A Multivariate Realized GARCH Model

Ilya Archakov[a] **Peter Reinhard Hansen**[c,d] **Asger Lunde**[b*]

[a] *University of Vienna*

[b] *CREATES, Aarhus University*

[c] *University of North Carolina*

[d] *Copenhagen Business School*

February 18, 2020

Abstract

We propose a novel class of multivariate Realized GARCH models that utilize realized measures of volatility and correlations. The key property of the model is a convenient parametrization of the correlation matrix that requires no additional structure to ensure positive definiteness. The correlation matrix is characterized by a vector, that can vary freely in the real vector space. A more parsimonious structure is often desired in practice, in particularly in high dimensional systems, and the framework facilitates simple and intuitive dimension reductions. We apply the model to returns of nine assets and illustrate a dimension reduction that arises from a natural block equicorrelation structure. Interestingly, we find that the empirical distribution of the transformed realized correlations is approximately Gaussian.

Keywords: Financial Volatility; Realized GARCH; High Frequency Data; Multivariate Modeling.

JEL Classification: G11, G17, C32, C58

[*]We acknowledge financial support by the Center for Research in Econometric Analysis of Time Series, CREATES, funded by the Danish National Research Foundation.

1 Introduction

Univariate GARCH models have had much empirical success since the ARCH model was introduced by Engle (1982). A large number of univariate GARCH-type models have been proposed in the literature, whereas the literature on multivariate GARCH models is less voluminous. It is not entirely obvious how the univariate GARCH structure is naturally generalized to higher dimensions. In a multivariate setting, the main object of interest is the conditional covariance matrix, $H_t = \mathrm{var}(r_t|\mathcal{F}_{t-1})$, and there are two main challenges to modeling H_t. First, a model must produce a positive (semi) definite matrix, and this entails nonlinear cross restrictions on the elements of H_t. Second, the number of covariance terms for a vector of dimension p is proportional to p^2, and this can become computationally difficult, unless p is relatively small. Many multivariate GARCH models impose a structure on H_t that serves to address these two issues. 1) Ensure that H_t is positive semidefinite; and 2) Make the model parsimonious. A convenient way to model H_t, which we adopt in this paper, is to separately model variances and correlations. This was the structure in the seminal paper by Bollerslev (1990), who proposed the constant correlation model. It is also the structure in the Dynamic Conditional Correlation (DCC) model by Engle (2002a), see also Engle and Sheppard (2001), Shephard et al. (2020), Aielli (2013), and Engle et al. (2019). In our empirical section we adopt the correlation structure used in Engle and Kelly (2012). The proposed the dynamic equicorrelation (DECO) models as well as the dynamic block equicorrelation model (block-DECO). The structure of the block-DECO arises as a special case in our framework. However, we model the correlations in a a novel way that guarantees positive definiteness, and in our empirical section we make use of realized measures of volatility. We also establish theoretical results that contributes to the existing literature on block-DECO models.

The main contributions in this paper are the following. First, we develop a new class of multivariate GARCH models that facilitates a flexible modeling of the correlation structure, while positive definiteness is assured without further constraints. Second, we develop a framework for dimension reduction for large correlation matrices. This dimension reduction is demonstrated for the case with block equicorrelation matrices. Specifically, we show that a block equicorrelation structure is equivalent to a particular factor structure. A block equicorrelation structure may be motivated by sector classifications, or some other appropriate categorization, of the assets being modeled. Third, we contribute to the exiting literature on block-DECO models in Theorem 2. We present simple expressions for the inverse correlation matrix of any (invertible) block equicorrelation correlation matrix, as well as its determinant. The results in Theorem 2 applies to block equicorrelation matrices with an arbitrary number of blocks. Previously, the only known expression for the inverse was for the case with two blocks, see Engle and Kelly (2012, lemma 2.3). Theorem 2 also characterizes the eigenvalues of any block equicorrelation matrix, and consequently, it fully characterizes the set of positive definite block equicorrelation matrices. Fourth, we demonstrate the usefulness of the framework in an empirical application. Specifically, we apply the block equicorrelation structure to nine asset returns from three different sectors. The new framework improves the empirical fit of the vector of returns both in-sample and out-of-sample, relative to simpler models such as the constant correlation model, Moreover, the predicted covariance matrices may be used for optimal portfolio construction, and we find that portfolio variance is reduced by a factor of two relative to a naive equal weights portfolio.

Another contribution of our empirical analysis is evidence that the transformed realized correlations, $\hat{\varrho}$ (defined below), are approximately Gaussian distributed. This result is analogous to existing results for the logarithmically transformed realized variances, see Andersen et al. (2001a). Combined, these results provide justification for the Gaussian specification used to define the likelihood function for the multivariate Realized GARCH model.

The new class of multivariate GARCH models is based on a convenient parametrization of the conditional correlation matrix, C_t, that corresponds to H_t. The parametrization is given by the vector $\varrho_t = g(C_t)$, where the mapping $\varrho = g(C)$ is defined by taking the matrix logarithm to the correlation matrix, C, and stacking the off-diagonal elements of $\log C$ into the vector $\varrho \in \mathbb{R}^d$, where $d = p(p-1)/2$. This parametrization was recently introduced by Archakov and Hansen (2020), who showed that it has many interesting properties. For instance, this mapping is one-to-one between the set of non-singular correlation matrices and $\mathbb{R}^d$, so regardless of how the vector ϱ_t is modeled or regulated, it will always map back to a unique positive definite correlation matrix. In other words, this parametrization guarantees positive definiteness, but need not impose any additional restrictions on H_t. It is, however, straight forward to impose additional structure on the correlation matrix if required. A more parsimonious structure on H_t will typically be needed unless p is small, and this can be achieved with a factor structure for ϱ_t, as we will show in Section 2.3. A situation where a factor structure for ϱ_t emerges naturally, is when the correlation matrix has a block structure with identical correlations within blocks.

Early multivariate GARCH models solely relied on daily returns in the dynamic of volatility, we will make use of realized measures of volatilities and correlations. This is beneficial, because the realized measures provide accurate signals about volatility, which is valuable in the dynamic modeling of conditional variances and correlations. The use of realized measures was popularized by results in Andersen and Bollerslev (1998), and key theoretical results were subsequently established in Andersen et al. (2001b), Barndorff-Nielsen and Shephard (2002a), Andersen et al. (2003), Barndorff-Nielsen and Shephard (2004a), see also Hansen and Lunde (2011) and references therein. Realized measures were initially used to evaluate and compare the performance of GARCH models, see Andersen and Bollerslev (1998). A natural next step was to incorporate realized measures in GARCH models and Engle (2002b) was one of the first to include realized measures as an exogenous variable in GARCH models. Complete models, that also specify a model for the realized measures, soon followed, including the MEM by Engle and Gallo (2006), the HEAVY model by Shephard and Sheppard (2010), and the Realized GARCH model by Hansen et al. (2012). Multivariate extensions of these models were proposed in Noureldin et al. (2012), Hansen et al. (2014), Dumitrescu and Hansen (2017), and Gorgi et al. (2019). Another way to incorporated realized measures in multivariate GARCH models is explored in Bauwens et al. (2012), who build on the Conditional Autoregressive Wishart model of Golosnoy et al. (2012).

Our approach to modeling the correlation structure could, with some adaptation, be implemented in a conventional manner, using daily returns only. However, it is advantageous to include realized measures in the modeling. Realized measures are computed from high frequency data, and these provide accurate signals about the key quantities of the model. Therefore, including realized measures in GARCH models make them more responsive to sudden changes in volatility and correlations, which improve the empirical fit and model predictions, see e.g. Hansen and Huang (2016). The proposed framework makes it easy to incorporate realized measures of volatility in the modeling. This part of the model builds on the Realized GARCH framework of Hansen et al. (2012), and the proposed model is the first multivariate generalization of Realized GARCH framework that does need not impose any restrictions on the covariance structure. Moreover, the realized GARCH framework admits a dependence between volatility shocks and returns shocks, that is well-documented in the empirical literature. This dependence is commonly referred to as the leverage effect.

The parametrization of the correlation matrix, C_t, involves the matrix logarithm of C_t. The matrix logarithm has previously been used in the modeling of covariance matrices in Chiu et al. (1996). In the context of the multivariate GARCH it was used in Kawakatsu (2006) and Asai and So (2015) applied it to the DCC model. The transformation has also been used in stochastic volatility models, see Ishihara et al. (2016), and in reduced-form models of realized covariance matrices, see e.g. Bauer and Vorkink (2011) and Weigand (2014).[1] Here we apply the matrix logarithm to the correlation matrix, which differs from applying it to the covariance matrix in important ways, see Archakov and Hansen (2020). In the present context, it enables us to model the conditional variances separately from the conditional correlations using a familiar GARCH structure for each of the univariate conditional variances. Moreover, this model structure enables us to explicitly model the empirically important leverage effect.

We proceed as follows. In Section 2 we introduce notation, present the modeling framework, and discuss how a factor structure can be imposed on the correlation matrix. We also derive results that are specific to block equicorrelation matrices. Section 3 details the estimation of the model and how the model can be used for forecasting. An extensive empirical analysis of nine asset returns from three sectors is presented in Section 4. We conclude in Section 5 and complete the paper with three appendices with proofs, step-by-step estimation instructions, and additional empirical results, respectively.

2 The Multivariate Realized GARCH Model

In this section we present the multivariate GARCH model that can utilize realized measures of variances and correlations. The key novelty in this model is the way in which the correlation structure is modeled. We apply a convenient vector parametrization of the correlation matrix, $\varrho_t = g(C_t)$ that does not impose any structure beyond positive definiteness of C_t. The formulation allows for additional structure to be imposed, for instance by modeling the vector ϱ_t with a factor structure. We explore, in details, one particular way to impose a factor structure on C_t, which is motivated by a block equicorrelation structure.

2.1 Notation and Preliminaries

We let r_t denote a p-dimensional vector of returns in period t, where t represents a generic unit of time – typically a trading day. The conditional mean is denoted by $\mu_t = \mathbb{E}(r_t|\mathcal{F}_{t-1})$ and the conditional variance by

$$H_t = \text{var}(r_t|\mathcal{F}_{t-1}),$$

where $\{\mathcal{F}_t\}$ is the natural filtration for (r_t, RM_t). Here RM_t denotes an ex-post empirical measure of H_t, such as the (multivariate) realized variance, see Barndorff-Nielsen and Shephard (2004b), or

[1] Additional related literature include the work by Liu (2009), Chiriac and Voev (2011), Golosnoy et al. (2012), and Bauwens et al. (2012).

the multivariate realized kernel, see Barndorff-Nielsen et al. (2011).

Following Engle (2002a) we decompose the conditional covariance matrix into variances and correlations,

$$H_t = \Lambda_{h_t}^{1/2} C_t \Lambda_{h_t}^{1/2}, \tag{1}$$

where $\Lambda_{h_t} = \mathrm{diag}(h_{1,t}, \ldots, h_{p,t})$ with $h_{i,t} = [H_t]_{ii}$, $i = 1, \ldots, p$. Thus $h_{i,t}$ is the conditional variance of $r_{i,t}$ (the i-th element of r_t) and $C_t = \mathrm{corr}(r_t|\mathcal{F}_{t-1})$ is the conditional correlation matrix of r_t. The structure in (1) is the basis for the Dynamic Conditional Correlation framework, see Engle (2002a) and Engle and Sheppard (2001). This formulation enables us to disentangle the dynamic properties of the conditional variances from that of the conditional correlation. In contrast to the conventional DCC model, we will incorporate realized measures of variances and correlations into the modeling, and we employ a different parametrization of C_t, which we detail below.

The central element in GARCH model is the equation that specifies the dynamic properties of H_t and how these are influenced by lagged returns. This equation can be enhanced to make use of realized measures of volatility. The Realized GARCH models are characterized by a measurement equation that specifies how the contemporaneous realized measures related to the conditional moments, such as H_t.

In this paper, we will split the modeling of H_t in two parts: The modeling of the conditional variances and the modeling of the conditional correlations. This leads to two sets of GARCH and measurement equations, that both utilize appropriate realized measures. From the $p \times p$ empirical measure of the covariance matrix in period t, RM_t, we extract the diagonal elements $x_t = \mathrm{diag}(\mathrm{RM}_t)$ and the corresponding correlation matrix denoted by

$$Y_t = \Lambda_{x_t}^{-1/2} \mathrm{RM}_t \Lambda_{x_t}^{-1/2}.$$

Here $\Lambda_{x_t} = \mathrm{diag}(x_{1,t}, \ldots, x_{p,t})$ denotes the diagonal matrix with the elements of x_t on the diagonal. The realized measure of the covariance matrix, RM_t, will be assumed to be positive definite so that Y_t is a correlation matrix with $\det Y_t > 0$. In summary, x_t and Y_t are the observed empirical measures of the latent variables, h_t and C_t, respectively. The realized measure, RM_t, will typically be consistent for the quadratic variation. The quadratic variation is not identical to the conditional variance, H_t, so we will need to entertain non-trivial measurement errors in how x_t and Y_t relate to h_t and C_t, respectively.

We let I_s denote the $s \times s$ identity matrix, let $1_{\{\cdot\}}$ denotes the indicator function, which is one if the expression within the curly brackets is true, and zero otherwise.

2.1.1 Parametrizing the Correlation Matrix

An invertible matrix, C_t, needs to satisfy two properties in order to be a proper correlation matrix: It must be positive definite and each of its diagonal elements must be equal to one. In some models these two requirements are satisfied by imposing a structure that implies even stronger conditions. For instance, an equicorrelation structure with the common correlation confined to the interval $(-\frac{1}{p-1}, 1)$.

In this paper we adopt a vector representation of the correlation matrix that was proposed by Archakov and Hansen (2020). This parametrization is based on the following mapping:

$$\varrho = g(C) = \mathrm{vecl}(\log C),$$

where $\log C = \sum_{k=1}^{\infty}(-1)^k(C-I)^k/k$ is the matrix logarithmically transformed correlation matrix, and $\mathrm{vecl}(\cdot)$ extracts and vectorizes the elements below the diagonal. To illustrate this parametrization, consider the following example,

$$g(\begin{bmatrix} 1.0 & \bullet & \bullet \\ 0.8 & 1.0 & \bullet \\ 0.0 & 0.2 & 1.0 \end{bmatrix}) = \mathrm{vecl}(\log \begin{bmatrix} 1.0 & \bullet & \bullet \\ 0.8 & 1.0 & \bullet \\ 0.0 & 0.2 & 1.0 \end{bmatrix}) = \mathrm{vecl}(\begin{bmatrix} -0.53 & \bullet & \bullet \\ 1.14 & -0.57 & \bullet \\ -0.13 & 0.28 & -0.03 \end{bmatrix}) = \begin{pmatrix} 1.14 \\ -0.13 \\ 0.28 \end{pmatrix}.$$

In Archakov and Hansen (2020), it is shown that $g(C)$ is a one-to-one mapping between $\mathcal{C}_{p\times p}$ and $\mathbb{R}^d$, where $\mathcal{C}_{p\times p}$ is the space of positive definite $p\times p$ correlation matrices and $d = p(p-1)/2$. Thus, a non-singular $p\times p$ correlation matrix can be represented and modeled as a vector in $\mathbb{R}^{p(p-1)/2}$. In the bivariate case, $p = 2$, where $C_{11} = C_{22} = 1$ and $C_{12} = C_{21} = \rho$, it can be shown that $g(C) = \frac{1}{2}\log\frac{1-\rho}{1+\rho}$, which is the Fisher transformation of the correlation coefficient, ρ. In this case, it is straight forward to invert the transformation, because $\rho = (e^{2\varrho}-1)/(e^{2\varrho}+1)$. An algorithm for the inverse mapping, $C = g^{-1}(\varrho)$, for an arbitrary vector, $\varrho \in \mathbb{R}^{p(p-1)/2}$ is given in Archakov and Hansen (2020).

