U_t}\frac{\partial U_t}{\partial H_t}=2\begin{bmatrix} h_{1t}^{-\frac{3}{2}} & 0 & \dots & 0 \\ 0 & h_{1t}^{-\frac{3}{2}} & \dots & 0 \\ \vdots & \vdots & \vdots & \vdots \\ 0 & \dots & 0 & h_{1t}^{-\frac{3}{2}} \end{bmatrix}\begin{bmatrix} h_{1t}^{\frac{1}{2}} \\ h_{2t}^{\frac{1}{2}} \\ \vdots \\ h_{kt}^{\frac{1}{2}} \end{bmatrix}-2\begin{bmatrix} h_{1t}^{-\frac{3}{2}} & 0 & \dots & 0 \\ 0 & h_{1t}^{-\frac{3}{2}} & \dots & 0 \\ \vdots & \vdots & \vdots & \vdots \\ 0 & \dots & 0 & h_{1t}^{-\frac{3}{2}} \end{bmatrix}\tilde{r_t}R_t^{-1}\tilde{r_t}\begin{bmatrix} h_{1t}^{-\frac{1}{2}} \\ h_{2t}^{-\frac{1}{2}} \\ \vdots \\ h_{kt}^{-\frac{1}{2}} \end{bmatrix}$$

And since we have $E_{t-1}(r_{it}^2)=h_{it}$, we can rewrite as

$$\frac{\partial QL_2}{\partial U_t}\frac{\partial U_t}{\partial H_t}=2\begin{bmatrix} h_{1t}^{-1} \\ h_{2t}^{-1} \\ \vdots \\ h_{kt}^{-1} \end{bmatrix}-2\begin{bmatrix} h_{1t}^{-1} & 0 & \dots & 0 \\ 0 & h_{1t}^{-1} & \dots & 0 \\ \vdots & \vdots & \vdots & \vdots \\ 0 & \dots & 0 & h_{1t}^{-1} \end{bmatrix}\tilde{e_t}R_t^{-1}e_t$$

by dividing the r_t by h_t. Now that R_t is an estimate of the correlation between standardized residuals, and from scoring conditions on the estimate of R_t, we have $E_{t-1}(\tilde{e_t}R_t^{-1}e_t)=\iota$, and finally differentiating with respect to θ, we have

$$\frac{\partial QL_2}{\partial U_t}\frac{\partial U_t}{\partial H_t}\frac{\partial H_t}{\partial \theta}=2\left(\begin{bmatrix} h_{1t}^{-1} \\ h_{2t}^{-1} \\ \vdots \\ h_{kt}^{-1} \end{bmatrix}-\begin{bmatrix} h_{1t}^{-1} & 0 & \dots & 0 \\ 0 & h_{2t}^{-1} & \dots & 0 \\ \vdots & \vdots & \vdots & \vdots \\ 0 & \dots & 0 & h_{kt}^{-1} \end{bmatrix}\iota\right)\frac{\partial H_t}{\partial \theta}=0$$

References

[1] C. Alexander, *A Primer on the Orthogonal GARCH Model*. University of Reading, February 2000.

[2] D. W. K. Andrews and W. Ploberger, *Optimal Tests when a Nuisance Parameter is Present only under the Alternative*, Econometrica, 62 (1994), pp. 1383–1414.

[3] T. Bollerslev, *Generalized Autoregressive Conditional Heteroskedasticity*, Journal of Econometrics, 31 (1986), pp. 307–27.

[4] ——, *Modelling the Coherence in Short Run Nominal Exchange Rates: A Multivariate generalized ARCH Model*, The Review of Economics and Statistics, 72 (1990), pp. 498–505.

[5] T. Bollerslev, R. F. Engle, and D. B. Nelson, *ARCH Models*, in Handbook of Econometrics, vol. 4, Elsevier North Holland, 1994.

[6] T. Bollerslev, R. F. Engle, and J. M. wooldridge, *A Capital Asset Pricing Model with Time-Varying Covariances*, Journal of Political Economy, 96 (1988), pp. 116–131.

[7] T. Bollerslev and J. M. wooldridge, *Quasi-maximum Likelihood Estimation and Inference in Dynamic Models with Time-Varying Covariances*, Econometric Reviews, 11 (1992), pp. 143–172.

[8] J. Y. Camble, A. W. Lo, and A. C. MacKinlay, *The Econometrics of Financial Markets*, Princeton University Press, Princeton, New Jersey, 1997.

