Parallel Bayesian inference for high dimensional dynamic factor copulas

Hoang Nguyen, M. Concepción Ausín and Pedro Galeano

Department of Statistics

Universidad Carlos III de Madrid, Spain

February 27, 2017

Abstract

Copula densities are widely used to model the dependence structure of financial time series. However, the number of parameters involved becomes explosive in high dimensions which results in most of the models in the literature being static. Factor copula models have been recently proposed for tackling the curse of dimensionality by describing the behaviour of return series in terms of a few common latent factors. To account for asymmetric dependence in extreme events, we propose a class of dynamic one factor copula where the factor loadings are modelled as generalized autoregressive score (GAS) processes. We perform Bayesian inference in different specifications of the proposed class of dynamic one factor copula models. Conditioning on the latent factor, the components of the return series become independent, which allows the algorithm to run in a parallel setting and to reduce the computational cost needed to obtain the conditional posterior distributions of model parameters. We illustrate our approach with the analysis of a simulated data set and the analysis of the returns of 150 companies listed in the S&P500 index.

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

1 Introduction

Many scholars have strived to model the dependence between financial time series. For instance, several multivariate GARCH models have been used for stock returns and exchange rates, see the Vector GARCH (VEC) models in Bollerslev et al. (1988) and Ding and Engle (2001), the BEKK GARCH models in Engle and Kroner (1995) and Ding and Engle (2001), the Constant Conditional Correlation (CCC) model by Bollerslev (1990), the Dynamic Conditional Correlation (DCC) models in Engle (2002), Tse and Tsui (2002), Capiello et al. (2006) and Galeano and Ausín (2010), the Orthogonal GARCH (OGARCH) models in Alexander and Chibumba and van der Wiede (2002), and the Factor GARCH models in Engle et al. (1990), Vrontos et al. (2003) and Lanne and Saikkonen (2007), among many others. However, these multivariate GARCH models encounter several problems in high dimensional settings for several reasons. Firstly, traditional dynamic correlations are not very useful in high dimensions. Secondly, the number of parameters of these models usually explodes faster than the number of dimensions, leading to intractable estimation problems. Thirdly, these models usually assume that standardized innovations have either a multivariate Gaussian or a multivariate Student-t distribution. Therefore, they assume the same parametric structure for each marginal distribution. Furthermore, the assumption that only a few parameters account for extreme financial events can be very restrictive. Shocks often affect to a group of assets in some circumstances rather than to each individual equally.

One possibility to avoid these issues is that of assuming a different univariate volatility model for each individual return series and a copula model for their dependence structure, see, for instance, Jondeau and Rockinger (2006), Patton (2006a) and Ausín and Lopes (2010), among others. 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. Particularly, copulas have been frequently used in finance and econometrics, see, for instance, Cherubini et al. (2011), Patton (2012) and Fan and Patton (2014). There are well known standard copula families, such as the elliptical copulas and the Archimedean copulas. Nevertheless, when the dimension is large, the use of these standard copula distributions is also 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 recorded when there is greater correlation during the bear market than during the bull market, see e.g. Erb et al. (1994), Longin and Solnik (2001), Ang and Chen (2002) and Lucas et al. (2014).

Alternatively, one can make use of factor copula models which assumes that the dependence structure of observable variables depends on a few latent variables. This is the approach taken in Hull and White (2004), van der Voort (2007), Murray et al. (2013), Krupskii and Joe (2013), Creal and Tsay (2015), Oh and Patton (2017a) and Oh and Patton (2017b), among others. In particular, Hull and White (2004) proposed 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 n^{th} to default swaps. Andersen and Sidenius (2004) and Oh and Patton (2017a) improve the model by considering a non-linear approach and the Student-t copula. On the other hand, Krupskii and Joe (2013) proposed a general class of factor copulas where the copula variables are conditionally independent given a few latent variables. The dependence structure is determined by bivariate linking 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. On the other hand, 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 function are static.

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 AR-GARCH processes to model the individual returns. Then, the series of standardized innovations are converted into series of copula observations, using inverse cumulative distribution functions, that are assumed to have a copula distributions. Particularly, as in Creal and Tsay (2015), we consider several copula specifications including Gaussian, Student-t and Hyperbolic Skew Student-t copulas. To handle a large number of returns, we assume a one factor structure that, first, drastically reduces the the number of parameters as they scales linearly with the dimension, and, second, provides natural economic interpretations. Additionally, we model the dynamic factor loadings as GAS processes, see Creal et al. (2013) and Harvey (2013). The GAS process is an observation driven process in which the dynamic behaviour depends on the complete density of the process rather than their first or second moments. Particularly, Koopman et al. (2016) find evidence of such assertion with several simulations and empirical data analysis. Importantly, we assume that the dynamic factor loadings equation depends on the copula density conditional on the factor rather than the unconditional copula density. The main benefit of such approach is that it allows us to perform parallel inference which heavily reduces the computational cost needed to obtain the conditional posterior distributions of model parameters. Finally, note that the proposed methodology allows for different tail behaviour and asymmetric dependence among financial returns.

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 by different types of factor copula models. We present our parallel Bayesian inference strategy in Section 3. Section 4 illustrates the performance of factor copula models with a simulated example. Then, we analyze a large series of stock returns listed in S&P 500 in Section 5 with the proposed method. Finally, conclusions are drawn in Section 6.

2 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-GARCH structure 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 $AR\left(k_i\right) - GARCH\left(p_i, q_i\right)$ model given by:

$$
\begin{aligned}
r_{it} &= c_i + \phi_{i1} r_{i,t-1} + \ldots + \phi_{ik_i} r_{i,t-k_i} + a_{it} \\
a_{it} &= \sigma_{it} \eta_{it} \\
\sigma_{it}^2 &= \omega_i + \alpha_{i1} a_{i,t-1}^2 + \ldots + \alpha_{ip_i} a_{i,t-p_i}^2 + \beta_{i1} \sigma_{i,t-1}^2 + \ldots + \beta_{iq_i} \sigma_{i,t-q_i}^2
\end{aligned}
$$

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 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 $\alpha_{i1}, \ldots, \alpha_{ip_i}, \beta_{i1}, \ldots, \beta_{iq_i}$ are GARCH parameters verifying the usual stationarity conditions. We note that the previous AR-GARCH 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 GARCH specification can be replaced with an EGARCH or a GJR-GARCH process.

Once 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 uniform $U(0,1)$ 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, as mentioned in the introduction, it is questionable whether such standard copula functions are appropriate for u_t when d is large enough. One plausible alternative is to assume, as in Krupskii and Joe (2013), a copula factor 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 correlation matrix parameter, R_t, characterized by a factor structure, somehow coming back to standard factor models widely analyzed in the literature. As in Creal and Tsay (2015), we model the dynamic factor loadings as GAS processes, but we assume that the dynamic factor loadings equation depends on the copula density conditional on the 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 Gaussian, Student-t and generalized hyperbolic skew Student-t one factor copula models and present some of their advantages over the existing alternatives.

