# Multi-step estimation of Multivariate GARCH models

Kevin Sheppard*
University of California
at San Diego

October 25, 2003

**Abstract**

Estimation of large time varying covariance matrices has proven difficult since the original multivariate volatility models were introduced. In this paper, we develop the empirical properties of a new class of MV-GARCH models capable of estimating large time-varying covariance matrices. We show that the problem of multivariate conditional variance estimation can be reduced to a series of univariate GARCH processes plus an additional conditional correlation estimator. We use the model to estimate a conditional covariance of up to 100 assets using S&P 500 Sector Indies and Dow Jones Industrial Average stocks. This new estimator demonstrates very strong performance especially considering the estimators easy of implementation.

## 1 Introduction

While multivariate GARCH models have found considerable empirical success, the problem of estimation of large time-varying matrices has restricted researchers to estimating models with limited scope or considerable restrictions. Large time-varying covariance matrices can be found in models of the term structure of treasuries or commodities, large vector auto-regressions, or in portfolio management and optimization. In this paper we describe a model which can be used to estimate extremely large time-varying covariance matrices and describe theoretical properties of the Dynamic Conditional Correlation(DCC) Multi-Variate GARCH model. This class of MV-GARCH models differs from other specifications in that univariate GARCH models are estimated for each asset series, then, using the standardized residuals resulting from the first step, a time varying correlation matrix is generated using a simple specification. This parameterization preserves the simple interpretation of a univariate GARCH models with an easy to compute correlation step. It will also be shown in that paper that the standard errors computed using the Bollerslev-Wooldridge method for each univariate GARCH model remain consistent, and only the error estimates 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 provides full generality. 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 forces only lagged own effects and cross products. The diagonal specification requires $O(k^2)$ parameters be estimated, namely $\frac{3}{2}k(k+1)$ need to be estimated for the (1,1) model. The diagonal specification allows for a relatively straightforward interpretation, as each series has the familiar standard GARCH specification.

In the constant conditional correlation multivariate GARCH specification proposed by Bollerslev(1990), the correlation matrix can be estimated

*Mailing address: 9500 Gilman Drive 0505, Department of Economics, UCSD, La Jolla, CA 92037 email: kksheppard@ucsd.edu

separately from the univariate GARCH formulations used for each series using the standard multivariate normal maximum likelihood estimator correlation estimator on the standardized residuals. The assumption of constant correlation makes estimating a large model feasible and ensures that the estimator is positive definite, which is difficult to achieve using the *vech* formulation. However, the constant correlation estimator, as proposed, does not describe how to construct consistent standard errors using the separate estimation process. Bollerslev finds the notion of constant correlation plausible, yet recent papers by Tsui and Yu(1999) have 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 LM test.

The BEKK formulation, proposed in Engle and Kroner(1995), developed a general quadratic form for the covariance equation which eliminated the positive definiteness problems of the original *vech* model. The authors developed the conditions necessary BEKK model to be fully general as well as for the BEKK and the vech representations to be equivalent. In order for the BEKK model to be fully general, the number of parameters needing to be estimated is $O(k^4)$, but the standard BEKK estimation will involve $O(k^2)$ parameters. Other more tractable formulations of the BEKK model include diagonal and scalar where the parameters are restricted to be either diagonal matrices or to be scalars. In addition to the large number of parameters needing to be estimated for the general form, the exact interpretation and impact of the individual coefficients is difficult to discern.

Recently, Alexander(2000) has demonstrated the use of factor GARCH models as first outlined in Diebold and Nerlove(1989) and Engle and Rodrigues(1989) for estimation of large covariance matrices. Factor or Orthogonal MV-GARCH models provide a method for estimating any variance covariance matrix using only univariate GARCH models. Alexander shows how a limited number of factors can account for 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 individual equities.

Engle(2000) proposed a new class of estimator that both preserves the easy of estimation of Bollerslev's constant correlation model yet allows for correlations to change over time. Dynamic Conditional Correlation MV-GARCH preserves the parsimony of the univariate GARCH models of individual assets with a simple GARCH like time varying correlation. Further, the number of parameters estimated using maximum likelihood is $O(N)$, 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 this 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 restricting enforcing a mean reversion to the unconditional correlation in the correlation estimator, the number of parameters needing to be simultaneously estimated os $O(k^2)$ 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 as well as discusses the estimation procedure used. Section three provides empirical results where systems with up to 100 assets are estimated and section seven concludes as well as outlines area of future research.

