

TI 2021-083/IV
Tinbergen Institute Discussion Paper

Dynamic Conditional Correlations with Partial Information Pooling

Revision: March 2023

Bram van Os¹
Dick van Dijk¹

Tinbergen Institute is the graduate school and research institute in economics of Erasmus University Rotterdam, the University of Amsterdam and Vrije Universiteit Amsterdam.

Contact: discussionpapers@tinbergen.nl

More TI discussion papers can be downloaded at <https://www.tinbergen.nl>

Tinbergen Institute has two locations:

Tinbergen Institute Amsterdam
Gustav Mahlerplein 117
1082 MS Amsterdam
The Netherlands
Tel.: +31(0)20 598 4580

Tinbergen Institute Rotterdam
Burg. Oudlaan 50
3062 PA Rotterdam
The Netherlands
Tel.: +31(0)10 408 8900

Dynamic Conditional Correlations with Partial Information Pooling*

Bram van Os[†]

Department of Econometrics, Erasmus University Rotterdam
and

Dick van Dijk

Department of Econometrics, Erasmus University Rotterdam

March 27, 2023

Abstract

We propose a novel Dynamic Conditional Correlation model with Conditional Linear Information Pooling (CLIP-DCC) which endogenously determines an optimal degree of commonality in the correlation innovations. Effectively, this allows a part of the update of each individual correlation to parsimoniously depend on the information contained in all asset return pairs. In contrast to existing approaches, such as the Dynamic EquiCorrelation (DECO) model, the CLIP-DCC model does not restrict long-run behavior, thereby naturally complementing target correlation matrix shrinkage approaches. Empirical findings suggest substantial benefits for a minimum-variance investor in real-time. Combining the CLIP-DCC model with target shrinkage yields additive improvements, confirming that they address distinct parts of uncertainty of the conditional correlation matrix.

Keywords: Dynamic Equicorrelation, High-dimensional covariance matrices, Covariance shrinkage, Minimum-variance portfolio.

*We would like to thank Christian Brownlees, Massimiliano Caporin, Peter Hansen, Onno Kleen, Oliver Linton, Jordi Llorens-Terrazas, Rasmus Lönn, Daan Opschoor, Andrew Patton, Timo Teräsvirta, Michel van der Wel and the participants at the IAAE 2021 annual conference (online, June 2021), the 27th CEF conference (online, June 2021), the 4th annual workshop on Financial Econometrics (Örebro, November 2021), the 29th SNDE conference (online, March 2022), the VieCo conference (Copenhagen, June 2022), the 4th QFFE conference (Marseille, June 2022) and the 14th annual SoFiE conference (Cambridge, June 2022) for providing useful comments and suggestions. Any remaining errors are our own.

[†]Address for correspondence: Department of Econometrics, Erasmus University Rotterdam, Burge-meester Oudlaan 50, 3062 PA Rotterdam, The Netherlands. E-mail: vanos@ese.eur.nl

1 Introduction

The conditional covariance matrix of asset returns is of crucial importance for portfolio construction and risk management, among others. Multivariate ARCH-type models, in particular the Dynamic Conditional Correlation (DCC) model by Engle (2002), have become a popular tool to model and forecast time-varying conditional covariance matrices. However, the accuracy of covariance matrix estimates is known to deteriorate as the number of assets grows. A key contributing factor within the DCC model is the noisy nature of the outer product of the standardized returns that is used to update the correlations. As noted by Engle and Kelly (2012), this causes the DCC update of each individual conditional correlation to only draw from its own cross-product of asset returns, leaving the (potentially useful) information on the correlations between other assets untapped.

In this paper, we propose a DCC model with Conditional Linear Information Pooling (CLIP-DCC) to better handle a large cross-section of assets. Inspired by the Dynamic EquiCorrelation (DECO) model of Engle and Kelly (2012), the CLIP-DCC model introduces cross-sectional commonality in the update of the conditional correlations. Specifically, the CLIP-DCC model can be represented as a DCC model with an additional ‘pooled’ innovation term based on the average cross-product of all pairs of asset returns. This allows the CLIP-DCC update of any correlation to parsimoniously exploit the information available in the full cross-section of assets. The pooling strength can be determined endogenously during likelihood estimation. The CLIP-DCC model offers two important advantages over the DECO model. First, the CLIP-DCC model still allows for cross-sectional differences in the conditional correlations, instead of fully equating all of them, which may be empirically unrealistic. Second, the CLIP-DCC model does not impose (implicit) restrictions on the long-run correlation matrix, whereas the DECO model structures it to be an equicorrelation matrix as well. By preserving the long-run correlation matrix the CLIP-DCC model complements approaches that shrink the correlation targeting matrix (Engle et al., 2019). In fact, we may combine the CLIP-DCC model with target shrinkage, allowing for the separate and independent handling of the uncertainty of the long-run correlation matrix and the short-run movements around it.

Furthermore, we show that the pooling aspect of the CLIP-DCC model can be further

refined using a block structure, similar to how the Block-DECO model loosens the DECO constraint, see Engle and Kelly (2012). An important theoretical contribution is that we prove that the resulting block-based pooling operation preserves positive definiteness. Practically, this means that the CLIP-DCC framework can easily incorporate relevant group structure information if available.

We demonstrate in a Monte Carlo exercise that Composite Likelihood (CL) can be used to reliably estimate the parameters in the CLIP-DCC model, in line with the findings of Pakel et al. (2021) for the DCC model. We empirically evaluate the performance of our CLIP-DCC model in an application to daily US large-cap stock returns for the period February 1981 until December 2020, constructing global minimum variance portfolios (GMVP) in real-time. For a wide span of portfolio dimensions, ranging from 10 to 500 stocks, we find that the CLIP-DCC model significantly reduces out-of-sample portfolio variance compared to the DCC model, whereby the relative improvements increase as the portfolio dimension grows. In addition, the CLIP-DCC model greatly outperforms the DCC model during periods of financial turmoil, such as the 2008 Great Financial Crisis. Combining the CLIP-DCC model with non-linear shrinkage (NLS) of the target, using the method of Ledoit and Wolf (2020), further improves performance. For example, for the portfolio dimension $N = 500$ we obtain an out-of-sample GMVP annualized standard deviation of 6.40(6.62) and 6.12(6.37) for the DCC and the CLIP-DCC model with(out) NLS of the target, respectively. This confirms that these methods address different sources of uncertainty in the conditional correlation estimates and are in fact complementary.

This paper builds upon a rich literature on multivariate volatility modeling (e.g. Bauwens et al., 2006; Silvennoinen and Teräsvirta, 2009). In particular, this paper is connected to extensions of the DCC model that help accommodate large cross-sections. This includes linear and non-linear shrinkage of the target correlation matrix (Hafner and Reznikova, 2012; Engle et al., 2019), CL estimation (Pakel et al., 2021), the use of intraday high and low prices (De Nard et al., 2022) and projection methods (Llorens-Terrazas and Brownlees, 2022). Our paper is most closely related to the (Block-)DECO model by Engle and Kelly (2012).

The outline of this paper is as follows. Section 2 develops the CLIP-DCC model, showing how it can be obtained from the combination of a DCC model with an appropriately scaled

(Block-)DECO-type model. A simulation study to assess the quality of the CL estimator for the CLIP-DCC model is carried out in Section 3. Section 4 presents a real-time application to daily stock data and Section 5 concludes. Finally, proofs and additional results are provided in the online Appendix.

2 Methodology

2.1 The DCC and DECO Models

Let $r_{i,t}$ denote the return on asset i at time t for $i = 1, \dots, N$ and $t = 1, \dots, T$, and let $\mathbf{r}_t := (r_{1,t}, \dots, r_{N,t})'$. For simplicity, we follow the standard practice to assume that the returns have mean zero. We focus on models for the conditional covariance matrix of \mathbf{r}_t , denoted as $\boldsymbol{\Sigma}_t := \mathbb{E}[\mathbf{r}_t \mathbf{r}_t' | \mathcal{I}_{t-1}]$, where \mathcal{I}_{t-1} is the information set at time $t - 1$. Specifically, we adopt the Dynamic Conditional Correlation (DCC) framework of Engle (2002) using the decomposition

$$\boldsymbol{\Sigma}_t = \mathbf{D}_t \mathbf{R}_t \mathbf{D}_t, \quad (1)$$

where \mathbf{D}_t is an $N \times N$ diagonal matrix with the conditional volatilities of the asset returns on the diagonal, and \mathbf{R}_t is the $N \times N$ conditional correlation matrix. Decomposition (1) is convenient as it facilitates (i) flexible covariance dynamics by separately modeling the conditional volatilities and correlations and (ii) sequential estimation of the volatility models and the correlation model, preventing the estimation of many parameters simultaneously.

In the DCC approach, the (squares of the) diagonal elements of \mathbf{D}_t are typically modelled by conventional GARCH-type specifications (Bollerslev, 1986). For example, in the empirical application in Section 4, we use the GJR-GARCH model (Glosten et al., 1993) such that the conditional variance $\sigma_{i,t}^2 := \mathbb{E}[r_{i,t}^2 | \mathcal{I}_{t-1}]$ is given by

$$\sigma_{i,t}^2 = w_i + (a_i + c_i \mathbf{1}[r_{i,t-1} < 0]) r_{i,t-1}^2 + b_i \sigma_{i,t-1}^2, \quad (2)$$

with $\mathbf{1}[\mathcal{A}]$ an indicator function for the event \mathcal{A} and the coefficients w_i , a_i , c_i and b_i satisfying $w_i > 0$, $a_i > 0$, $b_i \geq 0$, and $a_i + c_i/2 + b_i < 1$.

The dynamics of the conditional correlation matrix \mathbf{R}_t in the DCC model make use of

the $N \times 1$ vector of standardized (or devolatilized) returns

$$\mathbf{z}_t = \mathbf{D}_t^{-1} \mathbf{r}_t. \quad (3)$$

Using (1) we find that $\mathbb{E}[\mathbf{z}_t \mathbf{z}_t' | \mathcal{I}_{t-1}] = \mathbf{R}_t$, such that the outer product of the devolatilized returns $\mathbf{z}_t \mathbf{z}_t'$ provides an unbiased ex-post proxy of the true conditional correlation matrix. The baseline DCC specification of Engle (2002) is given by

$$\mathbf{Q}_t^{DCC} = (1 - \alpha - \beta) \mathbf{C} + \alpha \mathbf{z}_{t-1} \mathbf{z}_{t-1}' + \beta \mathbf{Q}_{t-1}^{DCC}, \quad (4)$$

$$\mathbf{R}_t^{DCC} = \text{diag}(\mathbf{Q}_t^{DCC})^{-1/2} \mathbf{Q}_t^{DCC} \text{diag}(\mathbf{Q}_t^{DCC})^{-1/2}, \quad (5)$$

where $\text{diag}(\mathbf{A})$ denotes the function that returns a diagonal matrix containing the diagonal elements of the $N \times N$ matrix \mathbf{A} , α and β are scalar parameters satisfying $\alpha > 0$, $\beta \geq 0$ and $\alpha + \beta < 1$, and \mathbf{C} is a (symmetric) positive definite matrix. Note that $\mathbb{E}(\mathbf{R}_t^{DCC}) \approx \mathbf{C}$, such that \mathbf{C} can be interpreted as the long-run correlation matrix, whereby the approximate nature stems from the re-scaling in (5) to go from \mathbf{Q}_t^{DCC} to \mathbf{R}_t^{DCC} . This re-scaling is necessary because, while \mathbf{Q}_t^{DCC} is guaranteed to be positive definite, its diagonal elements are (close to but) not necessarily equal to unity. For this reason, \mathbf{Q}_t^{DCC} is typically referred to as a *pseudo*-correlation matrix. Aielli (2013) proposes a correction for this scaling step, but Engle et al. (2019) among others note that the effects are not relevant in empirical applications; we therefore do not pursue this correction here. An interesting alternative for the re-scaling step is the projection method of Llorens-Terrazas and Brownlees (2022).

While the DCC innovation $\mathbf{z}_t \mathbf{z}_t'$ is an unbiased estimator of \mathbf{R}_t , it nevertheless provides only highly noisy updates of the individual conditional correlations. The Dynamic Equicorrelation (DECO) model by Engle and Kelly (2012) dramatically restricts the correlation matrix by assuming that all conditional correlations are equal. Practically, this restriction is implemented by applying a transformation to the conditional correlation matrix provided by the DCC-recursion given by (4) and (5) as follows

$$\rho_t = \frac{1}{N(N-1)} \iota_N' (\mathbf{R}_t^{DCC} - \mathbf{I}_N) \iota_N, \quad (6)$$

$$\mathbf{R}_t^{DECO} = \rho_t \mathbf{J}_N + (1 - \rho_t) \mathbf{I}_N, \quad (7)$$

where ι_N is an $N \times 1$ vector of ones, \mathbf{I}_N is the $N \times N$ identity matrix and \mathbf{J}_N is an $N \times N$ matrix of ones. The form of (6) and (7) shows that the dynamic equicorrelation matrix

\mathbf{R}_t^{DECO} is constructed using the (cross-sectional) average of the conditional correlations of \mathbf{R}_t^{DCC} , denoted by ρ_t . Because of the linearity of (6) and (7), this means that all the conditional correlations of the DECO model are approximately innovated by the average of the cross-products of $\mathbf{z}_t\mathbf{z}'_t$. The main practical allure of this model is that analytic expressions are available for the inverse and determinant of \mathbf{R}_t^{DECO} . This dramatically speeds up likelihood estimation of the parameters α and β , see Engle and Kelly (2012) for details.