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 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 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, $D_t = diag\left(\sqrt{1-\rho_t^2}\right)$ is a diagonal matrix whose elements are $\sqrt{1-\rho_{it}^2}$, for $i = 1, \ldots, d$, and $\epsilon_t = (\epsilon_{1t}, \ldots, \epsilon_{dt})'$, the noise, is a sequence of independent and identically standard multivariate Gaussian random variables. 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 $R_t = \rho_t \rho_t' + D_t D_t' = I_d + \rho_t \rho_t' - diag\left(\rho_t^2\right)$, where I_d is the d-dimensional unit diagonal matrix and $diag\left(\rho_t^2\right)$ stands for the diagonal matrix with elements $\rho_{1t}^2, \ldots, \rho_{dt}^2$. 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,

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

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, where $|b| < 1$ to guarantee stationarity, and $p\left(u_t | z_t, f_t, \mathcal{F}_t, \theta\right)$ 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 through 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 $f_{i,t}$ depends linearly on $f_{i,t-1}$ 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 in its first or second moment. Blasques et al. (2015) proved that the use of the score s_{it} leads to 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 type 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 computation burden as the score s_{it} has a closed form expression that allow us to parallelize the derivation of the d processes $s_{1t}, \ldots, s_{dt}$. Specifically, as showed 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 because, as will be shown in the next subsections, s_{it} has a similar structure to the given in (3) for the dynamic Student-t and generalized hyperbolic skew Student-t one factor copula models.

The main difference of our proposed model with respect to the the dynamic GAS model defined in Oh and Patton (2017b) is that we define the score in Equation (2) as the density derivative conditioning on the latent variable, z_t, while in Oh and Patton (2017b), the definition of s_{it} is not conditioned on the latent factor. 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} are obtained by numerical differentiation of the joint copula density, which is much more computationally expensive.

2.3 Dynamic Student-t one factor copula model

Next, we extend the dynamic Gaussian one factor copula model to the Student copula case. The standard Student-t distribution depends on the degrees of freedom parameter ν, that controls the generation of extreme events and consequently, the Student-t copula allows for tail dependencies which are not possible with the Gaussian copula assumption. Here, we impose that the joint distribution function of u_t is given by $C(u_{1t}, \ldots, u_{dt} \mid R_t, \nu) = F_{MSt}\left(F_{St}^{-1}(u_{1t} \mid \nu), \ldots, F_{St}^{-1}(u_{1t} \mid \nu) \mid R_t, \nu\right)$, where $F_{St}(\cdot \mid \nu)$ denotes the univariate standard Student-t cdf with degrees of freedom ν and $F_{MSt}(\cdot \mid R_t, \nu)$ denotes the multivariate Student-t cdf with R_t correlation matrix and degrees of freedom ν. In this case, the inverse cdf transformations, $x_t = (x_{1t}, \ldots, x_{dt})'$, where $x_{it} = F_{St}^{-1}(u_{it} \mid \nu)$, for each $i = 1, \ldots, d$, follows a multivariate Student-t distribution with zero mean, correlation matrix R_t and degrees of freedom ν. Then, we assume a dynamic Student-t one factor copula model for x_t given by:

$$x_t = \sqrt{\zeta_t}\left(\rho_t z_t + D_t \epsilon_t\right)$$

where z_t, ϵ_t and ρ_t, for $t = 1, \ldots, T$, 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. Accordingly, the components of the multivariate random vector $x_t = (x_{1t}, \ldots, x_{dt})'$ are conditionally independent given z_t, ρ_t and ζ_t.

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

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

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, $\theta = (\nu, a, b, f_{1c}, \ldots, f_{dc})'$. Again, this model setting is influenced by the developments in Creal and Tsay (2015) for dynamic stochastic copulas. However, one advantage of our proposal is that the observation driven process remains similar. As shown in

Appendix A.2, if we let $\dot{x}_{it} = \frac{x_{it}}{\sqrt{\zeta_t}}$, the score function is,

$$s_{it} = \frac{\partial \log p(u_t|z_t, \zeta_t, f_t, \mathcal{F}_t, \theta)}{\partial f_{it}} = \frac{1}{2}\dot{x}_{it}z_t + \frac{1}{2}\rho_{it} - \rho_{it}\frac{\dot{x}_{it}^2 + z_t^2 - 2\rho_{it}\dot{x}_{it}z_t}{2\left(1-\rho_{it}^2\right)} \tag{4}$$

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 rather different from the Student 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 easy attainable conditional cdf for x_t and therefore, it is computationally expensive to derive the score s_{it}, as mentioned before.

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

Next, as in Creal and Tsay (2015), we use the generalized hyperbolic skew Student-t (GSt) distribution proposed by Aas and Haff (2006) to extend the Gaussian and Student-t models. The GSt distribution depends on two parameters, ν and γ, that controls the generation of extremes events and skewness, respectively. Particularly, the GSt distribution reduces to the Student-t distribution when $\gamma = 0$ and reduces to the Gaussian distribution when $\gamma = 0$ and $\nu \to \infty$. Additionally, the multivariate generalized hyperbolic skew Student-t (MGSt) also depends on a correlation matrix, R_t.

Here, we assume that the joint distribution function of u_t is given by $C(u_{1t}, \ldots, u_{dt} \mid R_t, \nu, \gamma) = F_{MGSt}\left(F_{GSt}^{-1}(u_{1t} \mid \nu, \gamma), \ldots, F_{GSt}^{-1}(u_{1t} \mid \nu, \gamma) \mid R_t, \nu\right)$, 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 multivariate GSt cdf with parameters ν and γ and correlation matrix R_t. Here, the vector of inverse cdf transformations, $x_t = (x_{1t}, \ldots, x_{dt})'$, where $x_{it} = F_{GSt}^{-1}(u_{it} \mid \nu, \gamma)$, for each $i = 1, \ldots, d$, follows a MGSt with zero mean vector, correlation 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:

$$x_t = \gamma\zeta_t + \sqrt{\zeta_t}\left(\rho_t z_t + D_t \epsilon_t\right) \tag{5}$$

for $i=1,\ldots,d$, where ζ_t, z_t, ϵ_t and ρ_t, for $t=1,\ldots,T$, are as in the Gaussian and Student-t cases. Consequently, $x_{1t},\ldots,x_{dt}$ are conditionally independent given z_t, ρ_t and ζ_t.

