CHICAGO BOOTH

The University of Chicago Booth School of Business

Working Paper No. 12-07
Fama-Miller Paper Series

Dynamic Equicorrelation

Robert Engle
New York University

Bryan Kelly
University of Chicago Booth School of Business

This paper also can be downloaded without charge from the
Social Science Research Network Electronic Paper Collection:
http://ssrn.com/abstract= 1354525

Dynamic Equicorrelation*

Robert Engle

New York University

Stern School of Business

Bryan Kelly

University of Chicago

Booth School of Business

June 2011

Abstract

A new covariance matrix estimator is proposed under the assumption that at every time period all pairwise correlations are equal. This assumption, which is pragmatically applied in various areas of finance, makes it possible to estimate arbitrarily large covariance matrices with ease. The model, called DECO, involves first adjusting for individual volatilities and then estimating correlations. A quasi-maximum likelihood result shows that DECO provides consistent parameter estimates even when the equicorrelation assumption is violated. We demonstrate how to generalize DECO to block equicorrelation structures. DECO estimates for US stock return data show that (block) equicorrelated models can provide a better fit of the data than DCC. Using out-of-sample forecasts, DECO and Block DECO are shown to improve portfolio selection compared to an unrestricted dynamic correlation structure.

Key words: Multivariate GARCH, equicorrelation, dynamic conditional correlation, conditional covariance

*Engle: 44 W. 4th St. Suite 9-190, New York, NY 10012, (212) 998-0710, rengle@stern.nyu.edu. Kelly: 5807 S. Woodlawn Ave., Chicago, IL 60637, (773) 702-8359, bryan.kelly@chicagobooth.edu.

1 Introduction

Since the first volatility models were formulated in the early eighties there have been efforts to estimate multivariate models. The specification of these models developed over the past 25 years with a range of papers surveyed by Bollerslev, Engle and Nelson (1994) and more recently by Bauwens, Laurent and Rombouts (2006) and Silvennoinen and Terasvirta (2008). A general conclusion from this analysis is that it is difficult to estimate multivariate GARCH models with more than a half dozen return series because the specifications are so complicated.

Recently, Engle (2002) proposed Dynamic Conditional Correlation (DCC), greatly simplifying multivariate specifications. DCC is designed for high-dimensional systems but has only been successfully applied to up to 100 assets by Engle and Sheppard (2005). As the size of the system grows, estimation becomes increasingly cumbersome. For cross sections of hundreds or thousands of stocks, as are common in asset pricing applications, estimation can break down completely.

Approaches exist to address the high dimension problem, though each has limitations. One type of approach is to impose structure on the system such as a factor model. Univariate GARCH dynamics in factors can generate time-varying correlations while keeping the residual correlation matrix constant through time. This idea motivated the Factor ARCH models of Engle, Ng and Rothschild (1990, 1992) and Engle's (2009b) Factor Double ARCH. The benefit of these models is their feasibility for large numbers of variates: If there are n dependent variables and k factors, estimation requires only $n+k$ GARCH models. Furthermore, it uses a full likelihood and will be efficient under appropriate conditions. One drawback is that it is not always clear what the factors are, or factor data may not be available. Another is that correlation dynamics can exist in residuals even after controlling for the factors, as in the case of US equity returns (Engle 2009; Engle and Rangel 2011). Addressing either of these problems leads back to the unrestricted DCC specification, and thus to the dimensionality dilemma.

A second solution uses the method of composite likelihood. This method was recently proposed by Engle, Shephard and Sheppard (2008) to estimate unrestricted DCC for vast cross sections. Composite likelihood overcomes the dimension limitation by breaking a large system into many smaller sub-systems in a way that generalizes the "MacGyver" method of Engle (2009a,b). This approach possesses great flexibility, but will generally be inefficient due to its reliance on a partial likelihood.

The contrast of Factor ARCH and composite likelihood highlights a fundamental trade-off in large system conditional covariance modeling. Imposing structure on the covariance can make estimation feasible and, if correctly specified, efficient; but it sacrifices generality and can suffer from break-downs due to misspecification. On the other hand, less structured models like composite likelihood break the curse of dimensionality while maintaining a general specification. However, its cost is a loss of efficiency from using a partial likelihood.

We propose a solution to this trade-off that selectively combines simplifying structural assumptions and composite likelihood versatility. We consider a system in which all pairs of returns have the same correlation on a given day, but this correlation varies over time. The model, called Dynamic Equicorrelation (DECO), eliminates the computational and presentational difficulties of high dimension systems. Because equicorrelated matrices have simple analytic inverses and determinants, likelihood calculation is dramatically simplified and optimization becomes feasible for vast numbers of assets.

DECO's structure can be substantially weakened by using block equicorrelated matrices, while maintaining the simplicity and robustness of the basic DECO formulation. A block model may capture, for instance, industry correlation structures. All stocks within an industry share the same correlation while correlations between industries take another value. In the two-block setting, analytic inverses and determinants are still available and fairly simple, thus optimization for the two-Block DECO model is as easy as the one block case. The two-block structure can also be combined with the method of composite likelihood to estimate Block DECO with an arbitrary number of blocks. Since the subsets of assets used in Block DECO are pairs of blocks rather than pairs of assets, a larger portion of the likelihood (and therefore more information) is used for optimization. As a result, the estimator can be more efficient than unrestricted composite likelihood DCC.

Another way to enrich dependence beyond equicorrelation is to combine DECO with Factor (Double) ARCH. To understand how this may work, consider a model in which one factor is observable and the dispersion of loadings on this factor is high. Further, suppose each asset loads roughly the same on a second, latent, factor. The first factor contributes to diversity among pairwise correlations, and is clearly not driven by noise. Thus, DECO may be a poor candidate for describing raw returns. However, residuals from a regression of returns on only the first factor *will* be well described by DECO. One way to model this data set is to use

Factor Double ARCH with DECO residuals. Such a model is estimated by first using GARCH regression models for each stock, then applying DECO to the standardized residuals.

What will occur if DECO is applied to variables that are *not* equicorrelated? If (block) equicorrelation is violated, DECO can still provide consistent parameter estimates. In particular, we prove quasi-maximum likelihood results showing that if DCC is a consistent estimator, then DECO and Block DECO will be consistent also. This means that when the true model is DCC, DECO makes estimation feasible when the dimension of the system may be otherwise too large for DCC to handle. While DECO is closely related to DCC, the two models are non-nested: DECO is not simply a restricted version of DCC, but a competing model. Indeed, DECO possesses some subtle, though important, features lacking in DCC. A key example is that DECO correlations between any pair of assets i and j depend on the return histories of *all* pairs. For the analogous DCC specification (ie., using the same number of parameters), the i, j correlation depends on the histories of i and j alone. In this sense, DECO parsimoniously draws on a broader information set when formulating the correlation process of each pair. To the extent that true correlations are affected by realizations of all assets, the failure of DCC to capture the information pooling aspect of DECO can disadvantage DCC as a descriptor of the data generating process.

In a one-factor world, the relation between the return on an asset and the market return is

$$r_j = \beta_j r_m + e_j, \qquad \sigma_j^2 = \beta_j^2 \sigma_m^2 + v_j.$$

If the cross-sectional dispersion of β_j is small and idiosyncrasies have similar variance each period, then the system is well-described by Dynamic Equicorrelation. A natural application for this one-factor structure lies in the market for credit derivatives such as collateralized debt obligations, or CDOs. A key feature of the risk in loan portfolios is the degree of correlation between default probabilities. A simple industry valuation model allows this correlation to be one number if firms are in the same industry and a different and smaller number if they are in different industries. Hence, within each industry an equicorrelation assumption is being made.

More broadly, to price CDOs, an assumption is often made that these are large homogeneous portfolios (LHPs) of corporate debt. As a consequence, each asset will have the same variance, the same covariance with the market factor and the same idiosyncratic variance. Thus, in an LHP, the j subscripts disappear. The correlation between any pair of assets then becomes

$$\rho = \frac{\beta^2 \sigma_m^2}{\beta^2 \sigma_m^2 + v}.$$

In fact, the LHP assumption implies equicorrelation.

The equicorrelation assumption also surfaces in derivatives trading. For instance, a popular position is to buy an option on a basket of assets and then sell options on each of the components, sometimes called a dispersion trade. By delta hedging each option, the value of this position can be seen to depend solely on the correlations. Let the basket have weights given by the vector w, and let the implied covariance matrix of components of the basket be given by the matrix $\boldsymbol{S}$. Then the variance of the basket can be expressed as $\sigma^2 = \boldsymbol{w}'\boldsymbol{S}\boldsymbol{w}$. In general we only know about the variances of implied distributions, not the covariances. Hence it is common to assume that all correlations are equal, giving $\sigma^2 = \sum_{j=1}^{n} w_j^2 s_j^2 + \rho \sum_{i \neq j} w_i w_j s_i s_j$, which can be solved for the implied correlation

$$\rho = \frac{\sigma^2 - \sum_{j=1}^{n} w_j^2 s_j^2}{\sum_{i \neq j} w_i w_j s_i s_j}. \tag{1}$$

As a consequence, the value of this position depends upon the evolution of the implied correlation. When each of the variances is a variance swap made up of a portfolio of options, the full position is called a correlation swap. As the implied correlation rises, the value of the basket variance swap rises relative to the component variance swaps.

There is a substantial history of the use of equicorrelation in economics. In early studies of asset allocation, Elton and Gruber (1973) found that assuming all pairs of assets had the same correlation reduced estimation noise and provided superior portfolio allocations over a wide range of alternative assumptions. Berndt and Savin (1975) study what could be called equi-autocorrelation matrices in production factor and consumer demand systems as a means of working with singular error variance matrices. Ledoit and Wolf (2004) use Bayesian methods for shrinking the sample correlation matrix to an equicorrelated target and show that this helps select portfolios with low volatility compared to those based on the sample correlation. Further prescriptions for avoiding the notorious noisiness of unrestricted sample correlations and betas abound in the literature (Ledoit and Wolf 2003, 2004; Michaud 1989; Jagannathan and Ma 2003; Jobson and Korkie 1980; and Meng, Hu and Bai 2011, among others). Ledoit and Wolf's Bayesian shrinkage and Elton and Gruber's parameter averaging are different approaches to noise reduction in unconditional correlation estimation. While the Bayesian method has not yet been employed for conditional variances, DECO makes it possible to incorporate Elton and Gruber's noise reduction technique into a dynamic setting. By averaging pairwise correlations, (Block) DECO smooths correlation estimates within groups. As long as this reduces estimation noise more than it compromises the true correlation structure, smoothing can be beneficial. Our empirical results suggest that the benefits of smoothing indeed extend to the conditional case. Across a range of first-stage factor models, (Block) DECO selects out-of-sample portfolios that have significantly lower volatilities than those chosen by unrestricted DCC.

The next section develops the DECO model and its theoretical properties. Section 3 presents Monte Carlo experiments that assess the model's performance under equicorrelated and non-equicorrelated generating processes. In Section 4, we apply DECO, Block DECO and DCC models to US stock return data. We find that DCC correlations between pairs of stocks have a large degree of comovement, suggesting that DECO may be beneficial in describing the system's correlation. Indeed we find that basic DECO, and DECO with 10 industry blocks, provide a better fit of the data than DCC. Lastly, we analyze the ability of (Block) DECO to construct optimal out-of-sample hedge portfolios. We find that (Block) DECO is the model that most often delivers MV portfolios with the lowest sample variance.

2 The Dynamic Equicorrelation Model

We begin by defining an equicorrelation matrix and present a result for its invertibility and positive definiteness that will be useful throughout the paper.

Definition 2.1 *A matrix $\boldsymbol{R}_t$ is an* equicorrelation matrix *of an $n \times 1$ vector of random variables if it is positive definite and takes the form*

$$\boldsymbol{R}_t = (1 - \rho_t)\boldsymbol{I}_n + \rho_t \boldsymbol{J}_n \tag{2}$$

where ρ_t is the equicorrelation, $\boldsymbol{I}_n$ denotes the n-dimensional identity matrix and $\boldsymbol{J}_n$ is the $n \times n$ matrix of *ones.*

Lemma 2.1 *The inverse and determinant of the equicorrelation matrix, $\boldsymbol{R}_t$, are given by*

$$\boldsymbol{R}_t^{-1} = \frac{1}{1-\rho_t}\boldsymbol{I}_n + \frac{-\rho_t}{(1-\rho_t)(1+[n-1]\rho_t)}\boldsymbol{J}_n \tag{3}$$

and

$$\det(\boldsymbol{R}_t) = (1-\rho_t)^{n-1}(1+[n-1]\rho_t). \tag{4}$$

Further, $\boldsymbol{R}_t^{-1}$ exists if and only if $\rho_t \neq 1$ and $\rho_t \neq \frac{-1}{n-1}$, and $\boldsymbol{R}_t$ is positive definite if and only if $\rho_t \in (\frac{-1}{n-1}, 1)$.

Proofs are provided in the online appendix.

Definition 2.2 *A time series of $n \times 1$ vectors $\{\tilde{\boldsymbol{r}}_t\}$ obeys a* Dynamic Equicorrelation (DECO) *model if $Var_{t-1}(\tilde{\boldsymbol{r}}_t) = \boldsymbol{D}_t\boldsymbol{R}_t\boldsymbol{D}_t$, where $\boldsymbol{R}_t$ is given by Equation 2 for all t and $\boldsymbol{D}_t$ is the diagonal matrix of conditional standard deviations of $\tilde{\boldsymbol{r}}_t$. The dynamic equicorrelation is ρ_t.*

2.1 Estimation

Like many covariance models, a two-stage quasi-maximum likelihood (QML) estimator of DECO will be consistent and asymptotically normal under broad conditions including many forms of model misspecification. We provide asymptotic results here for a general framework that includes several standard multivariate GARCH models as special cases, including DECO and the original DCC model. The development is a slightly modified reproduction of the two-step QML asymptotics of White (1994). After presenting the general result, we elaborate on practical estimation of DECO using Gaussian returns and GARCH covariance evolution.

