Parallel Bayesian inference for high dimensional dynamic factor copulas

Hoang Nguyen[(a)], M. Concepción Ausín[(a,b)] and Pedro Galeano[(a,b)]

[(a)] Department of Statistics - Universidad Carlos III de Madrid, Spain

[(b)] UC3M-BS Institute of Financial Big Data (IFiBiD)

November 5, 2018

Abstract

To account for asymmetric dependence in extreme events, we propose a dynamic generalized hyperbolic skew Student-t factor copula where the factor loadings follow Generalized Autoregressive Score (GAS) processes. Conditioning on the latent factor, the components of the return series become independent, which allows us to run Bayesian estimation in a parallel setting. Hence, Bayesian inference on different specifications of dynamic one factor copula models can be done in a few minutes. Finally, we illustrate the performance of our proposed models on the returns of 140 companies listed in the S&P500 index. We compare the prediction power of different competing models using Value-at-Risk (VaR), and Conditional Value-at-risk (CVaR), and show how to obtain optimal portfolios in high dimensions based on minimum CVaR.

Keywords: Bayesian inference; Factor copula models; GAS model; Generalized hyperbolic skew Student-t factor copula; Parallel estimation.

1 Introduction

Copulas have become an essential tool for modelling non-standard multivariate distributions as they allow for skewness and fat tails in the marginal distributions and a non-linear dependence structure, see Cherubini et al. (2011), Patton (2012) and Fan and Patton (2014), among others. With a few time series, standard copula families such as the elliptical and the Archimedean copulas are usually applied. However, when the dimension is large, the use of these standard copula distributions is problematic. For instance, the Student-t copula is only able to fit well a few time series, see Demarta and McNeil (2005) and Creal and Tsay (2015). Also, asymmetric dependence is often found, revealing greater correlation during bear markets than during bull markets, see e.g. Erb et al. (1994), Longin and Solnik (2001), Ang and Chen (2002), and Lucas et al. (2014).

The aim of this paper is to propose a parallel Bayesian procedure for handling a large set of financial returns using factor copula models. For that, we use EGARCH processes to model the individual returns. Then, the series of standardized innovations are converted into a series of Uniform(0,1) observations, using cumulative distribution functions, that are assumed to have a copula distribution. To handle a large number of returns, we assume a one factor structure that, first, drastically reduces the number of parameters as they scale linearly with the dimension, and, second, provides natural economic interpretations. To account for asymmetric dependence in extreme events, we propose a group dynamic multivariate generalized hyperbolic skew Student-t (MGSt) factor copula where the factor loadings follow Generalized Autoregressive Score (GAS) processes, see Creal et al. (2013) and Harvey (2013). Importantly, we assume that the dynamic factor loadings equation depends on the copula density conditional on the factor rather than the unconditional copula density, as proposed in Oh and Patton (2017b). The main benefit of our approach is that it allows us to perform parallel inference which greatly reduces the computational cost. Hence, a sizable problem can be fitted from a few minutes up to one hour with a personal computer. The MGSt copula allows for different tail behavior and asymmetric dependence among financial returns.

Factor copulas models have been used previously in the literature as a solution for the curse of dimensionality. For instance, Hull and White (2004) propose a model based on combining linearly the common factor risk and idiosyncratic risk for the Gaussian copula for valuing tranches of collateralized debt obligations and nth to default swaps. Andersen and Sidenius (2004), and van der Voort (2007) improve the model by considering a non-linear factor structure while Murray et al. (2013) extend for multi-factor Gaussian copulas. Oh and Patton (2017a) consider non-normal distributions for the latent factors and the innovations. On the other hand, Krupskii and Joe (2013) propose a general class of factor copulas where the dependence structure is decomposed into a sequence of bivariate copulas and conditional bivariate copulas between each copula variable and the latent factors. Specifically, if bivariate Gaussian linking copulas are used, then the factor copula model can be seen as a copula version of the multivariate Gaussian distribution with a correlation matrix that has a factor structure. Otherwise, if bivariate non-Gaussian linking copulas are used, the model is able to model tail asymmetry and tail dependence, that are important characteristics of financial returns. Nevertheless, Krupskii and Joe (2013) consider the case in which the parameters of the copula functions are static. Close to our research line, Creal and Tsay (2015) propose a class of dynamic stochastic factor copulas. They employ sequential Monte Carlo methods to draw paths of factor loadings and show an empirical illustration with multi-group Student-t one factor copulas. However, the computation cost is expensive even in a parallel setting. Oh and Patton (2017b) offer a dynamic skewed Student-t one factor copula. They apply a variance target strategy to reduce computation cost and use maximum likelihood estimation to estimate the parameters.

Recently, several dynamic copula GARCH models have been proposed for stock returns and exchange rates, where the time-varying parameters are driven by a mean or higher moment functions of lagged information, see Jondeau and Rockinger (2006), Patton (2006a), Ausín and Lopes (2010) and Christoffersen et al. (2012), among others. However, these dynamic models encounter several problems such as specifying how many lag periods and what appropriate dynamic functions should be used. Alternatively, the GAS process, proposed by Creal et al. (2013) and Harvey (2013), is an observation driven process in which the dynamic behavior depends on the score of the predictive likelihood. Particularly, Koopman et al. (2016) find evidence with several simulations and empirical data analysis that the GAS model outperforms models based on moment updating. Oh and Patton (2017b) derive the factor loadings of copula GAS models based on the numerical differentiation of the joint copula density which is computationally expensive in high dimensions. Instead, we consider the latent variables as a source of exogenous information and modify the score based on the predictive likelihood conditional on this exogenous source. This strategy has been used in Lucas et al. (2018) for analyzing high-dimensional banking data through mixture models. Furthermore, we account for the asymmetric dependence in upper and lower tails using the MGSt copula and divide the assets into group sectors. Hence, shocks can affect a group of assets in some circumstances rather than jointly to all series.

In general, the multivariate generalized hyperbolic distribution could be formulated as a location-scale mixture of multivariate Gaussian distributions, see McNeil et al. (2010). Additionally, the multivariate generalized hyperbolic skew Student-t has been proposed for modelling dynamic multivariate stock returns in Mencía and Sentana (2005), for dynamic copula models in Lucas et al. (2017) and for factor stochastic volatility models in Li and Scharth (2018), among others. In comparison to Lucas et al. (2017), we assume a factor structure to reduce the number of parameters in high dimensions. Moreover, we divide variables into groups to model the tail dependence among them. Different from Li and Scharth (2018), we allow for flexible marginal distributions for the returns and concentrate on the dependence structure rather than jointly model the dynamic of conditional mean and conditional volatility.

We compare our proposed dynamic factor copula models with the Exponential Weight Moving Average (EWMA) and Dynamic Conditional Correlation models (DCC), see Engle and Kelly (2012). We find that our proposal performs better for high dimensional time series generated in different stress test scenarios. We also consider several copula specifications including the Gaussian and the Student-t as special cases of the generalized hyperbolic Skew Student-t copulas. We show an empirical example of 140 asset returns for companies listed in S&P 500 index. We found the strongest lower tail dependence among stocks in the Insurance and Finance sectors while other sectors such as Food and Beverage, Pharmacy, and Retail only reveal weak lower tail dependence. We also perform optimal portfolio allocation based on minimization of the CVaR. We use the penalized quantile regression method to prevent extreme positive and negative weights. It also overcomes the computational difficulties in comparison with traditional optimization methods.

The rest of the paper is organized as follows. Section 2 introduces the model for univariate marginal returns and specifies our proposal to model the dependence structure with different types of dynamic factor copula models. We present our parallel Bayesian inference strategy in Section 3 and describe how to perform return prediction and risk management in Section 4. Section 5 illustrates the performance of factor copula models with simulated examples. In Section 6, we analyze a large series of stock returns listed in S&P 500 and compare the prediction power of models using VaR and CVaR. Section 6 also compares the optimal portfolio allocation based on minimizing variance and minimizing CVaR. Finally, conclusions are drawn in Section 7.

2 Dynamic factor copula models

In this section, we introduce our modeling strategy based on the spirit of Creal and Tsay (2015), Oh and Patton (2017a) and Oh and Patton (2017b). For that, the first step is to assume a simple $AR-EGARCH$ structure (Nelson, 1991) on the individual returns and then assume a one factor copula structure on the transformed standardized innovations.

2.1 Model specification

Let $r_t = (r_{1t}, \ldots, r_{dt})'$, for $t = 1, \ldots, T$, be a d-dimensional financial return time series. We assume that each individual return, r_{it}, for $i = 1, \ldots, d$, follows a stationary $AR\,(k_i) - EGARCH\,(p_i, q_i)$ model given by:

$$
\begin{aligned}
r_{it} &= c_i + \phi_{i1} r_{i,t-1} + \ldots + \phi_{ik_i} r_{i,t-k_i} + a_{it} \\
a_{it} &= \sigma_{it}\eta_{it} \\
log(\sigma_{it}^2) &= \omega_i + \sum_{j=1}^{p_i} \beta_{ij} log(\sigma_{i,t-j}^2) + \sum_{j=1}^{q_i} [\alpha_{ij}\eta_{i,t-j} + \gamma_{ij}(|\eta_{i,t-j}| - E|\eta_{i,t-j}|)]
\end{aligned}
$$

where c_i is a constant, $\phi_{i1}, \ldots, \phi_{ik_i}$ are autoregressive parameters verifying the usual stationarity conditions, a_{it} is a sequence of innovations or shocks, σ_{it}^2 is the conditional volatility of the return r_{it}, η_{it} is a sequence of independent standardized innovations with continuous distribution function F_{η_i}, ω_i is a constant, and $\alpha_{i1}, \ldots, \alpha_{iq_i}, \beta_{i1}, \ldots, \beta_{ip_i}, \gamma_{i1}, \ldots, \gamma_{iq_i}$ are $EGARCH$ parameters verifying the usual stationarity conditions. Hence, the $EGARCH$ model takes into account the negative correlation between stock returns and changes in return volatility. We note that the previous $AR-EGARCH$ model can be replaced with any other appropriate specification. For instance, the autoregressive process may be reduced to a simple constant or replaced with an $ARMA$ process, while the $EGARCH$ specification can be replaced with an $GARCH$ (Bollerslev, 1986) or a $GJR-GARCH$ process (Glosten et al., 1993).

Once appropriate models have been specified for all the return series, we can make use of copulas to model their dependence structure. For that, it is well known that $u_{it} = F_{\eta_i}(\eta_{it})$, for each $i = 1, \ldots, d$, is a sequence of independent random variables with a $U(0,1)$ distribution and the dependence structure among the variables in the vector $u_t = (u_{1t}, \ldots, u_{dt})'$ is given by an unknown copula function. A standard approach is to assume that u_t has either a Gaussian copula or a Student-t copula distribution. Nevertheless, it is questionable whether such copula functions are appropriate. One plausible alternative is to assume, as in Krupskii and Joe (2013), a factor copula model in which $u_{1t}, \ldots, u_{dt}$ are conditionally independent given a small set of latent variables. Nevertheless, we consider instead an approach in the spirit of Creal and Tsay (2015), Oh and Patton (2017a) and Oh and Patton (2017b). The idea is to focus on a family of copula models including, among others, the Gaussian, Student-t and generalized hyperbolic skew Student-t copulas, which depend on a conditional scale matrix parameter, R_t, characterized by a factor structure, somehow coming back to standard factor models widely analyzed in the literature. As in Oh and Patton (2017b), we model the dynamic factor loadings as GAS processes, but we assume that the dynamic factor loading equations depend on the copula density conditional on the latent factor rather than the unconditional copula density that allows us to perform parallel inference which heavily reduces the computational cost needed to obtain the conditional posterior distributions of model parameters.

In the next subsections, we describe in detail our proposed dynamic generalized hyperbolic skew Student-t one factor copula model which reduces to Gaussian and Student-t as special cases. We also present some of their advantages over existing alternatives. To simplify, we first present the Gaussian case and then the most general case.

2.2 Dynamic Gaussian one factor copula model

In this subsection, we assume that u_t follows a Gaussian copula with correlation matrix parameter R_t and joint distribution function $C(u_{1t}, \ldots, u_{dt} \mid R_t) = \Phi_d\left(\Phi^{-1}(u_{1t}), \ldots, \Phi^{-1}(u_{dt}) \mid R_t\right)$, where $\Phi(\cdot)$ denotes the univariate standard Gaussian cdf and $\Phi_d(\cdot \mid R_t)$ denotes the multivariate Gaussian cdf with zero mean vector and correlation matrix R_t. Therefore, the vector of inverse cdf transformations, $x_t = (x_{1t}, \ldots, x_{dt})'$, where $x_{it} = \Phi^{-1}(u_{it})$, for each $i = 1, \ldots, d$, follows a multivariate Gaussian distribution with zero mean vector and correlation matrix R_t. We assume a dynamic

Gaussian one factor copula model for x_t given by:

$$x_t = \rho_t z_t + D_t \epsilon_t, \tag{1}$$

where z_t, the latent factor, is a sequence of independent and identically standard Gaussian distributed random variables, $\rho_t = (\rho_{1t}, \ldots, \rho_{dt})'$, is the vector of factor loadings, D_t is a diagonal matrix with elements $\sqrt{1-\rho_{it}^2}$, for $i = 1, \ldots, d$, and $\epsilon_t = (\epsilon_{1t}, \ldots, \epsilon_{dt})'$, is a sequence of independent and identically standard multivariate Gaussian random variables. The latent factor z_t, the idiosyncratic noise ϵ_t, and the dynamic correlations ρ_t are contemporaneously independent. However, the dynamic correlations ρ_t are derived based on the past information of the latent and copula data at time $t-1$. Consequently, the components of the multivariate random vector $x_t = (x_{1t}, \ldots, x_{dt})'$ are conditionally independent given the latent factor z_t and the factor loading vector ρ_t, whose elements, ρ_{it}, represent the correlation between x_{it} and z_t, for $t = 1, \ldots, T$. Therefore, the conditional correlation matrix is given by $R_t = \rho_t \rho_t' + D_t D_t'$. Observe that for the static case, the described model coincides with the one factor Gaussian copula model proposed in Krupskii and Joe (2013). In a dynamic framework, we allow the components of the correlation vector $\rho_t = (\rho_{1t}, \ldots, \rho_{dt})'$ to vary across time as follows,

$$\begin{aligned} \rho_{it} &= \frac{1-\exp(-f_{it})}{1+\exp(-f_{it})} \\ f_{i,t+1} &= (1-b) f_{ic} + a s_{it} + b f_{it} \\ s_{it} &= \frac{\partial \log p(u_t | z_t, f_t, \mathcal{F}_t, \theta)}{\partial f_{it}} \end{aligned} \tag{2}$$

for $i = 1, \ldots, d$, where f_{it} is an observation driven process which fluctuates around a constant value f_{ic}, a and b are two parameters that are assumed to be constant across assets, such that $|b| < 1$ to guarantee stationarity, and $p(u_t | z_t, f_t, \mathcal{F}_t, \theta)$ is the conditional probability density function of u_t given the latent variable, z_t, the random vector $f_t = (f_{1t}, \ldots, f_{dt})'$, the set of all information available at time t, denoted by $\mathcal{F}_t = \{U^{t-1}, F^{t-1}\}$, where $U^{t-1} = \{u_1, \ldots, u_{t-1}\}$ and $F^{t-1} = \{f_0, \ldots, f_{t-1}\}$, and the vector of static parameters, $\theta = (a, b, f_{1c}, \ldots, f_{dc})'$. Note that ρ_{it} is assumed to follow a modified logistic transformation, used also in Dias and Embrechts (2010), Patton (2006b) and Creal et al. (2013), to guarantee that $\rho_{it} \in (-1, 1)$. Also observe that $f_{i,t+1}$ depends linearly on f_{it} and the adjustment term s_{it}. Clearly, this model reduces to a Gaussian time-invariant one factor copula model, see Murray et al. (2013) and Oh and Patton (2017a), when $a = b = 0$.

