TI 2019-013/IV
Tinbergen Institute Discussion Paper

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

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

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

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

Contact: discussionpapers@tinbergen.nl

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

Tinbergen Institute has two locations:

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

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

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 Universiteit Rotterdam*

This version: February 6, 2019

Abstract

We develop new multi-factor copula models with time-varying dependence structures via factor loadings with observation-driven dynamics. The new models are highly flexible, scalable to high dimensions, and ensure positivity of covariance and correlation matrices. The model retains a closed-form likelihood expression, thus allowing 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 appears crucial for parsimoniously describing the dependence dynamics in high-dimensional stock return data, particularly when compared to the typically used single-factor models with dynamic loadings. The new factor models also improve on recently proposed benchmarks in terms of one-step-ahead copula density forecasts and global minimum variance portfolio performance.

Keywords: factor copulas, factor structure, realized correlation, score-driven dynamics.
JEL: C32, C58, G17.

* We thank Andrew Patton and participants at the 10^{th} Annual SoFiE conference and seminar participants at Tinbergen Institute Amsterdam, Lund University, and Vrije Universiteit Amsterdam for helpful comments.
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 modern econometric 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 much scarcer, 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) provide 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 variables by a smaller set of latent variables with time-varying factor 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 so far mainly been implemented for the single-factor case; see the references above. This seems to be mainly driven by computational reasons. Though adding more factors with dynamic loadings is in principle possible, it would increase the computational burden substantially. In Oh and Patton (2018) the increased computational burden 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 new 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 more expensive as the number of factors with dynamic factor loadings grows.

Though restricting the number of factors to just a single factor 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 a multi-factor copula model based on observed characteristics. In particular, we assume that the cross-sectional units can be grouped using observable characteristics, such as the industry of the firm, its headquarters location, etc. Each of these groups is possibly subject to one or more common factors as well as to group-specific factors. 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 static parameters in our most flexible multi-factor copula model. The new multi-factor model can easily be extended to 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 compare the multi-factor copula models with three popular multivariate GARCH (MGARCH) models: the cDCC model of Engle (2002); 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 also consider economic performance of the models. We find that for our panel of equity returns, both the within-industry and between-industry dependence dynamics are key data features that need to be matched. Single-factor models and the standard DECO model have difficulty fitting all these dynamic features simultaneously. Our multi-factor specifications, by contrast, outperform all benchmarks considered in terms of density forecasts, both in-sample and one-step-ahead out-of-sample. For economic criteria, the picture changes somewhat. In that case our one-factor specification with heterogeneous dynamic loadings has the best ex-post variance of the Global Minimum Variance Portfolio (GMVP). We attribute this difference to the character of the *minimum* GMVP 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.

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 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. In Section 2, we introduce the multi-factor copula model and review the cDCC and (Block) DECO models, as well as the estimation procedures under thin-tailed and fat-tailed errors. 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.

2 The modeling framework

In this section we first describe the general copula setup in Section 2.1. In Section 2.2, we then introduce the class of closed-form dynamic factor copulas with score-driven time-varying loadings, particularly in a multi-factor setting. Section 2.3 covers the more familiar benchmark models we use from the multivariate GARCH class of copulas, in particular the DECO, Block DECO, and DCC models. Sections 2.4 and 2.5 finally discuss the specification of the marginal distributions and the details on parameter estimation, either using full likelihood, a two-step targeting approach, or composite likelihood methods.

2.1 The copula set-up

Let $\boldsymbol{y}_t = (y_{1,t}, \ldots, y_{N,t})^\top \in \mathbb{R}^N$ denote a vector of asset returns over day t, $t = 1, \ldots, T$. We decompose the conditional joint distribution $\boldsymbol{F}_t(\boldsymbol{y}_t)$ of $\boldsymbol{y}_t$ 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 $\mathcal{F}_{t-1}$ is the information set containing all information up to and including time $t-1$, $\boldsymbol{C}_t(\,\cdot\,; \boldsymbol{\theta}_{C,t})$ is the conditional copula given $\mathcal{F}_{t-1}$ and the time-varying copula parameter vector $\boldsymbol{\theta}_{C,t}$, and $F_i(y_{i,t}; \boldsymbol{\theta}_{M,i,t})$ $i = 1, \ldots, N$, 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 come back 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 $y_{i,t}$, 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 dimension N is large, splitting the modeling task into modeling the marginals and the copula may substantially reduce the computational burden as parameters can be estimated using a 2-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 $\mathbf{C}_t(\,\cdot\,;\boldsymbol{\theta}_{C,t})$ given the PITs. This can be done using factor copulas or multivariate GARCH models like the DCC or DECO models. We consider these two frameworks in the next two subsections.

2.2 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 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.

We start from the factor copula structure

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

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). The vector $\tilde{\boldsymbol{\lambda}}_{i,t}$ and scalar $\sigma_{i,t}$ are defined as

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

for an unrestricted $k\times 1$ vector $\boldsymbol{\lambda}_{i,t}$, where the $\exp(\,\cdot\,)$ operates element wise. This ensures that $x_{i,t}$ has zero mean and unit variance by design. 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

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

As k is typically much smaller than N, computing the inverse of the $k\times k$ matrix $\mathbf{I}_k+\tilde{\boldsymbol{L}}_t^\top \boldsymbol{D}_t^{-1}\tilde{\boldsymbol{L}}_t$ 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 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.

2.2.1 The factor structure

Our main goal in this paper is to develop feasible factor structures that allow for multiple factors in a flexible, dynamic way but 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).

We assume that we can split the N assets into G groups according to an observed characteristic such as industry, region, etc. Each group is subject to several factors. For the sake of exposition, we take the example of $G = 4$ groups with 2 firms in each group throughout this whole subsection. In reality, however, the number of groups and 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. The small setting of $G = 4$ with 2 firms per group, however, allows us to clearly illustrate the structure of the factor loadings matrices.

In our most general specification the loading matrix consists of a lower-triangular matrix with columns containing group-specific loadings. The loadings 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 an Cholesky decomposition of the quasi correlation matrix that contains within and between group correlations. 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 common factor with (common) equi-loadings, and a set of group-specific factors with corresponding group-specific loadings. The loading matrix changes into

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

For $G \geq 4$ and at least 2 firms in each group, this model is identified. To see this, note that the correlation matrix $\boldsymbol{R}_t$ for $G = 4$ has 4 within-group correlations and 6 between-group correlations, hence 10 free positions for the 5 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. The first (equi)factor with common loadings $\tilde{\lambda}_{0,t}$ 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 with a common factor with group-specific loadings. The loading matrix $\tilde{\boldsymbol{L}}_t^\top$ is now 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 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 dependence as $\tilde{\lambda}_{2,j,t}$ and $\tilde{\lambda}_{2,j+1,t}$ determine both. We label the model in equation (8) the 2-Factor (2F) copula model. Omitting the factor $\tilde{\lambda}_{1,t}$ in (8) leads to the 1-Factor-Group (1F-Gr) model, which consists of a single factor but with G different group loadings. This model has also been used in Lucas et al. (2017) and Oh and Patton (2018).

Finally, the MF model from equation (7) has an interesting special case if we omit the group-specific factors: we then obtain a single-factor model with common loadings. This results in a DECO correlation structure as in Engle and Kelly (2012) (see Section 2.4), 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' (factor loading) for all assets. We label this special case the 1F-Equi copula model; see also the the single-factor

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 factor 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	$N \times G$

copula structures of Oh and Patton (2018) and Creal and Tsay (2015).

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 $\tilde{\boldsymbol{L}}_t^\top$.

2.2.2 Distributional assumptions

Given the various factor structures proposed in Section 2.2.1, 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. We discuss the latter in the next subsection.

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

$$
\begin{aligned}
u_{i,t} &= T(x_{i,t};\nu_C), \qquad\qquad i=1,\dots N, \\
x_{i,t} &= \sqrt{\zeta_t}\left(\tilde{\boldsymbol{\lambda}}_{i,t}^{\top}\boldsymbol{z}_t+\sigma_{i,t}\epsilon_{i,t}\right), \qquad \boldsymbol{z}_t\sim \mathrm{N}(\mathbf{0},\mathbf{I}_p), \quad \epsilon_{i,t}\sim \mathrm{N}(0,1), \\
\zeta_t &\sim \text{Inv-Gamma}\left(\nu_C/2,\nu_C/2\right).
\end{aligned} \tag{9}
$$

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 the studies of 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 fit empirical data.

Further generalizations to (9) are easily possible by changing the distribution of ζ_t to the Generalized Inverse Gaussian distribution and adding a multiple of ζ_t to the level equation for $x_{i,t}$. This gives rise to the Generalized Hyperbolic (GH) copula that allows for both thin and fat tails as well as skewness. A special case is the GH skewed Student's t copula; see for instance Azzalini and Capitanio (2003) and Lucas et al. (2014, 2017). An initial estimate of our model with a DECO (1F-Equi) specification as in Engle and Kelly (2012) and a skewed Student's t density as in Azzalini and Capitanio (2003) only showed a marginal increase of the likelihood compared to the symmetric Student's t copula. We therefore leave such further generalizations to future research.

2.2.3 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) or 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 dynamic copula 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 likelihood fit at time t. Information theoretic optimality results for this approach can be found in Blasques et al. (2015); see also the generalizations in Creal et al. (2018).

For instance, for the 1-Factor equicorrelation model we have $\tilde{\boldsymbol{L}}_t^\top = \tilde{\lambda}_t \boldsymbol{\iota}$ for a scalar parameter $\tilde{\lambda}_t = (1+\exp(-\lambda_t))^{-1}$, where $\boldsymbol{\iota}$ denotes an $N \times 1$ vector filled with ones. The score-driven dynamics for λ_t are given by

$$\lambda_{t+1} = \omega + A\, s_t + B\, \lambda_t, \qquad s_t = \partial \log \boldsymbol{c}(\boldsymbol{x}_t; \lambda_t, \nu_C)/\partial \lambda_t, \tag{10}$$

with ω, A and B scalar parameters, and $\boldsymbol{c}(\,\cdot\,; \lambda_t, \nu_C)$ the Student's t copula density. Unless stated otherwise, we assume that A and B are scalers in our factor copula models. 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 more involved than (10) and are given in the appendix.