First, define the (scaled) log quasi-likelihood of the model as $L(\{\tilde{\boldsymbol{r}}_t\}, \boldsymbol{\theta}, \boldsymbol{\phi}) = \frac{1}{T}\sum_{t=1}^{T} \log f_t(\tilde{\boldsymbol{r}}_t, \boldsymbol{\theta}, \boldsymbol{\phi})$, which is parameterized by vector $\boldsymbol{\gamma} = (\boldsymbol{\theta}, \boldsymbol{\phi}) \in \boldsymbol{\Gamma} = \boldsymbol{\Theta} \times \boldsymbol{\Phi}$. The two-step estimation problem may be written as

$$\max_{\boldsymbol{\theta}\in\boldsymbol{\Theta}} L_1(\{\tilde{\boldsymbol{r}}_t\}, \boldsymbol{\theta}) = \frac{1}{T}\sum_{t=1}^{T} \log f_{1,t}(\tilde{\boldsymbol{r}}_t, \boldsymbol{\theta}) \tag{5}$$

$$\max_{\boldsymbol{\phi}\in\boldsymbol{\Phi}} L_2(\{\tilde{\boldsymbol{r}}_t\}, \hat{\boldsymbol{\theta}}, \boldsymbol{\phi}) = \frac{1}{T}\sum_{t=1}^{T} \log f_{2,t}(\tilde{\boldsymbol{r}}_t, \hat{\boldsymbol{\theta}}, \boldsymbol{\phi}) \tag{6}$$

where $\hat{\boldsymbol{\theta}}$ is the solution to (5). The full two-stage QML estimator for this problem is $\hat{\boldsymbol{\gamma}} = (\hat{\boldsymbol{\theta}},\hat{\boldsymbol{\phi}})$, where $\hat{\boldsymbol{\phi}}$ is the second stage maximizer solving (6) given $\hat{\boldsymbol{\theta}}$. Under the technical assumptions listed in the online appendix, White (1994) proves the following result for consistency and asymptotic normality of $\hat{\boldsymbol{\gamma}}$.

Conjecture 2.1 (White 1994, Theorem 6.11) *Under Assumptions B.1 through B.6 in the online appendix,*

$$\sqrt{T}(\hat{\boldsymbol{\gamma}} - \boldsymbol{\gamma}^*) \overset{A}{\sim} N(\mathbf{0}, \boldsymbol{A}^{*-1}\boldsymbol{B}^*\boldsymbol{A}^{*-1}),$$

where

$$\boldsymbol{A}^* = \begin{pmatrix} \nabla_{\boldsymbol{\theta\theta}} E[L_1(\tilde{\boldsymbol{r}}_t, \boldsymbol{\theta}^*)] & \mathbf{0} \\ \nabla_{\boldsymbol{\theta\phi}} E[L_2(\tilde{\boldsymbol{r}}_t, \boldsymbol{\theta}^*, \boldsymbol{\phi}^*)] & \nabla_{\boldsymbol{\phi\phi}} E[L_2(\tilde{\boldsymbol{r}}_t, \boldsymbol{\theta}^*, \boldsymbol{\phi}^*)] \end{pmatrix} \textit{ and } \boldsymbol{B}^* = \mathrm{Var}\big(T^{-1/2}\sum_t(\boldsymbol{s}^{*\,\prime}_{1,t}, \boldsymbol{s}^{*\,\prime}_{2,t})\big)$$

where $\boldsymbol{s}^*_{1,t} = \nabla_{\boldsymbol{\theta}} L_1(\tilde{\boldsymbol{r}}_t, \boldsymbol{\theta}^*)$ *and* $\boldsymbol{s}^*_{2,t} = \nabla_{\boldsymbol{\phi}} L_2(\tilde{\boldsymbol{r}}_t, \boldsymbol{\theta}^*, \boldsymbol{\phi}^*)$.

In the remainder of the section we assume that both DECO and DCC log densities and their derivatives satisfy Assumptions B.1-B.4 and B.6 of Conjecture 2.1. These are high-level assumptions about continuity, differentiability, boundedness and the applicability of central limit theorems. Further, we assume that the DCC model is identified, this takes the form of Assumption B.5. Note that we have not verified these assumptions for the generating processes we consider in this paper. The dynamic covariance literature uniformly appeals to QML asymptotic theory when performing inference, though satisfaction of high-level assumptions has not been established for *any* model that has been proposed in this area. A rigorous analysis of asymptotic theory for multivariate GARCH processes remains an important unanswered question. For this reason, we refer to any result that follows immediately from White's (1994) Theorem 6.11 as a conjecture. We refer readers to the online appendix for more detail on this point, as well as simulation evidence that supports the satisfaction of these high-level assumptions for the DECO model.

DECO is adopted for individual applications by specifying a conditional volatility model (i.e., defining the process for $\boldsymbol{D}_t$) and a ρ_t process. We assume that each conditional volatility follows a GARCH model. We work with volatility-standardized returns, denoted by omitting the tilde, $\boldsymbol{r}_t = \boldsymbol{D}_t^{-1}\tilde{\boldsymbol{r}}_t$, so that $Var_{t-1}(\boldsymbol{r}_t) = \boldsymbol{R}_t$.

The basic ρ_t specification we consider derives from the DCC model of Engle (2002) and its cDCC modification proposed by Aielli (2009). The correlation matrix of standardized returns, $\boldsymbol{R}_t^{DCC}$, is given by

$$\boldsymbol{Q}_t = \bar{\boldsymbol{Q}}(1-\alpha-\beta) + \alpha\tilde{\boldsymbol{Q}}_{t-1}^{\frac{1}{2}}\boldsymbol{r}_{t-1}\boldsymbol{r}_{t-1}'\tilde{\boldsymbol{Q}}_{t-1}^{\frac{1}{2}} + \beta\boldsymbol{Q}_{t-1}. \tag{7}$$

$$\boldsymbol{R}_t^{DCC} = \tilde{\boldsymbol{Q}}_t^{-\frac{1}{2}}\boldsymbol{Q}_t\tilde{\boldsymbol{Q}}_t^{-\frac{1}{2}} \tag{8}$$

where $\tilde{\boldsymbol{Q}}_t$ replaces the off-diagonal elements of $\boldsymbol{Q}_t$ with zeros but retains its main diagonal and $\bar{Q}$ is the unconditional covariance matrix of standardized residuals.

DECO sets ρ_t equal to the average pairwise DCC correlation.

$$\boldsymbol{R}_t^{DECO} = (1-\rho_t)\boldsymbol{I}_n + \rho_t\boldsymbol{J}_{n\times n} \tag{9}$$

$$\rho_t = \frac{1}{n(n-1)}\left(\boldsymbol{\iota}'\boldsymbol{R}_t^{DCC}\boldsymbol{\iota} - n\right) = \frac{2}{n(n-1)}\sum_{i>j}\frac{q_{i,j,t}}{\sqrt{q_{i,i,t}q_{j,j,t}}} \tag{10}$$

where $q_{i,j,t}$ is the i,jth element of $\boldsymbol{Q}_t$. The following assumption and lemma ensure that DECO possesses certain properties important for dynamic correlation models.

Assumption 2.1 *The matrix $\bar{\boldsymbol{Q}}$ is positive definite, $\alpha+\beta<1$, $\alpha>0$ and $\beta>0$.*

Lemma 2.2 *Under Assumption 2.1, the correlation matrices generated by every realization of a DECO process according to Equations 7 through 10 are positive definite and the process is mean reverting.*

The result states that, for any correlation matrix, the transformation to equicorrelation shown in Equations 9 and 10 results in a positive definite matrix. The bounds $(\frac{-1}{n-1}, 1)$ for ρ_t are not assumptions of the model, but are guaranteed by this transformation as long as $\boldsymbol{R}_t^{DCC}$ is positive definite. As will be seen in our later discussion of Block DECO, bounds on permissible correlation values become looser as the number of blocks increases.

We estimate DECO with Gaussian quasi-maximum likelihood, which embeds it in the framework of Conjecture 2.1. Conditional on past realizations, the return distribution is $\tilde{\boldsymbol{r}}_{t|t-1} \sim N(0, \boldsymbol{H}_t)$, $\boldsymbol{H}_t = \boldsymbol{D}_t\boldsymbol{R}_t\boldsymbol{D}_t$. To establish notation, we use superscripted densities f_t^{DECO} and f_t^{DCC} to indicate the Gaussian density of $\tilde{\boldsymbol{r}}_{t|t-1}$ assuming the covariance specifically obeys DECO or DCC, respectively. Similarly, superscripted log likelihoods L^{DECO} and L^{DCC} represent the log likelihood of DECO or DCC. Omission of superscripts will be used to discuss densities and log likelihoods without specific assumptions on the dynamics or structure of covariance matrices.

The multivariate Gaussian log likelihood function L can be decomposed (suppressing constants) as

$$\begin{aligned} L &= \frac{1}{T}\sum_t \log(f_t) \\ &= -\frac{1}{T}\sum_t (\log|\boldsymbol{H}_t| + \tilde{\boldsymbol{r}}_t'\boldsymbol{H}_t^{-1}\tilde{\boldsymbol{r}}_t) \\ &= -\frac{1}{T}\sum_t \big(\log|\boldsymbol{D}_t|^2 + \tilde{\boldsymbol{r}}_t'\boldsymbol{D}_t^{-2}\tilde{\boldsymbol{r}}_t - \boldsymbol{r}_t'\boldsymbol{r}_t\big) - \frac{1}{T}\sum_t \big(\log|\boldsymbol{R}_t| + \boldsymbol{r}_t'\boldsymbol{R}_t^{-1}\boldsymbol{r}_t\big) \end{aligned}$$

Let $\boldsymbol{\theta} \in \boldsymbol{\Theta}$ and $\boldsymbol{\phi} \in \boldsymbol{\Phi}$ denote the vector of univariate volatility parameters and the vector of correlation parameters, respectively. The above equation says that the log likelihood can be separated additively into two terms. The first term, which we call $L_{Vol}(\boldsymbol{\theta})$, depends on the parameters of the univariate GARCH processes which affect only the $\boldsymbol{D}_t$ matrices and are independent of $\boldsymbol{R}_t$ and $\boldsymbol{\phi}$. The second term, which we call $L_{Corr}(\boldsymbol{\phi}, \boldsymbol{\theta})$, depends on both the univariate GARCH parameters (embedded in the $\boldsymbol{r}_t$ terms) as well as the correlation parameters. Engle (2002) and Engle and Sheppard (2005) note that correlation models of this form satisfy the assumptions of Conjecture 2.1. In particular, $L_{Vol}(\boldsymbol{\theta})$ corresponds to L_1 in the conjecture and $\hat{\boldsymbol{\theta}}$ is the vector of first-stage volatility parameter estimates. $L_{Corr}(\hat{\boldsymbol{\theta}}, \boldsymbol{\phi})$ corresponds to the second-step likelihood, L_2, so $\hat{\boldsymbol{\phi}}$ is the maximizer of $L_{Corr}(\hat{\boldsymbol{\theta}}, \boldsymbol{\phi})$. Note that the model ensures stationarity and that the covariance matrix remains positive definite in all periods whenever Assumption 2.1 holds and the univariate GARCH models are stationary.

The vector $\hat{\boldsymbol{\gamma}}^{DECO}$ is the two-stage Gaussian estimator when the second stage likelihood obeys DECO. It assumes returns are Gaussian and the correlation process obeys Equations 9 and 10. The following conjectures specifically apply Conjecture 2.1 to the Gaussian DECO and DCC models.

Conjecture 2.2 *Assuming that f_t^{DECO} and L^{DECO} satisfy Assumptions B.1-B.6 with corresponding unique maximizer $\boldsymbol{\gamma}^*$, then $\hat{\boldsymbol{\gamma}}^{DECO}$ is consistent and asymptotically normal for $\boldsymbol{\gamma}^*$.*

The analogous corollary for $\hat{\boldsymbol{\gamma}}^{DCC}$, the two-stage Gaussian DCC estimator, will be useful in developing our later comparison between DECO and DCC.

Conjecture 2.3 *Assuming that f_t^{DCC} and L^{DCC} satisfy Assumptions B.1-B.6 with corresponding unique maximizer $\boldsymbol{\gamma}^*$, then $\hat{\boldsymbol{\gamma}}^{DCC}$ is consistent and asymptotically normal for $\boldsymbol{\gamma}^*$.*

In addition, the asymptotic covariance matrices for $\hat{\boldsymbol{\gamma}}^{DECO}$ and $\hat{\boldsymbol{\gamma}}^{DCC}$ corresponding to Conjectures 2.2 and 2.3 take the form stated in Conjecture 2.1, replacing L_1 and L_2 with the appropriately superscripted likelihoods $L_{Vol}^{(\cdot)}$ and $L_{Corr}^{(\cdot)}$.