The dynamic equation (2) is inspired by the GAS model, see Creal et al. (2013) and Harvey (2013), in which the score s_{it} depends on the complete density of u_t rather than on its first or second moment. Blasques et al. (2015) proved that the use of the score s_{it} leads to the minimum Kullback-Leibler divergence between the true conditional density and the model-implied conditional density, while Koopman et al. (2016) showed some empirical examples where the GAS model outperforms other observation driven models. In addition, we consider here the latent variable z_t as a source of exogenous information and derive the observation density conditional on this source. The main reason for such a choice is to reduce dramatically the computational burden as the score s_{it} has a closed form expression that allows us to parallelize the derivation of the d processes $s_{1t}, \ldots, s_{dt}$. Specifically, as shown in Appendix A.1, s_{it} is given by,

$$s_{it} = \frac{1}{2} x_{it} z_t + \frac{1}{2} \rho_{it} - \rho_{it} \frac{x_{it}^2 + z_t^2 - 2\rho_{it} x_{it} z_t}{2\left(1 - \rho_{it}^2\right)}, \tag{3}$$

for $i = 1, \ldots, d$. Therefore, s_{it} depends on the values of the pseudo observable x_{it}, the latent variable z_t, and their mutual correlation ρ_{it}. The model is also attractive, as will be shown in the next subsections, s_{it} has a similar structure to the one given in (3) for the dynamic Student-t and generalized hyperbolic skew Student-t one factor copula models.

As noted before, the main difference of our proposed model with respect to the dynamic GAS model defined in Oh and Patton (2017b) is that the score in (2) is conditioned on the latent variable, z_t. We show in Appendix B that

$$s_{it}^{OP} = \frac{\partial \log p\left(u_t | f_t, \mathcal{F}_t, \theta\right)}{\partial f_{it}} = \mathbf{E}_{z_t}\left[\left.\frac{\partial \log p\left(u_t | z_t, f_t, \mathcal{F}_t, \theta\right)}{\partial f_{it}}\right| u_t, f_t, \theta\right] = \mathbf{E}_{z_t}\left[s_{it} \right| u_t, f_t, \mathcal{F}_t, \theta\left.\right].$$

Thus, the score function (3) is the expectation of the proposal score s_{it} over z_t where z_t follows $p(z_t | u_t, f_t, \mathcal{F}_t, \theta)$ distribution. Therefore, since z_t is sampled from its posterior $p(z_t | x_t, f_t, \mathcal{F}_t, \theta)$, one should expect the use of s_{it} in (3) to be similar to the use of the score function in Oh and Patton (2017b). As mentioned before, our proposed specification allows us to obtain the expressions for s_{it} in parallel for $i = 1, \ldots, d$, reducing the computational burden. This contrasts with Oh and Patton (2017b) where the expressions for s_{it}^{OP} are obtained by the numerical differentiation of the joint copula density, which is much more computationally expensive.

2.3 Dynamic generalized hyperbolic skew Student-t one factor copula model

Next, we use the generalized hyperbolic skew Student-t (GSt) distribution proposed by Aas and Haff (2006) to extend the Gaussian factor copula model. The GSt distribution depends on two parameters, ν and γ, which control the generation of extremes events and skewness, respectively. The GSt distribution reduces to the Student-t distribution when $\gamma = 0$ and reduces to the Gaussian distribution when $\gamma = 0$ and $\nu \to \infty$.

Here, we assume that the joint distribution function of u_t is given by $C(u_{1t}, \dots, u_{dt} \mid R_t, \nu, \gamma) = F_{MGSt}\left(F_{GSt}^{-1}(u_{1t} \mid \nu, \gamma), \dots, F_{GSt}^{-1}(u_{dt} \mid \nu, \gamma) \mid R_t, \nu, \gamma\right)$, where $F_{GSt}(\,\cdot \mid \nu, \gamma)$ denotes the univariate standard GSt cdf with degrees of freedom ν and skewness parameter γ, and $F_{MGSt}(\,\cdot \mid R_t, \nu, \gamma)$ denotes the MGSt cdf with parameters ν and γ and scale matrix R_t. Hence, the MGSt copula allows for asymmetric tail dependence which are not possible with the Gaussian copula assumption. Here, the vector of inverse cdf transformations, $x_t = (x_{1t}, \dots, x_{dt})'$, where $x_{it} = F_{GSt}^{-1}(u_{it} \mid \nu, \gamma)$, for each $i = 1, \dots, d$, follows a MGSt with zero location vector, scale matrix R_t, degrees of freedom ν, and skewness parameter γ. Then, we assume a dynamic generalized hyperbolic skew Student-t one factor copula model for x_t given by:

$$x_t = \gamma\zeta_t + \sqrt{\zeta_t}\,(\rho_t z_t + D_t \epsilon_t) \tag{4}$$

for $i = 1, \dots, d$, where z_t, ϵ_t and ρ_t, for $t = 1, \dots, T$, are as in the Gaussian case, and ζ_t is a sequence of independent and identically inverse Gamma distributed random variables with parameters $\left(\frac{\nu}{2}, \frac{\nu}{2}\right)$, denoted by $IG\left(\frac{\nu}{2}, \frac{\nu}{2}\right)$, and independent of z_t, ϵ_t and ρ_t. Particularly, when $\gamma = 0$, x_t follows multivariate Student-t distribution as a special case. In any case, the components of the multivariate random vector $x_t = (x_{1t}, \dots, x_{dt})'$ are contemporaneously independent at time t given z_t, ρ_t and ζ_t. However, note that ρ_t depends on the past values of x_t, z_t and ζ_t through the GAS process.

As in the Gaussian case, the vector $\rho_t = (\rho_{1t}, \dots, \rho_{dt})'$ is allowed to vary across time as in (2), but replacing the value of the score s_{it} with,

$$s_{it} = \frac{\partial \log p\,(u_t | z_t, \zeta_t, f_t, \mathcal{F}_t, \theta)}{\partial f_{it}},$$

where $p\left(u_t|z_t,\zeta_t,f_t,\mathcal{F}_t,\theta\right)$ is the conditional probability density function of u_t given z_t, ζ_t, f_t, $\mathcal{F}_t$, and the parameters of the copula function, $\theta=(a,b,\nu,\gamma,f_{1c},\ldots,f_{dc})'$. Again, this model setting is influenced by the developments in Creal and Tsay (2015) for stochastic factor copulas and Oh and Patton (2017b) for dynamic factor copulas. However, one advantage of our proposal is that the observation driven process remains similar. As shown in Appendix A.2, if we let $\tilde{x}_{it}=\frac{x_{it}-\gamma\zeta_t}{\sqrt{\zeta_t}}$, the score function is,

$$s_{it}=\frac{\partial\log p(u_t|z_t,\zeta_t,f_t,\mathcal{F}_t,\theta)}{\partial f_{it}}=\frac{1}{2}\tilde{x}_{it}z_t+\frac{1}{2}\rho_{it}-\rho_{it}\frac{\tilde{x}_{it}^2+z_t^2-2\rho_{it}\tilde{x}_{it}z_t}{2\left(1-\rho_{it}^2\right)},\tag{5}$$

which is similar to the score function in (3). Consequently, we enjoy here the same computational advantages described in the Gaussian case. On the other hand, this proposed model is different from the skew Student-t factor copula model in Oh and Patton (2017a) and Oh and Patton (2017b) since these authors consider different symmetric and asymmetric Student-t distributions for z_t and ϵ_t. Their models do not lead to an easily attainable conditional cdf for x_t and therefore, it is computationally expensive to derive the score s_{it}, as mentioned before.

Demarta and McNeil (2005) noted that the marginal univariate GSt only has finite variance when $\nu>4$ in comparison with the Student-t distribution which requires $\nu>2$. They also differ in the tail decay. While the Student-t density has the tail decay as $x^{-\nu-1}$, the GSt density has a heaviest tail decay as $x^{-\nu/2-1}$ and the lightest tail as $x^{-\nu/2-1}\exp\left(-2|\gamma x|\right)$ (for $\gamma\neq 0$). We obtain the tail dependence of the dynamic MGSt one factor copula model using a numerical approximation of the joint quantile exceedance probability, see Appendix C. Finally, Demarta and McNeil (2005) suggested several extensions for more complex copula functions. For example, when ζ_t follows a generalized inverse Gaussian distribution, x_{it} is generalized hyperbolic distributed. Also, one could propose different distributions of the type $x_{it}=\gamma_g h\left(\zeta_t\right)+\sqrt{\zeta_t}\left(\rho_{it}z_t+\sqrt{1-\rho_{it}^2}\epsilon_{it}\right)$, where $h\left(\zeta_t\right)$ is a function of ζ_t. However, the properties of x_{it} would generally be intractable.

2.4 Dynamic group generalized hyperbolic skew Student-t one factor copulas

One potential drawback of the previous models is that only a few parameters control all of the tail co-movements which can be very restrictive for high dimensional returns. In order to relax this assumption, our strategy is to split the d assets into G groups in such a way that returns in the same group have similar characteristics.

Therefore, we write $u_t = (u'_{1t}, \ldots, u'_{Gt})'$, where $u_{gt} = \left(u_{1gt}, \ldots, u_{n_g gt}\right)'$, for $g = 1, \ldots, G$ and $\sum_{g=1}^{G} n_g = d$. In the most general case of the MGSt copula, we define $x_{igt} = F_{GSt}^{-1}(u_{igt}|\nu_g, \gamma_g)$ for each asset i, for $i = 1, \ldots, n_g$, belonging to group g, where $g = 1, \ldots, G$, such that,

$$x_{gt} = \gamma_g \zeta_{gt} + \sqrt{\zeta_{gt}}\left(\rho_{gt} z_t + D_{gt}\epsilon_{gt}\right) \tag{6}$$

where $x_{gt} = \left(x_{1gt}, \ldots, x_{n_g gt}\right)'$ is the vector of inverse transformations in group g, $\rho_{gt} = \left(\rho_{1gt}, \ldots, \rho_{n_g gt}\right)'$ is the vector of factor loadings in group g, and D_{gt} is the diagonal matrix with elements $\sqrt{1-\rho_{igt}^2}$ and $\epsilon_{gt} = \left(\epsilon_{1gt}, \ldots, \epsilon_{n_g gt}\right)'$ are, respectively, the corresponding diagonal matrix and noise vector in group g.

Observe that the set of mixing variables $\zeta_t = (\zeta_{1t}, \ldots, \zeta_{Gt})'$ create G multivariate MGSt distributions with degrees of freedom parameters $\nu_1, \ldots, \nu_G$ and skewness parameters $\gamma_1, \ldots, \gamma_G$, respectively. Then, the dynamic of the i-th the scale parameters in group g is given by:

$$\begin{aligned} \rho_{igt} &= \frac{1-\exp\left(-f_{igt}\right)}{1+\exp\left(-f_{igt}\right)} \\ f_{ig,t+1} &= \left(1-b_g\right) f_{igc} + a_g s_{igt} + b_g f_{igt} \end{aligned} \tag{7}$$

where the set of parameters $a = (a_1, \ldots, a_G)'$ and $b = (b_1, \ldots, b_G)'$ adjust the dynamic behavior of the scale parameters in each group g. Here, the i-th score in group g is given by:

$$s_{igt} = \frac{1}{2}\tilde{x}_{igt} z_t + \frac{1}{2}\rho_{igt} - \rho_{igt}\frac{\tilde{x}_{igt}^2 + z_t^2 - 2\rho_{igt}\tilde{x}_{igt} z_t}{2\left(1-\rho_{igt}^2\right)} \tag{8}$$

where $\tilde{x}_{igt} = \frac{x_{igt} - \gamma_g \zeta_{gt}}{\sqrt{\zeta_{gt}}}$. Note that when $G = 1$, the model reduces to the copula specification proposed in the previous section.

The model becomes extremely flexible by assuming that each series has its own dynamic group. Indeed, the model is able to capture the different behaviors in the upper and lower tail dependence for those assets in the same group. However, note that the assets in different groups show no tail dependence due to the independence assumption among the components of ζ_t. Also, the pseudo observable $x_{igt} = F_{GSt}^{-1}(u_{igt}|\nu_g, \gamma_g)$ requires an intensive computation as long as ν_g and γ_g receive new trial values. A parallel Bayesian algorithm is implemented in the next section to speed up calculations.

3 Bayesian inference

In this section, we present our parallel Bayesian inference strategy to obtain the posterior distribution of the model parameters of the dynamic one-factor copula models presented in Section 2.

3.1 Prior distributions

We focus on defining a prior distribution for the copula parameters. In all cases, we use proper but uninformative prior assumptions. We describe the prior for the most general proposed model, the group MGSt factor copula, which contains all other models as particular cases. First, we assume uniform priors for all the elements in $f_c = \{f_{igc} : g = 1, \ldots, G; i = 1, \ldots, n_g\}$. More precisely, we assume a priori that $f_{igc} \sim U(-5, 5)$, so that the value of ρ_{igc} ranges between $(-0.9866, 0.9866)$. Additionally, f_{11c} is restricted to be positive to guarantee model identifiability. Second, as usual in GAS models, we assume uniform priors for all the elements in $a = \{a_g : g = 1, \ldots, G\}$ and $b = \{b_g : g = 1, \ldots, G\}$. More precisely, we assume a priori that $a_g \sim U(-0.5, 0.5)$ and $b_g \sim U(0, 1)$. Third, we assume a prior shifted Gamma distributions for all the degrees of freedom parameters in $\nu = \{\nu_g : g = 1, \ldots, G\}$, such that $\nu_g = 4 + \tilde{\nu}_g$, where $\tilde{\nu}_g \sim G(2, 2.5)$, in order that the variance of the pseudo observations, x_{it}, is finite. Fourth, we assume a priori a standard Gaussian distribution for all the skewness parameters in $\gamma = \{\gamma_g : g = 1, \ldots, G\}$, i.e., a priori $\gamma_g \sim N(0, 1)$, for $g = 1, \ldots, G$. In the particular case of a Student-t copula, we assume that ν_g follows a priori a shifted Gamma distribution with $\nu_g = 2 + \tilde{\nu}_g$, such that the variance of x_{it} is finite and set the skewness parameter $\gamma_g = 0$.

Finally, the latent states $z = \{z_t : t = 1, \ldots, T\}$ are treated as nuisance independent parameters following independent $N(0, 1)$ distributions, as considered in the model assumptions. Additionally, the elements of $\zeta = \{\zeta_{gt} : g = 1, \ldots, G; t = 1, \ldots, T\}$ are nested as nuisance parameters for the realization of the pseudo observations x_{it} and depend on the respective elements of ν.

3.2 Posterior inference

Given a sample of return data, $r = \{r_t : t = 1, \ldots, T\}$, and the priors defined before, we are interested in the posterior of the model parameters given by the set of marginal parameters, $\vartheta_i = (c_i, \phi_{i1}, \ldots, \phi_{ik_i}, \omega_i, \alpha_{i1}, \ldots, \alpha_{ip_i}, \beta_{i1}, \ldots, \beta_{iq_i}, \gamma_{i1}, \ldots, \gamma_{ip_i})'$, and the set of factor copula parameters, $\vartheta_c = (a, b, \nu, \gamma, z, \zeta, f_c)'$. The likelihood is given by,

$$l\left(\vartheta_1, \ldots, \vartheta_d, \vartheta_c \mid r\right) = \prod_{t=1}^{T} c(F_{\eta_1}(\eta_{1t} \mid \vartheta_1), \ldots, F_{\eta_d}(\eta_{dt} \mid \vartheta_d) \mid \vartheta_c) \prod_{i=1}^{d} f_{\eta_i}(\eta_{it} \mid \vartheta_i),$$

where $c\left(\cdot \mid \vartheta_c\right)$ denotes the copula density function with parameters ϑ_c and $f_{\eta_i}(\eta_i \mid \vartheta_i)$ is the marginal density function of the standardized innovations, η_{it}. Given this decomposition of the likelihood, we follow the standard two-stage estimation procedure for copulas where, in a first step, we estimate the marginal parameters, $\hat{\vartheta}_i$, independently using the maximum likelihood for each $i = 1, \ldots, d$, and, in a second step, we obtain an approximate sample of the copula observations, $u = \{u_t : t = 1, \ldots, T\}$, where $u_{it} = F_{\eta_i}\left(\eta_{it} \mid \hat{\vartheta}_i\right)$, for $t = 1, \ldots, T$ and for each $i = 1, \ldots, d$. This two-stage estimation procedure has been shown to be statistically efficient by Joe (2005) and Chen and Fan (2006) in case of parametric and semi-parametric distributions for standardized residuals. Alternatively, a fully Bayesian approach where the joint posterior distribution is approximated in a single step would be done but the two-step approach simplifies enormously the computational burden in the high dimensional setting that we are considering.