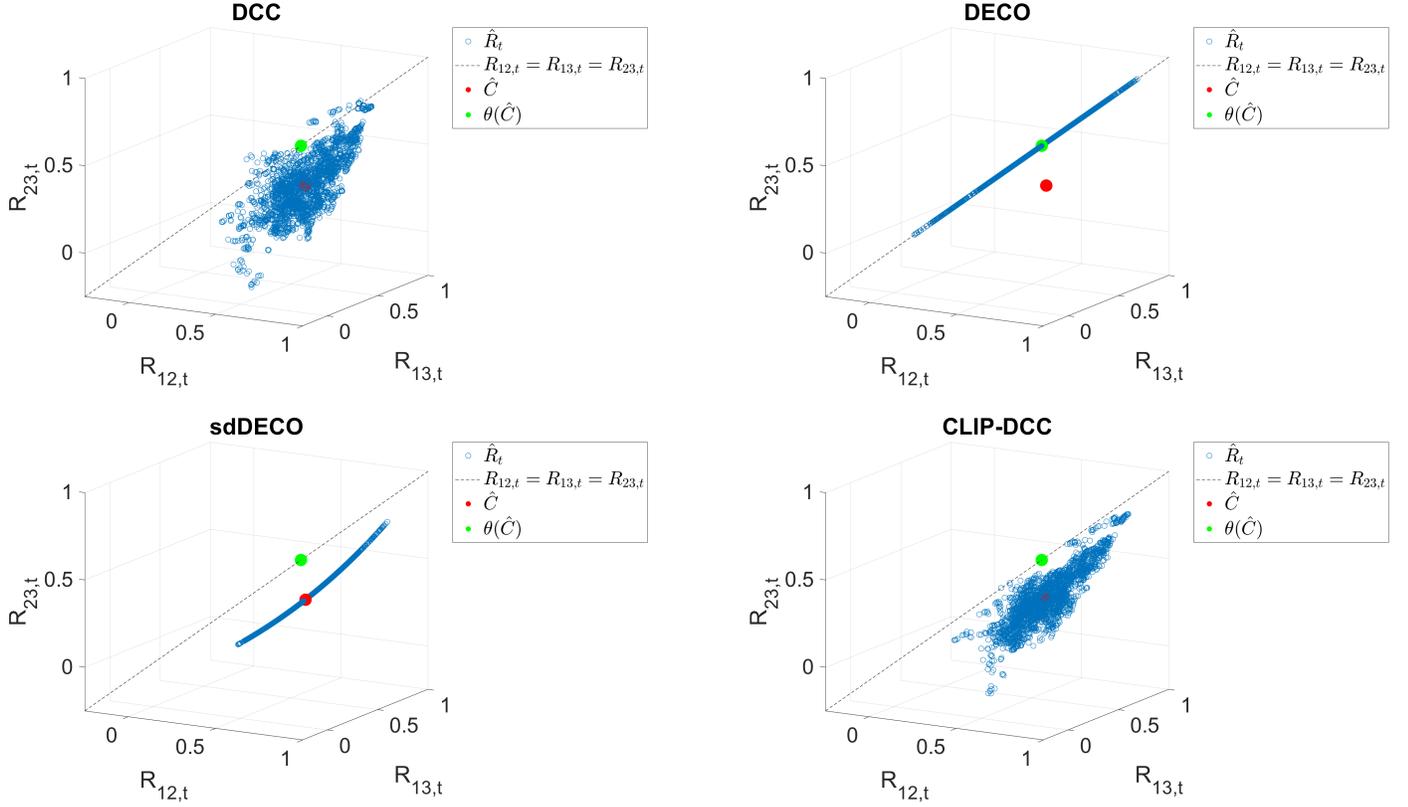
To visualize the differences and implications of the DCC and DECO models, we consider an empirical example and estimate both models on daily returns of the 30 industry portfolios constructed by Kenneth French, from January 2, 2000 until December 31, 2009¹. The top panels of Figure 1 depict the dynamic conditional correlations between the first three industries, which are Food, Beer and Smoke.

Figure 1 shows that almost all correlation estimates lie in the positive domain, reflecting positive co-movement of these industry portfolios. Comparing the top two panels, we make two important observations. First, the dynamic correlation estimates of the DECO model lie on the 3-dimensional ‘45 degree’ line due to the equicorrelation structure, while those of the DCC model form a fairly dispersed cloud. The DCC and DECO models are thus extremes on the bias-variance spectrum. By this we mean that, on the one hand the DCC model is innovated using $\mathbf{z}_t\mathbf{z}'_t$, which is an unbiased, yet very noisy, ex-post proxy of the true conditional correlation matrix \mathbf{R}_t . On the other hand, the DECO model equates all conditional correlations and is therefore approximately innovated by the average of the cross-products of $\mathbf{z}_t\mathbf{z}'_t$. This obviously reduces variance but potentially introduces bias in case the equicorrelation assumption is not appropriate. Therefore, it might be worthwhile exploring a setup that can fill this gap and provide a more ‘optimized’ level of commonality in the update-mechanism.

Second, we observe that the DECO model also implicitly warps the unconditional correlation matrix as a direct consequence of the structure imposed on the conditional correlations. That is, the DCC model is centered around an unrestricted long-run, estimated here using the sample covariance matrix of the devolatilized returns $\hat{\mathbf{C}} = T^{-1} \sum_{t=1}^T \mathbf{z}_t\mathbf{z}'_t$ with

¹Data is obtained from: https://mba.tuck.dartmouth.edu/pages/faculty/ken.french/data_library.html

Figure 1: Selection of dynamic correlation estimates of the DCC, DECO, sdDECO and CLIP-DCC models using thirty industry portfolios, January 2000 until December 2009.



Note: This figure contains the dynamic conditional correlation estimates of the Food (1), Beer (2) and Smoke (3) industry portfolios. The correlation models are estimated using Composite Likelihood (except the DECO model, which is estimated using the full likelihood) on the devolatilized returns obtained from GJR-GARCH(1,1) specifications and with target $\hat{\mathbf{C}} = T^{-1} \sum_{t=1}^T \mathbf{z}_t \mathbf{z}_t'$.

$[\hat{c}_{12} \hat{c}_{13} \hat{c}_{23}] = [0.621 \ 0.436 \ 0.344]$, which is the default estimator of \mathbf{C} in the literature. In contrast, the DECO model is centered around an equicorrelation matrix with estimated equicorrelation parameter $\hat{\rho} = 0.537$. To make this warping explicit, note that

$$\mathbb{E}[\mathbf{R}_t^{DECO}] = \mathbb{E}[\rho_t \mathbf{J}_N + (1 - \rho_t) \mathbf{I}_N] = \bar{\rho} \mathbf{J}_N + (1 - \bar{\rho}) \mathbf{I}_N, \quad (8)$$

$$\bar{\rho} = \mathbb{E}[\rho_t] = \frac{1}{N(N-1)} \iota_N' (\mathbb{E}[\mathbf{R}_t^{DCC}] - \mathbf{I}_N) \iota_N, \quad (9)$$

where $\bar{\rho}$ denotes the long-run average correlation. Due to the linearity of the transformation from \mathbf{R}_t^{DCC} to \mathbf{R}_t^{DECO} as given by (6) and (7), the long-run $\mathbb{E}[\mathbf{R}_t^{DECO}]$ is thus structured in the same way as the conditional correlation matrices \mathbf{R}_t^{DECO} .

Intuitively, however, the conditional and unconditional correlation matrix could likely benefit from different levels of structure. This stems from the differences in effective sample size used for their estimation. Namely, estimation of the long-run correlation matrix can draw from the entire sample. Conversely, the conditional correlations can only tap from a much smaller selection of recent observations. Within our model framework this is reflected by the fact that $\hat{\mathbf{C}}$ uses the entire sample of T observations, while the DCC recursion in (4) amounts to an exponentially weighted moving average. In particular, if the concentration ratio N/T is relatively small, then the estimator $\hat{\mathbf{C}}$ will be fairly accurate, demanding little additional structure. On the other hand, if N itself is not small or if the autocorrelation of the conditional correlations is low, then the conditional correlation matrix can likely benefit substantially from structure. Decoupling the structure imposed on the conditional and unconditional correlations would essentially allow us to separately deal with uncertainty of the long-run correlation matrix and the deviations therefrom.

The main contribution of this paper, namely the DCC model with Conditional Linear Information Pooling (CLIP-DCC), is the result of addressing these two concerns. Specifically, we first introduce a scaled DECO-type model to undo the warping of the long-run dynamics in Section 2.2 and, second, consider a mixture setup to provide a more appropriate level of cross-sectional structure for the conditional correlations in Section 2.3. Due to the linearity of the operations, we show that the CLIP-DCC model amounts to a DCC model with an additional ‘pooled’ innovation term, such that implementation is straightforward. Section 2.4 then shows how the CLIP-DCC model may be extended to incorporate relevant group structure information using a block structure.

2.2 The Scaled Direct DECO Model

In this section, we introduce a model similar to the DECO model, but with the structure imposed at the pseudo-correlation level \mathbf{Q}_t . Moreover, the model involves a recentering step to undo the implicit warping of the long-run correlation matrix. We argue that it is generally more convenient to apply structure in the pseudo-correlation space \mathbf{Q}_t than at the correlation level \mathbf{R}_t for two reasons. First, one can consider transformations that preserve positive definiteness but do not necessarily maintain the unit diagonal, without having to

standardize twice. The re-centering step is an example of such a transformation. Second, imposing structure at the pseudo-correlation level can come with large computational gains.

To facilitate the discussion we begin by reviewing the concept of a compound symmetric (CS) matrix, which is closely related to the equicorrelation matrix.

Definition 1 (Compound-symmetric matrix) *A square matrix \mathbf{S} of dimension $N \times N$ is a compound symmetric matrix if and only if it can be written as*

$$\mathbf{S} = o\mathbf{J}_N + (d - o)\mathbf{I}_N, \quad (10)$$

that is, all diagonal elements of \mathbf{S} are equal to $d \in \mathbb{R}$, while all off-diagonal elements are equal to $o \in \mathbb{R}$.

Proposition 1 (Positive-definiteness of CS matrices) *A CS matrix \mathbf{S} of dimension $N \times N$ with $N \geq 2$, diagonal element $d \in \mathbb{R}$ and off-diagonal element $o \in \mathbb{R}$ is positive definite if and only if $d > 0$ and $\frac{o}{d} \in (\frac{-1}{N-1}, 1)$. Additionally, if \mathbf{S} is positive definite then $\text{diag}(\mathbf{S})^{-1/2} \mathbf{S} \text{diag}(\mathbf{S})^{-1/2}$ is an equicorrelation matrix with equicorrelation $\frac{o}{d}$.*

Next, we introduce a transformation that can be used to construct a CS matrix from an arbitrary square matrix.

Definition 2 (Compound-symmetry transformation) *For any $\mathbf{A} \in \mathbb{R}^{N \times N}$ we define the mapping $\theta(\mathbf{A}) : \mathbb{R}^{N \times N} \rightarrow \mathbb{R}^{N \times N}$ as*

$$\theta(\mathbf{A}) := \theta^O(\mathbf{A})\mathbf{J}_N + [\theta^D(\mathbf{A}) - \theta^O(\mathbf{A})]\mathbf{I}_N, \quad (11)$$

where the scalars $\theta^D(\mathbf{A})$ and $\theta^O(\mathbf{A})$ are the diagonal and off-diagonal averages of \mathbf{A} , that is for $N \geq 2$,

$$\theta^D(\mathbf{A}) := \frac{1}{N} \sum_{i=1}^N a_{ii}, \quad (12)$$

$$\theta^O(\mathbf{A}) := \frac{1}{N(N-1)} \sum_{i=1}^N \sum_{j=1, j \neq i}^N a_{ij}, \quad (13)$$

where a_{ij} is the ij -th element of \mathbf{A} . If $N = 1$, we set $\theta(\mathbf{A}) = \theta^D(\mathbf{A}) = \theta^O(\mathbf{A}) = \mathbf{A}$.

Note that $\mathbf{R}_t^{DECO} = \theta(\mathbf{R}_t^{DCC})$, such that $\theta(\cdot)$ can be interpreted as an intuitive generalization of the DECO transformation in (6) and (7) that accommodates covariance matrices, instead of only correlation matrices. The mapping $\theta(\cdot)$ has several convenient properties, as summarized in the following proposition.

Proposition 2 (Linearity and positivity of the CS transformation) $\theta(\cdot)$ is a linear mapping that preserves positive (semi-)definiteness, that is for $\mathbf{A} \in \mathbb{R}^{N \times N}$ we have that

1. $\theta(\mathbf{A} + \mathbf{B}) = \theta(\mathbf{A}) + \theta(\mathbf{B})$, $\forall \mathbf{B} \in \mathbb{R}^{N \times N}$,
2. $\theta(k\mathbf{A}) = k\theta(\mathbf{A})$, $\forall k \in \mathbb{R}$,
3. If \mathbf{A} is positive (semi-)definite, then $\theta(\mathbf{A})$ is positive (semi-)definite.

Applying the compound symmetry mapping $\theta(\cdot)$ to \mathbf{Q}_t^{DCC} as given in (4), we can construct a ‘direct’ DECO (dDECO) model. That is, we define $\mathbf{Q}_t^{dDECO} := \theta(\mathbf{Q}_t^{DCC})$ and using Proposition 2 we obtain the recursion

$$\mathbf{Q}_t^{dDECO} = (1 - \alpha - \beta)\theta(\mathbf{C}) + \alpha\theta(\mathbf{z}_{t-1}\mathbf{z}'_{t-1}) + \beta\mathbf{Q}_{t-1}^{dDECO}, \quad (14)$$

where \mathbf{Q}_t^{dDECO} is positive definite if \mathbf{Q}_t^{DCC} is positive definite. Standardization of \mathbf{Q}_t^{dDECO} by dividing by its diagonal elements will yield a valid equicorrelation matrix (see again Proposition 1). This model is therefore highly similar to the DECO model as it is based on the same equal pairwise correlations assumption. The difference between the two models is a slightly different implementation of this constraint, namely, the dDECO model switches the order of standardization and pooling compared to the DECO model.