To appreciate the payoff from making the equicorrelation assumption, consider the second step likelihood under DECO.

$$
\begin{aligned}
L_{Corr}^{DECO}(\hat{\boldsymbol{\theta}}, \boldsymbol{\phi}) &= -\frac{1}{T}\sum_t (\log |\boldsymbol{R}_t^{DECO}| + \hat{\boldsymbol{r}}_t' \boldsymbol{R}_t^{DECO^{-1}} \hat{\boldsymbol{r}}_t) \\
&= -\frac{1}{T}\sum_t \left[\log\left([1-\rho_t]^{n-1}[1+(n-1)\rho_t] \right) + \frac{1}{1-\rho_t}\left(\sum_i (\hat{r}_{i,t}^2) - \frac{\rho_t}{1+(n-1)\rho_t}(\sum_i \hat{r}_{i,t})^2 \right) \right]
\end{aligned}
$$

where $\hat{\boldsymbol{r}}_t$ are returns standardized for first stage volatility estimates, $\hat{\boldsymbol{r}}_t = \boldsymbol{D}_t(\hat{\boldsymbol{\theta}})^{-1}\tilde{\boldsymbol{r}}_t$, and ρ_t obeys Equation 10. In DCC, the conditional correlation matrices must be recorded and inverted for all t and their determinants calculated; further, these T inversions and determinant calculations are repeated for each of the many iterations required in a numeric optimization program. This is costly for small cross sections and potentially infeasible for very large ones. In contrast, DECO reduces computation to n-dimensional vector outer products with no matrix inversions or determinants required, rendering the likelihood optimization problem manageable even for vast-dimensional systems. The likelihood at time t can be calculated from just three statistics, the average cross-sectional standardized return, the average cross-sectional squared standardized return, and the predicted correlation. This is a simple calculation in all settings considered in this paper.

2.2 Differences Between DECO and DCC

The transformation from DCC correlations to DECO in Equation 10 introduces subtle differences between the DECO and DCC likelihoods. The two models are non-nested; they share the same number of parameters and there is no parameter restriction that makes the models identical. The correlation matrix for DCC is non-equicorrelated in all realizations, while, by definition, DECO is always exactly equicorrelated.

Both models build off of the $\boldsymbol{Q}$ process in (7). On a given day, $\boldsymbol{Q}$ is updated as a function of the lagged return vector and the lagged $\boldsymbol{Q}$ matrix. From here, $\boldsymbol{Q}$ is transformed to $\boldsymbol{R}^{DCC} = \tilde{\boldsymbol{Q}}_t^{-\frac{1}{2}} \boldsymbol{Q}_t \tilde{\boldsymbol{Q}}_t^{-\frac{1}{2}}$. $\boldsymbol{R}^{DCC}$ is the correlation matrix that enters the DCC likelihood. Note that the i, j element of this matrix is $q_{i,j,t}/\sqrt{q_{i,i,t}q_{j,j,t}}$. Clearly, the information about pair i, j's correlation at time t depends on the history of assets i and j alone. On the other hand, the correlation between i and j under DECO is $\rho_t = \frac{2}{n(n-1)} \sum_{i>j} \frac{q_{i,j,t}}{\sqrt{q_{i,i,t}q_{j,j,t}}}$, which depends on the history of *all* pairs. The failure of DCC to capture this information pooling aspect of DECO correlations hinders the ability of the DCC likelihood to provide a good description of the data generating process, resulting in poor estimation performance, as will be seen in the Monte Carlo results of Section 3.

There is also a key difference that arises from the need to estimate DCC with a partial, composite likelihood. When DECO is the true model, DCC estimation is akin to estimating the correlation of a single pair, sampled $n(n-1)/2$ times. The difference between each pair is measurement error. DECO, by averaging pairwise correlation at each step, attenuates this measurement error. It also uses the full cross-sectional likelihood rather than a partial one like composite likelihood.

2.3 DECO as a Feasible DCC Estimator

Often the equicorrelation assumption fails so that there is cross-sectional variation in pairwise correlations, as in DCC. In this case the DECO model remains a powerful tool. The following result shows that as long as DCC is a Gaussian QML estimator, DECO will be also.

Proposition 2.1 *Assume that f_t^{DECO} and L^{DECO} satisfy Assumptions B.1-B.4 and B.6, and assume that the DCC model is identified and therefore satisfies Assumption B.5. Then Assumption B.5 is guaranteed to be satisfied for the DECO model, so that $\hat{\boldsymbol{\gamma}}^{DECO}$ is identified and hence DECO is a consistent and asymptotically normal estimator for $\boldsymbol{\gamma}^*$.*

Suppose that DCC is the true model and that the high-level regularity conditions on the densities of DCC and DECO (differentiability and boundedness, etc.) are satisfied, but that DECO is not *assumed* to be identified. Then Proposition 2.1 *guarantees* that DECO is identified, therefore it consistently estimates the true DCC parameters. Furthermore, the asymptotic covariance matrix for $\hat{\boldsymbol{\gamma}}^{DECO}$ takes the form stated in Conjecture 2.1 (replacing L_1 and L_2 with L_{Vol}^{DECO} and L_{Corr}^{DECO}) since Proposition 2.1 in turn satisfies the conditions of Conjecture 2.2.

2.4 Estimation Structure Versus Fit Structure

How useful is Proposition 2.1 in practice? Suppose, for instance, the system is so large that DCC estimation based on full maximum likelihood is infeasible. The result says that one can consistently estimate DCC parameters using DECO despite its misspecification. The estimated parameters can then be plugged into Equation 7 to reconstruct the unrestricted DCC fitted process. In short, DECO, like composite likelihood, provides feasible estimates for a DCC model that may be otherwise computationally infeasible.

The flexibility of DECO goes beyond its ability to fit unrestricted DCC processes. The logic of Proposition 2.1 ensures that DECO can consistently estimate block equicorrelation processes as well. To do this, α and β are estimated with DECO, then data is run through the evolution equation in (7), plugging in $\hat{\alpha}^{DECO}$ and $\hat{\beta}^{DECO}$. The resulting DCC correlation series, based on DECO estimates, can be used to construct any fitted block correlation structure after the fact by averaging pairwise DCC correlations within blocks (see schematic in Figure 1).

Throughout the remaining sections we will refer to "estimation structures" and "fit structures," and it is important to draw the distinction between them. The estimation structure is the structure that the correlation matrix takes *within the likelihood*. When DECO is used, the estimation structure is a single block. The fit structure, on the other hand, refers to the structure of the final, fitted correlation matrices. It is achieved by averaging DCC correlations within blocks *after* estimation. The resulting block structure can be different than the estimation structure and have one block, many blocks, or be unrestricted (as in DCC).

In the next section, we present an alternative estimation approach called Block DECO. Block DECO directly models the block correlation structure ex ante and makes use of it within the estimation procedure. In this case, the estimation structure will be allowed to have multiple blocks. As with DECO, ex post block averaging can be used to generate a different desired correlation fit structure. With Block DECO as the estimator, fitted correlations can have the same, more, or fewer blocks than the estimation structure.

Using DECO with ex post averaging to achieve block correlations is, from an implementation standpoint, simpler than using full-fledged Block DECO estimation. As will be shown, Block DECO estimation involves composite likelihood and thus is operationally more complex. Ex post averaging achieves the same outcome of dynamic block correlations with the simplicity of DECO's Gaussian QML estimation. The advantage of more complicated Block DECO estimation is that it can potentially be more efficient. We turn to that model now.

2.5 The Block Dynamic Equicorrelation Model

While DECO will be consistent even when equicorrelation is violated, it is possible that a loosening of the structure to block equicorrelation can improve maximum likelihood estimates. In this vein, we extend DECO to take the block structure into account ex ante and thus incorporate it into the estimation procedure.

As an example of Block DECO's usefulness, consider modeling correlation of stock returns with particular interest in intra- and inter-industry correlation dynamics. This may be done by imposing equicorrelation within and between industries. Each industry has a single dynamic equicorrelation parameter and each industry pair has a dynamic cross-equicorrelation parameter. With block equicorrelation, richer cross-sectional variation is accommodated while still greatly reducing the effective dimensionality of the correlation matrix.

This section presents the class of block Dynamic Equicorrelation models and examines their properties.

Definition 2.3 $\boldsymbol{R}_t$ *is a* K-block equicorrelation matrix *if it is positive definite and takes the form*

$$\boldsymbol{R}_t = \begin{pmatrix} (1-\rho_{1,1,t})\boldsymbol{I}_{n_1} & 0 & \cdots \\ 0 & \ddots & 0 \\ \vdots & 0 & (1-\rho_{K,K,t})\boldsymbol{I}_{n_K} \end{pmatrix} + \begin{pmatrix} \rho_{1,1,t}\boldsymbol{J}_{n_1} & \rho_{1,2,t}\boldsymbol{J}_{n_1\times n_2} & \cdots \\ \rho_{2,1,t}\boldsymbol{J}_{n_2\times n_1} & \ddots & \\ \vdots & & \rho_{K,K,t}\boldsymbol{J}_{n_K} \end{pmatrix} \tag{11}$$

where $\rho_{l,m,t} = \rho_{m,l,t} \;\; \forall l, m.$

Block DECO specifies that, conditional on the past, each variable is Gaussian with mean zero, variance one, and correlations taking the structure in Equation 11. The return vector $\boldsymbol{r}_t$ is partitioned into K sub-vectors; each sub-vector $\boldsymbol{r}_l$ contains n_l returns. The Block DECO correlation matrix, $\boldsymbol{R}_t^{BD}$, allows distinct processes for each of the K diagonal blocks and $K(K-1)/2$ unique off-diagonal blocks. Blocks on the main diagonal have equicorrelations following $\rho_{l,l,t}$ while blocks off the main diagonal follow $\rho_{l,m,t}$, where

$$\rho_{l,l,t} = \frac{1}{n_l(n_l-1)} \sum_{i\in l, j\in l, i\neq j} \frac{q_{i,j,t}}{\sqrt{q_{i,i,t}q_{j,j,t}}} \quad \text{and} \quad \rho_{l,m,t} = \frac{1}{n_l n_m} \sum_{i\in l, j\in m} \frac{q_{i,j,t}}{\sqrt{q_{i,i,t}q_{j,j,t}}}. \tag{12}$$

The i, jth element of the matrix in (7) is $q_{i,j,t}$, so Block DECO correlations are calculated as the average DCC correlation within each block. Despite the block structure of equicorrelations, Equation 7 remains the underlying DCC model, thus the parameters α and β do not vary across blocks. Densities, log likelihoods and parameter estimates corresponding to Block DECO model are superscripted with BD in congruence with notation for the base DECO model.

The following results show the consistency and asymptotic normality of Block DECO. In analogy to DECO, $\hat{\boldsymbol{\gamma}}^{BD}$ is the two-stage Gaussian Block DECO estimator assuming returns are Gaussian and the correlation process obeys Equation 12.

Conjecture 2.4 *Assuming that f_t^{BD} and L^{BD} satisfy Assumptions B.1-B.6 with corresponding unique maximizer $\boldsymbol{\gamma}^*$, then $\hat{\boldsymbol{\gamma}}^{BD}$ is consistent and asymptotically normal for $\boldsymbol{\gamma}^*$.*

Further, like DECO, Block DECO is a QML estimator of DCC models.

Proposition 2.2 *Assume that f_t^{BD} and L^{BD} satisfy Assumptions B.1-B.4 and B.6, and assume that the DCC model is identified and therefore satisfies Assumption B.5. Then Assumption B.5 is guaranteed to be satisfied for the Block DECO model, so that $\hat{\boldsymbol{\gamma}}^{BD}$ is identified and hence Block DECO is a consistent and asymptotically normal estimator for $\boldsymbol{\gamma}^*$.*

The proof follows the same argument as the proof of Proposition 2.1. The asymptotic covariance matrix for $\hat{\boldsymbol{\gamma}}^{BD}$ takes the form stated in Conjecture 2.1 (replacing L_1 and L_2 with L_{Vol}^{BD} and L_{Corr}^{BD}).

Block DECO balances the flexibility of unrestricted correlations with the structural simplicity of DECO. However, when the number of blocks is greater than two, the analytic forms for the inverse and determinant of the Block DECO matrix begin to lose their tractability. A special case that remains simple regardless of the number of blocks occurs when each of the blocks on the main diagonal are equicorrelated, but all off-diagonal block equicorrelations are forced to zero. Each diagonal block constitutes a small DECO submodel, and therefore its inverse and determinant are known. The full inverse matrix is the block diagonal matrix of inverses for the DECO sub-models, and its determinant is the product of the sub-model determinants.

Conveniently, the composite likelihood method can be used to estimate Block DECO in more general cases. The composite likelihood is constructed by treating each pair of blocks as a sub-model, then calculating the quasi-likelihoods of each sub-model, and finally summing quasi-likelihoods over all block pairs. As discussed in Engle, Shephard and Sheppard (2008), each pair provides a valid, though only partially informative, quasi-likelihood. A model for any number of blocks requires only the analytic inverse and determinant for a two-block equicorrelation matrix when using the method of composite likelihood. The following lemma establishes the analytic tractability provided by two-block equicorrelation. We suppress t subscripts as all terms are contemporaneous.