# 2 Model

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

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

and

$$H_t = D_t R_t D_t$$

where $D_t$ is the $k \times k$ diagonal matrix composed from univariate GARCH series 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:

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

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

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

for $i = 1, 2, \ldots, k$ with the usual GARCH restrictions for stationarity being imposed such as non-negativity of coefficients and $\sum_{p=1}^{P_i} \alpha_{ip} + \sum_{q=1}^{Q_i} \beta_{iq} < 1$[2]. 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. The proposed dynamic correlation structure is:

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

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

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

$$Q_t^{*-1} = \begin{bmatrix} \frac{1}{\sqrt{q_{11}}} & 0 & 0 & \ldots & 0 \\ 0 & \frac{1}{\sqrt{q_{22}}} & 0 & \ldots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & \ldots & \frac{1}{\sqrt{q_{kk}}} \end{bmatrix}$$

so that $Q_t^*$ is a diagonal matrix with the square root of the diagonal elements on the diagonal. . The typical element of $R_t$ will be of the form $\rho_{ijt} = \frac{q_{ijt}}{\sqrt{q_{ii}q_{jj}}}$. The immediate implication from this form is that $R_t$ will necessarily be a correlation matrix by Cauchy-Schwartz.[3]

# 3 Data and Estimation Results

The data used in this paper consist of the 30 Dow Jones Industrial Average stocks plus the average as well as 100 S&P 500 Sector Indices including the S&P 500 Composite. All DOW data were from January 1, 1990 to December 31, 1999 and were taken from Datastream. The S&P sector indices were from January 21 ,1991 when Standards and

[1] The standard errors of the model will depend on the choice of filtration(ARMA, demeaning), but exact form of the standard errors will be evident from the theory presented in the paper for zero mean residuals.

[2] Exact conditions for non-negativity of GARCH(p,q) processes can be can be found in Nelson and Cao(1992)

[3] A correlation matrix is defined as a real, symmetric positive definite matrix, with ones on the diagonal and all other terms less than or equal to one in absolute value.

Poor began reporting daily sector indices to December 31, 1999 and were also taken from Datastream. All days the market were closed were removed, with the number of days removed varying between seven and nine depending on the year[4]. After removing these days there were 2274 days for the S&P 500 indices and 2528 days dow the Dow Jones Industrial Average Stocks.

The model used in the for the empirical section was a simple DCC(1,1)-MVGARCH where each of the univariate GARCH models estimated was a GARCH(1,1)[5]. In addition, an integrated form of this model was estimated where $\lambda = \beta = 1 - \alpha$ was imposed. Table 1 summarizes the estimated $\hat{\alpha}$ and $\hat{\beta}$ of the mean reverting model for different numbers of assets using the S&P indices, as well as the statistic derived from a likelihood ratio test that tests $H_0 : \alpha = 0, \beta = 0$ and thus there was no dynamic correlation. T-stats are reported for each coefficient using the mispecification robust standard errors in parenthesis.

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.[6]. Table 1 also reports the estimated $\hat{\lambda}$ and reports the $\chi^2$ statistic of the test of the null of an integrated model against an alternative of a mean reverting model.[7] For two of the sets of assets, the estimated $\hat{\lambda}$ was on the boundary, and the test statistic is most likely not distributed $\chi^2$. All of the parameters with the exception of the two from the integrated model were significant at the 1% level. 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% for all models.

Table 2 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 always in all instances but one selected at the boundary of 1. We interpret this as a failing of the integrated model when the dynamics of the correlation between assets are less in magnitude. The typical estimated set of parameters was highly persistent($\hat{\beta} > .97$) with a slow innovation process($\hat{\alpha} < .01$). For all models, we reject the null of constant correlation in favor of dynamic conditional correlation at the 1% level based on the likelihood ratio test statistic.

Finally, to insure 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 a 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 0.012 to 0.005 while the $\hat{\beta}$ range from 0.975 to 0.99. None of the models produced parameters which were constrained on the boundary of $\alpha + \beta \leq 1$. Comparing this to the estimated coefficient of 0.163 and 0.979 for the 10 asset model estimated including the S&P 500 index, these have a slightly less dynamic structure, most likely due to the correlation structure of the different sector indices and the S&P 500 index. Figure 2 contains the estimated $\hat{\lambda}$ of the same ten models. For five of these models, the parameter value is on the boundary. We feel that the evidence does not provide support for the integrated version as all models considered preferred the mean reverting model.