2.3 Copula-MGARCH class of models

We compare the dynamic factor copula models from the previous subsection against 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

$$\begin{aligned} \boldsymbol{Q}_{t+1} &= \boldsymbol{\Omega} + A\,\boldsymbol{Q}_t^*\boldsymbol{x}_t\boldsymbol{x}_t^\top\boldsymbol{Q}_t^* + B\,\boldsymbol{Q}_t \qquad (11) \\ \boldsymbol{R}_t^{cDCC} &= \boldsymbol{Q}_t^{*-1}\boldsymbol{Q}_t\boldsymbol{Q}_t^{*-1} \end{aligned}$$

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:

$$\boldsymbol{R}_t^{DECO} = \rho_t\boldsymbol{J}_{N\times N} + (1-\rho_t)\boldsymbol{I}_N \qquad (12)$$

$$\rho_t = \frac{1}{N(N-1)}(\boldsymbol{\iota}^\top\boldsymbol{R}_t^{cDCC}\boldsymbol{\iota} - N) \qquad (13)$$

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 $\rho_{g,g}$, $g = 1,...G$, 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

$$\begin{aligned}\boldsymbol{R}_t^{BL-DECO} &= \begin{pmatrix} (1-\rho_{1,1,t})\boldsymbol{I}_{n_1} & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & (1-\rho_{G,G,t})\boldsymbol{I}_{n_G} \end{pmatrix} \\ &+ \begin{pmatrix} \rho_{1,1,t}\boldsymbol{J}_{n_1\times n_1} & \cdots & \rho_{1,G,t}\boldsymbol{J}_{n_1\times n_G} \\ \vdots & \ddots & \vdots \\ \rho_{1,G,t}\boldsymbol{J}_{n_G\times n_1} & \cdots & \rho_{G,G,t}\boldsymbol{J}_{n_G\times n_G} \end{pmatrix}. \end{aligned} \tag{14}$$

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

$$\rho_{g,g,t} = \frac{1}{n_g(n_g-1)} \sum_{i\in g, j\in g, i\neq j} \frac{q_{i,j,t}}{\sqrt{q_{i,i,t}q_{j,j,t}}}, \tag{15}$$

$$\rho_{l,m,t} = \frac{1}{n_l n_m} \sum_{i\in l, j\in m} \frac{q_{i,j,t}}{\sqrt{q_{i,i,t}q_{j,j,t}}}, \quad l \neq m, \tag{16}$$

where $q_{i,j,t}$ is the i,j-th element of the matrix $\boldsymbol{Q}_t$ from the cDCC model in (11). 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 $\mathbf{\Omega}$.

2.4 Marginals

To operationalize the copula approach, we first need to model the conditional marginal distributions. We use the same marginal models for both the MGARCH copulas and the multi-factor copulas in order to ensure a clean comparison of the copula performance in high dimensions.

We model the marginal distributions by means of univariate volatility models using daily returns. To account for the fat-tailed nature of daily returns, we use the t-GAS model of Creal et al. (2011, 2013). 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:

$$y_{i,t} = \phi_{0,i} + \sum_{q=1}^{Q} \phi_{q,i} y_{i,t-q} + \epsilon_{i,t}, \qquad \epsilon_{i,t} \sim t(0, h_{i,t}, \nu_i), \tag{17}$$

$$h_{i,t+1} = \omega_i + \alpha_i \left(w_{i,t} \epsilon_{i,t}^2 - h_{i,t} \right) + \beta_i \, h_{i,t}, \qquad w_{i,t} = \frac{\nu_i + 1}{\nu_i - 2 + {h_{i,t}}^{-1} \epsilon_{i,t}^2}, \tag{18}$$

with $h_{i,t}$ the conditional variance of asset i 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 $h_{i,t}$. 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 $\epsilon_{i,t}^2$ are downweighted by $w_{i,t}$, since possible outliers (jumps) might are not only attributed to an increase in variance, but also to the fat-tailed nature of the return data.

2.5 Parameter estimation

Parameter estimation requires some further details, both for the factor copula and the MGARCH copula models. To estimate the model's 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 marginals. This approach follows directly from decomposing the joint likelihood as

$$\begin{aligned}\mathcal{L}(\boldsymbol{\theta}) \equiv & \sum_{t=1}^{T} \log \boldsymbol{f}_t(\boldsymbol{y}_t; \boldsymbol{\theta}) = \sum_{i=1}^{N} \sum_{t=1}^{T} \log f_{i,t}(y_{i,t}; \boldsymbol{\theta}_M) \\ & + \sum_{t=1}^{T} \log \boldsymbol{c}_t \left(F_{1,t}(y_{1,t}; \boldsymbol{\theta}_M), \ldots, F_{N,t}(y_{N,t}; \boldsymbol{\theta}_M); \boldsymbol{\theta}_C\right)\end{aligned} \tag{19}$$

with $\boldsymbol{\theta} = \{\boldsymbol{\theta}_M, \boldsymbol{\theta}_C\}$. 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 the dependencies. For the factor copula specifications, inverses and determinants of $\boldsymbol{R}_t$ are available in closed form, which substantially reduces the computational burden in high dimensions. It 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 values of ω. A computational challenge may arise if G increases. Assuming that the lambda process is covariance stationary, and defining the unconditional mean of λ_t as $\bar{\lambda}$, we have

$$\mathbb{E}[\lambda_{t+1}] = \omega + B\,\mathbb{E}[\lambda_t] \quad \Leftrightarrow \quad \bar{\lambda} = \frac{\omega}{1-B}, \tag{20}$$

where we omit the subscript for the sake of simplicity. We therefore estimate copula parameters in two steps. In the first step, we match each $\bar{\lambda}$ with the empirical within- and between-group correlation using a moment estimator. Let $\boldsymbol{R}_M$ denote the $G \times G$ unconditional correlation matrix based on $x_{it} = N^{-1}(u_{it})$, where the off-diagonal (i,j) equals the average correlation between assets from group i and j, and the diagonal contains the average sample correlations within each group of assets ($i = j$).[1] The moment estimator is obtained by minimizing the squared distance $L_{M,dist}$:

$$L_{M,dist} = \boldsymbol{R}_M - \bar{\boldsymbol{L}}\bar{\boldsymbol{L}}' \tag{21}$$

with $\bar{\boldsymbol{L}}'$ a $G \times G$ lower triangular matrix. In case $G = 4$, we have

$$\bar{\bar{\boldsymbol{L}}}' = \begin{pmatrix} \bar{\lambda}_{1,1} & 0 & 0 & 0 \\ \bar{\lambda}_{1,2} & \bar{\lambda}_{2,2} & 0 & 0 \\ \bar{\lambda}_{1,3} & \bar{\lambda}_{2,3} & \bar{\lambda}_{3,3} & 0 \\ \bar{\lambda}_{1,4} & \bar{\lambda}_{2,4} & \bar{\lambda}_{3,4} & \bar{\lambda}_{4,4} \end{pmatrix} \tag{22}$$

In a second step, we only have to estimate the A and B parameters. Given the estimated $\hat{\bar{\lambda}}_{i,j}$, the intercepts in the transition equation for lambda are easily obtained using (20). This two-step targeting procedure substantially decreases the computational burden, as the moment-estimator is very fast. In the second step, only the two remaining parameters A and B have to be estimated.

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 Gaussian (copula) log-likelihood values to estimate A and B. 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

[1] Note that we use the inverse Normal here. Of course, we could use the t copula as well. As we will show in the next section, even if the data is Student's t distributed, using $\boldsymbol{R}_M$ based on the inverse Normal distribution will not affect the quality of the moment estimator of ω

$i, j \in \{1 \ldots, G\}$, i.e.,

$$\mathcal{L}_{i,j}^{Stud} = \sum_{t=1}^{T} \left[-\frac{1}{2} \log |\boldsymbol{R}_t| - \frac{\nu + n_i + n_j}{2} \log \left(1 + \frac{\boldsymbol{x}_t^\top \boldsymbol{R}_t^{-1} \boldsymbol{x}_t}{\nu - 2} \right) \right], \tag{23}$$

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 $i > j$,

$$\max L_{CL} = \max \sum_{i>j} L_{i,j}^{Stud}, \tag{24}$$

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

In this section, we present three Monte Carlo experiments. The first experiment investigates the finite sample properties of the maximum likelihood estimator for $\boldsymbol{\theta}_C$ of the Multi-factor copula model. In the second experiment, we establish the finite sample properties of the 2-step moment-based estimator of the MF-LT model. Finally, we assess the accuracy of the filtered factor loadings in case of misspecficiation of the number of factors.

In the first experiment, we simulate $N = 100$ dimensional time series of length $T \in \{500, 1000\}$ 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

$$
\begin{aligned}
x_{i,t} &= \sqrt{\zeta_t}\left(\tilde{\boldsymbol{\lambda}}_{i,t}^{\top}\boldsymbol{z}_t + \sigma_{i,t}\epsilon_{i,t}\right), && (25)\\
\tilde{\boldsymbol{L}}_t &= \begin{pmatrix} \tilde{\lambda}_{1,t}^{eq} & \tilde{\lambda}_{2,1,t}^{gr,f} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ \tilde{\lambda}_{1,t}^{eq} & 0 & \cdots & \tilde{\lambda}_{2,G,t}^{gr,f} \end{pmatrix} \otimes \boldsymbol{\iota}_{N/G} \\
\lambda_{1,t+1}^{eq} &= \omega^{eq} + A^{eq}\, s_t^{eq} + B\,\lambda_{1,t}^{eq}, && (26)\\
\lambda_{2,g,t+1}^{gr,f} &= \omega_g^{gr,f} + A^{gr,f}\, s_{g,t}^{gr,f} + B\,\lambda_{2,g,t}^{gr,f}, \qquad g = 1,\ldots,G, && (27)
\end{aligned}
$$

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 the appendix.