First imposing structure and then standardizing, as in the dDECO model, greatly reduces computational demands. This is because \mathbf{Q}_t^{dDECO} has only two unique elements, namely its diagonal element $d_t := \theta^D(\mathbf{Q}_t^{DCC})$ and its off-diagonal element $o_t := \theta^O(\mathbf{Q}_t^{DCC})$, whereby $\theta^D(\cdot)$ and $\theta^O(\cdot)$ are linear mappings. Instead of tracking the $N \times N$ matrix \mathbf{Q}_t^{DCC} in full, as needed for the DECO model, we thus only need to consider the two scalar processes for d_t and o_t given by

$$d_t = (1 - \alpha - \beta)\theta^D(\mathbf{C}) + \alpha\theta^D(\mathbf{z}_{t-1}\mathbf{z}'_{t-1}) + \beta d_{t-1}, \quad (15)$$

$$o_t = (1 - \alpha - \beta)\theta^O(\mathbf{C}) + \alpha\theta^O(\mathbf{z}_{t-1}\mathbf{z}'_{t-1}) + \beta o_{t-1}, \quad (16)$$

which are driven by the diagonal and off-diagonal mean of the DCC innovation $\mathbf{z}_{t-1}\mathbf{z}'_{t-1}$, respectively. Afterwards, \mathbf{Q}_t^{dDECO} can be obtained from d_t and o_t as in (10).

Due the linearity of $\theta(\cdot)$, we have that both the DECO model and the dDECO model have (approximately) displaced the long-run correlation matrix from \mathbf{C} to $\theta(\mathbf{C})$, an equicorrelation matrix if \mathbf{C} has a unit diagonal. We therefore propose a multiplicative re-centering step to undo the warping of the long-run correlation matrix. Specifically, we construct the scaled dDECO (sdDECO) model as follows

$$\mathbf{Q}_t^{sdDECO} := [\mathbf{C}^{1/2}\theta(\mathbf{C})^{-1/2}]\mathbf{Q}_t^{dDECO}[\theta(\mathbf{C})^{-1/2}\mathbf{C}^{1/2}], \quad (17)$$

where $\mathbf{C}^{1/2}$ and $\theta(\mathbf{C})^{-1/2}$ denote the symmetric square-root of \mathbf{C} and $\theta(\mathbf{C})^{-1}$ obtained using the eigenvalue decomposition. Symmetry and positive definiteness of \mathbf{Q}_t^{sdDECO} is inherited from the positive definiteness of \mathbf{Q}_t^{dDECO} and immediately apparent.

Inserting (14) in (17), the update recursion for \mathbf{Q}_t^{sdDECO} can be written directly as

$$\mathbf{Q}_t^{sdDECO} = (1 - \alpha - \beta)\mathbf{C} + \alpha\mathbf{Z}_{t-1} + \beta\mathbf{Q}_{t-1}^{sdDECO}, \quad (18)$$

where \mathbf{Z}_{t-1} is a positive semi-definite innovation term (see again Proposition 2) given by

$$\mathbf{Z}_{t-1} = [\mathbf{C}^{1/2}\theta(\mathbf{C})^{-1/2}]\theta(\mathbf{z}_{t-1}\mathbf{z}'_{t-1})[\theta(\mathbf{C})^{-1/2}\mathbf{C}^{1/2}], \quad (19)$$

where $\mathbb{E}(\theta(\mathbf{z}_{t-1}\mathbf{z}'_{t-1})) = \theta(\mathbb{E}(\mathbf{z}_{t-1}\mathbf{z}'_{t-1})) \approx \theta(\mathbf{C})$, such that $\mathbb{E}(\mathbf{Z}_{t-1}) \approx \mathbf{C}$. Here the approximate nature stems from minor differences arising from standardization. From there it is straightforward to see that the sdDECO model, by construction, has again center of movement \mathbf{C} . Effectively, the conditional correlations of the sdDECO model are therefore updated using the average of the cross-products of $\mathbf{z}_{t-1}\mathbf{z}'_{t-1}$, similar to the (d)DECO model, but with an additional scaling to preserve the long-run correlation matrix. Because the sdDECO model leaves the unconditional correlations unaltered, the pooling can be interpreted to be ‘conditional’.

2.3 The CLIP-DCC Model

Although the sdDECO model preserves the long-run correlation target, it still imposes a large amount of structure on the conditional correlations as all time-variation is generated

by two scalar processes, see again (15)-(17). Therefore, to overcome our other concern and provide a more nuanced amount of commonality in the individual correlation updates, we now present the DCC model with Conditional Linear Information Pooling (CLIP-DCC). Specifically, for the pseudo-correlation process $\mathbf{Q}_t^{CLIP-DCC}$ we consider a convex combination of \mathbf{Q}_t^{DCC} and \mathbf{Q}_t^{sdDECO} , that is

$$\mathbf{Q}_t^{CLIP-DCC} = (1 - \omega)\mathbf{Q}_t^{DCC} + \omega\mathbf{Q}_t^{sdDECO}, \quad (20)$$

where $\omega \in [0, 1]$ is the mixture weight. Using (4) and (18) and assuming α and β are the same for both processes, we may directly write the update recursion for $\mathbf{Q}_t^{CLIP-DCC}$ as

$$\mathbf{Q}_t^{CLIP-DCC} = (1 - \alpha - \beta)\mathbf{C} + \alpha[(1 - \omega)\mathbf{z}_{t-1}\mathbf{z}'_{t-1} + \omega\mathbf{Z}_{t-1}] + \beta\mathbf{Q}_{t-1}^{CLIP-DCC}, \quad (21)$$

where \mathbf{Z}_{t-1} is given by (19). It is straightforward to show that $\mathbf{Q}_t^{CLIP-DCC}$ is positive definite if \mathbf{Q}_t^{DCC} is positive definite and has center of movement \mathbf{C} . Furthermore, reparameterization with $\alpha_1 := \alpha(1 - \omega)$ and $\alpha_2 := \alpha\omega$ yields,

$$\mathbf{Q}_t^{CLIP-DCC} = (1 - \alpha_1 - \alpha_2 - \beta)\mathbf{C} + \alpha_1\mathbf{z}_{t-1}\mathbf{z}'_{t-1} + \alpha_2\mathbf{Z}_{t-1} + \beta\mathbf{Q}_{t-1}^{CLIP-DCC}, \quad (22)$$

which reveals that, effectively, we are simply adding a pooling term \mathbf{Z}_{t-1} to the DCC recursion as an additional explanatory variable. As noted by Engle and Kelly (2012), the correlations of the DCC model evolve essentially independently while those of the (d)DECO model co-move perfectly. In contrast, the CLIP-DCC model allows for a more nuanced level of cross-sectional dependence, with magnitude determined by the mixture weight ω .

Although intuitively one may think of the sdDECO component in (20) as a shrinkage target for the conditional correlation matrix, our approach differs from traditional shrinkage methods. Namely, we determine the optimal pooling intensity endogenously by estimating ω simultaneously with the DCC model parameters (based on the likelihood, as discussed in detail in Section 2.5), instead of relying on asymptotic theory or cross-validation. Separation of the structure imposed on the conditional and unconditional correlation matrices is crucial for this estimation process. This is because the sample correlation matrix $\hat{\mathbf{C}}$ used for targeting already provides the best estimate of the long-run \mathbf{C} in-sample, such that its optimal pooling intensity using maximum likelihood is near zero. As a result, a setup where a dDECO component is used instead of an sdDECO component in (20), yields very

small estimates of the mixture weight ω , effectively reducing the model to the base DCC model. Structuring the long-run correlation matrix thus demands a different approach as will be outlined in Section 2.5.

To illustrate the appeal of the CLIP-DCC framework, we again consider our empirical example in Figure 1. The bottom left panel shows that the sdDECO model yields a line-type movement pattern similar to the DECO model, but with center of movement $\hat{\mathbf{C}}$ as in the DCC model. The slight curvature of the sdDECO line is the result of the standardization step from \mathbf{Q}_t^{sdDECO} to \mathbf{R}_t^{sdDECO} . In addition, we observe in the bottom right panel that the variability of the CLIP-DCC model is between that of the DCC and sdDECO models. The estimated mixture weight $\hat{\omega}$ is equal to 0.514, such that the CLIP-DCC update relies roughly equally on the unstructured innovation $\mathbf{z}_{t-1}\mathbf{z}'_{t-1}$ and the pooled innovation \mathbf{Z}_{t-1} , see again (21).

Finally, it is interesting to remark that the estimated autoregressive coefficient $\hat{\beta}$ is lower for the CLIP-DCC model than for the DCC model (0.944 and 0.965, respectively), while the estimated innovation coefficient $\hat{\alpha}$ is higher for the CLIP-DCC model compared to the DCC model (0.056 and 0.029, respectively). For both models the sum $\hat{\alpha} + \hat{\beta}$ is close to but strictly smaller than 1. Bollerslev et al. (2020), among others, find a similar pattern when replacing $\mathbf{z}_t\mathbf{z}'_t$ in (4) with more accurate realized estimators based on intraday data. Because more aggressive updating leads to added variability, this explains why the visual differences between the CLIP-DCC and DCC models are substantial, but perhaps slightly smaller than expected from the large mixture weight $\hat{\omega}$. More importantly, this means that the CLIP-DCC model is able to be more responsive to new data by countering the added uncertainty with cross-sectional structure.

2.4 The Block-CLIP-DCC Model

In the presence of a clear group structure of the assets, one would like to use this information in the estimation of the conditional correlation matrix. While the CLIP-DCC model does not need any such information, we may extend our framework to exploit it when available. That is, the mixture setup of the CLIP-DCC model and the block structure of the Block-DECO model of Engle and Kelly (2012) can be seen as different solutions to the same

cross-sectional structuring problem, which are not mutually exclusive.

Given a K -group partition n_1, n_2, \dots, n_K of the cross-section of N assets such that $n_j \in \mathbb{N}^+, j = 1, 2, \dots, K$ and $\sum_{j=1}^K n_j = N$, we assume without loss of generality that the data is ‘sorted’ such that the assets $i = 1, \dots, n_1$ are in the first group, the assets $i = n_1 + 1, \dots, n_1 + n_2$ are in the second group and so forth. For notational convenience, we summarize this K -group information in the $K \times 1$ vector $G := [n_1, n_2, \dots, n_K]'$. We now introduce a mapping $\theta^{BL}(\cdot, G)$ that can be used to turn a square matrix into a ‘ K -block compound symmetric matrix’ with group structure G .

Definition 3 (Block-CS transformation) For $\mathbf{A} \in \mathbb{R}^{N \times N}$ with $N \geq 2$, we define

$$\theta^{BL}(\mathbf{A}, G) = \begin{bmatrix} \theta(\mathbf{A}_{11}^*) & \tau(\mathbf{A}_{12}^*) & \cdots & \tau(\mathbf{A}_{1K}^*) \\ \tau(\mathbf{A}_{21}^*) & \theta(\mathbf{A}_{22}^*) & \cdots & \tau(\mathbf{A}_{2K}^*) \\ \vdots & \vdots & \ddots & \vdots \\ \tau(\mathbf{A}_{K1}^*) & \tau(\mathbf{A}_{K2}^*) & \cdots & \theta(\mathbf{A}_{KK}^*) \end{bmatrix}, \quad (23)$$

$$\tau(\mathbf{V}) = \left(\frac{1}{m_1 m_2} \iota_{m_1}' \mathbf{V} \iota_{m_2} \right) \mathbf{J}_{m_1 \times m_2}, \quad \forall \mathbf{V} \in \mathbb{R}^{m_1 \times m_2}, \quad (24)$$

where \mathbf{A}_{ij}^* for $i, j = 1, 2, \dots, K$ is the ij -th $n_i \times n_j$ block of \mathbf{A} using the block partition $G := [n_1, n_2, \dots, n_K]'$, $n_1, n_2, \dots, n_K \in \mathbb{N}^+$ and $\sum_{j=1}^K n_j = N$.

In Definition 3, we observe that the diagonal blocks of $\theta^{BL}(\mathbf{A}, G)$ are obtained by applying the compound symmetry transformation $\theta(\cdot)$ to the diagonal blocks of \mathbf{A} . For the off-diagonal blocks, we use the function $\tau(\cdot)$, which returns a constant matrix with the mean of the input matrix everywhere. Proposition 3 summarizes some useful properties of $\theta^{BL}(\cdot, G)$, matching the properties of $\theta(\cdot)$ listed in Proposition 2.

Proposition 3 (Linearity and positivity of the block-CS transformation) $\theta^{BL}(\cdot, G)$, with $G := [n_1, n_2, \dots, n_K]'$, $n_1, n_2, \dots, n_K \in \mathbb{N}^+$ and $\sum_{j=1}^K n_j = N$, is a linear mapping that preserves positive (semi-)definiteness, that is, for $\mathbf{A} \in \mathbb{R}^{N \times N}$ with $N \geq 2$ we have that

1. $\theta^{BL}(\mathbf{A} + \mathbf{B}, G) = \theta^{BL}(\mathbf{A}, G) + \theta^{BL}(\mathbf{B}, G), \quad \forall \mathbf{B} \in \mathbb{R}^{N \times N},$
2. $\theta^{BL}(k\mathbf{A}, G) = k\theta^{BL}(\mathbf{A}, G), \quad \forall k \in \mathbb{R},$
3. If \mathbf{A} is positive (semi-)definite, then $\theta^{BL}(\mathbf{A}, G)$ is positive (semi-)definite.