Figure 3 contains a plot of the cumulative returns of 4 S&P 500 indices, while figure 4 plot the estimated dynamic correlation of these assets. Notice for the first 1500 days of the sample, that the S&P 500 index is highly correlated with the S&P 500 Auto Parts Index, but in the last 500 days of the sample this relationship changes dramatically. The correlations range form 0.2 to 0.8 over the course of the sample. Figure 5 contains the estimated

[4] 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

[5] The data were not filtered other than simple demeaning

[6] 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 $31^{s}t$ asset in the DJIA models. There was no perfectly redundant assets as the DJIA uses variable weights

[7] 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.

| No.of Assets | $\hat{\alpha}$ | $\hat{\beta}$ | $\chi^2$ value | P-value | $\hat{\lambda}$ | $\chi^2$ value | P-value |
|---|---|---|---|---|---|---|---|
| 2 | 0.0262 | 0.9664 | 107.36 | < 0.1% | 0.9742 | 97.04 | < 0.1% |
| | (5.47) | (144.30) | | | (242.47) | | |
| 3 | 0.0176 | 0.9714 | 104.36 | < 0.1% | 0.9834 | 95.5 | < 0.1% |
| | (6.16) | (175.77) | | | (460.53) | | |
| 4 | 0.0143 | 0.9734 | 126.7 | < 0.1% | 0.9864 | 117.68 | < 0.1% |
| | (7.31) | (226.17) | | | (649.67) | | |
| 5 | 0.0118 | 0.9747 | 131.32 | < 0.1% | 0.9895 | 123.78 | < 0.1% |
| | (7.69) | (258.54) | | | (837.71) | | |
| 6 | 0.0098 | 0.9763 | 129.28 | < 0.1% | 0.9927 | 121.38 | < 0.1% |
| | (7.74) | (272.59) | | | (976.01) | | |
| 7 | 0.0098 | 0.9769 | 129.28 | < 0.1% | 0.9931 | 180.1 | < 0.1% |
| | (9.16) | (330.63) | | | (1194.19) | | |
| 8 | 0.0094 | 0.9744 | 193.96 | < 0.1% | 1 | 193.42 | < 0.1% |
| | (9.29) | (290.40) | | | (0.55) | | |
| 9 | 0.0088 | 0.9724 | 193.44 | < 0.1% | 1 | 190.98 | < 0.1% |
| | (9.15) | (265.59) | | | (0.51) | | |
| 10 | 0.0163 | 0.9791 | 191 | < 0.1% | 0.9913 | 1206.82 | < 0.1% |
| | (20.3) | (906.26) | | | (4233.08) | | |
| 15 | 0.0131 | 0.9856 | 1406.04 | < 0.1% | 0.9906 | 4302.08 | < 0.1% |
| | (33.03) | (2151.57) | | | (5157.59) | | |
| 20 | 0.0109 | 0.9891 | 4590.94 | < 0.1% | 0.9895 | 5614.18 | < 0.1% |
| | (52.25) | (5118.26) | | | (6370.25) | | |
| 25 | 0.0087 | 0.9911 | 5943.44 | < 0.1% | 0.9912 | 4398.4 | < 0.1% |
| | (59.95) | (6906.42) | | | (7752.5) | | |
| 50 | 0.0056 | 0.9905 | 4707.28 | < 0.1% | 0.9952 | 5490.42 | < 0.1% |
| | (49.8) | (3476.96) | | | (14037.06 | | |
| 75 | 0.0041 | 0.9895 | 6157.6 | < 0.1% | 0.9998 | 5892.94 | < 0.1% |
| | (46.55) | (3131.71) | | | (0.56) | | |
| 100 | 0.0036 | 0.9884 | 5980.08 | < 0.1% | 0.9998 | 7304.02 | < 0.1% |
| | (47.95) | (2789.92) | | | (52995) | | |

Table 1: Parameters Estimated on the S&P 500 Indices. The numbers in parentheses are robust T-stats. The leftmost $\chi^2$ value is for the null of constant correlation against the alternative of dynamic conditional correlation, while the rightmost is for the null of an integrated process against an alternative of a dynamic conditional correlation.

variance and covariances for these stocks over the sample period. The variances for each series are simply the result of the a univariate GARCH specification. Figure 6, contains the minimum variance portfolio weights for these same four assets. For the first 1500 days 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 which include a short position in the S&P 500 index and greater the 1 weight in the S&P 500 Auto Parts index. Figure 7 contains the dynamic correlations of the first four Dow Jones Industrial Average stocks. Consistent with the lower innovation parameter estimated for this model, the conditional correlations are less in magnitude than the four asset S&P 500 model. However, there does appear to be periods where the correlation is above the unconditional correlation for 100's of days.