Now, considering the G different asset groups, we assume that the matrix sample of copula observations, $u = \{u_t : t = 1, \ldots, T\}$, is such that $u_t = (u'_{1t}, \ldots, u'_{Gt})'$, where $u_{gt} = \left(u_{1gt}, \ldots, u_{n_g gt}\right)'$, for $g = 1, \ldots, G$. Then, the likelihood of the MGSt copula is given by:

$$l\left(\vartheta_c \mid u\right) = \prod_{t=1}^{T} p\left(u_t | z_t, \zeta_t, f_t, \mathcal{F}_t, \theta\right),$$

where $f_t = (f_{1t}, \ldots, f_{Gt})'$ with $f_{gt} = \left(f_{1gt}, \ldots, f_{n_g gt}\right)$, for $g = 1, \ldots, G$. Recall that $\mathcal{F}_t = \{U^{t-1}, F^{t-1}\}$, where $U^{t-1} = \{u_1, \ldots, u_{t-1}\}$ and $F^{t-1} = \{f_0, \ldots, f_{t-1}\}$, and $\theta = (a, b, \nu, \gamma, f_c)'$ is the vector of static parameters. Therefore, given the conditional density (11) in Appendix A.2, the likelihood is given by:

$$p\left(u|z,\zeta,f_c,a,b,\nu,\gamma\right)=\prod_{t=1}^{T}\prod_{g=1}^{G}\prod_{i=1}^{n_g}\frac{\phi\left(\frac{F_{GSt}^{-1}(u_{igt}|\nu)-\gamma_g\zeta_t}{\sqrt{\zeta_{gt}}}\mid\rho_{igt}z_t,\sqrt{1-\rho_{igt}^2}\right)}{f_{GSt}\left(F_{GSt}^{-1}\left(u_{igt}\mid\nu_g,\gamma_g\right)\mid\nu_g,\gamma_g\right)\sqrt{\zeta_{gt}}}.$$

As a result, the joint posterior density of the group dynamic MGSt factor copula parameters can be written as follows:

$$\begin{aligned}p\left(z,\zeta,f_c,a,b,\nu,\gamma|u\right)\propto&\prod_{t=1}^{T}\prod_{g=1}^{G}\prod_{i=1}^{n_g}\frac{\phi(\tilde{x}_{igt}|\rho_{igt}z_t,\sqrt{1-\rho_{igt}^2})}{f_{GSt}\left(x_{igt}|\nu_g,\gamma_g\right)\sqrt{\zeta_{gt}}}\prod_{t=1}^{T}\phi(z_t|0,1)\\&\times\prod_{t=1}^{T}\prod_{g=1}^{G}IG\left(\zeta_{gt}|\frac{\nu_g}{2},\frac{\nu_g}{2}\right)\prod_{g=1}^{G}G\left(\nu_g-4|2,2.5\right)\prod_{g=1}^{G}\phi\left(\gamma_g|0,1\right),\end{aligned}\tag{9}$$

where $\tilde{x}_{igt}=\frac{x_{igt}-\gamma_g\zeta_{gt}}{\sqrt{\zeta_{gt}}}$ and $x_{igt}=F_{GSt}^{-1}\left(u_{it}\mid\nu_g,\gamma_g\right)$.

The computation of the pseudo observable $x_{igt}=F_{GSt}^{-1}(u_{igt}|v_g,\gamma_g)$ is often time-consuming especially when the value of v_g and γ_g change in each MCMC iteration. We create a sequence of $m=1000$ values with equal increment in the range $x_{seq}=[x_{Low},x_{High}]$ and find their exact cdf $u_{seq}=F_{GSt}(x_{seq}|v_g,\gamma_g)$. The approximate values of x_{igt} is calculated as the linear interpolation between two nearest neighbors in the sequence. We employ the algorithm in the SkewHyperbolic package (Scott and Grimson (2015)) to find out the reasonable range $[x_{Low},x_{High}]$ which guarantees to cover all the values of x_{igt} and also that the relative difference between the approximate and the exact value of x_{igt} is no more than 1%.

3.3 MCMC algorithm

Here, a parallel algorithm is exploited to obtain a posterior sample of the model parameters. Due to the fact that the conditional posterior of z_t is Gaussian, we can make fast inference for each latent variable at time $t=1,\ldots,T$. Also, the conditional posterior of a_g, b_g, ν_g, γ_g, and ζ_{gt} can be sampled in parallel for the groups $g=1,\ldots,G$, where G is usually a moderate number. Finally, since conditional on z_t, each component of x_t is independent, we can create a parallel estimation procedure for f_{igc} for $i=1,\ldots,n_g$ and $g=1,\ldots,G$. Thus, the algorithm is scalable in high dimensional returns.

1. Set initial values for $\vartheta^{(0)} = \left(z^{(0)}, f_c^{(0)}, a^{(0)}, b^{(0)}, \nu^{(0)}, \gamma^{(0)}, \zeta^{(0)}\right)$.

2. For iteration $j = 1, \dots, N$, obtain $\rho_{igt}^{(j)}$ for $i = 1, \dots, n_g$, $g = 1, \dots, G$ and $t = 1, \dots, T$:

   (a) For $t = 1, \dots, T$, sample $z_t^{(j)} \sim p\left(z_t | u, a^{(j-1)}, b^{(j-1)}, f_c^{(j-1)}, \nu^{(j-1)}, \gamma^{(j-1)}, z_{1:(t-1)}^{(j)}, \zeta^{(j-1)}\right)$.

   (b) Parallel for $i = 1, \dots, n_g$ and $g = 1, \dots, G$, sample

$$f_{igc}^{(j)} \sim p\left(f_{igc} | u, a^{(j-1)}, b^{(j-1)}, z^{(j)}, \nu^{(j-1)}, \gamma^{(j-1)}, \zeta^{(j-1)}\right).$$

   (c) Parallel for $g = 1, \dots, G$, sample $a_g^{(j)} \sim p\left(a_g | u, b^{(j-1)}, f_c^{(j)}, z^{(j)}, \nu^{(j-1)}, \gamma^{(j-1)}, \zeta^{(j-1)}\right)$.

   (d) Parallel for $g = 1, \dots, G$, sample $b_g^{(j)} \sim p\left(b_g | u, a^{(j)}, f_c^{(j)}, z^{(j)}, \nu^{(j-1)}, \gamma^{(j-1)}, \zeta^{(j-1)}\right)$.

   (e) Parallel for $g = 1, \dots, G$, sample $\nu_g^{(j)} \sim p\left(\nu_g | u, a^{(j)}, b^{(j)}, f_c^{(j)}, z^{(j)}, \gamma^{(j-1)}, \zeta^{(j-1)}\right)$.

   (f) Parallel for $g = 1, \dots, G$, sample $\gamma_g^{(j)} \sim p\left(\gamma_g | u, a^{(j)}, b^{(j)}, f_c^{(j)}, z^{(j)}, \nu^{(j)}, \zeta^{(j-1)}\right)$.

   (g) Parallel for $g = 1, \dots, G$, sample $\zeta_{gt}^{(j)} \sim p\left(\zeta_{gt} | u, a^{(j)}, b^{(j)}, f_c^{(j)}, z^{(j)}, \nu^{(j)}, \gamma^{(j)}, \zeta_{g,1:(t-1)}^{(j)}\right)$ for $t = 1, \dots, T$.

The conditional posterior distributions for all the parameters are given in Appendix D. In the algorithm, we apply the Gibbs sampler for step $2a$ and the Adaptive Random Walk Metropolis Hasting (ARWMH) (see Roberts and Rosenthal (2009)) for steps $2b$ to $2f$. As suggested by Creal and Tsay (2015), we use the independent MH in step $2g$ to generate new values of $\log\left(\zeta_{gt}^{(j)}\right)$ from a Student-t distribution with 4 degrees of freedom with mean equal to the mode and scale equal to the inverse Hessian at the mode. Logarithms guarantee that $\zeta_{gt}^{(j)}$ is positive. Thus, for each time period t, we accept $\zeta_{gt}^{(j)}$ with probability:

$$\min\left\{1, \frac{p\left(\zeta_{gt}^{(j)} | u, a^{(j)}, b^{(j)}, f_c^{(j)}, z^{(j)}, \nu^{(j)}, \gamma^{(j)}, \zeta_{g,1:(t-1)}^{(j)}\right) q\left(\zeta_{gt}^{(j-1)}\right)}{p\left(\zeta_{gt}^{(j-1)} | u, a^{(j)}, b^{(j)}, f_c^{(j)}, z^{(j)}, \nu^{(j)}, \gamma^{(j)}, \zeta_{g,1:(t-1)}^{(j)}\right) q\left(\zeta_{gt}^{(j)}\right)}\right\}.$$

Observe that this Bayesian algorithm reduces to steps $2a$ to $2d$ for the dynamic Gaussian one factor copula. Also, step $2f$ is omitted for the dynamic Student-t one factor copula since $\gamma = 0$. The codes and implementation of the algorithm are available at https://github.com/hoanguc3m/FactorCopula.

4 Prediction of returns and risk management

In this section, we illustrate how the estimated copula models help to predict returns and measure the risk of the portfolio such as portfolio variance, quantile of the portfolio's profit/loss distribution for a given horizon (VaR) and conditional expected loss above a quantile (CVaR). Finally, we employ a simulation procedure to allocate an optimal portfolio based on minimum variance and minimum CVaR.

4.1 Prediction of returns

Based on the MCMC samples from the conditional posterior distribution of copula parameters $\vartheta_c^{(n)} = \left(a^{(n)}, b^{(n)}, \nu^{(n)}, \gamma^{(n)}, z^{(n)}, \zeta^{(n)}, f_c^{(n)}\right)$, for $n = 1, \ldots, N$, we can obtain the distribution of the predicted return $r_t = \{r_{i,t} : i = 1, \ldots, d\}$ at time $t = T + 1$. For the sake of simplicity, we consider $AR(1) - EGARCH(1,1)$ for the marginal and generate replications of one-step-ahead predicted return $(r_{1t}^{(n)}, \ldots, r_{dt}^{(n)})$ as follows,

$$\begin{aligned}
r_{it}^{(n)} &= \hat{c}_i + \hat{\phi}_{i1} r_{i,t-1} + a_{it}^{(n)} \\
a_{it}^{(n)} &= \sigma_{it} \eta_{it}^{(n)} \\
log(\sigma_{it}^2) &= \hat{\omega}_i + \hat{\alpha}_{i1} \eta_{i,t-1} + \hat{\gamma}_{i1}(|\eta_{i,t-1}| - E|\eta_{i,t-1}|) + \hat{\beta}_{i1} log(\sigma_{i,t-1}^2)
\end{aligned}$$

where $\hat{\vartheta}_i = \left(\hat{c}_i, \hat{\phi}_{i1}, \hat{\omega}_i, \hat{\alpha}_{i1}, \hat{\beta}_{i1}, \hat{\gamma}_{i1}\right)'$ is the set of marginal parameters in $AR(1) - EGARCH(1,1)$ model. The standardized innovation is obtained as $\eta_{it}^{(n)} = F_{\eta_i}^{-1}(u_{igt}^{(n)}) = F_{\eta_i}^{-1}(F_{GSt}(x_{igt}^{(n)} | \vartheta_c^{(n)}))$ and the value of $x_{igt}^{(n)}$ is generated from Equations (6 - 8) where $\zeta_{gt}^{(n)} \sim IG(\frac{\nu^{(n)}}{2}, \frac{\nu^{(n)}}{2})$, $z_t^{(n)} \sim N(0,1)$, and $\epsilon_{igt}^{(n)} \sim N(0,1)$, i.e.,

$$
\begin{aligned}
x_{igt}^{(n)} &= \gamma_g^{(n)}\zeta_{gt}^{(n)} + \sqrt{\zeta_{gt}^{(n)}}\left(\rho_{igt}^{(n)} z_t^{(n)} + \sqrt{(1-\rho_{igt}^{(n)2})}\epsilon_{igt}^{(n)}\right) \\
\rho_{igt}^{(n)} &= \frac{1-\exp\left(-f_{igt}^{(n)}\right)}{1+\exp\left(-f_{igt}^{(n)}\right)} \\
f_{igt}^{(n)} &= \left(1-b_g^{(n)}\right) f_{igc}^{(n)} + a_g^{(n)} s_{ig,t-1}^{(n)} + b_g^{(n)} f_{ig,t-1}^{(n)} \\
s_{ig,t-1}^{(n)} &= \frac{1}{2}\tilde{x}_{ig,t-1}^{(n)} z_{t-1}^{(n)} + \frac{1}{2}\rho_{ig,t-1}^{(n)} - \rho_{ig,t-1}^{(n)} \frac{\tilde{x}_{ig,t-1}^{(n)2} + z_{t-1}^{(n)2} - 2\rho_{ig,t-1}^{(n)}\tilde{x}_{ig,t-1}^{(n)} z_{t-1}^{(n)}}{2\left(1-\rho_{ig,t-1}^{(n)2}\right)} \\
\tilde{x}_{ig,t-1}^{(n)} &= \frac{x_{ig,t-1}^{(n)} - \gamma_g^{(n)}\zeta_{g,t-1}^{(n)}}{\sqrt{\zeta_{g,t-1}^{(n)}}} = \frac{F^{-1}(u_{ig,t-1}|\gamma_g^{(n)},\nu_g^{(n)}) - \gamma_g^{(n)}\zeta_{g,t-1}^{(n)}}{\sqrt{\zeta_{g,t-1}^{(n)}}}
\end{aligned}
$$

We can also obtain the distribution of predicted return at time $T+h$, where $h>1$, conditional on the return information at time $t=T+h-1$. As the return prediction needs information about the latent variables z_t and ζ_{gt}, we choose z_t and ζ_{gt} as the maximum a posteriori of its conditional posterior distribution when obtaining new data.

4.2 Risk measurement

Assume that we have a portfolio constructed with the return series $r_{1t},\dots,r_{dt}$. Then, the total return at time t is calculated as,

$$r_t = \sum_{i=1}^{d} \delta_{it} r_{it}$$

where $\delta_t = \{\delta_{it}\}_{i=1}^d$ is the set of asset weights in the portfolio at time t such that $\sum_{i=1}^d \delta_{it} = 1$. The $q\%$ VaR is the threshold loss value such that the probability of a loss exceeds VaR is q, over the time horizon t, i.e.,

$$q = \mathtt{Pr}\left(\sum_{i=1}^{d} \delta_{it} r_{it} \le -\mathtt{VaR}_{q,t}\right).$$

Similarly, the CVaR is the conditional expected loss above $q\%$ VaR, i.e.,

$$\mathtt{CVaR}_{q,t} = -\mathtt{E}\left(\sum_{i=1}^{d} \delta_{it} r_{it} \,\middle|\, \sum_{i=1}^{d} \delta_{it} r_{it} \le -\mathtt{VaR}_{q,t}\right).$$

Here, we estimate the one-step-ahead $VaR_{q,t}$ and $CVaR_{q,t}$ for the portfolio of equal weight. In the previous section, we obtain the distribution of one-step-ahead predicted return $\{(r_{1,t}^{(n)},\dots,r_{d,t}^{(n)})\}_{t=T+1}^{T+H}$.