Having introduced the transformed conditional correlations, $\varrho_t = g(C_t)$, we denote the corresponding transformed realized correlations by,

$$y_t = g(Y_t).$$

Both ϱ_t and y_t are $p(p-1)/2$-dimensional vectors. Instead of modeling the conditional correlation matrix, C_t, that is subject to positive (semi-) definite constraints, we will model the vector ϱ_t

that varies freely in $\mathbb{R}^{p(p-1)/2}$. The corresponding empirical measures are subjected to the same transformation, and y_t is the empirical measure that we use to model the dynamic properties of ϱ_t. Much like the way that squared returns are used to model the dynamic of the conditional variance in a standard GARCH model.

The model presented above, can be viewed as a natural generalization of the bivariate structure in Hansen et al. (2014), because the two models coincide when $p = 2$, in which case ϱ_t emerges as the Fisher transformed correlation. The approach in Hansen et al. (2014) was to construct a multivariate model (for dimensions higher than 2) by fusing bivariate models to form a larger system. The fusing of bivariate models implies a restricted single-factor structure on C_t, that is not imposed in this paper.

2.2 The Model

With the required notation in place, we are now ready to introduce the multivariate realized GARCH model which consists of return equations, GARCH equations, and measurement equations. The return equation for the i-th asset at times t takes the form

$$r_{i,t} = \mu_i + h_{i,t}^{1/2} z_{i,t}, \qquad i = 1, \dots, p, \quad t = 1, \dots, T,$$

where we have assumed that $\mathbb{E}(r_{i,t}|\mathcal{F}_{t-1})$ is constant, as is often the case in GARCH models. It follows that the standardized return $z_{i,t} = h_{i,t}^{-1/2}(r_{i,t} - \mu_i)$ is such that $\mathbb{E}(z_{i,t}|\mathcal{F}_{t-1}) = 0$ and $\mathrm{var}(z_{i,t}|\mathcal{F}_{t-1}) = 1$. However, the standardized returns, $z_t = \Lambda_{h_t}^{-1/2}(r_t - \mu)$, are not uncorrelated since $\mathrm{var}(z_t|\mathcal{F}_{t-1}) = C_t$. In our likelihood analysis, we will specify the distributional properties of z_t, $t = 1, \dots, T$, to be distributed as i.i.d.$N(0, C_t)$.

Next, we specify the GARCH equations state how H_t depends on past observable variables, and we make use of lagged values of both returns and realized measures. The dynamics for the vector of conditional variances and the vector representation of the conditional correlations are as follows:

$$\begin{aligned} \log h_t &= \omega + \beta \log h_{t-1} + \tau(z_{t-1}) + \gamma \log x_{t-1}, \\ \varrho_t &= \tilde{\omega} + \tilde{\beta} \varrho_{t-1} \qquad\qquad\qquad\quad + \tilde{\gamma} y_{t-1}, \end{aligned} \tag{2}$$

where ω is an $p \times 1$ vector, β and γ are $p \times p$ matrices, $\tau(\cdot)$ is an leverage function that we elaborate on below. Similarly $\tilde{\omega}$ is an $d \times 1$ vector and $\tilde{\beta}$ and $\tilde{\gamma}$ are $d \times d$ matrices with $d = p(p-1)/2$.

Realized GARCH models are characterized by measurement equations that relate latent variables to their corresponding empirical quantities. In the present context, where we have realized measures of the variances and the (transformed) correlations, we adopt the following measurement equations:

$$\begin{aligned} \log x_t &= \xi + \Phi \log h_t + \delta(z_t) + v_t, \\ y_t &= \tilde{\xi} + \tilde{\Phi}\varrho_t + \tilde{v}_t, \end{aligned} \tag{3}$$

where ξ is an $p \times 1$ vector, Φ is an $p \times p$ matrix, $\tilde{\xi}$ is an $d \times 1$ vector, $\tilde{\Phi}$ is an $d \times d$ matrix, and $\delta(\cdot)$ is (like $\tau(\cdot)$) a leverage function that captures dependencies between return and volatility innovations. This dependency is known to be empirically important, and is often referred to as the leverage effect, see Black (1976), Christie (1982), Engle and Ng (1993).[2] In our empirical analysis we simplify the structure further by specifying the matrices β, $\tilde{\beta}$, γ, $\tilde{\gamma}$, Φ, and $\tilde{\Phi}$ to be diagonal matrices.

The measurement equations involves the "error" terms v_t and $\tilde{v}_t$, which we stack into the vector $u_t = (v_t', \tilde{v}_t')'$. In our likelihood analysis we specify u_t to be iid$N(0, \Sigma)$ and independent of z_t. The Gaussian specification is less likely to be at odds with data because the measurement equations are formulated with the logarithmically transformed variances, $\log x_t$ and the transformed correlations, $y_t = g(Y_t)$. For instance, Andersen et al. (2001a) and Barndorff-Nielsen and Shephard (2002b) found that the logarithm of the realized variance can be well approximated by a Gaussian distribution, and the transformation, g, can be interpreted as a generalization of the Fisher transformation of a single correlation, see Archakov and Hansen (2020). The implication is that y_t is approximately Gaussian distributed, as is the case for $\log x_t$, and this is desirable in the present context where we adopt a Gaussian specification. Empirical justification for the Gaussian specification will be presented in Section 4.3.

The measurement errors are likely to be correlated in practice, so that the covariance matrix Σ is expected to be a non-diagonal matrix. The assumption that u_t and z_t are independent is obviously restrictive, however the inclusion of the leverage function, $\delta(\cdot)$, served to eliminate some forms of dependencies, because it is sought to capture the conditional mean, $\delta(z_t) = \mathbb{E}[\log x_t - \xi - \Phi \log h_t | z_t]$, which would imply mean-independence, $\mathbb{E}[u_t | z_t] = 0$. Following Hansen et al. (2012), we introduce a parametric leverage function given by,

$$\delta(z_t) = \delta_1 z_t + \delta_2 (z_t \circ z_t - \iota_p)$$

where ι_p denotes the $p \times 1$ vector of ones, δ_1 and δ_2 are $p \times p$ coefficient matrices. This leverage

[2]The leverage effect is sometimes used to refer to a linear dependence, i.e. the (usually negative) correlation between the return volatility and the actual returns.

function defines a multivariate version of the leverage function introduced in Hansen et al. (2012). We parametrize τ similarly, $\tau(z) = \tau_1 z + \tau_2(z \circ z - \iota_p)$, where τ_1 and τ_2 are $p \times p$ matrices. This structure is motivated by results in Hansen et al. (2012) and Hansen and Huang (2016) who found that a parsimonious leverage function written as a second-order Hermite polynomial is sufficient for capturing the asymmetry dependence between return shocks and volatility shocks. In our empirical analysis we also impose a diagonal matrix structure on the four $p \times p$ matrices δ_1, δ_2, τ_1, and τ_2, to reduce the number of free parameters in the model.

In the next Section we discuss way to further simplify the model by reducing the number of latent variables that drive the variation in the correlation matrix.

2.3 Dimension Reduction of the Correlation Structure

The formulation above permits a flexible modeling of the correlation matrix. Any non-singular correlation matrix, C_t, maps to a unique vector $\varrho_t = g(C_t)$ and any vector ϱ_t (of proper dimension) maps to a unique correlation matrix, $C_t = g^{-1}(\varrho_t)$.

This can be useful for systems of modest dimensions. The number of latent variables embedded in the correlation matrix, C_t, is the dimension of ϱ_t which equals $d(p) = p(p-1)/2$. Thus $d(p) = 1, 3, 6, 10, 15, 21, 28, 36,$ for $p = 2, \ldots, 9$. So the number of latent variables becomes unmanageable unless p is relatively small, and this necessitates some additional structure in the model.

We can impose structure on C_t by modeling ϱ_t with a smaller number of latent variables, specifically by letting $\varrho_t = \varrho(\zeta_t)$ be a function of a lower dimensional vector $\zeta_t \in \mathbb{R}^k$, where $k < d$. In this section, we consider the case were ϱ_t is a linear function of ζ_t:

$$\varrho_t = A\zeta_t,$$

for some $d \times k$ matrix A. The implication is that the variation in C_t is driven by k factors, where the factors are the elements in ζ_t. This enables us to reduce the dimensions of the GARCH and measurement equations. The linear restriction befits the lower dimensional GARCH equation,

$$\zeta_t = \check{\omega} + \check{\beta}\zeta_{t-1} + \check{\gamma}\check{y}_{t-1}, \tag{4}$$

where $\check{y}_t = (A'A)^{-1}A'y_t \in \mathbb{R}^k$. The reason that $\check{y}_t$ is the natural signal about ζ_t follows from well known projection arguments. If A has full column rank, k, then there exists an $d \times (d-k)$ matrix, $A_\perp$ so that $(A, A_\perp)$ is a full rank matrix, and $A'A_\perp = 0$. Thus $\varrho_t = A\zeta_t$ implies $A'_\perp\varrho_t = 0$ and the identity $I_d = A(A'A)^{-1}A' + A_\perp(A_\perp' A_\perp)^{-1}A_\perp'$ (the $d \times d$ identity matrix) establishes that the factor structure $\varrho_t = A\zeta_t$ implies $\varrho_t = A(A'A)^{-1}A'\varrho_t$. Recall that the vector of transformed realized correlations, y_t, is the our empirical "signal" about ϱ_t. The identity,

$$y_t = [A(A'A)^{-1}A' + A_\perp(A_\perp' A_\perp)^{-1}A_\perp']y_t = A\check{y}_t + A_\perp(A_\perp' A_\perp)^{-1}A_\perp' y_t,$$

shows that $\check{y}_t$ is the the natural signal about ζ_t when $\varrho_t = A\zeta_t$. This motivates the GARCH equation (4) and the corresponding measurement equation is given by

$$\check{y}_t = \check{\xi} + \check{\Phi}\zeta_t + \check{v}_t. \tag{5}$$

So the linear restrictions enable us to reduce the dimension of both the GARCH and measurement equation to k from d.

The dimension reduction (outlined here) requires a $d \times k$ matrix, A. This matrix may be chosen ex-ante, but it can also be a matrix that is determined empirically. An empirically estimated A (or rather the subspace spanned by its columns) including the number of columns in A (the number of factors), would require proper inference methods to be derived for this purpose. In this paper, we will focus on the case where A is known, and leave the case with an empirically determined A for future research. We also leave more general factor structures, $\varrho_t = \varrho(\zeta_t)$, for future research. The linear structure, $\varrho_t = A\zeta_t$, considered in this paper is the structure that emerges from a block equicorrelation matrix.

Specific choices for A can be motivated by assuming specific structures apply to C_t, and we will in our empirical analysis adopt an equicorrelation structure and a block equicorrelation structure that define the matrix A. Imposing the block equicorrelation structure on C_t can be achieved with a particular matrix, A, as we will show in the next subsection.

2.4 Dynamic Block Equicorrelation

The transformation, $\varrho_t = g(C_t)$, has interesting properties, and preserves certain structures in C_t, see Archakov and Hansen (2020). In this paper, we present a rigorous proof that the block equicorrelation structure is preserved in the transformed matrix $\log C_t$. The result can be illustrated with the following example:

$$\underbrace{\begin{pmatrix} 1.0 & 0.4 & 0.4 & 0.2 & 0.2 & 0.2 \\ 0.4 & 1.0 & 0.4 & 0.2 & 0.2 & 0.2 \\ 0.4 & 0.4 & 1.0 & 0.2 & 0.2 & 0.2 \\ 0.2 & 0.2 & 0.2 & 1.0 & 0.6 & 0.6 \\ 0.2 & 0.2 & 0.2 & 0.6 & 1.0 & 0.6 \\ 0.2 & 0.2 & 0.2 & 0.6 & 0.6 & 1.0 \end{pmatrix}}_{=C} \qquad \underbrace{\begin{pmatrix} -.16 & .349 & .349 & .104 & .104 & .104 \\ .349 & -.16 & .349 & .104 & .104 & .104 \\ .349 & .349 & -.16 & .104 & .104 & .104 \\ .104 & .104 & .104 & -.36 & .553 & .553 \\ .104 & .104 & .104 & .553 & -.36 & .553 \\ .104 & .104 & .104 & .553 & .553 & -.36 \end{pmatrix}}_{=\log C}. \tag{6}$$

The property that the block structure in C is carried over to $\log C$ holds in general, regardless of the number of blocks and block sizes.

Next we introduce the notion of a quasi-block matrix that is similar to a block matrix except for its diagonal.

Definition 1. A matrix, B, is said to be a quasi-block matrix, with partition $s_1, s_2, \ldots, s_K$, if

$$B = \begin{bmatrix} B_{[1,1]} & B_{[1,2]} & \cdots & B_{[1,K]} \\ B_{[2,1]} & B_{[2,2]} & & B_{[2,K]} \\ \vdots & & \ddots & \vdots \\ B_{[K,1]} & B_{[K,2]} & \cdots & B_{[K,K]} \end{bmatrix},$$

where $B_{[i,j]}$ is an $s_i \times s_j$ matrix where all elements are identical if $i \neq j$, and (if $i = j$) the off-diagonal elements of $B_{[i,i]}$ have an identical common value, whereas the diagonal elements of $B_{[i,i]}$ may have a different common value.

A $p \times p$ correlation matrix, C, with a quasi-block structure, is known as block equicorrelation matrices, see Engle and Kelly (2012). The underlying structure is a classification of variables into groups where the correlation between two variables is defined by their group membership. By sorting the variables according to their group membership, the block equicorrelation structure emerges, as exemplified in (6) for the case with two groups. If all variables belong to the same group, then all correlation coefficients in C are identical, and this case corresponds to the equicorrelation matrix. A dynamic equicorrelation model is know as the DECO model, see Engle and Kelly (2012).

For block equicorrelation matrices we have an interesting result that the matrix-logarithmically transformed correlation matrix preserves the same block structure, as illustrated in (6). This facilitates a parsimonious modeling of correlation matrices with a quasi-block structure, because the elements of the transformed matrix can be modeled in an unrestricted way, without compromising the required structure for a correlation matrix. The formal result that justifies this approach to modeling is the following.