The Multi-Factor copula model has two different types of λs, 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 $\omega^{eq} = 0.01$ and let ω_g be equally spaced on the interval $[-0.05\,, 0.01]$. We set $A^{eq} = 0.01$ and $A^{gr,f} = 0.02$. For the copula's tail behavior, we use $\nu_C \in \{30, \infty\}$, where $\nu_C \to \infty$ corresponds to the Gaussian factor copula. Finally, in line with our empirical results later on we set $B = 0.92$ for normally distributed factors ($\nu_C \to \infty$) and $B = 0.97$ for the Student's t case ($\nu_C = 30$).

Table 2 presents the results based on 1,000 replications. All parameters are estimated near their true values and the standard deviation decreases in general when the sample size T increases. There seems to be a small bias away from zero in the group specific intercepts $\omega_i^{gr,f}$ for $T = 500$. The bias shrinks when the sample size increases to $T = 1,000$. 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

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 (25)–(27). $B(N)$ 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 in parentheses based on 1,000 Monte Carlo replications.

		T = 500		T = 1,000	
Coef.	True	MF-N	MF-t	MF-N	MF-t
ω^{eq}	0.010	0.012 (0.005)	0.011 (0.003)	0.011 (0.003)	0.011 (0.002)
$\omega_1^{gr,f}$	-0.050	-0.058 (0.014)	-0.057 (0.011)	-0.054 (0.009)	-0.055 (0.009)
$\omega_2^{gr,f}$	-0.043	-0.051 (0.012)	-0.049 (0.010)	-0.047 (0.008)	-0.047 (0.007)
$\omega_3^{gr,f}$	-0.037	-0.043 (0.010)	-0.042 (0.009)	-0.039 (0.007)	-0.040 (0.006)
$\omega_4^{gr,f}$	-0.030	-0.035 (0.009)	-0.034 (0.007)	-0.032 (0.006)	-0.033 (0.005)
$\omega_5^{gr,f}$	-0.023	-0.027 (0.007)	-0.027 (0.006)	-0.025 (0.005)	-0.026 (0.004)
$\omega_6^{gr,f}$	-0.017	-0.019 (0.006)	-0.019 (0.005)	-0.018 (0.004)	-0.018 (0.003)
$\omega_7^{gr,f}$	-0.010	-0.012 (0.005)	-0.011 (0.003)	-0.011 (0.003)	-0.011 (0.003)
$\omega_8^{gr,f}$	-0.003	-0.004 (0.005)	-0.004 (0.003)	-0.004 (0.003)	-0.004 (0.002)
$\omega_9^{gr,f}$	0.003	0.004 (0.005)	0.004 (0.003)	0.004 (0.003)	0.004 (0.002)
$\omega_{10}^{gr,f}$	0.010	0.012 (0.005)	0.011 (0.003)	0.011 (0.003)	0.011 (0.002)
A^{eq}	0.010	0.010 (0.001)	0.010 (0.002)	0.010 (0.001)	0.010 (0.001)
$A^{gr,f}$	0.020	0.019 (0.004)	0.018 (0.004)	0.019 (0.003)	0.019 (0.003)
$B(N)$	0.920	0.907 (0.020)		0.915 (0.013)	
$B(t)$	0.970		0.966 (0.007)		0.967 (0.005)
ν_C	30		30.37 (2.33)		30.27 (1.50)

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 $t(30)$ distributed errors. Based on empirical parameter estimates, we set A and B equal to 0.025 and 0.95 respectively and allow for $(10 \times 11)/2 = 55$ different ω parameters, ranging from -1.4 to 0.2.

Table 3 presents the results based on 1,000 replications. The Monte Carlo averages of all estimated ω parameters 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. We further note that the assumed distribution does not have a large impact on the moment estimator of ω.

In the final Monte Carlo experiment, we assess the impact of misspecification of the factor

Table 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 in parentheses 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.

Coef.	True	MF-LT N	MF-LT t
ω_1	-0.101	-0.101 (0.052)	-0.099 (0.053)
ω_4	-0.480	-0.476 (0.057)	-0.474 (0.061)
ω_7	-0.596	-0.591 (0.081)	-0.592 (0.083)
ω_{10}	-0.189	-0.189 (0.078)	-0.188 (0.084)
ω_{13}	-1.114	-1.092 (0.080)	-1.090 (0.082)
ω_{16}	-0.857	-0.848 (0.103)	-0.848 (0.104)
ω_{19}	-0.600	-0.594 (0.116)	-0.599 (0.112)
ω_{22}	-1.179	-1.161 (0.088)	-1.157 (0.088)
ω_{25}	-1.115	-1.094 (0.128)	-1.097 (0.134)
ω_{28}	-1.052	-1.006 (0.056)	-1.007 (0.057)
ω_{31}	-0.988	-0.976 (0.117)	-0.976 (0.113)
ω_{34}	-0.924	-0.924 (0.159)	-0.917 (0.168)
ω_{37}	-0.861	-0.848 (0.099)	-0.845 (0.102)
ω_{40}	-0.797	-0.790 (0.142)	-0.789 (0.152)
ω_{43}	-0.733	-0.727 (0.094)	-0.727 (0.095)
ω_{46}	-0.670	-0.653 (0.055)	-0.656 (0.056)
ω_{49}	-0.606	-0.606 (0.093)	-0.604 (0.094)
ω_{52}	-0.542	-0.548 (0.079)	-0.556 (0.083)
ω_{55}	-0.780	-0.713 (0.057)	-0.717 (0.059)
A	0.025	0.025 (0.002)	0.025 (0.002)
B	0.950	0.945 (0.005)	0.944 (0.006)
ν_C	30		30.45 (1.636)

structure on the fitted dependence structure. For this study, we simulate 1,000 time-series of length $T = 1,000$ and dimension $N = 25$ using the MF model with Student's $t(30)$ distributed errors. We use $G = 5$ different groups, each containing $N/G = 5$ cross-sectional units. As our statistical model, we use each of the factor copula specifications from Table 1 based on either the Gaussian or Student's $t(\nu_C)$ distribution. Using the statistical models, we estimate the model-implied dependence structure from the fitted correlation matrices $\hat{\boldsymbol{R}}_t$. These can be compared to the true $\boldsymbol{R}_t$ in each simulation run. We measure the fit by the average squared Euclidian distance

$$Ł = \frac{1}{T} \sum_{t=1}^{T} \text{vech}\,(\hat{\boldsymbol{R}}_t - \boldsymbol{R}_t)^\top \text{vech}\,(\hat{\boldsymbol{R}}_t - \boldsymbol{R}_t), \tag{28}$$

Table 4: Performance of misspecified factor copulas
This table summarizes the mean and standard deviation of the average Euclidian distance between a simulated correlation matrix $\boldsymbol{R}_t$ ($t = 1, \ldots, 1000$) from the MF-model with $t(30)$-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	2F	1F-Group	1F-Equi
Student's t	0.035	2.104	2.879	2.606
	(0.023)	(0.208)	(0.186)	(0.155)
Gaussian	0.059	1.944	2.655	2.663
	(0.027)	(0.156)	(0.170)	(0.175)

which is a consistent loss function according to Laurent et al. (2013).

Table 4 presents the mean and the standard deviation of the discrepancies between $\boldsymbol{R}_t$ and $\hat{\boldsymbol{R}}_t$ over all 1,000 replications. The distributional assumptions appear to have little effect: differences all fall within the simulation error bands. However, given the MF copula DGP with one equi-factor and five industry factors, the effect of underestimating the number of the factors can be substantial. The models with one or two factors seem to result in substantially different dynamic dependence structures than the DGP. This clearly shows that the effect of misspecification can be substantial if one sticks to one-factor or two-factor copula models when the DGP has more factors. It is therefore important to allow for more factors in our empirical work than the typical one-factor implementations as found in Creal and Tsay (2015) and Oh and Patton (2018).

4 Empirical application

4.1 Data

In our high-dimensional empirical application we investigate the daily open-to-close returns of 100 constituents of the S&P 500. These 100 stocks are randomly chosen from different industries. Table 5 provides an overview of the tickers of each company, grouped into 10 different industries. The data spans the period January 2, 2001 until December 31, 2014 and contains $T = 3,521$ days. The Financial industry covers most companies (i.e. 19), followed by

Table 5: S&P 500 constituents

This table lists ticker symbols of 100 companies listed at the S&P 500 index during the period January 2, 2001 until December 31, 2014. All 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	Finance	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

the Consumer Services and Energy industries, respectively. Each industry covers at least four companies.

To model the marginal behavior, we estimate a univariate t GAS volatility model from equation (17) for each of the 100 series. For the conditional mean, we find significance for at most the first two AR lags for some of the return series and therefore set $Q = 2$ in (17). Table 6 shows the cross-sectional mean and quantiles of the estimated parameters of the marginal distribution across all 100 stocks. The mean value of $\nu = 8.22$ underlines the fat-tailed nature of stock returns, even after filtering for time-varying 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 $\bar{x}_{i,t} = \Phi^{-1}(\hat{u}_{i,t})$. We subsequently test each series $\bar{x}_{i,t}$, $t = 1, \ldots, T$, for normality using the Kolmogorov-Smirnov test. Across the 100 firms, we only reject the null hypothesis for 5 series. We conclude that the marginal distributional assumptions are adequate for our subsequent analysis.

