TI 2019-013/IV
Tinbergen Institute Discussion Paper

Closed-Form Multi-Factor Copula Models with Observation-Driven Dynamic Factor Loadings

Revision: October 23, 2019

Anne Opschoor[1]
André Lucas1[1]
Istvan Barra[2]
Dick van Dijk[3]

[1] Vrije Universiteit Amsterdam and Tinbergen Institute
[2] Independent Research
[3] Erasmus University Rotterdam

Closed-Form Multi-Factor Copula Models with Observation-Driven Dynamic Factor Loadings*

Anne Opschoor[a,b], André Lucas[a,b], Istvan Barra[c], Dick van Dijk[d,b]

[a] *Vrije Universiteit Amsterdam*

[b] *Tinbergen Institute*

[c] *Independent research*

[d] *Erasmus University Rotterdam*

This version: October 23, 2019

Abstract

We develop new multi-factor dynamic copula models with time-varying factor loadings and observation-driven dynamics. The new models are highly flexible, scalable to high dimensions, and ensure positivity of covariance and correlation matrices. A closed-form likelihood expression allows for straightforward parameter estimation and likelihood inference. We apply the new model to a large panel of 100 U.S. stocks over the period 2001–2014. The proposed multi-factor structure is much better than existing (single-factor) models at describing stock return dependence dynamics in high-dimensions. The new factor models also improve one-step-ahead copula density forecasts and global minimum variance portfolio performance. Finally, we investigate different mechanisms to allocate firms into groups and find that a simple industry classification outperforms alternatives based on observable risk factors, such as size, value or momentum.

Keywords: factor copulas, factor structure, score-driven dynamics, multivariate density forecast.

JEL: C32, C58, G17.

* *Corresponding author*: Anne Opschoor, Vrije Universiteit Amsterdam, De Boelelaan 1105, 1081 HV, Amsterdam, The Netherlands. *E-mail*: `a.opschoor@vu.nl`. *Phone*: +31(0)20-5982663.

1 Introduction

Copulas are a key ingredient in many current applications in economics and finance (see for example Patton, 2009; Cherubini et al., 2011; Fan and Patton, 2014; McNeil et al., 2015). In particular, time-varying copulas have turned out to be an important and flexible tool to describe dependence dynamics in an unstable environment (see Patton, 2006; Manner and Reznikova, 2012; Lucas et al., 2014). Most copula applications deal with a cross-sectional dimension that is small to moderate (for an overview, see Patton, 2013). Applications to high-dimensional data sets are scarce, mainly due to the 'curse of dimensionality': the number of parameters grows rapidly when the dimension increases.

Recently, Creal and Tsay (2015), Oh and Patton (2017, 2018), and Lucas et al. (2017) put forward a general approach to modeling time-varying dependence in high cross-sectional dimensions using a factor copula structure. The factor copula structure describes the dependence between a large number of observed variables by a smaller set of latent variables (or factors) with time-varying loadings. This allows one to considerably limit the number of parameters required to flexibly describe the dynamics of high-dimensional dependence structures.

Dynamic factor copulas have mainly been implemented for the single-factor case; see the references above. This is predominantly driven by computational reasons. Though adding more factors with dynamic loadings is in principle possible, it would increase the computational burden substantially. In the approach of Oh and Patton (2018) this results from the fact that the densities of the common latent factors and of the idiosyncratic factors do not convolute easily. The copula density is then not available in closed form and additional numerical methods are required for estimation. This requires considerable computational effort, particularly if multiple factors are used. Creal and Tsay (2015) face a different challenge as they use a standard parameter driven recurrence equation for the factor loading dynamics. This introduces additional stochastic components into the model that need to be integrated out (see also Hafner and Manner, 2012). Bayesian simulation techniques are used for this integration step, which again becomes computationally expensive as the number of factors with dynamic loadings grows.

Though restricting the number of factors to just a single one is understandable from a computational point of view, it seems too restrictive for most empirical applications. For instance, when modeling panels of equity returns a minimum of three to five factors seems to be the standard (see Fama and French, 1993, 2016). A computationally simple yet flexible approach that can easily deal with both the multi-factor setting and dynamic loadings thus seems to be called for.

In this paper we develop exactly such a multi-factor copula model with dynamic loadings. For this purpose, we assume that the cross-sectional units can be grouped using observable characteristics, such as the industry of the firm, its headquarters location, or risk characteristics such as firm size, its book-to-market value, etc. Each of these groups is possibly subject to one or more common factors as well as to group-specific factors. We keep the number of parameters in the model limited by assuming that all units in a particular group have identical factor loadings. We allow the loadings for each of the factors to vary over time using score-driven dynamics as introduced by Creal et al. (2013) and Harvey (2013). Using appropriate distributional assumptions for the latent common and group-specific factors as well as for the idiosyncratic components, we obtain a model with a tractable, closed-from likelihood expression. Hence, parameter estimation and inference are straightforward using Maximum Likelihood (ML) and the computational burden is kept to a minimum. In particular, a two-step targeting approach that combines a moment-based estimator and the ML approach leads to fast estimation of the parameters in our most flexible multi-factor copula model. The new multi-factor model can easily be implemented for high dimensions. In addition, the model easily allows for the inclusion of exogenous variables that help to describe the dynamics of the factor loadings.

As a typical high-dimensional financial data set, we consider a panel of 100 U.S. daily equity returns across 10 different industries over the period 2001–2014. We group the stocks according to industry, and consider various single- and multi-factor specifications, with Gaussian and Student's t copulas. We compare the factor copula models with three popular multivariate

GARCH (MGARCH) models: the cDCC model of Engle (2002) and Aielli (2013), and the DECO and Block DECO models of Engle and Kelly (2012). Our comparison is based on in-sample and out-of-sample (density) forecasts, using the Model Confidence Set approach of Hansen et al. (2011). In addition, we consider the economic performance of the models when used for constructing Global Minimum Variance Portfolios (GMVPs).

We find that for our panel of equity returns, both within-industry and between-industry dependence dynamics are key data features that need to be accommodated. Single-factor models and the standard DECO model have difficulty matching all these within- and between-industry dynamics simultaneously. Our multi-factor specification with Student's t copula, by contrast, outperforms all benchmarks considered in terms of density forecasts, both in-sample and one-step-ahead out-of-sample. When considering the joint lower tail of the multivariate distribution, we again find that the multi-factor Student's t copula model always belongs to the Model Confidence Set.

For economic criteria, simpler models prevail, though the multi-factor model still belongs to the Model Confidence Set. Meanwhile, our one-factor specification with heterogeneous dynamic loadings has the best ex-post variance of the GMVP. We attribute this difference to the character of the global *minimum* variance criterion: differences in *minimum* variance are harder to obtain and typically smaller, such that the increased flexibility of more complex models does not offset the associated estimation risk of the additional parameters used. This contrasts with the criterion based on the full density forecasts, where all dynamics play a more dominant role and the multi-factor specifications work best in-sample and out-of-sample.

As a final novelty in this paper, we investigate whether industry classification provides the best grouping structure. As alternatives, we consider other classifications based on observable risk characteristics such as firm size, value, or momentum. This provides a further modeling challenge, as the group structure is allowed to vary over time, with corresponding changes in the factor loadings matrix. We find that group classifications based on observable risk characteristics do not outperform the simpler, static classification based on industry.

This paper relates to various strands of the literature. First, there is an extensive literature on factor models and the computation of large covariance matrices, see for example Fan et al. (2008, 2011, 2016). Engle et al. (1990) develop factor ARCH models with an application to asset pricing with many assets. However, the benefit of factor copulas is the flexibility in choosing the factor structure and distributional assumptions, both with respect to the marginals and the copula structure. Second, factor copulas have recently been introduced by Krupskii and Joe (2013); Oh and Patton (2017), among others. Oh and Patton (2018) and Lucas et al. (2017) are the first to introduce the score-driven framework of Creal et al. (2013) within factor copulas. Compared to their work, we consider specifications that yield closed-form densities and use a parametrization that is easily scalable to many factors and high cross-sectional dimensions. Third, we relate to a strand of literature on Copula-MGARCH models, such as Christoffersen et al. (2012, 2014), who combine a skewed Student's t copula with a DCC model to study diversification benefits in a panel of more than 200 asset returns. These models suffer in general from the curse of dimensionality mentioned earlier. In addition, (large) covariance or correlation matrices need to be inverted many times during parameter estimation, which becomes computationally cumbersome and numerically problematic.

The rest of this paper is organized as follows. Section 2 presents the multi-factor copula model with dynamic loadings. We carefully lay out the different aspects of our modeling approach, including various possible common factor specifications and the loading dynamics. We also discuss important details concerning parameter estimation, using either full likelihood, a two-step targeting approach, or composite likelihood methods. Section 3 studies the performance of the multi-factor copula models in a controlled environment. Section 4 provides the results for the empirical application. Section 5 concludes. An online appendix to this paper provides more details on some of the derivations, as well as more empirical and simulation results for the new models.

2 The modeling framework

In this section, we develop the class of closed-form dynamic multi-factor copulas with score-driven loadings, as an approach for describing time-varying dependence that remains manageable yet versatile in high-dimensional settings. Our aim is to characterize the conditional joint distribution $\boldsymbol{F}_t(\boldsymbol{y}_t)$ of the vector $\boldsymbol{y}_t = (y_{1,t}, \ldots, y_{N,t})^\top \in \mathbb{R}^N$ of asset returns in period t, $t = 1, \ldots, T$, where N is possibly large. We decompose $\boldsymbol{F}_t(\boldsymbol{y}_t)$ into N marginals and a conditional copula as in Patton (2006),

$$\boldsymbol{y}_t|\mathcal{F}_{t-1} \sim \boldsymbol{F}_t(\boldsymbol{y}_t) = \boldsymbol{C}_t\Big(F_{1,t}(y_{1,t};\boldsymbol{\theta}_{M,1,t}) \ , \ \ldots \ , \ F_{N,t}(y_{N,t};\boldsymbol{\theta}_{M,N,t}) \ ; \ \boldsymbol{\theta}_{C,t}\Big), \tag{1}$$

where $\boldsymbol{C}_t(\,\cdot\,;\boldsymbol{\theta}_{C,t})$ is the conditional copula given the information set $\mathcal{F}_{t-1} = \sigma(\boldsymbol{y}_{t-1}, \boldsymbol{y}_{t-2}, \ldots)$ and the time-varying copula parameter vector $\boldsymbol{\theta}_{C,t}$, and $F_{i,t}(y_{i,t};\boldsymbol{\theta}_{M,i,t})$, $i = 1, \ldots, N$, denotes the conditional marginal distribution of asset i given $\mathcal{F}_{t-1}$ and the time-varying marginal distribution parameter vector $\boldsymbol{\theta}_{M,i,t}$. We return to the choice of the marginals later. Note that the conditional copula $\boldsymbol{C}_t(\,\cdot\,;\boldsymbol{\theta}_{C,t})$ can also be interpreted as the conditional distribution $\boldsymbol{C}_t(\boldsymbol{u}_t;\boldsymbol{\theta}_{C,t})$ of the probability integral transforms (PITs) $\boldsymbol{u}_t = (u_{1,t}, \ldots, u_{N,t})^\top$ of $\boldsymbol{y}_t$, where $u_{i,t} \equiv F_{i,t}(y_{i,t};\boldsymbol{\theta}_{M,i,t})$ for $i = 1, \ldots, N$.

As is well known, decomposing the multivariate (conditional) distribution $\boldsymbol{F}_t(\boldsymbol{y}_t)$ into its marginals and copula has several advantages. Particularly when the cross-sectional dimension N is large, splitting the modeling task into specifying the marginals and the copula may substantially reduce the computational burden as parameters can be estimated using a two-step approach. As modeling the univariate marginal distributions is relatively simple and fast even for large N, the main remaining challenge is to parsimoniously specify the conditional copula $\boldsymbol{C}_t(\,\cdot\,;\boldsymbol{\theta}_{C,t})$. This can be done using factor copulas or multivariate GARCH models like the DCC or DECO models.

2.1 Observation-driven dynamic factor copulas

The general literature on copula modeling is extensive; see for instance Patton (2009, 2013) or Fan and Patton (2014) for partial overviews. However, the literature on how to deal with copulas in large cross-sectional dimensions is rather scarce. The main challenge in high dimensions is to keep the parameter space manageable, but at the same time to allow for sufficient flexibility in the dependence structure. To strike this balance, we use a multi-factor copula structure that we endow with score-driven parameter dynamics. Furthermore, we assume that the N asset returns can be clustered into G groups, with assets in the same group having identical factor loadings.

We start from the factor copula structure

$$\begin{aligned} u_{i,t} &= D_{\boldsymbol{x},i}(x_{i,t}; \tilde{\boldsymbol{\lambda}}_{i,t}, \sigma_{i,t}, \boldsymbol{\psi}_C), \qquad i = 1, \ldots N, \\ x_{i,t} &= \tilde{\boldsymbol{\lambda}}_{i,t}^{\top} \boldsymbol{z}_t + \sigma_{i,t}\epsilon_{i,t}, \\ \boldsymbol{z}_t &\overset{\text{i.i.d.}}{\sim} D_{\boldsymbol{z}}(\boldsymbol{z}_t \mid \boldsymbol{\psi}_C), \qquad \epsilon_{i,t} \overset{\text{i.i.d.}}{\sim} D_{\epsilon}(\epsilon_{i,t} \mid \boldsymbol{\psi}_C), \end{aligned} \tag{2}$$

where $\tilde{\boldsymbol{\lambda}}_{i,t}$ is a $k \times 1$ vector of scaled factor loadings, $\boldsymbol{z}_t$ is a $k \times 1$ vector of common latent factors, $\epsilon_{i,t}$ is an idiosyncratic shock, $\boldsymbol{z}_t$ and $\epsilon_{i,t}$ are cross-sectionally and serially independent with distributions $D_{\boldsymbol{z}}$ and D_{ϵ}, respectively, characterized by zero means and identity covariance matrix and static shape parameter vector $\boldsymbol{\psi}_C$, and $D_{\boldsymbol{x},i}(\,\cdot\,)$ denotes the implied marginal distribution of $x_{i,t}$; see Creal and Tsay (2015). We define the vector $\tilde{\boldsymbol{\lambda}}_{i,t}$ and scalar $\sigma_{i,t}$ as

$$\tilde{\boldsymbol{\lambda}}_{i,t} = \frac{\boldsymbol{\lambda}_{i,t}}{\sqrt{1 + \boldsymbol{\lambda}_{i,t}^{\top}\boldsymbol{\lambda}_{i,t}}}, \qquad \sigma_{i,t}^2 = \frac{1}{1 + \boldsymbol{\lambda}_{i,t}^{\top}\boldsymbol{\lambda}_{i,t}} \tag{3}$$

for an unrestricted $k \times 1$ vector $\boldsymbol{\lambda}_{i,t}$, such that $x_{i,t}$ has zero mean and unit variance by design. Further parameterization details can be added to ensure for instance that some elements of $\boldsymbol{\lambda}_{i,t}$ are positive by design, e.g., by taking exponential functions or a multinomial logit parameterization. Further details can be found in the online appendix.

The correlation matrix of $\boldsymbol{x}_t = (x_{1,t}, \ldots, x_{N,t})^\top$ equals

$$\boldsymbol{R}_t = \tilde{\boldsymbol{L}}_t^\top \tilde{\boldsymbol{L}}_t + \boldsymbol{D}_t, \qquad \tilde{\boldsymbol{L}}_t = \left(\tilde{\boldsymbol{\lambda}}_{1,t}, \ldots, \tilde{\boldsymbol{\lambda}}_{N,t}\right), \qquad \boldsymbol{D}_t = \operatorname{diag}\left(\sigma_{1,t}^2, \ldots, \sigma_{N,t}^2\right), \tag{4}$$

which satisfies all requirements of a correlation matrix, namely positive semi-definiteness and ones on the diagonal. The copula parameter vector gathers all free parameters in $\boldsymbol{\theta}_{C,t}^\top = (\boldsymbol{\lambda}_{1,t}^\top, \ldots, \boldsymbol{\lambda}_{N,t}^\top, \boldsymbol{\psi}_C^\top)$.

The factor copula structure in (2) comes with an important computational advantage, namely that the inverse and determinant of $\boldsymbol{R}_t$ are available in closed form as

$$\boldsymbol{R}_t^{-1} = \boldsymbol{D}_t^{-1} - \boldsymbol{D}_t^{-1}\tilde{\boldsymbol{L}}_t^\top \left(\mathbf{I}_k + \tilde{\boldsymbol{L}}_t \boldsymbol{D}_t^{-1} \tilde{\boldsymbol{L}}_t^\top\right)^{-1} \tilde{\boldsymbol{L}}_t \boldsymbol{D}_t^{-1}, \qquad |\boldsymbol{R}_t| = \left|\mathbf{I}_k + \tilde{\boldsymbol{L}}_t \boldsymbol{D}_t^{-1} \tilde{\boldsymbol{L}}_t^\top\right| \cdot |\boldsymbol{D}_t|, \tag{5}$$

where $\mathbf{I}_k$ denotes the k-dimensional identity matrix. As the number of common latent factors k is typically much smaller than the number of observed assets N, computing the inverse of the $k \times k$ matrix $\mathbf{I}_k + \tilde{\boldsymbol{L}}_t \boldsymbol{D}_t^{-1} \tilde{\boldsymbol{L}}_t^\top$ is much faster than computing the inverse of the $N \times N$ matrix $\boldsymbol{R}_t$.