Then, it is easy to obtain the predictive $\mathrm{VaR}_{q,t}$ and $\mathrm{CVaR}_{q,t}$ using the return simulation. The estimated VaR_q and CVaR_q are the average of $\{\mathrm{VaR}_{q,t}\}_{t=T+1}^{T+H}$ and $\{\mathrm{CVaR}_{q,t}\}_{t=T+1}^{T+H}$ along the prediction period. We compare the prediction powers of the proposed copula models using backtesting for VaR. The expected number of days that the realized portfolio return goes below the $VaR_{q,t}$ threshold is qH.

4.3 Optimal portfolio allocation

Next, we can take advantage of the predicted returns above for active portfolio allocation. Classically, Markowitz (1952) introduces portfolio allocation theory based on the mean-variance approach. The optimal weight for the minimum mean-variance problem is obtained by solving

$$\hat{\delta}_t = \arg\min_{\delta_t} \left\{ \delta_t^{'} \Sigma_t \delta_t : \delta_t^{'} \mathbf{1} = 1, \delta_t^{'} \mu_t = \mu_0 \right\}$$

where μ_t and Σ_t are the expected return and the covariance matrix of the assets in the portfolio at time t, and μ_0 is the expected return. Jagannathan and Ma (2003) recommend imposing nonnegative constraints on portfolio weights ($\delta_t > 0$). This strategy is not only commonly used by practitioners but also improves the efficiency of optimal portfolios using sample moments. In the empirical illustration, we show an example of an optimal portfolio using minimum variance, as follows:

$$\hat{\delta}_t^{(Var)} = \arg\min_{\delta_t} \left\{ V\left(\sum_{i=1}^{d} \delta_{it} r_{it} \right) : \delta_t^{'} \mathbf{1} = 1, \delta_t^{'} \geq 0 \right\}$$

Alternatively, the common optimization problem is to obtain the portfolio with minimum VaR or CVaR. Alexander and Baptista (2004) compare the portfolio selection using VaR and CVaR and recommend CVaR as a more appropriate tool for risk management. However, the minimum CVaR portfolio is often time consuming in high dimensions and results in extreme asset weights. Xu et al. (2016) deal with this issue by proposing a weight constraint on the minimum CVaR portfolio,

$$\hat{\delta}_t^{(CVaR)} = \arg\min_{\delta_t} \left\{ \mathtt{CVaR}_{q,t} + \lambda_t \sum_{i=1}^{d} Pen(\delta_{it}) : \delta_t^{'} \mathbf{1} = 1 \right\}$$

where λ_t is a penalty parameter and the Pen function can be chosen as the LASSO (Tibshirani (1996)), or SCAD (Fan and Li (2001)) penalty functions, among others. Following Bassett Jr et al.

(2004), the CVaR can be written as,

$$\texttt{CVaR}_{q,t} = q^{-1} \arg\min_{\xi_t} E\rho_q \left[r_t - \xi_t \right] - \mu_t$$

where ξ_t is the q quantile of r_t. The quantile loss function $\rho_q [u] = u(q - I(q < 0))$ as defined in Koenker and Bassett Jr (1978). Note that

$$r_t = \sum_{i=1}^{d} \delta_{it} r_{it} = r_{1t} - \sum_{i=2}^{d} \delta_{it}(r_{1t} - r_{it})$$

Let $Y_t = r_{1t}$ and $X_{it} = r_{1t} - r_{it}$. Then, it is straightforward to write the optimal portfolio problem with LASSO penalty as a Lasso penalized quantile regression,

$$\hat{\delta}_t^{(CVaR)} = \arg\min_{\delta_t, \xi_t} E\rho_q \left[Y_t - \sum_{i=2}^{d} \delta_{it} X_{it} - \xi_t \right] - \lambda_t \sum_{i=1}^{d} |\delta_{it}|$$

where the factor q is absorbed into the penalty term, and μ_t is the constant at time t. We choose a λ_t for each period based on the minimum BIC value for the penalized quantile regression (see Lee et al. (2014))

$$\hat{\lambda}_t = \arg\min_{\lambda_t} \log \left(\sum_{n=1}^{N} \rho_\tau (Y_t^{(n)} - \sum_{i=2}^{d} \delta_{it} X_{it}^{(n)} - \xi_t) \right) + |S| \frac{\log N}{2N}$$

where N is the number of return simulation and $|S|$ is the number of points in the set S such that $S = \{i : \hat{\delta}_{it,\lambda} \neq 0, i \in [2, p]\}$. We substitute the optimal weights in each period to obtain $CVaR_{q,t}$

5 Simulation study

5.1 Simulated data

In this section, we illustrate the proposed Bayesian methodology using simulated data from the MGSt one factor copula in Section 2.4. We generate a random sample of $d = 100$ time series with $G = 10$ groups of different sizes and a time length $T = 1000$ from Equations (6) to (8). The value of the parameters $\vartheta_c = (a, b, \nu, \gamma, z, \zeta, f_c)'$ are randomized. More precisely, a is generated from a $U(0.05, 0.10)$ distribution, b is generated from a $U(0.95, 0.985)$, ν is generated from a $U(6, 18)$,

γ is generated from a $U(-1,0)$ distribution, z_t is generated from a $\Phi(0,1)$ distribution, and ζ_{gt} is generated from an $IG(\nu_g/2, \nu_g/2)$ where $t = 1, \dots, T, g = 1, \dots, G$. The expected correlation between pseudo observation x_t and the latent factor z_t are sampled from a $U(0.1, 0.9)$ distribution, which results in values for f_{igc} ranging in the interval $(0.2, 3)$.

We estimate the set of true parameters, ϑ_c, using 20.000 MCMC iterations where the first 10.000 are discarded as burn-in iterations. The algorithm seems to perform adequately and convergence is fast. Practically, all the posteriors reached convergence after 1000 iterations. We retain every 10-th iterations to reduce autocorrelation. The algorithm takes around 25 minutes, 70 minutes and 90 minutes for the Gaussian, Student and MGSt one factor copula model, respectively, on an Intel Core i7-4770 processor (4 cores - 8 threads - 3.4GHz).

Figure 1 shows the box plots of the posterior sample from the MCMC together with the true values of the model parameters. Observe that the true values of a_g, b_g, ν_g and γ_g lie between the first and the third quantile of the credible intervals in 50% of the cases and never reach out of their whiskers. The posterior distributions of b_g are skewed to the left with heavier tails. Also, the posterior samples show larger variances for higher values of the degree of freedom parameters ν_g. We have observed that there is a negative correlation between MCMC samples of ν_g and γ_g which means that if the posterior mean of ν_g underestimates its true value, the value of γ_g will overestimate its true value. However, the effect is weakly observed. We select some values of f_c and z_t to illustrate the comparison between the posterior mean of f_{igc} versus its true value, for $i = 1, \dots, d$ and $g = 1 \dots, G$, and z_t versus its true value, for $t = 1, \dots, T$. We obtain quite accurate results. The posterior variance of z_t also reduces when the dimension increases. We obtain a smaller posterior standard deviation of ρ_c when its true value is high. In general, most of the parameters which govern the dynamic dependence in each group are correctly estimated. We perform a Monte Carlo study in Online Appendix.

5.2 Comparison of estimators

Next, we compare several dynamic correlation models in different scenarios based on Engle (2002) and Creal et al. (2011)'s proposal. We generate $d = 10$ time series from a multivariate Gaussian,

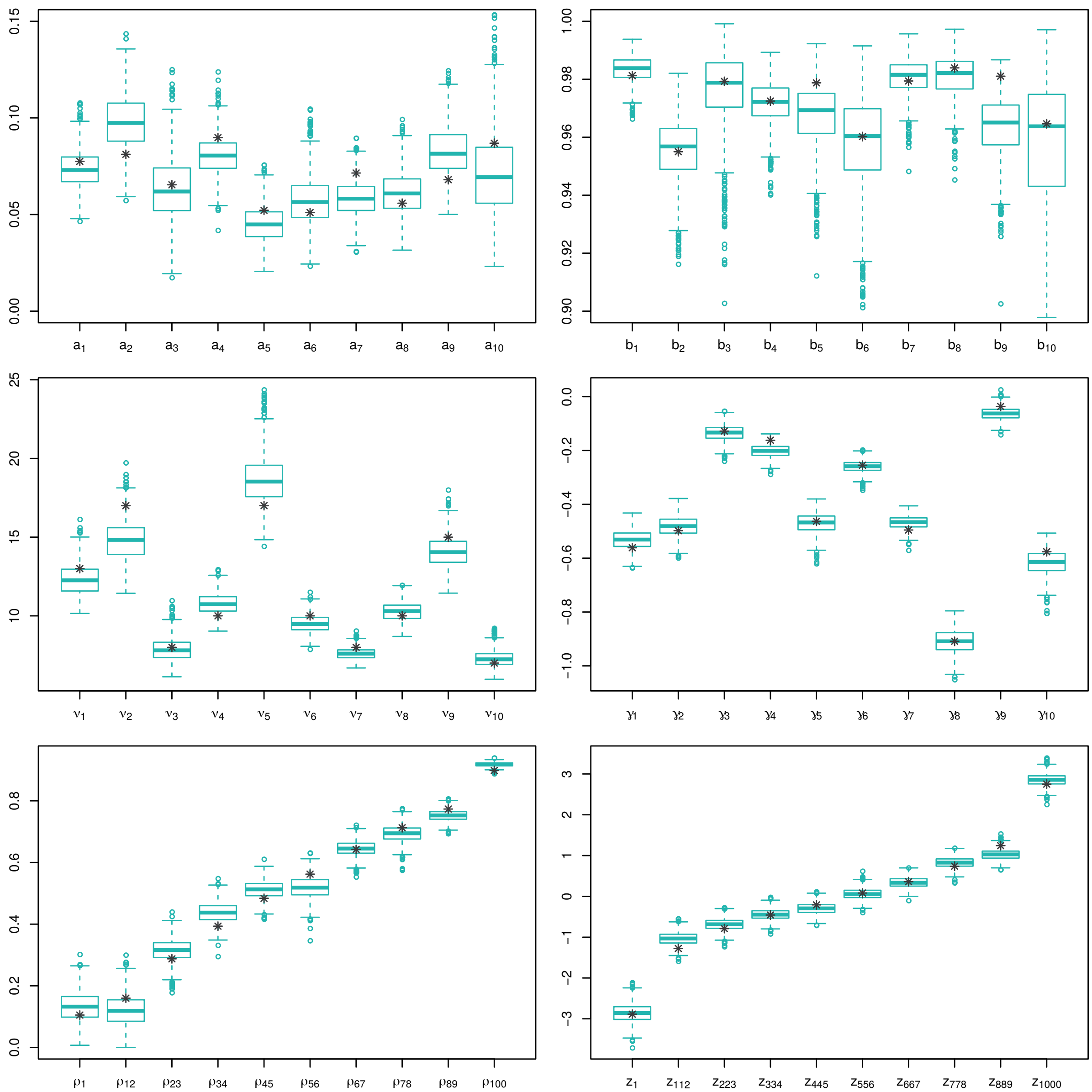

Figure 1: Box plots for the posterior samples of $(a, b, \nu, \gamma, \rho_c, z)$ and true values (stars)
The figure shows the box plot for the posterior samples from simulated data. We select a few parameters ρ_c and z based on their ranks of values to conserve space. We observe that the true values of (a, b, ν, γ) never reach out of their whiskers. The online version of this figure is in color.

Student-t, GSt distribution with zero location vector and scale matrix R_t such that,

$$\text{Multivariate Gaussian: } y_t = R_t^{1/2}\epsilon_t$$

$$\text{Multivariate Student-t : } y_t = \sqrt{\zeta_t}R_t^{1/2}\epsilon_t$$

$$\text{Multivariate GSt: } y_t = \gamma\zeta_t + \sqrt{\zeta_t}R_t^{1/2}\epsilon_t$$

where $\epsilon_t \sim \Phi_d(0, I_d)$, $\zeta_t \sim IG(\nu/2, \nu/2)$, we let $\nu = 10$ and $\gamma = -0.1$ and the time varying scale matrix

$$R_t = \begin{bmatrix} 1 & ... & r_{1dt} \\ ... & 1 & ... \\ r_{d1t} & ... & 1 \end{bmatrix} = \rho_t \rho_t' + D_t D_t' \tag{10}$$

where D_t is a diagonal matrix with elements $1 - \rho_t^2$. We consider six models to account for different behaviors of $\rho_t = (\rho_{1t}, \dots, \rho_{dt})'$ in high dimensions, for $\forall i = 1, \dots, d$:

1. Constant: $\rho_{it} = \sqrt{0.9}$.
2. Sine: $\rho_{it} = \sqrt{0.5 + 0.4cos(2\pi t/200 + \varphi i/20)}$.
3. Fast sine: $\rho_{it} = \sqrt{0.5 + 0.4cos(2\pi t/20 + \varphi i/20)}$.
4. Step: $\rho_{it} = \sqrt{0.4 + 0.5I(t > 500)}$.
5. Ramp: $\rho_{it} = \sqrt{mod((t + \varphi i)/200)}$.
6. Model: $\rho_{it} = \sqrt{\frac{\exp(h_{it})}{1+\exp(h_{it})}}$ where $h_{it} = -0.4(1 - 0.99) + 0.99h_{i,t-1} + 0.14\eta_{it}$ and $\eta_{it} \sim \Phi(0, 1)$.

Here, φ is used to control for the co-movement of the joint dependence of time series. When $\varphi = 0$, we have equivalent-scale models in sine, fast sine, and ramp. Figure 2 shows the r_{ijt} process for selected elements in the scale matrix R_t with $\varphi = 10$.

We generate 100 datasets for each multivariate distribution and estimate the scale matrix R_t using the EWMA, the DCC (Engle (2002)) and the GAS models. For EWMA, we let

$$\Sigma_t = \phi \Sigma_{t-1} + (1 - \phi) y_{t-1} y_{t-1}' \text{ where } \phi = 0.96,$$
$$R_t = diag(\Sigma_t)^{-1/2} \Sigma_t diag(\Sigma_t)^{-1/2}.$$

We generate 1100 time periods and measure the accuracy of each model based on the mean absolute error (MAE) and the mean squared error (MSE) for the last $T = 1000$ observations,

$$MAE = \frac{1}{T} \frac{\sum_{t=1}^{T} \sum_{i=1}^{d} \sum_{j=1}^{d} (|\hat{\rho}_{ijt} - \rho_{ijt}|)}{d^2 - d},$$
$$MSE = \frac{1}{T} \frac{\sum_{t=1}^{T} \sum_{i=1}^{d} \sum_{j=1}^{d} (\hat{\rho}_{ijt} - \rho_{ijt})^2}{d^2 - d}.$$