Table 6: Marginal distribution parameter estimates

This table reports summaries of the maximum likelihood parameter estimates of the GAS t volatility models of (17) for 100 different 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			

4.2 Full-sample comparison

After estimating the parameters of the marginal distributions, we estimate the parameters of the score-driven factor copula models as well as the benchmark MGARCH models using the full sample of 3,521 observations. The groups for the factor copulas are based on the firm industries as laid out in Table 5.

Table 7 shows the parameter estimates and maximized log-likelihood values for all models. The A.1 and A.2 panels contain the results for our 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 loadings matrix, allowing for cross-exposures between groups of the group factors. The B.1 and B.2 panels contain the Gaussian and t benchmark copulas from the MGARCH class, respectively: the cDCC model, the DECO model, and the Block DECO model. In both multi-factor copula models, we assume that the B parameter is the same for all factors. While in the MF-LT model we also assume a common A parameter for all factors, the A parameter is allowed to differ between the common factor and the industry-specific factors in the MF model. Within the group of 10 industry factors,

Table 7: 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 equity returns of 100 assets listed at 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 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, as well as the Akaike Information Criteria (AIC) 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
Panel A.1: Gaussian factor copula's								
1F-Equi	-0.037	0.008			0.890		65,991	-131,976
	(0.007)	(0.001)			(0.019)			
1F-Group				0.013	0.971		68,133	-136,241
				(0.001)	(0.006)			
2F	-0.027	0.009		0.029	0.929		73,295	-146,562
	(0.001)	(0.000)		(0.001)	(0.002)			
MF	-0.026	0.009	0.033		0.914		82,422	-164,815
	(0.005)	(0.001)	(0.002)		(0.012)			
MF-LT		0.029			0.933		**83,319**	**-166,525**
		(0.001)			(0.008)			
Panel A.2: t-factor copula's								
1F-Equi	-0.024	0.024			0.925	36.49	69,699	-139,391
	(0.005)	(0.002)			(0.013)	(1.91)		
1F-Group				0.009	0.984	32.01	72,309	-144,592
				(0.001)	(0.003)	(1.04)		
2F	-0.032	0.030		0.023	0.983	38.62	75,686	-151,342
	(0.008)	(0.003)		(0.002)	(0.001)	(1.62)		
MF	-0.011	0.023	0.030		0.958	45.27	84,952	-169,873
	(0.004)	(0.002)	(0.003)		(0.010)	(2.07)		
MF-LT		0.019			0.969	37.46	**86,346**	**-172,575**
		(0.001)			(0.005)	(2.51)		
Panel B.1: Gaussian copula-MGARCH models								
cDCC (CL)		0.017			0.968		74,263	-138,623
		(0.000)			(0.001)			
DECO		0.071			0.929		64,474	-119,044
		(0.001)			(0.001)			
Block DECO		0.030			0.957		**83,087**	**-156,270**
		(0.000)			(0.000)			
Panel B.2: t copula-MGARCH models								
DECO		0.106			0.894	34.43	69,314	-128,721
		(0.000)			(0.000)	(0.80)		
Block DECO		0.032			0.955	22.51	**86,222**	**-162,537**
		(0.000)			(0.001)	(0.14)		

however, we use a common A parameter for parsimony. To save space, we do not report all the different intercepts ω_g for all groups for the factor copulas with group-specific loadings. All standard errors are based on the sandwich (robust covariance matrix) estimator $A_0^{-1}B_9A_0^{-1}$ with B_0 the inverse Hessian of the likelihood evaluated at the optimum (information matrix), and A_0 the expected value of the outer product of the gradients at the optimum.

Three main results emerge from Table 7. 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 and block DECO). The gain is not only with respect to the total log-likelihood, but also with respect to the AIC, which takes into account the number of estimated parameters.

Second, multi-factor models appear to provide a much better fit than one-factor copula models. For 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. This increases the log likelihood by almost 10,000 points in both the Gaussian and Student's t case. An additional 1,500 points is gained by going to the triangular loadings structure and allowing more freedom for the between-industry correlations.

Third, Table 7 shows two additional interesting findings. As is usual, the Student's t factor copulas fit considerably better than their Gaussian counterparts. Log-likelihood differences range between 2,000 and 4,000 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 fat-tailedness. We also see a 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. Both findings confirm the empirical analysis 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 is much lower for the Block DECO t specification than for the MF-LT t model or the DECO model. It seems that there is empirically some efficiency loss or bias effect due to the use of the Composite Likelihood method for parameter estimation.

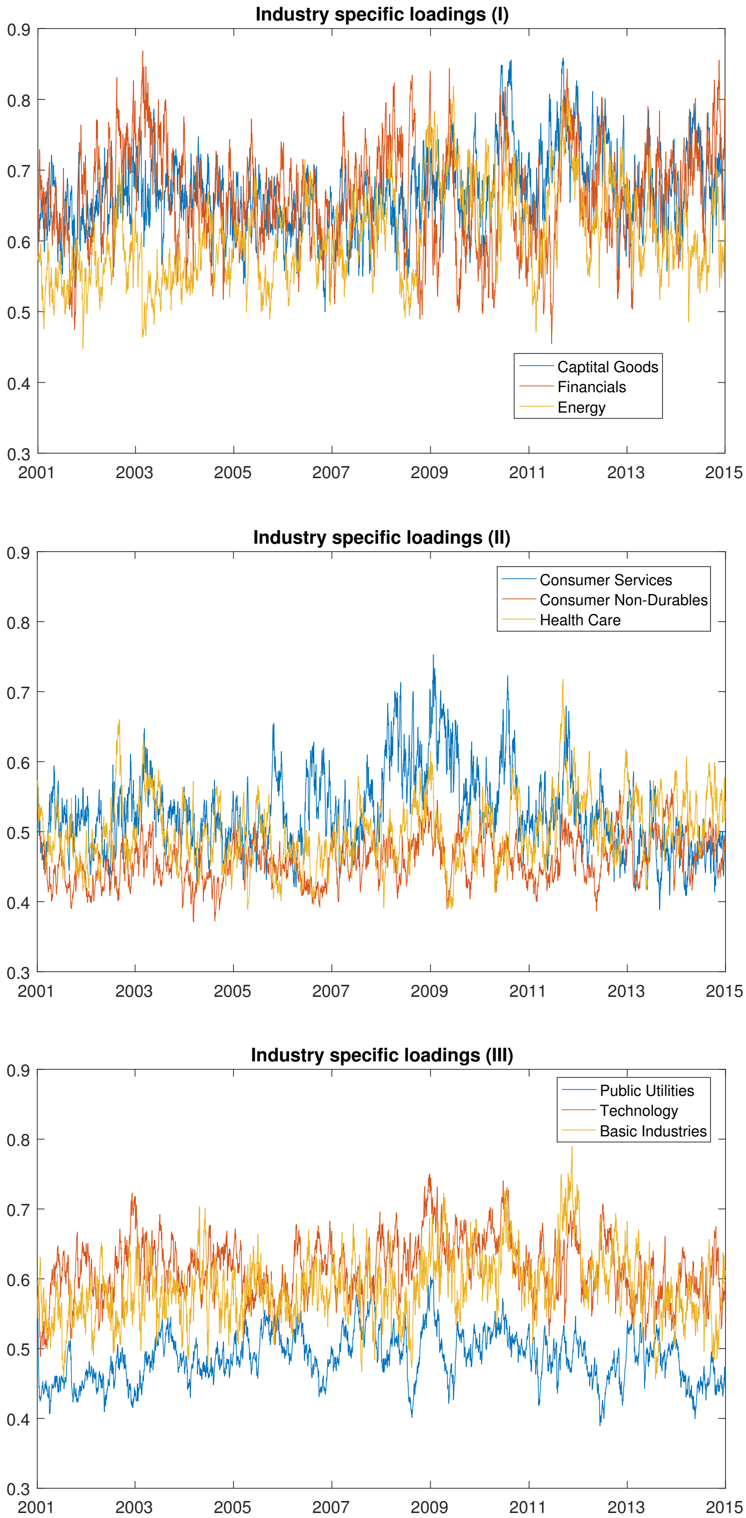

Figure 1: Factor loadings of the Multi-Factor LT t-copula

This figure shows factor loadings through time for the first factor of the MF-LT t copula model. We show 9 different factor loadings, corresponding to Financials, Capital Goods, Energy, Consumer Services, Consumer Non-Durables, Health Care, Public Utilities, Technology, and Basic Industries. The sample spans the period from January 2, 2001 until December 31, 2014 ($T = 3,521$ days).

Figure 1 shows the different factor loadings corresponding to the MF-LT t copula model, which has the richest factor structure and the best fit. We concentrate on the loadings of the first factor. The top, middle and bottom panels show three industry loadings each. We note that the loadings of the Financials and Energy (Consumer Services and Non-Durables) industries are higher than the Capital Goods (Health Care) loadings. Further, Financial industry loadings exceed the loadings of the Energy industry during the years 2001 and 2013. Huge upward spikes are visible as well for the Financials during the heat of the global financial crisis (2007-2009): the loading increases rapidly from 0.5 to 0.8. The main take-away is that the nine industry factors pick up significant group-specific dynamics that cannot be captured by a single factor model. Hence allowing for different industry factors with their own GAS loadings captures additional dynamic effects present in the data and results in substantial increases in the statistical fit of the model, as shown before.

Figure 2 plots the implied within correlation of Financials and Capital Goods industry companies for the MF-LT t and the Block DECO-t model. The two panels show a similar pattern for the within industry correlations. Correlations are high around the same times for both models, as are the differences between the industry correlations. The main feature evident from the figure is that the MF-LT specification seems to be much more responsive to incoming data, whereas the Block DECO seems to be much smoother. Whether one or the other pattern is more helpful for out-of-sample forecasting is a question that we address next. For the in-sample analysis, we conclude that the MF-LT t model has the best total statistical fit, which is even obtained using a considerably smaller number of static parameters. Furthermore, the fit of the multi-factor model is considerably better than the fit of one-factor models.