The class of factor copulas is very flexible. We can vary the number and types of factors, the distributional assumptions of the common factors $\boldsymbol{z}_t$ and idiosyncratic shocks $\epsilon_{i,t}$, and the dynamics of the factor loadings $\boldsymbol{\lambda}_{i,t}$. The following subsections discuss each of these choices in more detail.

2.1.1 The factor structure

Our main goal in this paper is to develop feasible dependence structures that allow for multiple factors in a flexible, dynamic way while still giving rise to a closed-form likelihood expression. With our focus on multiple factors, we extend earlier papers that focus on single-factor implementations, such as Oh and Patton (2018) and Creal and Tsay (2015).

A key aspect of our approach is the assumption that we can split the N assets into G groups according to an observed characteristic such as industry, region, or riskiness, etc. Each group may be subject to several factors, where the factor loadings are assumed to be identical within each group. Resembling the Block DECO model of Engle and Kelly (2012), this implies that all assets within a group share the same dependence structure, while the dependence between any two assets in two different groups is also the same (but still varying across the possible combinations of groups). This yields a flexible, yet highly parsimonious set-up.

For the sake of exposition, we take the example of $G = 4$ groups with 2 firms in each group throughout this subsection. In reality, of course, the number of groups and the number of firms per group is typically much larger. For instance, in our application in Section 4 we have $G = 10$ groups with up to 19 firms per group.

In our most general specification the loading matrix is obtained from a lower-triangular matrix with columns containing group-specific loadings. The loading matrix then takes the form

$$\tilde{\boldsymbol{L}}_t^\top = \begin{pmatrix} \tilde{\lambda}_{1,1,t} & 0 & 0 & 0 \\ \tilde{\lambda}_{1,2,t} & \tilde{\lambda}_{2,2,t} & 0 & 0 \\ \tilde{\lambda}_{1,3,t} & \tilde{\lambda}_{2,3,t} & \tilde{\lambda}_{3,3,t} & 0 \\ \tilde{\lambda}_{1,4,t} & \tilde{\lambda}_{2,4,t} & \tilde{\lambda}_{3,4,t} & \tilde{\lambda}_{4,4,t} \end{pmatrix} \otimes \begin{pmatrix} 1 \\ 1 \end{pmatrix}, \tag{6}$$

where $\otimes$ denotes the Kronecker product. The first column vector can be interpreted as a common-factor with group-specific loadings, like different market betas. Overall, the loading matrix could also be seen as a Cholesky decomposition of the quasi correlation matrix that contains within and between group correlations. Note that the Cholesky decomposition could be sensitive to the different ordering of the groups. We show in Section 4 that in our empirical application this effect is small: estimated dependence measures hardly change when we reorder the variables. We label the model with the factor structure in (6) as the Multi-Factor Lower-Triangular (MF LT) copula model.

A second, much more restricted version of our general specification combines a single common factor with (common) equi-loadings, and a set of G group-specific factors with corresponding group-specific loadings. This results in the loading matrix

$$\tilde{\boldsymbol{L}}_t^\top = \begin{pmatrix} \tilde{\lambda}_{1,t} & \tilde{\lambda}_{2,1,t} & 0 & 0 & 0 \\ \tilde{\lambda}_{1,t} & 0 & \tilde{\lambda}_{2,2,t} & 0 & 0 \\ \tilde{\lambda}_{1,t} & 0 & 0 & \tilde{\lambda}_{2,3,t} & 0 \\ \tilde{\lambda}_{1,t} & 0 & 0 & 0 & \tilde{\lambda}_{2,4,t} \end{pmatrix} \otimes \begin{pmatrix} 1 \\ 1 \end{pmatrix}. \tag{7}$$

For $G \geq 3$ and at least 2 firms in each group, this model meets the necessary requirement for identification. To see this, note that the correlation matrix $\boldsymbol{R}_t$ for $G = 3$ has 3 within-group correlations and 3 between-group correlations, hence 6 free positions for the 4 different parameters in $\tilde{\boldsymbol{L}}_t$. For more groups and firms, the number of positions in $\boldsymbol{R}_t$ increases quadratically, whereas the number of parameters in $\tilde{\boldsymbol{L}}_t$ increases linearly, thus allowing for overidentification. The first (equi)factor with common loadings $\tilde{\lambda}_{1,t}$ affects both the within-group and the between-group correlations. The group-specific factors with their group-specific loadings, on the other hand, only affect the within-group correlations and not the between-group correlations. We label this model as the Multi-Factor (MF) copula model.

A third specification is obtained by replacing the group-specific factors in (7) with a common factor with group-specific loadings. The loading matrix $\tilde{\boldsymbol{L}}_t^\top$ is then given by

$$\tilde{\boldsymbol{L}}_t^\top = \begin{pmatrix} \tilde{\lambda}_{1,t} & \tilde{\lambda}_{2,1,t} \\ \tilde{\lambda}_{1,t} & \tilde{\lambda}_{2,2,t} \\ \tilde{\lambda}_{1,t} & \tilde{\lambda}_{2,3,t} \\ \tilde{\lambda}_{1,t} & \tilde{\lambda}_{2,4,t} \end{pmatrix} \otimes \begin{pmatrix} 1 \\ 1 \end{pmatrix}. \tag{8}$$

From an asset pricing point of view, this second common factor has different betas for each group. There is now less freedom, however, to capture the differences between within-group and between-group effects as $\tilde{\lambda}_{2,g,t}$ and $\tilde{\lambda}_{2,h,t}$ determine both, for $g, h = 1, \ldots, G$. Note that $\tilde{\lambda}_{2,1,t}$ cannot be rotated to zero without destroying the equi-loading structure of the first column of $\tilde{\boldsymbol{L}}_t^\top$. This underlines that the model is identified. We label the model in equation (8) the 2-Factor (2F) copula model. Omitting the factor corresponding to $\tilde{\lambda}_{1,t}$ in (8) leads

Table 1: Various factor structures and their properties

This table summarizes the various factor structures that are proposed given that there are N assets allocated to G different groups. We show the number of factors, the number of different scaled factor loadings, the dimension of the scaled factor loading matrix and the existence of an equi-factor, group-specific factors and/or group-specific loadings.

Name	# factors	# λs	common factor with (equi) common loading	common factor with group loadings	group factors with group loadings	dim $\tilde{\boldsymbol{L}}_t^\top$
1F-Equi	1	1	yes	no	no	$N \times 1$
1F-Group	1	G	no	yes	no	$N \times 1$
2F	2	$G+1$	yes	yes	no	$N \times 2$
MF	$G+1$	$G+1$	yes	no	yes	$N \times (G+1)$
MF-LT	G	$G(G+1)/2$	no	yes	yes	$N \times G$

to the 1-Factor-Group (1F-Gr) model, which consists of a single factor but with G different group loadings. The 1F-Gr model has also been used in Lucas et al. (2017) and Oh and Patton (2018). Similarly, if instead we drop the factor corresponding to $\tilde{\lambda}_{2,g,t}$ in (8), we obtain a single-factor model with common loadings. We label this special case the 1F-Equi copula model; see also the the single-factor copula structures of Oh and Patton (2018) and Creal and Tsay (2015). It corresponds to a DECO correlation structure as in Engle and Kelly (2012), where each pairwise asset correlation is assumed to be the same. From an asset pricing perspective, the single factor can be seen as the market factor, with an identical beta for all assets.

Table 1 lists all the factor structures considered in this paper with their corresponding properties, such as the number of factors, the number of different $\tilde{\lambda}$s, and the associated dimension of $\tilde{\boldsymbol{L}}_t^\top$.

2.1.2 Distributional assumptions

Given the various factor structures proposed in Section 2.1.1, the next step is to specify a distribution for the common, group-specific, and idiosyncratic factors in (2). Oh and Patton (2018) assume a skewed and symmetric Student's t density for the common factor z_t and the idiosyncratic shock $\epsilon_{i,t}$, respectively. As a result, their copula density for $x_{i,t}$ is not available in closed form. Hence, likelihood evaluation and parameter estimation become computationally involved. Also Creal and Tsay (2015) do not have a likelihood in closed form due to their choice of a new stochastic component in the transition equation for the factor loading $\lambda_{i,t}$. They solve the issue by employing Bayesian (numerical) techniques to estimate the parameters. Again, this is computationally costly for increasing dimensions, particularly in multi-factor settings.

In contrast to the above approaches, we retain tractability of the model and a closed form of the likelihood by two particular choices. First, we make convenient distributional assumptions for the factors $\boldsymbol{z}_t$ and $\epsilon_{i,t}$. Second, we consider a score-driven transition equation for the factor loadings $\lambda_{i,t}$. We discuss the latter in the next subsection.

To model $\boldsymbol{z}_t$ and $\epsilon_{i,t}$, we use the Student's t copula,

$$u_{i,t} = T(x_{i,t}; \nu_C), \qquad i = 1, \dots N, \tag{9a}$$

$$x_{i,t} = \sqrt{\zeta_t}\left(\tilde{\boldsymbol{\lambda}}_{i,t}^{\top}\boldsymbol{z}_t + \sigma_{i,t}\epsilon_{i,t}\right), \qquad \boldsymbol{z}_t \sim \mathrm{N}(\boldsymbol{0}, \mathbf{I}_k), \quad \epsilon_{i,t} \sim \mathrm{N}(0,1), \tag{9b}$$

$$\zeta_t \sim \text{Inv-Gamma}\left(\nu_C/2, \nu_C/2\right). \tag{9c}$$

where $T(\,\cdot\,; \nu_C)$ denotes the cdf of the univariate Student's t distribution with ν_C degrees of freedom, location zero, and unit scale, and ζ_t denotes an independent Inverse-Gamma distributed random variable. Note that – in contrast to Creal and Tsay (2015) and Oh and Patton (2018) – our proposed factor structures of the previous subsection easily fit into the distributional framework above, while the copula density (and thus the likelihood) retains its analytical closed-form expression. For the special case $\nu_C \to \infty$, we obtain $\zeta_t \equiv 1$ and a Gaussian copula setting. The Gaussian copula, however, has no tail dependence (see McNeil et al., 2015) and may therefore be less suitable to describe the dependence structure in empirical applications.

The copula in (9a)–(9c) is symmetric. Oh and Patton (2017, 2018), by contrast, develop a 1-factor asymmetric copula model that allows for skewness. By adding an additional term $\boldsymbol{\gamma}\zeta_t$ for $\boldsymbol{\gamma} \in \mathbb{R}^N$ to the right-hand side of (9b) and letting ζ_t be Generalized Inverse Gaussian, we obtain the Generalized Hyperbolic (GH) copula class with skewness parameter $\boldsymbol{\gamma}$. A special case is the GH skewed Student's t copula as used in for instance Lucas et al. (2014, 2017). Such a generalization would come at a substantial increase in computational burden as the copula requires the numerical inversion of the marginal cdfs at each point in time for all coordinates. Preliminary experiments for the simplest model structures and a skewed t copula did not result in major in-sample likelihood increases or in substantial changes in the paths of the fitted dynamic dependence parameters. Therefore, we leave such further generalizations for future research and in this paper concentrate on the value-added of the multi-factor structures.

2.1.3 Score-driven factor loading dynamics

To complete our dynamic factor copula specification, we formulate the dynamics of the factor loadings $\lambda_{i,t}$. In general, there are two approaches to modeling time-varying factor loadings. The first approach is parameter-driven and assumes $\lambda_{i,t}$ evolves as a stochastic process driven by its own innovation. This leads to so-called stochastic copula models as in Hafner and Manner (2012) and Creal and Tsay (2015). Estimating such models is typically more involved and requires integrating out the random innovations of the time-varying parameters in a numerically efficient way. The second approach is observation-driven and assumes the factor loadings depend on functions of past observables. Our proposal falls into the latter category and uses score-driven dynamics as introduced by Creal et al. (2013); see also Harvey (2013) and Oh and Patton (2018). As mentioned before, an advantage of the observation-driven approach is that the likelihood is available in closed-form via a standard prediction error decomposition. This substantially reduces the computational burden compared to a parameter-driven approach.

Score-driven dynamics use the score of the conditional copula density to drive $\lambda_{i,t}$. Intuitively, this adjust the loadings in a steepest ascent direction of the local log likelihood fit at time t. See the information theoretic optimality motivation for this approach in Blasques et al. (2015) and its generalizations in Creal et al. (2018). As an example in our context, consider a 1-Factor equicorrelation copula, such that $\tilde{\boldsymbol{L}}_t^\top = \tilde{\lambda}_t \boldsymbol{\iota}_N$ for a scalar parameter $\tilde{\lambda}_t = \lambda_t / \sqrt{1+\lambda_t^2}$, such that $\tilde{\lambda}_t \in [-1, 1]$ by design, where $\boldsymbol{\iota}_N$ denotes an $N \times 1$ vector filled with ones. Then the score-driven dynamics for λ_t are given by $\lambda_{t+1} = \omega + A\, s_t + B\, \lambda_t$, with $s_t = \partial \log \boldsymbol{c}(\boldsymbol{x}_t; \lambda_t, \nu_C)/\partial \lambda_t$, and $\boldsymbol{c}(\,\cdot\,; \lambda_t, \nu_C)$ the Student's t copula density. We assume the same type of factor loading dynamics in case λ_{t+1} is a vector, where we only allow the intercept

ω to be element-specific (collecting these in the vector $\boldsymbol{\omega}$) but continue to assume that A and B are scalars. Extensions to non-scalar A or B are straightforward, and some of these are investigated in the empirical application later on. Following Oh and Patton (2018), we use unit scaling for the score s_t in the sense of Creal et al. (2013) in order to reduce the computational burden of estimating a separate scaling function. Explicit expressions for the scores for all factor copula specifications used in our paper are provided in the online Appendix A.

2.2 Benchmarks, marginals, and parameter estimation

We compare the dynamic factor copula models developed above against MGARCH alternatives, in particular the cDCC model (Engle, 2002; with the correction of Aielli, 2013) and the (Block) DECO model of Engle and Kelly (2012); see online Appendix C for the implementation details of these models in our setting. To maintain a fair comparison in high dimensions, we also consider the MGARCH models in a copula framework and use the same marginal models for the MGARCH and multi-factor copulas.

For the marginal distributions, we use the univariate t-GAS volatility model of Creal et al. (2011, 2013). We also perform a robustness check with other marginals, such as univariate GARCH models with skewed t innovations. The results are qualitatively similar. For more details on the estimation results for the marginal distributions or for the copula results based on PITs from skewed marginal distributions, see online Appendices D and E, respectively.

Parameter estimation requires some further details, both for the factor copula and the MGARCH copula models. To estimate the model parameters, we use a two-step likelihood based approach. First, we estimate the parameters of each of the marginals (separately). Second, we estimate the copula parameters conditional on the marginal parameter estimates. This approach follows directly from decomposing the joint likelihood as

$$\begin{aligned}\mathcal{L}(\boldsymbol{\theta}) \equiv& \sum_{t=1}^{T} \log \boldsymbol{f}_t(\boldsymbol{y}_t; \boldsymbol{\theta}_t) = \sum_{i=1}^{N} \sum_{t=1}^{T} \log f_{i,t}(y_{i,t}; \boldsymbol{\theta}_{M,t}) \\ &+ \sum_{t=1}^{T} \log \boldsymbol{c}_t \left(F_{1,t}(y_{1,t}; \boldsymbol{\theta}_{M,1,t}), \ldots, F_{N,t}(y_{N,t}; \boldsymbol{\theta}_{M,N,t}); \boldsymbol{\theta}_{C,t} \right)\end{aligned} \tag{10}$$

with $f_{i,t}(\cdot;\boldsymbol{\theta}_{M,i,t})$ denoting the conditional marginal density corresponding to $F_{1,t}(\cdot;\boldsymbol{\theta}_{M,i,t})$, and $\boldsymbol{\theta}_t = \{\boldsymbol{\theta}_{M,t}, \boldsymbol{\theta}_{C,t}\}$. According to Patton (2013), the implied efficiency loss of the two-step compared to the one-step approach is small.

We assume a Student's t and Gaussian copula to model dependence, as discussed before. For the factor copula specifications, inverses and determinants of $\boldsymbol{R}_t$ are given in closed form by (5), which substantially reduces the computational burden in high dimensions. This enables us to estimate these models by full maximum likelihood.