Lemma 2.3 *If $\boldsymbol{R}$ is a two-block equicorrelation matrix, that is, if*

$$\boldsymbol{R} = \begin{bmatrix} (1-\rho_{1,1})\boldsymbol{I}_{n_1} & 0 \\ 0 & (1-\rho_{2,2})\boldsymbol{I}_{n_2} \end{bmatrix} + \begin{bmatrix} \rho_{1,1}\boldsymbol{J}_{n_1\times n_1} & \rho_{1,2}\boldsymbol{J}_{n_1\times n_2} \\ \rho_{2,1}\boldsymbol{J}_{n_2\times n_1} & \rho_{2,2}\boldsymbol{J}_{n_2\times n_2} \end{bmatrix}$$

then,

i. the inverse is given by

$$\boldsymbol{R}^{-1} = \begin{bmatrix} b_1\boldsymbol{I}_{n_1} & 0 \\ 0 & b_2\boldsymbol{I}_{n_2} \end{bmatrix} + \begin{bmatrix} c_1\boldsymbol{J}_{n_1\times n_1} & c_3\boldsymbol{J}_{n_1\times n_2} \\ c_3\boldsymbol{J}_{n_2\times n_1} & c_2\boldsymbol{J}_{n_2\times n_2} \end{bmatrix}$$

where

$$b_i = \frac{1}{1-\rho_{i,i}},\ i = 1,2$$

$$c_1 = \frac{\rho_{1,1}\big(\rho_{2,2}(n_2-1)+1\big) - \rho_{1,2}^2 n_2}{(\rho_{1,1}-1)\big([\rho_{1,1}(n_1-1)+1][\rho_{2,2}(n_2-1)+1] - n_1 n_2 \rho_{1,2}^2\big)}$$

$$c_2 = \frac{\rho_{2,2}\big(\rho_{1,1}(n_1-1)+1\big) - \rho_{1,2}^2 n_1}{(\rho_{2,2}-1)\big([\rho_{1,1}(n_1-1)+1][\rho_{2,2}(n_2-1)+1] - n_1 n_2 \rho_{1,2}^2\big)}$$

$$c_3 = \frac{\rho_{1,2}}{n_1 n_2 \rho_{1,2}^2 - \big(\rho_{1,1}(n_1-1)+1\big)\big(\rho_{2,2}(n_2-1)+1\big)}$$

ii. the determinant is given by

$$\det(\boldsymbol{R}) = (1-\rho_{1,1})^{n_1-1}(1-\rho_{2,2})^{n_2-1}\big[(1+[n_1-1]\rho_{1,1})(1+[n_2-1]\rho_{2,2}) - \rho_{1,2}^2 n_1 n_2\big]$$

iii. $\boldsymbol{R}$ *is positive definite if and only if*

$$\rho_i \in \big(\frac{-1}{n_i-1}, 1\big), \; i = 1, 2$$

and

$$\rho_{1,2} \in \Bigg(-\sqrt{\frac{\big(\rho_{1,1}(n_1-1)+1\big)\big(\rho_{2,2}(n_2-1)+1\big)}{n_1 n_2}}, \sqrt{\frac{\big(\rho_{1,1}(n_1-1)+1\big)\big(\rho_{2,2}(n_2-1)+1\big)}{n_1 n_2}} \Bigg).$$

With this result in hand, the likelihood function of a two-Block DECO model can, as in the simple equicorrelation case, be written to avoid costly inverse and determinant calculations.

$$\begin{aligned}
L &= -\frac{1}{2}\sum_t \big(\log|\boldsymbol{R}_t| + \boldsymbol{r}_t'\boldsymbol{R}_t^{-1}\boldsymbol{r}_t\big) \\
&= -\frac{1}{2}\sum_t \Bigg[\log\Bigg((1-\rho_{1,1,t})^{n_1-1}(1-\rho_{2,2,t})^{n_2-1}\big[(1+[n_1-1]\rho_{1,1,t})(1+[n_2-1]\rho_{2,2,t}) - \rho_{1,2,t}^2 n_1 n_2\big]\Bigg) \\
&\quad + \boldsymbol{r}_t'\Bigg(\begin{bmatrix} b_1\boldsymbol{I}_{n_1} & 0 \\ 0 & b_2\boldsymbol{I}_{n_2}\end{bmatrix} + \begin{bmatrix} c_1\boldsymbol{J}_{n_1\times n_1} & c_3\boldsymbol{J}_{n_1\times n_2} \\ c_3\boldsymbol{J}_{n_2\times n_1} & c_2\boldsymbol{J}_{n_2\times n_2}\end{bmatrix}\Bigg)\boldsymbol{r}_t\Bigg] \\
&= -\frac{1}{2}\sum_t \Bigg[\log\Bigg((1-\rho_{1,1,t})^{n_1-1}(1-\rho_{2,2,t})^{n_2-1}\big[(1+[n_1-1]\rho_{1,1,t})(1+[n_2-1]\rho_{2,2,t}) - \rho_{1,2,t}^2 n_1 n_2\big]\Bigg) \\
&\quad + \Bigg(b_1\sum_i^{n_1} r_{i,1}^2 + b_2\sum_i^{n_2} r_{i,2}^2 + c_{1,t}(\sum_i^{n_1} r_{i,1})^2 + 2c_{3,t}(\sum_i^{n_1} r_{i,1})(\sum_i^{n_2} r_{i,2}) + c_{2,t}(\sum_i^{n_2} r_{i,2})^2\Bigg)\Bigg]
\end{aligned}$$

In the multi-block case, the above two-block log likelihood is calculated for each pair of blocks, and then these submodel likelihoods are summed, forming the objective function to be maximized.

3 Correlation Monte Carlos

3.1 Equicorrelated Processes

This section presents results from a series of Monte Carlo experiments that allow us to evaluate the performance of the DECO framework when the true data generating process is known. We begin by exploring the model's estimation ability when DECO is the generating process. Asset return data for 10, 30 or 100 assets are simulated over 1,000 or 5,000 periods according to Equations 7-10. We also consider a range of values for α and β. For each simulated data set, we estimate DECO and composite likelihood DCC. Here and throughout, we use a subset of n randomly chosen pairs of assets to form the composite likelihood in order to speed up computation. In unreported results, we run a subset of our simulations estimating composite likelihood with all $n(n-1)/2$ pairs, and results were virtually indistinguishable. Engle, Shephard and Sheppard (2008) find that the loss from using a subset of n pairs is negligible.

Simulations are repeated 2,500 times and summary statistics for the maximum likelihood parameter estimates are calculated. Table 1 reports the mean, median and standard deviation of α and β estimates, their average QML asymptotic standard errors (calculated using the "sandwich" covariance estimator of Bollerslev and Wooldridge 1992), and the root mean squared error (RMSE) for the true versus fitted average pairwise correlation process. Both models use correlation targeting, thus the intercept matrix is the same for both models and not reported.

The results show that across parameter values, cross section sizes, and sample lengths, DECO outperforms unrestricted DCC in terms of both accuracy and efficiency. Depending on simulation parameters, DECO is between two to ten times more accurate than DCC at matching the simulated average correlation path, as measured by RMSE. In small samples, DCC can fare particularly poorly. For example, when $T = 1000$, $n = 10$ and the $\beta = 0.97$, DCC's mean β estimate is 0.55, versus 0.93 for DECO. Also, DCC's QML standard errors can grossly underestimate the true variability of its estimates for all sample sizes. The simulations also show that when the data generating process is equicorrelation, increasing the number of assets in the cross section improves estimates.

What explains the poor performance of DCC? Some characteristics of the DECO likelihood are lacking in DCC, as discussed in Section 2.2. DCC updates pairwise correlations using pairwise data histories, rather than using the data history of all series as in DECO. Also, the partial information nature of composite likelihood DCC makes it even more difficult for DCC to estimate the parameters of a DECO process.

3.2 Non-Equicorrelated Processes

Proposition 2.1 highlights DECO's ability to consistently estimate DCC parameters despite violation of equicorrelation. To demonstrate the performance of DECO in this light, we simulate series using DCC as the data generating process (Equations 7 and 8). Thus, while equicorrelation is violated, the average pairwise correlation behaves according to DECO and the assumptions of Conjecture 2.1 are satisfied. In the correlation evolution we use an intercept matrix that is non-equicorrelated; the standard deviation of off-diagonal elements is 0.33, demonstrating that the differences in pairwise correlations for the simulated cross sections are substantial.

Again, we generate return data for 10, 30 or 100 assets over 1,000 or 5,000 periods using a range of values for α and β. Next, we estimate both DECO and composite likelihood DCC. Table 2 reports summary statistics. DECO exhibits a downward bias in its β estimates that is exacerbated at low values of T/N. For large T/N, the β bias nearly disappears. Composite likelihood performs comparatively well, though the difference in accuracy versus DECO is almost indistinguishable when T is 5,000. The superior performance of DCC is perhaps most clearly seen in its excellent precision. In all cases, the variability of DCC estimates are a fraction of DECO's.

It appears that DECO's performance under misspecification (Table 2) is overall better than DCC's performance under misspecification (Table 1). Small samples generally result in a downward bias in DECO estimates, but these estimates always manage to stay within one standard error of the estimates achieved by the correctly specified model. This is in contrast to the severe downward biases displayed by DCC in Table 1. Similarly, DECO's QML standard errors, while understated by an order of magnitude of roughly two, are only mildly biased compared to the performance of DCC's standard errors in Table 1.

4 Empirical Analysis

4.1 Data

Since DECO is motivated primarily as a means of estimating dynamic covariances for large systems, our sample includes constituents of the S&P 500 Index. A stock is included if it was traded over the full horizon

1995-2008 and was a member of the index at some point during that time. This amounts to 466 stocks. Data on returns and SIC codes (which will be used for block assignments) come from the CRSP daily file. In our Factor ARCH regressions and Block DECO estimation we use Fama-French three-factor return data and industry assignments (based on SICs) from Ken French's website. Precise definitions of portfolios can be found there.

We also compare average correlations for (Block) DECO and DCC to option implied correlations. For this analysis we use a 36 stock subset of the S&P sample that were continuously traded over 1995-2008 and were members of the Dow Jones Industrials at some point in that period. We also use daily option implied volatilities on these constituents and the index from October 1997 through September 2008 from the standardized options file of OptionMetrics.

Before proceeding to the results, we include a brief aside regarding estimation that will be important for the information criterion comparisons we make throughout. All second-stage correlation models that we estimate have the same number of parameters: an α estimate, a β estimate, and $n(n-1)/2$ unique elements of the intercept matrix. Each factor structure, however, has a different number of parameters. Residual GARCH models contain a total of $5n$ parameters. In addition, the loadings in a K-factor model (including a constant as one of the K factors) contribute an additional nK parameters. Also, the likelihoods from different composite likelihood methods are not directly comparable because they use sub-models of differing dimensions. Therefore, we use composite likelihood fitted parameters to evaluate the full joint Gaussian likelihood after the fact, which is directly comparable to the DECO likelihood.

4.2 Dynamic Equicorrelation in the S&P 500, 1995-2008

Our appraisal of DECO has thus far relied on simulated data, now we assess DECO estimates for the S&P 500 sample. As discussed in the section on model estimation, we use a consistent two-step procedure to estimate correlations. In the first stage we regress individual stock returns on a constant and specify residuals to be asymmetric GARCH(1,1) processes with Student-t innovations (Glosten, Jagannathan and Runkle 1993). GARCH regressions are estimated stock-by-stock via maximum likelihood, and then volatility-standardized residuals are given as inputs to the second-stage DECO model. Here and throughout, second-stage models are estimated using correlation targeting for the intercept matrix $\bar{\boldsymbol{Q}}$. The first column of Panel A in Table 3a shows estimates for the basic DECO specification, their standard errors, and the Akaike information criterion (AIC) for the full two-stage log likelihood.

We find $\hat{\alpha} = .021$ and $\hat{\beta} = .979$, thus the DECO parameters are in the range of typical estimates from GARCH models. Rounded to three decimals places, $\hat{\alpha}$ and $\hat{\beta}$ sum to one, indicating that the equicorrelation is nearly integrated. Figure 2a plots the fitted S&P DECO series against the price level of the S&P 500 Index. The clearest feature of the plot is the tendency for the average correlation to rise when the market is decreasing and fall when the market is increasing. This inverse relationship between market value and correlations has been documented previously in the literature. Longin and Solnik (1995, 2001) find that correlations between country level indices are higher during bear markets and in volatile periods. Ang and Chen (2002) find the same result for correlations between portfolios of US stocks and the aggregate market. Our results show that, over the past 15 years, correlations reached their highest level during the global crisis in the last four months of 2008, when the average correlation between S&P 500 stocks reached nearly 60%.

4.3 Factor ARCH DECO

As discussed in the introduction, DECO may be used to model residuals from a factor model of returns. As a simple example, consider a one factor model for returns: $r_j = \beta_j r_m + e_j$. If the factor r_m (and each idiosyncrasy e_j) obeys a univariate GARCH model and if the vector of idiosyncrasies e is dynamically equicorrelated, then we call this a Factor (Double) ARCH DECO model. (See Engle 2009b for additional detail on appending multivariate GARCH models to factor model residuals.) The log likelihood of a factor model decomposes additively since $\log f_{r,t}(\boldsymbol{r}_t) = \log f_{r,t}(\boldsymbol{r}_t|Factors_t) + \log f_{Factors,t}(Factors_t)$. An additive log likelihood can be maximized by maximizing each element of the sum separately, thus the volatility and correlations of factors can be estimated separately from the volatility and correlations of residuals and estimates will be consistent.