Theorem 1. *Suppose that C is a non-singular correlation matrix with a quasi-block structure. Then $\log C$ has the same quasi-block structure.*

The block equicorrelation structure is an intuitive way to impose structure on a correlation matrix, which reduces the number of free parameters. A $p \times p$ correlation matrix has $p(p-1)/2$ correlations, whereas a block equicorrelation matrix with b groups has at most $b(b-1)/2+b$ distinct correlations.[3] In our empirical analysis, we model returns for nine assets - three assets from three distinct sectors - so we employ a 3×3 block structure, which has 6 distinct correlations, whereas the unrestricted correlation matrix has $9 \times 8/2 = 36$ correlations.

The correlations in a block equicorrelation matrix are subject to non-linear cross restrictions, steaming from the positive definiteness requirement. These conditions are fully characterized below in Theorem 2

The result in Theorem 1 facilitates modeling of a block equicorrelation matrix using an unrestricted vector (the distinct off-diagonal elements of the transformed correlation matrix, $\log C$). These elements can vary freely, and will always map back to a unique positive definite correlation matrix. Bypassing the need to verify positive definiteness, makes the result in Theorem 1 very useful for the empirical implementation of the block DECO model by Engle and Kelly (2012).

It is useful to have closed-form expressions for the inverse of C. But in the existing literature, such have only been derived for the the case with two blocks, see Engle and Kelly (2012, lemma 2.3). The following Theorem provides general expressions for an arbitrary number of blocks.

Theorem 2. *Suppose that C is a block-equicorrelation matrix with block sizes $s_1, \ldots, s_K$, where ρ_{ij} is the correlation coefficient in block $C_{[i,j]}$. Then its blocks can be expressed as*

$$C_{[i,j]} = a_{ij} P_{[i,j]} + 1_{\{i=j\}}(1-\rho_{ii})(I_{s_i} - P_{[i,i]}), \qquad \text{for} \quad i,j = 1, \ldots, K,$$

for some constants $a_{ij} = \rho_{ij}\sqrt{s_i s_j}$ for $i \neq j$, $a_{ii} = 1 + (s_i - 1)\rho_{ii}$ and where $P_{[i,j]} = \frac{1}{\sqrt{s_i s_j}} \iota_{s_i} \iota_{s_j}'$, is such that $P_{[i,j]} P_{[j,k]} = P_{[i,k]}$.

The eigenvalues of C are those of the (symmetric) matrix $A = \{a_{ij}\}_{i,j=1,\ldots,K}$, and $1-\rho_{ii}$ (with multiplicity $s_i - 1$), $i = 1, \ldots, K$. So that $\det C = (\det A)(1-\rho_{11})^{s_1 - 1} \cdots (1-\rho_{KK})^{s_k - 1}$.

[3]The exact number of distinct correlation is $b(b-1)/2 + \tilde{b}$, where $\tilde{b}$ is the number of groups that contain two or more elements. The obvious reason being that a 1×1 diagonal block will not have a correlation coefficient.

The correlation matrix, C, is invertible if and only if A is invertible and $\rho_{ij} < 1$ for all $i, j = 1, \ldots, K$. In which case, the inverse, C^{-1}, is a quasi-block matrix, with blocks given by

$$C^{-1}_{[i,j]} = a^{\#}_{ij} P_{[i,j]} + 1_{\{i=j\}} \tfrac{1}{1-\rho_{ii}} (I_{s_i} - P_{[i,i]}), \qquad \textit{for} \quad i, j = 1, \ldots, K,$$

where $a^{\#}_{ij}$ is the ij-th element of A^{-1}.

In the special case where the block-sizes are identical ($s_i = s$ for all $i, j = 1, \ldots, K$), then C and C^{-1} can be expressed as Kronecker products, $C = A \otimes P + D \otimes P_\perp$ and $C^{-1} = A^{-1} \otimes P + D^{-1} \otimes P_\perp$ where $P = \iota_s(\iota_s' \iota_s)\iota_s'$ and $P_\perp = I - P$ are orthogonal projection matrices, and $D = \mathrm{diag}(1 - \rho_{11}, \ldots, 1 - \rho_{KK})$, and $\det C = \det A \cdot (\det D)^{s-1}$.

A valuable contribution of Theorem 2 is that it facilitates simple evaluation of the log-likelihood function for a block equicorrelation models with a Gaussian specification. Theorem 2 also shows that the matrix A characterizes the set of positive definite block equicorrelation matrices. For a correlation matrix to be positive definite, all correlation coefficients must be less than one in absolute value. This follows from the determinant of the 2×2 principal sub-matrices of C. The additional requirement that guarantees that a block equicorrelation matrices is positive definite, is that A is positive definite. The characterization of positive definite block equicorrelation matrices has independently been obtained in Roustant and Deville (2017), who formulate it with a different matrix, that A used here. An benefit of our formulation is that the eigenvalues of A are also eigenvalues of C. This is, for instance, useful in the evaluation of the log-likelihood with a Gaussian specification.[4]

The block structure offers a useful dimension reduction, but in order to make use of it, one has to specify the block structure. The block structure could be selected based on prior knowledge, where subsets of variables are naturally bundled together, such as assets from distinct economic sectors. Alternatively, one could use a data-driven approach to form the blocks. For instance, by using empirical measures of correlations to group assets with similar correlations together. In our empirical analysis, we study nine assets, where the block structure is defined by the sectors to which the assets belong.

[4]We were unaware of the result in Roustant and Deville (2017) until an anonymous referee pointed it out us. Surprisingly, Engle and Kelly (2012) is not cited in Roustant and Deville (2017).

2.5 Measurement Equation under Dimension Reduction

The original measurement equation (3) can be replaced with the condensed equation (5) under the factor structure, $\varrho_t = \varrho(\zeta_t)$, but one could also maintain the original equation. In many applications the primary objective is to obtain a good model for returns, whereas the model for the realized measures are of secondary interest. The main purpose of the measurement equations for the transformed variances and correlations is to facilitate a way to incorporate the realized measures into the modeling. When a factor structure is used, $\varrho_t = \varrho(\zeta_t)$, the lower dimensional measurement equation for ζ_t is sufficient to achieve this, and using a lower dimensional system of equations can be numerically advantageous.

It should be pointing out that there are situations where the original measurement equation (3) should be used. For instance, if multiple model specifications are to be compared in terms of their total log-likelihood, then it requires that all models employ the same measurement equations. Different factor structures could, in principle, be condensed to different measurement equations. But a comparison of the total likelihood for models based on different measurement equations is a comparison of models that model different (dimensions of) variables. So this would amount to a comparison of apples and oranges. Therefore, if multiple models are to be compared in terms of their total likelihood, then one should adopt a common set of measurement equations, such as (3). An alternative way to proceed, is to evaluate the model specifications in terms of their partial log-likelihood for returns, and ignore the part of the likelihood that relate to the realized measures. Since the primary objective of multivariate GARCH models is to model the conditional distribution of returns, this may be the preferred way to compare models. The relevant terms of the log-likelihood for this comparisons are detailed in the next section. In our empirical analysis we demonstrate how different model specifications can be compared in terms of their ability to model the conditional distribution of returns, using the return log-likelihood.

3 Estimation

The estimation problem is relatively simple because the model is an observation-driven model. The dynamics of all latent variables (conditional variances and correlations) are driven by observable variables (returns and realized measures). Moreover, the model has a structure that permits some simplifications in the expression for the likelihood function.

We can factorize the joint density of (r_t, x_t, y_t), conditional on the past, into the marginal density for returns and the density for realized variables, conditional on contemporaneous returns. Thus, the joint density is expressed as the product, $f_{t-1}(r_t, x_t, y_t) = f_{t-1}(r_t)f_{t-1}(x_t, y_t|r_t)$. The log-likelihood function can therefore be deduced from

$$\sum_{t=1}^{T} \log f_{t-1}(r_t, x_t, y_t) = \sum_{t=1}^{T} \log f_{t-1}(r_t) + \sum_{t=1}^{T} \log f_{t-1}(x_t, y_t|r_t). \tag{7}$$

We estimate parameters by quasi maximum likelihood estimation, where the likelihood function is obtained assuming Gaussian distributions for z_t and u_t. Specifically, that $\{z_t\}$ is a sequence of independent vectors distributed as $z_t \sim N_p(0, C_t)$, while $\{u_t\}$ is independent of $\{z_t\}$ and distributed as $u_t \sim i.i.d.N(0, \Sigma)$.

Let θ represents all unknown parameters in the model. The Gaussian specification and the structure (7) imply that the log-likelihood function is given by

$$\ell(\theta) = \sum_{t=1}^{T} \ell_{r,t}(\theta) + \sum_{t=1}^{T} \ell_{x,y|r,t}(\theta),$$

where

$$\begin{aligned} -2\ell_{r,t}(\theta) &= c_p + \sum_{k=1}^{p} \log h_{k,t} + \log |C_t| + z_t C_t^{-1} z_t, \\ -2\ell_{x,y|r,t}(\theta) &= c_{\frac{p(p+1)}{2}} + \log |\Sigma| + u_t' \Sigma^{-1} u_t, \end{aligned}$$

with $c_n = n \log 2\pi$ and

$$\begin{aligned} z_t &= \Lambda_{h_t}^{-1}(r_t - \mu), \quad \text{and} \\ u_t &= \begin{pmatrix} v_t \\ \tilde{v}_t \end{pmatrix} = \begin{pmatrix} \log x_t - \xi - \Phi \log h_t - \delta(z_t) \\ y_t - \tilde{\xi} - \tilde{\Phi} \varrho_t \end{pmatrix}. \end{aligned}$$

Here $|C_t|$ and $|\Sigma|$ denote the determinants of C_t and Σ, respectively.

For a particular value of θ it is straight forward to evaluate the log-likelihood function, using a recursive scheme. The values for h_t and C_t are computed with the GARCH equations, then the innovations, z_t and u_t, are computed from the measurement equations, and then one can proceed to compute the quantities for period $t + 1$. Finally, the likelihood function can be evaluated. Maximization methods undertake this computation repeatedly to obtain the maximum likelihood estimates. This computation does require starting values for the latent variables, h_1 and C_1. We recommend simply treating them as unknown parameters (as part of θ), which is a common approach in GARCH models. Another approach is to simply fix them to have a particular value, defined by appropriate empirical quantities.

The structure of the log-likelihood function permits an important simplification for the maximization problem. Given residuals, $\hat{u}_t$, $t = 1, \ldots, T$ it can be shown that the maximum likelihood estimator of Σ is $\hat{\Sigma} = T^{-1}\sum_{t=1}^{T}\hat{u}_t\hat{u}_t'$. This, in turns, simplifies a term in the log-likelihood function,

$$\sum_{t=1}^{T}\hat{u}_t'\hat{\Sigma}^{-1}\hat{u}_t = \mathrm{tr}\{\hat{\Sigma}^{-1}\sum_{t=1}^{T}\hat{u}_t\hat{u}_t'\} = \mathrm{tr}\{TI_{p+k}\} = T(p+k),$$

where I_{p+k} is the $(p+k) \times (p+k)$ identity matrix. So the objective to be maximized is (apart from a constant) given by

$$-\tfrac{1}{2}\sum_{t=1}^{T}\left\{\sum_{i=1}^{p}\log h_{i,t} + \log|C_t| + z_tC_t^{-1}z_t\right\} - \tfrac{T}{2}\log\left|\tfrac{1}{T}\sum_{t=1}^{T}u_tu_t'\right|, \tag{8}$$

where the omitted constant is $-\frac{T}{2}(c_p+c_d+p+k) = -\frac{p(p+1)}{4}T\log 2\pi - T\frac{p+k}{2}$. For block equicorrelation matrices, both the determinant, $|C_t|$, and the inverse, C_t^{-1}, is simple to evaluate using Theorem 2. More extensive details of the estimation method are described in greater detail in Appendix B.

3.1 Estimation with Structure Imposed on the Correlation Matrix

In this section, we add a few details about estimation in the situation where $\varrho_t = A\zeta_t$ is used to impose structure on the correlation matrix. The GARCH and measurement equations for the correlation parameters, must be revised to embody the structure $\varrho_t = A\zeta_t$, see Section 2.3. The appropriate GARCH equation is (4) and if the lower dimensional measurement equation (5) is adopted, $\check{y}_t = \check{\xi} + \check{\Phi}\zeta_t + \check{v}_t$ with $\check{y}_t = (A'A)^{-1}A'y_t$, then the vector of (all) measurement errors is redefined with $u_t = (v_t', \check{v}_t')'$.

To illustrate the features arising with five assets, with a block structure consisting of two blocks, two assets in the first block and three assets in the second block. In this case C_t has just 3 distinct correlations, a_t, b_t, and c_t, say. In this scenario we have by Theorem 1 the following structure of

$$C_t = \begin{pmatrix} 1 & & & & \\ a_t & 1 & & & \\ b_t & b_t & 1 & & \\ b_t & b_t & c_t & 1 & \\ b_t & b_t & c_t & c_t & 1 \end{pmatrix} \qquad \log C_t = \begin{pmatrix} * & & & & \\ \tilde{a}_t & * & & & \\ \tilde{b}_t & \tilde{b}_t & * & & \\ \tilde{b}_t & \tilde{b}_t & \tilde{c}_t & * & \\ \tilde{b}_t & \tilde{b}_t & \tilde{c}_t & \tilde{c}_t & * \end{pmatrix},$$

Thus $\varrho_t = (\tilde{a}_t, \tilde{b}_t, \tilde{b}_t, \tilde{b}_t, \tilde{b}_t, \tilde{b}_t, \tilde{b}_t, \tilde{c}_t, \tilde{c}_t, \tilde{c}_t)'$ and

$$\varrho_t = A\zeta_t \qquad \text{where} \qquad A' = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 1 & 1 & 1 & 1 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 1 \end{pmatrix}, \qquad \zeta_t = \begin{pmatrix} \tilde{a}_t \\ \tilde{b}_t \\ \tilde{c}_t \end{pmatrix}.$$

In this case, the dimension reduction is from $d = 10$ to $k = 3$. In our empirical analysis we consider a case with 9 assets, and a correlation matrix C_t with a 3×3 block structure. In this case the dimension reduction is from $d = 36$ correlations in C_t, to $k = 6$ distinct correlations with the block structure, and the dimension of A is therefore 36×6 in our empirical example. This defines a scalable model. Adding additional assets from sectors that are already being modeled, does not increase the dimension of the latent variable used to model the correlation structure, ζ_t. The block equicorrelation structure is therefore a promising starting point for scaling the model to dimensions higher than the nine dimensional system analyzed in the empirical section.

3.2 Forecasting

The one-step ahead forecasting of the return distributions from the model is straight forward because all dynamic variables are specified in the observation driven manner, and are simple functions of lagged variables. So from the observed variables in period t, all the conditional variances and correlations for period $t+1$ can be computed from the GARCH equations. The elements of H_{t+h}, are not predetermined beyond horizon $h = 1$, because they also depend on future realizations of z_t and u_t. It is nevertheless straight forward to compute a distributional forecasts for H_{t+h} using simulation methods or a bootstrap method. So multi-step ahead forecasts can be inferred from the estimated model, at any forecasting horizon. Forecasting schemes for the Realized GARCH models of this kind are detailed in Lunde and Olesen (2014) and Hansen et al. (2014). In this context, a bootstrap method is typically preferred because it does not rely on the distributional assumptions for z_t and u_t.