Using Proposition 3, we may repeat the steps in the previous sections and obtain Block-dDECO, Block-sdDECO and Block-CLIP-DCC models by replacing $\theta(\cdot)$ with $\theta^{BL}(\cdot, G)$.

In practice, we may be unsure of the ‘best’ group structure and have multiple candidate structures. Therefore, we may also allow for multiple (distinct) block structures simultaneously. That is, we can consider a mixture setup of the DCC model with L Block-sdDECO models, each with a distinct group structure G_l , $l = 1, 2, \dots, L$. This then yields L additional explanatory terms in the pseudo-correlation update recursion of the form

$$\mathbf{B}_{l,t-1} = [\mathbf{C}^{1/2}\theta^{BL}(\mathbf{C}, G_l)^{-1/2}]\theta^{BL}(\mathbf{z}_{t-1}\mathbf{z}'_{t-1}, G_l)[\theta^{BL}(\mathbf{C}, G_l)^{-1/2}\mathbf{C}^{1/2}], \quad (25)$$

where $\mathbf{B}_{l,t-1}$ is a positive semi-definite (by Proposition 3) explanatory term with group structure G_l , suitably scaled to preserve long-run dynamics (i.e. we have $\mathbb{E}(\mathbf{B}_{l,t-1}) \approx \mathbf{C}$).

2.5 Parameter Estimation and Target Shrinkage

We can estimate the parameters in the CLIP-DCC model using maximum likelihood in a three-step procedure, following Engle (2002). The same procedure can be applied to all other models considered in Section 2, including sdDECO and Block-CLIP-DCC. In the first step, we estimate univariate GJR-GARCH models for each asset.

Second, we estimate the intercept matrix \mathbf{C} , see again (4), with targeting by using the sample covariance of the devolatilized returns. This prevents the likelihood estimation of $\mathcal{O}(N^2)$ parameters. It is well established that the quality of the sample correlation matrix degrades as the concentration ratio N/T grows, see e.g. Ledoit and Wolf (2004). We may therefore also consider alternative targeting procedures. For example, in our empirical application, we consider the non-linear shrinkage (NLS) estimator by Ledoit and Wolf (2020)², which shrinks the sample correlation matrix at the eigenvalue level and has theoretical as well as empirical advantages compared to linear shrinkage estimators (Engle et al., 2019).

Third, we use Gaussian quasi-maximum likelihood (QML) to estimate the scalar parameters α , β and ω . In particular, following Pakel et al. (2021), we employ a Composite

²We make use of the code made available by Michael Wolf at https://www.econ.uzh.ch/dam/jcr:11d24ab0-7ec2-4b3f-8ef4-7affaa727d25/analytical_shrinkag.m.zip

Likelihood (CL) approach which approximates the full log likelihood using an average of bivariate log likelihoods constructed from pairs of asset returns. Specifically, we use CL estimation based on contiguous pairs, that is we pair asset 1 and 2, 2 and 3 and so forth. This results in $N - 1$ bivariate log likelihoods whose average will be maximized. For the DCC model, Pakel et al. (2021) find this yields highly similar parameter estimates as compared to CL based on all $N(N - 1)/2$ pairs, which is much more computationally intensive. Because CL avoids multiplications, inverses and determinants with $N \times N$ matrices it is computationally much cheaper than traditional full maximum likelihood (FML), especially for large N . Moreover, Pakel et al. (2021) find that the estimates of α and β of the DCC model are meaningfully biased towards 0 for realistic sample sizes when estimated using FML, but not for CL. This bias for FML is found to rapidly increase with the dimension N .

3 Monte Carlo Simulation

To assess the quality of the CL estimator for the CLIP-DCC model, particularly in view of the mixture weight ω , we conduct a Monte Carlo study using the setup of Pakel et al. (2021). We simulate data from a CLIP-DCC DGP assuming a conditional multivariate Gaussian distribution for $N \in \{10, 30, 100\}$, $T = 2000$, $\alpha = 0.05$, $\beta = 0.93$ and $\omega \in \{0, 0.25, 0.5, 0.75, 1\}$. This selection of mixture weights therefore includes the DCC model ($\omega = 0$) and the sdDECO model ($\omega = 1$). For simplicity, we set all $\sigma_{i,t} = 1$ and do not involve estimation of the univariate GARCH models. For the intercept matrix \mathbf{C} , we use $\mathbf{C}_{i,j} = \pi_i \pi_j$ for $i \neq j$ and $\mathbf{C}_{i,i} = 1$ for $i, j = 1, \dots, N$, where the π_i are drawn from a truncated normal distribution with mean 0.5 and standard deviation 0.1 and truncation interval $[0.1, 0.9]$.

Table 1 contains the Monte Carlo means and standard deviations of the parameter estimates of α , β and ω , obtained by estimating the CLIP-DCC model on the simulated data using CL estimation based on 500 replications. We observe for all considered settings and for all three parameters that the mean estimate is very close to the true parameter value. In particular, even for $\omega = 0$ and $\omega = 1$, when the CLIP-DCC model collapses to the DCC model and the sdDECO model, respectively, the CL approach performs satisfactory. Additionally, we find that the Monte Carlo standard deviations decrease as N increases.

Table 1: Monte Carlo means and standard deviations of the parameters of the CLIP-DCC model estimated using CL.

		$\omega = 0$	$\omega = 0.25$	$\omega = 0.5$	$\omega = 0.75$	$\omega = 1$
$N = 10$	α	0.052 (0.004)	0.050 (0.006)	0.051 (0.008)	0.050 (0.010)	0.051 (0.011)
	β	0.928 (0.005)	0.927 (0.008)	0.926 (0.012)	0.925 (0.017)	0.925 (0.022)
	ω	0.037 (0.051)	0.240 (0.097)	0.502 (0.087)	0.761 (0.068)	0.986 (0.026)
$N = 30$	α	0.051 (0.003)	0.050 (0.005)	0.050 (0.006)	0.051 (0.008)	0.050 (0.009)
	β	0.928 (0.003)	0.927 (0.005)	0.927 (0.007)	0.926 (0.012)	0.927 (0.016)
	ω	0.028 (0.040)	0.242 (0.081)	0.500 (0.064)	0.763 (0.040)	0.994 (0.012)
$N = 100$	α	0.051 (0.002)	0.050 (0.004)	0.050 (0.005)	0.051 (0.006)	0.050 (0.008)
	β	0.928 (0.002)	0.928 (0.003)	0.926 (0.005)	0.925 (0.010)	0.927 (0.014)
	ω	0.026 (0.034)	0.243 (0.069)	0.503 (0.047)	0.763 (0.029)	0.999 (0.004)

Note: This table contains the average parameter estimates (over the replications) of the CLIP-DCC model estimated using CL. The standard deviations of the parameter estimates are displayed in parentheses below the averages. The data is simulated from a CLIP-DCC model with $\alpha = 0.05$, $\beta = 0.93$, $\omega \in \{0, 0.25, 0.5, 0.75, 1\}$, $N \in \{10, 30, 100\}$ and $T = 2000$ using 500 replications.

This suggests that the potential efficiency loss of CL decreases as N becomes large, in line with the results of Pakel et al. (2021). Furthermore, we find in unreported additional simulations for different values of T that bias tends to decrease as T grows. Finally, we remark that the amount of skewness and excess kurtosis of the parameter estimates is mostly mild except when the true mixture weight ω is at either of the bounds ($\omega = 0$ or $\omega = 1$), which naturally compresses the distribution of $\hat{\omega}$. These findings reassure us that the CLIP-DCC model parameters may be effectively estimated using the CL approach.

4 Empirical Application

4.1 Data and Empirical Set-up

We examine the empirical usefulness of the CLIP-DCC model and its interaction with shrinkage of the target correlation matrix in an application to portfolio construction for US equity. We collect daily stock prices from the Center for Research in Security Prices (CRSP) database for the period from February 6, 1981 until December 31, 2020 and consider a wide range of portfolio sizes with $N \in \{10, 30, 50, 100, 300, 500\}$. The investment universe is (slowly) varying over time, in the sense that at each estimation date (see below) we select the N stocks with the highest market capitalization, similar to Engle et al. (2019). Note that for $N = 30, 100,$ and 500 this roughly picks the constituents of the Dow Jones Industrial Average (DJIA), the S&P 100 and the S&P 500 indexes, respectively.

The parameters in the CLIP-DCC and alternative correlation models are estimated using a rolling window of 2500 observations, using devolatilized returns obtained from univariate GJR-GARCH models as given in (2). For computational purposes the models are re-estimated every 21 trading days using CL based on contiguous pairs with alphabetically sorted stocks according to their ticker symbols, similar to Pakel et al. (2021). At each estimation date, we use the N largest stocks with the additional requirement that closing prices are available for both the complete estimation window as well as 21 days ahead for evaluation purposes. This yields exactly 360 estimation dates and 7560 evaluation days.

4.2 Evaluation

In the absence of high quality ex-post measures, we follow the common practice of indirect evaluation of the conditional correlation matrices by using them to construct portfolios, see e.g. Engle and Kelly (2012) and Engle et al. (2019). Specifically, we consider the global minimum variance portfolio (GMVP). This portfolio is popular due to its simplicity and its independence of the expected return, which empirically is often poorly estimated (Michaud, 1989). The analytical solution for the GMVP weight vector u_t is given as

$$u_t = \frac{\Sigma_t^{-1} l_N}{l_N' \Sigma_t^{-1} l_N}, \quad (26)$$

where Σ_t denotes the covariance matrix at time t . For our application, we consider a daily re-balancing approach whereby we make new GMVPs every day using one-step ahead forecasts of the conditional covariance matrix. These forecasts can be obtained by combining correlation matrix forecasts with the diagonal matrix of volatility forecasts obtained from the GJR-GARCH models, see again (1). Filling in the covariance matrix forecasts in (26) then yields feasible estimators of the GMVP weights.

In terms of evaluation, we consider the (annualized) average (AV) and the (annualized) standard deviation (SD) of the out-of-sample daily portfolio log returns as well as the corresponding information ratio $IR := AV/SD$. Naturally, since the objective of the GMVP is to minimize variance, we are mostly interested in the out-of-sample standard deviation. To assess whether the differences in SDs resulting from different correlation models are significant, we employ the test by Ledoit and Wolf (2011) based on heteroskedasticity-and-autocorrelation corrected standard errors. For robustness, we also consider two mean-variance (MV) portfolios and the quasi-likelihood (QLIKE) loss, see Patton and Sheppard (2009), in Appendix B. Qualitatively, our findings there are similar to the main results presented below.

4.3 Parameter Estimates

Table 2 contains average parameter estimates and average standard errors across the estimation dates of the different correlation models for the investment universe $N = 100$. To quantify time-variation, we also present the min-max ranges of the parameter estimates. Findings for the other portfolio sizes are highly similar and not displayed for brevity.

We observe that the average parameter estimates $\hat{\alpha}$ and $\hat{\beta}$ are fairly standard. That is, we find small values of $\hat{\alpha}$, large values of $\hat{\beta}$, and a sum close to 1. Comparing the parameter estimates of the DCC model to the estimates of the pooled models, we note that the latter admit a larger $\hat{\alpha}$ and a smaller $\hat{\beta}$. This reveals that the pooled models require less smoothing of the conditional correlations over time as a consequence of the imposed cross-sectional structure. In addition, for the CLIP-DCC model, we note that $\hat{\alpha}_1 = \hat{\alpha}(1 - \hat{w})$, the parameter for the DCC innovation term $\mathbf{z}_{t-1}\mathbf{z}'_{t-1}$, see (22), is comparable to the parameter estimate $\hat{\alpha}$ of the DCC model. Therefore, despite the addition of the pooled innovation

Table 2: Average parameter estimates for the different DCC models for $N = 100$, December 1990 until December 2020.

	DCC	dDECO	sdDECO	CLIP-DCC
$\hat{\alpha}$	0.015 [0.006, 0.050] (0.006)	0.039 [0.007, 0.112] (0.017)	0.043 [0.008, 0.138] (0.018)	0.040 [0.009, 0.136] (0.017)
$\hat{\beta}$	0.969 [0.858, 0.991] (0.014)	0.945 [0.794, 0.991] (0.025)	0.944 [0.795, 0.991] (0.025)	0.948 [0.800, 0.989] (0.022)
$\hat{\omega}$				0.653 [0.474, 0.767] (0.186)
$\hat{\alpha}_1$				0.012 [0.004, 0.035] (0.006)
$\hat{\alpha}_2$				0.028 [0.005, 0.102] (0.015)

Note: This table contains the average parameter estimates of the DCC, dDECO, sdDECO and CLIP-DCC models across the different estimation windows for $N = 100$. In addition, the minimum and maximum parameter estimates and the average estimated standard errors are displayed in brackets behind and in parentheses below the average parameter estimates, respectively. We estimate the models every 21-days using a moving window of length $T = 2500$ for a total of 360 estimation moments.

term \mathbf{Z}_{t-1} , the contribution of the DCC innovation term is largely maintained in the CLIP-DCC model, while the autoregressive parameter estimate $\hat{\beta}$ is reduced somewhat. This highlights the simultaneous bias-variance trade-off in the time dimension and the cross-sectional dimension as a result of the joint estimation of the smoothing parameters α and β and the pooling parameter ω .