Our next empirical result demonstrates the usefulness of DECO in capturing lingering dynamics among correlations of factor model residuals. We consider two factor structures for returns: the CAPM and the Fama-French (1993) three-factor model. In both cases, the first-stage models are regressions of individual stock returns on a set of factors where, as before, all factors and idiosyncrasies are modeled as asymmetric

GARCH(1,1). The second-stage is estimated with the basic DECO specification. Estimation results are shown in the first column of Panels B and C in Table 3a. CAPM residual correlations are slightly less persistent than the no factor case, with $\hat{\alpha} + \hat{\beta} = .992$. Adding the market factor substantially increases the log likelihood, even after accounting for its additional parameters (as seen by the decreased AIC versus column 1 of Panel A). CAPM residual correlations are plotted against S&P 500 Index price level in Figure 2b. On average, the residual correlation is very low, dropping to less than 2% for most of the sample from a time average of over 20% with no factor (Figure 2a). The most striking feature of this plot is the large increase in residual correlations from 1999 through late 2001, corresponding to the rise and fall of the technology bubble. It appears that, before and after the tech boom, the CAPM does a very good job of describing return correlations. During the tech episode, an additional factor seems to surface. The impact of this factor on dependence among assets is not captured by the CAPM, but is picked up by residual DECO.

Estimates for residual correlations using the Fama-French three-factor model show much weaker dynamics among residual correlations, as persistence drops to $\hat{\alpha} + \hat{\beta} = .814$. Including three factors further improves the AIC. Figure 2c shows that residual correlations are almost always about 1.5% and flat. The only exceptions are brief spikes to 3% during the peak of the tech bubble and the crisis of late 2008.

4.4 Comparing DECO and DCC Correlations

The previous subsections have explored DECO fits using no factor structure, the one-factor CAPM, and the Fama-French three-factor model. We now examine the fits of DCC for each of these three first-stage models to compare with DECO. DCC is estimated using composite likelihood with sub-models that are pairs of stocks; n of the possible $n(n-1)/2$ pairs are randomly selected as sub-models, where $n = 466$ for our sample. The use of a large subset of all pairs reduces computation time while negligibly degrading the performance of the estimator (as suggested by Engle, Shephard and Sheppard 2008). As a check of this point, we use all pairs for estimating composite likelihood DCC in a smaller cross section of 36 Dow Jones constituents (a subset of our S&P sample) and find that the results are nearly identical to the results when only 36 randomly selected pairs are used.

DCC parameter estimates are shown in the third column of Table 3a. When no factor is used, parameter estimates for DECO and DCC are similar and within two standard errors of each other. DECO achieves a lower AIC, making it the better model according to this criterion. Note, DECO and DCC use the same number of parameters, so the lower AIC for DECO is due solely to its better likelihood fit. We next evaluate how much pairwise DCC correlations deviate from the equicorrelation series of DECO. Figure 3a plots DECO against the 25th, 50th and 75th percentile of pairwise DCC correlations when the first stage model has no factor. These quartiles give a sense of the dispersion of pairwise correlations. As the figure shows, the upper and lower quartiles are almost always within 5% of the median, and the dynamic pattern of the quartiles closely track the equicorrelation. The similar correlation dynamics for pairs of stocks and for equicorrelation is consistent with DECO's ability to achieve a superior fit.

When the first-stage model includes the CAPM market factor, DCC α and β estimates again are very close to those of DECO. In this case, DCC fits the data better according to the Akaike criterion. To understand how DCC might provide a better fit, consider the DCC residual correlation quartiles shown in Figure 3b. We see first that the dispersion of correlations has increased relative to the average correlation. Residual equicorrelation is roughly 2-3% over time, while the 75th and 25th DCC percentiles are around 6% and -3% on average. Furthermore, other than during the technology bubble, there appears to be no systematic relationship between the time series pattern in equicorrelation and the pattern of pairwise correlations. This picture therefore suggests that the ability of DECO to describe residual CAPM correlations is limited, consistent with the AIC values we find.

Using the Fama-French model reinforces the notion that DCC is a more apt descriptor of factor model residuals due to the tendency for residual pairwise correlations to exhibit idiosyncratic dynamics. Table 3a, Panel C shows that DCC continues to find stronger dynamics in correlations than DECO in terms of α and β estimates, and pairwise DCC correlations in Figure 3c are quite distinct from the equicorrelation in their time series behavior. In summary, our results elucidate the conditions under which DECO can provide a good description of the data. When comovement among all pairs shows broadly similar time series dynamics, DECO fits well and outperforms DCC. Conversely, when dynamics in pairwise correlations are dissimilar, DCC may be a more appropriate model.

4.5 Block DECO

In our last description of correlations among S&P constituents we repeat the above analyses using 10-Block DECO as the correlation estimator. Stocks are assigned to blocks based on SIC codes according to the industry classification scheme for Ken French's 10 industry portfolios. We estimate 10-Block DECO using Gaussian composite likelihood with sub-models that are pairs of blocks. Due to the low number of blocks, all $10(10-1)/2$ pairs of industries are used to form the Block DECO composite likelihood. In particular, when formulating the likelihood contribution of industry pair i, j, a total of $n_i + n_j$ stocks are used in the sub-model.

When no factors are used in the first-stage GARCH regressions, Block DECO achieves a better AIC than both DCC and DECO, and finds similar parameter estimates. To get a sense of the flexibility Block DECO adds to the correlation structure, Figure 4 plots within-industry correlations for energy, telecom and health stocks. We choose only three of the 10 sectors to keep the plot legible while illustrating the richness a block structure can add to the cross section of correlations. A few interesting patterns emerge. First, the correlation among energy stocks has slowly trended upward over the entire sample. While correlations were low for the market as a whole over 2004-2007, energy correlations remained high and continued to climb. Telecom stocks, meanwhile, had the sharpest rise in correlations in the market downturn following the technology boom. Health stocks maintained relatively low correlations throughout the sample. All three groups, however, experience drastic increases in correlations during late 2008, at which time all groups saw their highest level of comovement.

We also estimate Block DECO on residuals from the CAPM and Fama-French model. While Block DECO achieves a better fit than DECO in these factor models, DCC maintains the superior AIC. Block DECO, like DCC, finds more persistent dynamics in correlations for Fama-French residuals than DECO.

4.6 Equicorrelation and Implied Correlations, Dow Jones Index

Options traded on an index and its members provide an opportunity to validate fits from correlation models against forward-looking implied correlations that are based solely on options prices. We briefly compare fitted correlations from DECO and DCC to option-implied correlations. Since options do not exist for all members of the S&P 500, we instead examine the Dow Jones Index, for which liquid options are traded on all consituents. Our sample of options data for the Dow Jones and its members begins in October 1997 (when Dow Jones Index options were introduced) through September 2008. Implied correlation is calculated from implied volatilities of the index and its constituents as in Equation 1. We use implied volatilities on call options standardized to have one month to maturity, available from OptionMetrics. We also estimate DECO, 10-Block DECO, and DCC using daily returns on Dow Jones stocks from 1995-2008. The first-stage model in all cases has no factors. Estimates are reported in Table 3b. Figure 5 plots the implied correlation against the average fitted pairwise correlation of each model. All three models broadly match the time series pattern of implied correlation. DECO seems to adjust more quickly and more dramatically during periods of sharp movements in the implied series. Implied correlations are almost always higher than model-based correlations, representing the correlation risk premium documented by Driessen, Maenhout and Vilkov (2009).

4.7 Out-of-Sample Hedging Performance

One way to evaluate the performance of DECO and DCC in an economically meaningful way is to use out-of-sample covariance forecasts to form minimum variance portfolios. A superior forecasting model should provide portfolios with lower variance than portfolios formed based on competing models. This type of comparison is motivated by the well-known mean-variance optimization setting of Markowitz (1952). Consider a collection of n stocks with expected return vector $\boldsymbol{\mu}$ and covariance matrix $\boldsymbol{\Sigma}$. Two hedge portfolios of interest are the global minimum variance (GMV) portfolio and the minimum variance portfolio subject to achieving an expected return of at least q. The GMV portfolio weights are the solution to the problem

$$\min_{\boldsymbol{\omega}} \boldsymbol{\omega}'\boldsymbol{\Sigma}\boldsymbol{\omega} \quad \text{s.t. } \boldsymbol{\omega}'\boldsymbol{\iota} = 1.$$

The MV portfolio is found by solving this problem subject to the additional constraint $\boldsymbol{\omega}'\boldsymbol{\mu} \geq q$. The expressions for optimal weights are

$$\boldsymbol{\omega}_{GMV} = \frac{1}{A}\boldsymbol{\Sigma}^{-1}\boldsymbol{\iota} \quad \text{and} \quad \boldsymbol{\omega}_{MV} = \frac{C - qB}{AC - B^2}\boldsymbol{\Sigma}^{-1}\boldsymbol{\iota} + \frac{qA - B}{AC - B^2}\boldsymbol{\Sigma}^{-1}\boldsymbol{\mu} \tag{13}$$

where $A = \boldsymbol{\iota}'\boldsymbol{\Sigma}^{-1}\boldsymbol{\iota}$, $B = \boldsymbol{\iota}'\boldsymbol{\Sigma}^{-1}\boldsymbol{\mu}$ and $C = \boldsymbol{\mu}'\boldsymbol{\Sigma}^{-1}\boldsymbol{\mu}$.

We focus on two forecasting questions. The first is motivated by Elton and Gruber (1973), who demonstrate that minimum variance portfolio choices can be improved by averaging pairwise correlations within groups. Our question extends this idea to the conditional setting, and is linked to the question of best correlation fit structure to employ with ex post averaging. Once DECO is estimated, it can be used to form out-of-sample unrestricted pairwise correlation forecasts (as in DCC). These pairwise can then be used to form different fitted correlation structures by averaging pairwise correlation forecasts within blocks as discussed in Section 2.4 and outlined in Figure 1. Ultimately, DECO estimates can be used to construct correlation forecasts that are equicorrelated, block equicorrelated, or unrestricted. By varying the choice of correlation structure in our forecasts we can evaluate the portfolio choice benefit of averaging pairwise correlations in a conditional setting (while keeping the estimation structure fixed as basic DECO).

Our experiment proceeds as follows. Using daily returns of the S&P cross section for the five-year estimation window beginning in January 1995 and ending December 1999, we

1. Estimate first-stage factor volatility models for each stock

2. Use estimates of regression/volatility models to form one-step ahead volatility forecasts for each stock

3. Using de-volatized residuals from the first stage, estimate the second-stage correlation model

4. Use correlation model parameter estimates to forecast unrestricted pairwise correlations one step ahead

5. Conduct ex post averaging of pairwise correlations to achieve each of the following correlation forecast fit structures (i) unrestricted, (ii) 30 industry blocks, (iii) 10 industry blocks and (iv) a single block

6. Combine correlation forecasts for each fit structure with regression/volatility model estimates and forecasts to construct the full covariance matrix forecast

7. Plug the resulting covariance forecast into (13) to find optimal portfolio weights (This step also requires an estimate of mean return $\boldsymbol{\mu}$. We set $\boldsymbol{\mu}$ equal to the historical mean and choose $q = 10\%$ annually)

8. Record realized returns for portfolios based on forecasts.

One-step ahead forecasts and portfolio choices are made in this manner for the next 22 days. After 22 days, the second-stage model is reestimated and the new parameters are used to generate the one-day ahead forecasts for the next 22 periods and new out-of-sample portfolio returns are calculated. This is repeated until all data through December 2008 have been used. The result is a set of 2,263 out-of-sample GMV and MV portfolio returns for each model.

After completing the forecasting procedure and recording portfolio returns, we calculate the realized daily variance for each ex post correlation fit structure. A superior model will produce optimal portfolios with lower variance realizations. We can test the significance of differences between portfolio variances for different correlation fit structures with a Diebold-Mariano test between the vectors of squared returns for each method. These tests are also related to the tests of Engle and Colacito (2006).

Out-of-sample GMV portfolio standard deviations when the first-stage has no factors model are reported in the first column of Table 4. A 30-block fit structure generates the lowest variance GMV portfolio with a standard deviation of 0.0093. This improves over the next best ex post structure, which used 10 blocks and achieves a standard deviation of 0.0096. The difference is significant at the 0.1% one-sided significance level. The same result is found for MV portfolios.

We repeat the hedging experiment using the CAPM and Fama-French factor structures. When the CAPM is used, ex post averaging over 10 blocks takes over the lowest variance position, significantly outperforming the second best (unrestricted) structure at the 2.5% one-sided significance level. The 10 block fit structure also achieves the lowest variance MV portfolio, though its improvement is not statistically significant over the next best model. For the Fama-French model, the unrestricted fit becomes the superior structure and significantly outperforms (block) equicorrelated structures.

DECO's minimum variance portfolio results so far suggest that there can be hedging benefits to varying the block structure of correlations after estimation. These results, however, do not speak to the estimation ability of different correlation models. If the same exercises were repeated using different correlation models than DECO to estimate α and β, how would portfolio choices fare?

Table 5 reports the standard deviations of GMV and MV portfolios when 10-Block DECO and DCC are used for estimation. For ease of reference and to conduct a new set of hypothesis tests, we reproduce DECO's results from Table 4. The table has a column for each first-stage factor model. Within a column, we report portfolio standard deviations achieved across different estimation and fit structures for correlations. Then, within columns (that is, for a given factor structure), Diebold-Mariano tests are performed to compare the variances achieved by DECO and Block DECO against the base case of DCC, holding the correlation fit structure fixed. For instance, the variance of DECO with an unrestricted fit structure is tested against DCC with an unrestricted structure, and 10-Block DECO with a 30-block fit is compared against DCC with a 30-block fit. In this way we can test whether (Block) DECO chooses significantly superior portfolios, holding the correlation fit structures fixed.