4 Empirical Analysis

4.1 Data Description

The data set for our empirical analysis spans the period from January 3, 2002 to December 29, 2017. After removing holidays and trading days with reduced trading hours, our sample includes

3975 trading days. The data consist of daily close-to-close returns and high-frequency data. The latter are used to compute the realized measures of volatility.

We include the nine stocks in our analysis. Three stocks from the energy sector, CVX, MRO, and OXY, three stocks from the Health Care sector, JNJ, LLY, and MRK, and three stocks from the Information Technology sector, AAPL, MU, and ORCL.

	Energy			Health Care			Information Tech.		
	CVX	MRO	OXY	JNJ	LLY	MRK	AAPL	MU	ORCL
	Daily returns (×100)								
Mean	0.055	0.058	0.078	0.037	0.022	0.027	0.136	0.052	0.053
Std.	1.598	2.486	2.065	1.099	1.499	1.676	2.206	3.381	2.016
Skewness	0.393	0.177	0.126	-0.226	0.046	-1.057	0.065	0.038	0.200
Kurtosis	17.503	11.329	13.040	22.055	10.397	27.775	8.096	7.405	8.279
Min	-12.489	-19.564	-18.493	-15.846	-12.348	-26.781	-17.920	-23.042	-14.509
Q-05%	-2.371	-3.731	-3.117	-1.582	-2.250	-2.281	-3.298	-5.101	-2.999
Q-25%	-0.758	-1.135	-0.931	-0.464	-0.713	-0.715	-0.956	-1.661	-0.885
Q-50%	0.086	0.088	0.053	0.016	0.039	0.028	0.090	0.000	0.024
Q-75%	0.877	1.315	1.107	0.564	0.749	0.818	1.235	1.780	1.001
Q-95%	2.270	3.633	3.033	1.669	2.230	2.347	3.707	5.302	3.089
Max	20.854	23.357	18.108	12.229	14.349	13.033	13.905	23.443	13.070
	Realized volatilities (in annual units)								
Mean	0.212	0.317	0.262	0.160	0.203	0.220	0.306	0.455	0.279
Std.	0.129	0.194	0.157	0.091	0.112	0.136	0.188	0.245	0.169
Skewness	6.058	3.646	4.595	4.058	4.065	4.123	3.602	3.037	2.463
Kurtosis	82.641	27.811	41.080	34.169	35.317	33.503	28.461	19.593	13.001
Min	0.072	0.095	0.073	0.051	0.063	0.064	0.054	0.115	0.066
Q-05%	0.106	0.146	0.130	0.083	0.106	0.107	0.124	0.230	0.114
Q-25%	0.140	0.203	0.173	0.109	0.139	0.144	0.187	0.307	0.173
Q-50%	0.181	0.265	0.225	0.134	0.174	0.181	0.265	0.393	0.235
Q-75%	0.244	0.360	0.301	0.179	0.231	0.250	0.367	0.520	0.324
Q-95%	0.411	0.645	0.499	0.317	0.401	0.456	0.622	0.900	0.627
Max	2.807	3.007	2.503	1.402	1.830	2.145	2.548	3.511	1.981

Table 1: Summary statistics of the daily returns and the realized volatilities (in annual units) computed from the realized kernel estimator. Statistics are based on the sample period from January 2nd, 2002 to December 29th, 2017 (3975 trading days in total).

We construct close-to-close daily returns for the individual stocks using close prices from the CRSP US Stock Database. These prices are adjusted for the stock splits and dividends. Intraday transaction data were obtained from the TAQ database and these were cleaned in accordance with the methodology suggested in Barndorff-Nielsen et al. (2009). From the high-frequency data we compute the 9×9 multivariate realized kernel estimator for each of the trading days in our sample, RM_t. The diagonal of RM_t defines the vector x_t, and the corresponding realized correlation matrix, Y_t, and the latter is transformed by $y_t = g(Y_t)$. Hence, Y_t is subjected to the same transformation as C_t, so that y_t can be interpreted as an empirical measurement of ϱ_t.

4.2 Summary Statistics

We present summary statistics for the nine return series and their corresponding realized variance measures in Table 1. The summary statistics are in line with figures that are typical in this context. We do observe that Health Care stocks (middle three columns) had the lowest volatility whereas IT stocks (the last three columns) had the largest average volatility in this sample period. This can be seen from the standard deviations in the second row, and the means and medians (Q-50%) of the Realized volatilities in the lower part of Table 1.

		Energy			Health Care			Information Tech.		
		CVX	MRO	OXY	JNJ	LLY	MRK	AAPL	MU	ORCL
Energy	CVX		0.490	0.511	0.155	0.116	0.129	0.161	0.107	0.158
			(0.167)	*(0.169)*	*(0.122)*	*(0.123)*	*(0.121)*	*(0.119)*	*(0.107)*	*(0.120)*
	MRO	0.554		0.483	0.069	0.077	0.085	0.123	0.119	0.118
		(0.145)		*(0.173)*	*(0.119)*	*(0.115)*	*(0.113)*	*(0.112)*	*(0.119)*	*(0.117)*
	OXY	0.566	0.543		0.087	0.078	0.088	0.122	0.094	0.118
		(0.143)	*(0.148)*		*(0.115)*	*(0.113)*	*(0.112)*	*(0.116)*	*(0.113)*	*(0.115)*
Health Care	JNJ	0.260	0.191	0.204		0.291	0.302	0.139	0.087	0.164
		(0.173)	*(0.171)*	*(0.168)*		*(0.145)*	*(0.157)*	*(0.118)*	*(0.112)*	*(0.120)*
	LLY	0.232	0.186	0.190	0.368		0.333	0.121	0.092	0.144
		(0.176)	*(0.169)*	*(0.167)*	*(0.158)*		*(0.160)*	*(0.109)*	*(0.105)*	*(0.113)*
	MRK	0.245	0.197	0.202	0.375	0.392		0.127	0.093	0.147
		(0.174)	*(0.166)*	*(0.166)*	*(0.168)*	*(0.164)*		*(0.112)*	*(0.105)*	*(0.117)*
Information Tech.	AAPL	0.282	0.243	0.244	0.244	0.228	0.234		0.208	0.279
		(0.166)	*(0.159)*	*(0.163)*	*(0.152)*	*(0.147)*	*(0.150)*		*(0.120)*	*(0.140)*
	MU	0.218	0.211	0.196	0.180	0.179	0.182	0.280		0.189
		(0.148)	*(0.150)*	*(0.151)*	*(0.140)*	*(0.133)*	*(0.137)*	*(0.133)*		*(0.121)*
	ORCL	0.281	0.238	0.240	0.267	0.250	0.255	0.350	0.267	
		(0.174)	*(0.169)*	*(0.168)*	*(0.157)*	*(0.154)*	*(0.158)*	*(0.157)*	*(0.140)*	

Table 2: Summary statistics for the realized correlations where the sector-based block structure is illustrated with the shaded regions. The average realized correlations and corresponding standard deviations (in parentheses) are shown in lower-left triangle. The corresponding statistics for the transformed elements, ϱ_t are shown in the upper-right triangle. Statistics are based on the sample period from January 2nd, 2002 to December 29th, 2017 (3975 trading days).

Summary statistics for the realized correlations are presented in Table 2. The block structure we explore for the correlation matrix is illustrated with the shaded regions in Table 2. The numbers below the diagonal are average correlations (the lower-left triangle of $\bar{Y} = T^{-1}\sum_t Y_t$), and the numbers above the diagonal are corresponding averages for the transformed quantities (the upper-right triangle of $T^{-1}\sum_t \log Y_t$). Note that the realized correlations within each of the blocks have similar averages. The three assets from the energy sector are highly correlated, with correlations of about 55% on average. The average "within sector" correlations for Health Care and Information Technology stocks are about 38% and 30%, respectively. The "between sector" correlations tend to be smaller and range from 18% to 28%. A similar pattern is observed for the corresponding elements of y_t, that are presented above the diagonal of Table 2.

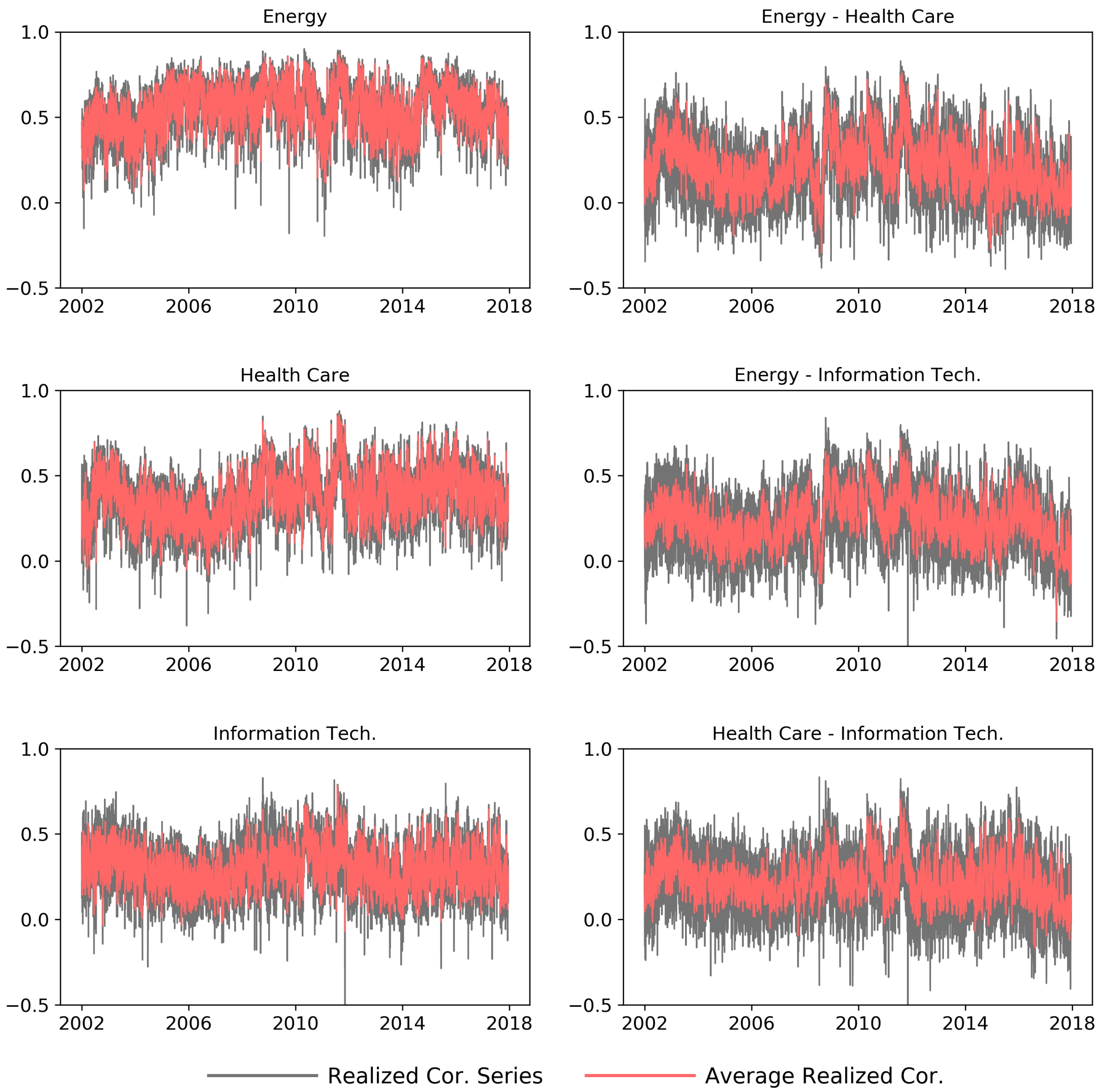

Figure 1: Daily realized correlations for the nine returns series over the full sample period. Left subplots present intra-sector correlations (gray lines) and their average (red line) for each sector. Right subplots present inter-sector correlations (gray lines) and their average (red line) for each pair of sectors.

The time series of realized correlation series are displayed in Figure 1. The left subplots present the within-sector correlations (gray lines) and their average daily correlation (red line) for each of the three sectors. The right subplots present the between-sector correlations (gray lines) and their daily averages (red line) for the three sector-pairs. The time series of 36 correlations are computed from the 9×9 multivariate Realized Kernel estimator. The correlations within each blocks tend to move closely together, and there are distinct differences between blocks, not only in terms of their average level, but also in the variation over time. For instance the inter-sector correlation between Health Care and Information Technology asset returns does not have a sharp decline in late 2008 as is the case for the two other inter-sector correlations that both involve Energy sector stocks. Figure 1 provides additional motivation for adopting a block structure of the correlation matrix.

4.3 Transformed Realized Correlations are Approximately Gaussian

The logarithmic realized variance of stock returns is approximately Gaussian distributed, as demonstrated in Andersen et al. (2001a). This help justify the Gaussian specification for the errors, v_t, in the measurement equation for the logarithm of realized variance measures, $\log x_t$. Interestingly, we find that components of logarithmically transformed realized correlation matrices are also approximately Gaussian distributed. This complementary result, help justify the Gaussian specification adopted for the errors, $\tilde{v}_t$ and $\check{v}_t$, in the measurement equations for y_t and $\check{y}_t$, respectively.

Figure 2 presents Q-Q plots for the empirical distribution of the transformed realized correlations against the normal distribution. For each element of y_t, the logarithmic transformed correlations, we have 3975 daily observations. The quantiles of their empirical distribution are plotted against the corresponding quantiles of the normal distribution. The left panels of Figure 2 are Q-Q plots for the series in the three diagonal blocks of Y_t, which have three series in each panel. Similarly, the right panels are for the three off-diagonal blocks of Y_t, which have six series in each panel. The red dots within each panel represent the Q-Q plots that for the corresponding elements of $\check{y}_t$, that are the daily averages within each block. The elements of $\check{y}_t$ are the relevant series in the model where we impose a block equicorrelation structure, see Section 2.4.

The Q-Q plots in Figure 2 show that the components of y_t have an empirical distribution that is well approximated by the Gaussian distribution, albeit with some deviations seen in the tail regions and this discrepancy is most pronounced for the between-sector blocks, seen in the right panels of Figure 2. The red dots that represent Q-Q plots for the block averages, which define the elements of $\check{y}_t$, show that their empirical distribution are very well approximated by the Gaussian distribution. This is not entirely unexpected since the elements of $\check{y}_t$ are defined as averages over elements of y_t.