For the CLIP-DCC model, we find that the average estimate of the mixture weight $\hat{\omega}$ is equal to 0.653, indicating that the sdDECO component is found to dominate the DCC component, see (20). In line with this, we have that $\hat{\alpha}_2 = \hat{\alpha}\hat{\omega}$ is more than twice as large as $\hat{\alpha}_1 = \hat{\alpha}(1 - \hat{\omega})$ on average, indicating that the pooled innovation term \mathbf{Z}_{t-1} is found to be more informative than the unstructured DCC innovation $\mathbf{z}_{t-1}\mathbf{z}'_{t-1}$, see again (22). All these findings are highly comparable with those of the empirical example in the methodology section using industry portfolios.

Finally, the min-max ranges indicate that the parameter estimates are subject to some time-variation. Although this is not necessarily insightful due to the changing investment universe, we do note a gradual increase of the mixture weight $\hat{\omega}$ over time combined with a

sharper decrease in the persistence $\hat{\alpha} + \hat{\beta}$ after the 2008 Great Financial Crisis. The former observation suggests a mild increase in correlation co-movement over time.

4.4 Portfolio Performance

Table 3 presents summary statistics of the daily out-of-sample log returns of the GMVPs constructed using the different models as well as the $1/N$ portfolio for comparison. First and foremost, we find that the CLIP-DCC model has the lowest out-of-sample SD for all considered portfolio sizes, outperforming all other correlation models and the $1/N$ portfolio. The SDs of the GMVPs decline as the portfolio dimension N grows, in line with the increasing possibilities for diversification. Moreover, we observe that the relative improvements of the CLIP-DCC model compared to the DCC model increase with N as well. That is, for $N = 10$ the improvements are relatively minor, with an SD of 15.392 for the CLIP-DCC model compared to 15.473 for the DCC model. By contrast, the improvements for $N = 500$ are much larger with an SD of 6.375 versus 6.621 for the CLIP-DCC and the DCC models, respectively. These reductions in SDs of the CLIP-DCC model compared to the DCC model are found to be highly significant for all but the smallest portfolio size $N = 10$, which is only significant at the ten percent level. This suggests that allowing for commonality in the dynamics of the conditional correlations becomes more important as the cross-sectional dimension increases.

Second, we observe that the dDECO model performs poorly compared to the other models in all metrics, while the sdDECO model delivers comparable performance to the DCC model. This suggests that imposing an equicorrelation structure on the long-run correlation matrix indeed incurs a too high bias for the reduction in variance obtained. It also indicates that the variance reduction by the structure of the sdDECO model roughly cancels against the bias incurred relative to the DCC model. The out-performance of the CLIP-DCC model is the result of allowing for an intermediate solution, providing a more optimized level of structure.

To further compare the DCC and CLIP-DCC models, we investigate the time course of their difference in performance. Figure 2 presents the difference in cumulative out-of-sample GMVP variance between the two models. Specifically, we subtract the cumulative

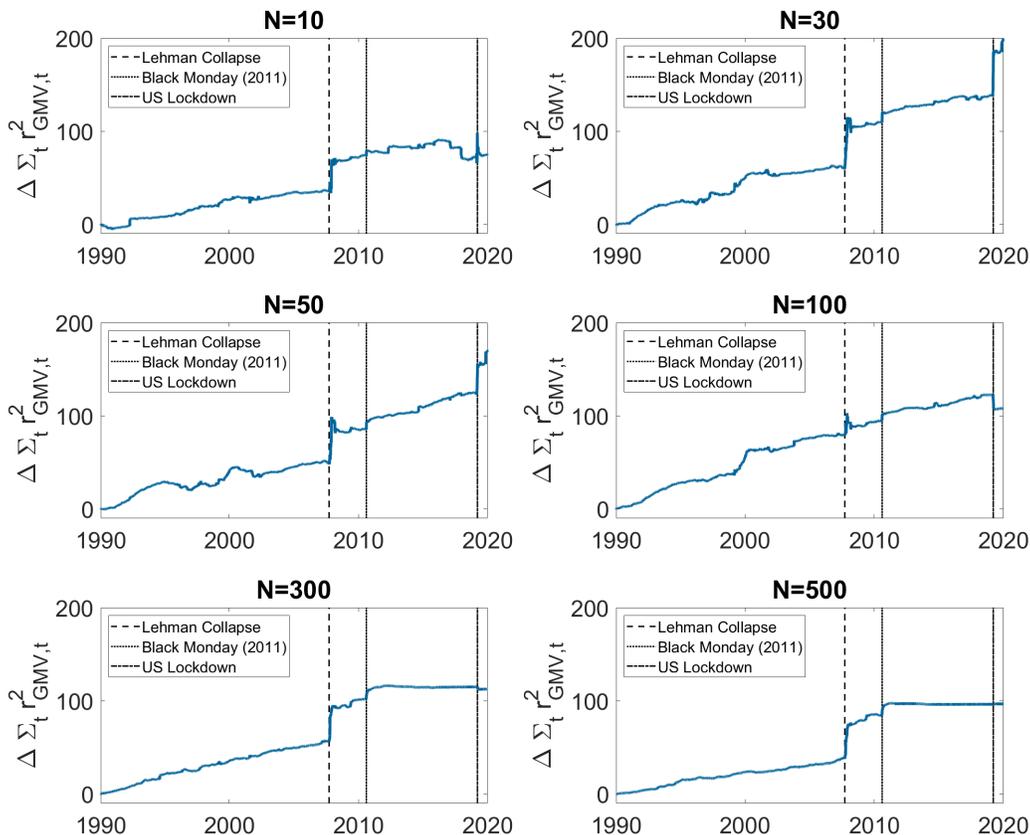
Table 3: Daily out-of-sample GMVP performance constructed using different DCC models for $N \in \{10, 30, 50, 100, 300, 500\}$, December 1990 until December 2020.

		DCC	dDECO	sdDECO	CLIP-DCC	1/N
$N = 10$	AV	4.856	3.758	5.778	5.272	7.283
	SD	15.473	15.745	15.480	15.392*	19.167
	IR	0.314	0.239	0.373	0.343	0.380
$N = 30$	AV	6.050	5.272	6.339	6.243	8.041
	SD	13.901	14.846	13.842	13.661***	18.457
	IR	0.435	0.355	0.458	0.457	0.436
$N = 50$	AV	4.506	3.198	3.792	4.221	8.458
	SD	13.060	14.448	13.151	12.842***	18.271
	IR	0.345	0.221	0.288	0.329	0.463
$N = 100$	AV	2.981	1.356	1.317	2.553	8.273
	SD	10.892	13.134	10.989	10.726***	18.225
	IR	0.274	0.103	0.120	0.238	0.454
$N = 300$	AV	6.239	1.654	2.514	5.418	9.034
	SD	7.986	11.163	7.876	7.750***	17.975
	IR	0.781	0.148	0.319	0.699	0.503
$N = 500$	AV	5.108	0.820	1.672	4.554	9.558
	SD	6.621	9.137	6.476	6.375***	17.961
	IR	0.771	0.090	0.258	0.714	0.532

Note: This table contains the annualized average (AV), standard deviation (SD) and information ratio (IR) of the out-of-sample daily log returns for the GMVPs constructed using different dynamic correlation models and the $1/N$ portfolio. The lowest SD per dimension size is highlighted in bold. The out-of-sample periods ranges from December 1990 until December 2020 for a total of 7560 days, using an estimation window of 2500 days and re-estimation of the parameters every 21 days. A significant decrease of the (logarithmic squared) SD of the CLIP-DCC model compared to the DCC model is indicated with a *,** and *** for a p -value below 0.1, 0.05 and 0.01, respectively, using the two-sided test by Ledoit and Wolf (2011) with HAC standard errors.

out-of-sample variance, as proxied by the sum of the GMVP squared log returns, of the CLIP-DCC model from the DCC model. A higher value therefore reflects a larger gain from the use of the CLIP-DCC framework relative to the DCC model.

Figure 2: Difference in cumulative out-of-sample GMVP variance between the DCC and CLIP-DCC model, December 1990 until December 2020.



Note: This figure contains for different investment universe sizes the evolution of the difference in cumulative out-of-sample GMVP variance between the DCC and CLIP-DCC model, denoted by $\Delta \Sigma_t r_{GMV,t}^2 = \sum_{t=1}^T [r_{DCC,t}^2 - r_{CLIP-DCC,t}^2]$, where $r_{DCC,t}$ and $r_{CLIP-DCC,t}$ denote the out-of-sample logarithmic returns of the GMVP constructed by the DCC and CLIP-DCC model, respectively. The vertical lines reflect economically relevant dates.

In Figure 2, we mostly find a steadily increasing cumulative benefit from using the CLIP-DCC model. However, during the peak of the 2008 Great Financial Crisis, we observe a large sudden increase. Although the movement may visually appear to be near instantaneous, it spans about a one to three month period. Furthermore, a second upward jump (albeit of a smaller magnitude) is located around Black Monday on August 8, 2011, following the downgrading of US sovereign debt by Standard and Poor's. This highlights the beneficial effects of the CLIP-DCC framework around periods of financial turmoil. One possible explanation is an increase in equity correlation co-movement associ-

ated with market downturns (Ang and Chen, 2002). This could reduce the bias incurred by the cross-sectional structure of the CLIP-DCC model. A second possible explanation may be an increase in uncertainty of the conditional correlations during these periods of high volatility, favoring the pooling aspect of the CLIP-DCC framework. A third reason may be that the lower weight on the past correlation matrix ($\hat{\beta}$) combined with a higher weight for innovations ($\hat{\alpha}$) for the CLIP-DCC model compared to the DCC model allows for a more rapid identification of this recessionary period.

In Figure 2, we also find that the CLIP-DCC and DCC models offer comparable performance between 2011-2020 for the largest investment universes $N = 300$ and 500 . This suggests that for large cross-sections the pooling function $\theta(\cdot)$ may be overly restrictive. To investigate more flexible pooling structures, Appendix B.1 considers Block-CLIP-DCC models as outlined in Section 2.4 using industry memberships. Although the Block-CLIP-DCC model outperforms the standard Block-DECO model, it does not outperform the CLIP-DCC. This may be due to industry memberships not reflecting a particularly strong clustering, see Oh and Patton (2023).

Finally, we find mixed effects of the COVID-19 pandemic around the date of the US lockdown in March 2020. In particular, we have that the CLIP-DCC framework provides benefits for the smaller investment universes $N = \{10, 30, 50\}$, losses for the medium sizes $N = \{100, 300\}$ and has almost no effect for $N = 500$. The impact of this pandemic recession is thus more ambiguous, likely due to its different nature.

Next, we investigate the effects of target shrinkage for the long-run correlation matrix \mathbf{C} and its interplay with the CLIP-DCC approach. Table 4 presents summary statistics of the daily out-of-sample log returns of the GMVPs of the DCC, dDECO, sdDECO and CLIP-DCC models using the NLS approach of Ledoit and Wolf (2020) for the target. Comparing Table 3 and Table 4, we observe that NLS of the target decreases portfolio SD in all cases. We find that the benefits of NLS are minor for the small portfolio sizes up to $N = 100$, but increase for the large portfolio sizes $N = 300$ and $N = 500$. Appendix B.4 examines the impact of a shorter estimation window of $T = 1250$ days, where we find even larger gains from using NLS. This is in line with intuition, because the quality of the sample covariance matrix rapidly deteriorates as the concentration ratio N/T rises (Ledoit and Wolf, 2004).

Table 4: Daily out-of-sample GMVP performance using different DCC models for $N \in \{10, 30, 50, 100, 300, 500\}$ and NLS of the target, December 1990 until December 2020.

		DCC	dDECO	sdDECO	CLIP-DCC	1/N
$N = 10$	AV	4.858	3.758	5.787	5.277	7.283
	SD	15.471	15.745	15.474	15.388*	19.167
	IR	0.314	0.239	0.374	0.343	0.380
$N = 30$	AV	6.064	5.271	6.377	6.261	8.041
	SD	13.891	14.845	13.826	13.650***	18.457
	IR	0.437	0.355	0.461	0.459	0.436
$N = 50$	AV	4.502	3.199	3.883	4.247	8.458
	SD	13.035	14.446	13.116	12.812***	18.271
	IR	0.345	0.221	0.296	0.332	0.463
$N = 100$	AV	2.979	1.356	1.358	2.517	8.273
	SD	10.846	13.132	10.956	10.682***	18.225
	IR	0.275	0.103	0.124	0.236	0.454
$N = 300$	AV	6.198	1.665	2.791	5.313	9.034
	SD	7.841	11.161	7.745	7.607***	17.975
	IR	0.791	0.149	0.360	0.698	0.503
$N = 500$	AV	4.865	0.820	1.778	4.305	9.558
	SD	6.400	9.137	6.206	6.122***	17.961
	IR	0.760	0.090	0.287	0.703	0.532

Note: This table contains the annualized average (AV), standard deviation (SD) and information ratio (IR) of the out-of-sample daily log returns for the GMVPs constructed using different dynamic correlation models and the $1/N$ portfolio. The lowest SD per dimension size is highlighted in bold. The out-of-sample periods ranges from December 1990 until December 2020 for a total of 7560 days, using an estimation window of 2500 days and re-estimation of the parameters every 21 days. A significant decrease of the (logarithmic squared) SD of the CLIP-DCC model compared to the DCC model is indicated with a *,** and *** for a p -value below 0.1, 0.05 and 0.01, respectively, using the two-sided test by Ledoit and Wolf (2011) with HAC standard errors.