As in the Student-t case, the correlation vector $\rho_t=(\rho_{1t},\ldots,\rho_{dt})'$ is allowed to vary across time as in (2), with s_{it} as in (4), where the parameters of the copula function are $\theta=(\nu,\gamma,a,b,f_{1c},\ldots,f_{dc})'$. As mentioned before, the score function shown in the Appendix A.3 is also similar to the previous models,

$$s_{it}=\frac{\partial\log p(u_t|z_t,\zeta_t,f_t,\mathcal{F}_t,\theta)}{\partial f_{it}}=\frac{1}{2}\tilde{x}_{it}z_t+\frac{1}{2}\rho_{it}-\rho_{it}\frac{\tilde{x}_{it}^2+z_t^2-2\rho_{it}\tilde{x}_{it}z_t}{2\left(1-\rho_{it}^2\right)} \tag{6}$$

where $\tilde{x}_{it}=\frac{x_{it}-\gamma\zeta_t}{\sqrt{\zeta_t}}$. Observe that again the conditional distribution does not change which makes easier to generate the dynamic process and reduces the computation burden.

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 a 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 numerical approximation of the joint quantile exceedance probability, see Appendix A.4. Finally, Demarta and McNeil (2005) suggest 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 be in general intractable.

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

One potential drawback of the previous models for high dimensional returns is that few parameters control all of the co-movements which can be very restrictive. 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 $x_{igt}=F_{GSt}^{-1}\left(u_{igt}|\nu_g,\gamma_g\right)$ for each asset i, for $i = 1, \ldots, n_g$, belonging to group g, where $g = 1, \ldots, G$, such that,

$$x_{gt} = \gamma_g \zeta_{gt} + \sqrt{\zeta_{gt}} \left(\rho_{gt} z_t + D_{gt} \epsilon_{gt} \right) \tag{7}$$

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 $D_{gt} = diag\left(\sqrt{1 - \rho_{gt}^2} \right)$ 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 correlation in group g is given by:

$$\begin{aligned} \rho_{igt} &= \frac{1 - \exp\left(-f_{igt}\right)}{1 + \exp\left(-f_{igt}\right)} \\ f_{igt} &= \left(1 - b_g\right) f_{igc} + a_g s_{i,g,t-1} + b_g f_{i,g,t-1} \end{aligned} \tag{8}$$

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

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

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 behaviour in the upper and lower tails 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 $x_{igt} = F_{GSt}^{-1}\left(u_{igt} | \nu_g, \gamma_g\right)$ 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 estimation

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

3.1 Prior distributions

Next, we define a prior distribution for both the marginal and copula parameters. In all cases, we use proper but uninformative prior assumptions.

For the marginal parameters, c_i, $\phi_{i1}, \ldots, \phi_{ik_i}$, ω_i, $\alpha_{i1}, \ldots, \alpha_{ip_i}$ $\beta_{i1}, \ldots, \beta_{iq_i}$, we assume uniform prior distributions restricted to the stationary region. Also, we need to define prior distributions for the parameters describing the innovation distributions, F_{η_i}. For example, in the empirical data example, we assume that the standardized innovations, η_{it}, follow univariate GSt distributions, defined in Section 2.4, with degrees of freedom $\nu_{i\eta}$ and skewness parameter $\gamma_{i\eta}$, for $i = 1, \ldots, d$. Then, as Creal and Tsay (2015) suggest, we assume prior shifted Gamma distributions for the degrees of freedom parameters, such that $\nu_{i\eta} = 4 + \tilde{\nu}_{i\eta}$, where $\tilde{\nu}_{i\eta} \sim G(2, 0.5)$, such that $\nu_{i\eta} > 4$, to ensure a finite variance of the GSt distribution. Also, we assume a priori that $\gamma_{i\eta} \sim N(0, 1)$.

For the copula parameters, we define a 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 the mean correlations 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 The 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, $\vartheta_i = (c_i, \phi_{i1}, \ldots, \phi_{ik_i}, \omega_i, \alpha_{i1}, \ldots, \alpha_{ip_i}, \beta_{i1}, \ldots, \beta_{iq_i}, \nu_{i\eta}, \gamma_{i\eta})'$, and the set of factor copula parameters, $\vartheta_c = (z, \zeta, f_c, a, b, \nu, \gamma)'$. The likelihood is given by,

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

where $c(\cdot \mid \vartheta_c)$ denotes the copula density function with parameters ϑ_c and $f_{\cdot\eta_i}(\mid \vartheta_i)$ 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 approximate the posterior distributions of the marginal parameters, ϑ_i, independently for each $i = 1, \ldots, d$, and, in a second step, we use the posterior means of the marginal parameters, $\bar{\vartheta}_i$, for $i = 1, \ldots, d$, to obtain an approximate sample of the copula observations, $u = \{u_t : t = 1, \ldots, T\}$, where $u_{it} = F_{\eta_i}\left(\eta_{it} \mid \bar{\vartheta}_i\right)$, for $t = 1, \ldots, T$ and for each $i = 1, \ldots, d$. 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:

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

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 $\theta = (f_c, a, b, \nu, \gamma)'$ is the vector of static parameters. Therefore, given the conditional density (10), 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 $x_{igt} = F_{GSt}^{-1}\left(u_{it} \mid \nu_g, \gamma_g\right)$.

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 only depends on the pseudo observations at time t, we can make parallel inference for each latent variable at time $t = 1, \ldots, T$. Also, the conditional posterior of a_g and 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 both high dimensional and long period time series 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 $j = 1, \ldots, N$, obtain $\rho_{igt}^{(j)}$ for $i = 1, \ldots, n_g$, $g = 1, \ldots, G$ and $t = 1, \ldots, T$:

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

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

The conditional posteriors distributions for all the parameters are given in the Appendix B. In the algorithm, we apply the Gibbs sampler for step $2a$ and the Adaptive Random Walk Metropolis Hasting (MH) (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 value time period t, we accept $\zeta_{gt}^{(j)}$ with probability:

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

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 one factor copula since $\gamma = 0$. The codes and implementation of the algorithm are available at https://github.com/hoanguc3m/FactorCopula.

4 Data simulation

In this section, we illustrate the proposed Bayesian methodology using simulated data from the MGSt one factor copula. We generate a random sample of $d = 100$ time series with $G = 10$ groups of the same size and a time length $T = 1000$ from (7). The value of the parameters a, b, ν, and γ 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 $U(6, 18)$ and γ is generated from a $U(-1, 0)$ distribution. 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, 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 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. Figure 2 illustrates the comparison between the posterior mean of f_{igc} versus its true value, for $i = 1, \dots, 10$ and $g = 1 \dots, 10$, and z_t versus its true value, for $t = 1, \dots, 1000$. We obtain quite accurate results. Observe that the latent variable z_t is generated from a standard Gaussian distribution and, for each t, we obtain a symmetric posterior distribution rather concentrated around the mean. The posterior variance of z_t also reduces when the dimension increases. In general, most of the parameters which govern the dynamic correlation in each group are correctly estimated, see the Online Appendix in the web site https://github.com/hoanguc3m/FactorCopula for additional results.