# 4 Conclusion

This paper presents a class of estimator which joins the simplicity and empirical success of univariate GARCH processes with an easy to interpret and estimate dynamic correlation estimator. We also feel that this estimator could be easily improved upon be considering different parameterizations of the conditional correlation estimator, although this is beyond the scope of the present paper.

# References

[1] C. Alexander, *A primer on the orthogonal garch model.* University of Reading, February 2000.

[2] T. Bollerslev, *Generalization of arch process*, Journal of Econometrics, 31 (1986), pp. 307–27.

| No.of Assets | $\hat{\alpha}$ | $\hat{\beta}$ | $\chi^2$ value | P-value | $\hat{\lambda}$ |
|---|---|---|---|---|---|
| 2 | 0.0089 | 0.9814 | 11.04 | $< 0.1\%$ | 0.9943 |
| | (2.08) | (92.23) | | | (325.11) |
| 3 | 0.0062 | 0.9747 | 7.96 | $< 0.1\%$ | 1.0000 |
| | (1.98) | (57.95) | | | (0.31) |
| 4 | 0.0059 | 0.9706 | 12.56 | $< 0.1\%$ | 1.0000 |
| | (2.75) | (78.22) | | | (0.29) |
| 5 | 0.0059 | 0.9776 | 30 | $< 0.1\%$ | 1.0000 |
| | (3.92) | (141.61) | | | (0.24) |
| 6 | 0.0056 | 0.9788 | 41.96 | $< 0.1\%$ | 1.0000 |
| | (3.91) | (110.70) | | | (0.55) |
| 7 | 0.0060 | 0.9815 | 83.34 | $< 0.1\%$ | 1.0000 |
| | (5.22) | (203.88) | | | (0.12) |
| 8 | 0.0062 | 0.9833 | 136 | $< 0.1\%$ | 1.0000 |
| | (7.26) | (346.63) | | | (0.47) |
| 9 | 0.0060 | 0.9822 | 149.78 | $< 0.1\%$ | 1.0000 |
| | (7.54) | (327.59) | | | (0.27) |
| 10 | 0.0060 | 0.9832 | 210.56 | $< 0.1\%$ | 1.0000 |
| | (8.65) | (398.27) | | | (0.10) |
| 15 | 0.0041 | 0.9826 | 175.24 | $< 0.1\%$ | 1.0000 |
| | (9.15) | (340.09) | | | (0.05) |
| 20 | 0.0039 | 0.9807 | 252.88 | $< 0.1\%$ | 1.0000 |
| | 9.83 | 371.42 | | | (0.23) |
| 25 | 0.0039 | 0.9764 | 304.66 | $< 0.1\%$ | 1.0000 |
| | 9.51 | 268.22 | | | (0.51) |
| 30 | 0.0036 | 0.9699 | 278.26 | $< 0.1\%$ | 1.0000 |
| | 9.18 | 205.39 | | | (0.26) |
| 31 | 0.0044 | 0.9798 | 735.38 | $< 0.1\%$ | 1 |
| | 14.49 | 490.68 | | | (0.09) |

Table 2: Parameters Estimated on the Dow Jones Industrial Average Stocks. The numbers in parentheses are robust T-stats. The $\chi^2$ value is for the null of constant correlation against the alternative of dynamic conditional correlation.


[3] ——, *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.

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

[5] 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.

[6] 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.

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

[8] R. F. Engle, *Dynamic conditional correlation - a simple class of multivariate garch models.* UCSD, May 2000.

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

[10] A. Hirsch, *Theory and practice of arch modelling.* Deutsch Bank, November 1997.

[11] A. Pagan, *Two stage and related estimators and their applications*, Review of Economic Studies, 53 (1986), pp. 517–538.