In case of our most general multi-factor copula model (the MF-LT), we potentially have $G(G+1)/2$ different elements in the vector of intercepts $\boldsymbol{\omega}$ in the score-driven dynamics of the factor loadings. A computational challenge may then arise if G becomes large. In that case, we suggest to estimate the copula parameters using the following two-step procedure. Assuming that the loading process is covariance stationary, and defining the unconditional mean of $\boldsymbol{\lambda}_t$ as $\bar{\boldsymbol{\lambda}}$, we have

$$\mathbb{E}[\boldsymbol{\lambda}_{t+1}] \;=\; \boldsymbol{\omega} + B\,\mathbb{E}[\boldsymbol{\lambda}_t] \quad \Leftrightarrow \quad \bar{\boldsymbol{\lambda}} = (1-B)^{-1}\,\boldsymbol{\omega}. \tag{11}$$

In the first step, we match $\bar{\boldsymbol{\lambda}}$ with the empirical within- and between-group correlations using a moment estimator. Let $\boldsymbol{R}_M$ denote the $G \times G$ quasi unconditional correlation matrix based on $x_{it} = \Phi^{-1}(u_{it})$, where the off-diagonal element (g,h) equals the average correlation between assets from group g and h, and the diagonal element (g,g) the asset correlations within group g, $g,h = 1,\ldots,G$, $g \neq h$. The moment estimator is obtained by minimizing

$$L_M = \operatorname{vech}(\boldsymbol{R}_M - \bar{\tilde{\boldsymbol{L}}}\bar{\tilde{\boldsymbol{L}}}^\top)^\top \operatorname{vech}(\boldsymbol{R}_M - \bar{\tilde{\boldsymbol{L}}}\bar{\tilde{\boldsymbol{L}}}^\top) \tag{12}$$

with $\bar{\tilde{\boldsymbol{L}}}^\top$ a $G \times G$ lower triangular matrix as in (6) depending on $\bar{\boldsymbol{\lambda}}$ in (11) via the same non-linear transformation that maps $\boldsymbol{L}_t = \boldsymbol{L}(\boldsymbol{\lambda}_t)$ into $\tilde{\boldsymbol{L}}_t = \tilde{\boldsymbol{L}}(\tilde{\boldsymbol{\lambda}}_t)$. In a second step, we estimate the remaining parameters A and B keeping $\bar{\boldsymbol{\lambda}}$ fixed and setting $\boldsymbol{\omega} = (1-B)\bar{\boldsymbol{\lambda}}$. This two-step targeting procedure substantially decreases the computational burden: The moment estimator in the first step is very fast, while in the second step only the two remaining parameters A

and B need to be estimated. Note that we use the inverse normal in the first step. For the Student's t copula, we could use the inverse Student's t cdf, but we show in the next section that the normal inverse cdf also works well for the Student's t copula case.

In contrast to the multi-factor models, inverses and determinants of $\boldsymbol{R}_t$ are not available in closed form for the Block DECO and cDCC specifications. We therefore estimate the cDCC model by means of the Composite Likelihood method of Engle et al. (2008). This technique is based on maximizing the sum of bivariate (copula) log-likelihood values to estimate A and B (and ν_C). In a second step the matrix $\boldsymbol{\Omega}$ is estimated by its sample analogue.

Finally, we also use a composite likelihood approach for the Block DECO model of Engle and Kelly (2012) by extending their proposal from the Gaussian to the Student's t case. They consider the joint log-likelihood of all the firms in two separate groups $i \neq j$, with $i, j \in \{1 \dots, G\}$, i.e.,

$$\mathcal{L}_{i,j}^{Stud} = \sum_{t=1}^{T} \left[-\frac{1}{2} \log |\boldsymbol{R}_t| - \frac{\nu + n_i + n_j}{2} \log \left(1 + \frac{\boldsymbol{x}_t^\top \boldsymbol{R}_t^{-1} \boldsymbol{x}_t}{\nu - 2} \right) \right], \tag{13}$$

where $|\boldsymbol{R}_t|$ and $\boldsymbol{R}_t^{-1}$ are given analytically for the 2-block case by Lemma 3.1 in Engle and Kelly (2012). The Composite Likelihood (CL) method now maximizes the sum of all log-likelihoods of each pair of blocks $i > j$,

$$\max \mathcal{L}_{CL} = \max \sum_{i>j} \mathcal{L}_{i,j}^{Stud}, \tag{14}$$

where the intercept $\boldsymbol{\Omega}$ is estimated by the unconditional correlation matrix of $\boldsymbol{x}_t$. Note that for $\nu \to \infty$, we recover the Gaussian Block DECO model, which is the specification used in most of the literature. As argued before, however, the Gaussian copula lacks tail dependence and may therefore be less suitable for fitting financial data.

3 Simulation experiment

We briefly report the results of three Monte Carlo experiments, conducted to study the properties of the new method. Full details can be found in the online Appendix B.2.

In the first experiment, we investigate the accuracy of estimation and inference in the new model. Panel A of Table 2 presents the outcomes for a set-up with an $N = 100$ dimensional time series of length $T = 1,000$ with $G = 10$ equally sized groups holding $N/G = 10$ individual cross-sectional units each. These sizes roughly correspond to the data dimensions in our empirical application. The data-generating process (DGP) is the Multi-Factor copula model from equation (7). We only report results for A, B, and ν_C. Results for ω and for smaller sample sizes can be found in the online Appendix and are qualitatively similar. We find that all parameters are estimated near their true values. Comparing results over sample sizes (in the online Appendix), we see that the standard deviation decreases in T. By comparing the Monte-Carlo standard error of the estimates (std column in Table 2) with the mean of the estimated standard error over all replications (mean(s.e.) column), we find that our computed standard errors fairly reflect estimation uncertainty. Overall, Panel A shows that the parameters and standard errors of the Gaussian and Student's t factor copulas with score-driven dynamic factor loadings can be accurately estimated if the model is correctly specified.

Panel B of Table 2 shows results for the MF-LT model from equation (6). The results for A, B, and ν_C are similar to those of the MF model. Our two-step approach for first targeting the ωs, and then estimating the remaining dynamic parameters thus appears to work well both for estimation and inference. Studying the ω estimates in the full tables in the online Appendix, we find that the standard deviations of moment-based estimators for ω are higher than the standard errors of the ML estimators for A,B, and ν_C. The two-step estimator thus implies a large computational gain at the expense of some efficiency loss in the estimation of the ωs. The assumed distribution does not appear to have a major impact on the performance.

Finally, we investigate the impact of misspecification of the factor structure on the fitted dependence structure. In this third experiment, we consider a DGP with $N = 25$, $T = 1,000$,

Table 2: Monte Carlo results of parameter estimates of the Multi-Factor-Copula
This table provides Monte Carlo results of parameter estimates using the multi-factor (MF) Gaussian and t-copula model as given in (7), and the MF-LT model based on (6). For full details, see the online Appendix B.2. B(N) and $B(t)$ denote the value of B in case of the Gaussian (N) and Student's t (t) factor copula model, respectively. The table reports the mean and standard deviation of the estimated coefficients, as well as the mean of the computed standard error. Results are based on 1,000 Monte Carlo replications.

Coef.	True	Normal			Student's t		
		mean	std	mean(s.e.)	mean	std	mean(s.e.)
Panel A: MF, $T = 1,000$							
A_{eq}(N)	0.0085	0.0085	0.0009	0.0008			
$A_{gr,f}$(N)	0.0095	0.0093	0.0018	0.0018			
$A_{eq}(t)$	0.0150				0.0149	0.0020	0.0019
$A_{gr,f}(t)$	0.0100				0.0096	0.0016	0.0016
B(N)	0.8700	0.8626	0.0221	0.0248			
$B(t)$	0.9200				0.9149	0.0129	0.0126
ν_C	35.00				35.1760	1.8629	1.8821
Panel B: MF-LT, $T = 1,000$							
A	0.015	0.0161	0.0006	0.0006	0.0161	0.0007	0.0006
B	0.970	0.9700	0.0025	0.0023	0.9697	0.0024	0.0024
ν_C	35.00				35.06	1.862	1.846

Table 3: Performance of misspecified factor copulas
This table summarizes the mean and standard deviation of the average Euclidian distance between a simulated correlation matrix $\boldsymbol{R}_t$ ($t = 1, \ldots, 1,000$) from the MF-LT model with $t(35)$-distributed errors and the estimated $\hat{\boldsymbol{R}}_t$ based on one, two, or multi-factor copula models with either Gaussian or a Student's t distributions. All results are based on 1,000 replications.

	MF-LT	MF	2F	1F-Group	1F-Equi
Student's t	0.058	1.692	1.600	0.891	2.541
	(0.023)	(0.266)	(0.222)	(0.107)	(0.308)
Gaussian	0.080	1.699	1.567	0.914	2.559
	(0.023)	(0.263)	(0.205)	(0.109)	(0.312)

and the MF factor structure, using Student's $t(35)$ distributed errors for $G = 5$ different groups, each containing $N/G = 5$ units. Using different (possibly mis-specified) factor copula models, we compute the time average of squared Frobenius norm of $\hat{\boldsymbol{R}}_t - \boldsymbol{R}_t$, which is a consistent loss function according to Laurent et al. (2013). The results in Table 3 clearly indicate that underestimating the number of factors causes substantial discrepancies between the true and the fitted dependence dynamics, particularly for one-factor models with an equi-loading structure, or for the multi-factor models that ignore the different between-group dependencies. This holds irrespective of the distribution used.

4 Empirical application

4.1 Data

In our high-dimensional empirical application we investigate the daily open-to-close returns of 100 randomly chosen constituents of the S&P 500 index, during the period from January 2, 2001 until December 31, 2014 ($T = 3,521$ days). Table B.1 in the online Appendix provides an overview of the ticker symbols of all stocks. The same Table shows the classification of the stocks into 10 groups, based on the firm's industry. For our cross-section, Financials is the largest group with 19 firms, followed by Consumer Services and Energy, respectively. Each industry group includes at least four firms.

To model the marginal characteristics of the daily stock returns, we estimate a univariate t-GAS volatility model as given in equations (D.1)-(D.2) in the online Appendix D. For the conditional mean, we find at most two significant autoregressive (AR) lags. We therefore use an AR(2) conditional mean specification in the marginal models for all 100 stocks. Estimation results are summarized in Table D.1 in the online Appendix D. We find a mean value of $\nu = 8.22$, underlining the fat-tailed nature of daily stock returns even after filtering for time-varying volatility. The mean estimate of β reflects the usual strong persistence in volatility. We follow Creal and Tsay (2015) and evaluate the fit of the marginal distributions by transforming the PITs $\hat{u}_{i,t}$ into Gaussian variables $\hat{x}_{i,t} = \Phi^{-1}(\hat{u}_{i,t})$, $t = 1, \ldots, T$. We subsequently test each series $\hat{x}_{i,t}$ for normality using the Kolmogorov-Smirnov test. Across the 100 firms, we only reject the null hypothesis for 5 series at the 5% significance level. We conclude that the marginal models are adequate for our subsequent analysis.

As an robustness check, we also estimated a GARCH model with the skewed Student's t distribution of Hansen (1994) and compare this to a GARCH-t model for all assets. The results of the skewed Student's t GARCH models are reported in online Appendix D. The comparison indicates that the average increase in the maximized log-likelihood relative to a GARCH t model equals only 1.3 points. Given this weak evidence for the presence of skewness, we therefore remain with the standard Student's t distribution for our main analysis.

4.2 Full-sample comparison

After estimating the parameters of the marginal distributions, we proceed to estimate the parameters of the score-driven factor copula models and the benchmark MGARCH copula models using the full sample of 3,521 observations. The factor copulas are based on grouping firms into industries as laid out in Table B.1 of the online Appendix.

Table 4 shows the parameter estimates and maximized log-likelihood values for all models. Panels A.1 and A.2 contain results for Gaussian and t factor copula specifications, respectively: a one-factor copula with homogeneous (1F-Equi) or with industry-specific (1F-Group) loadings, a two-factor copula (2F) with one factor with homogeneous loadings and one factor with industry-specific loadings, a multi-factor copula (MF) with 10 industry factors, and a multifactor model (MF-LT) with a triangular loading matrix. Panels B.1 and B.2 contain results for Gaussian and t benchmark copulas from the MGARCH class: the cDCC, DECO and Block DECO models. In both multi-factor copula models, we assume that the B parameter in the GAS dynamics of the factor loadings is the same for all parameters. For the MF-LT model we assume a common scalar A, while A is allowed to differ between the common factor and the industry-specific factors in the MF model. In the 2F model we also allow for different A values for the two common factors but assume B is the same. Finally, for the 1F-GR model we assume a common A and B parameter for all different groups. To save space, we do not report all the different intercepts ω_g for all groups for the factor copulas with group-specific loadings. These detailed results are provided in the online Appendix E. Standard errors are based on the sandwich (robust covariance matrix) estimator $\hat{\boldsymbol{H}}_0^{-1}\hat{\boldsymbol{G}}_0\hat{\boldsymbol{H}}_0^{-1}$ with $\hat{\boldsymbol{H}}_0$ the inverse Hessian of the likelihood, and $\hat{\boldsymbol{G}}_0$ the outer product of the gradient.

Five interesting results emerge from Table 4. First, in terms of the statistical fit, the MF-LT t model outperforms the other factor-copula models, as well as the MGARCH-copula models (cDCC, DECO and block DECO). The MF-LT model not only achieves the highest total log-likelihood value, but also performs best in terms of AIC, which takes into account the number of estimated parameters.

Second, multi-factor models provide a much better fit than one-factor copula models. For

Table 4: Parameter estimates of the full sample

This table reports maximum likelihood parameter estimates of various factor copula models, the (block) DECO model of Engle and Kelly (2012) and the cDCC model of Engle (2002), applied to daily returns of 100 stocks included in the S&P 500 index. We consider five different factor copula models, see Table 1 for the definition of their abbreviations. Panel A.1 presents the factor models with a Gaussian copula density, Panel A.2 presents the parameter estimates corresponding with the Student's t copula. Panel B.1 and B.2 present the estimates of the MGARCH class of models. In case of the cDCC and Block DECO models, the table shows parameters estimates obtained by the Composite Likelihood (CL) method. Standard errors are provided in parenthesis and based on the (sandwich) robust covariance matrix estimator. We report the copula log-likelihood,the Akaike Information Criteria (AIC) as well as the number of estimated parameters for all models. The sample comprises daily returns from January 2, 2001 until December 31, 2014 (3,521 observations).

Model	ω^{eq}	A^{eq}	A^{ind}	A^{gr}	B	ν_C	LogL	AIC	♯ para
Panel A.1: Gaussian factor copulas									
1F-Equi	0.017	0.005			0.975		65,934	-131,862	3
	(0.002)	(0.000)			(0.003)				
1F-Group				0.007	0.970		68,086	-136,148	12
				(0.001)	(0.006)				
2F	0.047	0.012		0.013	0.941		71,667	-143,306	14
	(0.006)	(0.000)		(0.001)	(0.009)				
MF	0.042	0.012	0.014		0.930		81,827	-163,626	14
	(0.005)	(0.000)	(0.001)		(0.009)				
MF-LT		0.009			0.964		**83,226**	**-166,339**	57
		(0.001)			(0.005)				
Panel A.2: t-factor copulas									
1F-Equi	0.062	0.012			0.918	36.52	69,679	-139,350	4
	(0.013)	(0.001)			(0.016)	(1.52)			
1F-Group				0.005	0.986	31.87	72,293	-144,560	13
				(0.000)	(0.001)	(1.11)			
2F	0.004	0.009		0.006	0.993	38.57	77,828	-155,627	15
	(0.002)	(0.001)		(0.001)	(0.002)	(1.64)			
MF	0.033	0.012	0.012		0.957	44.98	84,858	-169,687	15
	(0.002)	(0.001)	(0.001)		(0.002)	(1.78)			
MF-LT		0.004			0.990	36.22	**86,433**	**-172,749**	58
		(0.000)			(0.002)	(1.38)			
Panel B.1: Gaussian copula-MGARCH models									
cDCC (CL)		0.017			0.968		74,263	-138,623	4,952
		(0.001)			(0.003)				
DECO		0.071			0.929		64,474	-119,044	4,952
		(0.001)			(0.001)				
Block DECO		0.030			0.957		**83,087**	**-156,270**	4,952
		(0.002)			(0.003)				
Panel B.2: t copula-MGARCH models									
cDCC (CL)		0.018			0.968	14.17	82,688	-155,470	4,953
		(0.001)			(0.002)	(0.58)			
DECO		0.106			0.894	34.43	69,314	-128,721	4,953
		(0.000)			(0.000)	(0.80)			
Block DECO		0.032			0.955	22.51	**86,222**	**-162,537**	4,953
		(0.002)			(0.003)	(0.60)			

example, the log-likelihood difference between the MF-LT t copula and the 1F-Equi t copula is more than 15,000 points. The largest gain with respect to the factor structure is obtained by including industry factors, that is, extending the 1F-Equi model to the MF specification. This increases the log likelihood by 15,000 points in both the Gaussian and Student's t case. Note that allowing for industry specific loadings in the single-factor model leads to a much more modest improvement in log likelihood of 2,500 points. Extending the single-factor model with a second factor with industry-specific loadings performs better, but the increase in log-likelihood is still only half of the improvement achieved by the MF specification.

Third, the Student's t factor copulas fit considerably better than their Gaussian counterparts. Log-likelihood differences range between 3,000 and 6,000 points, depending on the specification. Differences for the multi-factor specifications are typically at the lower end of this range. This underlines that allowing for more than one factor also takes care of part of the tail clustering.

Fourth, we find strong persistence in the time-varying factor loadings with a value of $B \approx 0.97$ for most of the estimated (t-)factor copula models. This finding, as well as the previous one, confirms the empirical results of Oh and Patton (2018) using an entirely different dataset of log-differences of U.S. CDS spreads.

Finally, we note that the estimated degrees of freedom parameter ν_C is (much) lower for the Block DECO t and cDCC t specifications than for the MF-LT t model or the DECO model. It seems that there is empirically some bias effect due to the use of the Composite Likelihood method for parameter estimation.