5 Empirical data

In this section, we illustrate our approach with a series of $d = 150$ stock returns of companies listed in the S&P 500 index, from 01/01/2010 to 31/12/2015. The daily data are taken from Yahoo finance and contain $T = 1509$ days observed during the considered six-year period. Table 1 shows

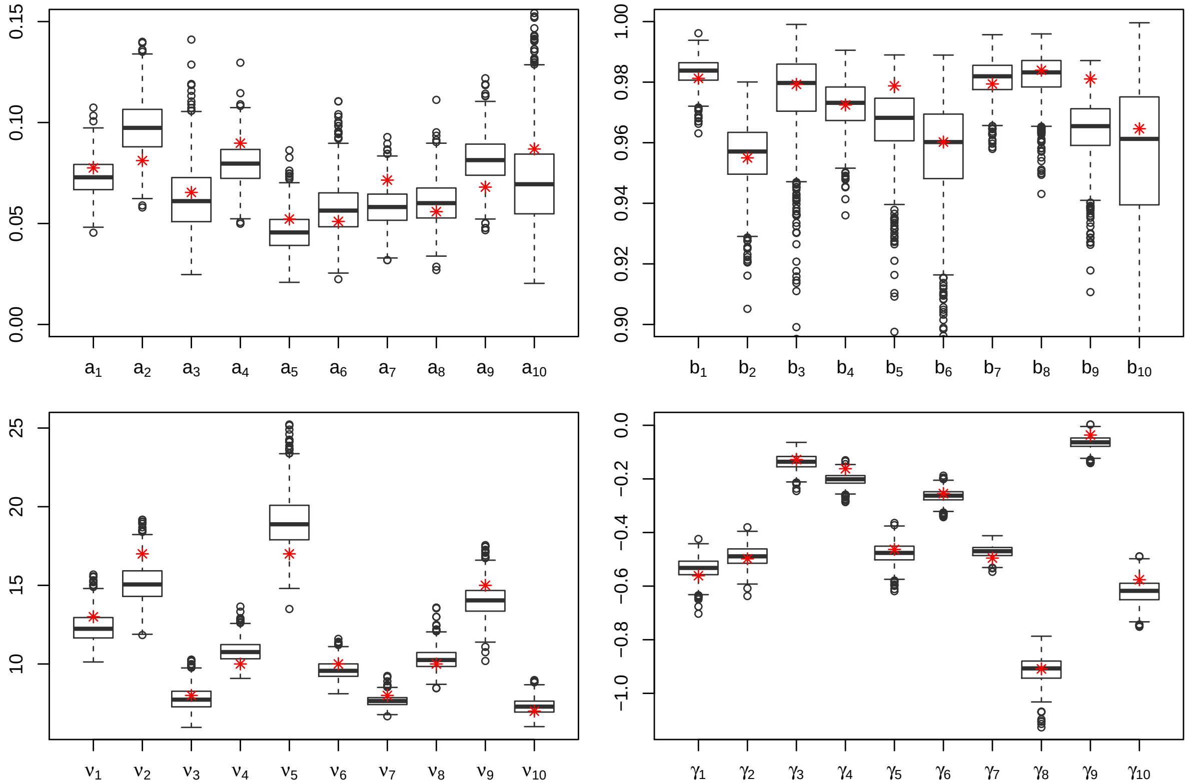

Figure 1: Box plots for the posterior samples of a_g, b_g, ν_g, and γ_g and true values (stars)

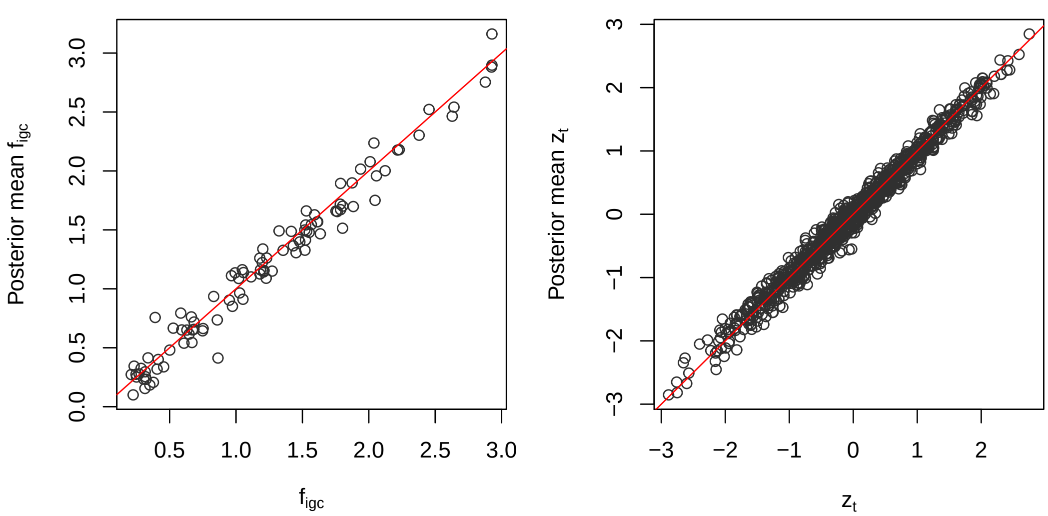

Figure 2: Posterior means of f_{igc} and z_t versus true values

Table 1: Summary statistics for cross-sectional daily returns of 150 firms listed in S&P 500

	Minimum	1st Qu.	Median	Mean	3rd Qu.	Maximum
Mean	-0.083	0.035	0.055	0.054	0.072	0.146
Minimum	-19.970	-13.020	-10.080	-10.620	-7.962	-4.528
1. Quartile	-1.599	-0.957	-0.784	-0.820	-0.645	-0.430
Median	-0.082	0.035	0.062	0.060	0.084	0.160
3. Quartile	0.515	0.811	0.923	0.951	1.070	1.778
Maximum	3.961	7.144	9.510	9.743	11.530	18.470
Skewness	-1.552	-0.302	-0.137	-0.155	0.053	0.552
Kurtosis	1.444	3.014	4.170	4.944	5.670	20.540

Summary statistics for cross-sectional daily returns of 150 firms listed in S&P 500 index. The distribution of the common statistics are shown by rows label using quantiles and mean

the summary statistics for the daily stock returns. The mean of daily earnings over the 150 stocks is 0.054%. The most extreme individual events are a worst one day crash at -19.97% and a best one day gain at 18.47%. The average skewness is -0.155, which reflects a slight asymmetry of the observed returns, and the average excess kurtosis is 4.944, which shows the extreme heavy tails of the return distributions.