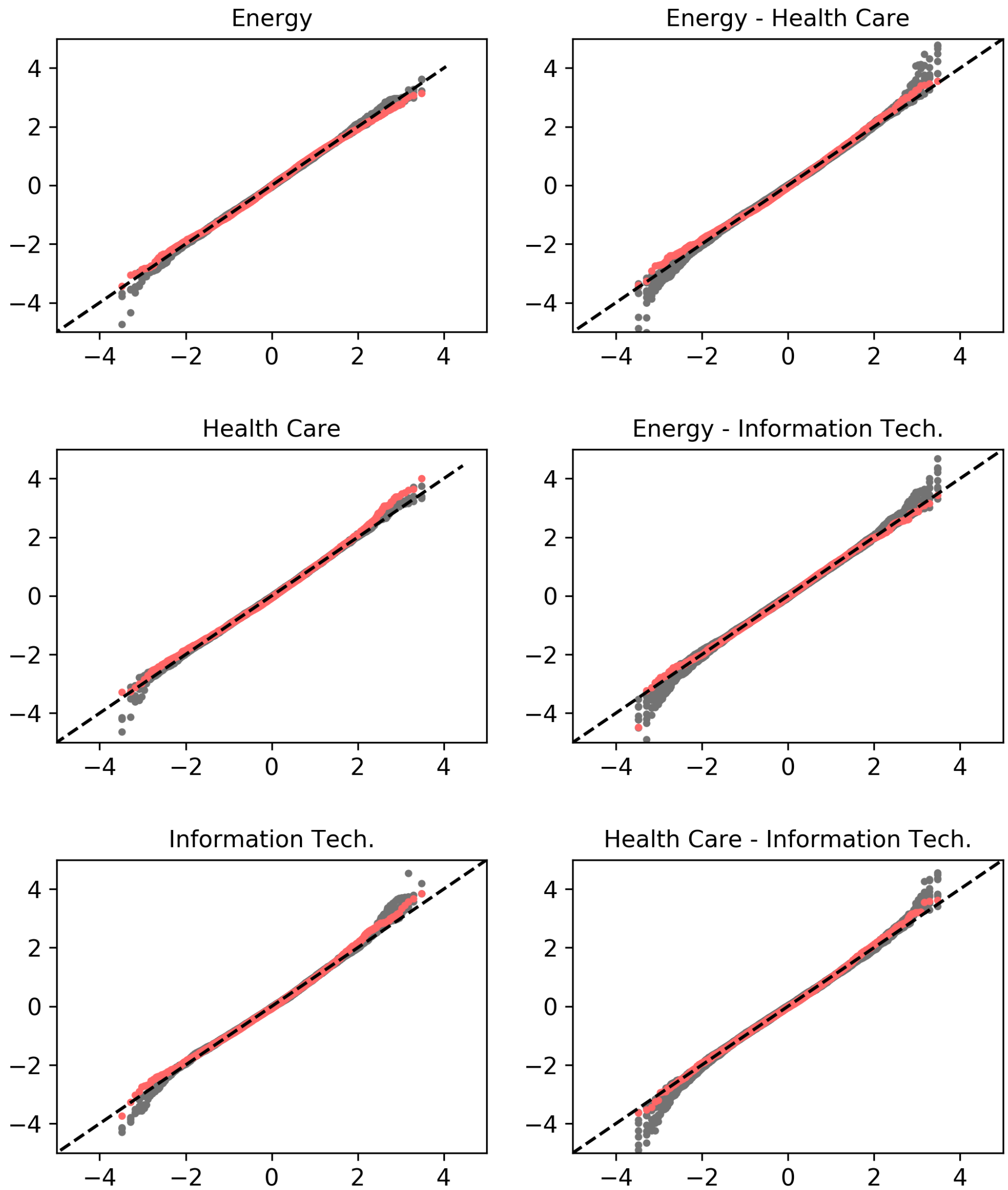

Figure 2: Q-Q plots for the components of parametrized realized correlarion matrices, y_t. Within each panel, black dots are related to the multiple components of the transformed realized correlation matrix, which positions correspond to a given block. Within each panel, red dots are related to the single component of the transformed realized block equicorrelation matrix (for which all realized correlations within each block are replaced by their average value), which corresponds to a given block. The Q-Q plots for all considered components are constructed using daily realized correlations for the period since January 2nd, 2002, until December 29th, 2017 (3975 trading days in total).

4.4 Empirical Analysis of the Multivariate Realized GARCH model

We estimate the model using six different specifications for the correlation matrix. The six specifications arise from the combinations of three structures for C_t: equicorrelation, block equicorrelation, and a "free" correlation structure, and the two ways in which we model the correlations: static and dynamic. The simplest model in our comparison is the static equicorrelation model which has a single correlation coefficient. This correlation coefficient is common for all correlations in C_t, and is constant across time. The most general specification is the dynamic correlation matrix with 36 unrestricted correlations (aside from the requirement that C_t be positive definite). In this model ϱ_t is a 36-dimensional vector, whereas ϱ_t is univariate in the dynamic equicorrelation model, and ϱ_t has dimension six in the dynamic block equicorrelation model.

In terms of modeling of the nine univariate conditional variances, the six specifications have the same structure. They all model each of the nine conditional variances, $h_{1,t}, \ldots, h_{9,t}$ with a Realized GARCH structure using the same information set. The six specifications differ in terms of the way the correlation structure is modeled. The dynamic equicorrelation model has one additional latent variable to model the time variation in the correlation matrix, the dynamic block equicorrelation model has six latent variables for this purpose, where as the dynamic "free" correlation model has 36 latent variables. Here $d = 36 = 10 \times 9/2$ is the number of distinct correlation coefficients in C_t when the correlation matrix is not subject to any restrictions beyond being positive definite.[5]

Parameter estimates for each of the six different specifications are reported in Table 3. The estimates are based on our full sample that spans the period from January 3, 2002 to December 29, 2017. Rather than reporting a very large number of point estimates, we report the range of estimates for each type of parameter. If the model only has one parameter of a particular type, then we report the point estimate. Each column in Table 3 corresponds to one of the six specifications. The first three columns are the three static models, and the last three columns are the three dynamic models.

In the upper panel, we present the point estimates for the part of the model that relates to the conditional variances, and these point estimates are in line with point estimates reported in the existing literature on GARCH models and Realized GARCH models. A simple measure of persistence of the conditional variance is given by $\pi = \beta + \gamma$, and our estimates are close to unity, which is to be expected since volatility is known to be highly persistent.

The lower panel of Table 3 presents the point estimates for the part of the specifications that define the correlation structure. For all the dynamic specifications we also observe a high level of persistence for the conditional correlations, as it is the case for the conditional variances.

[5] For the 9-dimensional vector of returns and 16 years of data, it took us about 30 minutes to estimate the dynamic equicorrelation model, about 8.5 hours to estimate the dynamic block equicorrelation model, and 82 hours to estimate the dynamic model with a free structure, using Matlab code on a standard desktop computer with an Intel Core i7-6700 (3.40GHz) processor.

	Parameter	Static correlation			Dynamic correlation		
		Equi	Block	Free	Equi	Block	Free
	μ	[0.026, 0.142]	[0.023, 0.140]	[0.022, 0.138]	[0.023, 0.135]	[0.014, 0.139]	[0.012, 0.148]
Variance parameters	ω	[-0.008, 0.229]	[-0.011, 0.205]	[-0.009, 0.201]	[-0.006, 0.209]	[-0.002, 0.188]	[0.000, 0.174]
	β	[0.483, 0.728]	[0.490, 0.728]	[0.488, 0.728]	[0.501, 0.723]	[0.507, 0.737]	[0.544, 0.763]
	Γ	[0.232, 0.428]	[0.231, 0.426]	[0.231, 0.426]	[0.241, 0.430]	[0.227, 0.418]	[0.201, 0.382]
	τ_1	[-0.031, 0.000]	[-0.031, 0.000]	[-0.031, -0.000]	[-0.032, 0.001]	[-0.031, 0.003]	[-0.029, 0.004]
	τ_2	[-0.007, 0.031]	[-0.007, 0.029]	[-0.007, 0.030]	[-0.007, 0.030]	[-0.007, 0.027]	[-0.007, 0.025]
	ξ	[-0.417, -0.002]	[-0.400, 0.009]	[-0.391, 0.003]	[-0.388, -0.004]	[-0.391, -0.015]	[-0.384, -0.028]
	δ_1	[-0.070, 0.005]	[-0.068, 0.005]	[-0.069, 0.005]	[-0.066, 0.012]	[-0.064, 0.012]	[-0.067, 0.013]
	δ_2	[0.015, 0.090]	[0.015, 0.088]	[0.015, 0.088]	[0.015, 0.091]	[0.014, 0.091]	[0.014, 0.084]
	σ_v^2	[0.153, 0.298]	[0.154, 0.298]	[0.154, 0.298]	[0.154, 0.297]	[0.155, 0.297]	[0.157, 0.299]
	π	[0.910, 0.975]	[0.912, 0.973]	[0.912, 0.973]	[0.920, 0.978]	[0.920, 0.977]	[0.926, 0.976]
Correlation parameters	$\bar{\rho}$	0.322	[0.227, 0.705]	[0.198, 0.710]			
	$\tilde{\omega}$				-0.009	[-0.008, 0.048]	[-0.016, 0.035]
	$\tilde{\beta}$				0.667	[0.713, 0.800]	[0.828, 0.958]
	$\tilde{\Gamma}$				0.405	[0.140, 0.320]	[0.021, 0.205]
	$\tilde{\xi}$				0.042	[-0.099, 0.076]	[-0.112, 0.140]
	$\tilde{\Phi}$				0.720	[0.640, 1.411]	[0.384, 1.663]
	$\sigma_{\tilde{v}}^2$				0.001	[0.002, 0.009]	[0.010, 0.020]
	$\tilde{\pi}$				0.959	[0.927, 0.982]	[0.966, 0.992]
	$\mathcal{L}$	-63851.475	-61912.446	-61828.382	-63466.324	-61419.596	-61345.306

Table 3: Parameter estimates for the six specifications. The first three columns with estimates are for the static correlation models, and the last three columns are for the dynamic correlation models. Within each specification we report the range of estimates for each "type" of parameter, unless the specification only as a single parameter of this type, in which case we report the point estimate. Persistence of volatilities and correlations are summarized with $\pi = \beta + \gamma$ and $\bar{\pi} = \bar{\beta} + \bar{\gamma}\bar{\Phi}$, respectively. The partial log-likelihood (for returns) is reported in the last row for each of the specifications.

It is not meaningful to compare the total log-likelihood of the different specifications, because they involve measurement equations with different dimensions. While all the specifications model the vector of returns, they differ in terms of the realized measures that are being modeled. For instance, the static specifications do not model any of the realized correlations. However, we can compare the models in terms of their partial log-likelihood for returns, $\ell_{r,t}(\theta)$, which we will refer to as the *return log-likelihood*. The return log-likelihood is an interesting metric because the objective is to obtain a good model for the conditional distribution of returns. We can compute likelihood ratio statistics based on the the partial log-likelihoods alone, and these will reflect which of the models has the best description of the conditional distribution of returns, in the Kullback-Leibler sense. But we can not compare these likelihood-ratio statistics to standard χ^2-based critical values, because the return log-likelihood is only one of two parts of the total log-likelihood.[6] The other term is the log-likelihood for realized measures of volatilities and correlations, and the dimension of the latter differ across specifications.

From the return log-likelihood it is evident that the equicorrelation structure is inferior to the more flexible structures. The return log-likelihood is about 2000 larger with block and free structures than the equicorrelation structure, which is a very substantial empirical improvement. Specifications with a dynamic structure rather than static also have far larger values of the return log-likelihood. For both the block equicorrelation and the free correlation structure, the dynamic specifications have a return log-likelihood that are about 415 larger than those of the static specifications. The dynamic model with a free correlation structure increases the partial likelihood ratio statistic by about 74 over the dynamic block equicorrelation model.

To illustrate the correlation structure that one of estimated models produces, we present the six conditional correlation series, produced by the estimated block equicorrelation model, in Figure 3. The figure also displays the average of the corresponding daily realized correlations. The left panels present the model-based equicorrelation (black line) for each of the three sectors, and the corresponding daily average of the realized correlation is represented with the red line. The right panels present the corresponding equicorrelation for each pair of sectors. The corresponding Figure for the models using the equicorrelation structure is presented in the appendix.

Based on the partial log-likelihood for the full sample, there are substantial gains from moving beyond the equicorrelation model and adopting a dynamic model for the correlations. In the next section, we evaluate the extent to which these improvements carry over to improved empirical fit in an out-of-sample comparison.

[6]Moreover, likelihood ratio statistics are not asymptotically χ^2-distributed under misspecified, see White (1994).

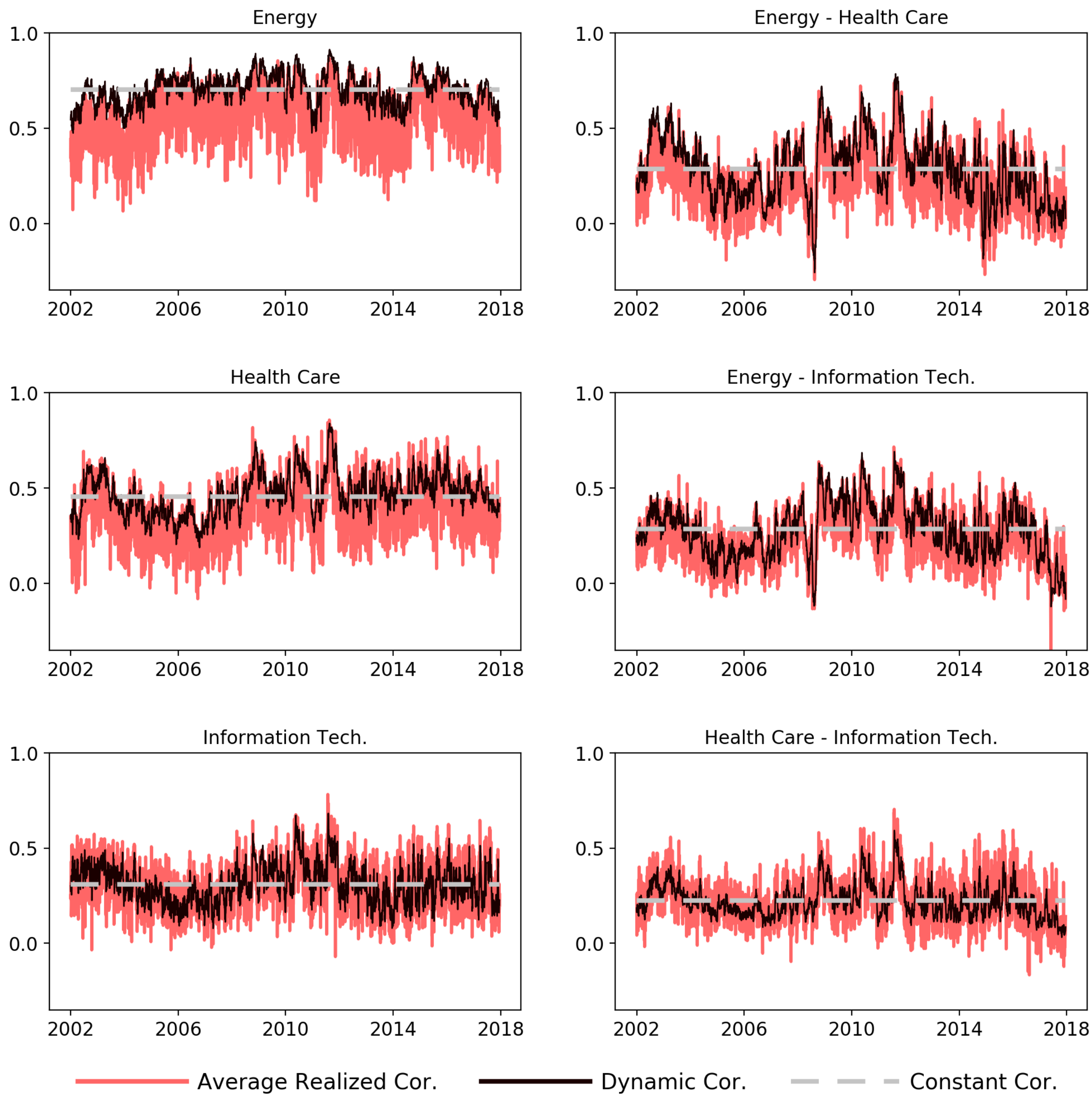

Figure 3: Intra-sectoral and inter-sectorial correlations. The left subplot present the model-based correlations for each of the thee sectors (black lines) and the corresponding daily averages of realized correlations (red lines). The estimates from the constant correlation model are indicated by white dashed lines. The right panels present the correlations across sectors.