Our main results are robust against two variations in the estimation set-up. First, we re-estimate all models based on PITs obtained from estimating a skewed Student's t GARCH model for the marginals. Second, we investigate the sensitivity of the MF-LT t model with respect to the ordering of the industries by re-estimating the MF-LT t model for 50 different random industry orderings. Online Appendix E shows the results for both robustness checks and confirm that our conclusions continue to hold.

Figure 1 shows an example of within and between industry correlations according tot he 1F-

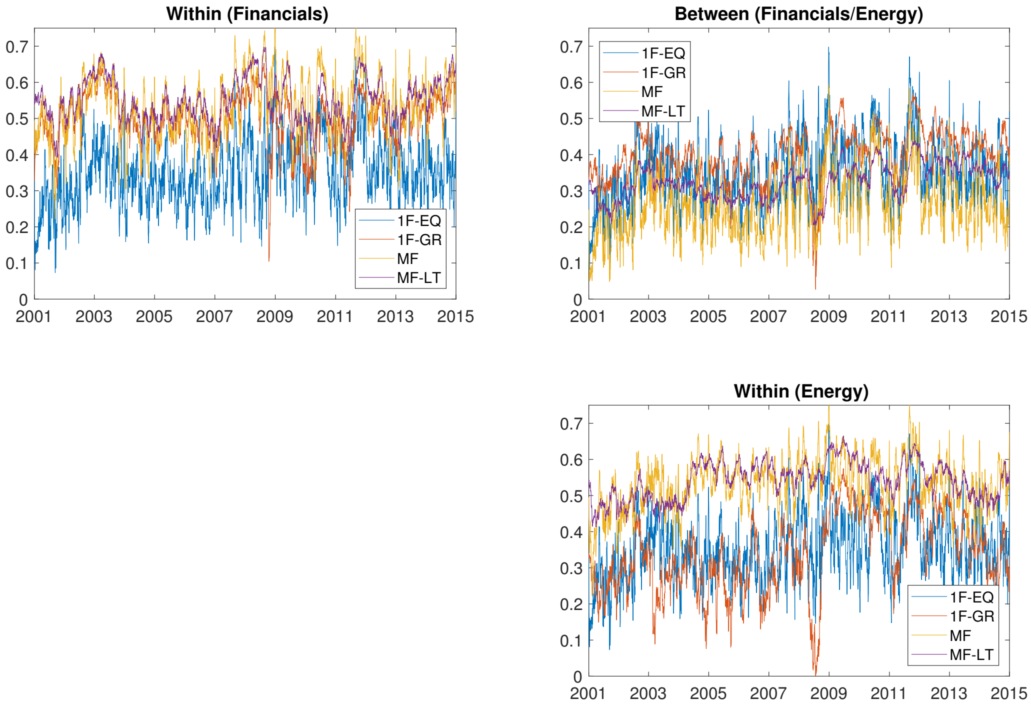

Figure 1: In-sample within and between industry correlations of the MF-LT t copula
This figure shows the fitted within industry correlation of Financials and Energy, as well as their between industry correlations according to the 1F-Equi t, 1F-Group t, MF t and MF-LT t models. The upper-left and lower-right panels show the within industry correlations for Financials and Energy companies, respectively, and the upper-right panel shows the between-industry correlation. The sample spans the period from January 2, 2001 until December 31, 2014 ($T = 3,521$ days).

Equi t, 1F-Group t, MF t and MF-LT t models. We see for instance that the one-factor copula models have a considerably different within industry correlation pattern than the multi-factor models, particularly for the Energy sector. The between correlations are much closer, though we still see that the MF-LT model results in the least noisy correlation estmates. The main take-away is that our multi-factor models picks up the within industry specific correlations that cannot be captured by a single factor models. This ability of the MF and MF-LT models explains their substantial increases in statistical fit as shown before.

4.3 Alternative groupings based on dynamic risk factors

So far, we have allocated firms into groups using their industry classification, but alternative grouping mechanisms are of course possible. Here we investigate an obvious candidate by

Table 5: Forming groups in the MF-LT model

This table reports maximum likelihood parameter estimates of the multi-factor copula model, applied to daily equity returns of 90 assets listed at the S&P 500 index. The 10 groups associated with the models are formed on different risk-factors, such as the book-to-market ratio, size and momentum. In addition, we consider the model based on industry groups. Standard errors are provided in parenthesis and based on the (sandwich) robust covariance matrix estimator. We report only the A and B of all estimated parameters (hence omitting the incercepts) and the copula log-likelihood for all models. The sample comprises daily returns from January 2, 2001 until December 31, 2014 (3,521 observations).

	Gaussian copula			t copula			
	A	B	LogL	A	B	ν_C	LogL
Value	0.009	0.965	61,025	0.004	0.995	29.16	64,950
	(0.001)	(0.005)		(0.000)	(0.001)	(0.98)	
Size	0.010	0.904	60,955	0.003	0.966	28.58	64,953
	(0.001)	(0.010)		(0.001)	(0.011)	(0.92)	
Momentum	0.009	0.963	61,368	0.003	0.992	29.33	65,144
	(0.001)	(0.003)		(0.001)	(0.002)	(1.01)	
Industry	0.010	0.964	**76,531**	0.005	0.989	34.49	**79,458**
	(0.001)	(0.006)		(0.000)	(0.002)	(1.22)	

forming groups based on key asset pricing risk factors, including firm size (market capitalization), value (book-to-market), and momentum, see Fama and French (1993) and Carhart (1997). Due to data availability, this reduces the sample from 100 to 90 assets. For the size and value factors, we form 10 new groups in July each year based on deciles of sorted market capitalization and book-to-market values of the previous fiscal year. Similarly, for momentum we sort stocks into deciles in January each year based on their sorted past 12-month returns. This mimics the way these factors are constructed in typical asset pricing studies. Using risk factors of this type to form groups comes with an additional challenge, namely that group composition can now change from one period to the next. This can be easily accommodated in the factor copula approach introduced in this paper, where it boils down to a straightforward but tedious bookkeeping exercise to account for possible switches in factor loadings of firms depending on their group allocation at time t. The use of such grouping criteria in a dynamic factor copula framework is new and can be seen as a separate contribution.

Table 5 shows the estimation results for our preferred in-sample model, the MF-LT copula, using different grouping criteria. Given the time-varying group composition, we alter our targeting approach for $\boldsymbol{\omega}$ by each year using the moments estimator (12) based on the unconditional correlation matrix $\boldsymbol{R}_M$ of $\hat{x}_{it} = \Phi^{-1}(\hat{u}_{it})$ for the 250 daily observations of the upcoming year. In a second step, we estimate the parameters A and B. The results show that the models with dynamic groups based on risk factors achieve a considerably worse statistical fit than the model with static industry groups. The minimum loss in log-likelihood exceeds 14,000 points. Among the three risk factors, momentum seems to performs best, but differences with size and value are small. The conclusions on the preferred grouping structure do not depend on the distributional assumption, and are similar for the Gaussian and the Student's t case.

4.4 Multivariate Density Forecasts

As we have closed-form copula density expressions, a natural way to compare the out-of-sample (OOS) forecasting performance of factor copula models and copula MGARCH models is to consider multivariate density forecasts as in Salvatierra and Patton (2015). Because we use the same marginal distributions in all models, the density forecast comparison actually boils down to an evaluation and comparison of the OOS copula density forecasts.

We using a moving estimation window of 1,000 observations (or roughly four calendar years), which leaves $P = 2,521$ observations for the out-of-sample period, starting 28 December 2004. Hence the OOS period includes the Great Financial Crisis. We re-estimate the parameters in all models after each 50 observations (or roughly 10 calendar weeks) and construct a one-step ahead density forecast each day.

We evaluate the copula density forecasts using two scoring rules. First, we consider their accuracy on their full support by means of the log scoring rule (see Mitchell and Hall, 2005; Amisano and Giacomini, 2007)

$$S_{l,t}(\hat{\boldsymbol{u}}_t, M_j) = \log c_t(\hat{\boldsymbol{u}}_t \mid \hat{\boldsymbol{\theta}}_{C,t}, M_j), \tag{15}$$

where $c_t(\cdot \mid \hat{\boldsymbol{\theta}}_{C,t}, M_j)$ is the Gaussian or Student's t conditional copula density obtained from model M_j and $\hat{\boldsymbol{u}}_t$ denote the corresponding PITs. Note that the PITs in $\boldsymbol{u}_t$ are based on the same marginal distributions for both model specifications in any log score comparison, and that the marginal densities therefore drop out from a difference in log scores between to models. We therefore omit the marginals from the log score expression in (15). This underlines that we are really comparing the forecasting quality of the copula part. Second, we focus the evaluation on the joint lower region of the copula support by using the conditional likelihood (cl) scoring rule proposed by Diks et al. (2014),

$$S_{cl,t}(\hat{\boldsymbol{u}}_t, M_j) = \Big(\log c_t(\hat{\boldsymbol{u}}_t \mid \hat{\boldsymbol{\theta}}_{C,t}, M_j) - \log C_t(\boldsymbol{q} \mid \hat{\boldsymbol{\theta}}_{C,t}, M_j)\Big) \times I[\hat{\boldsymbol{u}}_t < \boldsymbol{q}] \tag{16}$$

where $\boldsymbol{q}$ is an $N \times 1$ vector and $C_t(\cdot \mid \hat{\boldsymbol{\theta}}_{C,t}, M_j)$ is the conditional Copula function, and $I[\hat{\boldsymbol{u}}_t < \boldsymbol{q}] = \prod_{i=1}^{N} I[\hat{\boldsymbol{u}}_{i,t} < q_i]$ with $q_i \in [0,1]$, $i = 1, \ldots, N$. Hence (16) is the log-likelihood of model M_j conditional on $\boldsymbol{u}_t < \boldsymbol{q}$ (element-wise). For any $\boldsymbol{q} = (q_1, \ldots, q_N)$ this boils down to the joint lower region $[0, q_1] \times \cdots \times [0, q_N]$. Obviously, when $q_i = 1$ for all i, we recover the log scoring rule. We use a time-varying threshold $\boldsymbol{q}_t = (\bar{q}_t, \ldots, \bar{q}_t)$, where $\bar{q}_t$ is such that $\frac{1}{1000}\sum_{j=1}^{1000} I[\hat{\boldsymbol{u}}_{t-j} < \boldsymbol{q}_t] = q$, with $q = 0.01$ or 0.05, and T_1 and T_2 the bounds of the (rolling) estimation sample. We thus compare the copula density forecasts in the joint empirical lower 1% or 5% tail.

For both scoring rules, models that deliver higher values are preferred. We can test whether differences in the scoring rule values for models M_i and M_j are significant by defining the score differential

$$d_{x,ij,t} = S_{x,t}(\hat{\boldsymbol{u}}_t, M_i) - S_{x,t}(\hat{\boldsymbol{u}}_t, M_j), \qquad \text{with } x = l, cl. \tag{17}$$

The null hypothesis of equal predictive ability is equivalent to $H_0 : \mathbb{E}[d_{x,ij,t}] = 0$, which can be tested using a standard Diebold and Mariano (1995) test statistic. Since we deal with a substantial number of different models and factor structures and hence many different copula density forecasts, we consider the Model Confidence Set (MCS) of Hansen et al. (2011). The MCS automatically accounts for the dependence between model outcomes given that all models are based on the same data.

Table 6 shows results of the copula density forecast evaluation. We report the mean of the log scores and the conditional likelihood scores, as well as the p-values of the Model Confidence Set.

The table shows three interesting results. First, in line with our full-sample results, the MF-LT t model performs best in terms of predictive ability when evaluated over the full copula support using the log scoring rule. The same pattern emerges from the MCS. The MCS p-value equals 1 for the MF-LT t, whereas that of all other models is below 0.01. Second, similar to the in-sample results, most of the gain for the factor copulas is obtained by allowing for industry-specific factors. For example, changing the equifactor from fixed (1F-Equi t) to industry-specific loadings (1F-Group t) increases the average log-score by only 0.75 points (from 21.08 to 21.83). Allowing for different industry factors (MF t), however, implies an additional increase of almost 4 points to an average log-score of 25.60. Allowing for cross-exposures in the MF-LT specification results in yet a further increase by 0.5 points. Third, when we consider density forecasts in the joint lower tail, the MF-LT model is always part of the MCS. In that case, however, also the MGARCH specifications perform well and are included in the MCS, in particular the block DECO-t model. The differences in the conditional likelihood scores are small in these cases, however, and below 0.015 points.

Overall, we conclude that the flexibility provided by the new MF-LT t model is also important out of sample using density forecast criteria. The more flexible parametrization allows for a larger class of dependence matrices than more restrictive one-factor models. This extension appears to be empirically important in high dimensions.

4.5 Economic out-of-sample performance

Finally, we assess the forecasting performance of the different models from an economic perspective. We do so by considering the ex-post variance of the ex-ante global minimum variance portfolio (GMVP); compare Chiriac and Voev (2011) and Engle and Kelly (2012), among others. The best forecasting model should provide portfolios with the lowest ex-post variance.

Assuming that an investor aims to minimize the 1-step portfolio volatility at time t subject to being fully invested, the resulting GMVP weights $\boldsymbol{w}_{t+1|t}$ are obtained as the solution of the

Table 6: One-step ahead copula density forecasts

This table evaluates the accuracy of one-step ahead copula density forecasts (in the left tail) of daily return series for 100 stocks from the S&P500 index, obtained by various factor copula and copula MGARCH models, assuming a Gaussian or Student's t distribution (denoted by N or t). We consider a 1-factor model with equi-loadings (1F-Equi), a 1-factor model with group-specific loadings (1F-Group), a 2-Factor model with one equifactor and an additional factor with group-specific loadings (2F), a multi-factor copula model with one equi-factor plus G group-specific factors (MF), and the lower triangular multi-factor model (MF-LT). In addition, we show the results of the cDCC model of Engle (2002) and the (block) DECO model of Engle and Kelly (2012). The table presents the mean of the log score (S_{ls}) and the conditional (tail) likelihood score (S_{cl}) for the lower joint 1% and 5% tail. We present the p-value associated with the Model Confidence Set of Hansen et al. (2011) in parentheses. Bold numbers in this row represent models that belong to the model confidence set at a significance level of 5%. The out-of-sample period covers December 28, 2004 until December 31, 2014 and contains 2,521 observations.

	Full	1% tail	5% tail
Model	$S_{ls,t}$(p-val)	$S_{cl,t}$(p-val)	$S_{cl,t}$(p-val)
1F-Equi	20.07 (0.00)	1.401 (0.00)	4.022 (0.00)
1F-Equi t	21.08 (0.00)	1.443 (0.00)	4.142 (0.00)
1F-Group N	20.73 (0.00)	1.411 (0.00)	4.045 (0.00)
1F-Group t	21.83 (0.00)	1.445 (0.00)	4.177 (0.00)
2F N	22.52 (0.00)	1.436 (0.00)	4.138 (0.00)
2F t	23.53 (0.00)	1.469 (0.01)	4.267 (0.00)
MF N	24.95 (0.00)	1.466 (0.00)	4.284 (0.00)
MF t	25.60 (0.00)	**1.494 (0.22)**	4.373 (0.04)
MF-LT N	25.32 (0.00)	1.469 (0.00)	4.291 (0.00)
MF-LT t	**26.10 (1.00)**	**1.500 (0.40)**	**4.400 (0.36)**
cDCC N	22.42 (0.00)	**1.501 (0.50)**	**4.369 (0.36)**
cDCC t	24.37 (0.00)	**1.509 (1.00)**	**4.384 (0.36)**
DECO N	19.76 (0.00)	1.415 (0.00)	4.020 (0.00)
DECO t	21.01 (0.00)	1.447 (0.00)	4.143 (0.00)
Block DECO N	25.24 (0.00)	1.478 (0.03)	4.314 (0.00)
Block DECO t	26.02 (0.01)	**1.505 (0.76)**	**4.415 (1.00)**

quadratic programming problem

$$\min \boldsymbol{w}_{t+1|t}^{\top} \left(\boldsymbol{H}_{t+1|t}\boldsymbol{R}^{*}_{t+1|t}\boldsymbol{H}_{t+1|t}\right) \boldsymbol{w}_{t+1|t}, \qquad \text{s.t. } \boldsymbol{w}_{t+1|t}^{\top}\boldsymbol{\iota} = 1, \tag{18}$$

with $\boldsymbol{H}_{t+1|t}$ the 1-step forecasts of the variances based on the marginal models, and $\boldsymbol{R}^{*}_{t+1|t}$ the one-step ahead forecast of the correlation matrix. As the forecast of the correlation matrix $\boldsymbol{R}^{*}_{t+1|t}$ is not the same as the forecast of the copula dependence matrix $\boldsymbol{R}_{t+1|t}$, we obtain the former by simulating 20,000 returns from the joint distribution of returns as constructed from the marginals and the conditional copula. Following Chiriac and Voev (2011), we assess the predictive ability of the different models by comparing the results to the *ex-post* realizations of the conditional standard deviation $\sigma_{p,t}$, given by $\sigma_{p,t} = \sqrt{\boldsymbol{w}_{t+1|t}^{\top} \boldsymbol{RC}_{t+1} \boldsymbol{w}_{t+1|t}}$, with $\boldsymbol{RC}_{t+1}$ the realized covariance matrix obtained using 5-minute returns. We decompose this matrix into realized variances and a realized correlation matrix, where the latter is by definition ill-conditioned and not positive definite. We use the 'eigenvalue cleaning' method used by Hautsch et al. (2012) to get a positive definite correlation matrix. Having constructed the ex-post conditional portfolio standard deviation, we test model performance by means of the Model Confidence Set (MCS) approach with a significance level of 5%.