As described in Section 3, we use a two-stage procedure to model the dependence structure of the stock returns. In Subsection 5.1, we fit a $AR\,(1){-}GARCH\,(1,1)$ model for the conditional mean and variance of marginal returns. Then, we take out the standardized innovation and transform them into the copula observations using the corresponding cdf. In Subsection 5.2, we estimate the one factor copula model and illustrate some empirical findings.

5.1 Marginal distributions

For simplicity, we fit an $AR\,(1) - GARCH\,(1,1)$ model for each marginal return series. Thus, we assume that:

$$r_{it} = c_i + \phi_{i1} r_{i,t-1} + a_{it}$$

$$a_{it} = \sigma_{it} \eta_{it}$$

$$\sigma_{it}^2 = \omega_i + \alpha_{i1} a_{i,t-1}^2 + \beta_{i1} \sigma_{i,t-1}^2$$

where the standardized innovations, η_{it}, are assumed to follow univariate GSt distributions, introduced in Section 2.4, with degrees of freedom $\nu_{i\eta}$ and skewness parameter $\gamma_{i\eta}$, for $i = 1, \ldots, d$.

Then,

$$\eta_{it} = \left(\gamma_i \xi_{it} + \sqrt{\xi_{it}} \iota_{it} - \mu_{i\eta}\right) / \sigma_{i\eta}$$

where $\iota_{it} \sim_{iid} N(0,1)$, $\xi_{it} \sim_{iid} IG\left(\frac{\nu_{i\eta}}{2}, \frac{\nu_{i\eta}}{2}\right)$. As η_{it} is a standarized GSt, the mean $\mu_{i\eta}$ and the standard deviation $\sigma_{i\eta}$ are adapted for η_{it} to have zero mean and unit variance:

$$\mu_{i\eta} = \gamma_{i\eta} \frac{\nu_{i\eta}}{\nu_{i\eta} - 2}$$
$$\sigma_{i\eta} = \sqrt{\frac{2\gamma_{i\eta}^2 \nu_{i\eta}^2}{(\nu_{i\eta} - 2)^2 (\nu_{i\eta} - 4)} + \frac{\nu_{i\eta}}{\nu_{i\eta} - 2}}$$

Bayesian inference for this model is developed in Stan (see Stan Development Team (2015)). For each marginal, we sample 20.000 iterations from the model and discard the initial 10.000 iterations for the burn-in period. We also retain every 10-th iterations to reduce the autocorrelation.

Table 2 shows the summary statistics for the posterior mean estimations of the univariate $AR(1) - GARCH(1,1)$ model across the 150 firms. The effect of conditional autoregressive mean is weak as the average value of ϕ_{1i} is -0.021 and ϕ_{1i} 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 $GARCH(1,1)$ setting. The average of α_{i1} and β_{i1} are significantly different from 0 for most of the marginals. On average, the previous variance will contribute up to the 84% the value of the current variance which creates volatility clustering. The degrees of freedom for each marginal also diversifies between 4.2 and 10 which accounts for different kurtosis. The skewness parameter ranges from -0.373 to 0.111 and most of them are not significantly different from 0.

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 posterior mean estimations for each marginal, we obtain the standardized innovations, η_{it}, of the $AR(1) - GARCH(1,1)$ process and transform them into the copula observations $u_{it} = F_{GSt}\left(\eta_{it} | \bar{c}_i, \bar{\phi}_{1i}, \bar{\omega}_i, \bar{\alpha}_{1i}, \bar{\beta}_{1i}, \bar{\nu}_{i\eta}, \bar{\gamma}_{i\eta}\right)$. 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 a GSt distribution is suitable with univariate GARCH volatilities by performing the Kolmogorov-Smirnov goodness of fit test. All series passed the test with p-values larger than 0.05. We also tested for

Table 2: Estimation results of $AR\,(1) - GARCH\,(1,1)$-GSt for each marginal return

	Mean	Minimum	1. Quartile	Median	3. Quartile	Maximum
c_i	0.071	-0.065	0.052	0.072	0.090	0.180
ϕ_{i1}	-0.021	-0.103	-0.041	-0.021	0.000	0.045
ω_i	0.172	0.019	0.063	0.115	0.211	1.313
α_{i1}	0.097	0.022	0.075	0.093	0.113	0.220
β_{i1}	0.843	0.604	0.801	0.864	0.901	0.974
$\nu_{i\eta}$	5.834	4.163	4.828	5.589	6.491	10.121
$\gamma_{i\eta}$	-0.049	-0.373	-0.085	-0.042	-0.000	0.111

Summary statistics for the posterior mean estimations of the univariate $AR\,(1) - GARCH\,(1,1)$-GSt model across the 150 firms. The distribution of posterior mean estimations are described by mean and quantiles.

serial correlation between the innovation and did not find significant results.

5.2 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 (8) 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 $G = 14$ group industries of assets depending on their SIC codes, as in Creal and Tsay (2015). These are Oil & Construction , Food & Beverage, Pharmaceuticals, Plastic Material & Plant Chemical, Textile & Papers, Steel, Home Appliances & Automobile, Electronics, Transportation, Telecommunication, Retail & Distribution, Insurance, Finance (not contained in Insurance), Services & Others. The detail number of firms are reported in Table 2 of Online Appendix. On average, there are 11 firms in each sector group. In the single group model, $G = 1$ as described in (5), only a few number of parameters account for 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 behaviours in tail and unrestricted correlation as in (8). In all cases, we generate 20.000 iterations with 10.000 burn-in for each factor copula model and every 10-th draws are taken in order to prevent the autocorrelation of the MCMC chains.

Table 3 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 model selection criteria, obtained as explained

Table 3: 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	-139990	-140848	-141211	-168123	-153480	-168957	-168709	-153690	-169362
BIC	-139766	-140039	-140264	-167825	-152666	-167935	-168336	-152871	-168266
DIC	-139968	-140851	-141214	-168145	-153500	-169037	-169092	-153772	-170008
# params	42	152	178	56	153	192	70	154	206
a	[0.022,0.144]	0.039	[0.023,0.140]	[0.017,0.078]	0.026	[0.023,0.135]	[0.017,0.062]	0.025	[0.023,0.132]
b	[0.793,0.998]	0.982	[0.611,0.996]	[0.901,0.999]	0.993	[0.585,0.997]	[0.924,0.999]	0.993	[0.597,0.997]
ν				[7.734,20.802]	35.606	[7.766,21.253]	[8.242,35.468]	34.234	[8.121,35.857]
γ							[-1.678,-0.058]	-0.378	[-1.689,-0.036]
f_c	[1.303,1.767]	[0.935,2.439]	[0.913,2.476]	[1.366,1.834]	[0.953,2.521]	[0.933,2.520]	[1.361,1.825]	[0.949,2.519]	[0.907,2.518]

Posterior estimations for nine one factor copula models and model selection criteria. Three different models are considered (the block-equivalent, 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 Appendix C. The dynamic MGSt copula appears to show 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 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 shows a interesting result that the models with more parameters accounting for extreme events are preferable over the models that have limitations on these behaviours. 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.