4.5 Out-of-Sample Model Comparisons

In this section, we compare the six different specifications in terms of an out-of-sample performance. To this end, we split the sample period into an in-sample period and an out-of-sample period. The in-sample period lasts from January 2nd, 2002 to December 31st, 2013 and includes 2993 trading days. The out-of-sample period spans the period from January 2nd, 2014, to December 29th, 2017 and has 982 trading days. We employ the so-called fixed scheme, where each model is estimated once using the data from the in-sample period. Then the estimated models are evaluated and compared out-of-sample using a range of criteria.

We first compare the models in terms of their partial log-likelihood for returns, then we turn to the other criteria – a minimum variance portfolio selection.

4.5.1 Analysis of Log-Likelihood for Returns

Multivariate GARCH models seek to describe the conditional distribution of the vector of returns. This objective is defined by the log-likelihood function for returns that is used in the estimation. A natural starting point for comparing the different specification is therefore in terms of the return log-likelihood which measures how well the estimated model can explain the distribution of the vector of returns. So, in this subsection, we evaluate and compare the specifications in terms of their average value of $\ell_{r,t}(\hat{\theta})$ (both in-sample and out-of-sample), where the parameter vector $\hat{\theta}$ is estimated with the in-sample data. This type of model evaluation is equivalent to one-day-ahead density forecasting of the return vector with the mean predictive log-likelihood as a gain function, see Amisano and Giacomini (2007), Geweke and Amisano (2010), and references therein.

We compute the average in-sample and average out-of-sample return log-likelihoods, for each of the six specifications. Figure 4 is a bar chart with the relative return log-likelihood of each specification relative to the static equicorrelation structure. So a positive value (which is observed for all other specifications) corresponds to better performance than the static equicorrelation model. Both the in-sample and out-of-sample log-likelihood increases by adopting a more flexible correlation structure. The in-sample log-likelihood increases with the complexity of the model, as it to be expected. So the simplest model, the static equicorrelation model, has the smallest in-sample log-likelihood, whereas the "free" dynamic model with 36 latent variables to describe the correlation structure has the largest in-sample log-likelihood. Good in-sample fit need not translate into good out-of-sample fit. Nevertheless, we do find that the two specifications with the lowest in-sample fit (static and dynamic equicorrelation) also have the lowest out-of-sample fit. The two equicorrelation models have substantially lower performance than all of the the block and free-correlation specifications. Among the block equicorrelation models and the free correlation models it is evident that the dynamic models do better than the static models. This is not only true in-sample, but also out-of-sample. Interestingly, it is the dynamic block equicorrelation model that has the highest out-of-sample fit. So the better in-sample fit of the dynamic-free specification indeed does not translate into a better out-of-sample fit. In fact, it is barely better than the static block equicorrelation model. This is an indication that the most flexible specification with 36 latent variables to describe the correlation structure is overfitting the in-sample data.

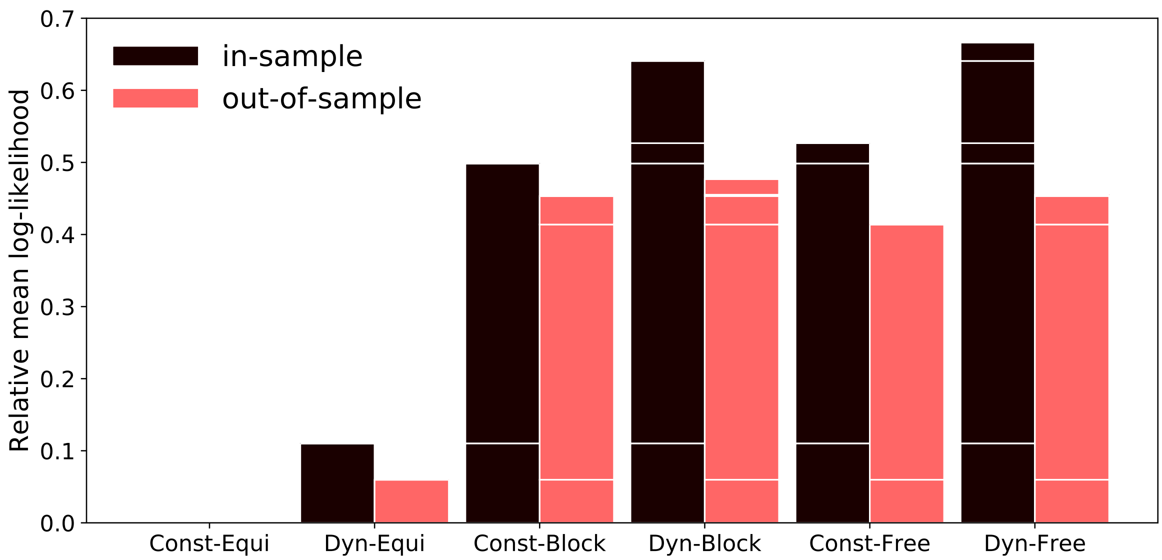

Figure 4: Average in-sample (black) and out-of-sample (red) return log-likelihoods, measured relative to the static equicorrelation specification. The horizontal white lines is the performance achieved by competing specifications. The horizontal white lines depict the performance of competing models. The in-sample period is January 2nd, 2002 to December 31st, 2013 (2993 trading days) and the out-of-sample period is January 2nd, 2014 to December 29th, 2017 (982 trading days).

Table 4 adds further information about the comparisons in terms of the return log-likelihood, and the statistical significance of the relative performance. We report the average per-period improvement of the return log-likelihood relative to the simplest specification with static equicorrelation. We evaluate the statistical significance of the relative return log-likelihoods using the model confidence set (MCS) by Hansen et al. (2011). We seek the specification with the largest expected out-of-sample partial log-likelihood, and the MCS is the subset of models that contains the best with a given level of confidence, after the original set of models has been trimmed by eliminating significantly inferior models in a sequential testing procedure. The MCS p-values are reported in parentheses and a small p-value is evidence that the model is significantly outperformed by other models in the comparison. The specifications in the 95% MCS are identified with bold font.

The relative average return log-likelihoods are reported for the in-sample period as a point of reference. Over the out-of-sample period, we find that dynamic block equicorrelation model has the best overall out-of-sample performance, but the MCS also includes the dynamic-free and the static-block specifications. The latter is included in the MCS because it performed particularly well in 2015. We also compare the specifications for each of calendar years in the out-of-sample period. For three of the four years the dynamic block equicorrelation specification has the best out-of-sample performance.

	Static correlation			Dynamic correlation		
Period	Equi	Block	Free	Equi	Block	Free
In-sample [2002-2013]	0.000	0.499	0.527	0.110	0.641	0.666
Out-of-sample [2014-2017]	0.000 *(0.00)*	**0.453** *(0.51)*	0.414 *(0.00)*	0.060 *(0.00)*	**0.477** *(1.00)*	**0.455** *(0.45)*
2014	0.000 *(0.00)*	**0.369** *(0.09)*	**0.360** *(0.09)*	0.029 *(0.01)*	**0.440** *(1.00)*	**0.420** *(0.52)*
2015	0.000 *(0.00)*	**0.663** *(1.00)*	0.570 *(0.00)*	-0.051 *(0.00)*	0.550 *(0.00)*	**0.561** *(0.19)*
2016	0.000 *(0.01)*	**0.366** *(0.95)*	**0.346** *(0.63)*	**0.070** *(0.06)*	**0.371** *(1.00)*	**0.336** *(0.63)*
2017	0.000 *(0.00)*	**0.412** *(0.24)*	**0.377** *(0.09)*	0.201 *(0.02)*	**0.551** *(1.00)*	**0.507** *(0.24)*

Table 4: Average out-of-sample partial log-likelihood relative to the constant equicorrelation model. Model confidence p-values are given in parentheses, and those significant at the 5% level are shaded gray. The comparisons are done for the entire out-of-sample period, 2014-2017, and for each of the years separately.

4.5.2 Global Minimum-Variance Portfolio

In this section, we evaluate and compare the ability of the estimated models to produce a low variance portfolio out-of-sample. At time $t-1$, we seek portfolio weights that minimize the variance of the portfolio return over the next period. Hence, for each of the specifications, we deduce the implied global minimum-variance (GMV) portfolio. The optimal portfolio weights solve

$$\min_{\omega_t \in \mathbb{R}^p} \omega_t' H_{(j),t} \omega_t, \qquad \text{s.t. } \omega_t' \iota = 1,$$

where $H_{(j),t}$ is the model-based conditional variance of the return vector, $r_t \in \mathbb{R}^p$, $j = 1, \ldots, 6$, and $\iota = (1, \ldots, 1)'$ is a p-dimensional vector of ones. In the absence of leverage constraints (such as no-shortening constraints) the well-know solution to this portfolio problem is:

$$\omega^*_{(j),t} = \frac{H^{-1}_{(j),t} \iota}{\iota' H^{-1}_{(j),t} \iota},$$

and the resulting portfolio returns are given by

$$R^{\mathrm{mv}}_{(j),t} = \omega^{*\prime}_{(j),t} r_t.$$

Because the different specifications produce different covariance matrices, $H_{(j),t}$, $j = 1,\ldots,6$, the resulting portfolio returns will differ, and may therefore have different variances and distributions. For illustrative purposes, we add the simple equal-weighted portfolio to the comparison. The returns of the equal-weighted portfolio are simply given by

$$R_{\mathrm{ew},t} = \frac{1}{p}\iota' r_t,$$

so that each asset is weighted by $1/p$, where $p = 9$ in this application.

We will compare the variance of the seven portfolios, to evaluate whether the different specifications produce substantially different portfolio variances. We report the in-sample and out-of-sample variances of the seven portfolios in the upper panel of Table 5. The first observation we make is that equal-weighted portfolio has a substantially larger variance than any of the model-based portfolios. This is not unexpected because the equal-weighted portfolio does not use any information about the covariance structure. On the other hand, the fact that the model-based portfolios can reduce the variance by a factor of two is evidence that all of the multivariate realized GARCH specifications produce sensible and valuable forecasts of the conditional covariance matrix, H_t.

Among the six model-based portfolios it is the most flexible specification (dynamic-free) that has the smallest out-of-sample variance, in the (full) out-of sample period. The portfolio based on the dynamic-block specification comes in as a close second, however all but the simplest static equicorrelation specification are found in the model confidence set. We also present results for each of the years in the out-of-sample period, and while there is some variation from year to year, the dynamic-free portfolio end up in the MCS every year.

A variance comparison is based on squared returns, which is sensitive to outliers. We therefore supplement the comparison with a comparison of absolute portfolio returns. These results are presented in the lower panel of Table 5, where we present the mean of absolute daily returns, measured in annualized units. These results, including the model confidence sets are almost identical to the results based on squared returns.

	Period	Equal weights	Static correlation Equi	Static correlation Block	Static correlation Free	Dynamic correlation Equi	Dynamic correlation Block	Dynamic correlation Free
Squared returns	In-sample [2002-2013]	2.149	1.117	1.091	1.077	1.063	0.990	0.991
	Out-of-sample [2014-2017]	0.952	0.506	**0.475**	**0.466**	**0.499**	**0.471**	**0.464**
		(0.00)	*(0.01)*	*(0.21)*	*(0.93)*	*(0.14)*	*(0.66)*	*(1.00)*
	2014	0.657	0.588	**0.553**	**0.546**	**0.577**	**0.542**	**0.525**
		(0.02)	*(0.02)*	*(0.25)*	*(0.25)*	*(0.10)*	*(0.25)*	*(1.00)*
	2015	1.393	0.695	**0.580**	**0.596**	0.712	**0.606**	**0.629**
		(0.00)	*(0.01)*	*(1.00)*	*(0.27)*	*(0.01)*	*(0.27)*	*(0.15)*
	2016	1.385	**0.471**	0.498	**0.465**	**0.491**	0.525	**0.496**
		(0.00)	*(0.86)*	*(0.02)*	*(1.00)*	*(0.83)*	*(0.02)*	*(0.86)*
	2017	0.324	0.251	0.254	0.240	**0.194**	**0.191**	**0.188**
		(0.00)	*(0.00)*	*(0.00)*	*(0.01)*	*(0.76)*	*(0.76)*	*(1.00)*
Abs. returns (annual)	In-sample [2002-2013]	0.158	0.113	0.111	0.110	0.111	0.107	0.106
	Out-of-sample [2014-2017]	0.114	0.085	**0.082**	**0.081**	0.082	**0.080**	**0.079**
		(0.00)	*(0.00)*	*(0.11)*	*(0.15)*	*(0.02)*	*(0.54)*	*(1.00)*
	2014	0.096	0.095	**0.090**	**0.089**	0.093	**0.087**	**0.087**
		(0.01)	*(0.00)*	*(0.18)*	*(0.18)*	*(0.01)*	*(0.57)*	*(1.00)*
	2015	0.145	0.099	**0.092**	**0.094**	0.100	**0.094**	**0.097**
		(0.00)	*(0.00)*	*(1.00)*	*(0.12)*	*(0.00)*	*(0.35)*	*(0.06)*
	2016	0.141	**0.081**	0.083	**0.080**	**0.081**	0.082	**0.078**
		(0.00)	*(0.40)*	*(0.01)*	*(0.41)*	*(0.40)*	*(0.04)*	*(1.00)*
	2017	0.071	0.061	0.062	0.060	**0.054**	**0.053**	**0.053**
		(0.00)	*(0.00)*	*(0.00)*	*(0.00)*	*(0.70)*	*(0.80)*	*(1.00)*

Table 5: The average of squared returns (top panel) and absolute returns (bottom panel) for the equal-weighted portfolio and the six global minimum portfolio. The in-sample period is January 2nd, 2002 to December 31st, 2013 (3021 trading days) and the out-of-sample period is January 2nd, 2014 to December 29th, 2017 (989 trading days). We also report results for each of the calendar years in the out-of-sample period. Numbers in parentheses are MCS p-values and bold numbers indicate specifications that are in the 95% MCS.