When the first-stage model includes no factors, the best overall GMV portfolio performance is achieved by estimating DECO, then using parameter estimates to construct 30-block correlation matrices ex post. This superior performance is statistically significant compared to DCC with 30-block fit structure. Block DECO also manages to significantly outperform DCC when the ex post structure has 30 blocks or is unrestricted. When a market factor or Fama-French factors are included, DECO continues to be the best estimation model, though its outperformance loses significance in the Fama-French case. The results are essentially the same for MV portfolios, with the exception that block-DECO becomes the best estimator for the first-stage model without a factor.

What qualitative assessments can we make from this analysis? First, the results highlight that tailoring the ex post block structure, regardless of the correlation estimator, can provide substantial improvement in hedging performance, corroborating evidence from a long strand of literature on unconditional portfolio choice. Second, it appears that (Block) DECO, besides offering a relatively good in-sample fit of the data as shown in Table 3a, provides statistically superior out-of-sample correlation forecasts compared with DCC. We interpret this as evidence that DECO can be a valuable way of dealing with noisy pairwise correlations during estimation.

5 Conclusion

DECO represents a major simplification in modeling time varying conditional covariance matrices for returns of an arbitrary number of assets. The equicorrelation assumption can be used to reduce noise and improve portfolio selection procedures, and it is a simplifying assumption that arises naturally in a variety of financial contexts. We extend the model to accommodate equicorrelated blocks which can be used to ease the restrictiveness of DECO while maintaining its simplicity and robustness. DECO and Block DECO are valuable models in the presence of non-equicorrelated variables. We prove quasi-maximum likelihood results that ensure (Block) DECO is a consistent estimator even when the true correlation process follows DCC. The theoretical properties of DECO are confirmed in experiments using simulated systems. For constituents of the S&P 500 Index, estimates show that DECO provides the a superior fit in the sense of information criteria relative to DCC. A test of forecasting performance shows that DECO can be used to construct out-of-sample hedge portfolios with significantly lower variance than those based on DCC.

Acknowledgements

The authors would like to thank two anonymous referees, the associate editor, and editors Arthur Lewbel and Jonathan Wright for helpful comments. Kelly's research was supported in part by the Neubauer Family Foundation.

Table 1. Monte Carlo with Equicorrelated Generating Process.

Using the DECO model of Equations 7-10, return data for 10, 30 or 100 assets are simulated over 1,000 or 5,000 periods using a range of values for α and β. Then, DECO is estimated with maximum likelihood and DCC is estimated using the (pairwise) composite likelihood of Engle, Shephard and Sheppard (2008). Simulations are repeated 2,500 times and summary statistics are calculated. The table reports the mean, median and standard deviation of α and β estimates, as well as their mean quasi-maximum likelihood asymptotic standard error estimates. Asymptotic standard errors are calculated using the "sandwich" covariance estimator of Bollerslev and Wooldridge (1992). We also calculate the root mean squared error (RMSE) for the true versus fitted average pairwise correlation process and report the average RMSE over all simulations. Correlation targeting is used in all cases, thus the intercept is the same for both models and not reported.

		DECO			Composite Likelihood			DECO			Composite Likelihood		
		α	β	RMSE	α	β	RMSE	α	β	RMSE	α	β	RMSE
		T=1000						T=5000					
		α = 0.10, β = 0.80											
n=10	Mean	0.100	0.785	0.015	0.016	0.685	0.053	0.099	0.797	0.007	0.014	0.865	0.051
	Median	0.099	0.793		0.015	0.824		0.099	0.799		0.014	0.876	
	Mean ASE	0.025	0.061		0.007	0.073		0.011	0.025		0.003	0.037	
	Std. Dev.	0.025	0.064		0.009	0.308		0.011	0.024		0.004	0.075	
n=30	Mean	0.098	0.793	0.010	0.009	0.684	0.045	0.099	0.799	0.005	0.009	0.870	0.044
	Median	0.098	0.797		0.009	0.810		0.098	0.800		0.009	0.879	
	Mean ASE	0.019	0.047		0.004	0.070		0.009	0.019		0.002	0.034	
	Std. Dev.	0.020	0.049		0.005	0.295		0.009	0.019		0.002	0.062	
n=100	Mean	0.099	0.796	0.008	0.007	0.718	0.041	0.099	0.799	0.004	0.007	0.876	0.040
	Median	0.099	0.797		0.007	0.811		0.099	0.800		0.007	0.879	
	Mean ASE	0.013	0.030		0.002	0.053		0.006	0.013		0.001	0.028	
	Std. Dev.	0.013	0.030		0.004	0.249		0.006	0.013		0.002	0.038	
		α = 0.05, β = 0.93											
n=10	Mean	0.050	0.919	0.015	0.012	0.709	0.053	0.050	0.929	0.007	0.009	0.955	0.049
	Median	0.049	0.926		0.011	0.927		0.050	0.929		0.009	0.966	
	Mean ASE	0.014	0.026		0.006	0.044		0.006	0.009		0.002	0.008	
	Std. Dev.	0.014	0.050		0.007	0.371		0.006	0.009		0.003	0.092	
n=30	Mean	0.049	0.924	0.010	0.007	0.706	0.045	0.049	0.930	0.005	0.006	0.960	0.044
	Median	0.049	0.927		0.007	0.922		0.049	0.930		0.006	0.968	
	Mean ASE	0.011	0.019		0.004	0.038		0.005	0.007		0.001	0.007	
	Std. Dev.	0.011	0.022		0.004	0.364		0.005	0.007		0.002	0.074	
n=100	Mean	0.049	0.928	0.008	0.005	0.741	0.042	0.049	0.930	0.004	0.005	0.966	0.041
	Median	0.049	0.928		0.005	0.926		0.049	0.930		0.005	0.968	
	Mean ASE	0.007	0.012		0.002	0.026		0.003	0.005		0.001	0.004	
	Std. Dev.	0.007	0.013		0.003	0.344		0.003	0.005		0.001	0.025	
		α = 0.02, β = 0.97											
n=10	Mean	0.022	0.928	0.015	0.008	0.548	0.035	0.020	0.969	0.007	0.005	0.824	0.033
	Median	0.020	0.963		0.006	0.704		0.020	0.969		0.004	0.980	
	Mean ASE	0.010	0.029		0.007	0.060		0.004	0.007		0.002	0.012	
	Std. Dev.	0.011	0.142		0.007	0.409		0.004	0.007		0.003	0.335	
n=30	Mean	0.021	0.953	0.011	0.004	0.490	0.030	0.020	0.969	0.005	0.003	0.835	0.029
	Median	0.020	0.966		0.003	0.545		0.020	0.970		0.003	0.980	
	Mean ASE	0.008	0.020		0.005	0.050		0.003	0.005		0.001	0.009	
	Std. Dev.	0.008	0.082		0.004	0.406		0.003	0.005		0.002	0.324	
n=100	Mean	0.020	0.966	0.008	0.002	0.486	0.027	0.020	0.970	0.004	0.002	0.902	0.026
	Median	0.020	0.968		0.002	0.547		0.020	0.970		0.002	0.981	
	Mean ASE	0.005	0.010		0.002	0.033		0.002	0.003		0.001	0.006	
	Std. Dev.	0.005	0.031		0.002	0.400		0.002	0.003		0.001	0.251	

Table 2. Monte Carlo with Non-Equicorrelated Generating Process.

Using the DCC model of Equations 7 and 8, return data for 10, 30 or 100 assets are simulated over 1,000 or 5,000 periods using a range of values for α and β. Then, DECO is estimated with maximum likelihood and DCC is estimated using the (pairwise) composite likelihood of Engle, Shephard and Sheppard (2008). Simulations are repeated 2,500 times and summary statistics are calculated. The table reports the mean, median and standard deviation of α and β estimates, as well as their mean quasi-maximum likelihood asymptotic standard error estimates. Asymptotic standard errors are calculated using the "sandwich" covariance estimator of Bollerslev and Wooldridge (1992). We also calculate the root mean squared error (RMSE) for the true versus fitted average pairwise correlation process and report the average RMSE over all simulations. Correlation targeting is used in all cases, thus the intercept is the same for both models and not reported.

		DECO			Composite Likelihood			DECO			Composite Likelihood		
		α	β	RMSE	α	β	RMSE	α	β	RMSE	α	β	RMSE
		T=1000						T=5000					
		α = 0.10, β = 0.80											
n=10	Mean	0.104	0.762	0.015	0.101	0.789	0.009	0.100	0.795	0.007	0.100	0.798	0.004
	Median	0.100	0.786		0.100	0.791		0.100	0.797		0.100	0.798	
	Mean ASE	0.038	0.098		0.008	0.020		0.016	0.036		0.004	0.008	
	Std. Dev.	0.038	0.128		0.012	0.027		0.016	0.036		0.005	0.011	
n=30	Mean	0.108	0.754	0.013	0.102	0.787	0.006	0.101	0.795	0.006	0.101	0.797	0.003
	Median	0.101	0.790		0.102	0.788		0.100	0.797		0.101	0.798	
	Mean ASE	0.047	0.123		0.005	0.011		0.020	0.045		0.002	0.005	
	Std. Dev.	0.049	0.154		0.007	0.016		0.019	0.045		0.003	0.007	
n=100	Mean	0.115	0.741	0.016	0.102	0.788	0.006	0.102	0.791	0.008	0.101	0.798	0.003
	Median	0.104	0.793		0.102	0.789		0.100	0.798		0.101	0.798	
	Mean ASE	0.056	0.133		0.002	0.006		0.025	0.057		0.001	0.003	
	Std. Dev.	0.059	0.184		0.006	0.013		0.025	0.060		0.003	0.006	
		α = 0.05, β = 0.93											
n=10	Mean	0.052	0.905	0.015	0.051	0.922	0.009	0.050	0.928	0.007	0.050	0.929	0.004
	Median	0.050	0.923		0.051	0.922		0.050	0.929		0.050	0.929	
	Mean ASE	0.021	0.042		0.005	0.008		0.009	0.014		0.002	0.003	
	Std. Dev.	0.022	0.095		0.006	0.010		0.008	0.014		0.003	0.004	
n=30	Mean	0.056	0.889	0.013	0.054	0.918	0.007	0.050	0.927	0.006	0.051	0.928	0.003
	Median	0.051	0.921		0.054	0.918		0.050	0.929		0.051	0.928	
	Mean ASE	0.028	0.057		0.003	0.004		0.011	0.017		0.001	0.002	
	Std. Dev.	0.030	0.131		0.004	0.007		0.011	0.019		0.002	0.003	
n=100	Mean	0.065	0.877	0.016	0.054	0.918	0.007	0.052	0.926	0.008	0.051	0.928	0.003
	Median	0.055	0.919		0.054	0.918		0.051	0.929		0.051	0.928	
	Mean ASE	0.036	0.073		0.001	0.002		0.014	0.022		0.001	0.001	
	Std. Dev.	0.041	0.148		0.003	0.005		0.014	0.026		0.001	0.002	
		α = 0.02, β = 0.97											
n=10	Mean	0.026	0.888	0.014	0.023	0.954	0.011	0.020	0.967	0.007	0.021	0.968	0.004
	Median	0.021	0.959		0.022	0.956		0.020	0.968		0.021	0.968	
	Mean ASE	0.017	0.093		0.003	0.009		0.005	0.011		0.001	0.002	
	Std. Dev.	0.019	0.209		0.004	0.013		0.005	0.014		0.002	0.003	
n=30	Mean	0.029	0.878	0.012	0.027	0.942	0.009	0.021	0.966	0.006	0.022	0.965	0.003
	Median	0.023	0.957		0.027	0.943		0.020	0.969		0.022	0.965	
	Mean ASE	0.020	0.065		0.002	0.006		0.007	0.013		0.001	0.001	
	Std. Dev.	0.025	0.220		0.003	0.010		0.007	0.015		0.001	0.002	
n=100	Mean	0.037	0.860	0.014	0.027	0.942	0.010	0.022	0.964	0.007	0.022	0.965	0.003
	Median	0.027	0.951		0.027	0.943		0.021	0.969		0.022	0.965	
	Mean ASE	0.027	0.072		0.001	0.003		0.009	0.017		0.000	0.001	
	Std. Dev.	0.035	0.230		0.002	0.007		0.009	0.025		0.001	0.002	

Table 3a. Full Sample Correlation Estimates for S&P 500 Constituents, 1995-2008.
The table presents estimation results for nine dynamic covariance models. Each model is a two-stage quasi-maximum likelihood estimator and is a combination of one of three first-stage models with one of three second-stage models. The first-stage models are GARCH regression models imposing a factor structure for the cross section of returns, in which the structures are no factor (Panel A), the Sharpe-Lintner one-factor CAPM (Panel B) and the Fama-French (1993) three-factor model (Panel C). The second-stage correlation models, estimated on standardized residuals from the first stage, are one- and 10-Block DECO and composite likelihood DCC. Below each estimate we report quasi-maximum likelihood asymptotic standard errors in italics. Asymptotic standard errors are calculated using the "sandwich" covariance estimator of Bollerslev and Wooldridge (1992). For each model we report the Akaike information criterion calculated using the sum of the first- and second-stage log likelihoods penalized for the number of parameters in both stages. The analysis is performed on the S&P 500 data set described in Section 4.