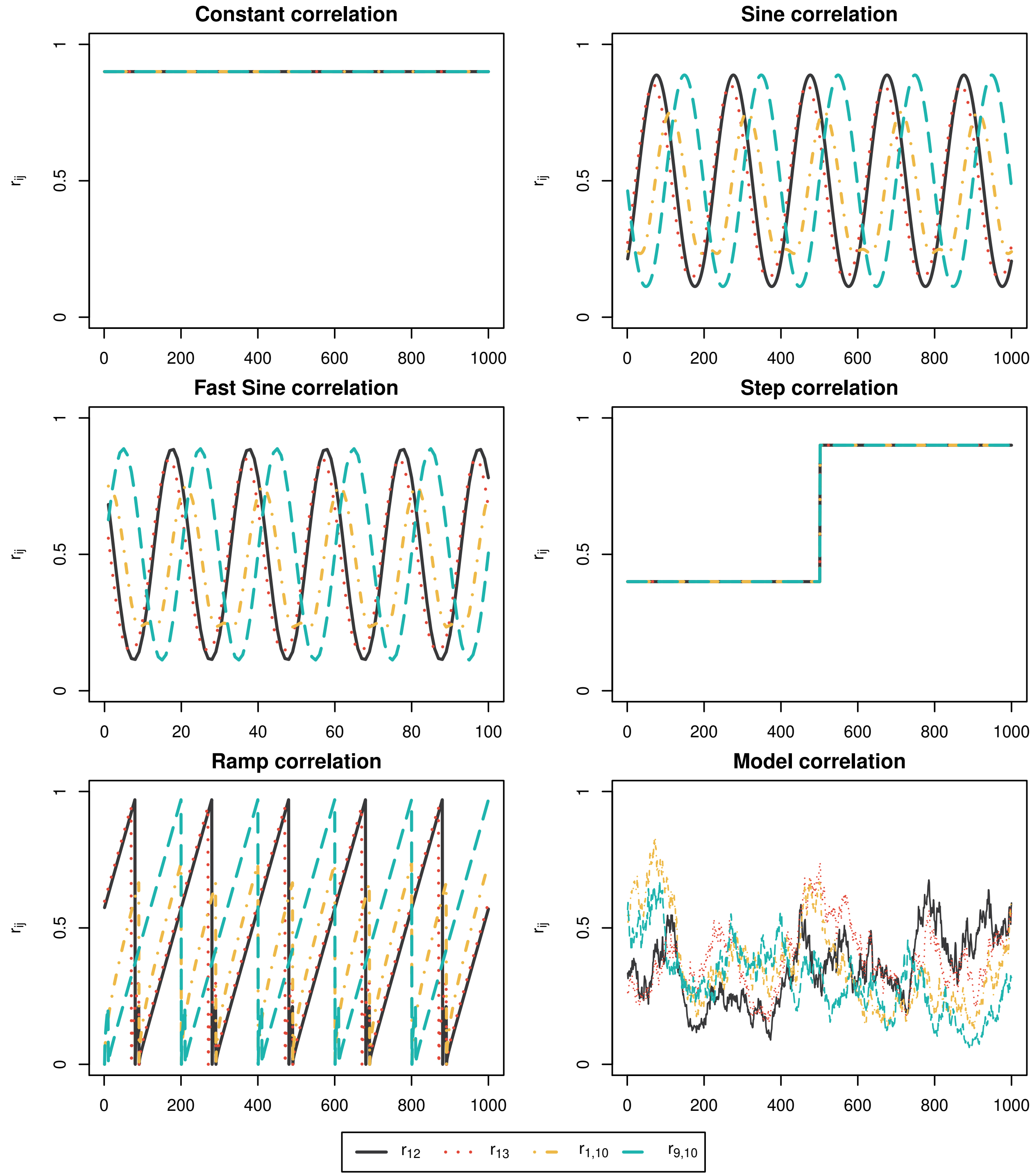

Figure 2: The r_{ij} processes for different stress tests

The figure shows selected r_{ij} processes for different stress scenarios. The R_t is equivalent-scale matrix in case of constant and step correlations. For sine, fast sine and ramp models, we set $\varphi = 10$ to account for lag in the correlation. The online version of this figure is in color.

Table 1 shows the comparison among the GAS and the EWMA to the benchmark DCC model. The relative values of MAE and MSE are measured based on the mean of relative MAE and MSE for each Monte Carlo dataset. For multivariate Gaussian simulation, the GAS model is preferred over the DCC model with an increase of accuracy at least 10% with $\varphi = 0$. The GAS model also outperforms the DCC model in most of the scenarios with multivariate Student-t and MGSt distribution for $\varphi = 10$.

Table 1: MAE and MSE results: in-sample

	Constant	Sine	Fast sine	Step	Ramp	Model
MAE - Multivariate Gaussian						
GAS	0.562	0.639	0.794	0.765	0.738	0.861
DCC	1.000	1.000	1.000	1.000	1.000	1.000
EWMA	5.153	1.048	1.068	1.068	0.921	1.199
MSE - Multivariate Gaussian						
GAS	0.384	0.563	0.721	0.646	0.809	0.829
DCC	1.000	1.000	1.000	1.000	1.000	1.000
EWMA	31.367	1.173	1.210	1.080	1.090	1.484
MAE - Multivariate Student-t						
GAS	0.506	0.792	0.883	0.871	0.805	0.947
DCC	1.000	1.000	1.000	1.000	1.000	1.000
EWMA	4.636	1.160	1.193	1.043	1.059	1.251
MSE - Multivariate Student-t						
GAS	0.321	0.738	0.843	0.793	0.876	0.985
DCC	1.000	1.000	1.000	1.000	1.000	1.000
EWMA	27.077	1.395	1.495	1.109	1.271	1.633
MAE - MGSt						
GAS	0.867	0.878	0.981	0.953	0.892	0.916
DCC	1.000	1.000	1.000	1.000	1.000	1.000
EWMA	5.135	1.107	1.194	1.137	1.045	1.181
MSE - MGSt						
GAS	0.483	0.869	0.936	0.943	0.964	0.905
DCC	1.000	1.000	1.000	1.000	1.000	1.000
EWMA	31.248	1.322	1.495	1.323	1.252	1.371

The table shows the MAE and MSE for the estimated dynamic scale matrix of EWMA, DCC, GAS models. The MAE and MSE are measured as relative values with respect to DCC model. For sine, fast sine, ramp model, we set $\varphi = 0$ for multivariate Gaussian distribution and $\varphi = 10$ for multivariate Student-t and multivariate GSt distribution.

6 Empirical data

In this section, we illustrate our approach with a series of $d = 140$ daily stock returns of companies listed in the S&P 500 index, from 01/01/2007 to 01/09/2014. This period includes the subprime mortgage crisis (2007-2009) and the European sovereign debt crisis (2010-2012). The data are taken from Datastream (2018) and contain $T = 2000$ days observed during the considered eight-year period. Table 2 shows the summary statistics for the daily stock returns. The mean daily return over the 140 stocks is 0.056%. The most extreme individual events are a one day crash of -53.8% and a one day gain of 87.0%. We calculate the robust skewness and robust kurtosis based on the quantile distribution of returns due to outliers, see Kim and White (2004). The average skewness is 0.05, which reflects a slight asymmetry of the observed returns, and the average excess kurtosis is 0.312, which shows the heavy tails of the return distributions.

As described in Section 3, we use a two-stage procedure to estimate the dependence structure of the stock returns. In Subsection 6.1, we fit a $AR(1) - EGARCH(1,1)$ model for the conditional mean and variance of marginal returns. Then, we take out the standardized innovations and transform them into the copula observations using the corresponding cdf. In Subsection 6.2, we estimate the one factor copula models and illustrate some empirical findings.

Table 2: Summary statistics for cross-sectional daily returns of 140 firms listed in S&P 500

Statistics	N	Mean	Minimum	1st Qu.	Median	3rd Qu.	Maximum
Mean	140	0.056	-0.028	0.038	0.053	0.071	0.168
Minimum	140	-16.559	-53.802	-19.929	-14.845	-11.470	-5.372
1. Quartile	140	-0.928	-1.801	-1.094	-0.890	-0.708	-0.417
Median	140	0.009	-0.000	0.000	0.000	0.011	0.066
3. Quartile	140	1.041	0.511	0.837	1.011	1.223	1.852
Maximum	140	19.435	7.795	13.015	17.941	22.854	86.983
Skewness	140	0.050	-0.045	0.035	0.054	0.071	0.131
Excess Kurtosis	140	0.312	0.124	0.256	0.303	0.359	0.661

Summary statistics for cross-sectional daily returns (in percentages) of 140 firms listed in S&P 500 index. The distribution of the common statistics are shown by rows label using quantiles and mean. The outliers can distort the sample statistics, hence we calculate the robust skewness (RS_2) and robust kurtosis (RK_2) based on the quantile distribution of returns, see Kim and White (2004)

6.1 Marginal distributions

For simplicity, we fit an $AR\,(1) - EGARCH\,(1,1)$ - skew Student-t model for each marginal return series using *rugarch* package, see Ghalanos (2018). The standardized innovation, η_{it}, is assumed to follow a univariate skew Student-t distributions (sStd), see Fernández and Steel (1998), with degrees of freedom $\nu_{i\eta}$ and skewness parameter $\xi_{i\eta}$, for $i = 1, \ldots, d$. Then,

$$\begin{aligned}
r_{it} &= c_i + \phi_{i1} r_{i,t-1} + \sigma_{it}\eta_{it} \\
log(\sigma^2_{it}) &= \omega_i + \alpha_{i1}\eta_{i,t-1} + \gamma_{i1}(|\eta_{i,t-1}| - E|\eta_{i,t-1}|) + \beta_{i1} log(\sigma^2_{i,t-1}) \\
\eta_{it} &\sim skew - Student - t(\nu_{i\eta}, \xi_{i\eta}) \\
u_{it} &= F_{sStd}(\eta_{it}|\nu_{i\eta}, \xi_{i\eta})
\end{aligned}$$

Table 3 shows the summary statistics for the posterior mean estimations of the univariate $AR\,(1) - EGARCH\,(1,1)$ - skew Student-t model across the 140 firms. The effect of the conditional autoregressive mean is weak as the average value of ϕ_{i1} is -0.035 and ϕ_{i1} is insignificantly different from 0 in most of marginals. This fact matches with other findings in the finance literature that the return levels are unpredictable. However, the variance returns are quite predictable through the $EGARCH\,(1,1)$ setting. The average of α_{i1}, β_{i1}, and γ_{i1} are significantly different from 0 for most of the marginals. On average, the volatility clustering is captured by the parameter β_{i1} standing at 0.989. The degrees of freedom for each marginal also diversifies between 2.66 and 9.86 which accounts for different kurtosis. The skewness parameter ranges from 0.871 to 1.109 and most of them are not significantly different from the one which represents for symmetric Student-t distribution. However, the leverage effect that negative return usually leads to an increase in the volatility of innovation is significantly observed for all marginal returns.

For the second stage, the standardized innovations are taken out and transformed to copula observations by applying the corresponding marginal cdfs. More specifically, using the maximum likelihood estimations for each marginal, we obtain the standardized innovations, η_{it}, of the $AR\,(1) - EGARCH\,(1,1)$ process and transform them into the copula observations $u_{it} = F_{sStd}\left(\eta_{it}|\hat{\vartheta}_i\right)$ where $\hat{\vartheta}_i = \{\hat{c}_i, \hat{\phi}_{i1}, \hat{\omega}_i, \hat{\alpha}_{i1}, \hat{\beta}_{i1}, \hat{\gamma}_{i1}, \hat{\nu}_{i\eta}, \hat{\xi}_{i\eta}\}$. This simplifies the computational burden in the high dimensional setting in which we only concentrate on estimating the copula parameters. Apart from that, we check if the choice of the sStd distribution is suitable with univariate GARCH volatilities

Table 3: Summary statistics of $AR(1) - EGARCH(1,1)$-skew Student-t for marginal returns

	Mean	Minimum	1. Quartile	Median	3. Quartile	Maximum
c_i	0.054	-0.028	0.033	0.050	0.067	0.170
ϕ_{i1}	-0.035	-0.127	-0.056	-0.037	-0.012	0.053
ω_i	0.011	-0.011	0.007	0.009	0.013	0.062
α_{i1}	-0.066	-0.131	-0.081	-0.067	-0.052	-0.022
β_{i1}	0.989	0.959	0.986	0.991	0.994	0.999
γ_{i1}	0.137	0.036	0.113	0.135	0.163	0.246
$\xi_{i\eta}$	0.998	0.871	0.972	0.995	1.022	1.109
$\nu_{i\eta}$	5.407	2.660	4.479	5.159	5.926	9.858

Summary statistics for the maximum likelihood estimations of the univariate $AR(1) - EGARCH(1,1)$-skew Student model across the 140 firms. The distribution of estimations are described by mean and quantiles.

by performing the Kolmogorov-Smirnov goodness of fit test as well as the Anderson-Darling test, the Neyman's smooth test of fit in Online Appendix. All series passed the test with p-values larger than 0.05. We also tested for the serial correlation of the innovations and did not find significant results.

6.2 Copula estimation

Next, we apply the proposed Bayesian approach to nine different one factor copula models. These are the dynamic Gaussian, Student-t and MGSt combined with particular cases of models referred as block equivalent mean correlation, single group and multiple-group. In the block equi-mean correlation model, the parameter f_{igc} in (7) is restricted to be the same for the assets belonging to the same group, i.e., $f_{1gc} = \ldots = f_{n_g gc}$, for all $g = 1, \ldots, G$. We classified $G = 12$ group industries of assets depending on their SIC codes, as in Creal and Tsay (2015), Oh and Patton (2017b), among others. These are Oil & Construction , Food & Beverage, Pharmaceuticals, Plastic Material & Plant Chemical, Textile & Papers, Steel, Home Appliances & Automobile, Electronics, Transportation & Communication, Retail & Distribution, Insurance, Finance (not contained in Insurance). The detailed number of firms are reported in Online Appendix. On average, there are 12 firms in each sector group. In the single group model, $G = 1$ as described in (4), only a few numbers of parameters account for the tail dependence, while the correlations are allowed to be different across the assets. Finally, the multiple-group dynamic model is the most flexible one with different behaviors in the tail dependence and unrestricted scale parameters as in (7). In all cases, we generate 40,000 iterations with 10,000 burn-in for each factor copula model and every 10-th draws are taken in order to prevent the autocorrelation in the MCMC chains. We also check for the convergence of the MCMC chains in Online Appendix.

Table 4 outlines the main estimation results for the nine one factor copula models. In particular, the table includes the value of the AIC, BIC, and DIC for the model selection, obtained as explained in Appendix E. The dynamic MGSt copula appears to show a better fit over the Gaussian and Student-t copula models. In general, the posterior means of the parameters a, b and f_c are similar across Gaussian, Student-t, and MGSt copulas, as shown for example in the model (3), (6) and (9). The block equi-mean correlation model reports a smaller posterior range for a and b. The degrees of freedom and skewness parameters are roughly similar between the block-equivalent and the multiple-group models. The model selection criteria show an interesting result that the models with more parameters accounting for extreme events are preferable over the models that have limitations on these behaviors. For example, in Gaussian copulas, the one group outperforms the block-equivalent due to the fact that they do not capture the extreme occurrences. However, it is preferable to use block equi-mean correlation in the Student-t and MGSt copulas rather than one group copula. The block-equivalent models could even be comparable with the multi-group models in all criteria AIC, BIC, and DIC. We also obtain that the group Student-t copula yields lower degrees of freedom than the single group Student-t copula. This finding confirms with Creal and Tsay (2015) due to the fact that when the number of assets in a group increases, the uncertainty reduces because the central limit theorem holds.

Table 4: Estimation results for alternative copula models

	Gaussian block equi (1)	Gaussian 1G (2)	Gaussian multi.group (3)	Student-t block equi (4)	Student-t 1G (5)	Student-t multi.group (6)	MGSt block equi (7)	MGSt 1G (8)	MGSt multi.group (9)
AIC	-163400	-164422	-164658	-192501	-189064	-192197	-197143	-189184	-196901
BIC	-163198	-163627	-163739	-192232	-188262	-191211	-196807	-188377	-195848
DIC	-166377	-167382	-167622	-211507	-208358	-212414	-212257	-208539	-213090
# params	36	142	164	48	143	164	60	144	188
a	[0.032,0.090]	0.070	[0.056,0.146]	[0.023,0.067]	0.026	[0.025,0.067]	[0.025,0.064]	0.026	[0.026,0.069]
b	[0.968,0.997]	0.971	[0.855,0.984]	[0.984,0.999]	0.992	[0.945,0.995]	[0.981,0.999]	0.992	[0.956,0.995]
ν				[6.814,11.907]	11.421	[6.816,11.870]	[7.778,23.391]	11.378	[7.885,23.086]
γ							[-1.230,-0.179]	-0.106	[-1.215,-0.184]
f_c	[1.215,1.858]	[0.957,2.360]	[0.965,2.369]	[1.278,1.997]	[1.286,2.727]	[1.030,2.442]	[1.250,1.986]	[1.288,2.716]	[1.015,2.432]

Posterior estimations for nine one factor copula models and model selection criteria. Three different models are considered (the block-equivalent mean, one group and multiple-group) for three different copula models (Gaussian, Student-t and MGSt). The table reports only the range of the posterior means for the group models and the point estimates for the one group model.