Based on this comparison, there is further evidence that a dynamic and flexible correlation modeling lead to significant improvements out-of-sample. Relative to the simplest specification, the static equicorrelation specification, we observe a variance reduction of about 50 basis points in annual volatility. In practice, one would also want to factor in portfolio turnover, because transaction costs may offset the the value of the variance reduction. This would entails a model comparison based on a different criterion. We leave this for future research.

5 Conclusion and Discussion

In this paper, we have introduced a novel framework for multivariate GARCH modeling. The framework is based on a new parametrization of correlation matrices, where the $p \times p$ correlation matrix, C_t is represented by the $d = \frac{p(p-1)}{2}$ dimensional vector, $\varrho_t = g(C_t)$. We make use of realized measures by adopting the structure of Realized GARCH models, where measurement equations are used to link realized measures of volatility with the corresponding latent variables. In addition to realized measures of variance, we also make use the realized correlation matrix in the dynamic modeling. We label the model a Multivariate Realized GARCH model, and observe that it is a natural extension of previous Realized GARCH models. In the special case where $p = 2$ (and hence $d = 1$), the proposed model is identical to the bivariate model Hansen et al. (2014), and the case with $p = 1$ (i.e. $d = 0$), we have a variant of the original univariate Realized GARCH model by Hansen et al. (2012).

In its most general form, the parametrization, $\varrho_t = g(C_t)$, does not impose any restrictions on C_t, aside from positive definiteness. However, in many situations it will be desirable to impose an additional structure on C_t in order to simplify the estimation problem and improve the out-of-sample performance of the estimated model. We have shown how this can be achieved with a simple linear structure. Specifically, the case $\varrho_t = A\zeta_t$, where A is a matrix and ζ_t is a vector of latent variables with a lower dimension than ϱ_t. This linear structure emerges naturally when C_t has a block equicorrelation structure, as we have shown in Theorem 1. Because the block structure of C_t carries over to $\log C_t$, the vector, ϱ_t, will have elements in common, which can be expressed as $A\zeta_t$, where A is a known matrix with zeros and ones. The proposed model is not specific to the case with a block equicorrelation matrix, but is more general. The proposed framework paves the way to explore a broad family of models for the correlation structure. For instance, other choices for A could be entertained, and different A-matrices will induce different (parsimonious) structures on C_t. It is also possible to explore data dependent choices for A, and it is also possible to explore factor structures, $\varrho_t = \varrho(\zeta_t)$, beyond the linear structure we have focused on in this paper. We leave these generalizations for future research.

We applied the Multivariate Realized GARCH model to nine assets from three different economic sectors, and compared static and dynamic correlation models using one of three structures for C_t: equicorrelation, block equicorrelation, and “free”. The empirical results favors a dynamic specifications C_t, and strongly prefer the block structure and the free structure over the equicorrelation structure. It is encouraging that the dynamic block equicorrelation specification performs well empirically, in-sample as well as out-of-sample, because this model can be scaled to higher dimensions. Scaling the dynamic free specification to high dimensions will be difficult, if not impossible, because the dimension of the latent variable, ϱ_t, is of order p^2.

If needed, there are many ways to speed up the estimation by tweaking the implementation or adopting a two-stage estimation method. The results presented in this paper were all based on maximum likelihood where all parameters are estimated jointly. A two-stage estimation method could greatly speed up the estimation by first estimating the conditional variance series using univariate realized GARCH models and then, in a second stage, estimate the parameters related to dynamic correlation structure while the first-stage estimates, and hence $h_{1,t}, \ldots, h_{d,t}$, $t = 1, \ldots, T$, are taken as given.

References

Aielli, G. P. (2013). Dynamic conditional correlation: on properties and estimation. *Journal of Business and Economic Statistics*, 31:282–299.

Amisano, G. and Giacomini, R. (2007). Comparing Density Forecasts via Weighted Likelihood Ratio Tests. *Journal of Business & Economic Statistics*, 25:177–190.

Andersen, T. G. and Bollerslev, T. (1998). Answering the skeptics: Yes, standard volatility models do provide accurate forecasts. 39(4):885–905.

Andersen, T. G., Bollerslev, T., Diebold, F. X., and Ebens, H. (2001a). The distribution of realized stock return volatility. *Journal of Financial Economics*, 61:43–76.

Andersen, T. G., Bollerslev, T., Diebold, F. X., and Labys, P. (2001b). The distribution of realized exchange rate volatility. *Journal of the American Statistical Association*, 96:42–55.

Andersen, T. G., Bollerslev, T., Diebold, F. X., and Labys, P. (2003). Modeling and forecasting realized volatility. *Econometrica*, 71:579–625.

Archakov, I. and Hansen, P. R. (2020). A new parametrization of correlation matrices. *https://sites.google.com/site/peterreinhardhansen/*.

Asai, M. and So, M. (2015). Long memory and asymmetry for matrix-exponential dynamic correlation processes. *Journal of Time Series Econometrics*, 7:69–74.

Barndorff-Nielsen, O. E., Hansen, P. R., Lunde, A., and Shephard, N. (2009). Realized kernels in practice: trades and quotes. *Econometrics Journal*, 12:C1–C32.

Barndorff-Nielsen, O. E., Hansen, P. R., Lunde, A., and Shephard, N. (2011). Multivariate realised kernels: consistent positive semi-definite estimators of the covariation of equity prices with noise and non-synchronous trading. *Jounal of Econometrics*, 162:149–169.

Barndorff-Nielsen, O. E. and Shephard, N. (2002a). Estimating quadratic variation using realized variance. *Journal of Applied Econometrics*, 17(5):457–477.

Barndorff-Nielsen, O. E. and Shephard, N. (2002b). How accurate is the asymptotic approximation to the distribution of realised variance. Economics Series Working Papers 2001-W16, University of Oxford, Department of Economics.

Barndorff-Nielsen, O. E. and Shephard, N. (2004a). Econometric analysis of realized covariation: High frequency based covariance, regression, and correlation in financial economics. *Econometrica*, 72:885–925.

Barndorff-Nielsen, O. E. and Shephard, N. (2004b). Econometric Analysis of Realized Covariation: High Frequency Based Covariance, Regression, and Correlation in Financial Economics. *Econometrica*, 72(3):885–925.

Bauer, G. H. and Vorkink, K. (2011). Forecasting multivariate realized stock market volatility. *Journal of Econometrics*, 160:93–101.

Bauwens, L., Storti, G., and Violante, F. (2012). Dynamic conditional correlation models for realized covariance matrices. *Working Paper*, (2012060).

Black, F. (1976). *Studies of Stock Price Volatility Changes*. Proceedings of the 1976 Meetings of the American Statistical Association, Business and Economic Statistics Section.

Bollerslev, T. (1990). Modelling the coherence in short-run nominal exchange rates: A multivariate generalized ARCH model. *The Review of Economics and Statistics*, 72:498–505.

Chiriac, R. and Voev, V. (2011). Modelling and forecasting multivariate realized volatility. *Journal of Applied Econometrics*, 26:922–947.

Chiu, T., Leonard, T., and Tsui, K.-W. (1996). The matrix-logarithmic covariance model. *Journal of the American Statistical Association*, 91:198–210.

Christie, A. A. (1982). The stochastic behavior of common stock variances : Value, leverage and interest rate effects. *Journal of Financial Economics*, 10:407–432.

Dumitrescu, E. and Hansen, P. (2017). Forecasting exchange rate volatility: Multivariate realized GARCH framework. *Working Paper*.

Engle, R. and Kelly, B. (2012). Dynamic Equicorrelation. *Journal of Business & Economic Statistics*, 30(2):212–228.

Engle, R. F. (1982). Autoregressive conditional heteroskedasticity with estimates of the variance of U.K. inflation. *Econometrica*, 45:987–1007.

Engle, R. F. (2002a). Dynamic Conditional Correlation: A Simple Class of Multivariate Generalized Autoregressive Conditional Heteroskedasticity Models. *Journal of Business & Economic Statistics*, 20(3):339–350.

Engle, R. F. (2002b). New frontiers for arch models. *Journal of Applied Econometrics*, 17(5):425–446.

Engle, R. F. and Gallo, G. (2006). A multiple indicators model for volatility using intra-daily data. *Journal of Econometrics*, 131:3–27.

Engle, R. F., Ledoit, O., and Wolf, M. (2019). Large dynamic covariance matrices. *Journal of Business and Economic Statistics*, 37:363–375.

Engle, R. F. and Ng, V. K. (1993). Measuring and Testing the Impact of News on Volatility. *Journal of Finance*, 48(5):1749–78.

Engle, R. F. and Sheppard, K. (2001). Theoretical and Empirical properties of Dynamic Conditional Correlation Multivariate GARCH. NBER Working Papers 8554, National Bureau of Economic Research, Inc.

Geweke, J. and Amisano, G. (2010). Comparing and evaluating Bayesian predictive distributions of asset returns. *International Journal of Forecasting*, 26(2):216–230.

Golosnoy, V., Gribisch, B., and Liesenfeld, R. (2012). The conditional autoregressive Wishart model for multivariate stock market volatility. *Journal of Econometrics*, 167(1):211–223.

Gorgi, P., Hansen, P. R., Janus, P., and Koopman, S. J. (2019). Realized Wishart-GARCH: A score-driven multi-asset volatility model. *Journal of Financial Econometrics*, 17:1–32.

Hansen, P. R. and Huang, Z. (2016). Exponential GARCH modeling with realized measures of volatility. *Journal of Business & Economic Statistics*, 34(2):269–287.

Hansen, P. R., Huang, Z., and Shek, H. (2012). Realized GARCH: A joint model of returns and realized measures of volatility. *Journal of Applied Econometrics*, 27:877–906.

Hansen, P. R. and Lunde, A. (2011). *Forecasting Volatility Using High-Frequency Data.* The Oxford Handbook of Economic Forecasting.

Hansen, P. R., Lunde, A., and Nason, J. M. (2011). The Model Confidence Set. *Econometrica*, 79(2):453–497.

Hansen, P. R., Lunde, A., and Voev, V. (2014). Realized beta GARCH: A multivariate GARCH model with realized measures of volatility. *Journal of Applied Econometrics*, 29:774–799.

Ishihara, T., Omori, Y., and Asai, M. (2016). Matrix exponential stochastic volatility with cross leverage. *Computational Statistics & Data Analysis*, 100:331–350.

Kawakatsu, H. (2006). Matrix exponential GARCH. *Journal of Econometrics*, 134:95–128.

Liu, Q. (2009). On portfolio optimization: How and when do we benefit from high-frequency data? *Journal of Applied Econometrics*, 24(4):560–582.

Lunde, A. and Olesen, K. V. (2014). Modeling and forecasting the volatility of energy forward returns. *Working Paper.*

Noureldin, D., Shephard, N., and Sheppard, K. (2012). Multivariate high-frequency-based volatility (HEAVY) models. *Journal of Applied Econometrics*, 27:907–933.

Roustant, O. and Deville, Y. (2017). On the validity of parametric block correlation matrices with constant within and between group correlations. *arXiv*, math.ST/1705.09793.

Shephard, N. and Sheppard, K. (2010). Realising the future: forecasting with high-frequency-based volatility (HEAVY) models. *Journal of Applied Econometrics*, 25(2):197–231.

Shephard, N., Sheppard, K., and Engle, R. F. (2020). Fitting vast dimensional time-varying covariance models. *Journal of Business and Economic Statistics.*

Weigand, R. (2014). Matrix Box-Cox models for multivariate realized volatility. *Working Paper.*

White, H. (1994). *Estimation, Inference and Specification Analysis.* Cambridge University Press, Cambridge.

A Proofs

Lemma A.1. *If A and B are $p \times p$ quasi block matrices with identical block structures, then $A+B$ and AB are quasi block matrices with the same block structure as A and B. Moreover, if A is invertible then A^{-1} is also a quasi block matrix with the same block structure as A.*

Proof. The result for $A+B$ is obvious. The result for the product, AB, can be established as follows. Let $s_i \times s_j$ denote the dimension of the (i, j)-th block. Then it follows (using the same reasoning as in the proof of Theorem 2) that the blocks of A and B have the following representation,

$$A_{[i,j]} = a_{ij} P_{[i,j]} + 1_{\{i=j\}} \alpha_i P_{[i,i]}^{\perp} \qquad \text{and} \qquad B_{[i,j]} = b_{ij} P_{[i,j]} + 1_{\{i=j\}} \beta_i P_{[i,i]}^{\perp},$$

where $P_{[i,j]} = \frac{1}{\sqrt{s_i s_j}} \iota_{s_i} \iota_{s_j}'$ and $P_{[i,i]}^{\perp} = I_{s_i} - P_{[i,i]}$, and a_{ij}, b_{ij}, α_i, and β_i, are some real scalars $i, j = 1, \ldots, K$. The (i, j)-th block of the matrix product, AB, is given by $(AB)_{[i,j]} = \sum_{k=1}^{K} A_{[i,k]} B_{[k,j]}$. From the structure above it follows that we for $i \neq j$ have

$$\begin{aligned} (AB)_{[i,j]} &= \sum_{k=1}^{K} a_{ik} P_{[i,k]} b_{kj} P_{[k,j]} + \alpha_i P_{[i,i]}^{\perp} b_{ij} P_{[i,j]} + a_{ij} P_{[i,j]} \beta_j P_{[j,j]}^{\perp} \\ &= \left(\sum_{k=1}^{K} a_{ik} b_{kj} \right) P_{[i,j]}, \end{aligned}$$

where we used that $P_{[i,i]} P_{[i,i]}^{\perp} = 0$ and that $P_{[i,k]} P_{[k,j]} = P_{[i,j]}$. This shows that all entries of $(AB)_{[i,j]}$ are identical and equal to $\frac{1}{\sqrt{s_i s_j}} \sum_{k=1}^{K} a_{ik} b_{kj}$. Similarly, for $i = j$ we have

$$\begin{aligned} (AB)_{[i,i]} &= \sum_{k=1}^{K} a_{ik} P_{[i,k]} b_{ki} P_{[k,i]} + \alpha_i P_{[i,i]}^{\perp} \beta_i P_{[i,i]}^{\perp} \\ &= \left(\sum_{k=1}^{K} a_{ik} b_{ki} \right) P_{[i,i]} + \alpha_i \beta_i P_{[i,i]}^{\perp} = (d_{ii} - \delta_i) P_{[i,i]} + \delta_i I_{s_i}, \end{aligned}$$

where $d_{ii} = \sum_{k=1}^{K} a_{ik} b_{ki}$ and $\delta_i = \alpha_i \beta_i$. So it follows that $(AB)_{[i,i]}$ has the required quasi block structure, where off-diagonal elements equal $(d_{ii} - \delta_i)/s_i$ and diagonal elements equal $(d_{ii} - \delta_i)/s_i + \delta_i$.