Alongside the GMVP's volatility, we also calculate a number of other relevant quantities, such as portfolio turnover (TO_t), concentration (CO_t), and the total short position (SP_t) for each competing model at time t. Turnover at time t is defined as

$$TO_t = \sum_{i=1}^{N} \left| w_{i,t+1|t} - w_{i,t|t-1} \frac{1 + y_{i,t}}{1 + \boldsymbol{w}_{t|t-1}^{\top} \boldsymbol{y}_t} \right|, \tag{19}$$

where $w_{i,t|t-1}$ is the i-th element of the weight vector $\boldsymbol{w}_{t|t-1}$. It measures the value of the portfolio that is bought/sold when rebalancing the portfolio to its new optimal position from time t to $t+1$. A model that produces more stable correlation matrix forecasts implies in general less turnover and hence, less transaction costs. Portfolio concentration and total portfolio short position both measure the amount of extreme portfolio allocations. Again, more stable forecasts of $\boldsymbol{R}_{t+1|t}^{*}$ should result in less extreme portfolio weights. Portfolio concentration is defined as

$$CO_t = \left(\sum_{i=1}^{N} w_{i,t|t-1}^2 \right)^{1/2}, \tag{20}$$

while the total portfolio short position SP_t is given by

$$SP_t = \sum_{i=1}^{N} w_{i,t|t-1} \cdot I[w_{i,t|t-1} < 0], \tag{21}$$

with $I[\cdot]$ an indicator function that takes the value 1 if the i-the element of the weight vector is lower than zero.

Table 7: Minimum variance portfolio results

This table reports results on a global minimum variance portfolio strategy, based on 1-step ahead predictions of the daily covariance matrix, according to four different type of factor copulas, the cDCC model of Engle (2002) and the (block) DECO model of Engle and Kelly (2012). The columns represent two types of one-factor copulas (one equi-factor or one factor with group-specific loadings, denoted by 1F-eq and 1F-gr), one 2-Factor model (one equi-factor plus an additional factor with group-specific loadings) and two types of multi-factor copula models (one equi-factor plus G group-specific factors and the MF-LT model). Each type of model is further discriminated across distribution (Gaussian vs. a Student's t) For each model, we show the mean of the ex-post portfolio standard deviation, the p-value corresponding with the Model Confidence Set of Hansen et al. (2011), using a significance level of 5%, and the mean of the portfolio turnover (TO), concentration (CO) and the total portfolio short positions (SP). Bold numbers indicate the models that stay within the MCS, or the lowest (absolute) portfolio turnover, concentration and total portfolio short positions. The out-of-sample period goes from December 28, 2004 until December 31, 2014 and contains 2,521 observations.

Model	$\bar{\sigma}_p$ (p-val)	TO	CO	SP
1F-Equi N	0.527 (0.00)	0.359	0.336	-0.595
1F-Equi t	0.528 (0.00)	0.351	0.339	-0.604
1F-Group N	**0.513 (0.97)**	0.332	**0.307**	**-0.530**
1F-Group t	**0.513 (1.00)**	**0.319**	0.310	-0.541
2F N	0.537 (0.00)	0.395	0.339	-0.604
2F t	0.542 (0.00)	0.377	0.343	-0.615
MF N	0.537 (0.00)	0.418	0.357	-0.636
MF t	0.538 (0.00)	0.407	0.359	-0.642
MF-LT N	**0.523 (0.07)**	0.407	0.353	-0.622
MF-LT t	**0.522 (0.11)**	0.397	0.353	-0.625
cDCC N	0.556 (0.00)	0.763	0.470	-1.032
cDCC t	0.564 (0.00)	0.774	0.477	-1.054
DECO N	0.525 (0.01)	0.345	0.335	-0.596
DECO t	0.525 (0.00)	0.346	0.335	-0.598
Block DECO N	0.531 (0.00)	0.398	0.352	-0.622
Block DECO t	0.532 (0.00)	0.399	0.352	-0.623

Table 7 reports the economic out-of-sample performance results. As for the density forecast results, the factor copulas also perform best in terms of economic performance. Again, the MF-LT model is included in the MCS. There are however also a number of remarkable differences. In terms of the ex-post variance of the GMVP, the 1-factor copulas with industry specific loadings now perform best. This contrasts with the density forecast setting, where the MF-LT t model performed best in-sample and out-of-sample. The multi-factor models also now still (marginally) outperform the block-DECO model in terms of ex-post variance of the GMVP. The 1-factor models, however, have the best performance, both in terms of ex-post variance, turnover, concentration, and total short positions. Unlike the density forecast setting, we also note that the choice of the distribution plays a less important role in Table 7.

To reconcile the findings in terms of economic performance with those of the density forecast evaluation from the previous subsection, it is important to note that the GMVP evaluation takes a very specific perspective. The GMVP focuses attention on an area of the forecast distribution where differences are more concentrated by design: all models focus on a portfolio with ex-ante *minimum* variance. If the different models are any good, differences in this concentrated performance measure are harder to obtain. This is corroborated by the results in Table 7. Although the results are sometimes statistically significantly different, they are all quite close in economic terms (with the possible exception of the cDCC, which lags somewhat more). Using such a performance measure, introducing models with more parameters and with the associated estimation risk typically deteriorates overall forecasting performance. This explains why the simpler 1-factor models do better here. By contrast, if the full density or complete tail area is taken into account as in the previous subsection, the additional flexibility of the more complex factor models has a beneficial effect on performance, particularly in the current high-dimensional setting.

5 Conclusions

We have introduced various factor structures within the class of closed-form factor copula models for high dimensions. The new factor copula model is computationally tractable with score-driven dynamics, implying a closed from copula density. Parameters can be estimated in a straightforward way by Maximum Likelihood and/or a fast two-step approach that combines a moment-based estimator and the ML approach.

The factor structures are based on group-specific characteristics such as industry classification. In addition, an important feature of our model is that it allows for more than one factor. Extensions to the model are also easily possible, such as the inclusion of covariates to describe the factor loading dynamics, and different time-varying group structures, such as the risk-based groups (size, value, momentum) also considered in this paper. This can be done without any difficulty for the positive definiteness of the implied dependence matrix.

Empirically, we modeled the dependence across 100 equity returns from the S&P 500 index over the period 2001-2014. We found that our factor copula models outperform multivariate GARCH (MGARCH) based counterparts, such as the (c)DCC and (block) DECO. In-sample, the multi-factor copula model has a better fit than one-factor models and benchmarks such as the cDCC and (Block-)DECO. Out-of sample, the good performance of multi-factor copula models persists. A simple static industry-based group structure for the copula appears better statistically than risk-based groups based on size, value, or momentum. Measured in terms of density forecasts, the multi-factor models perform best, whereas in terms of the global minimum variance portfolio variance, simpler 1-factor models outperform other models. In all settings, we thus find score-driven factor copulas to describe the dynamics of the data well. Given their computational ease and closed-form likelihood expression, they thus provide a useful tool for modeling high-dimensional dynamic dependence structures.

Acknowledgements

We thank Andrew Patton, the associate editor, two anonymous referees, and participants at the 10^{th} Annual SoFiE conference and seminar participants at Tinbergen Institute Amsterdam, Lund University, Heidelberg University, Maastricht University, and Vrije Universiteit Amsterdam for helpful comments.

References

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

Amisano, G. and R. Giacomini (2007). Comparing density forecasts via weighted likelihood ratio tests. *Journal of Business and Economic Statistics 25*, 177–190.

Blasques, F., S. Koopman, and A. Lucas (2015). Information-theoretic optimality of observation-driven time series models for continuous responses. *Biometrika 102*, 325–343.

Carhart, M. M. (1997). On persistence in mutual fund performance. *The Journal of Finance 52*(1), 57–82.

Cherubini, U., S. Mulinacci, F. Gobbi, and S. Romagnoli (2011). *Dynamic Copula Methods in Finance*. John Wiley and Sons.

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

Christoffersen, P., V. Errunza, K. Jacobs, and X. Jin (2014). Correlation dynamics and international diversification benefits. *International Journal of Forecasting 30*, 807–824.

Christoffersen, P., V. Errunza, K. Jacobs, and H. Langlois (2012). Is the potential for international diversification disappearing? a dynamic copula approach. *Review of Financial Studies 25*, 3711–3751.

Creal, D., S. Koopman, and A. Lucas (2011). A dynamic multivariate heavy-tailed model for time-varying volatilities and correlations. *Journal of Business Economic Statistics 29*, 552–563.

Creal, D., S. Koopman, and A. Lucas (2013). Generalized autoregressive score models with applications. *Journal of Applied Econometrics 28*, 777–795.

Creal, D., S. Koopman, A. Lucas, and M. Zamojski (2018). Generalized autoregressive Method of Moments. Tinbergen Institute Discussion Paper, TI 14-029/III (updated).

Creal, D. and R. Tsay (2015). High dimensional dynamic stochastic copula models. *Journal of Econometrics 189*, 335–345.

Diebold, F. and R. Mariano (1995). Comparing predictive accuracy. *Journal of Business and Economic Statistics 13*, 253–263.

Diks, C., V. Panchenko, O. Sokolinskiy, and D. van Dijk (2014). Comparing the accuracy of multivariate density forecasts in selected regions of the copula support. *Journal of Economic Dynamics and Control 48*, 79–94.

Engle, R. (2002). Dynamic conditional correlation: A simple class of multivariate generalized autoregressive conditional heteroskedasticity models. *Journal of Business and Economic Statistics 20*, 339–350.

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

Engle, R., V. Ng, and M. Rothschild (1990). Asset pricing with a factor-arch covariance structure: Empirical estimates for treasury bills. *Journal of Econometrics 45*, 213–237.

Engle, R., N. Shephard, and K. Sheppard (2008). Fitting and testing vast dimensional time-varying covariance models. Working Paper.

Fama, E. and K. French (1993). Common risk factors in the returns on stocks and bonds. *Journal of Financial Economics 33*, 3–56.

Fama, E. and K. French (2016). Dissecting anomalies with a five-factor model. *Review of Financial Studies 29*, 69–103.

Fan, J., Y. Fan, and J. Lv (2008). High dimensional covariance matrix estimation using a factor model. *Journal of Econometrics 147*, 186–197.

Fan, J., Y. Liao, and H. Liu (2016). An overview of the estimation of large covariance and precision matrices. *The Econometrics Journal 19*, C1–C32.

Fan, J., Y. Liao, and M. Mincheva (2011). High dimensional covariance matrix estimation in approximate factor models. *Annals of Statistics 39*, 3320.

Fan, Y. and A. Patton (2014). Copulas in econometrics. *Annual Review of Economics 6*, 179–200.

Hafner, C. and H. Manner (2012). Dynamic stochastic copula models: Estimation, inference and applications. *Journal of Applied Econometrics 27*, 269–295.

Hansen, B. (1994). Autoregressive conditional density estimation. *International Economic Review 35*, 705–730.

Hansen, P., A. Lunde, and J. Nason (2011). The model confidence set. *Econometrica 79*, 453–497.

Harvey, A. (2013). *Dynamic Models for Volatility and Heavy Tails: With Applications to Financial and Economic Time Series*. Cambridge University Press.

Hautsch, N., L. Kyj, and R. Oomen (2012). A blocking and regularization approach to high-dimensional realized covariance estimation. *Journal of Applied Econometrics 27*, 625–645.

Krupskii, P. and H. Joe (2013). Factor copula models for multivariate data. *Journal of Multivariate Analysis 120*, 85–101.

Laurent, S., J. Rombouts, and F. Violante (2013). On loss functions and ranking forecasting performances of multivariate volatility models. *Journal of Econometrics 173*, 1–10.

Lucas, A., B. Schwaab, and X. Zhang (2014). Conditional euro area sovereign default risk. *Journal of Business and Economic Statistics 32*, 271–284.

Lucas, A., B. Schwaab, and X. Zhang (2017). Modeling financial sector joint tail risk in the euro area. *Journal of Applied Econometrics 32*, 171–191.

Manner, H. and O. Reznikova (2012). A survey on time-varying copulas: Specification, simulations, and application. *Econometric Reviews 31*, 654–687.

McNeil, A., R. Frey, and P. Embrechts (2015). *Quantitative risk management: Concepts, techniques and tools*. Princeton university press.

Mitchell, J. and S. Hall (2005). Evaluating, comparing and combining density forecasts using the klic with an application to the bank of england and niesr fan-charts of inflation. *Oxford Bulletin of Economics and Statistics 67*, 995–1033.

Oh, D. and A. Patton (2017). Modeling dependence in high dimensions with factor copulas. *Journal of Business and Economic Statistics 35*, 139–154.

Oh, D. and A. Patton (2018). Time-varying systemic risk: Evidence from a dynamic copula model of cds spreads. *Journal of Business and Economic Statistics 36*, 181–195.

Patton, A. (2006). Modelling asymmetric exchange rate dependence. *International Economic Review 47*, 527–556.

Patton, A. (2009). Copula-based models for financial time series. In *Handbook of financial time series*, pp. 767–785. Springer.

Patton, A. (2013). Copula methods for forecasting multivariate time series. In *Handbook of economic forecasting*, Volume 2, pp. 899–960. Elsevier.

Salvatierra, I. and A. Patton (2015). Dynamic copula models and high frequency data. *Journal of Empirical Finance 30*, 120–135.

Online Appendix to:
Closed-Form Multi-Factor Copula Models with Observation-Driven Dynamic Factor Loadings

Anne Opschoor[a,b], Istvan Barra[c], André Lucas[a,b], Dick van Dijk[d,b]

[a] *Vrije Universiteit Amsterdam*

[b] *Tinbergen Institute*

[c] *Independent research*

[d] *Erasmus University Rotterdam*

A Derivations of the score

A.1 General set-up

The general set-up of the (multi) factor copulas is given by equation (2). We are interested in the score s_t, defined as

$$\boldsymbol{s}_t = \partial \log \boldsymbol{c}_t(\boldsymbol{x}_t; \boldsymbol{R}_t, \psi_C)/\partial \boldsymbol{\lambda}_t \tag{A.1}$$

where $\boldsymbol{\lambda}_t$ holds all dynamic factor loadings, and with a slight abuse of notation $\boldsymbol{c}_t(\cdot)$ is related to the conditional copula density. Note that the dimension of $\boldsymbol{\lambda}_t$ (and hence $\boldsymbol{s}_t$) depends on the chosen factor structure.

We consider a Student's t and a Gaussian copula density for $\boldsymbol{x}_t = (x_{1,t}, \ldots, x_{N,t})^\top = (T_\nu^{-1}(u_{1,t}), \ldots, T_\nu^{-1}(u_{N,t}))^\top$ for the vector of PITs $(u_{1,t}, \ldots, u_{N,t})^\top$, with $T_\nu^{-1}(\,\cdot\,)$ the standard inverse Student's t cdf with ν degrees of freedom, where $\nu \to \infty$ for the Gaussian case. We have the following specifications:

$$\log \boldsymbol{c}_{Stud,t}(\boldsymbol{x}_t; \boldsymbol{R}_t, \psi_C) = -\frac{1}{2}\log|\boldsymbol{R}_t| - \frac{\nu+N}{2}\log\left(1+\frac{\boldsymbol{x}_t^\top \boldsymbol{R}_t^{-1}\boldsymbol{x}_t}{\nu-2}\right) + a_{Stud}(\nu), \tag{A.2}$$

$$\log \boldsymbol{c}_{Gauss,t}(\boldsymbol{x}_t; \boldsymbol{R}_t, \psi_C) = -\frac{1}{2}\log|\boldsymbol{R}_t| + -\frac{1}{2}\boldsymbol{x}_t^\top \boldsymbol{R}_t^{-1}\boldsymbol{x}_t + a_{Gaus}, \tag{A.3}$$

where $a_{Stud}(\nu)$ and a_{Gaus} are constants that do not depend on $\boldsymbol{R}_t$. Further, the dependence matrix $\boldsymbol{R}_t$ is modeled as

$$\boldsymbol{R}_t = \tilde{\boldsymbol{L}}_t^\top \tilde{\boldsymbol{L}}_t + \boldsymbol{D}_t, \qquad \tilde{\boldsymbol{L}}_t = \left(\tilde{\boldsymbol{\lambda}}_{1,t}, \ldots, \tilde{\boldsymbol{\lambda}}_{N,t}\right), \qquad \boldsymbol{D}_t = \operatorname{diag}\left(\sigma_{1,t}^2, \ldots, \sigma_{N,t}^2\right), \tag{A.4}$$

with

$$\tilde{\boldsymbol{\lambda}}_{i,t} = \frac{\boldsymbol{\lambda}_{i,t}}{\sqrt{1+\boldsymbol{\lambda}_{i,t}^\top \boldsymbol{\lambda}_{i,t}}} = \boldsymbol{\lambda}_{i,t}\cdot\sigma_{it}, \qquad \sigma_{it}^2 = \frac{1}{1+\boldsymbol{\lambda}_{i,t}^\top \boldsymbol{\lambda}_{i,t}} \tag{A.5}$$

for a vector $\boldsymbol{\lambda}_{i,t} \in \mathbb{R}^{k\times 1}$. This ensures that $x_{i,t}$ has unit variance by design.