Furthermore, also when using target shrinkage the CLIP-DCC model achieves the lowest out-of-sample SD among all models. In fact, we find that the improvements of NLS and the CLIP-DCC model over the base DCC model are additive. This confirms their theoretical

complementary nature, acting on different components of the uncertainty of the conditional correlation matrix. In sum, we recommend using the CLIP-DCC model combined with NLS of the target to model conditional correlation matrices.

5 Conclusion

We augment the Dynamic Conditional Correlation (DCC) model of Engle (2002) by allowing for partial commonality in the update of the conditional correlations. Specifically, we propose the DCC model with Conditional Linear Information Pooling (CLIP-DCC) which allows the correlation innovation of each pair to parsimoniously depend on the information contained in all asset return pairs. Additionally, the CLIP-DCC model preserves long-run dynamics, thereby naturally complementing methods that shrink the correlation target. We demonstrate in a Monte Carlo study that the parameters of the CLIP-DCC model, including the pooling intensity, can be effectively estimated using composite likelihood. A real-time empirical application to a large selection of US stocks from 1981 until 2020 finds significant benefits of the CLIP-DCC model for a minimum-variance investor. Combining the CLIP-DCC model with the shrinkage estimator of Ledoit and Wolf (2020) for the target yields complementary benefits, confirming that these methods target different sources of uncertainty of the conditional correlation matrix.

References

- Aielli, G. P. (2013). Dynamic conditional correlation: On properties and estimation. *Journal of Business & Economic Statistics* 31(3), 282–299.
- Ang, A. and J. Chen (2002). Asymmetric correlations of equity portfolios. *Journal of Financial Economics* 63(3), 443–494.
- Bauwens, L., S. Laurent, and J. V. Rombouts (2006). Multivariate garch models: A survey. *Journal of Applied Econometrics* 21(1), 79–109.
- Bollerslev, T. (1986). Generalized autoregressive conditional heteroskedasticity. *Journal of Econometrics* 31(3), 307–327.

- Bollerslev, T., A. J. Patton, and R. Quaadvlieg (2020). Multivariate leverage effects and realized semicovariance GARCH models. *Journal of Econometrics* 217(2), 411–430.
- De Nard, G., R. F. Engle, O. Ledoit, and M. Wolf (2022). Large dynamic covariance matrices: Enhancements based on intraday data. *Journal of Banking & Finance* 138, 106426.
- Engle, R. F. (2002). 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. and B. Kelly (2012). Dynamic equicorrelation. *Journal of Business & Economic Statistics* 30(2), 212–228.
- Engle, R. F., O. Ledoit, and M. Wolf (2019). Large dynamic covariance matrices. *Journal of Business & Economic Statistics* 37(2), 363–375.
- Glosten, L. R., R. Jagannathan, and D. E. Runkle (1993). On the relation between the expected value and the volatility of the nominal excess return on stocks. *The Journal of Finance* 48(5), 1779–1801.
- Hafner, C. M. and O. Reznikova (2012). On the estimation of dynamic conditional correlation models. *Computational Statistics & Data Analysis* 56(11), 3533–3545.
- Ledoit, O. and M. Wolf (2004). Honey, I shrunk the sample covariance matrix. *The Journal of Portfolio Management* 30(4), 110–119.
- Ledoit, O. and M. Wolf (2011). Robust performances hypothesis testing with the variance. *Wilmott* 2011(55), 86–89.
- Ledoit, O. and M. Wolf (2020). Analytical nonlinear shrinkage of large-dimensional covariance matrices. *The Annals of Statistics* 48(5), 3043–3065.
- Llorens-Terrazas, J. and C. Brownlees (2022). Projected dynamic conditional correlations. *International Journal of Forecasting*, In Press.

- Michaud, R. O. (1989). The Markowitz optimization enigma: Is ‘optimized’ optimal? *Financial Analysts Journal* 45(1), 31–42.
- Oh, D. H. and A. J. Patton (2023). Dynamic factor copula models with estimated cluster assignments. *Journal of Econometrics*, In Press.
- Pakel, C., N. Shephard, K. Sheppard, and R. F. Engle (2021). Fitting vast dimensional time-varying covariance models. *Journal of Business & Economic Statistics* 39(3), 652–668.
- Patton, A. J. and K. Sheppard (2009). Evaluating volatility and correlation forecasts. In T. Andersen, R. Davis, J.-P. Kreiss, and T. Mikosch (Eds.), *Handbook of Financial Time Series*, pp. 801–838. Springer Verlag.
- Silvennoinen, A. and T. Teräsvirta (2009). Multivariate GARCH models. In T. Andersen, R. Davis, J.-P. Kreiss, and T. Mikosch (Eds.), *Handbook of Financial Time Series*, pp. 201–229. Springer Verlag.

Online Appendix to Dynamic Conditional Correlations with Partial Information Pooling

A Proofs	1
A.1 Proposition 1	1
A.2 Proposition 2	1
A.3 Proposition 3	2
B Additional Results	6
B.1 Block-based models	6
B.2 Mean-variance portfolios	8
B.3 QLIKE losses	8
B.4 Robustness estimation length	12

A Proofs

A.1 Proposition 1

First, we note that $\frac{1}{d}\mathbf{S}$ with $d > 0$ is positive definite if and only if \mathbf{S} is positive definite. Next, we may write $\frac{1}{d}\mathbf{S} = \rho\mathbf{J}_N + (1 - \rho)\mathbf{I}_N$ with $\rho = \frac{\rho}{d}$, which may be shown to be positive definite if and only if $\frac{\rho}{d} \in (\frac{-1}{N-1}, 1)$. This follows from the fact that $\frac{1}{d}\mathbf{S}$ has two distinct eigenvalues, $1 + (N - 1)\rho$ and $1 - \rho$, which can be seen to both be positive if and only if $\frac{\rho}{d} \in (\frac{-1}{N-1}, 1)$. This also reveals that standardizing a positive definite compound symmetric matrix \mathbf{S} yields a valid equicorrelation matrix $\frac{1}{d}\mathbf{S}$ with equicorrelation $\frac{\rho}{d}$, see also Definition 2.1 and Proposition 2.1 of Engle and Kelly (2012).

A.2 Proposition 2

For the case that $N = 1$ we define $\theta(\cdot)$ to be the identity mapping, such that the properties of Proposition 2 trivially hold. We now proceed with the case $N \geq 2$. By noting that $\theta(\cdot)$ is a linear function of $\theta^D(\cdot)$ and $\theta^O(\cdot)$, which are sums and thus clearly linear themselves, we have that $\theta(\cdot)$ is a linear mapping and properties 1 and 2 immediately follow.

Using Proposition 1, it suffices for the positive definiteness part of property 3 to show that $\theta^D(\mathbf{A}) > 0$ and $\frac{\theta^O(\mathbf{A})}{\theta^D(\mathbf{A})} \in (\frac{-1}{N-1}, 1)$. By definition we have that if \mathbf{A} is positive definite that $\mathbf{x}'\mathbf{A}\mathbf{x} > 0, \forall \mathbf{x} \in \mathbb{R}^N$ with $\mathbf{x} \neq 0_N$. By selecting \mathbf{x} to be a vector containing a single 1 at the i -th position and 0 elsewhere, we observe that $a_{ii} > 0$ for all $i \in 1, \dots, N$. This implies that all diagonal elements are strictly larger than 0, such that clearly $\theta^D(\mathbf{A}) > 0$. In addition, by selecting \mathbf{x} to be a vector with 1 on the i -th position, -1 on the j -th position and 0 elsewhere, we obtain $a_{ii} + a_{jj} > a_{ij} + a_{ji}$ for all $i, j \in 1, \dots, N$, where $i \neq j$. Taking sums over i and j , using $a_{ij} = a_{ji}$ by symmetry and using the definitions of $\theta^D(\cdot)$ and $\theta^O(\cdot)$, we obtain

$$\begin{aligned}
& \sum_{i=1}^N \sum_{j=1}^N (a_{ii} + a_{jj}) > \sum_{i=1}^N \sum_{j=1}^N (a_{ij} + a_{ji}) \\
N \sum_{i=1}^N a_{ii} + N \sum_{j=1}^N a_{jj} &> 2 \sum_{i=1}^N \sum_{j=1}^N a_{ij} \\
N \sum_{i=1}^N a_{ii} &> \sum_{i=1}^N \sum_{j=1}^N a_{ij} \\
N \sum_{i=1}^N a_{ii} &> \sum_{i=1}^N \sum_{j=1, j \neq i}^N a_{ij} + \sum_{i=1}^N \sum_{j=i}^N a_{ij} \\
N \sum_{i=1}^N a_{ii} &> \sum_{i=1}^N \sum_{j=1, j \neq i}^N a_{ij} + \sum_{i=1}^N a_{ii} \\
(N-1) \sum_{i=1}^N a_{ii} &> \sum_{i=1}^N \sum_{j=1, j \neq i}^N a_{ij} \\
\frac{1}{N} \sum_{i=1}^N a_{ii} &> \frac{1}{N(N-1)} \sum_{i=1}^N \sum_{j=1, j \neq i}^N a_{ij} \\
\theta^D(\mathbf{A}) &> \theta^O(\mathbf{A}),
\end{aligned} \tag{A.1}$$

such that $\frac{\theta^O(\mathbf{A})}{\theta^D(\mathbf{A})} < 1$. Finally, by selecting $\mathbf{x} = \iota_N$ we obtain $N\theta^D(\mathbf{A}) + N(N-1)\theta^O(\mathbf{A}) > 0$ which in turn may be rewritten to show that $\frac{\theta^O(\mathbf{A})}{\theta^D(\mathbf{A})} > \frac{-1}{N-1}$. Together, this entails that $\frac{\theta^O(\mathbf{A})}{\theta^D(\mathbf{A})} \in (\frac{-1}{N-1}, 1)$, which concludes the proof.

For the positive semi-definite case of property 3, we note that if \mathbf{A} is positive semi-definite then by definition we have that $\mathbf{x}'\mathbf{A}\mathbf{x} \geq 0, \forall \mathbf{x} \in \mathbb{R}^N$. Therefore we may use the same arguments as for the positive definite case but lose the strictness of the inequalities. Here we note that if $\theta^D(\mathbf{A}) = 0$, this implies that \mathbf{A} and also $\theta(\mathbf{A})$ are an $N \times N$ matrix of zeros, which is positive semi-definite (all eigenvalues are 0). Therefore for arguments that utilize $\frac{\theta^O(\mathbf{A})}{\theta^D(\mathbf{A})}$ we can consider the case that $\theta^D(\mathbf{A}) > 0$ and obtain that $\frac{\theta^O(\mathbf{A})}{\theta^D(\mathbf{A})} \in [\frac{-1}{N-1}, 1]$, such that $\theta(\mathbf{A})$ is positive semi-definite.

A.3 Proposition 3

From Proposition 2 we have that $\theta(\cdot)$ is a linear function. It can be straightforwardly verified that also the matrix averaging transformation $\tau(\cdot)$ is a linear function as it can be

written as a sum. Therefore, since $\theta^{BL}(\cdot, G)$ is composed of $\theta(\cdot)$ and $\tau(\cdot)$ operations, it is straightforward to show that properties 1 and 2 of Proposition 3 hold.

With regards to property 3, we follow the proof structure of Theorem 1 from Roustant and Deville (2017). Specifically, we first assume that \mathbf{A} is positive semi-definite and define the block averaging function $\tau^{BL}(\mathbf{A}, G)$ for $\mathbf{A} \in \mathbb{R}^{N \times N}$ with $N \geq 2$ and block structure G ,

$$\tau^{BL}(\mathbf{A}, G) = \begin{bmatrix} \tau(\mathbf{A}_{11}^*) & \tau(\mathbf{A}_{12}^*) & \cdots & \tau(\mathbf{A}_{1K}^*) \\ \tau(\mathbf{A}_{21}^*) & \tau(\mathbf{A}_{22}^*) & \cdots & \tau(\mathbf{A}_{2K}^*) \\ \vdots & \vdots & \ddots & \vdots \\ \tau(\mathbf{A}_{K1}^*) & \tau(\mathbf{A}_{K2}^*) & \cdots & \tau(\mathbf{A}_{KK}^*) \end{bmatrix}, \quad (\text{A.2})$$

which may also be written as

$$\tau^{BL}(\mathbf{A}, G) = \mathbf{L}(G)\mathbf{A}\mathbf{L}(G)', \quad (\text{A.3})$$

$$\mathbf{L}(G) = \begin{bmatrix} 1/n_1 \mathbf{J}_{n_1} & \mathbf{O}_{n_1 \times n_2} & \cdots & \mathbf{O}_{n_1 \times n_K} \\ \mathbf{O}_{n_2 \times n_1} & 1/n_2 \mathbf{J}_{n_2} & \cdots & \mathbf{O}_{n_2 \times n_K} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{O}_{n_K \times n_1} & \mathbf{O}_{n_K \times n_2} & \cdots & 1/n_K \mathbf{J}_{n_K} \end{bmatrix}, \quad (\text{A.4})$$

where $\mathbf{O}_{n_j \times n_k}$ is an $n_j \times n_k$ matrix of zeros and \mathbf{A}_{ij}^* for $i, j = 1, 2, \dots, K$ is the ij -th block of \mathbf{A} using the block partition $G := [n_1, n_2, \dots, n_K]'$. Here $\mathbf{L}(G)$ can be seen to be positive semi-definite as it is a block diagonal matrix with positive semi-definite diagonal blocks. From the quadratic form it can be seen that $\tau^{BL}(\mathbf{A}, G)$ is positive semi-definite if \mathbf{A} is positive semi-definite. We then consider the difference $\Delta(\mathbf{A}, G) := \theta^{BL}(\mathbf{A}, G) - \tau^{BL}(\mathbf{A}, G)$ which admits the following form

$$\Delta(\mathbf{A}, G) = \begin{bmatrix} \theta(\mathbf{A}_{11}^*) - \tau(\mathbf{A}_{11}^*) & \mathbf{O}_{n_1 \times n_2} & \cdots & \mathbf{O}_{n_1 \times n_K} \\ \mathbf{O}_{n_2 \times n_1} & \theta(\mathbf{A}_{22}^*) - \tau(\mathbf{A}_{22}^*) & \cdots & \mathbf{O}_{n_2 \times n_K} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{O}_{n_K \times n_1} & \mathbf{O}_{n_K \times n_2} & \cdots & \theta(\mathbf{A}_{KK}^*) - \tau(\mathbf{A}_{KK}^*) \end{bmatrix}, \quad (\text{A.5})$$

that is, $\Delta(\mathbf{A}, G)$ is a block diagonal matrix consisting of the matrices $\theta(\mathbf{A}_{jj}^*) - \tau(\mathbf{A}_{jj}^*)$ for