For the inverse, A^{-1}, we simply verify that

$$D_{[i,j]}^{-1} = a_{ij}^{\#} P_{[i,j]} + 1_{\{i=j\}} \tfrac{1}{\alpha_i} P_{[i,i]}^{\perp},$$

is the inverse of $a_{ij} P_{[i,j]} + 1_{\{i=j\}} \alpha_i P_{[i,i]}^{\perp}$ where $a_{ij}^{\#}$ is the (i, j)-th elements of A^{-1} – the inverse of $A = \{a_{ij}\}_{i,j=1,\ldots,K}$.

For the diagonal blocks

$$\begin{aligned}(AD^{-1})_{[i,i]} &= \sum_{k=1}^{K} a_{ik} P_{[ik]} a^{\#}_{ki} P_{[k,i]} + \alpha_i P^{\perp}_{[i,i]} \tfrac{1}{\alpha_i} P^{\perp}_{[i,i]} \\ &= \sum_{k=1}^{K} a_{ik} a^{\#}_{ki} P_{[i,i]} + P^{\perp}_{[i,i]} = P_{[i,i]} + P^{\perp}_{[i,i]},\end{aligned}$$

which equals I_{s_i} as needed. For the off-diagonal blocks

$$\begin{aligned}(AD^{-1})_{[i,j]} &= \sum_{k=1}^{K} a_{ik} P_{[ik]} a^{\#}_{kj} P_{[k,j]} + \alpha_i P^{\perp}_{[i,i]} a^{\#}_{ij} P_{[i,j]} + a_{ij} P_{[i,j]} \frac{1}{\alpha_j} P^{\perp}_{[j,j]} \\ &= \sum_{k=1}^{K} a_{ik} a^{\#}_{kj} P_{[i,j]} + 0 + 0 = 0,\end{aligned}$$

as required. This completes the proof. $\square$

Proof of Theorem 1. If C is a quasi block matrix, then so is $C-I$, and it follows by Lemma A.1 that so is $(C-I)(C-I)$. By induction, so is $(C-I)^k$. Lemma also tells us that $(I+C)^{-1}$ and hence $(I+C)^{-k}$, have the same block structure as C. Now $\log C = 2\sum_{k\geq 1} \frac{1}{2k+1}[(C-I)(C+I)]^{2k+1}$, see Hickham, and since each of the terms $[(C-I)(C+I)]^{2k+1}$ has the same block structure as C, it follows that $\log C$ also has the same block structure as C. $\square$

Proof of Theorem 2. The elements of $P_{[i,j]}$ are all identical and equal to $\frac{1}{\sqrt{s_i s_j}}$. For $i \neq j$ the common elements of $C_{[i,j]}$ is ρ_{ij}, so it must equal $a_{ij}/\sqrt{s_i s_j}$, which proves that $a_{ij} = \rho_{ij}\sqrt{s_i s_j}$ for $i \neq j$. To simplify the notation, we set $b_i = 1 - \rho_{ii}$. For $i = j$ we have $C_{[i,i]} = (a_{ii} - b_i)P_{[i,i]} + b_i I_{s_i}$, whose off-diagonal elements and diagonal must equal ρ_{ii} and 1, respectively. The former implies $\rho_{ii} = (a_{ii} - b_i)/s_i$ and the latter implies $(a_{ii} - b_i)/s_i + b_i = 1$. Therefore it follows that $b_i = 1 - \rho_{ii}$ and $a_{ii} = 1 + (s_i - 1)\rho_{ii}$, as stated in the Theorem.

The eigenvalues and determinant of C are obtained by a rotation of C using the following orthonormal matrix. Let

$$Q = \begin{bmatrix} \frac{1}{\sqrt{s_1}}\iota_{s_1} & 0 & \cdots & & i_{s_1\perp} & 0 & \cdots & 0 \\ 0 & \frac{1}{\sqrt{s_2}}\iota_{s_2} & & & 0 & i_{s_1\perp} & & \vdots \\ \vdots & & \ddots & & & & \ddots & \\ 0 & \cdots & & \frac{1}{\sqrt{s_K}}\iota_{s_K} & 0 & \cdots & & i_{s_1\perp} \end{bmatrix},$$

where $\iota_{s_i\perp}$ is some $s_i \times (s_i - 1)$ matrix for which $\iota'_{s_i}\iota_{s_i\perp} = 0$ and $\iota'_{s_i\perp}\iota_{s_i\perp} = I_{s_i-1}$. Then it follows that $Q'Q = QQ' = I$, so that the eigenvalues of C are identical to those of,

$$Q'CQ = \begin{bmatrix} a_{11} & \cdots & a_{1K} & 0 & \cdots & & 0 \\ \vdots & \ddots & \vdots & \vdots & & & \vdots \\ a_{K1} & \cdots & a_{KK} & 0 & & & \\ 0 & \cdots & 0 & (1-\rho_{11})I_{s_1-1} & \ddots & & \\ \vdots & & & \ddots & \ddots & 0 & \\ & & & & (1-\rho_{K-1,K-1})I_{s_{K-1}-1} & & 0 \\ 0 & \cdots & & & & 0 & (1-\rho_{kk})I_{s_K-1} \end{bmatrix}.$$

This also established the result for the determinant of C, which can also be obtained directly with $\det C = \det(CQQ') = \det(Q'CQ) = (\det A) \cdot (1-\rho_{11})^{s_1-1} \cdots (1-\rho_{KK})^{s_k-1}$.

Finally, we verify the expression for C^{-1} . We multiplying C with the block matrix

$$D_{[i,j]} = a^{\#}_{i,j} P_{[i,j]} + 1_{\{i=j\}} \tfrac{1}{1-\rho_{ii}} P^{\perp}_{[i,i]}, \qquad i,j = 1,\ldots,K,$$

where $P^{\perp}_{[i,i]} = I_{s_i} - P_{[i,i]}$, and verify that the outcome is the required identity matrix. Let $M_{[i,j]}$ denote the (i,j)-th block of CD. Then we have

$$M_{[i,i]} = \sum_{k=1}^{K} a_{ik} P_{[ik]} a^{\#}_{ki} P_{[k,i]} + (1-\rho_{ii}) P^{\perp}_{[i,i]} \tfrac{1}{1-\rho_{ii}} P^{\perp}_{[i,i]} = \sum_{k=1}^{K} a_{ik} a^{\#}_{ki} P_{[i,i]} + P^{\perp}_{[i,i]},$$

where we used that $P_{[i,i]}$ and $P^{\perp}_{[i,i]} = I - P_{[i,i]}$ are orthogonal projection matrices, so that $P_{[i,i]} P^{\perp}_{[i,i]} = P_{[i,i]}(I - P_{[i,i]}) = 0$. Since $a^{\#}_{ki}$ are the elements of the A^{-1} we have $\sum_{k=1}^{K} a_{ik} a^{\#}_{ki} = 1$, so that $M_{[i,i]} = I_{s_i}$. Next, for $i \neq j$ we have

$$M_{[i,j]} = \sum_{k=1}^{K} a_{ik} P_{[i,k]} a^{\#}_{kj} P_{[k,j]} + \frac{a_{ij}}{b_j} P_{ij} P^{\perp}_{[j,j]} + b_i a^{\#}_{ij} P^{\perp}_{[i,i]} P_{[i,j]} = \sum_{k=1}^{K} a_{ik} a^{\#}_{kj} P_{[i,j]},$$

where we used that $P_{[i,k]} P_{[k,j]} = P_{[i,j]}$ and

$$P_{ij} P^{\perp}_{[j,j]} = P_{[i,j]}(I_{s_j} - P_{[j,j]}) = \tfrac{1}{\sqrt{s_i s_j}} \iota_{s_i} \iota'_{s_j} (I_{s_j} - \tfrac{1}{s_j} \iota_{s_j} \iota'_{s_j}) = \tfrac{1}{\sqrt{s_i s_j}} \iota_{s_i} (\iota'_{s_j} - \iota'_{s_i}) = 0.$$

The second result now follows since $a^{\#}_{ki}$ are the elements of the A^{-1}, so that $\sum_{k=1}^{K} a_{ik} a^{\#}_{kj} = 0$ for $i \neq j$, and consequently, $M_{[i,j]} = 0$ for $i \neq j$. This completes the proof. $\square$

B Estimation Procedure in Details

Here we provide a detailed description of the estimation procedure for the Multivariate Realized GARCH model in the situation where $\varrho_t = A\zeta_t$. In the unrestricted case ($A = I$), we have $\varrho_t = \zeta_t$. The data consists of the p-dimensional return vector, r_t, for $t = 1, ..., T$ and the corresponding transformed realized measures. The realized measures consist of the p-dimensional vector of realized variances, x_t, and the k-dimensional vector of transformed correlations $\check{y}_t = (A'A)^{-1}A'y_t$, where the "free" specification ($A = I$) has $\check{y}_t = y_t$ and $k = d = p(p-1)/2$.

The log-likelihood is evaluated using the following steps:

1. Initialize all parameters except Σ, and let these be represented by θ. Also initialize values for the conditional variances and the transformed conditional correlations, h_1 and ζ_1. We will maximize the log-likelihood with respect to (θ, h_1, ζ_1). An alternative approach is to fix values for h_1 and ζ_1, such as their unconditional mean.

2. Compute $h_t(\theta)$ and $\zeta_t(\theta)$ using the GARCH equations, and map the latter to $\varrho_t = A\zeta_t$ and then to $C_t(\theta)$, for $t = 1, ..., T$. The mapping from ϱ_t to C_t can be done with the algorithm provided in Archakov and Hansen (2020, corollary 1).

3. Compute the standardized returns, $z_t(\theta)$, and the measurement errors $u_t(\theta)$, for $t = 1, ..., T$.

4. Compute $\hat{\Sigma}(\theta) = T^{-1}\sum_t u_t(\theta)u_t(\theta)'$ and evaluate the log-likelihood using (8).

The model can now be estimated by maximizing (8) with respect to θ, h_1, and ζ_1, by repeating steps 2, 3 and 4 every time (θ, h_1, ζ_1) has been updated.

C Additional Empirical Results

In Figure C.1, we plot the 36 realized correlation series (grey lines) as well as the average realized correlation (red line). While it is not possible to identify the paths of the individual correlation series in the Figure, it does reveals a great deal of dispersion across the individual correlations relative to the average correlation series.

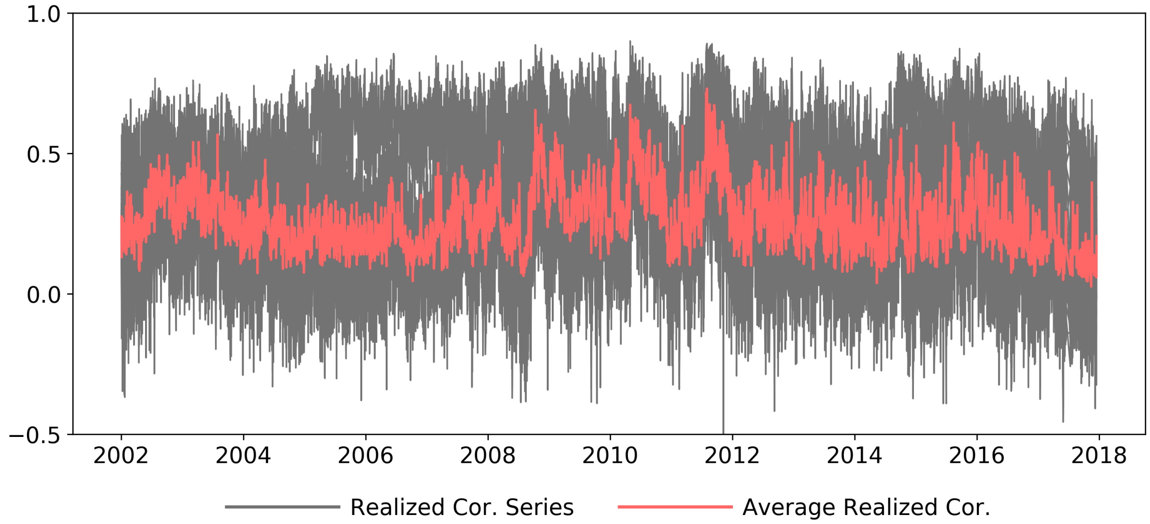

Figure C.1: Daily realized correlations for the full sample period. The realized correlations for the 36 correlations series (grey lines) are computed from the multivariate Realized Kernel estimator. The red line is the time series with the daily average correlation.

The estimated time series for equicorrelation matrix is presented in Figure C.2. The red line is the average realized correlation, and the black line is the time series of ρ_t, as produced by the estimated dynamic equicorrelation model. The estimate of the constant equicorrelation model is illustrated with the dashed line.

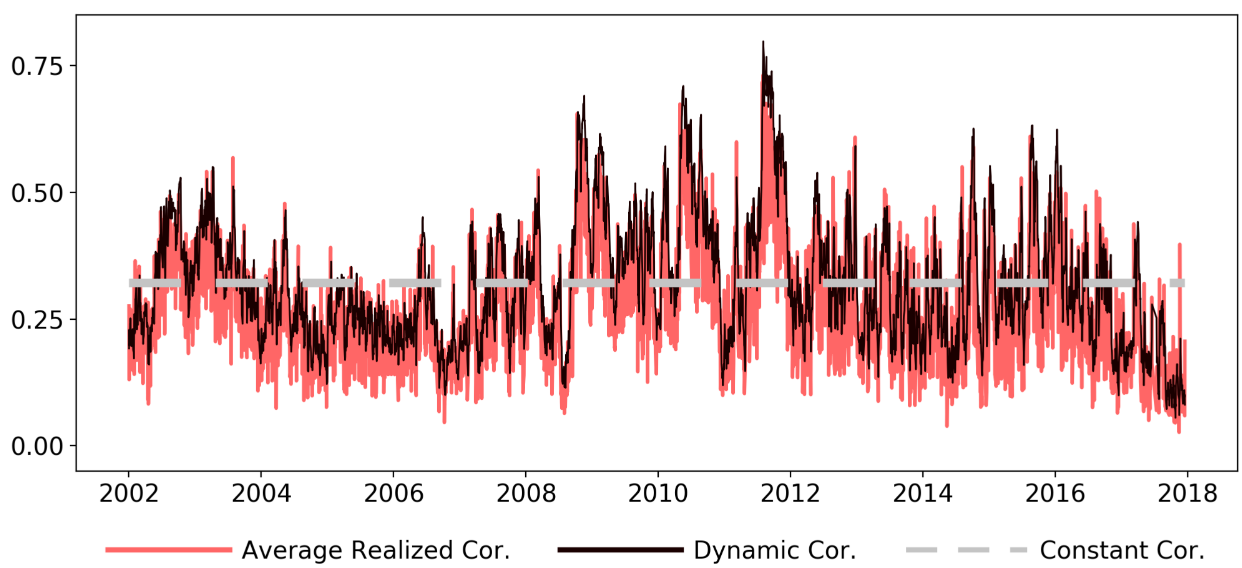

Figure C.2: The common correlation coefficient deduced from the estimated equicorrelation model (black line) and the daily average of the 36 realized correlations (red line). The white dashed line represents the estimate of the common correlation coefficient produced by the constant equicorrelation structure.