In Tables 5 and 6, we report respectively details of the estimation for the dynamic group

Student-t and dynamic group MGSt copulas. The posterior means of a and b show different dynamic behaviors in each group sector. The values of f_c are depicted as interval range of the posterior means of f_{igc}, for assets $i = 1, \dots, n_g$ belonging to each group $g = 1, \dots, G$. The values in parentheses are the average values of the posterior standard deviations. The posterior means of ν are quite different among groups as well as between models. The standard deviations of ν are small in the case of the group Student-t and seem to be higher in the MGSt model. The lowest degree of freedom parameter in the dynamic group Student-t is in Oil industries standing at 6.8. Despite that, it is strongly negative skewed in the MGSt copula model. The other groups that have low degrees of freedom such as Paper, Insurance, and Finance reveal a slight skewness. Although the posterior variance of the degrees of freedom in some industries are higher in the case of group MGSt copula, it still supports for the hypothesis that the lower tail is heavier and the distribution is highly asymmetric rather than there is symmetry in both upper tail and lower tail. We show different tail dependence in each group sector by calculating the average of penultimate tail dependence (Manner and Segers (2011)) of bivariate copulas of the assets belonging to the same sector at quantile 0.5%. The strongest lower tail dependence is 0.280 from the Finance sector with also strong upper tail dependence.

Figure 3 describes the posterior mean of the dynamic conditional Kendall-τ correlation among group sectors using the Student and MGSt copula. This posterior mean Kendall-τ between sector g_1 and g_2 is calculated based on the average of the scale parameters over iterations and group members as $\frac{1}{n_{g_1} n_{g_2}} \sum_{ij} \rho_{it}\rho_{jt}$ where i, j belong to group g_1 and g_2 respectively, and $i \neq j, g_1 \neq g_2$. And then, we calculate the Kendall-τ correlation equivalent of the factor copula models using Monte Carlo simulation. As we can see, a common pattern is that the correlation increased over time during the subprime mortgage crisis (2008-2009) and the European sovereign debt crisis (peak in 2012). Finance and Insurance sectors suffered most as the correlations go up during the crisis while in other sectors such as Food and Retail, the correlation is less volatile.

Figure 4 shows the posterior distribution of the conditional variance and conditional Kendall-t correlation of several companies including Citigroup, Goldman Sachs GP., McDonalds, Johnson & Johnson, Apple, and Intel using MGSt copula. The first two columns illustrate the conditional variance and the last column depicts the conditional Kendall-t correlation between the couple. As mentioned above, the $2007-2009$ and $2010-2012$ period experienced a high volatility and a rise in

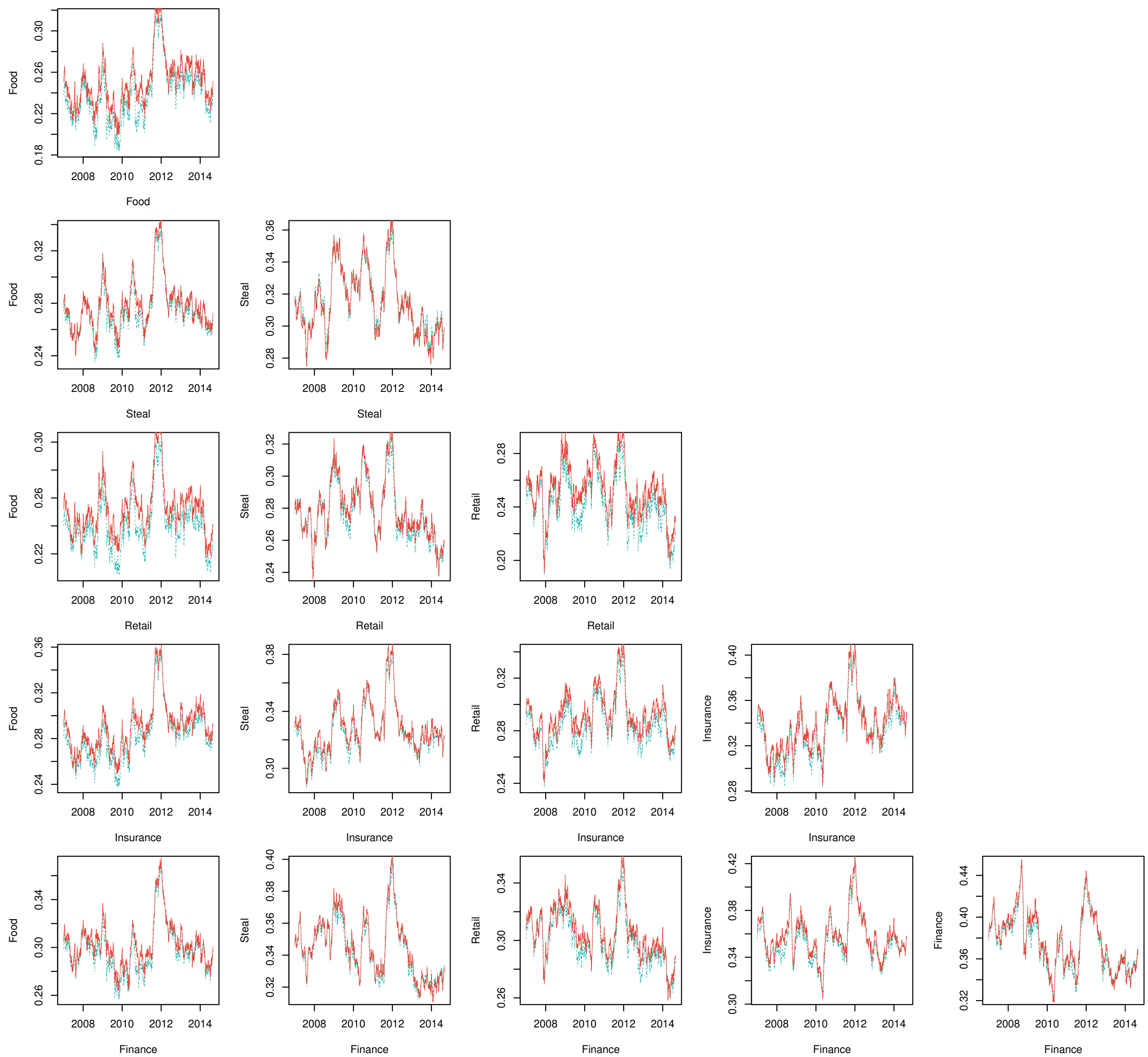

Figure 3: The Kendall-τ correlation among group sectors

The figure shows the Kendall-τ correlation among sectors using group Student and MGSt factor copulas. The Kendall-τ correlation for Student model is the blue dash line and Kendall-τ for MGSt model is the red solid line. We select some sectors for illustrating and conserve space. The Kendall-τ correlation among sectors increased during crisis, reached a peak in 2009 and 2012. The online version of this figure is in color.

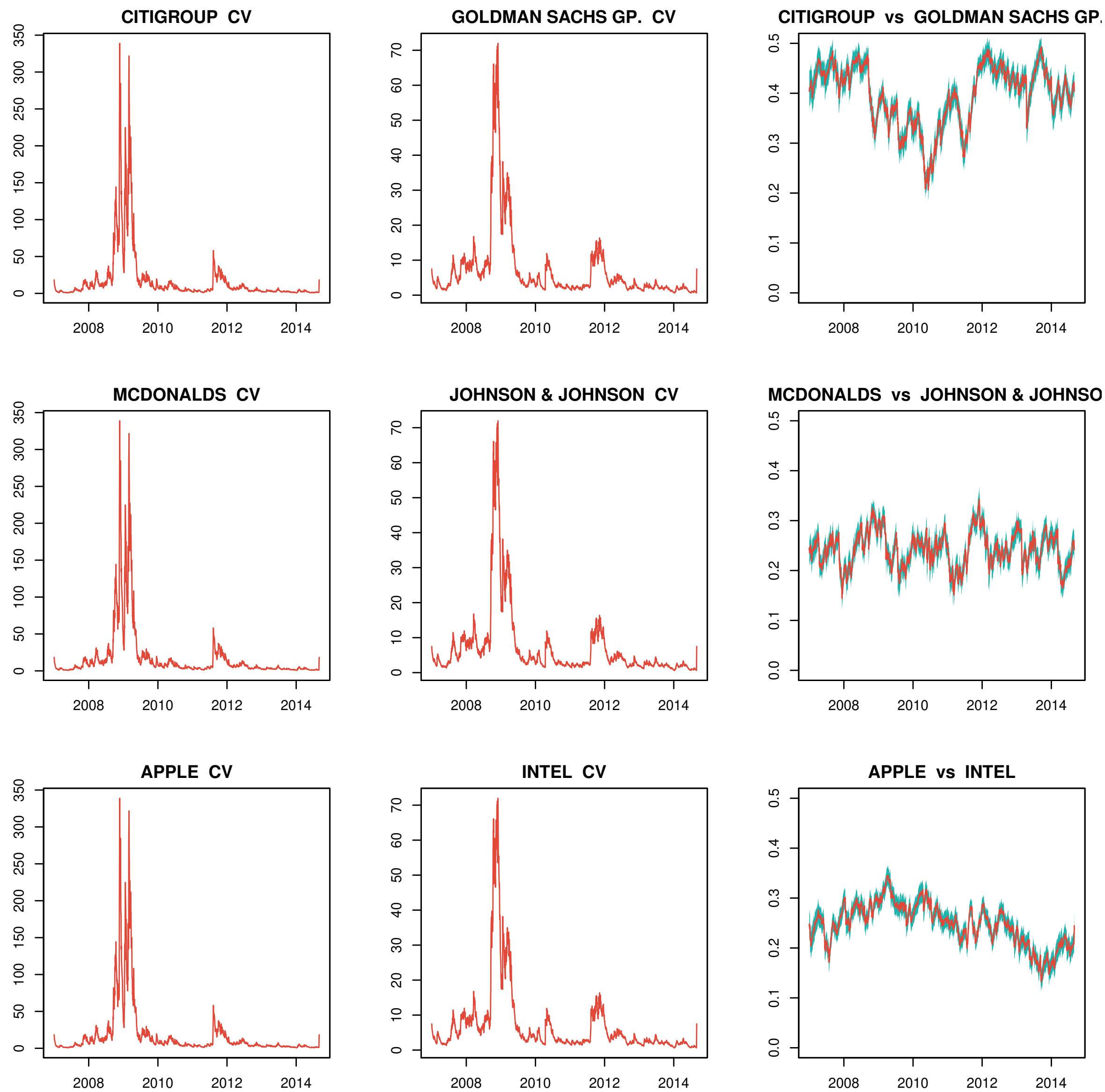

Figure 4: Posterior Kendall-τ correlation among time series

The first two columns describe the conditional variance and the last column depicts the dynamic Kendall-τ correlation together with the 95% credible interval using MGSt copula. First row: Citigroup, Goldman Sachs GP., second row: McDonalds, Johnson & Johnson, third row: Apple, Intel. The online version of this figure is in color.

Table 5: Results for the group Student-t copula with time-varying factor loadings.

	Oil	Food & Bev.	Pharma.	Plastics	Paper	Steel
a	0.067	0.043	0.059	0.052	0.067	0.034
	(0.005)	(0.007)	(0.011)	(0.014)	(0.006)	(0.004)
b	0.985	0.987	0.977	0.945	0.976	0.986
	(0.002)	(0.005)	(0.009)	(0.037)	(0.004)	(0.003)
ν	6.816	9.827	9.690	8.954	9.292	11.870
	(0.215)	(0.606)	(0.644)	(0.471)	(0.377)	(0.421)
f_c	[1.46,1.75]	[1.17,2.07]	[1.16,1.50]	[1.20,2.12]	[1.19,1.97]	[1.23,2.22]
	(0.098)	(0.085)	(0.076)	(0.056)	(0.076)	(0.070)
# firms	13	7	7	7	12	17
$\lambda_L = \lambda_U$	0.177	0.115	0.098	0.160	0.148	0.151
	Home App.	Electronics	Transportation	Retail	Insurance	Finance
a	0.038	0.025	0.049	0.045	0.041	0.041
	(0.006)	(0.004)	(0.010)	(0.006)	(0.004)	(0.004)
b	0.985	0.995	0.975	0.985	0.988	0.992
	(0.005)	(0.002)	(0.013)	(0.004)	(0.002)	(0.002)
ν	9.457	11.001	9.706	10.659	8.001	7.032
	(0.331)	(0.657)	(0.533)	(0.530)	(0.224)	(0.186)
f_c	[1.33,2.32]	[1.03,1.77]	[1.37,2.03]	[1.20,1.72]	[1.03,2.44]	[1.32,2.44]
	(0.073)	(0.105)	(0.068)	(0.082)	(0.086)	(0.103)
# firms	15	8	7	11	18	18
$\lambda_L = \lambda_U$	0.179	0.122	0.181	0.111	0.194	0.238

Posterior estimations for the interest parameters of the group Student-t factor copula. This includes the posterior means and standard deviations for (a, b, ν) and the values of f_c are depicted as interval range of the posterior means together with the average posterior standard deviations. The tail dependence are calculated as the average of the penultimate tail dependence (Manner and Segers (2011)) of bivariate Student-t copula with the mean correlation of the assets belonging to the same sector at quantile 0.5% (see Appendix C).

correlation among all examples due to the financial crisis. The cross Kendall-τ correlation between financial series are more volatile than other sectors. They were even highly dependent before crisis happened in 2007.

6.3 Risk measures and portfolio allocation

Table 7 shows the average of VaR, CVaR and standard deviation of the predicted returns for the equally weighted portfolio. We choose $H = 400$ ahead trading days. All the models perform quite well in terms of risk measure except the one group Student and one group MGSt copulas. In all other cases, the numbers of days that the realized return of the portfolio exceeds the threshold VaR are close to their expected numbers. The block equi-mean correlation and the group MGSt factor copulas also captures correctly quantile dependence at 1% level. The value of CVaR not only depends on the VaR threshold but also depends on the copula types. For copulas that have no or only one parameter to control for the tail dependence, the CVaR is often higher than those with flexible tail dependence.

Table 6: Results for the group MGSt copula with time-varying factor loadings.

	Oil	Food & Bev.	Pharma.	Plastics	Paper	Steel
a	0.061	0.045	0.066	0.051	0.069	0.035
	(0.006)	(0.008)	(0.014)	(0.013)	(0.006)	(0.004)
b	0.984	0.984	0.970	0.956	0.977	0.986
	(0.003)	(0.006)	(0.013)	(0.023)	(0.004)	(0.003)
ν	23.086	13.967	17.039	10.022	9.875	12.354
	(1.926)	(1.671)	(3.107)	(0.648)	(0.436)	(0.460)
γ	-1.215	-0.450	-0.682	-0.251	-0.236	-0.264
	(0.078)	(0.072)	(0.124)	(0.032)	(0.023)	(0.018)
f_c	[1.38,1.75]	[1.16,2.01]	[1.13,1.48]	[1.18,2.09]	[1.16,1.94]	[1.20,2.18]
	(0.090)	(0.082)	(0.071)	(0.058)	(0.078)	(0.073)
# firms	13	7	7	7	12	17
λ_L	0.222	0.159	0.149	0.204	0.186	0.191
λ_U	0.071	0.065	0.044	0.113	0.103	0.117
	Home App.	Electronics	Transportation	Retail	Insurance	Finance
a	0.034	0.026	0.052	0.047	0.047	0.042
	(0.005)	(0.004)	(0.009)	(0.007)	(0.005)	(0.004)
b	0.989	0.995	0.976	0.985	0.985	0.992
	(0.003)	(0.002)	(0.009)	(0.005)	(0.003)	(0.002)
ν	10.230	12.028	9.997	13.994	8.750	7.885
	(0.399)	(0.819)	(0.629)	(1.051)	(0.296)	(0.297)
γ	-0.257	-0.270	-0.246	-0.418	-0.212	-0.184
	(0.020)	(0.034)	(0.029)	(0.044)	(0.017)	(0.022)
f_c	[1.32,2.29]	[1.01,1.75]	[1.34,1.98]	[1.21,1.71]	[1.02,2.43]	[1.31,2.42]
	(0.082)	(0.110)	(0.070)	(0.082)	(0.081)	(0.104)
# firms	15	8	7	11	18	18
λ_L	0.225	0.157	0.230	0.156	0.239	0.280
λ_U	0.130	0.086	0.139	0.068	0.147	0.184

Posterior estimations for the interest parameters of group MGSt factor copula. This includes the posterior means and standard deviations for (a, b, ν, γ) and the values of f_c are depicted as interval range of the posterior means together with the average posterior standard deviations. The tail dependences are calculated as the average of the penultimate tail dependence (Manner and Segers (2011)) of bivariate MGSt copula with the mean correlation of the assets belonging to the same sector at quantile 0.5%. (see Appendix C).