$j = 1, \dots, K$, which may be written as

$$\begin{aligned}
\theta(\mathbf{A}_{jj}^*) - \tau(\mathbf{A}_{jj}^*) &= \theta(\mathbf{A}_{jj}^*) - \frac{1}{n_j^2} \mathbf{l}'_{n_j} \mathbf{A}_{jj}^* \mathbf{l}_{n_j} \mathbf{J}_{n_j} \\
&= \theta(\mathbf{A}_{jj}^*) - \frac{1}{n_j^2} [n_j(n_j - 1)\theta^O(\mathbf{A}_{jj}^*) + n_j\theta^D(\mathbf{A}_{jj}^*)] \mathbf{J}_{n_j} \\
&= \theta^O(\mathbf{A}_{jj}^*) \mathbf{J}_{n_j} + [\theta^D(\mathbf{A}_{jj}^*) - \theta^O(\mathbf{A}_{jj}^*)] \mathbf{I}_{n_j} - \frac{1}{n_j^2} [n_j(n_j - 1)\theta^O(\mathbf{A}_{jj}^*) + n_j\theta^D(\mathbf{A}_{jj}^*)] \mathbf{J}_{n_j} \\
&= [\theta^D(\mathbf{A}_{jj}^*) - \theta^O(\mathbf{A}_{jj}^*)] \mathbf{I}_{n_j} + \frac{1}{n_j} [\theta^O(\mathbf{A}_{jj}^*) - \theta^D(\mathbf{A}_{jj}^*)] \mathbf{J}_{n_j} \\
&= [\theta^D(\mathbf{A}_{jj}^*) - \theta^O(\mathbf{A}_{jj}^*)] [\mathbf{I}_{n_j} - \frac{1}{n_j} \mathbf{J}_{n_j}],
\end{aligned} \tag{A.6}$$

where $\theta^D(\mathbf{A}_{jj}^*) - \theta^O(\mathbf{A}_{jj}^*)$ is a non-negative scalar (see again the proof of Proposition 2) and $\mathbf{I}_{n_j} - \frac{1}{n_j} \mathbf{J}_{n_j}$ a positive semi-definite matrix. The latter fact may be derived from Proposition 1 by noting that it is a CS matrix or from noticing that is in fact a projection matrix. We now have that $\Delta(\mathbf{A}, G)$ is positive semi-definite because all its diagonal matrices are positive semi-definite. Since $\theta^{BL}(\mathbf{A}, G) = \Delta(\mathbf{A}, G) + \tau^{BL}(\mathbf{A}, G)$, it follows that $\theta^{BL}(\mathbf{A}, G)$ is also positive semi-definite as it is the sum of two positive semi-definite matrices.

Finally, we show that if \mathbf{A} is positive definite that $\theta^{BL}(\mathbf{A}, G)$ is also positive definite. For $\mathbf{x} \in \mathbb{R}^N$ we consider $h(\mathbf{x}) := \mathbf{x}' \theta^{BL}(\mathbf{A}, G) \mathbf{x}$, which may also be written as

$$h(\mathbf{x}) = \mathbf{x}' \Delta(\mathbf{A}, G) \mathbf{x} + \mathbf{x}' \tau^{BL}(\mathbf{A}, G) \mathbf{x}, \tag{A.7}$$

where $h(\mathbf{x})$ is 0 if and only if $\mathbf{x}' \Delta(\mathbf{A}, G) \mathbf{x}$ and $\mathbf{x}' \tau^{BL}(\mathbf{A}, G) \mathbf{x}$ are both equal to 0 (as neither can be negative due to positive semi-definiteness). First, we have for $\mathbf{x}' \tau^{BL}(\mathbf{A}, G) \mathbf{x}$ that

$$\mathbf{x}' \tau^{BL}(\mathbf{A}, G) \mathbf{x} = \mathbf{x}' \mathbf{L}(G) \mathbf{A} \mathbf{L}(G)' \mathbf{x} = \mathbf{x}^m(G)' \mathbf{A} \mathbf{x}^m(G), \tag{A.8}$$

$$\mathbf{x}^m(G) = \left[\left(\frac{1}{n_1} \mathbf{l}'_{n_1} \mathbf{x}_1^* \right) \mathbf{l}'_{n_1} \quad \left(\frac{1}{n_2} \mathbf{l}'_{n_2} \mathbf{x}_2^* \right) \mathbf{l}'_{n_2} \quad \cdots \quad \left(\frac{1}{n_K} \mathbf{l}'_{n_K} \mathbf{x}_K^* \right) \mathbf{l}'_{n_K} \right]', \tag{A.9}$$

where \mathbf{x}_j^* for $j = 1, \dots, K$ is the j -th subvector of \mathbf{x} based on the group structure G . Because \mathbf{A} is assumed positive definite we have that $\mathbf{x}' \tau^{BL}(\mathbf{A}, G) \mathbf{x}$ is 0 if and only if $\mathbf{x}^m(G) = \mathbf{0}_N$. We also observe that $\mathbf{x}^m(G)$ can be viewed as a vector of group means, such that $\mathbf{x}^m(G) = \mathbf{0}_N$ if and only if the group means of the vector \mathbf{x} based on the group structure G are all 0. That is, $\mathbf{x}' \tau^{BL}(\mathbf{A}, G) \mathbf{x} = 0$ if and only if $\frac{1}{n_j} \mathbf{l}'_{n_j} \mathbf{x}_j^* = 0$ for all $j = 1, \dots, K$. Note that if $n_j = 1$, i.e. for groups with only one member, this directly implies that that $\mathbf{x}_j^* = 0$.

Second, we have for $\mathbf{x}'\Delta(\mathbf{A}, G)\mathbf{x}$ that

$$\mathbf{x}'\Delta(\mathbf{A}, G)\mathbf{x} = \sum_{j=1}^K [\theta^D(\mathbf{A}_{jj}^*) - \theta^O(\mathbf{A}_{jj}^*)] \mathbf{x}_j^{*'} [\mathbf{I}_{n_j} - \frac{1}{n_j} \mathbf{J}_{n_j}] \mathbf{x}_j^*, \quad (\text{A.10})$$

which is 0 if and only if $\mathbf{x}_j^{*'} [\mathbf{I}_{n_j} - \frac{1}{n_j} \mathbf{J}_{n_j}] \mathbf{x}_j^* = 0$ for all $j = 1, \dots, K$, because $[\theta^D(\mathbf{A}_{jj}^*) - \theta^O(\mathbf{A}_{jj}^*)] > 0$ for all $j = 1, \dots, K$ with $n_j \geq 2$ if \mathbf{A} is positive definite. Using the symmetric square root of $\mathbf{I}_{n_j} - \frac{1}{n_j} \mathbf{J}_{n_j}$, we may show that $\mathbf{x}_j^{*'} [\mathbf{I}_{n_j} - \frac{1}{n_j} \mathbf{J}_{n_j}] \mathbf{x}_j^* = 0$ implies that $[\mathbf{I}_{n_j} - \frac{1}{n_j} \mathbf{J}_{n_j}]^{1/2} \mathbf{x}_j^* = 0_{n_j}$, which in turn implies $[\mathbf{I}_{n_j} - \frac{1}{n_j} \mathbf{J}_{n_j}] \mathbf{x}_j^* = 0_{n_j}$. From there, we observe that the condition $\mathbf{x}_j^* - \iota_{n_j} (\frac{1}{n_j} \iota_{n_j}' \mathbf{x}_j^*) = 0_{n_j}$ only holds if \mathbf{x}_j^* is a scalar multiple of ι_{n_j} . If we combine this with the requirements for $\mathbf{x}'\tau^{BL}(\mathbf{A}, G)\mathbf{x} = 0$, we then get that $h(\mathbf{x})$ is 0 if and only if $\mathbf{x}_j^* = 0_{n_j}$ for all $j = 1, \dots, K$, which is equivalent to $\mathbf{x} = 0_N$. This means that $\mathbf{x}'\mathbf{A}\mathbf{x} > 0$ for all $\mathbf{x} \in \mathbb{R}^N$ where $\mathbf{x} \neq 0_N$. We conclude that if \mathbf{A} is positive definite then $\theta^{BL}(\mathbf{A}, G)$ is also positive definite.

B Additional Results

B.1 Block-based models

We examine the performance of the block-based CLIP-DCC models as outlined in (23)-(25) and follow Engle and Kelly (2012) by using industry group membership based on SIC codes to impose a block structure. Specifically, we use the $K = 5$ and $K = 10$ Fama-French industry categorisation and estimate the BL-dDECO, BL-sdDECO and BL-CLIP-DCC models for $N \in \{100, 300, 500\}$, also employing NLS of the target. Table B.1 summarizes the out-of-sample performance of the GMVPs constructed using these models.

In Table B.1, we observe large improvements of the BL-dDECO model over the dDECO model, some improvements of the BL-sdDECO model over the sdDECO model and nearly no benefit of the BL-CLIP-DCC model over the CLIP-DCC model. Comparing the block structures, we find $K = 10$ to be superior for the BL-dDECO and BL-sdDECO models, while $K = 5$ is slightly better for the BL-CLIP-DCC model. This confirms that the poor performance of the dDECO model is mainly the result of imposing too much structure on the long-run correlation matrix and loosening it increases performance. For example, for $N = 500$ we find an SD of 9.137 for the dDECO model, an SD of 8.394 and 7.347 for the $K = 5$ and $K = 10$ BL-dDECO models and an SD of 6.400 for the DCC model. Furthermore, the fact that the BL-CLIP-DCC model appears to offer no benefit over the CLIP-DCC model suggests that correlation movement information within industries is not much more informative than information from assets in other industries. For a further discussion of the appropriateness of clustering based on SIC codes, we refer to Oh and Patton (2023). It would be interesting for future research to consider an empirical application with data that admits a very clear group structure. For example, by considering many assets from a few very differently behaving asset classes.

Table B.1: Daily out-of-sample GMVP performance constructed using different industry block-based DCC models for $N \in \{100, 300, 500\}$ and NLS of the target, December 1990 until December 2020.

$K = 5$		BL-dDECO	BL-sdDECO	BL-CLIP-DCC	1/N
$N = 100$	AV	1.112	1.063	2.233	8.273
	SD	12.116	10.890	10.675	18.225
	IR	0.092	0.098	0.209	0.454
$N = 300$	AV	1.594	2.818	5.250	9.034
	SD	10.130	7.711	7.590	17.975
	IR	0.157	0.365	0.692	0.503
$N = 500$	AV	1.023	1.938	4.328	9.558
	SD	8.394	6.208	6.109	17.961
	IR	0.122	0.312	0.708	0.532

$K = 10$		BL-dDECO	BL-sdDECO	BL-CLIP-DCC	1/N
$N = 100$	AV	2.028	1.842	2.693	8.273
	SD	11.657	10.873	10.694	18.225
	IR	0.174	0.169	0.252	0.454
$N = 300$	AV	2.889	3.084	5.336	9.034
	SD	9.080	7.730	7.597	17.975
	IR	0.318	0.399	0.702	0.503
$N = 500$	AV	1.628	2.096	4.339	9.558
	SD	7.347	6.234	6.121	17.961
	IR	0.222	0.336	0.709	0.532

Note: This table contains the annualized average (AV), standard deviation (SD) and information ratio (IR) of the out-of-sample daily log returns for the GMVPs constructed using different block-based dynamic correlation models and the $1/N$ portfolio. The lowest SD per dimension size is highlighted in bold. The out-of-sample periods ranges from December 1990 until December 2020 for a total of 7560 days, using an estimation window of 2500 days and re-estimation of the parameters every 21 days. Five ($K = 5$) and ten ($K = 10$) Fama-French industry memberships are used to impose a block-structure.

B.2 Mean-variance portfolios

This section considers mean-variance (MV) portfolios using a similar re-balancing strategy as for the GMVPs in the main text. For the MV portfolios the variance is minimized subject to a return constraint. Specifically, we mimic the strategies of Engle and Kelly (2012) and Engle et al. (2019), which use the sample mean and a momentum signal for the return constraint, respectively. The results are contained in Appendix Table B.2 and B.3. Here we also find the CLIP-DCC model to significantly reduce the SD compared to the DCC model, although the effect on the mean differs across portfolio sizes. Maximum IR portfolios or portfolios with leverage constraints are left for future research.