4.3 Multivariate Density Forecasts

As we have closed-form copula densities, 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). Using the log score rule (see Mitchell and Hall, 2005; Amisano and Giacomini, 2007), the multivariate one-step-ahead

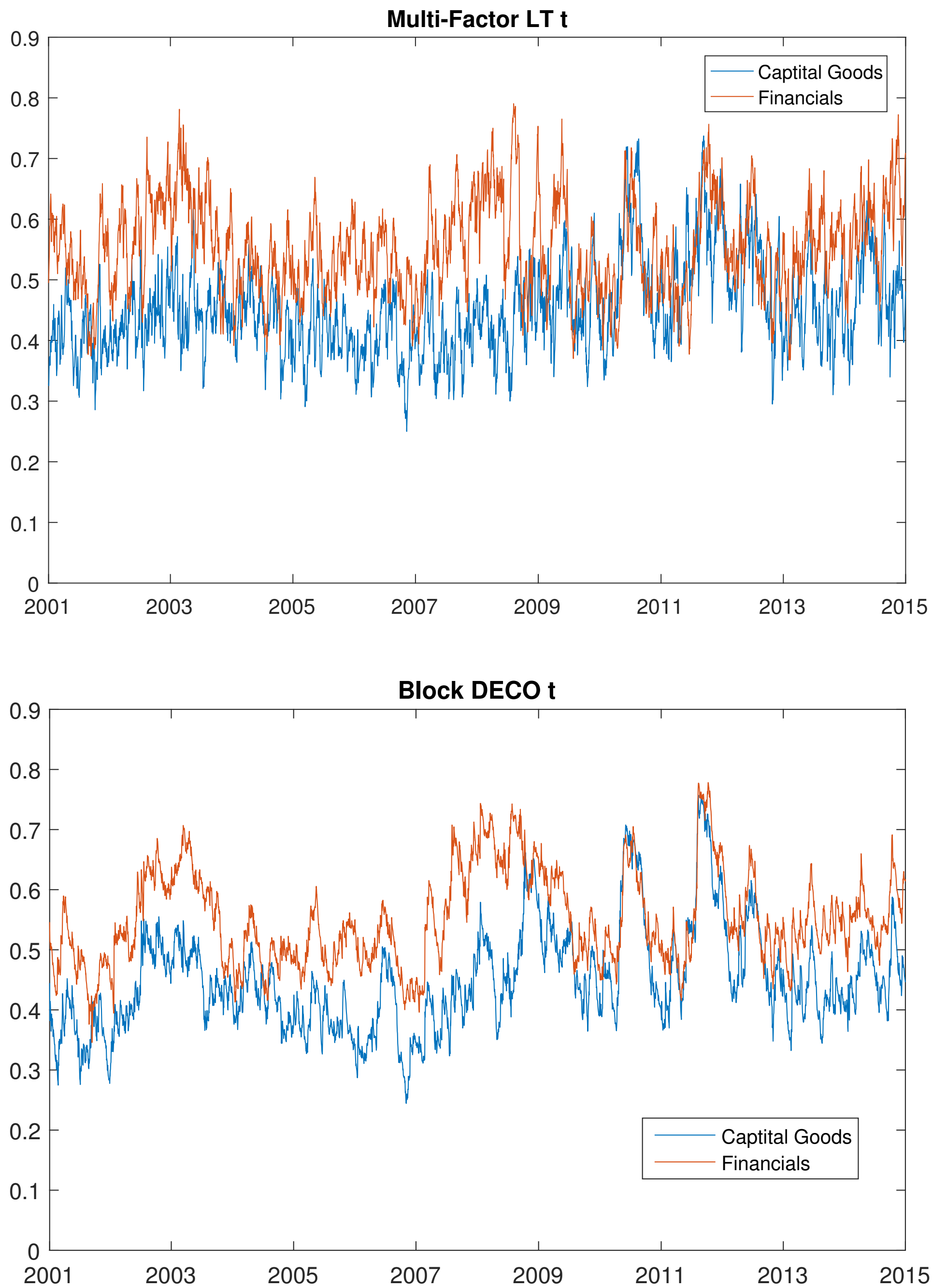

Figure 2: Within group dependencies

This figure shows the within dependence between stock returns of 19 Financial companies (red line) and 11 Capital Goods companies (blue line) through time, according to the MF-LT t-copula model (upper panel) and the Block DECO t model (lower panel). The sample spans the period from January 2, 2001 until December 31, 2014 ($T = 3,521$ days).

density forecasts require the evaluation of the OOS copula log-likelihood.

Using a moving-window with an in-sample period of 1,000 observations (or roughly four calendar years) 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 each model after 50 observations, which corresponds to roughly 10 calendar weeks.

Define the difference in the log score (ls) between two copula density forecasts M_1 and M_2 as

$$d_{ls,t} = S_{ls,t}(\boldsymbol{u}_t, M_1) - S_{ls,t}(\boldsymbol{u}_t, M_2) \tag{29}$$

for $t = 1001, 1002, \ldots, T$, with $S_{ls,t}(\boldsymbol{u}_t, M_j)$ the log score at time t of the density forecast corresponding to model M_j for $j = 1, 2$, i.e.,

$$S_{ls,t}(\boldsymbol{u}_t, M_j) = \log c_t(\boldsymbol{u}_t, R_t \mid \mathcal{F}_{t-1}, M_j), \tag{30}$$

where $c_t(\cdot)$ is the Gaussian or Student's t copula density. Note that the marginal distributions based on identical specifications for all models drop out in (30).

The null and alternative hypotheses of equal predictive ability are given by

$$H_0 : \mathbb{E}[d_{ls,t}] = 0, \qquad\qquad H_A : \mathbb{E}[d_{ls,t}] \neq 0. \tag{31}$$

This hypothesis can be tested by means of a Diebold and Mariano (1995) test statistic

$$DM_{ls} = \frac{\bar{d}}{\sqrt{\hat{\sigma}^2/P}}, \tag{32}$$

where $\bar{d}$ is the out-of-sample average of the log score differences and $\hat{\sigma}^2$ a HAC estimator of the true variance σ^2 of $d_{ls,t}$. This test-statistic is asymptotically $\mathrm{N}(0,1)$ distributed under the assumptions of Giacomini and White (2006). A significantly positive value of DM_{ls} means that model M_1 has superior forecast performance over model M_2.

Since we deal with a substantial number of different models and factor structures and hence many different density forecasts, we also consider the Model Confidence Set (MCS) of

Hansen et al. (2011). To build the MCSs, we compute (minus) the log score values and use a significance level of 5%. The MCS automatically accounts for the dependence between model outcomes given that all models are based on the same data.

Table 8 compares copula density forecasts over the whole region (Panel A) and over the joint lower tails (Panels B.1-B.4) for the factor copulas and the MGARCH copula models. We report the mean of the log score as well as the pair-wise DM test statistics of the MF-LT t model against all other models. Finally, we show the p-values of the Model Confidence Set approach.

The table shows two interesting results. First, in line with our full-sample results, Panel A shows that the MF-LT t model performs best in terms of predictive ability. The pairwise comparisons via the DM test-statistics reveal a rejection of the null hypothesis of equal predictive ability for all models. Only when comparing the ML-LT t model to the Block DECO t model, the results is less strong and only significant at a 10% level.

The same pattern emerges from MCS approach. The MCS p-value equals 1 for the MF-LT t, whereas that of all other models equals 0, except the Block DECO t model. The latter stays within the MCS, but with a much more modest p-value of 22%. 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 0.7 points (from 21.09 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 24.77. Allowing for cross exposures in the MF-LT specification results in yet a further increase by 1.3 points.

Overall, we conclude that the flexibility provided by the new MF-LT t model also seems important in a forecasting context. The more flexible parametrization allows for a larger class of dependence matrices. This extension appears to be empirically important in high dimensions.

Table 8: One-step ahead copula density forecasts

This table provides the accuracy of one-step ahead copula density forecasts of 100 daily returns of the S&P500 index, obtained by various factor copula and 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 DM-statistic of a pairwise test on predictive accuracy of the MF-LT t model against all other models. A positive DM-statistic implies that the MF-LT t model has superior density forecasts against the alternative models. Finally, we present the p-value associated with the Model Confidence Set of Hansen et al. (2011), based on a significance level of 5%. Bold numbers in this row represent those models that belong to the model confidence set. The out-of-sample period goes from January 2005 until December 2014 and contains 2,519 observations.

Model	$S_{ls,t}$	DM_{ls}	MCS p-val
1F-Equi N	20.10	33.86	(0.00)
1F-Equi t	21.09	35.72	(0.00)
1F-Group N	20.73	34.03	(0.00)
1F-Group t	21.83	36.78	(0.00)
2F N	21.79	29.08	(0.00)
2F t	23.28	29.22	(0.00)
MF N	23.03	14.62	(0.00)
MF t	24.77	13.22	(0.00)
MF-LT N	25.35	11.32	(0.00)
MF-LT t	**26.07**		**(1.00)**
cDCC(CL)	22.42	14.10	(0.00)
DECO N	19.76	32.38	(0.00)
DECO t	21.01	36.73	(0.00)
Block DECO N	25.24	8.51	(0.00)
Block DECO t	26.02	1.68	**(0.22)**

4.4 Economic out-of-sample performance

Finally, we assess the forecasting performance of the different models from an economic point of view. 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 the investor's aim is to minimize the 1-step portfolio volatility at time t subject to a fully invested portfolio, the resulting GMVP weights $\boldsymbol{w_{t+1|t}}$ are obtained by the solution of the quadratic programming problem

$$\min \boldsymbol{w_{t+1|t}}' \left(\boldsymbol{D_{t+1|t}} \boldsymbol{R}^*_{t+1|t} \boldsymbol{D_{t+1|t}}\right) \boldsymbol{w_{t+1|t}}, \qquad \text{s.t. } \boldsymbol{w_{t+1|t}}'\boldsymbol{\iota} = 1, \tag{33}$$

with $\boldsymbol{D_{t+1|t}}$ the 1-step ahead forecasted variances from the GAS t model, $\boldsymbol{R}^*_{t+1|t}$ the one-step ahead forecasted correlation matrix according to our different factor copula models, and $\boldsymbol{\iota}$ a $k \times 1$ vector of ones.[2] 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, which are given by $\sigma_{p,t} = \sqrt{\boldsymbol{w_{t+1|t}}'\boldsymbol{RC_{t+1}}\boldsymbol{w_{t+1|t}}}$, 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 again 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