Figure 5 shows the smoothed weight for the global optimal portfolio. The figure on the left is obtained by finding the minimum variance portfolio for all assets. We only show the top five assets that have an average weight larger than 5%. The figure on the right illustrates the weight of the minimum CVaR portfolio at 5%. The collections of heavy weight assets chosen in both portfolios are quite similar. The portfolio contains assets that are robust to high volatility such as those in the group Retail, Paper, Pharmaceuticals. We see the similar pattern in both optimal portfolios with holding less "Kellogg" and increasing "Johnson & Johnson", despite that the weights are different. On average, about 60 assets are included in the minimum CVaR in each period while the minimum variance portfolio contains about 15 assets.

Table 7: Risk measure for alternative copula models

	Gaussian block equi (1)	Gaussian 1G (2)	Gaussian multi.group (3)	Student-t block equi (4)	Student-t 1G (5)	Student-t multi.group (6)	MGSt block equi (7)	MGSt 1G (8)	MGSt multi.group (9)
VaR_5%	1.664(22)	1.678(18)	1.665(22)	1.617(24)	1.857(16)	1.658(22)	1.59(23)	1.875(19)	1.618(22)
VaR_1%	2.556(4)	2.574(4)	2.556(4)	2.382(5)	3.014(2)	2.456(5)	2.336(5)	3.059(1)	2.379(5)
CVaR_5%	2.237	2.252	2.239	2.108	2.614	2.165	2.059	2.653	2.093
CVaR_1%	3.14	3.152	3.142	2.83	3.879	2.92	2.758	3.96	2.802
Std	1.066	1.074	1.067	1.028	1.212	1.049	1.03	1.21	1.045

Risk measure for nine one factor copula models. The number in the bracket is the number of days that realized return of the portfolio exceeds the threshold VaR. For $H = 400$ trading days, the expected number of violations are 20 days, 4 days at 5%, 1% level correspondingly.

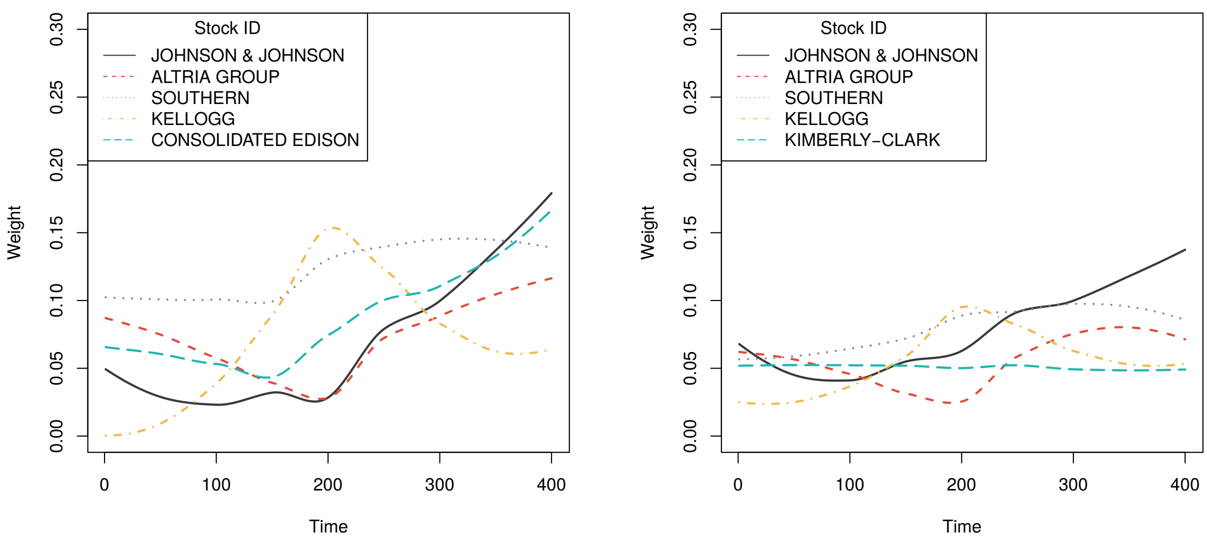

Figure 5: Portfolio allocation among time series based on min-variance and min-CVaR

The figure on the left shows the smooth weight for “Johnson & Johnson”, “Altria Group”, “Southern”, “Kellogg”, and “Consolidated Edison” in the global minimum variance portfolio. The figure on the right shows the smooth weights on global minimum CVaR portfolio. In the optimal variance and CVaR, there are similar assets that follow a similar trend, the weights are quite different.

7 Conclusion

In this paper, we have proposed a family of one factor copula models and developed a Bayesian algorithm to make parallel inference on the model parameters. In our proposed models, the time series become independent conditioning on the latent factor which allows us to introduce an estimation strategy in a parallel setting. Furthermore, the factor loadings have been modeled as GAS processes which imposes a dynamic dependence structure in their densities. Using multiple-group MGSt copulas, we obtain different types of tail and asymmetric dependence. The models are extendible since the number of parameters scales linearly with the dimension. As an extension, more complex copula functions can be build based on the distribution of ζ_g. However, this also may require the computational cost to obtain the inverse cdf. Also, we might consider factor models using the family of Archimedean copulas that only allow for lower tail dependence, due to the empirical finding that half of the groups only show weak evidence of upper tail dependence. Finally, one factor models may not be enough for the high dimensional dependence as Oh and Patton (2017a) and Nguyen et al. (2018) suggest. One future direction could be to extend the proposed approach to dynamic multi factor models.

Acknowledgment

We thank Michael Wiper, Andrew J. Patton, the Associate Editor and two anonymous referees for their helpful comments. The first author acknowledge financial support from the Spanish Ministry of Economy and Competitiveness, project numbers ECO2015-66593-P. E-mail address: hoang.nguyen@uc3m.es

Appendix

A Score update for the factor copula model

A.1 Dynamic Gaussian one factor copula

The conditional cdf of $u_t = (u_{1t}, \dots, u_{dt})'$, where $u_{it} = \Phi(x_{it})$, is given by:

$$F(u_{1t}, \dots, u_{dt} \mid z_t, f_t, \mathcal{F}_t, \theta) = \Pr(U_{1t} \le u_{1t}, \dots, U_{dt} \le u_{dt} \mid z_t, f_t, \mathcal{F}_t, \theta)$$

$$= \Pr\left(X_{1t} \le \Phi^{-1}(u_{1t}), \ldots, X_{dt} \le \Phi^{-1}(u_{dt}) \mid z_t, f_t, \mathcal{F}_t, \theta\right)$$
$$= \prod_{i=1}^{d} \Pr\left(X_{it} \le \Phi^{-1}(u_{it}) \mid z_t, f_t, \mathcal{F}_t, \theta\right).$$

Now, note that, given $\{z_t, f_t, \mathcal{F}_t, \theta\}$, the correlation ρ_{it} is known and X_{it} follows a Gaussian distribution with mean $\rho_{it} z_t$ and standard deviation $\sqrt{1-\rho_{it}^2}$. Then, the conditional density of u_t is,

$$p(u_t \mid z_t, f_t, \mathcal{F}_t, \theta) = \frac{\partial^d F(u_{1t}, \ldots, u_{dt} \mid z_t, f_t, \mathcal{F}_t, \theta)}{\partial u_{1t} \ldots \partial u_{dt}} = \prod_{i=1}^{d} \frac{\phi\left(\Phi^{-1}(u_{it}) \mid \rho_{it} z_t, \sqrt{1-\rho_{it}^2}\right)}{\phi\left(\Phi^{-1}(u_{it}) \mid 0, 1\right)},$$

where $\phi(\cdot \mid \mu, \sigma)$ denotes a normal pdf with mean, μ, and standard deviation, σ. Then, the proposed dynamic process is based on the derivative of the log conditional density wrt the dynamic f_{it}, i.e.:

$$\begin{aligned}
s_{it} &= \frac{\partial \log p(u_t \mid z_t, f_t, \mathcal{F}_t, \theta)}{\partial f_{it}} = \frac{\partial \log p(u_t \mid z_t, f_t, \mathcal{F}_t, \theta)}{\partial \rho_{it}} \frac{\partial \rho_{it}}{\partial f_{it}} \\
&= \frac{\partial \sum_{i=1}^{d} \left(\log \phi\left(\Phi^{-1}(u_{it}) \mid \rho_{it} z_t, \sqrt{1-\rho_{it}^2}\right) - \log \phi\left(\Phi^{-1}(u_{it}) \mid 0, 1\right)\right)}{\partial \rho_{it}} \frac{1-\rho_{it}^2}{2} \\
&= \frac{\partial \left(-\frac{1}{2}\log(2\pi) - \frac{1}{2}\log(1-\rho_{it}^2) - \frac{1}{2}\frac{(\Phi^{-1}(u_{it}) - \rho_{it} z_t)^2}{1-\rho_{it}^2}\right)}{\partial \rho_{it}} \frac{1-\rho_{it}^2}{2} \\
&= \left(\frac{\rho_{it}}{(1-\rho_{it}^2)} + \frac{z_t(\Phi^{-1}(u_{it}) - \rho_{it} z_t)}{1-\rho_{it}^2} - \frac{\rho_{it}(\Phi^{-1}(u_{it}) - \rho_{it} z_t)^2}{(1-\rho_{it}^2)^2}\right) \frac{1-\rho_{it}^2}{2} \\
&= \frac{1}{2}\Phi^{-1}(u_{it}) z_t + \frac{1}{2}\rho_{it} - \rho_{it} \frac{\Phi^{-1}(u_{it})^2 + z_t^2 - 2\rho_{it}\Phi^{-1}(u_{it}) z_t}{2(1-\rho_{it}^2)},
\end{aligned}$$

which leads to the expression given in (3).

A.2 Dynamic generalized hyperbolic skew Student-t one factor copula

The conditional cdf of $u_t = (u_{1t}, \ldots, u_{dt})$, where $u_{it} = F_{GSt}(x_{it} \mid \nu, \gamma)$, is:

$$\begin{aligned}
F(u_{1t}, \ldots, u_{dt} \mid z_t, \zeta_t, f_t, \mathcal{F}_t, \theta) &= \Pr\left(X_{1t} \le F_{GSt}^{-1}(u_{1t} \mid \nu, \gamma), \ldots, X_{dt} \le F_{GSt}^{-1}(u_{dt} \mid \nu, \gamma) \mid z_t, \zeta_t, f_t, \mathcal{F}_t, \theta\right) \\
&= \prod_{i=1}^{d} \Pr\left(\tilde{X}_{it} \le \frac{F_{GSt}^{-1}(u_{it} \mid \nu, \gamma) - \gamma \zeta_t}{\sqrt{\zeta_t}} \mid z_t, \zeta_t, f_t, \mathcal{F}_t, \theta\right),
\end{aligned}$$

where $\tilde{X}_{it} = (X_{it} - \gamma\zeta_t)/\sqrt{\zeta_t}$. Similarly, given $\{z_t, \zeta_t, f_t, \mathcal{F}_t, \theta\}$, the correlation ρ_{it} is known and $\tilde{X}_{it}$ follows a Gaussian distribution with mean $\rho_{it} z_t$ and standard deviation $\sqrt{1-\rho_{it}^2}$. Then, the conditional density of u_t is,

$$p\left(u_t \mid z_t, \zeta_t, f_t, \mathcal{F}_t, \theta\right) = \frac{\partial^d F\left(u_{1t}, \ldots, u_{dt} \mid z_t, \zeta_t, f_t, \mathcal{F}_t, \theta\right)}{\partial u_{1t} \ldots \partial u_{dt}} = \prod_{i=1}^{d} \frac{\phi\left(\frac{F_{GSt}^{-1}(u_{it}|\nu)-\gamma\zeta_t}{\sqrt{\zeta_t}} \mid \rho_{it} z_t, \sqrt{1-\rho_{it}^2}\right)}{f_{GSt}\left(F_{GSt}^{-1}\left(u_{it} \mid \nu, \gamma\right) \mid \nu, \gamma\right)\sqrt{\zeta_t}}, \tag{11}$$

where $f_{GSt}\left(\cdot \mid \nu, \gamma\right)$ denotes the standard generalized hyperbolic skew Student-t with ν degrees of freedom and γ skewness parameter. Thus, the equation for s_{it} remains,

$$\begin{aligned}
s_{it} &= \frac{\partial \log p\left(u_t \mid z_t, \zeta_t, f_t, \mathcal{F}_t, \theta\right)}{\partial f_{it}} = \frac{\partial \log \phi\left(\frac{F_{GSt}^{-1}(u_{it}|\nu)-\gamma\zeta_t}{\sqrt{\zeta_t}} \mid \rho_{it} z_t, \sqrt{1-\rho_{it}^2}\right)}{\partial \rho_{it}} \frac{1-\rho_{it}^2}{2} \\
&= \frac{1}{2} \frac{F_{GSt}^{-1}\left(u_{it} \mid \nu\right) - \gamma\zeta_t}{\sqrt{\zeta_t}} z_t + \frac{1}{2}\rho_{it} - \rho_{it} \frac{\left(\frac{F_{GSt}^{-1}(u_{it}|\nu)-\gamma\zeta_t}{\sqrt{\zeta_t}}\right)^2 + z_t^2 - 2\rho_{it} \frac{F_{GSt}^{-1}(u_{it}|\nu)-\gamma\zeta_t}{\sqrt{\zeta_t}} z_t}{2(1-\rho_{it}^2)},
\end{aligned}$$

which leads to the expression given in (5).

B Equivalence of predictive density

Here, we show that our GAS update equation is similar to Lucas et al. (2018) where the value of score s_{it} is the likelihood conditional on the unobservable mixing variable. For that,

$$\begin{aligned}
s_{it}^{OP} &= \frac{\partial}{\partial f_t} \log p\left(u_t | f_t, \mathcal{F}_t, \theta\right) = \frac{\partial}{\partial f_t} \mathtt{log} \int p(u_t, z_t | f_t, \mathcal{F}_t, \theta) dz_t \\
&= \left[\int p(u_t, z_t | f_t, \mathcal{F}_t, \theta) dz\right]^{-1} \int \frac{\partial}{\partial f_t} p(u_t, z_t | f_t, \mathcal{F}_t, \theta) dz_t \\
&= \left[\int p(u_t, z_t | f_t, \mathcal{F}_t, \theta) dz\right]^{-1} \int \frac{\partial}{\partial f_t} p(u_t | z_t, f_t, \mathcal{F}_t, \theta) p(z_t) dz_t \\
&= \left[\int p(u_t, z_t | f_t, \mathcal{F}_t, \theta) dz\right]^{-1} \int \frac{1}{p(u_t | z_t, f_t, \mathcal{F}_t, \theta)} \frac{\partial p(u_t | z_t, f_t, \mathcal{F}_t, \theta)}{\partial f_t} p(u_t | z_t, f_t, \mathcal{F}_t, \theta) p(z_t) dz_t \\
&= \int \frac{\partial \mathtt{log} p(u_t | z_t, f_t, \mathcal{F}_t, \theta)}{\partial f_t} \frac{p(u_t | z_t, f_t, \mathcal{F}_t, \theta) p(z_t)}{\int p(u_t, z_t | f_t, \mathcal{F}_t, \theta) dz} dz_t \\
&= \int \frac{\partial \mathtt{log} p(u_t | z_t, f_t, \mathcal{F}_t, \theta)}{\partial f_t} p(z_t | u_t, f_t, \mathcal{F}_t, \theta) dz_t \\
&= \mathbf{E}_{z_t}\left[\left.\frac{\partial \mathtt{log} p(u_t | z_t, f_t, \mathcal{F}_t, \theta)}{\partial f_t}\right| u_t, f_t, \mathcal{F}_t, \theta\right].
\end{aligned}$$

Here, the value of standard score s_{it}^{OP} is the expectation of our proposal score s_{it} over z_t, where z_t has the pdf $p(z_t|u_t, f_t, \mathcal{F}_t, \theta)$ distribution.