In Tables 4 and 5, we report respectively the detail estimations for the dynamic group Student-t and dynamic group MGSt copulas. The posterior means of a and b shows different dynamic behaviours 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

Table 4: Results for the group Student-t copula with time-varying factor loadings.

	Oil	Food & Bev.	Pharma.	Plastics	Paper	Steel	Home App.
a	0.035	0.045	0.079	0.037	0.043	0.032	0.041
	(0.006)	(0.009)	(0.031)	(0.007)	(0.006)	(0.003)	(0.009)
b	0.996	0.989	0.902	0.988	0.986	0.995	0.960
	(0.002)	(0.006)	(0.146)	(0.005)	(0.004)	(0.001)	(0.017)
ν	12.562	13.103	7.766	14.271	14.916	17.733	15.883
	(1.215)	(1.207)	(0.596)	(1.041)	(0.906)	(1.005)	(0.804)
f_c	[1.28,1.83]	[1.10,1.99]	[1.26,1.41]	[1.21,1.97]	[1.13,1.90]	[1.33,2.03]	[1.40,2.50]
	(0.177)	(0.119)	(0.070)	(0.096)	(0.093)	(0.134)	(0.060)
# firms	5	6	5	8	12	13	14
$\lambda_L = \lambda_U$	0.031	0.022	0.072	0.026	0.019	0.019	0.038
	Electronics	Transportation	Telecom.	Retail	Insurance	Finance	Services
a	0.135	0.042	0.061	0.023	0.052	0.051	0.034
	(0.017)	(0.007)	(0.014)	(0.003)	(0.007)	(0.008)	(0.004)
b	0.585	0.978	0.973	0.997	0.972	0.984	0.991
	(0.064)	(0.007)	(0.014)	(0.002)	(0.008)	(0.006)	(0.002)
ν	17.344	9.287	8.134	21.253	9.792	9.483	20.681
	(1.106)	(0.388)	(0.470)	(1.963)	(0.375)	(0.356)	(1.013)
f_c	[1.39,2.35]	[1.00,2.12]	[0.93,2.10]	[1.10,1.87]	[1.30,2.48]	[1.45,2.52]	[1.31,2.24]
	(0.049)	(0.075)	(0.086)	(0.153)	(0.074)	(0.099)	(0.097)
# firms	12	12	7	8	14	17	17
$\lambda_L = \lambda_U$	0.026	0.086	0.087	0.007	0.106	0.099	0.011

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 dependences are calculated using bivariate Student-t copula with the mean correlation of the assets belonging to the same sector.

in the case of the group Student-t and seems to be higher in the MGSt model. The lowest degree of freedom parameter in the dynamic group Student-t is in Pharmaceutical industries standing at 7.8. However, in the MGSt copula model, it is strongly negative skewed which results in a higher posterior estimation for the degrees of freedom. While the other groups that have low degrees of freedom such as Telecommunication, Insurance, and Finance reveal a slight skewness. Although the posterior variation 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 a symmetry in both upper tail and lower tail. We show different tail dependences in each group sector by calculating the mean correlation of all assets in the groups and derive the upper tail and lower tail of bivariate MGSt copulas. The strongest lower tail dependence is 0.42 from the Pharmaceutical sector despite of no upper tail dependence.

Figure 3 describes the posterior mean of the conditional correlation among time series using

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

	Oil	Food & Bev.	Pharma.	Plastics	Paper	Steel	Home App.
a	0.034	0.045	0.049	0.037	0.045	0.033	0.038
	(0.006)	(0.009)	(0.019)	(0.007)	(0.006)	(0.003)	(0.012)
b	0.996	0.991	0.976	0.989	0.985	0.995	0.957
	(0.002)	(0.005)	(0.023)	(0.005)	(0.004)	(0.001)	(0.036)
ν	12.971	14.069	35.857	15.368	15.248	17.982	17.012
	(1.389)	(1.554)	(5.343)	(1.504)	(1.062)	(1.104)	(1.036)
γ	-0.109	-0.237	-1.689	-0.202	-0.146	-0.173	-0.242
	(0.075)	(0.075)	(0.163)	(0.073)	(0.044)	(0.048)	(0.043)
f_c	[1.27,1.84]	[1.09,1.95]	[1.27,1.43]	[1.21,1.97]	[1.12,1.89]	[1.33,2.02]	[1.40,2.48]
	(0.174)	(0.127)	(0.099)	(0.098)	(0.092)	(0.132)	(0.062)
# firms	5	6	5	8	12	13	14
λ_L	0.047	0.056	0.423	0.051	0.035	0.039	0.086
λ_U	0.016	0.005	0.000	0.008	0.008	0.008	0.010
	Electronics	Transportation	Telecom.	Retail	Insurance	Finance	Services
a	0.132	0.042	0.065	0.023	0.054	0.052	0.034
	(0.017)	(0.007)	(0.017)	(0.004)	(0.007)	(0.009)	(0.003)
b	0.597	0.974	0.969	0.997	0.971	0.982	0.991
	(0.060)	(0.008)	(0.019)	(0.002)	(0.007)	(0.007)	(0.002)
ν	17.560	10.220	8.121	21.196	9.845	9.405	21.333
	(1.133)	(1.028)	(0.515)	(1.885)	(0.402)	(0.345)	(1.151)
γ	-0.195	-0.214	-0.075	-0.208	-0.109	-0.036	-0.318
	(0.043)	(0.089)	(0.035)	(0.055)	(0.037)	(0.029)	(0.040)
f_c	[1.38,2.33]	[0.99,2.10]	[0.91,2.08]	[1.09,1.85]	[1.29,2.46]	[1.44,2.52]	[1.30,2.23]
	(0.049)	(0.070)	(0.082)	(0.154)	(0.074)	(0.091)	(0.097)
# firms	12	12	7	8	14	17	17
λ_L	0.057	0.200	0.130	0.016	0.176	0.119	0.035
λ_U	0.010	0.022	0.054	0.003	0.060	0.082	0.002

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 dependence are calculated numerically using bivariate MGSt copula with the mean correlation of the assets belonging to the same sector (see Appendix A.4).