[2] Note that $\boldsymbol{R}^*_{t+1|t} \neq \boldsymbol{R}_{t+1|t}$ as we need the correlations of $\boldsymbol{y_{t+1}}$ instead of $\boldsymbol{x_t}$. We compute $\boldsymbol{R}^*_{t+1|t}$ by means of simulating 10,000 returns from the joint distribution of returns as constructed from the marginals and conditional copula.

each competing model at time t. Turnover at time t is defined as

$$TO_t = \sum_{i=1}^{k} \left| w_{t+1|t}^{(i)} - w_{t|t-1}^{(i)} \frac{1 + y_t^{(i)}}{1 + \boldsymbol{w_{t|t-1}}' \boldsymbol{y}_t} \right|, \tag{34}$$

where $w_{t|t-1}^{(i)}$ and $y_t^{(i)}$ the i-th element of the weight vector $\boldsymbol{w_{t|t-1}}$ and return vector $\boldsymbol{y}_t$. 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

$$CO_t = \left(\sum_{i=1}^{k} w_t^{(i)2} \right)^{1/2}, \tag{35}$$

while the total portfolio short position SP_t is given by

$$SP_t = \sum_{i=1}^{k} w_t^{(i)} \cdot I[w_t^{(i)} < 0], \tag{36}$$

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 9 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. The 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 was the best in-sample and out-of-sample performer. Note however, that also here the MF-LT t model still comes close, as indicated by the MCS p-value of 0.03. The multi-factor models also still outperform the block-DECO model in terms of ex-post variance of the GMVP. The 1-factor models, however, has the best performance, both in terms of ex-post variance, turnover, concentration, and total

Table 9: 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 January 2005 until December 2014 and contains 2,519 observations.

	$\bar{\sigma}_p$	p-val	TO	CO	SP
Factor-copula models					
1F-Equi N	0.522	(0.00)	0.353	0.336	-0.593
1F-Equi t	0.524	(0.00)	0.352	0.338	-0.600
1F-Group N	0.510	**(1.00)**	0.343	**0.307**	**-0.527**
1F-Group t	0.510	**(0.95)**	**0.321**	0.310	-0.537
2F N	0.525	(0.00)	0.399	0.330	-0.581
2F t	0.535	(0.00)	0.376	0.344	-0.619
MF N	0.534	(0.00)	0.406	0.357	-0.634
MF t	0.534	(0.00)	0.404	0.359	-0.641
MF-LT N	0.523	(0.00)	0.422	0.350	-0.619
MF-LT t	0.522	(0.03)	0.407	0.351	-0.621
Copula-MGARCH models					
cDCC (CL)	0.554	(0.00)	0.761	0.471	-1.030
DECO N	0.523	(0.00)	0.345	0.336	-0.592
DECO t	0.523	(0.00)	0.346	0.337	-0.594
Block DECO N	0.531	(0.00)	0.398	0.352	-0.619
Block DECO t	0.533	(0.00)	0.399	0.352	-0.621

number of short positions. Unlike the density forecast setting, we also note that the choice of the distribution plays a less important role in Table 9.

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 focusses attention on an area of the forecast distribution where differences are more concentrated: all models focus on a portfolio with ex-ante *minimum* variance. If the models are any good, differences in this concentrated performance measure are hard to obtain. This is corroborated by the results in Table 9. Although the results are sometimes statistically significantly different, they are all quite close in economic terms (with the possible exception of the cDCC). Using such a performance measure, introducing models with more parameters and their associated estimation risk typically deteriorates overall performance. This explains why the simpler 1-factor models do better here. By contrast, if the full density 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.

Summarizing, we conclude that the one-step-ahead copula density forecasts over the whole support of the MF-LT t model are superior to those of the other factor copula models and GARCH based copula models considered. In terms of economic performance of the GMVP, however, simpler factor models appear to work better. In particular, a one-factor model with industry specific loadings turns out to have the lowest conditional portfolio standard deviation. This model also has significantly (in terms of MCS) lower turnover, lower concentration, and smaller short positions than the other factor models, and even more so than the MGARCH models considered like the cDCC and (block) DECO.

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. 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. 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. 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 in high-dimensional settings. Given their computational ease and closed-form likelihood expression, they thus provide a useful tool for modeling high-dimensional dynamic dependence structures.

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.

Azzalini, A. and A. Capitanio (2003). Distributions generated by perturbation of symmetry with emphasis on a multivariate skew t-distribution. *Journal of the Royal Statistical Society: Series B (Statistical Methodology) 65*, 367–389.

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

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.

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.

Giacomini, R. and H. White (2006). Tests of conditional predictive ability. *Econometrica 74*, 1545–1578.

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

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 Universiteit 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

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

with $\boldsymbol{\lambda}_{i,t}$ a $k \times 1$ vector of time-varying factor loadings, and $\log \boldsymbol{c}(\cdot)$ the log-copula density. Note that the dimension of $\boldsymbol{\lambda}_{i,t}$ (and hence $\boldsymbol{s}_t$) depends on the chosen factor structure.

We assume a Student's t and a Gaussian copula density for $\boldsymbol{x}_t = (x_{1,t}, \ldots, x_{N,t})^\top$. The corresponding copula log-likelihood at time t is defined as

$$l_{Stud,t} = -\frac{1}{2}\log|\boldsymbol{R}_t| - \frac{\nu+n}{2}\log\left(1 + \frac{\boldsymbol{x}_t^\top \boldsymbol{R}_t^{-1}\boldsymbol{x}_t}{\nu-2}\right), \tag{A.2}$$

$$l_{Gaus,t} = -\frac{1}{2}\log|\boldsymbol{R}_t| + -\frac{1}{2}\boldsymbol{x}_t^\top \boldsymbol{R}_t^{-1}\boldsymbol{x}_t, \tag{A.3}$$

omitting terms that do not depend on $\boldsymbol{R}_t$. Further, the dependence matrix $\boldsymbol{R}_t$ of $\boldsymbol{x}_t$ is modeled as

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

with

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

for an unrestricted $k \times 1$ vector $\boldsymbol{\lambda}_{i,t}$. This ensures that $x_{i,t}$ has zero mean and unit variance by design.

Hence, using the chain rule

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

Given (A.4), we have

$$\mathrm{d} \operatorname{vec} \boldsymbol{R}_t = \mathrm{d} \operatorname{vec}(\tilde{\boldsymbol{L}}_t^\top \tilde{\boldsymbol{L}}_t + \boldsymbol{D}_t) = (\mathbf{I}_{n^2} + K_n)\left(\mathbf{I}_n \otimes \tilde{\boldsymbol{L}}_t^\top\right) \mathrm{d} \operatorname{vec} \tilde{\boldsymbol{L}}_t + \mathrm{d} \operatorname{vec} \boldsymbol{D}_t, \tag{A.7}$$

wher K_n is such that $\operatorname{vec}(A) = K_n \operatorname{vec}(A^\top)$ for a general $n \times n$ matrix A. We obtain

$$\frac{\partial \operatorname{vec} \boldsymbol{R}_t}{\partial \boldsymbol{\lambda}_{i,t}^\top} = (\mathbf{I}_{n^2} + K_n)\left(\mathbf{I}_n \otimes \tilde{\boldsymbol{L}}_t^\top\right) \frac{\partial \operatorname{vec} \tilde{\boldsymbol{L}}_t}{\partial \boldsymbol{\lambda}_{i,t}^\top} + \frac{\partial \operatorname{vec} \boldsymbol{D}_t}{\partial \boldsymbol{\lambda}_{i,t}^\top}. \tag{A.8}$$

Hence to summarize, we need to derive $\frac{\partial \log \boldsymbol{c}(\boldsymbol{x}_t; \boldsymbol{\lambda}_{i,t}, \psi_C)}{\partial \operatorname{vec}(\boldsymbol{R}_t)^\top}$, $\frac{\partial \operatorname{vec} \tilde{\boldsymbol{L}}_t}{\partial \boldsymbol{\lambda}_{i,t}^\top}$ and $\frac{\partial \operatorname{vec} \boldsymbol{D}_t}{\partial \boldsymbol{\lambda}_{i,t}^\top}$. Note that the first quantity does NOT depend on any factor structure, only on the chosen copula density, while the latter two do depend on the factor structure. We will therefore derive $\frac{\partial \log \boldsymbol{c}(\boldsymbol{x}_t; \boldsymbol{\lambda}_{i,t}, \psi_C)}{\partial \operatorname{vec}(\boldsymbol{R}_t)^\top}$ in case of the Student's t and Gaussian copula density first, and then derive the other two quantities in the following subsections.