[12] Y. K. Tse, *A test for constant correlations in a multivariate garch model*, Journal of Econometric, 98 (2000), pp. 107–127.

[13] Y. K. Tse and A. K. Tsui, *A multivariate garch model with time-varying correlations.* National University of Singapore, December 1998.

[14] 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.

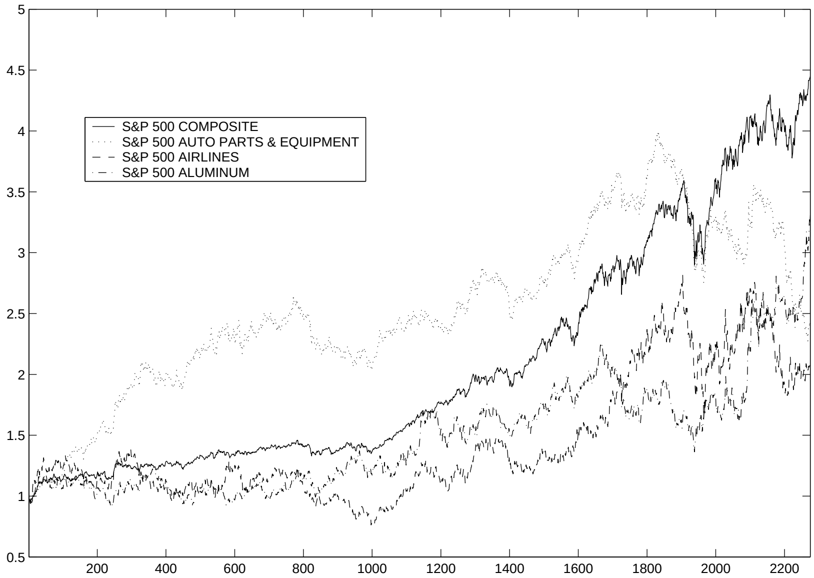


Figure 1: Graph showing the daily cumulative return of the 4 assets used in the S&P 500 indices estimation. Note the seemingly high correlation between the S&P 500 Auto Parts and the S&P 500 Index early in the sample and the break down of this relationship during the last two years of the series.

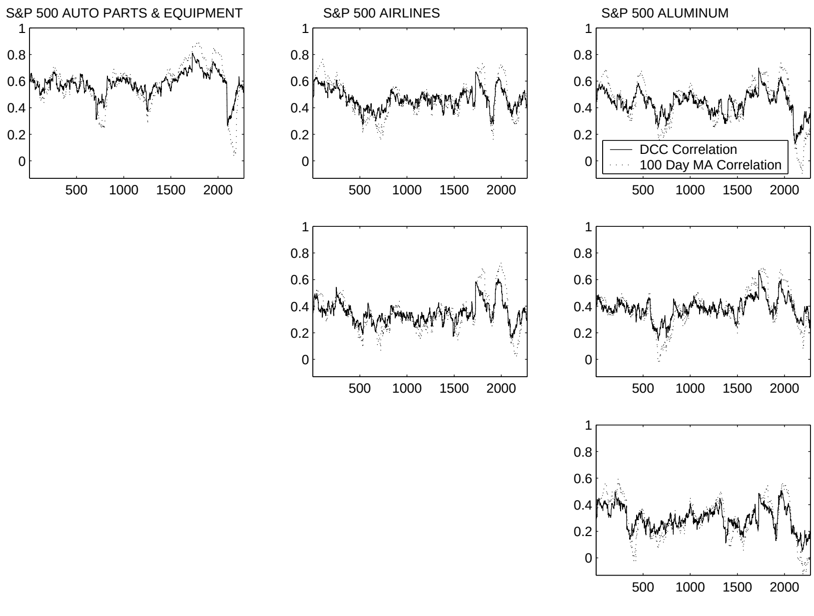


Figure 2: Graph of the Dynamic Conditional Correlation and a 100 day moving average correlation for 4 S&P 500 Indices. The DCC is mush more responsive and does not posses the typical boxcar shape of the MA correlation. The first asset in this model was the S&P 500 index, and thus the left most box on the first row is the correlation of the S&P 500 index and the S&P 500 Auto Parts sector index, and left most box on the second row is the correlation between the S&P 500 Auto Parts sector index and the S&P 500 Airlines sector index, and so forth.