	DECO	10-Block	DCC
Panel A: No Factor			
α	0.021	0.023	0.014
	0.003	*0.007*	*0.005*
β	0.979	0.975	0.975
	0.003	*0.008*	*0.034*
AIC	-6,802.0	-6,861.6	-6,735.0
Panel B: CAPM Market Factor			
α	0.006	0.008	0.005
	0.005	*0.010*	*0.004*
β	0.986	0.982	0.986
	0.016	*0.028*	*0.019*
AIC	-6,851.4	-6,921.9	-7,182.2
Panel C: Fama-French 3 Factors			
α	0.010	0.002	0.004
	0.005	*0.001*	*0.003*
β	0.804	0.998	0.984
	0.114	*0.001*	*0.029*
AIC	-6,899.2	-6,955.2	-7,270.2

Table 3b. Full Sample Correlation Estimates for Dow Jones Constituents, 1995-2008.
This table repeats the analysis of Table 3a, Panel A (no factor) for the subsample of 36 Dow Jones constituents.

	DECO	10 Block	DCC
α	0.034	0.023	0.019
	0.005	*0.010*	*0.005*
β	0.964	0.971	0.970
	0.005	*0.013*	*0.007*
AIC	-572.7	-581.3	-577.4

Table 4. DECO Out-of-Sample Minimum Variance Portfolio Comparison (by Ex Post Fit Structure). The table presents the results of an out-of-sample portfolio formation experiment to test covariance forecasting ability. The following procedure is used to create sequential, non-overlapping covariance forecasts, which are then used to form portfolios. First, the covariance model is estimated using a cross section of daily returns for the five-year estimation window beginning January 1995 and ending December 1999, and one-step ahead covariance forecasts for the next 22 days are formed. Using each day's forecast, we construct the global minimum variance (GMV) portfolio and the minimum variance (MV) portfolio subject to a 10% required annual return (see Section 4) and record the return for each portfolio that day. Data from the 22-day forecast period is then added to the estimation sample and the model is re-estimated. The new estimates are used to generate covariance forecasts for the subsequent 22 days and new out-of-sample portfolio returns are calculated. This is repeated until all data through December 2008 has been used. The result is a set of 2,263 out-of-sample GMV and MV portfolio returns. The table reports standard deviations of the portfolio return time series. Three different first-stage models, which are factor structures for the cross section of returns, are considered: 1) no factor, 2) the Sharpe-Lintner CAPM and 3) the Fama-French (1993) three-factor model. The second-stage model is DECO, estimated on standardized residuals from the first stage. Portfolios are formed with four different ex post correlation fit structures based on industry assignment, as described in Section 4. Within each column, the correlation forecasting method that achieves the lowest standard deviation portfolio is shown in bold. A Diebold-Mariano test is calculated for the significance of the best model against the second best model in the same column. The best model is accompanied by *, **, or *** if it achieves a lower standard deviation than the next best model at the 2.5%, 1.0% or 0.1% (one-sided) level, respectively. The analysis is performed on the S&P 500 data set described in Section 4.

		GMV			MV		
Estimation Structure	Fit Structure	No Factor	CAPM	FF3	No Factor	CAPM	FF3
DECO	Unrestricted	0.0117	0.0090	**0.0073** ***	0.0106	0.0086	**0.0073** ***
	30-Block	**0.0093** ***	0.0093	0.0091	**0.0088** ***	0.0090	0.0091
	10-Block	0.0096	**0.0090** **	0.0087	0.0090	**0.0086** *	0.0086
	1-Block	0.0101	0.0100	0.0091	0.0096	0.0098	0.0091

Table 5. Out-of-Sample Minimum Variance Portfolio Comparison (by Correlation and Factor Model). The table presents the results of an out-of-sample portfolio formation experiment to test covariance forecasting ability. The following procedure is used to create sequential, non-overlapping covariance forecasts, which are then used to form portfolios. First, a covariance model is estimated using a cross section of daily returns for the five-year estimation window beginning January 1995 and ending December 1999, and one-step ahead covariance forecasts for the next 22 days are formed. Using each day's forecast, we construct the global minimum variance (GMV) portfolio and the minimum variance (MV) portfolio subject to a 10% required annual return (see Section 4) and record the return for each portfolio that day. Data from the 22-day forecast period is then added to the estimation sample and the model is re-estimated. The new estimates are used to generate covariance forecasts for the subsequent 22 days and new out-of-sample portfolio returns are calculated. This is repeated until all data through December 2008 has been used. The result is a set of 2,263 out-of-sample GMV and MV portfolio returns. The table reports standard deviations of the portfolio return time series. Three different first-stage models, which are factor structures for the cross section of returns, are considered: 1) no factor, 2) the Sharpe-Lintner CAPM and 3) the Fama-French (1993) three-factor model. Three second-stage models, DECO, 10-Block DECO and DCC, are estimated on standardized residuals from the first stage. Portfolios are formed with four different ex post correlation fit structures based on industry assignment, as described in Section 4. Within each column, the correlation forecasting method that achieves the lowest standard deviation portfolio is shown in bold. A Diebold-Mariano test is calculated for the significance of difference between each structure/fit pair (for DECO and Block DECO) against the same fit structure based on DCC estimates. Values accompanied by *, **, or *** achieve a lower standard deviation than their DCC counterpart at the 2.5%, 1.0% or 0.1% (one-sided) level, respectively. The analysis is performed on the S&P 500 data set described in Section 4.

		GMV			MV		
Estimation Structure	Fit Structure	No Factor	CAPM	FF3	No Factor	CAPM	FF3
DECO	Unrestricted	0.0117	0.0090	**0.0073**	0.0106	0.0086	**0.0073**
	30-Block	**0.0093** ***	0.0093 **	0.0091 **	0.0088 *	0.0090 **	0.0091 ***
	10-Block	0.0096	**0.0090** **	0.0087 **	0.0090	**0.0086** ***	0.0086 **
	1-Block	0.0101	0.0100	0.0091	0.0096	0.0098	0.0091
10-Block DECO	Unrestricted	0.0112 *	0.0097	0.0077	0.0102	0.0091	0.0077
	30-Block	0.0094 *	0.0093 *	0.0091	**0.0088**	0.0090	0.0092
	10-Block	0.0096	0.0091	0.0087	0.0090	0.0087	0.0088
	1-Block	0.0101	0.0101	0.0091	0.0096	0.0099	0.0091
DCC	Unrestricted	0.0118	0.0091	0.0073	0.0105	0.0086	0.0073
	30-Block	0.0094	0.0093	0.0091	0.0088	0.0090	0.0091
	10-Block	0.0096	0.0091	0.0087	0.0089	0.0087	0.0087
	1-Block	0.0101	0.0100	0.0092	0.0096	0.0098	0.0091

Figure 1. Procedure for Generating Fitted Correlation Structures.
The schematic diagram summarizes how a correlation structure used as part of the maximum likelihood estimation procedure, what we call the "estimation structure," can differ from the "fit structure" of the correlation series eventually generated from the estimated model.

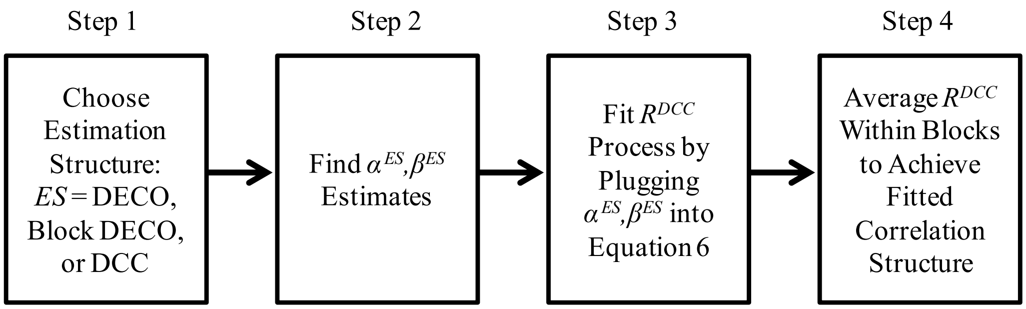

Figure 2. Fitted Dynamic Equicorrelation by Factor Model, S&P 500 Constituents, 1995-2008.
The figure shows fitted residual equicorrelations of S&P 500 constituents estimated with the DECO model (black line) and the S&P 500 Index level (grey area). Equicorrelation fits are based on model estimates in the first column of Table 3. The graphs correspond to the following factor schemes: a) no factor, b) the Sharpe-Lintner CAPM and c) the Fama-French (1993) three-factor model.

Figure 2a. S&P 500 Index Level vs. DECO with First-Stage Model with No Factor

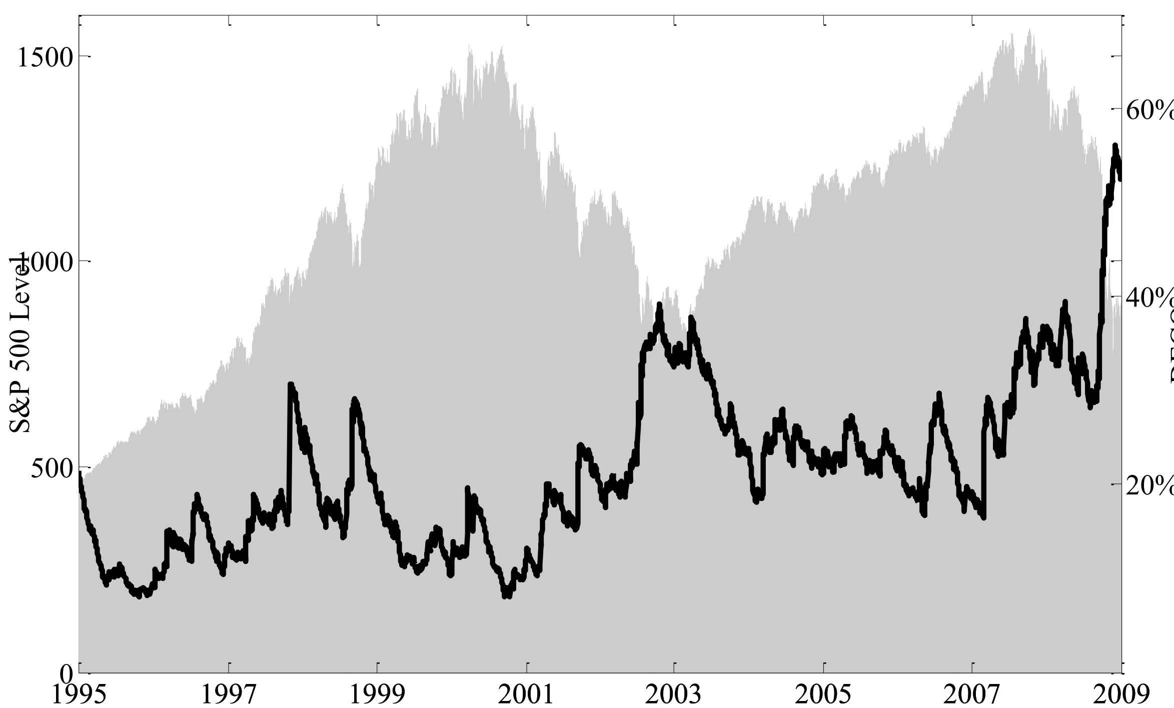

Figure 2. Continued.

Figure 2b. S&P 500 Index Level vs. DECO with First-Stage CAPM One-Factor Model

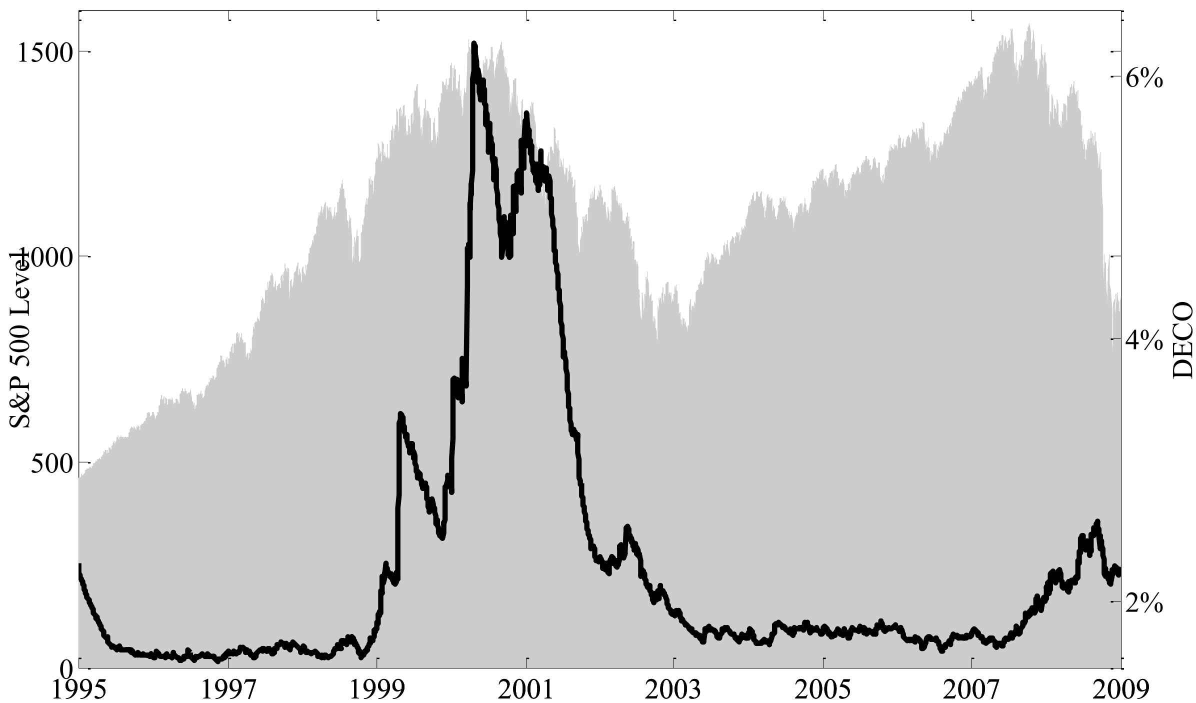

Figure 2c. S&P 500 Index Level vs. DECO with First-Stage Fama-French Three-Factor Model