For the t-copula, $\mathrm{d}\, l_{Stud,t}$ is given by

$$\begin{aligned}
\mathrm{d}\, l_{Stud,t} &= -\frac{1}{2} \operatorname{tr}\left(\boldsymbol{R}_t^{-1} \, \mathrm{d} \boldsymbol{R}_t\right) - \frac{\nu+n}{2} \frac{1}{1+\frac{\boldsymbol{x}_t^\top \boldsymbol{R}_t^{-1} \boldsymbol{x}_t}{\nu-2}} \mathrm{d}\left(\frac{\boldsymbol{x}_t^\top \boldsymbol{R}_t^{-1} \boldsymbol{x}_t}{\nu-2}\right) \\
&= -\frac{1}{2}\left(\operatorname{vec} \boldsymbol{R}_t^{-1}\right)^\top \mathrm{d} \operatorname{vec} \boldsymbol{R}_t + \frac{1}{2}\left[\frac{\nu+n}{\nu + \boldsymbol{x}_t^\top \boldsymbol{R}_t^{-1} \boldsymbol{x}_t}\right] \boldsymbol{x}_t^\top \boldsymbol{R}_t^{-1} (\mathrm{d} \boldsymbol{R}_t) \boldsymbol{R}_t^{-1} \boldsymbol{x}_t \\
&= -\frac{1}{2}\left(\operatorname{vec} \boldsymbol{R}_t^{-1}\right)^\top \mathrm{d} \operatorname{vec} \boldsymbol{R}_t + \frac{1}{2}\left[\frac{\nu+n}{\nu + \boldsymbol{x}_t^\top \boldsymbol{R}_t^{-1} \boldsymbol{x}_t} \boldsymbol{x}_t^\top \boldsymbol{R}_t^{-1} \otimes \boldsymbol{R}_t^{-1} \boldsymbol{x}_t\right]^\top \mathrm{d} \operatorname{vec} \boldsymbol{R}_t \\
&= \left\{-\frac{1}{2}\left(\operatorname{vec} \boldsymbol{R}_t^{-1}\right)^\top + \frac{1}{2}\left[\frac{\nu+n}{\nu + \boldsymbol{x}_t^\top \boldsymbol{R}_t^{-1} \boldsymbol{x}_t} \operatorname{vec}\left(\boldsymbol{R}_t^{-1} \boldsymbol{x}_t \boldsymbol{x}_t^\top \boldsymbol{R}_t^{-1}\right)\right]^\top\right\} \mathrm{d} \operatorname{vec} \boldsymbol{R}_t, \qquad (A.9)
\end{aligned}$$

hence

$$\frac{\partial \log \boldsymbol{c}(\boldsymbol{x}_t; \tilde{\boldsymbol{\lambda}}_{i,t}, \nu_C)}{\partial \operatorname{vec}(\boldsymbol{R}_t)^\top} = -\frac{1}{2}\left(\operatorname{vec} \boldsymbol{R}_t^{-1}\right)^\top + \frac{1}{2}\left[\frac{\nu+n}{\nu + \boldsymbol{x}_t^\top \boldsymbol{R}_t^{-1} \boldsymbol{x}_t} \operatorname{vec}\left(\boldsymbol{R}_t^{-1} \boldsymbol{x}_t \boldsymbol{x}_t^\top \boldsymbol{R}_t^{-1}\right)\right]^\top. \tag{A.10}$$

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

$$\mathrm{d}\, l_{Gaus,t} = \left\{ -\frac{1}{2} \left(\operatorname{vec} \boldsymbol{R}_t^{-1} \right)^\top + \frac{1}{2} \operatorname{vec} \left[\boldsymbol{R}_t^{-1} \boldsymbol{x}_t \boldsymbol{x}_t^\top \boldsymbol{R}_t^{-1} \right]^\top \right\} \mathrm{d} \operatorname{vec} \boldsymbol{R}_t, \tag{A.11}$$

such that

$$\frac{\partial \log \boldsymbol{c}(\boldsymbol{x}_t; \tilde{\boldsymbol{\lambda}}_{i,t})}{\partial \operatorname{vec}(\boldsymbol{R}_t)^\top} = -\frac{1}{2} \left(\operatorname{vec} \boldsymbol{R}_t^{-1} \right)^\top + \frac{1}{2} \operatorname{vec} \left[\boldsymbol{R}_t^{-1} \boldsymbol{x}_t \boldsymbol{x}_t^\top \boldsymbol{R}_t^{-1} \right]^\top. \tag{A.12}$$

A.2: 1-Equi-Factor

In the 1-Factor equi-copula, we have $\tilde{\boldsymbol{L}}_t^\top = \tilde{\lambda}_t \boldsymbol{\iota}_N$ with $\tilde{\lambda}_t = \exp(\lambda_t)/\sqrt{1+\exp(2\lambda_t)}$, and $\lambda_t \in \mathbb{R}$. Then $\partial \operatorname{vec} \tilde{\boldsymbol{L}}_t / \partial \lambda_t$ is given by

$$\frac{\partial \operatorname{vec} \tilde{\boldsymbol{L}}_t}{\partial \lambda_t} = \boldsymbol{\iota}_N \frac{\partial}{\partial \lambda_t} \tilde{\lambda}_t = \boldsymbol{\iota}_N \left[\tilde{\lambda}_t - \frac{\exp(3\lambda_t)}{(1+\exp(2\lambda_t))^{3/2}} \right] = \boldsymbol{\iota}_N \tilde{\lambda}_t (1 - \tilde{\lambda}_t^2). \tag{A.13}$$

Second, given that $\boldsymbol{D}_t = \frac{1}{1+\exp(2\lambda_t)} \mathbf{I}_n$ it holds that

$$\frac{\partial \operatorname{vec} \boldsymbol{D}_t}{\partial \lambda_t} = \operatorname{vec} \mathbf{I}_N \frac{\partial}{\partial \lambda_t} \frac{1}{1+\exp(2\lambda_t)} = \operatorname{vec} \mathbf{I}_N \frac{-2\exp(2\lambda_t)}{(1+\exp(2\lambda_t))^2} = \operatorname{vec} \mathbf{I}_N \, \big(-2\tilde{\lambda}_t^2 (1 - \tilde{\lambda}_t^2) \big). \tag{A.14}$$

Hence $\partial \log \boldsymbol{c}(\boldsymbol{x}_t; \boldsymbol{\lambda}_{i,t}, \psi_C)/\partial \boldsymbol{\lambda}_t$ is obtained by combining (A.10) or (A.12) with (A.8), (A.13) and (A.14).

A.3: 1-Factor model with heterogeneous loadings

Let $\boldsymbol{S}^L$ denote a selection matrix of size $N \times G$ such that $S_{i,g}^L = 1$ if asset i belongs to group g, and 0 else, where $i = 1, \ldots, N$ and $g = 1, \ldots, G$. Further, let $\boldsymbol{S}^{D1}$ denote a selection matrix of size $N^2 \times G$ such that $S_{i+(i-1)N,g}^D = 1$ if asset i belongs to group g, and 0 else, where $i = 1, \ldots, n$, and $g = 1, \ldots, G$.

Then $\tilde{\boldsymbol{L}}_t^\top = \boldsymbol{S}^L \tilde{\boldsymbol{\lambda}}_t^{gr}$ and $\operatorname{vec} \boldsymbol{D}_t = \boldsymbol{S}^{D1} \tilde{\boldsymbol{\lambda}}_t^{gr}$, with $\tilde{\boldsymbol{\lambda}}_t^{gr} = [\tilde{\lambda}_{t,1}, \tilde{\lambda}_{t,2} \ldots, \tilde{\lambda}_{t,G}]^\top$ a $G \times 1$ vector with the G different group loadings. Further, $\tilde{\lambda}_{t,g} = \exp(\lambda_{t,g})/\sqrt{1+\exp(2\lambda_{t,g})}$.

The objects of interest are $\partial \operatorname{vec} \tilde{\boldsymbol{L}}_t / \partial (\boldsymbol{\lambda}_t^{gr})^\top$ and $\partial \operatorname{vec} \boldsymbol{D}_t / \partial (\boldsymbol{\lambda}_t^{gr})^\top$. We have

$$\frac{\partial \operatorname{vec} \tilde{\boldsymbol{L}}_t}{\partial (\boldsymbol{\lambda}_t^{gr})^\top} = \boldsymbol{S}^L \frac{\partial \operatorname{vec} \tilde{\boldsymbol{\lambda}}_t^{gr}}{\partial (\boldsymbol{\lambda}_t^{gr})^\top} = \boldsymbol{S}^L \cdot \operatorname{diag}\big(\tilde{\lambda}_{t,1}(1 - \tilde{\lambda}_{t,1}^2) \ , \ \ldots \ , \ \tilde{\lambda}_{t,G}(1 - \tilde{\lambda}_{t,G}^2) \big). \tag{A.15}$$

Similarly,

$$\frac{\partial \operatorname{vec} \boldsymbol{D}_t}{\partial (\boldsymbol{\lambda}_t^{gr})^\top} = -2\boldsymbol{S}^{D1} \cdot \operatorname{diag}\big(\tilde{\lambda}_{t,1}^2 (1 - \tilde{\lambda}_{t,1}^2) \ , \ \ldots \ , \ \tilde{\lambda}_{t,G}^2 (1 - \tilde{\lambda}_{t,G}^2) \big). \tag{A.16}$$

A.4: MF model

The MF model consists of two types of factors: an equi-factor, and G industry factors, each with a group-specific loading. Note that the equifactor is in fact group-specific *after scaling*, i.e., even if λ_t^{eq} is scalar, $\tilde{\lambda}_{t,g}^{eq} = \exp(\lambda_t^{eq})/\sqrt{1+\exp(2\lambda_t^{eq})+\exp(2\lambda_{t,g}^{ind})}$ depends on the group via the unscaled industry factor loading.

We define $\tilde{\boldsymbol{\lambda}}_t^{eq} = \boldsymbol{S}^L(\tilde{\lambda}_{t,1}^{eq},\ldots,\tilde{\lambda}_{t,G}^{eq})^\top$, and $\tilde{\boldsymbol{\lambda}}_t^{ind} = \boldsymbol{S}^L\mathrm{diag}(\tilde{\lambda}_{t,1}^{ind},\ldots,\tilde{\lambda}_{t,G}^{ind})$, using the same definition for $\boldsymbol{S}^L$ as in Section A.3. The corresponding loading matrix equals $\tilde{\boldsymbol{L}}_t^\top = [\tilde{\boldsymbol{\lambda}}_t^{gr} \quad \tilde{\boldsymbol{\lambda}}_t^{ind}] \in \mathbb{R}^{N\times(G+1)}$.