C Tail dependence for the generalized hyperbolic skew Student-t copula

Consider the bivariate GSt copula. We derive the penultimate tail dependence of a pair of pseudo observables x_{igt} and x_{jgt} in a same group, g, at time t as:

$$\begin{aligned} C_{GSt}(u, u|R_t, \nu_g, \gamma_g) &= F_{MGSt}(F_{GSt}^{-1}(u), F_{GSt}^{-1}(u)|R_t, \nu_g, \gamma_g) \\ &= \int_{-\infty}^{F_{GSt}^{-1}(u)} \int_{-\infty}^{F_{GSt}^{-1}(u)} f_{MGSt}(x_1, x_2|R_t, \nu_g, \gamma_g) dx_1 dx_2 \end{aligned} \tag{12}$$

where

$$\begin{aligned} f_{MGSt}(x_1, x_2|R_t, \nu_g, \gamma_g) &= c \frac{K_{\frac{\nu+2}{2}}\left(\sqrt{(\nu + Q(x))Q(\gamma)}\right) \exp(x' R_t^{-1} \gamma)}{\sqrt{(\nu + Q(x))Q(\gamma)}^{-\frac{\nu+2}{2}} \left(1 + \frac{Q(x)}{\nu}\right)^{\frac{\nu+2}{2}}}, \\ c &= \frac{2^{1-\frac{\nu+2}{2}}}{\Gamma(\frac{\nu}{2})\pi\nu|R_t|^{0.5}}, \\ Q(x) &= x' R_t^{-1} x', \\ Q(\gamma) &= \gamma' R_t^{-1} \gamma', \end{aligned} \tag{13}$$

and $K_\lambda(.)$ is the modified Bessel function of the third kind with index λ. We obtain $C(u, u|R_t, \nu_g, \gamma_g)$ as the numerical integral. Then, we take the average of $C(u, u|R_t, \nu_g, \gamma_g)$ over T observations for the mean correlation of all assets in the same group. The tail penultimate dependence are,

$$\begin{aligned} \lambda_L &= \sum_{t=1}^{T} \frac{1}{T} \frac{C_{GSt}(u, u|R_t, \nu_g, \gamma_g)}{u}, \text{ and,} \\ \lambda_U &= \sum_{t=1}^{T} \frac{1}{T} \frac{1 - 2u + C_{GSt}(1 - u, 1 - u|R_t, \nu_g, \gamma_g)}{u} \end{aligned}$$

Table (6) reports the penultimate tail dependence of bivariate MGSt copula at $u = 0.005$.

D Posterior inference

From the joint posterior of the dynamic MGSt factor copula model in (9), we derive the conditional posterior for each parameters as follows:

$$
\begin{aligned}
p(z_t|u,a,b,f_c,z_{1:(t-1)},\nu,\gamma,\zeta) &\propto \prod_{g=1}^{G}\prod_{i=1}^{n_g}\phi(\tilde{x}_{igt}\mid \rho_{igt}z_t,\sqrt{1-\rho_{igt}^2})\phi(z_t\mid 0,1)\\
p(f_{igc}|u,a,b,z,\nu,\gamma,\zeta) &\propto \prod_{t=1}^{T}\prod_{i=1}^{n_g}\phi(\tilde{x}_{igt}\mid \rho_{igt}z_t,\sqrt{1-\rho_{igt}^2})\\
p(a_g|u,b,f,z,\nu,\gamma,\zeta) &\propto \prod_{t=1}^{T}\prod_{i=1}^{n_g}\phi(\tilde{x}_{igt}\mid \rho_{igt}z_t,\sqrt{1-\rho_{igt}^2})\\
p(b_g|u,a,f,z,\nu,\gamma,\zeta) &\propto \prod_{t=1}^{T}\prod_{i=1}^{n_g}\phi(\tilde{x}_{igt}\mid \rho_{igt}z_t,\sqrt{1-\rho_{igt}^2})\\
p(\nu_g|u,f,a,b,z,\gamma,\zeta) &\propto \prod_{t=1}^{T}\prod_{i=1}^{n_g}\frac{\phi(\tilde{x}_{igt}\mid \rho_{igt}z_t,\sqrt{1-\rho_{igt}^2})}{f_{GSt}(x_{igt}|\nu_g,\gamma_g)}\\
&\times\prod_{t=1}^{T}IG\left(\zeta_{gt}|\frac{\nu_g}{2},\frac{\nu_g}{2}\right)G(\nu_g-4\mid 2,2.5)\\
p(\gamma_g|u,f,a,b,z,\nu,\zeta) &\propto \prod_{t=1}^{T}\prod_{i=1}^{n_g}\frac{\phi(\tilde{x}_{igt}\mid \rho_{igt}z_t,\sqrt{1-\rho_{igt}^2})}{f_{GSt}(x_{igt}|\nu_g,\gamma_g)}\phi(\gamma_g\mid 0,1)\\
p(\zeta_{gt}|u,f,a,b,z,\nu,\gamma,\zeta_{g,1:(t-1)}) &\propto \prod_{i=1}^{n_g}\frac{\phi(\tilde{x}_{igt}\mid \rho_{igt}z_t,\sqrt{1-\rho_{igt}^2})}{\sqrt{\zeta_{gt}}}IG\left(\zeta_{gt}\mid\frac{\nu_g}{2},\frac{\nu_g}{2}\right)
\end{aligned}
$$

As the conditional posterior of z_t depends on the pseudo observations at time t and the GAS process ρ_{igt}, it is fast to sequentially sample from the Gaussian conjugate distribution. Also, the conditional posteriors of a_g, b_g, ν_g, γ_g and ζ_{gt} only depend on the pseudo observations in group g. Then, we can make parallel inference for $g = 1, \ldots, G$. Finally, conditional on z_t, each time series is independent, for $i = 1, \ldots, n_g$ and $g = 1, \ldots, G$. Then, we also create a parallel estimation procedure for f_{igc}.

E Model selection

The statistics of model selection are calculated based on the average of the log-likelihood. We take the average of the log likelihood after MCMC iterations at the posterior mean of the parameters of interest, $\theta_{int} = \{a, b, f_c, \nu, \gamma\}$, as the integral over the nuisance parameter space, $\theta_{nui} = \{z, \zeta\}$. Then, the AIC, BIC and DIC are given,

$$
\begin{aligned}
AIC &= -2\mathbb{E}_{\theta_{nui}}\left[\log\left(p\left(u|\bar{\theta}_{int}, f, \mathcal{F}\right)\right)\right] + 2k \\
BIC &= -2\mathbb{E}_{\theta_{nui}}\left[\log\left(p\left(u|\bar{\theta}_{int}, f, \mathcal{F}\right)\right)\right] + k\log T \\
DIC &= -4\mathbb{E}_{\theta}\left[\log p\left(u|\theta, f, \mathcal{F}\right)|u\right] + 2E_{\theta_{nui}}\left[\log\left(p\left(u|\bar{\theta}_{int}, f, \mathcal{F}\right)\right)\right]
\end{aligned}
$$

where k is the number of parameters of interest.

References

K. Aas and I. H. Haff. The generalized hyperbolic skew students t-distribution. *Journal of Financial Econometrics*, 4(2):275–309, 2006.

G. J. Alexander and A. M. Baptista. A comparison of VaR and CVaR constraints on portfolio selection with the mean-variance model. *Management Science*, 50(9):1261–1273, 2004.

L. Andersen and J. Sidenius. Extensions to the Gaussian copula: Random recovery and random factor loadings. *Journal of Credit Risk*, 1(1):29–70, 2004.

A. Ang and J. Chen. Asymmetric correlations of equity portfolios. *Journal of Financial Economics*, 63(3):443–494, 2002.

M. C. Ausín and H. F. Lopes. Time-varying joint distribution through copulas. *Computational Statistics & Data Analysis*, 54(11):2383–2399, 2010.

G. W. Bassett Jr, R. Koenker, and G. Kordas. Pessimistic portfolio allocation and choquet expected utility. *Journal of Financial Econometrics*, 2(4):477–492, 2004.

F. Blasques, S. J. Koopman, and A. Lucas. Information-theoretic optimality of observation-driven time series models for continuous responses. *Biometrika*, 102(2):325–343, 2015.

T. Bollerslev. Generalized autoregressive conditional heteroskedasticity. *Journal of Econometrics*, 31(3):307–327, 1986.

X. Chen and Y. Fan. Estimation and model selection of semiparametric copula-based multivariate dynamic models under copula misspecification. *Journal of Econometrics*, 135(1-2):125–154, 2006.

U. Cherubini, S. Mulinacci, F. Gobbi, and S. Romagnoli. *Dynamic copula methods in finance*. John Wiley & Sons, 2011.

P. Christoffersen, V. Errunza, K. Jacobs, and H. Langlois. Is the potential for international diversification disappearing? a dynamic copula approach. *The Review of Financial Studies*, 25(12): 3711–3751, 2012.

D. Creal, S. J. Koopman, and A. Lucas. A dynamic multivariate heavy-tailed model for time-varying volatilities and correlations. *Journal of Business & Economic Statistics*, 29(4):552–563, 2011.

D. Creal, S. J. Koopman, and A. Lucas. Generalized autoregressive score models with applications. *Journal of Applied Econometrics*, 28(5):777–795, 2013.

D. D. Creal and R. S. Tsay. High dimensional dynamic stochastic copula models. *Journal of Econometrics*, 189(2):335 – 345, 2015.

T. Datastream. Thomson reuters datastream. *In Constituents of the S&P 500*, (Retrieved: March 2018 from Datastream database), 2018.

S. Demarta and A. J. McNeil. The t copula and related copulas. *International Statistical Review*, 73(1):111–129, 2005.

A. Dias and P. Embrechts. Modeling exchange rate dependence dynamics at different time horizons. *Journal of International Money and Finance*, 29(8):1687–1705, 2010.

R. Engle. Dynamic conditional correlation: A simple class of multivariate generalized autoregressive conditional heteroskedasticity models. *Journal of Business & Economic Statistics*, 20(3):339–350, 2002.

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

C. B. Erb, C. R. Harvey, and T. E. Viskanta. Forecasting international equity correlations. *Financial Analysts Journal*, 50(6):32–45, 1994.

J. Fan and R. Li. Variable selection via nonconcave penalized likelihood and its oracle properties. *Journal of the American Statistical Association*, 96(456):1348–1360, 2001.

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

C. Fernández and M. F. Steel. On Bayesian modeling of fat tails and skewness. *Journal of the American Statistical Association*, 93(441):359–371, 1998.

A. Ghalanos. *rugarch: Univariate GARCH models.*, 2018. R package version 1.4-0.

L. R. Glosten, R. Jagannathan, and D. E. Runkle. On the relation between the expected value and the volatility of the nominal excess return on stocks. *The Journal of Finance*, 48(5):1779–1801, 1993.

A. C. Harvey. *Dynamic models for volatility and heavy tails: with applications to financial and economic time series*, volume 52. Cambridge University Press, 2013.

J. Hull and A. White. Valuation of a CDO and an n-th to default CDS without monte carlo simulation. *The Journal of Derivatives*, 12(2):8–23, 2004.

R. Jagannathan and T. Ma. Risk reduction in large portfolios: Why imposing the wrong constraints helps. *The Journal of Finance*, 58(4):1651–1683, 2003.

H. Joe. Asymptotic efficiency of the two-stage estimation method for copula-based models. *Journal of Multivariate Analysis*, 94(2):401–419, 2005.

E. Jondeau and M. Rockinger. The Copula-GARCH model of conditional dependencies: An international stock market application. *Journal of International Money and Finance*, 25(5):827–853, 2006.

T.-H. Kim and H. White. On more robust estimation of skewness and kurtosis. *Finance Research Letters*, 1(1):56–73, 2004.

R. Koenker and G. Bassett Jr. Regression quantiles. *Econometrica*, pages 33–50, 1978.

S. J. Koopman, A. Lucas, and M. Scharth. Predicting time-varying parameters with parameter-driven and observation-driven models. *Review of Economics and Statistics*, 98(1):97–110, 2016.

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

E. R. Lee, H. Noh, and B. U. Park. Model selection via bayesian information criterion for quantile regression models. *Journal of the American Statistical Association*, 109(505):216–229, 2014.

M. Li and M. Scharth. Leverage, asymmetry and heavy tails in the high-dimensional factor stochastic volatility model. *UTS Working Paper Series*, (49), 2018.

F. Longin and B. Solnik. Extreme correlation of international equity markets. *Journal of Finance*, 56(2):649–676, 2001.

A. Lucas, B. Schwaab, and X. Zhang. Conditional euro area sovereign default risk. *Journal of Business & Economic Statistics*, 32(2):271–284, 2014.

A. Lucas, B. Schwaab, and X. Zhang. Modeling financial sector joint tail risk in the euro area. *Journal of Applied Econometrics*, 32(1):171–191, 2017.

A. Lucas, J. Schaumburg, and B. Schwaab. Bank business models at zero interest rates. *Journal of Business & Economic Statistics*, (to appear), 2018.

H. Manner and J. Segers. Tails of correlation mixtures of elliptical copulas. *Insurance: Mathematics and Economics*, 48(1):153–160, 2011.

H. Markowitz. Portfolio selection. *The Journal of Finance*, 7(1):77–91, 1952.

A. J. McNeil, R. Frey, and P. Embrechts. *Quantitative risk management: concepts, techniques, and tools*. Princeton university press, 2010.

J. Mencía and E. Sentana. Estimation and testing of dynamic models with generalized hyperbolic innovations. *CEMFI Working Paper 0411*, 2005.

J. S. Murray, D. B. Dunson, L. Carin, and J. E. Lucas. Bayesian Gaussian copula factor models for mixed data. *Journal of the American Statistical Association*, 108(502):656–665, 2013.

D. B. Nelson. Conditional heteroskedasticity in asset returns: A new approach. *Econometrica*, pages 347–370, 1991.

H. Nguyen, M. C. Ausín, and P. Galeano. Variational inference for high dimensional structured factor copulas. *UC3M Working Papers Statistics and Econometrics*, WP18-05, 2018.

D. H. Oh and A. J. Patton. Modelling dependence in high dimensions with factor copulas. *Journal of Business & Economic Statistics*, 35(1), 2017a.

D. H. Oh and A. J. Patton. Time-varying systemic risk: Evidence from a dynamic copula model of CDS spreads. *Journal of Business & Economic Statistics*, 2017b. doi: 10.1080/07350015.2016.1177535.

A. J. Patton. Estimation of copula models for time series of possibly different lengths. *Journal of Applied Econometrics*, 21(2):147–173, 2006a.

A. J. Patton. Modelling asymmetric exchange rate dependence. *International Economic Review*, 47(2):527–556, 2006b.

A. J. Patton. A review of copula models for economic time series. *Journal of Multivariate Analysis*, 110:4–18, 2012.

G. O. Roberts and J. S. Rosenthal. Examples of adaptive MCMC. *Journal of Computational and Graphical Statistics*, 18(2):349–367, 2009.

D. Scott and F. Grimson. Skewhyperbolic: The skew hyperbolic student t-distribution. 2015. URL `https://CRAN.R-project.org/package=SkewHyperbolic`. R package version 0.3-2.

R. Tibshirani. Regression shrinkage and selection via the lasso. *Journal of the Royal Statistical Society, Series B (Methodological)*, pages 267–288, 1996.

M. van der Voort. Factor copulas: External defaults. *The Journal of Derivatives*, 14(3):94–102, 2007.

Q. Xu, Y. Zhou, C. Jiang, K. Yu, and X. Niu. A large CVaR-based portfolio selection model with weight constraints. *Economic Modelling*, 59:436–447, 2016.