Define $\boldsymbol{L}_t = (\boldsymbol{\lambda}_{1,t}, \ldots, \boldsymbol{\lambda}_{N,t}) \in \mathbb{R}^{k\times N}$, then using the chain rule we obtain

$$\frac{\partial \log \boldsymbol{c}_t(\boldsymbol{x}_t; \boldsymbol{R}_t, \psi_C)}{\partial \boldsymbol{\lambda}_t^\top} = \frac{\partial \log \boldsymbol{c}_t(\boldsymbol{x}_t; \boldsymbol{R}_t, \psi_C)}{\partial \operatorname{vec}(\boldsymbol{R}_t)^\top} \cdot \frac{\partial \operatorname{vec}(\boldsymbol{R}_t)}{\partial \operatorname{vec}(\boldsymbol{L}_t)^\top} \cdot \frac{\partial \operatorname{vec}(\boldsymbol{L}_t)}{\partial \boldsymbol{\lambda}_t^\top}. \tag{A.6}$$

The first two factors in (A.6) are generic for any factor structure. The last factor, by contrast, strongly depends on the factor structure and will be dealt with in separate subsections. A further factor might be added in case some elements of $\boldsymbol{\lambda}_t$ are restricted to be positive, or lie in some range. This can for instance be obtained by specifying that element of $\boldsymbol{\lambda}_t$ as the exponential function of a new, unrestricted time varying parameter, and by taking the derivative with respect to this new parameter. The derivative of this last type of transformation can be added as a final chain rule factor in (A.6) and will typically take the form of a simple, diagonal matrix.

The first factor in (A.6) only depends on the conditional copula density specification. For the Student's t case, we obtain

$$\begin{aligned}
\mathrm{d}\log \boldsymbol{c}_{Stud,t}(\boldsymbol{x}_t;\boldsymbol{R}_t,\psi_C) &= -\frac{1}{2}\operatorname{tr}\left(\boldsymbol{R}_t^{-1}\,\mathrm{d}\,\boldsymbol{R}_t\right) - \frac{\nu+N}{2}\frac{1}{1+\frac{\boldsymbol{x}_t^\top\boldsymbol{R}_t^{-1}\boldsymbol{x}_t}{\nu-2}}\,\mathrm{d}\left(\frac{\boldsymbol{x}_t^\top\boldsymbol{R}_t^{-1}\boldsymbol{x}_t}{\nu-2}\right)\\
&= -\frac{1}{2}\left(\operatorname{vec}(\boldsymbol{R}_t)^{-1}\right)^\top \mathrm{d}\operatorname{vec}(\boldsymbol{R}_t) + \frac{1}{2}\left(\frac{\nu+N}{\nu+\boldsymbol{x}_t^\top\boldsymbol{R}_t^{-1}\boldsymbol{x}_t}\right)\boldsymbol{x}_t^\top\boldsymbol{R}_t^{-1}(\mathrm{d}\,\boldsymbol{R}_t)\boldsymbol{R}_t^{-1}\boldsymbol{x}_t\\
&= -\frac{1}{2}\left(\operatorname{vec}(\boldsymbol{R}_t)^{-1}\right)^\top \mathrm{d}\operatorname{vec}(\boldsymbol{R}_t) + \frac{1}{2}\left(\frac{\nu+N}{\nu+\boldsymbol{x}_t^\top\boldsymbol{R}_t^{-1}\boldsymbol{x}_t}\boldsymbol{x}_t^\top\boldsymbol{R}_t^{-1}\otimes\boldsymbol{R}_t^{-1}\boldsymbol{x}_t\right)^\top \mathrm{d}\operatorname{vec}(\boldsymbol{R}_t)\\
&= \left(-\frac{1}{2}\left(\operatorname{vec}(\boldsymbol{R}_t)^{-1}\right)^\top + \frac{1}{2}\left(\frac{\nu+N}{\nu+\boldsymbol{x}_t^\top\boldsymbol{R}_t^{-1}\boldsymbol{x}_t}\operatorname{vec}\left(\boldsymbol{R}_t^{-1}\boldsymbol{x}_t\boldsymbol{x}_t^\top\boldsymbol{R}_t^{-1}\right)\right)^\top\right)\mathrm{d}\operatorname{vec}(\boldsymbol{R}_t), \qquad (A.7)
\end{aligned}$$

and hence

$$\frac{\partial\log \boldsymbol{c}_{Stud,t}(\boldsymbol{x}_t;\boldsymbol{R}_t,\nu_C)}{\partial\operatorname{vec}(\boldsymbol{R}_t)^\top} = -\frac{1}{2}\left(\operatorname{vec}(\boldsymbol{R}_t)^{-1}\right)^\top + \frac{1}{2}\left(\frac{\nu+N}{\nu+\boldsymbol{x}_t^\top\boldsymbol{R}_t^{-1}\boldsymbol{x}_t}\operatorname{vec}\left(\boldsymbol{R}_t^{-1}\boldsymbol{x}_t\boldsymbol{x}_t^\top\boldsymbol{R}_t^{-1}\right)\right)^\top. \qquad (A.8)$$

For the Gaussian case, we let $\nu_C\to\infty$ and obtain

$$\mathrm{d}\log \boldsymbol{c}_{Gaus,t}(\boldsymbol{x}_t;\boldsymbol{R}_t,\psi_C) = \left(-\frac{1}{2}\left(\operatorname{vec}(\boldsymbol{R}_t)^{-1}\right)^\top + \frac{1}{2}\operatorname{vec}\left(\boldsymbol{R}_t^{-1}\boldsymbol{x}_t\boldsymbol{x}_t^\top\boldsymbol{R}_t^{-1}\right)^\top\right)\mathrm{d}\operatorname{vec}(\boldsymbol{R}_t), \qquad (A.9)$$

such that

$$\frac{\partial\log \boldsymbol{c}_{Gaus,t}(\boldsymbol{x}_t;\boldsymbol{R}_t,\psi_C)}{\partial\operatorname{vec}(\boldsymbol{R}_t)^\top} = -\frac{1}{2}\left(\operatorname{vec}(\boldsymbol{R}_t)^{-1}\right)^\top + \frac{1}{2}\operatorname{vec}\left(\boldsymbol{R}_t^{-1}\boldsymbol{x}_t\boldsymbol{x}_t^\top\boldsymbol{R}_t^{-1}\right)^\top. \qquad (A.10)$$

For the second factor in (A.6) we obtain

$$\mathrm{d}\operatorname{vec}(\boldsymbol{R}_t) = \mathrm{d}\operatorname{vec}(\tilde{\boldsymbol{L}}_t^\top\tilde{\boldsymbol{L}}_t+\boldsymbol{D}_t) = (\mathbf{I}_{N^2}+K_N)\left(\mathbf{I}_N\otimes\tilde{\boldsymbol{L}}_t^\top\right)\mathrm{d}\operatorname{vec}(\tilde{\boldsymbol{L}}_t)+\mathrm{d}\operatorname{vec}(\boldsymbol{D}_t), \qquad (A.11)$$

where K_N is the commutation matrix, i.e., $\operatorname{vec}(A)=K_N\operatorname{vec}(A^\top)$ for a general $N\times N$ matrix A. As a result, we obtain

$$\frac{\partial\operatorname{vec}(\boldsymbol{R}_t)}{\partial\operatorname{vec}(\boldsymbol{L}_t)^\top} = (\mathbf{I}_{N^2}+K_N)\left(\mathbf{I}_N\otimes\tilde{\boldsymbol{L}}_t^\top\right)\cdot\frac{\partial\operatorname{vec}(\tilde{\boldsymbol{L}}_t)}{\partial\operatorname{vec}(\boldsymbol{L}_t)^\top}+\frac{\partial\operatorname{vec}(\boldsymbol{D}_t)}{\partial\operatorname{vec}(\boldsymbol{L}_t)^\top}. \qquad (A.12)$$

Let S^D be an $N^2 \times N$ selection matrix, such that for a diagonal $N \times N$ matrix A with the $N \times 1$ vector a on the diagonal we have $\mathrm{vec}(A) = S^D \cdot a$. Then

$$\frac{\partial\,\mathrm{vec}(\boldsymbol{D}_t)}{\partial\,\mathrm{vec}(\boldsymbol{L}_t)^\top} = S^D\ \frac{\partial \mathrm{diag}(\boldsymbol{D}_t)}{\partial\,\mathrm{vec}(\boldsymbol{L}_t)^\top} = -2\,S^D\ D^2 \begin{pmatrix} \boldsymbol{\lambda}_{1,t}^\top & 0 & \cdots & 0 \\ 0 & \boldsymbol{\lambda}_{2,t}^\top & \cdots & 0 \\ \vdots & & \ddots & \vdots \\ 0 & 0 & \cdots & \boldsymbol{\lambda}_{N,t}^\top \end{pmatrix}, \tag{A.13}$$

where $\mathrm{diag}(A) \in \mathbb{R}^{N\times 1}$ holds the diagonal elements of the $N \times N$ matrix A. Similarly, we obtain

$$\frac{\partial\,\mathrm{vec}(\tilde{\boldsymbol{L}}_t)}{\partial\,\mathrm{vec}(\boldsymbol{L}_t)^\top} = \begin{pmatrix} \boldsymbol{Q}_{1,t} & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \boldsymbol{Q}_{N,t} \end{pmatrix}, \qquad \boldsymbol{Q}_{i,t} = \frac{\mathbf{I}_k}{(1-\boldsymbol{\lambda}_{i,t}^\top\boldsymbol{\lambda}_{i,t})^{1/2}} - \frac{\boldsymbol{\lambda}_{i,t}\boldsymbol{\lambda}_{i,t}^\top}{(1-\boldsymbol{\lambda}_{i,t}^\top\boldsymbol{\lambda}_{i,t})^{3/2}}, \tag{A.14}$$

for $i = 1, \ldots, N$. Note due to the special structure in (A.14), we have that

$$\left(\mathbf{I}_N \otimes \tilde{\boldsymbol{L}}_t^\top\right) \frac{\partial\,\mathrm{vec}(\tilde{\boldsymbol{L}}_t)}{\partial\,\mathrm{vec}(\boldsymbol{L}_t)^\top} = \begin{pmatrix} \tilde{\boldsymbol{L}}_t^\top \boldsymbol{Q}_{1,t} & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \tilde{\boldsymbol{L}}_t^\top \boldsymbol{Q}_{N,t} \end{pmatrix}. \tag{A.15}$$

We now turn to the last factor in (A.6) for the factor models considered in this paper.

A.2 1-Equi-Factor

In the 1-Factor equi-copula, we have $\boldsymbol{L}_t = \lambda_t \boldsymbol{\iota}_N^\top$ and $\boldsymbol{\lambda}_t = (\lambda_t) \in \mathbb{R}^{1\times 1}$. We then have

$$\frac{\partial\,\mathrm{vec}(\boldsymbol{L}_t)}{\partial \boldsymbol{\lambda}_t^\top} = \boldsymbol{\iota}_N. \tag{A.16}$$

The result is now obtained by combining (A.16) above, (A.6), (A.8) or (A.10), and (A.12)–(A.15).

A.3 1-Factor model with heterogeneous loadings

In this case we have $\boldsymbol{\lambda}_t = (\lambda_{t,1}, \lambda_{t,2} \ldots, \lambda_{t,G})^\top \in \mathbb{R}^{G\times 1}$ and $\boldsymbol{L}_t = \boldsymbol{\lambda}_t^\top (S_1^{gr})^\top$ with

$$S_1^{gr} = \begin{pmatrix} \boldsymbol{\iota}_{N_1} & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \boldsymbol{\iota}_{N_G} \end{pmatrix} \in \mathbb{R}^{N\times G}, \tag{A.17}$$

where N_g for $g = 1, \ldots, G$ is the number of firms in group g. We then have

$$\frac{\partial \operatorname{vec}(\boldsymbol{L}_t)}{\partial \boldsymbol{\lambda}_t^\top} = \frac{\partial \operatorname{vec}\left(\boldsymbol{\lambda}_t^\top (S_1^{gr})^\top\right)}{\partial \boldsymbol{\lambda}_t^\top} = S_t^{gr} \frac{\partial \operatorname{vec}(\boldsymbol{\lambda}_t^\top)}{\partial \boldsymbol{\lambda}_t^\top} = S_t^{gr} \frac{\partial \operatorname{vec}(\boldsymbol{\lambda}_t)}{\partial \boldsymbol{\lambda}_t^\top} = S_t^{gr} . \tag{A.18}$$

A.4 2-factor model

The 2F model consists of an equi-loading vector, and a set of heterogeneous loadings. In this case we have $\boldsymbol{\lambda}_t = (\lambda_{t,0}, \lambda_{t,1}, \lambda_{t,2}, \ldots, \lambda_{t,G})^\top \in \mathbb{R}^{(G+1)\times 1}$. Let $\delta_{i,j}$ be the kronecker delta, i.e., $\delta_{i,j} = 1$ if $i = j$ and zero otherwise. Also define

$$S_i^{2f} = \begin{pmatrix} \delta_{0,i}\boldsymbol{\iota}_{N_1} & \delta_{1,i}\boldsymbol{\iota}_{N_1} \\ \vdots & \vdots \\ \delta_{0,i}\boldsymbol{\iota}_{N_G} & \delta_{G,i}\boldsymbol{\iota}_{N_G} \end{pmatrix} \in \mathbb{R}^{N\times 2}. \tag{A.19}$$

Then

$$\boldsymbol{L}_t = \sum_{i=0}^{G} \lambda_{t,i} \cdot (S_i^{2f})^\top, \tag{A.20}$$

and

$$\frac{\partial \operatorname{vec}(\boldsymbol{L}_t)}{\partial \boldsymbol{\lambda}_t^\top} = \Big(\operatorname{vec}\left((S_0^{2f})^\top\right) , \operatorname{vec}\left((S_1^{2f})^\top\right) , \ \ldots \ , \operatorname{vec}\left((S_G^{2f})^\top\right) \Big) \in \mathbb{R}^{2N\times(G+1)}. \tag{A.21}$$

A.5 MF model

The MF model consists of two types of factors: an equi-factor, and G industry factors, each with a group-specific loading. In this case we have $\boldsymbol{\lambda}_t = (\lambda_{t,0}, \lambda_{t,1}, \lambda_{t,2}, \dots, \lambda_{t,G})^\top \in \mathbb{R}^{(G+1)\times 1}$. Define

$$S_i^{mf} = \begin{pmatrix} \delta_{0,i}\boldsymbol{\iota}_{N_1} & \delta_{1,i}\boldsymbol{\iota}_{N_1} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ \delta_{0,i}\boldsymbol{\iota}_{N_G} & 0 & \cdots & \delta_{G,i}\boldsymbol{\iota}_{N_G} \end{pmatrix} \in \mathbb{R}^{N\times(G+1)}. \tag{A.22}$$

Then

$$\boldsymbol{L}_t = \sum_{i=0}^{G} \lambda_{t,i} \cdot (S_i^{mf})^\top, \tag{A.23}$$

and

$$\frac{\partial\,\mathrm{vec}(\boldsymbol{L}_t)}{\partial \boldsymbol{\lambda}_t^\top} = \Big(\ \mathrm{vec}\left((S_0^{mf})^\top\right) \ , \ \mathrm{vec}\left((S_1^{mf})^\top\right) \ , \ \dots \ , \ \mathrm{vec}\left((S_G^{mf})^\top\right) \ \Big) \in \mathbb{R}^{(G+1)N\times(G+1)}. \tag{A.24}$$

A.6 MF LT model

For the MF LT model, we have $\boldsymbol{\lambda}_t = (\lambda_{t,1}, \lambda_{t,2}, \dots, \lambda_{t,G(G+1)/2})^\top \in \mathbb{R}^{G(G+1)/2\times 1}$. Also define

$$S_i^{lt} = \begin{pmatrix} \delta_{1,i}\boldsymbol{\iota}_{N_1} & 0 & \cdots & 0 \\ \delta_{2,i}\boldsymbol{\iota}_{N_2} & \delta_{3,i}\boldsymbol{\iota}_{N_2} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ \delta_{\frac{1}{2}G(G-1)+1,i}\boldsymbol{\iota}_{N_G} & \delta_{\frac{1}{2}G(G-1)+2,i}\boldsymbol{\iota}_{N_G} & \cdots & \delta_{\frac{1}{2}G(G+1),i}\boldsymbol{\iota}_{N_G} \end{pmatrix} \in \mathbb{R}^{N\times G}. \tag{A.25}$$

Then

$$\boldsymbol{L}_t = \sum_{i=1}^{G(G+1)/2} \lambda_{t,i} \cdot (S_i^{lt})^\top, \tag{A.26}$$

and

$$\frac{\partial\,\mathrm{vec}(\boldsymbol{L}_t)}{\partial \boldsymbol{\lambda}_t^\top} = \Big(\;\mathrm{vec}\left((S_1^{lt})^\top\right)\;,\;\mathrm{vec}\left((S_2^{lt})^\top\right)\;,\;\ldots\;,\;\mathrm{vec}\left((S_{\frac{1}{2}G(G+1)}^{lt})^\top\right)\;\Big) \in \mathbb{R}^{G\,N\times\frac{1}{2}G(G+1)}. \tag{A.27}$$

B Sample composition

B.1 Sample composition

Table B.1: Selected S&P500 constituents

This table lists ticker symbols of the 100 stocks in our dataset. All stocks are included in the S&P 500 index. Tickers are grouped per industry.