The scaled lambdas are given by

$$
\begin{aligned}
\tilde{\lambda}_{t,g}^{eq} &= \exp(\lambda_t^{eq})/\sqrt{X_{t,g}}, &\text{(A.17)}\\
\tilde{\lambda}_{t,g}^{ind} &= \exp(\lambda_{t,g}^{ind})/\sqrt{X_{t,g}}, &\text{(A.18)}\\
X_{t,g} &= 1+\exp(2\lambda_t^{eq})+\exp(2\lambda_{t,g}^{ind}). &\text{(A.19)}
\end{aligned}
$$

We have the following four types of derivatives:

$$
\begin{aligned}
\frac{\partial\tilde{\lambda}_{t,g}^{eq}}{\partial\lambda_t^{eq}} &= \tilde{\lambda}_{t,g}^{eq}\big(1-(\tilde{\lambda}_{t,g}^{eq})^2\big), &\text{(A.20)}\\
\frac{\partial\tilde{\lambda}_{t,g}^{eq}}{\partial\lambda_{t,g}^{ind}} &= -\tilde{\lambda}_{t,g}^{eq}(\tilde{\lambda}_{t,g}^{ind})^2, &\text{(A.21)}\\
\frac{\partial\tilde{\lambda}_{t,g}^{ind}}{\partial\lambda_t^{eq}} &= -\tilde{\lambda}_{t,g}^{ind}(\tilde{\lambda}_{t,g}^{eq})^2, &\text{(A.22)}\\
\frac{\partial\tilde{\lambda}_{t,g}^{ind}}{\partial\lambda_{t,g}^{ind}} &= \tilde{\lambda}_{t,g}^{ind}\big(1-(\tilde{\lambda}_{t,g}^{ind})^2\big). &\text{(A.23)}
\end{aligned}
$$

Using these results, we have

$$
\frac{\partial\,\mathrm{vec}\,\tilde{\boldsymbol{L}}_t}{\partial\lambda_t} = \mathrm{vec}\left(\boldsymbol{S}^L\cdot\begin{pmatrix}\tilde{\lambda}_{t,1}^{eq}(1-(\tilde{\lambda}_{t,1}^{eq})^2)\\ \vdots \\ \tilde{\lambda}_{t,G}^{eq}(1-(\tilde{\lambda}_{t,G}^{eq})^2)\end{pmatrix} \qquad -\boldsymbol{S}^L\cdot\mathrm{diag}\big(\tilde{\lambda}_{t,1}^{ind}(\tilde{\lambda}_{t,1}^{eq})^2,\ldots,\tilde{\lambda}_{t,G}^{ind}(\tilde{\lambda}_{t,G}^{eq})^2\big)\right), \tag{A.24}
$$

and

$$
\frac{\partial\,\mathrm{vec}\,\tilde{\boldsymbol{L}}_t}{\partial\lambda_{t,g}^{ind}} = \mathrm{vec}\left(-\boldsymbol{S}^L\cdot\begin{pmatrix}\tilde{\lambda}_{t,1}^{eq}(\tilde{\lambda}_{t,1}^{ind})^2\\ \vdots \\ \tilde{\lambda}_{t,G}^{eq}(\tilde{\lambda}_{t,G}^{ind})^2\end{pmatrix} \qquad \tilde{\lambda}_{t,g}^{ind}(1-(\tilde{\lambda}_{t,g}^{eq})^2)\boldsymbol{S}^L\odot e_g\right), \tag{A.25}
$$

where e_g is the gth row from the unit matrix $\mathbf{I}_G$, and $\odot$ is the element-by-element Hadamard product. Note that both matrices above are very sparse.

To derive $\frac{\partial \operatorname{vec} \boldsymbol{D}_t}{\partial \boldsymbol{\lambda}_t} \in \mathbb{R}^{N^2 \times (G+1)}$, we note that

$$\begin{aligned} \frac{\partial \sigma^2_{t,g}}{\partial \lambda^{eq}_t} &= -2(\tilde{\lambda}^{eq}_{t,g})^2 \sigma^2_{t,g}, & \text{(A.26)} \\ \frac{\partial \sigma^2_{t,g}}{\partial \lambda^{ind}_{t,g}} &= -2(\tilde{\lambda}^{ind}_{t,g})^2 \sigma^2_{t,g}. & \text{(A.27)} \end{aligned}$$

Define $\boldsymbol{S}^{D2}$ as a selection matrix of size $N^2 \times G$ such that $\boldsymbol{S}^{D2}_{i+(i-1)N,g} = 1$ for $i = 1 \dots N$ if firm i belongs to group g, and zero else. Also define $\boldsymbol{\sigma}^2_t = (\sigma^2_{t,1}, \dots, \sigma^2_{t,G})^\top$. We then have $\operatorname{vec} \boldsymbol{D}_t = \boldsymbol{S}^{D2} \cdot \boldsymbol{\sigma}^2_t$, and thus

$$\frac{\partial \operatorname{vec} \boldsymbol{D}_t}{\partial \boldsymbol{\lambda}_t^\top} = -2 \boldsymbol{S}^{D2} \cdot \boldsymbol{\sigma}^2_t \cdot \left(\tilde{\boldsymbol{\lambda}}_t \odot \tilde{\boldsymbol{\lambda}}_t \right)^\top. \tag{A.28}$$

A.5: MF LT model

For the MF LT model, we have

$$\sigma^2_{t,g} = \left(1 + \sum_{i=1}^{g} \exp(2\lambda_{t,g,i}) \right)^{-1}, \tag{A.29}$$

$$\frac{\partial \sigma^2_{t,g}}{\partial \lambda_{t,j,k}} = \begin{cases} -2\sigma^2_{t,g} \lambda^2_{t,g,k}, & \text{if } j = g \text{ and } k \leq g, \\ 0, & \text{else,} \end{cases} \tag{A.30}$$

$$\frac{\partial \tilde{\lambda}_{t,i,j}}{\partial \lambda_{t,k,\ell}} = \begin{cases} \tilde{\lambda}_{t,i,j}(1 - \tilde{\lambda}^2_{t,i,j}) & \text{if } i = k,\ j = \ell, \text{ and } j \leq i, \\ -\tilde{\lambda}_{t,i,j} \tilde{\lambda}^2_{t,i,\ell} & \text{if } i = k,\ j \neq \ell,\ j \leq i, \text{ and } \ell \leq k, \\ 0, & \text{else.} \end{cases} \tag{A.31}$$

Define a selection matrix $\boldsymbol{S}^L \in \mathbb{R}^{N \cdot G \times G(G+1)/2}$ such that $\operatorname{vec} \tilde{\boldsymbol{L}}_t = \boldsymbol{S}^L \tilde{\boldsymbol{\lambda}}_t$ for the vectors

$$\tilde{\boldsymbol{\lambda}}_t^\top = \Big(\tilde{\lambda}_{t,1,1} \ , \ \tilde{\lambda}_{t,2,1}, \tilde{\lambda}_{t,2,2} \ , \ \tilde{\lambda}_{t,3,1}, \dots, \tilde{\lambda}_{t,3,3} \ , \ \tilde{\lambda}_{t,4,1}, \dots, \tilde{\lambda}_{t,4,4} \ , \ \dots \ , \ \tilde{\lambda}_{t,G,1}, \dots, \tilde{\lambda}_{t,G,G} \Big), \tag{A.32}$$

$$\boldsymbol{\lambda}_t^\top = \Big(\lambda_{t,1,1} \ , \ \lambda_{t,2,1}, \lambda_{t,2,2} \ , \ \lambda_{t,3,1}, \dots, \lambda_{t,3,3} \ , \ \lambda_{t,4,1}, \dots, \lambda_{t,4,4} \ , \ \dots \ , \ \lambda_{t,G,1}, \dots, \lambda_{t,G,G} \Big), \tag{A.33}$$

and a second selection matrix $\boldsymbol{S}^{D2} \in \mathbb{R}^{N^2 \times G}$ such that $\operatorname{vec} \boldsymbol{D}_t = \boldsymbol{S}^{D2} \boldsymbol{\sigma}_t^2$ for vector $\boldsymbol{\sigma}_t^2 = (\sigma_{t,1}^2, \ldots, \sigma_{t,G}^2)^\top$. Then the required derivative expressions follow directly from equations (A.29)–(A.31) as

$$\frac{\partial \operatorname{vec} \tilde{\boldsymbol{L}}_t}{\partial \boldsymbol{\lambda}_t^\top} = \boldsymbol{S}^L \cdot \begin{pmatrix} \tilde{\lambda}_{t,1,1}(1 - \tilde{\lambda}_{t,1,1}^2) & 0 & 0 & \ldots & 0 \\ \tilde{\lambda}_{t,2,1}(1 - \tilde{\lambda}_{t,2,1}^2) & -\tilde{\lambda}_{t,2,1}\tilde{\lambda}_{t,2,2}^2 & 0 & \ldots & 0 \\ -\tilde{\lambda}_{t,2,1}^2 \tilde{\lambda}_{t,2,2} & \tilde{\lambda}_{t,2,2}(1 - \tilde{\lambda}_{t,2,2}^2) & 0 & \ldots & 0 \\ \vdots & & & & \vdots \\ -\tilde{\lambda}_{t,G,1}^2 \tilde{\lambda}_{t,G,G} & \ldots & \ldots & -\tilde{\lambda}_{t,G,G-1}^2 \tilde{\lambda}_{t,G,G} & \tilde{\lambda}_{t,G,G}(1 - \tilde{\lambda}_{t,G,G}^2) \end{pmatrix}, \quad \text{(A.34)}$$

and $\partial \operatorname{vec} \boldsymbol{D}_t / \partial \boldsymbol{\lambda}_t^\top$ as in (A.28).