MGSt copula. This posterior mean is calculated as the average over iterations of the mean correlation between time series i and j as $\frac{1}{d(d-1)/2} \sum_{ij} \rho_{it} \rho_{jt}$ where $i, j = 1, \dots, d, i \neq j$. In the top left figure, we compare the mean correlation using block-equivalent, single group and multiple-group MGSt copula. The correlation path generated by block-equivalent model is close to that by multiple-group model, while the single group correlation path does not move much. Combined with the evidence of the model selection, we conclude that few parameters controlling for the dynamic behaviour of correlation is restrictive and the block equivalent mean correlation can approximate well the correlation structure in comparison to the multiple-group dynamic model. In the bottom left and bottom right figure, we illustrate the mean correlation together with the 95% credible interval of Student-t and MGSt copulas. They are almost the same in both point estimation as well as the confident brand. In the top right, we describe the distribution of mean correlation R_{ijt} across the 150 time series. The one factor dynamic MGSt copula could capture different behaviour among the cross correlation. As we can see, a common pattern is that the correlation increased over time in the end of 2011. This finding is similar to Creal and Tsay (2015) for stochastic copula model because of the financial crisis in $2010 - 2011$.

Figure 4 shows the posterior distribution of conditional variance and conditional correlation of several companies including PepsiCo, McDonalds, JP Morgan Chase, Morgan Stanley, Boeing, and Exxon Mobil using MGSt copula. The first two columns illustrate the conditional variance and the last column depicts the conditional correlation between the couple. As mentioned above, the $2010 - 2011$ period experienced a high volatility and a rise in correlation among all examples due to the financial crisis. The cross correlation also goes up recently but not for every series which might be due to the shocks to separated sectors rather than the whole economy.

6 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 that allows us to introduce an estimation strategy in a parallel setting. Furthermore, the factor loadings have been modelled as GAS processes which imposes a dynamic dependence structure in their densities. Using MGSt copulas,

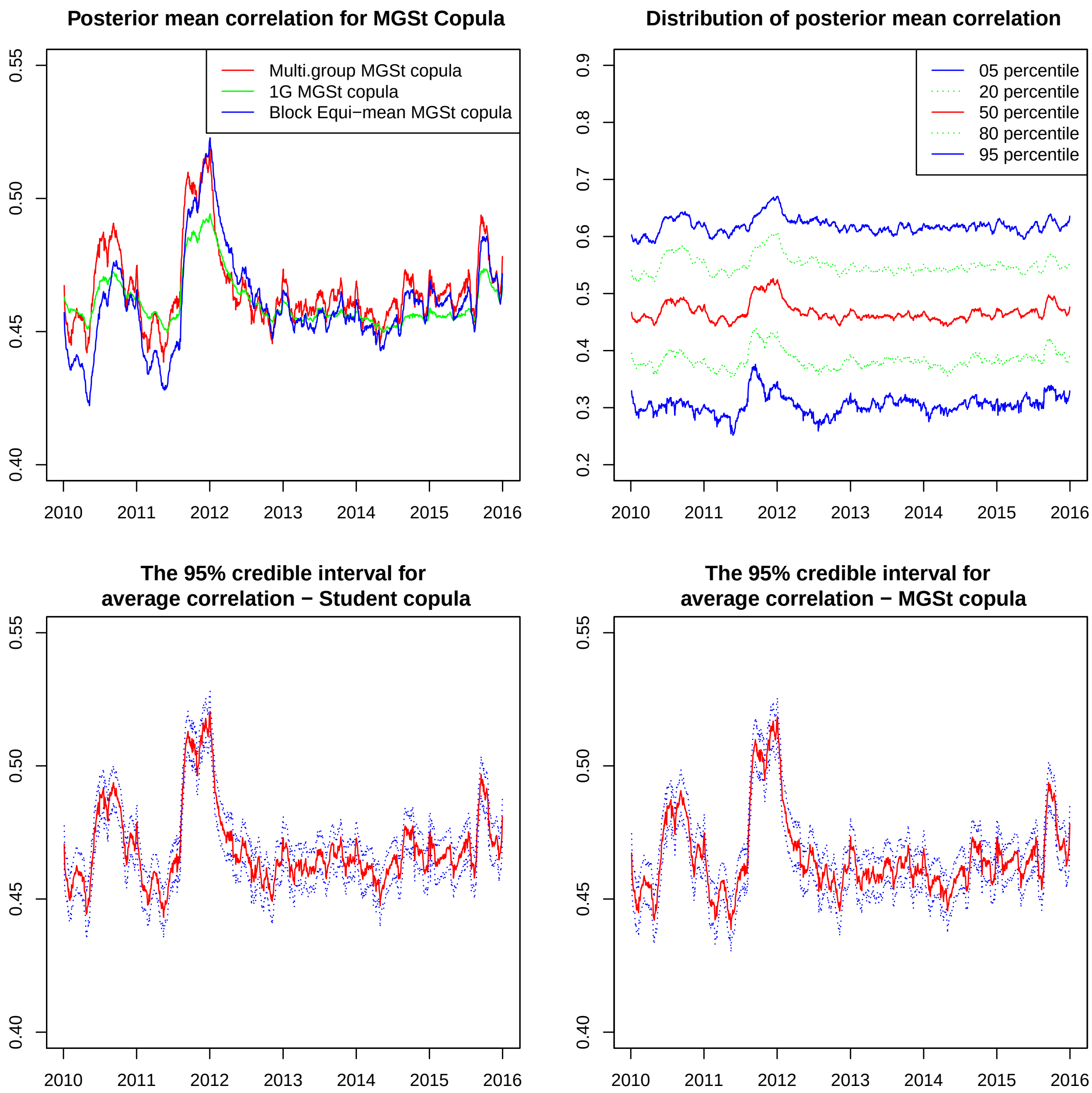

Figure 3: The distribution of posterior correlation among factor copula models

The mean correlation using block equi-mean correlation, single group and multiple-group MGSt copula shown in the top left. The top right figure shows the distribution of mean correlation across 150 time series. The two bottom figures illustrate the mean correlation together with the 95% credible interval of multiple-group Student and MGSt copula.