B.3 QLIKE losses

For robustness, we use the quasi-likelihood (QLIKE) loss of Patton and Sheppard (2009) to directly evaluate the one-step ahead covariance forecasts made using the different correlation models. Specifically, Appendix Table B.4 summarizes the QLIKE performance of the models using the average (AV) loss, the standard deviation (SD) of the losses and the proportion of improvement (PI), which reflects the share of dates that the model has a lower QLIKE than the base DCC model. Overall, we conclude again that the CLIP-DCC model outperforms the DCC model and the dDECO model looking at both the AV and PI values. Comparing the dDECO and DCC model, we find the latter to outperform for all dimension sizes, except $N = 500$. This indicates that the structure of the dDECO model may only be useful in the very large dimensional setting. Interestingly, the sdDECO model is the best performing in terms of the AV for the larger dimensions ($N = 50, 100, 300, 500$), while the CLIP-DCC model is the best in terms of the PI. This result for the PI paints a similar picture as Figure 2, where we also observe that the CLIP-DCC model consistently beats the DCC model by a small amount, which adds up over time.

Table B.2: Daily out-of-sample MV portfolio (standard) performance constructed using different DCC models for $N \in \{10, 30, 50, 100, 300, 500\}$, December 1990 until December 2020.

		DCC	dDECO	sdDECO	CLIP-DCC	1/N
$N = 10$	AV	5.028	4.261	5.497	5.201	7.283
	SD	16.317	16.666	16.374	16.268	19.167
	IR	0.308	0.256	0.336	0.320	0.380
$N = 30$	AV	6.300	5.822	6.584	6.492	8.041
	SD	14.016	15.176	13.983	13.796***	18.457
	IR	0.449	0.384	0.471	0.471	0.436
$N = 50$	AV	4.819	3.600	4.127	4.578	8.458
	SD	13.069	14.585	13.167	12.853***	18.271
	IR	0.369	0.247	0.313	0.356	0.463
$N = 100$	AV	3.765	2.117	2.226	3.413	8.273
	SD	11.012	13.178	11.129	10.846***	18.225
	IR	0.342	0.161	0.200	0.315	0.454
$N = 300$	AV	7.023	2.570	3.133	6.261	9.034
	SD	8.115	10.985	8.060	7.895***	17.975
	IR	0.865	0.234	0.389	0.793	0.503
$N = 500$	AV	5.450	1.756	1.806	4.950	9.558
	SD	6.789	9.149	6.697	6.559***	17.961
	IR	0.803	0.192	0.270	0.755	0.532

Note: This table contains the annualized average (AV), standard deviation (SD) and information ratio (IR) of the out-of-sample daily log returns for the MV portfolios constructed using different dynamic correlation models and the $1/N$ portfolio. Specifically, we use the geometric sample mean and a 10 percent annual return target, similar to Engle and Kelly (2012). The lowest SD per dimension size is highlighted in bold. The out-of-sample periods ranges from December 1990 until December 2020 for a total of 7560 days, using an estimation window of 2500 days and re-estimation of the parameters every 21 days. A significant decrease of the (logarithmic squared) SD of the CLIP-DCC model compared to the DCC model is indicated with a *,** and *** for a p -value below 0.1, 0.05 and 0.01, respectively, using the two-sided test by Ledoit and Wolf (2011) with HAC standard errors.

Table B.3: Daily out-of-sample MV portfolio (momentum) performance constructed using different DCC models for $N \in \{10, 30, 50, 100, 300, 500\}$, December 1990 until December 2020.

		DCC	dDECO	sdDECO	CLIP-DCC	1/N
$N = 10$	AV	5.211	3.960	5.727	5.379	7.283
	SD	16.671	16.977	16.727	16.630	19.167
	IR	0.313	0.233	0.342	0.323	0.380
$N = 30$	AV	7.157	6.780	7.191	7.273	8.041
	SD	14.603	15.210	14.480	14.353***	18.457
	IR	0.490	0.446	0.497	0.507	0.436
$N = 50$	AV	5.522	5.098	4.693	5.268	8.458
	SD	13.609	14.507	13.647	13.378***	18.271
	IR	0.406	0.351	0.344	0.394	0.463
$N = 100$	AV	4.402	3.180	2.317	3.866	8.273
	SD	11.374	12.921	11.402	11.178***	18.225
	IR	0.387	0.246	0.203	0.346	0.454
$N = 300$	AV	7.396	3.530	3.527	6.646	9.034
	SD	8.601	11.025	8.442	8.318***	17.975
	IR	0.860	0.320	0.418	0.799	0.503
$N = 500$	AV	5.987	2.229	2.419	5.488	9.558
	SD	7.246	9.215	7.081	6.953***	17.961
	IR	0.826	0.242	0.342	0.789	0.532

Note: This table contains the annualized average (AV), standard deviation (SD) and information ratio (IR) of the out-of-sample daily log returns for the MV portfolios constructed using different dynamic correlation models and the $1/N$ portfolio. Specifically, we use a momentum signal taking the geometric mean over the previous 252 days, excluding the most recent 21 days, similar to Engle et al. (2019). The target return is set to the arithmetic mean of this mean vector. The lowest SD per dimension size is highlighted in bold. The out-of-sample periods ranges from December 1990 until December 2020 for a total of 7560 days, using an estimation window of 2500 days and re-estimation of the parameters every 21 days. A significant decrease of the (logarithmic squared) SD of the CLIP-DCC model compared to the DCC model is indicated with a *, ** and *** for a p -value below 0.1, 0.05 and 0.01, respectively, using the two-sided test by Ledoit and Wolf (2011) with HAC standard errors.

Table B.4: Daily out-of-sample QLIKE using different DCC models for $N \in \{10, 30, 50, 100, 300, 500\}$, December 1990 until December 2020.

		DCC	dDECO	sdDECO	CLIP-DCC
$N = 10$	AV	14.449	15.109	14.432	14.357
	SD	13.215	13.044	12.896	12.889
	PI		0.308	0.460	0.481
$N = 30$	AV	41.882	46.497	41.326	41.136
	SD	34.642	32.841	32.895	33.283
	PI		0.208	0.495	0.578
$N = 50$	AV	71.932	78.882	69.883	69.905
	SD	51.529	47.931	47.930	48.773
	PI		0.218	0.560	0.661
$N = 100$	AV	153.837	164.419	145.324	146.550
	SD	97.043	86.354	86.832	89.776
	PI		0.279	0.683	0.796
$N = 300$	AV	515.756	517.575	452.134	464.607
	SD	307.922	237.937	253.667	270.666
	PI		0.465	0.892	0.945
$N = 500$	AV	973.402	905.116	823.696	851.357
	SD	550.597	378.756	434.620	464.127
	PI		0.611	0.949	0.974

Note: This table contains the average (AV) and standard deviation (SD) of the daily out-of-sample QLIKE score for the different DCC models. In addition, the proportion of improvement (PI) denotes the share of dates that the model has a lower QLIKE score than the DCC model. The out-of-sample periods ranges from December 1990 until December 2020 for a total of 7560 days, using an estimation window of 2500 days and re-estimation of the parameters every 21 days.

B.4 Robustness estimation length

Appendix Table B.5 and B.6 contain the daily out-of-sample performance of the GMPVs constructed using the different correlation models with an estimation window length of $T = 1250$ and without and with NLS of the target, respectively. Comparing Appendix Table B.5 and B.6 with Table 3 and Table 4, we note that the SD is higher for the DCC, sdDECO and CLIP-DCC models for $T = 1250$ than for $T = 2500$, while the reverse is true for the dDECO model. This suggests that models that have an unrestricted long-run can benefit from a longer estimation sample, which makes sense as the concentration ratio N/T greatly determines the quality of the sample covariance estimator (Ledoit and Wolf, 2004).

Furthermore, our overall conclusions are highly similar here to those when using the longer estimation window $T = 2500$. That is, the dDECO model performs worst, the DCC and sdDECO model perform relatively similar and the CLIP-DCC model performs the best in terms of out-of-sample SD. Comparing Table B.5 and B.6, we also find again that NLS of the target decreases the SD values, such that the CLIP-DCC model with NLS has the lowest SD. In terms of the relative benefits of both methods, we find NLS and the CLIP-DCC model to decrease the SD about the same compared to the DCC model for $N = 100$. For higher values of N the NLS appears to add more value, while for smaller N the CLIP-DCC model is found to be most useful. Because the benefits of both methods are additive (or even slightly synergistic for large N), we recommend using a combined approach.

Table B.5: Daily out-of-sample GMVP performance constructed using different DCC models for $N \in \{10, 30, 50, 100, 300, 500\}$ and $T = 1250$, December 1990 until December 2020.

		DCC	dDECO	sdDECO	CLIP-DCC	1/N
$N = 10$	AV	4.225	3.167	4.746	4.486	7.283
	SD	15.589	15.909	15.592	15.521*	19.167
	IR	0.271	0.199	0.304	0.289	0.380
$N = 30$	AV	5.220	4.279	4.571	5.029	8.041
	SD	14.045	14.796	13.945	13.854***	18.457
	IR	0.372	0.289	0.328	0.363	0.436
$N = 50$	AV	2.891	2.068	1.992	2.560	8.458
	SD	13.374	14.499	13.309	13.148***	18.271
	IR	0.216	0.143	0.150	0.195	0.463
$N = 100$	AV	3.239	0.853	0.978	2.727	8.273
	SD	11.142	13.104	11.172	10.991***	18.225
	IR	0.291	0.065	0.088	0.248	0.454
$N = 300$	AV	5.256	1.241	2.038	4.760	9.034
	SD	8.366	10.989	8.374	8.205***	17.975
	IR	0.628	0.113	0.243	0.580	0.503
$N = 500$	AV	4.937	0.883	1.926	4.505	9.558
	SD	7.165	9.022	7.146	7.005***	17.961
	IR	0.689	0.098	0.269	0.643	0.532

Note: This table contains the annualized average (AV), standard deviation (SD) and information ratio (IR) of the out-of-sample daily log returns for the GMVPs constructed using different dynamic correlation models and the $1/N$ portfolio. The lowest SD per dimension size is highlighted in bold. The out-of-sample periods ranges from December 1990 until December 2020 for a total of 7560 days, using an estimation window of 1250 days and re-estimation of the parameters every 21 days. A significant decrease of the (logarithmic squared) SD of the CLIP-DCC model compared to the DCC model is indicated with a *, ** and *** for a p -value below 0.1, 0.05 and 0.01, respectively, using the two-sided test by Ledoit and Wolf (2011) with HAC standard errors.

Table B.6: Daily out-of-sample GMVP performance constructed using different DCC models for $N \in \{10, 30, 50, 100, 300, 500\}$, $T = 1250$ and NLS of the target, December 1990 until December 2020.

		DCC	dDECO	sdDECO	CLIP-DCC	1/N
$N = 10$	AV	4.227	3.169	4.732	4.477	7.283
	SD	15.583	15.908	15.582	15.513*	19.167
	IR	0.271	0.199	0.304	0.289	0.380
$N = 30$	AV	5.284	4.280	4.736	5.145	8.041
	SD	14.010	14.791	13.897	13.816***	18.457
	IR	0.377	0.289	0.341	0.372	0.436
$N = 50$	AV	2.946	2.071	2.229	2.687	8.458
	SD	13.308	14.495	13.226	13.081***	18.271
	IR	0.221	0.143	0.169	0.205	0.463
$N = 100$	AV	3.183	0.855	1.088	2.737	8.273
	SD	11.027	13.100	11.055	10.881***	18.225
	IR	0.289	0.065	0.098	0.252	0.454
$N = 300$	AV	5.377	1.241	2.474	4.829	9.034
	SD	7.979	10.987	7.977	7.813***	17.975
	IR	0.674	0.113	0.310	0.618	0.503
$N = 500$	AV	4.577	0.883	2.038	4.059	9.558
	SD	6.589	9.021	6.491	6.370***	17.961
	IR	0.695	0.098	0.314	0.637	0.532

Note: This table contains the annualized average (AV), standard deviation (SD) and information ratio (IR) of the out-of-sample daily log returns for the GMVPs constructed using different dynamic correlation models and the $1/N$ portfolio. The lowest SD per dimension size is highlighted in bold. The out-of-sample periods ranges from December 1990 until December 2020 for a total of 7560 days, using an estimation window of 1250 days and re-estimation of the parameters every 21 days. A significant decrease of the (logarithmic squared) SD of the CLIP-DCC model compared to the DCC model is indicated with a *,** and *** for a p -value below 0.1, 0.05 and 0.01, respectively, using the two-sided test by Ledoit and Wolf (2011) with HAC standard errors.

References

- Engle, R. F. and B. Kelly (2012). Dynamic equicorrelation. *Journal of Business & Economic Statistics* 30(2), 212–228.
- Engle, R. F., O. Ledoit, and M. Wolf (2019). Large dynamic covariance matrices. *Journal of Business & Economic Statistics* 37(2), 363–375.
- Ledoit, O. and M. Wolf (2004). Honey, I shrunk the sample covariance matrix. *The Journal of Portfolio Management* 30(4), 110–119.
- Ledoit, O. and M. Wolf (2011). Robust performances hypothesis testing with the variance. *Wilmott* 2011(55), 86–89.
- Ledoit, O. and M. Wolf (2020). Analytical nonlinear shrinkage of large-dimensional covariance matrices. *The Annals of Statistics* 48(5), 3043–3065.
- Oh, D. H. and A. J. Patton (2023). Dynamic factor copula models with estimated cluster assignments. *Journal of Econometrics*, In Press.
- Patton, A. J. and K. Sheppard (2009). Evaluating volatility and correlation forecasts. In T. Andersen, R. Davis, J.-P. Kreiss, and T. Mikosch (Eds.), *Handbook of Financial Time Series*, pp. 801–838. Springer Verlag.
- Roustant, O. and Y. Deville (2017). On the validity of parametric block correlation matrices with constant within and between group correlations. *Working paper, available at <https://arxiv.org/abs/1705.09793>*.