Ind Nr.	Industry	# Comp.	Tickers
1	Capital Goods	10	AA,BA,CAT,HON,F,NOC,UTX,A,IR,GD
2	Financials	19	AXP,JPM,AIG,BAC,C,KEY,MTB,COF,USB, BBT,STI,WFC,GS,MS,MMC,HIG,PNC, XL,MCO
3	Energy	12	GE,XOM,BHI,MUR,SLB,CVX,HAL,OXY, APC,SU,CNX,PXD
4	Consumer Services	14	HD,MCD,WMT,TGT,BXP,DIS,JCP,NLY, ANF,EQR,WY,RCL,WSM,TV
5	Consumer Non-Durables	9	KO,MO,SYY,PEP,CL,AVP,GIS,CPB,EL
6	Health Care	11	PFE,ABT,BAX,JNJ,LLY,THC,MMM,MRK,BMY, MDT,CI
7	Public Utilities	7	AEP,AEE,DUK,SO,WMB,VZ,EXC
8	Technology	5	IBM,DOV,HPQ,TSM,CSC
9	Basic Industries	9	PG,DD,FLR,DOW,AES,IP,ATI,LPX,POT
10	Transportation	4	LUV,UPS,NSC,FDX

B.2 Full simulation results

This supplementary appendix presents the full details of the three Monte Carlo experiments from Section 3.

In the first experiment, we simulate $N = 100$ dimensional time series of length $T = 1,000$ with $G = 10$ equally sized groups holding $N/G = 10$ individual cross-sectional units each. These sizes roughly correspond to the data dimensions in our empirical application.

As our data-generating process (DGP), we take the Multi-Factor copula

$$
\begin{aligned}
x_{i,t} &= \sqrt{\zeta_t}\left(\tilde{\boldsymbol{\lambda}}_{i,t}^{\top}\boldsymbol{z}_t + \sigma_{i,t}\epsilon_{i,t}\right), && \text{(B.1)}\\
\tilde{\boldsymbol{L}}_t &= \begin{pmatrix} \tilde{\lambda}_{1,t}^{eq} & \tilde{\lambda}_{2,1,t}^{gr,f} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ \tilde{\lambda}_{1,t}^{eq} & 0 & \cdots & \tilde{\lambda}_{2,G,t}^{gr,f} \end{pmatrix} \otimes \boldsymbol{\iota}_{N/G} \\
\lambda_{1,t+1}^{eq} &= \omega^{eq} + A^{eq}\, s_t^{eq} + B\,\lambda_{1,t}^{eq}, && \text{(B.2)}\\
\lambda_{2,g,t+1}^{gr,f} &= \omega_g^{gr,f} + A^{gr,f}\, s_{g,t}^{gr,f} + B\,\lambda_{2,g,t}^{gr,f}, \qquad g = 1,\ldots,G, && \text{(B.3)}
\end{aligned}
$$

with $\otimes$ denoting the Kronecker product, and where $\boldsymbol{z}_t \sim \mathrm{N}(\mathbf{0}, \mathbf{I}_{G+2})$, $\epsilon_{i,t} \sim \mathrm{N}(0,1)$, and $\zeta_t \sim \text{Inv-Gamma}\left(\frac{1}{2}\nu_C, \frac{1}{2}\nu_C\right)$, where the mapping from $\boldsymbol{\lambda}_t$ to $\tilde{\boldsymbol{\lambda}}_t$ is given in (3). The expressions for the scores s_t^{eq} and $s_{g,t}^{gr,f}$ can be found in Supplementary Appendix A.

The Multi-Factor copula model has two different types of factor loadings, each with its own score-driven dynamics: one $\lambda_{1,t}^{eq}$ for the common equi-factor, and G different $\lambda_{2,g,t}^{gr,f}$s for each of the group-specific factors. Each of these 11 loadings has its own intercept. We use a pooled persistence parameter B common to all factor loadings, and type-specific score parameters A^{eq} and $A^{gr,f}$.

Guided by the empirical application, we set $\omega^{eq} = 0.07$ and let ω_g be equally spaced on the interval $[0.01\,, 0.07]$. For the Gaussian copulas, we set $A^{eq} = 0.0085$ and $A^{gr,f} = 0.0095$, while for the t-copula these parameters equal 0.015 and 0.01, respectively. For the copula's tail behavior, we use $\nu_C \in \{35, \infty\}$, where $\nu_C \to \infty$ corresponds to the Gaussian factor copula. Finally, in line with our empirical results later on we set $B = 0.87$ for normally distributed factors ($\nu_C \to \infty$) and $B = 0.92$ for the Student's t case ($\nu_C = 35$).

Table B.2 presents the results based on 1,000 replications. All parameters are estimated near their true values. The standard deviations decrease in T. We also observe that the mean of the estimated standard error over all simulation runs matches closely the Monte-Carlo standard error of the estimates, indicating that computed standard errors fairly reflect estimation uncertainty. Overall, we conclude that the parameters of the Gaussian and Student's t factor copulas with score-driven dynamic factor loadings can be accurately estimated if the model is correctly specified.

In the second Monte Carlo experiment, we investigate the two-step approach of estimating the copula parameters of the Multi-Factor LT model. For this study, we simulate 1,000 time-series of length $T = 1,000$ and dimension $N = 100$ with $G = 10$ equally sized groups holding $N/G = 10$ assets using the MF-LT model with Normal and Student's $t(35)$ distributed errors. Based on empirical parameter estimates, we set A and B equal to 0.015 and 0.97 respectively and allow for $(10 \times 11)/2 = 55$ different ω parameters, ranging from -0.10 to 0.9. Table B.3 presents the results based on 1,000 replications. The Monte Carlo averages of almost all parameters again lie close to their true values. Note that the standard deviations of moment-based estimators for ω are considerably higher than the standard errors of the ML estimators for A,B, and ν_C. Using the two-step estimator thus implies a huge computational gain, but at the cost of some efficiency loss. The average estimated standard errors for A, B for the MF-LT N and for the MF-LT t model again lie close to their Monte-Carlo counterparts, such that standard errors correctly reflect the estimation uncertainty. We further note that the assumed distribution does not have a large impact on the moment estimator of ω.

Table B.2: Monte Carlo results of parameter estimates of the Multi-Factor-Copula
This table provides Monte Carlo results of parameter estimates using the multi-factor (MF) Gaussian and t-copula model as given in (B.1)–(B.3). $B(N)$ and $B(t)$ denote the value of B in case of the Gaussian (N) and Student's t (t) factor copula model, respectively. The table reports the mean and standard deviation of the estimated coefficients, as well as the mean of the computed standard error. Results are based on 1,000 Monte Carlo replications.

Panel A: T = 500							
		MF N			MF t		
Coef.	True	mean	std	mean(s.e.)	mean	std	mean(s.e.)
ω_{eq}	0.0700	0.0761	0.0145	0.0159	0.0752	0.0121	0.0127
ω_1	0.0100	0.0094	0.0074	0.0097	0.0097	0.0050	0.0051
ω_2	0.0167	0.0172	0.0073	0.0076	0.0179	0.0044	0.0042
ω_3	0.0233	0.0251	0.0070	0.0073	0.0250	0.0049	0.0049
ω_4	0.0300	0.0326	0.0076	0.0081	0.0323	0.0060	0.0059
ω_5	0.0367	0.0400	0.0090	0.0093	0.0395	0.0068	0.0070
ω_6	0.0433	0.0471	0.0097	0.0106	0.0468	0.0079	0.0081
ω_7	0.0500	0.0542	0.0108	0.0120	0.0540	0.0088	0.0092
ω_8	0.0567	0.0617	0.0125	0.0134	0.0612	0.0100	0.0104
ω_9	0.0633	0.0691	0.0135	0.0149	0.0684	0.0111	0.0115
ω_{10}	0.0700	0.0763	0.0150	0.0162	0.0755	0.0121	0.0127
$A_{eq}(N)$	0.0085	0.0085	0.0012	0.0011			
$A_{gr,f}(N)$	0.0095	0.0089	0.0027	0.0025			
$A_{eq}(t)$	0.0150				0.0146	0.0027	0.0026
$A_{gr,f}(t)$	0.0100				0.0090	0.0025	0.0023
$B(N)$	0.8700	0.8584	0.0269	0.0296			
$B(t)$	0.9200				0.9136	0.0137	0.0144
ν_C	35.000				35.445	2.828	2.702
Panel B: T = 1000							
ω_{eq}	0.0700	0.0739	0.0119	0.0134	0.0744	0.0115	0.0112
ω_1	0.0100	0.0094	0.0058	0.0068	0.0104	0.0033	0.0032
ω_2	0.0167	0.0173	0.0050	0.0054	0.0176	0.0036	0.0033
ω_3	0.0233	0.0246	0.0051	0.0057	0.0248	0.0043	0.0042
ω_4	0.0300	0.0314	0.0059	0.0065	0.0319	0.0052	0.0052
ω_5	0.0367	0.0387	0.0069	0.0077	0.0389	0.0063	0.0061
ω_6	0.0433	0.0458	0.0078	0.0088	0.0461	0.0073	0.0071
ω_7	0.0500	0.0529	0.0088	0.0100	0.0531	0.0084	0.0081
ω_8	0.0567	0.0599	0.0100	0.0112	0.0602	0.0094	0.0091
ω_9	0.0633	0.0667	0.0110	0.0123	0.0673	0.0103	0.0101
ω_{10}	0.0700	0.0738	0.0121	0.0136	0.0744	0.0115	0.0112
$A_{eq}(N)$	0.0085	0.0085	0.0009	0.0008			
$A_{gr,f}(N)$	0.0095	0.0093	0.0018	0.0018			
$A_{eq}(t)$	0.0150				0.0149	0.0020	0.0019
$A_{gr,f}(t)$	0.0100				0.0096	0.0016	0.0016
$B(N)$	0.8700	0.8626	0.0221	0.0248			
$B(t)$	0.9200				0.9149	0.0129	0.0126
ν_C	35.00				35.1760	1.8629	1.8821

Table B.3: Monte Carlo results of parameter estimates of the MF-LT model

This table provides Monte Carlo results of parameter estimates using the multi-factor (MF) LT Gaussian and t-copula model with a loading matrix given in (6). The table reports the mean and standard deviation based on 1,000 Monte Carlo replications. Since we have 55 different values of ω, we only report $\omega_1, \omega_4, \ldots, \omega_{55}$ in addition to A, B and ν_C.

		MF-LT N		MF-LT t	
Coef.	True	mean	std	mean	std
ω_1	0.893	0.884	0.0530	0.886	0.0539
ω_4	0.621	0.606	0.0538	0.607	0.0541
ω_7	0.560	0.550	0.0526	0.551	0.0511
ω_{10}	0.845	0.832	0.0524	0.830	0.0529
ω_{13}	0.187	0.185	0.0485	0.186	0.0495
ω_{16}	0.146	0.144	0.0491	0.149	0.0490
ω_{19}	0.119	0.122	0.0478	0.127	0.0487
ω_{22}	0.009	0.012	0.0473	0.013	0.0475
ω_{25}	0.003	0.007	0.0492	0.005	0.0477
ω_{28}	0.310	0.385	0.0434	0.384	0.0397
ω_{31}	0.163	0.127	0.0508	0.127	0.0511
ω_{34}	0.156	0.119	0.0483	0.123	0.0475
ω_{37}	0.250	0.229	0.0477	0.230	0.0479
ω_{40}	0.026	0.041	0.0466	0.042	0.0458
ω_{43}	0.011	0.011	0.0453	0.007	0.0479
ω_{46}	0.591	0.626	0.0447	0.622	0.0435
ω_{49}	-0.020	-0.009	0.0466	-0.008	0.0457
ω_{52}	0.016	0.017	0.0491	0.018	0.0493
ω_{55}	0.347	0.504	0.0436	0.497	0.0431
A	0.015	0.016	0.0006	0.016	0.0007
B	0.970	0.970	0.0025	0.970	0.0024
ν_C	30.00			35.06	1.862

C Recap of benchmark MGARCH models

In this appendix we give a brief recap of the MGARCH benchmark models we use, in particular the cDCC model (Engle, 2002) (with the correction of Aielli (2013)) and the (Block) DECO model of Engle and Kelly (2012) in high dimensions. To maintain a fair comparison between both classes of models, we also cast the MGARCH models into a copula framework. Hence the innovations in these models are $x_{i,t} = P^{-1}(u_{i,t})$, with $u_{i,t}$ estimated in a first step by the same marginals, and $P^{-1}(\,\cdot\,)$ the inverse marginal CDF corresponding to the copula specification at hand.

The cDCC model is given by

$$\begin{aligned} \boldsymbol{Q}_{t+1} &= \boldsymbol{\Omega} + A\,\boldsymbol{Q}_t^{*}\boldsymbol{x}_t\boldsymbol{x}_t^{\top}\boldsymbol{Q}_t^{*} + B\,\boldsymbol{Q}_t \qquad (\text{C.1}) \\ \boldsymbol{R}_t^{cDCC} &= \boldsymbol{Q}_t^{*-1}\boldsymbol{Q}_t\boldsymbol{Q}_t^{*-1} \end{aligned}$$

with $\boldsymbol{Q}_t^{*}$ a diagonal matrix with entries $q_{ii,t}$, A and B scalars and $\boldsymbol{\Omega}$ a $N \times N$ matrix. The DECO model assumes that the dependence between all assets is the same (equi-dependence) and takes the average of all pairwise DCC correlations:

$$\boldsymbol{R}_t^{DECO} = \rho_t \boldsymbol{J}_{N\times N} + (1-\rho_t)\boldsymbol{I}_N \qquad (\text{C.2})$$

$$\rho_t = \frac{1}{N(N-1)}(\boldsymbol{\iota}^{\top}\boldsymbol{R}_t^{cDCC}\boldsymbol{\iota} - N) \qquad (\text{C.3})$$

where $\boldsymbol{J}_{N\times N}$ denotes a $N \times N$ matrix of ones. As noted earlier, the DECO model corresponds to a one-factor model, though the DECO and score-driven dynamics are different.

A third variant is the Block DECO model that allows for different intra-block correlations $\rho_{g,g}$, and inter-block correlations $\rho_{g,h}$ with $g \neq h$. Similar to the multi-factor models, the size of each block may differ. The Block DECO correlation matrix is defined as

$$\begin{aligned}\boldsymbol{R}_t^{BL-DECO} &= \begin{pmatrix} (1-\rho_{1,1,t})\boldsymbol{I}_{n_1} & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & (1-\rho_{G,G,t})\boldsymbol{I}_{n_G} \end{pmatrix} \\ &+ \begin{pmatrix} \rho_{1,1,t}\boldsymbol{J}_{n_1\times n_1} & \cdots & \rho_{1,G,t}\boldsymbol{J}_{n_1\times n_G} \\ \vdots & \ddots & \vdots \\ \rho_{1,G,t}\boldsymbol{J}_{n_G\times n_1} & \cdots & \rho_{G,G,t}\boldsymbol{J}_{n_G\times n_G} \end{pmatrix}. \end{aligned} \tag{C.4}$$

The Block DECO model allows for G distinct within-group correlations $\rho_{g,g,t}$, $g = 1,\ldots,G$ as well as for $G(G-1)/2$ unique (off-diagonal) between-group correlations $\rho_{g,h,t}$ for $g \neq h$. The dynamic correlations are computed as

$$\rho_{g,g,t} = \frac{1}{n_g(n_g-1)} \sum_{i\in g, j\in g, i\neq j} \frac{q_{i,j,t}}{\sqrt{q_{i,i,t}q_{j,j,t}}}, \tag{C.5}$$

$$\rho_{g,h,t} = \frac{1}{n_g n_h} \sum_{i\in g, j\in h} \frac{q_{i,j,t}}{\sqrt{q_{i,i,t}q_{j,j,t}}}, \quad g \neq h, \tag{C.6}$$

where $q_{i,j,t}$ is the i,j-th element of the matrix $\boldsymbol{Q}_t$ from the cDCC model in (C.1). Put differently, the correlations of the Block DECO model are obtained by averaging all DCC correlations within each block.

Similar to the multi-factor dynamic copula models, the Block DECO model allows for different within-group and between-group correlations. This model comes with the additional flexibility: via the matrix Ω each between-group correlation has its own intercept, while in the factor copula approach the between-group correlations are spanned by a smaller set of parameters. This flexibility comes at two important costs. First, it is hard to impose ex-ante that the dynamic correlations from the Block DECO give rise to a positive definite correlation matrix. Though in practice a maximum likelihood type approach will steer the parameters away from a region where the predicted dependence matrix is indefinite, this is not guaranteed by the structure of the model. By contrast, the factor copula models with score-driven dynamics automatically ensure a positive semi-definite correlation matrix at all times, which is particularly relevant when using the model for forecasting. Second, the Block DECO model averages DCC correlations, which means that it relies heavily on the A and B parameters from the cDCC model and its unconditional $N \times N$ intercept $\mathbf{\Omega}$.