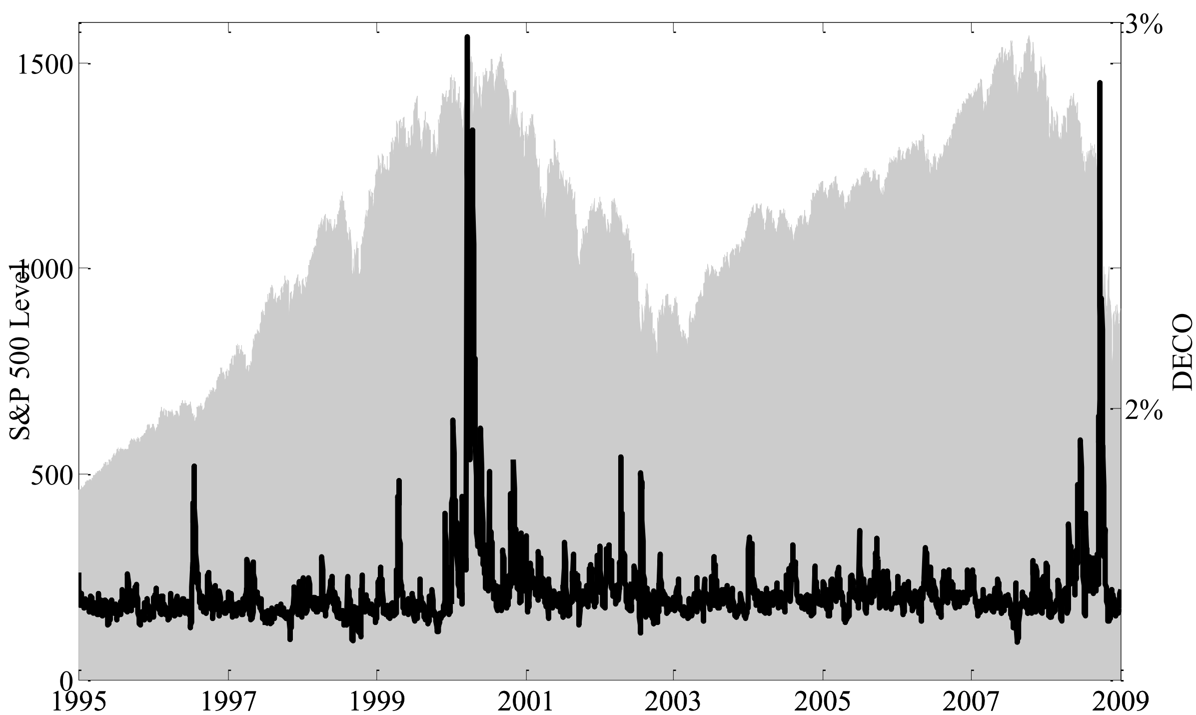

Figure 3. Fitted DECO and DCC Correlations, S&P 500 Constituents, 1995-2008.
The figure shows fitted correlations of S&P 500 constituents estimated with DECO and DCC. Correlation fits are based on model estimates in Table 3.The graphs correspond to the following first-stage factor schemes: a) no factor, b) the Sharpe-Lintner CAPM and c) the Fama-French (1993) three-factor model. Each plot shows the fitted one-block equicorrelation and the 25th, 50th and 75th percentile of pairwise DCC correlations in each period.

Figure 3a. DECO and DCC with First-Stage Model with No Factor

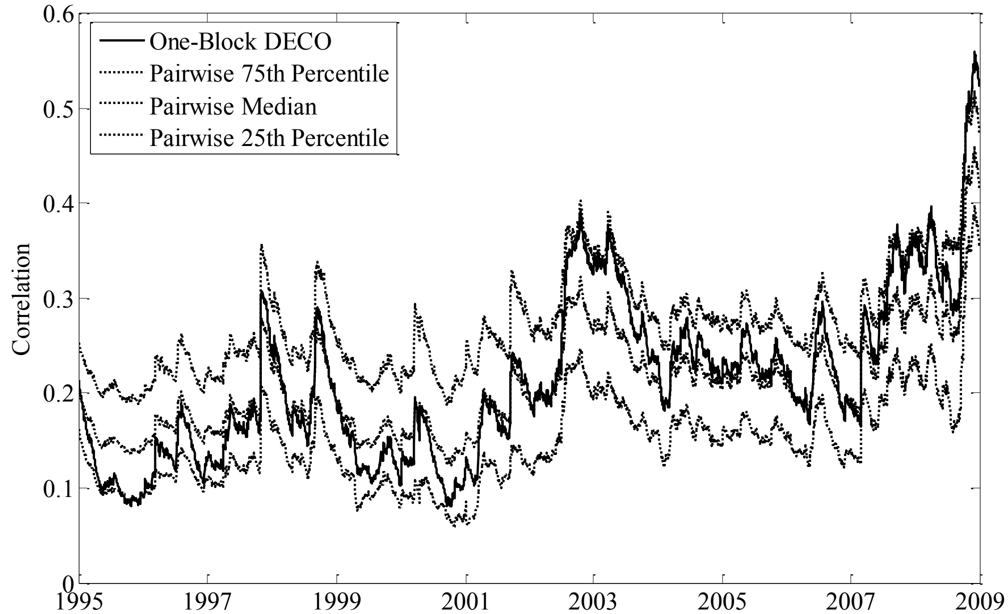

Figure 3b. DECO and DCC with First-Stage CAPM One-Factor Model

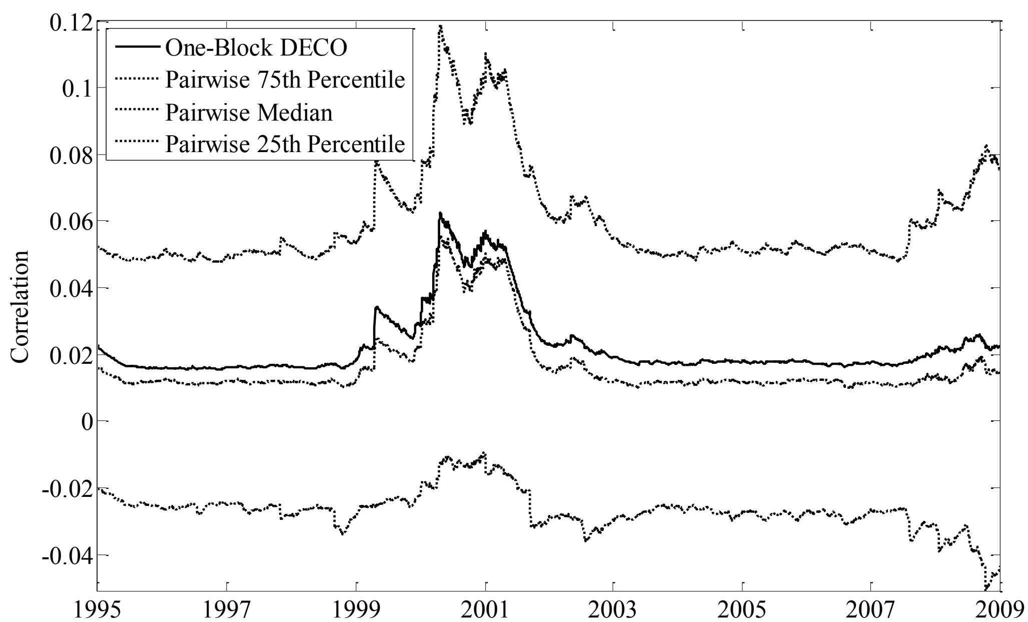

Figure 3. Continued

Figure 3c. DECO and DCC with First-Stage Fama-French Three-Factor Model

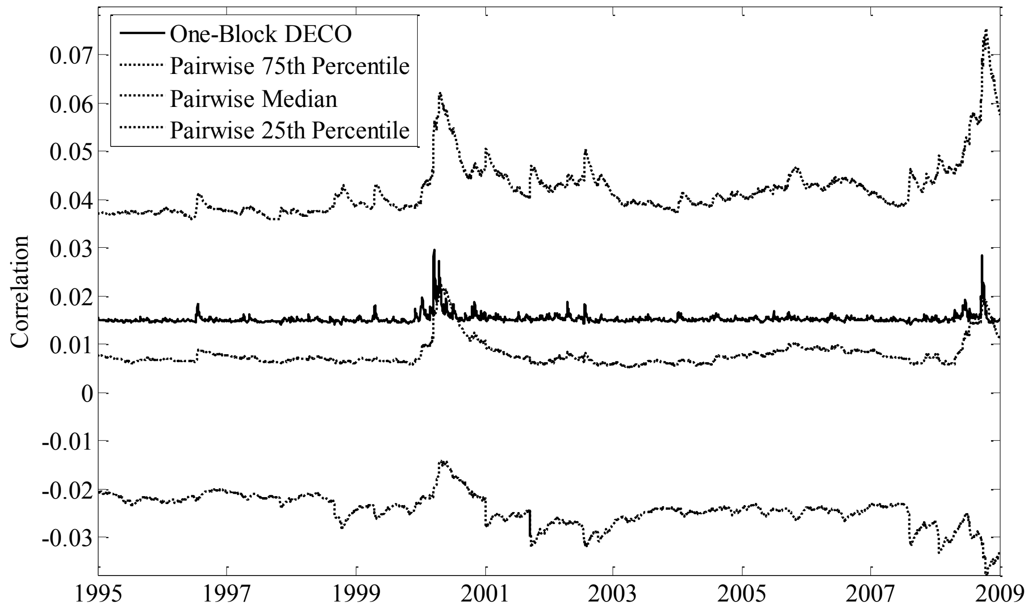

Figure 4. Selected Block Equicorrelations for the S&P 500, 1995-2008.
The figure shows within-block equicorrelations of energy, telecom and health stocks (according to the 10-industry assignments on Ken French's website) in our S&P 500 sample. Estimates are made with Block DECO composite likelihood using a first-stage model with no factor. Correlation fits correspond to model estimates in Table 3.

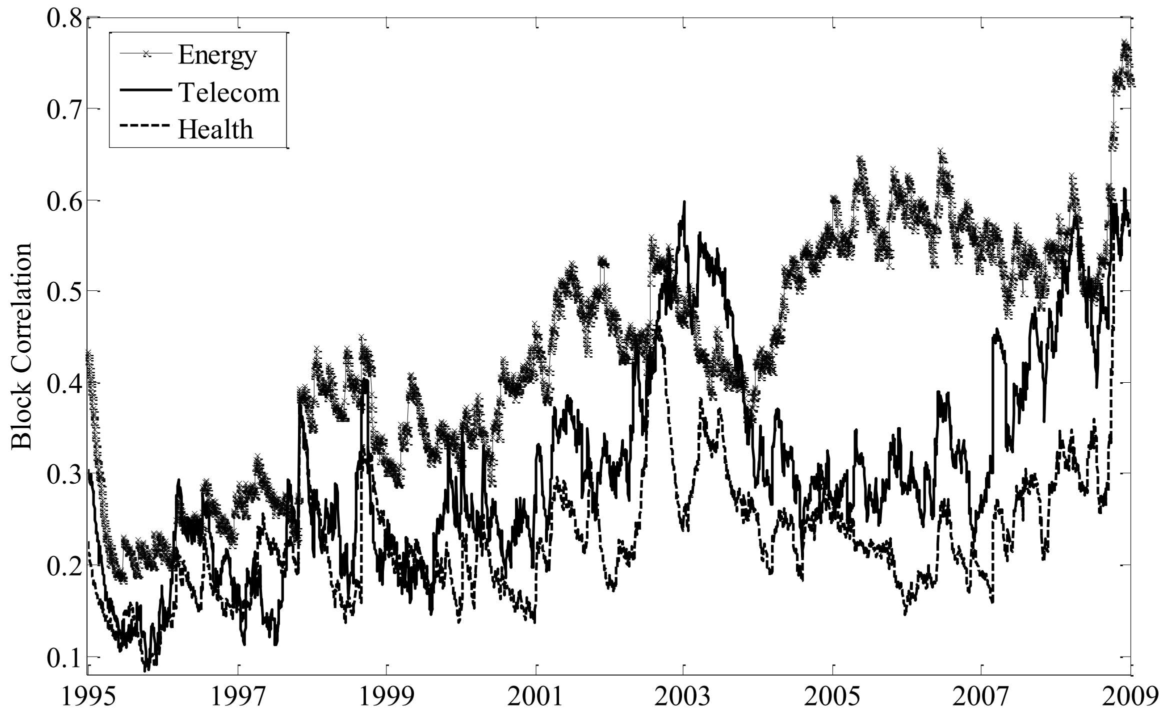

Figure 5. Dow Jones Index Option-Implied Correlations and Model-Based Average Correlations. The figure shows the option implied correlation of Dow Jones stocks and the average pairwise correlation estimated using DECO, 10-Block DECO and DCC. Model-based correlations correspond to parameter estimates shown in Table 3b. The options sample horizon covers October 1997 to September 2008, and the correlation model fits are estimated using data from January 1995 to December 2008.