[9] G. Casella and R. L. Berger, *Statistical Inference*, Duxbury Press, Belmont, California, 1990.

[10] J. Conlisk, *Matrix Algebra for Economics Lecture Notes*, 1998.

[11] A. Demos and E. Sentana, *Testing for GARCH Effects: A One-sided Approach*, Journal of Econometrics, 86 (1998), pp. 97–127.

[12] R. F. Engle, *Dynamic Conditional Correlation - A Simple Class of Multivariate GARCH Models*. UCSD, May 2001.

[13] R. F. Engle and K. F. Kroner, *Multivariate Simultaneous Generalized ARCH*, Econometric Theory, 11 (1995), pp. 122–150.

[14] R. F. Engle and S. Manganelli, *CaViaR: Conditional Autoregressive Value at Risk by Regression Quantiles*. University of California at San Diego Working Paper, July 1999.

[15] R. F. Engle, V. Ng, and M. Rothschild, *Asset Pricing with a Factor ARCH Covariance Structure: Empirical Estimates for Treasury Bills*, Journal of Econometrics, 45 (1990), pp. 213–237.

[16] R. V. Foutz and R. C. Srivastava, *The Performance of the Likelihood Ratio Test when the Model is Incorrect*, The Annals of Statistics, 5 (1977), pp. 1183–94.

[17] A. R. Gallant and H. White, *A Unified Theory for Estimation and Inference for Nonlinear Dynamic Models*, Basil Blackwell, New York, 1988.

[18] C. Gourieroux, A. Monfort, and A. Trognon, *Pseudo Maximum Likelihood Methods: Theory*, Econometrica, 52 (1984), pp. 681–700.

[19] A. Harvey, E. Ruiz, and N. Shephard, *Multivariate Stochastic Variance Models*, Review of Economic Studies, 61 (1994), pp. 247–264.

[20] A. Hirsch, *Theory and Practice of ARCH Modelling*. Deutsch Bank, November 1997.

[21] T. Jeantheau, *Strong Consistency of Estimators for Multivariate GARCH Models*, Econometric Theory, 14 (1998), pp. 70–86.

[22] O. Ledoit, *Improved Estimation of the Covariance Matrix of Stock Returns with an Application to Portfolio Selection*. Anderson Graduate School of Management, March 1997.

[23] J. H. H. Lee, *A Lagrange Multiplier Test for GARCH Models*, Economics Letters, 37 (1991), pp. 265–272.

[24] W. K. Li and T. Mak, *On the Squared Residual Autocorrelations in Non-linear time series with Conditional Heteroskedasticity*, Journal of Time Series Analysis, 15 (1994), pp. 627–636.

[25] K.-Y. Liang and S. G. Self, *On the Asymptotic Behavior of the Pseudolikelihood Ratio Test Statistic*, Journal of the Royal Statistical Society B, 58 (1996), pp. 785–96.

[26] G. F. Loudon, W. H. Watt, and P. K. Yadav, *An Empirical Analysis of Alternative Parametric Arch Models*, Journal of Applied Econometrics, 15 (2000), pp. 117–136.

[27] J. R. Magnus and H. Neudecker, *Matrix Differential Calculus with Application in Statistics and Econometrics*, John Wiley and Sons, New York, 1988. Wiley Series in Probability and Mathematical Statistics.

[28] D. Nelson and C. Cao, *Inequality Constraints in the univariate GARCH model*, Journal of Business and Economic Statistics, 10 (1992), pp. 229–35.

[29] W. K. Newey and D. McFadden, *Large Sample estimation and Hypothesis Testing*, in Handbook of Econometrics, vol. 4, Elsevier North Holland, 1994.

[30] A. Pagan, *Two Stage and Related Estimators and Their Applications*, Review of Economic Studies, 53 (1986), pp. 517–538.

[31] Y. K. Tse, *A Test for Constant Correlations in a Multivariate GARCH Model*, Journal of Econometric, 98 (2000), pp. 107–127.

[32] Y. K. Tse and A. K. Tsui, *A Multivariate GARCH Model with Time-Varying Correlations.* National University of Singapore, December 1998.

[33] A. K. Tsui and Q. Yu, *Constant Conditional Correlation in a Bivariate GARCH model: Evidence from the stock market in China*, Mathematics and Computers in Simulation, 48 (1999), pp. 503–509.

[34] H. White, *Estimation, Inference, and Specification Analysis*, Cambridge University Press, Cambridge, 1996. Econometric Society Monographs.