D Models for the marginals

In our main analysis, we use the univariate t-GAS volatility model of Creal et al. (2011, 2013) for the marginal distributions. That is, we assume a Student's t distribution for the individual returns $y_{i,t}$ with ν_i degrees of freedom with the following return and volatility dynamics (omitting the i for the sake of exposition)

$$y_t = \phi_0 + \sum_{j=1}^{p} \phi_j y_{t-j} + \epsilon_t, \qquad \epsilon_t \sim t(0, h_t, \nu), \tag{D.1}$$

$$h_{t+1} = \omega + \alpha\,(w_t \epsilon_t^2 - h_t) + \beta\, h_t, \qquad w_t = \frac{\nu + 1}{\nu - 2 + {h_t}^{-1} \epsilon_t^2}, \tag{D.2}$$

with h_t the conditional variance at time t. This model updates the conditional variance by the (scaled) score, i.e., the partial derivative of the log Student's t density with respect to the variance h_t. We follow Creal et al. (2011, 2013) and scale the score by the inverse conditional Fisher information matrix. The interpretation of the scaled score is highly intuitive in this model: Large values of ϵ_t^2 are downweighted by w_t, since possible outliers (jumps) might not only be attributed to an increase in variance, but also to the fat-tailed nature of the return data. The estimation results for the marginal models are summarized in Table D.1.

Table D.1: Marginal distribution parameter estimates

This table reports summaries of the maximum likelihood parameter estimates of the t-GAS volatility models in (D.1)-(D.2) for 100 daily time series of equity returns. The columns present the mean and quantiles of the cross-sectional distribution of each parameter. Data are observed over the period January 2, 2001 until December 31, 2014 ($T = 3,521$ trading days).

	Mean	5%	25%	Med	75%	95%
ϕ_0	0.027	-0.030	0.010	0.025	0.046	0.091
ϕ_1	-0.009	-0.049	-0.027	-0.008	0.008	0.026
ϕ_2	-0.012	-0.044	-0.028	-0.011	0.001	0.020
ω	0.025	0.009	0.014	0.021	0.029	0.060
α	0.091	0.062	0.077	0.088	0.104	0.129
β	0.991	0.983	0.988	0.992	0.995	0.998
ν	8.22	5.53	6.77	8.21	9.25	11.41
KS test for Student's t dist of std. residuals						
Number of rejections		5				

As a robustness check, we also considered marginal distributions based on a GARCH model with a skewed Student's t distribution for the innovations. The specification of that model is

$$
\begin{aligned}
y_t &= \phi_0 + \sum_{j=1}^{p} \phi_j y_{t-j} + \epsilon_t, \qquad \epsilon_t \sim St(0, h_t, \nu, \lambda), && \text{(D.3)}\\
h_{t+1} &= \omega + \alpha\, \epsilon_t^2 + \beta\, h_t, && \text{(D.4)}
\end{aligned}
$$

where the pdf of the skewed Student's t distribution of Hansen (1994) for a zero mean variable $z_t = (y_t - \mu_t)/\sqrt{h_t}$ with $\mu_t = \mathbb{E}_{t-1}[y_t]$ is given by

$$
f(z_t; \lambda, \nu) = \begin{cases} bc\left(1 + \frac{1}{\nu-2}\left(\frac{bz_t+a}{1-\lambda}\right)^2\right)^{-\frac{\nu+1}{2}} & \text{if} \quad z_t < -\frac{a}{b} \\ bc\left(1 + \frac{1}{\nu-2}\left(\frac{bz_t+a}{1+\lambda}\right)^2\right)^{-\frac{\nu+1}{2}} & \text{if} \quad z_t \geq -\frac{a}{b} \end{cases} \tag{D.5}
$$

with

$$
a = 4\lambda c \frac{\nu-2}{\nu-1}, \quad b^2 = 1 + 3\lambda^2 - a^2, \text{and} \quad c = \frac{\Gamma(\frac{\nu+1}{2})}{\sqrt{\pi(\nu-2)}\Gamma(\frac{\nu}{2})}
$$

such that $f(y_t|\mu_t, h_t, \nu, \lambda) = 1/h_t f(z_t; \lambda, \nu)$. Further, λ is the skewness parameter and ν again represents the degrees of freedom. A (positive) negative value of λ indicates (positive) negative skewness.

The results for these marginals are given in Table D.2 and result in qualitatively similar conclusions as the main results in Table 4 (see Appendix E).

Table D.2: Marginal distribution parameter estimates (skewed Student's t distribution)

This table reports summaries of the maximum likelihood parameter estimates of the GARCH skewed Student's t volatility models in (D.3)-(D.4) for 100 daily time series of equity returns. The columns present the mean and quantiles of the cross-sectional distribution of each parameter. Data are observed over the period January 2, 2001 until December 31, 2014 ($T = 3,521$ trading days).

	Mean	5%	25%	Med	75%	95%
ϕ_0	0.025	-0.026	0.008	0.022	0.041	0.095
ϕ_1	-0.009	-0.051	-0.030	-0.007	0.009	0.025
ϕ_1	-0.013	-0.043	-0.029	-0.013	-0.001	0.018
ω	0.026	0.008	0.014	0.021	0.029	0.064
α	0.067	0.043	0.055	0.063	0.079	0.098
β	0.924	0.887	0.909	0.928	0.939	0.950
ν	7.84	5.28	6.47	7.74	8.92	11.13
λ	-0.011	-0.071	-0.037	-0.010	0.013	0.055

KS test for Student's t dist of std. residuals	
Number of rejections	2

E Additional In-sample Factor Copula results

This appendix shows supplementary results our multi-factor Copula results with respect to three issues: 1) the non-reported different ω_i of our (multi-) factor copulas listed in Table 4, 2) parameter estimates of all factor models when the PITs are based on skewed Student's t GARCH model and 3) the sensitivity of MF-LT t model with respect to the different ordering of the industries.

The intercepts reported in Table E.1 can be further interpreted. For instance, the industry intercepts of the MF model show that the within Financial and Energy correlations are unconditionally much higher than for example within Capital Goods and Basic Industries correlations. This holds for both the Gaussian and t copula models.

The second part of this appendix holds a robustness check with respect to the assumed marginal distributions. As written in Appendix D, in our main analysis we use the univariate t-GAS volatility model of Creal et al. (2011, 2013) for the marginal distributions. In the same appendix, we also presented the results for skewed Student's t GARCH marginals. Table E.2 contains the results for all Factor and MGARCH Copula models if the PITs of the skewed t GARCH marginals are used in the copula analysis. The results confirm the analysis of the main text, and the statistical ordering of the different copula specifications.

The third part of this appendix contains a robustness check with respect to the ordering of the industries when estimating the MF-LT model. We re-estimated the MF-LT t model for 50 different random orderings to investigate the model's sensitivity to this. Table E.3 presents the average, minimum and maximum values of the estimates of A, B, and ν, as well as the log-likelihood value.

We find that the estimated parameters are very stable with respect to the ordering chosen. There appears some limited variation ($< 1\%$) in the maximized log-likelihood value, such that some further small gains in likelihood might be possible by optimizing over the ordering of the industries.

To conclude, Figure E.1 shows the fitted within and between dependencies of Capital

Table E.1: Intercept parameter estimates of (multi-) Factor Copula models

This table reports the maximum likelihood parameter estimation results of intercepts ω_i for the 1F-Group, 2F and MF copula models listed in Table 4. The different intercepts correspond to the Capital Goods, Finance, Energy, Consumer Services, Consumer Non-Durables, Health Care, Public Utilities, Technology, Basic Industries, and Transportation industries respectively. We also list again the estimated B parameter. Panel A (B) corresponds with the Gaussian (t) copula likelihood. Data are observed over the period January 2, 2001 until December 31, 2014 ($T = 3,521$ trading days).

	1F-Group		2F		MF	
Parameter	$\hat{\omega}_g$	s.e.	$\hat{\omega}_g$	s.e.	$\hat{\omega}_g$	s.e.
Panel A: Gaussian factor copulas						
ω_{eq}			0.047	(0.005)	0.042	(0.005)
$\omega_{CapGoods}$	0.025	(0.005)	0.026	(0.004)	0.011	(0.001)
ω_{Fin}	0.030	(0.006)	0.054	(0.002)	0.056	(0.008)
ω_{Energy}	0.020	(0.005)	0.011	(0.003)	0.055	(0.007)
$\omega_{ConsSer}$	0.020	(0.004)	0.017	(0.007)	0.021	(0.003)
$\omega_{ConsNon-Dur}$	0.016	(0.004)	0.013	(0.005)	0.033	(0.005)
ω_{Health}	0.018	(0.004)	0.005	(0.007)	0.029	(0.004)
$\omega_{PublUtil}$	0.017	(0.004)	0.012	(0.003)	0.041	(0.005)
ω_{Tech}	0.023	(0.005)	0.011	(0.002)	0.021	(0.003)
$\omega_{BasicInd}$	0.022	(0.005)	0.009	(0.009)	0.008	(0.001)
$\omega_{Transport}$	0.024	(0.005)	0.014	(0.005)	0.038	(0.006)
B	0.970	(0.006)	0.941	(0.004)	0.930	(0.009)
Panel B: t factor copulas						
ω_{eq}			0.004	(0.002)	0.033	(0.002)
$\omega_{CapGoods}$	0.012	(0.001)	0.002	(0.001)	0.015	(0.001)
ω_{Fin}	0.014	(0.002)	0.006	(0.002)	0.034	(0.002)
ω_{Energy}	0.010	(0.001)	-0.001	(0.001)	0.033	(0.002)
$\omega_{ConsSer}$	0.009	(0.001)	0.002	(0.001)	0.011	(0.001)
$\omega_{ConsNon-Dur}$	0.008	(0.001)	0.001	(0.001)	0.018	(0.001)
ω_{Health}	0.009	(0.001)	0.002	(0.001)	0.017	(0.001)
$\omega_{PublUtil}$	0.008	(0.001)	0.000	(0.000)	0.026	(0.001)
ω_{Tech}	0.011	(0.001)	0.003	(0.001)	0.012	(0.001)
$\omega_{BasicInd}$	0.010	(0.001)	0.002	(0.001)	0.008	(0.001)
$\omega_{Transport}$	0.011	(0.001)	0.003	(0.001)	0.020	(0.001)
B	0.986	(0.001)	0.993	(0.002)	0.957	(0.002)

Goods, Financials and Health companies according to our empirical specification against a randomly chosen order. Again, the differences are hardly noticeable. We conclude that the ordering of the groups does not materially affect our results.

Table E.2: Parameter estimates of the full sample based on skewed Student's t errors

This table reports maximum likelihood parameter estimates of various factor copula models, the (block) DECO model of Engle and Kelly (2012) and the cDCC model of Engle (2002), applied to daily equity returns of 100 assets listed at the S&P 500 index. The marginals are modeled assuming a skewed t GARCH model. We consider five different factor copula models, see Table 1 for the definition of their abbreviations. Panel A.1 presents the factor models with a Gaussian copula density, Panel A.2 presents the parameter estimates corresponding with the t-factor copula. Panel B.1 and B.2 present the estimates of the MGARCH class of models. In case of the cDCC and Block DECO models, the table shows parameters estimates obtained by the Composite Likelihood (CL) method. Standard errors are provided in parenthesis and based on the (sandwich) robust covariance matrix estimator. We report the copula log-likelihood,the Akaike Information Criteria (AIC) as well as the number of estimated parameters for all models. The sample comprises daily returns from January 2, 2001 until December 31, 2014 (3,521 observations).

Model	ω^{eq}	A^{eq}	A^{ind}	A^{gr}	B	ν	LogL	AIC	♯ para
Panel A.1: Gaussian factor copulas									
1F-Equi	0.018	0.005			0.973		66,055	-132,105	3
	(0.002)	(0.000)			(0.003)				
1F-Group				0.007	0.969		68,221	-136,419	12
				(0.001)	(0.009)				
2F	0.058	0.006		0.008	0.913		73,380	-146,733	14
	(0.007)	(0.000)		(0.001)	(0.010)				
MF	0.080	0.007	0.005		0.896		82,329	-164,630	14
	(0.004)	(0.000)	(0.001)		(0.005)				
MF-LT		0.009			0.962		**83,401**	**-166,688**	57
		(0.001)			(0.006)				
Panel A.2: t-factor copulas									
1F-Equi	0.060	0.012			0.920	34.39	69,790	-139,571	4
	(0.010)	(0.001)			(0.013)	(1.29)			
1F-Group				0.004	0.986	30.12	72,420	-144,815	13
				(0.000)	(0.000)	(0.99)			
2F	0.035	0.011		0.011	0.946	36.80	76,607	-153,184	15
	(0.004)	(0.001)		(0.001)	(0.005)	(1.52)			
MF	0.070	0.014	0.010		0.909	42.83	84,804	-169,578	15
	(0.002)	(0.001)	(0.001)		(0.003)	(1.48)			
MF-LT		0.004	0.991			34.30	**86,603**	**-173,091**	58
		(0.000)	(0.002)			(1.24)			
Panel B.1: Gaussian copula-MGARCH models									
cDCC (CL)		0.017			0.967		76,210	-142,515	4,952
		(0.001)			(0.003)				
DECO		0.031			0.957		65,034	-120,165	4,952
		(0.003)			(0.005)				
Block DECO		0.030			0.956		**83,306**	**-156,707**	4,952
		(0.002)			(0.003)				
Panel B.2: t copula-MGARCH models									
cDCC (CL)		0.018			0.967	13.92	84356	-158,807	4,953
		(0.001)			(0.003)	(0.56)			
DECO		0.038			0.949	29.97	69630	-129,354	4,953
		(0.003)			(0.005)	(1.13)			
Block DECO		0.031			0.955	21.82	**86,450**	**-162,995**	4,953
		(0.002)			(0.003)	(0.57)			

Table E.3: Estimated parameters of the MF-LT t model

This table contains summary statistics of the estimate parameters of a MF-LT t model with 50 different random ordering of groups. We show the average, minimum and maximum values of the parameters and the maximized log-likelihood over the 50 generated random orderings. The first row (*current*) corresponds with the ordering used in the paper. Results are based on the full sample.

	A	B	ν	LogL
current	0.004	0.990	36.22	86,433
mean	0.005	0.990	36.35	86,463
min	0.004	0.988	35.82	86,314
max	0.006	0.992	36.79	86,551

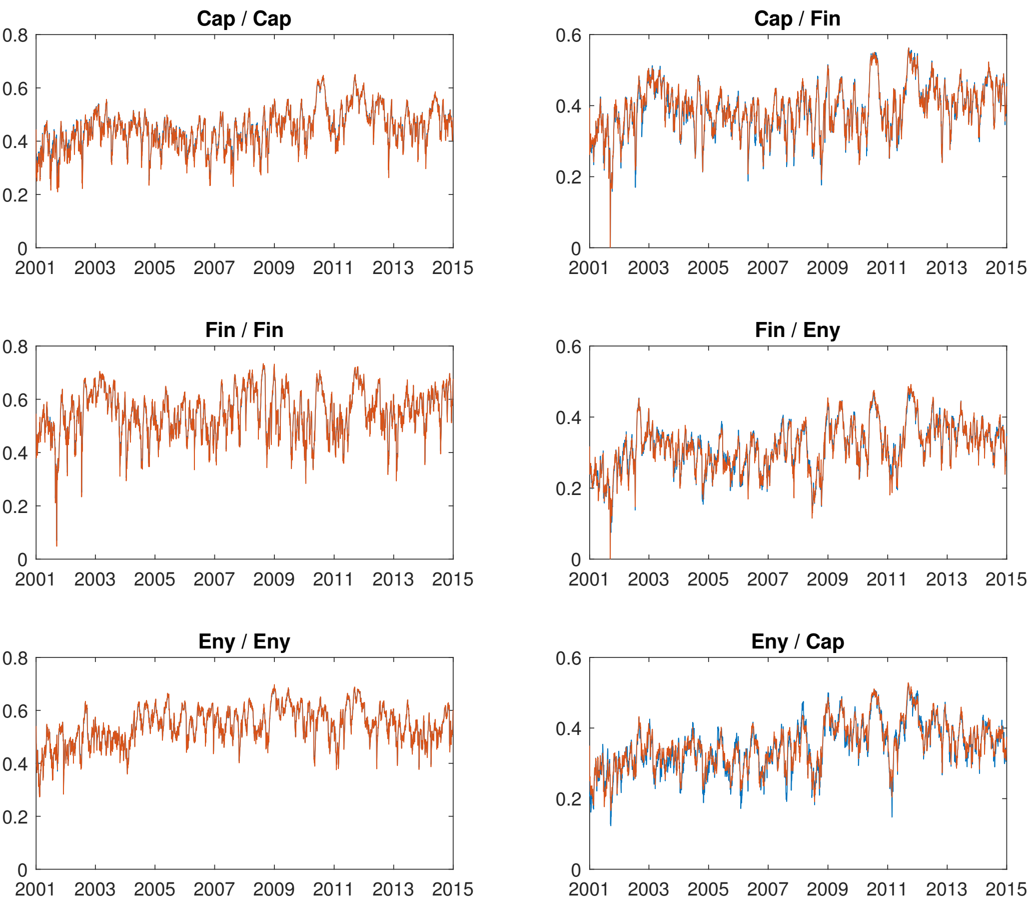

Figure E.1: Fitted dependencies of the MF-LT t model according to two different orders

This figure shows within and between dependencies of Financials, Capital Goods and Energy according to the MF-LT model. The red line is based on the group ordering used in the paper, while the blue line corresponds with a randomly chosen group ordering. The sample spans the period from January 2, 2001 until December 31, 2014 ($T = 3,521$ days).