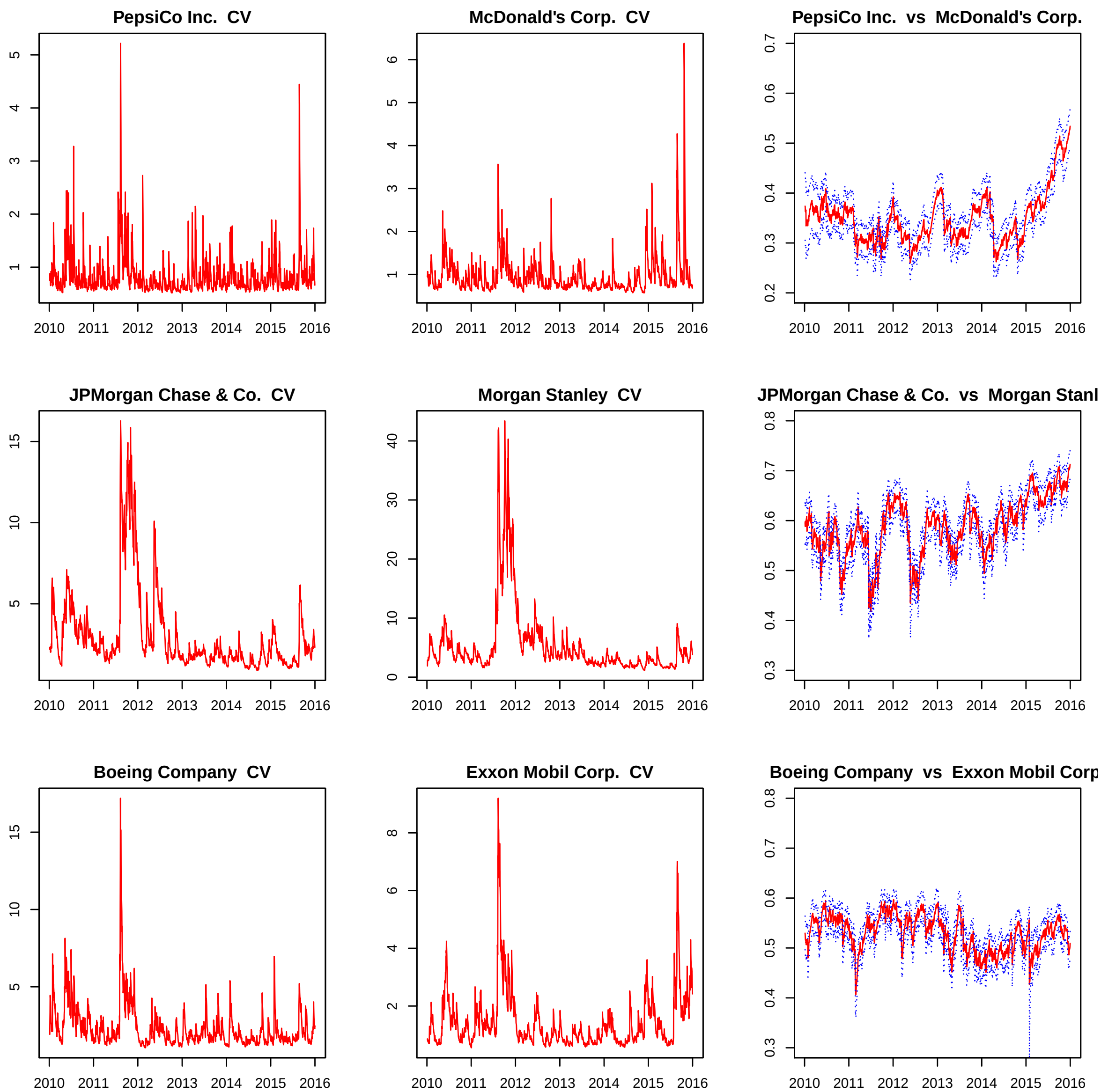

Figure 4: Posterior correlation among each time series

The first two columns describe the conditional variance and the last column depicts the conditional correlation together with the 95% credible interval using MGSt copula. First row: PepsiCo, McDonald, second row: JP Morgan Chase, Morgan Stanley, third row: Boeing, Exxon Mobil we obtain different types of tail and asymmetric dependence. The models are extendible since the number of parameters scales linearly with the dimension. As extension, more complex copula functions can be build based on the distribution of ζ_g. However, this also may require more computational cost to obtain the inverse cdf. Also, we might consider factor models using the family of Archimedean copulas, whose have only 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) suggest. One future direction could be to extend the proposed approach to dynamic multi factor models.

Appendix

A Score update for one factor copula model

A.1 Dynamic Gaussian one factor copula

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

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

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},

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

which leads to the expression given in (3).

A.2 Dynamic Student one factor copulas

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

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

where $\dot{X}_{it} = X_{it}/\sqrt{\zeta_t}$. Similarly, given $\{z_t, \zeta_t, f_t, \mathcal{F}_t, \theta\}$, the correlation ρ_{it} is known and $\dot{X}_{it}$ 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,

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

where $f_{St}(\cdot \mid \nu)$ denotes the standard Student-t with ν degrees of freedom. Thus, the equation for s_{it} remains,

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

which leads to the expression given in (4).

A.3 Dynamic hyperbolic skew Student 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:

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

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 $\dot{X}_{it}$ 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,

$$p\left(u_t \mid z_t, \zeta_t, f_t, \mathcal{F}_t, \theta\right) = \frac{\partial^d F\left(u_{1t}, \ldots, u_{dt} \mid z_t, \zeta_t, f_t, \mathcal{F}_t, \theta\right)}{\partial u_{1t} \ldots \partial u_{dt}} = \prod_{i=1}^{d} \frac{\phi\left(\frac{F_{GSt}^{-1}(u_{it}|\nu)-\gamma\zeta_t}{\sqrt{\zeta_t}} \mid \rho_{it} z_t, \sqrt{1-\rho_{it}^2}\right)}{f_{GSt}\left(F_{GSt}^{-1}\left(u_{it} \mid \nu, \gamma\right) \mid \nu, \gamma\right)\sqrt{\zeta_t}}, \tag{10}$$

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 (6).

A.4 Tail dependence for hyperbolic skew Student copula

Consider the bivariate MGSt copula. We derive the tail dependence of a pair of pseudo observables x_{igt} and x_{jtg} in a same group,g, from the Equation (7) as:

$$\begin{aligned} x_{igt} &= \gamma_g \zeta_g + \sqrt{\zeta_g}\left(\rho_{igt} z_t + \sqrt{1-\rho_{igt}^2}\epsilon_{igt}\right) \\ x_{jgt} &= \gamma_g \zeta_g + \sqrt{\zeta_g}\left(\rho_{jgt} z_t + \sqrt{1-\rho_{jgt}^2}\epsilon_{jgt}\right) \end{aligned} \tag{11}$$

Demarta and McNeil (2005) suggest that the joint quantile exceedance probability is obtained as the integral over the nuisance parameters (z_t, ζ_{gt}),

$$\begin{aligned} C(u,u|\nu_g,\gamma_g) &= \Pr(x_{igt} \leq F_{GSt}^{-1}(u), x_{jgt} \leq F_{GSt}^{-1}(u)) \\ &= \mathbb{E}\left(\Pr\left(\epsilon_{igt} \leq \frac{F_{GSt}^{-1}(u)-\gamma_g\zeta_{gt}}{\sqrt{\zeta_{gt}}\sqrt{1-\rho_{igt}^2}} - \frac{\rho_{igt} z_t}{\sqrt{1-\rho_{igt}^2}}, \epsilon_{jgt} \leq \frac{F_{GSt}^{-1}(u)-\gamma_g\zeta_{gt}}{\sqrt{\zeta_{gt}}\sqrt{1-\rho_{jgt}^2}} - \frac{\rho_{jgt} z_t}{\sqrt{1-\rho_{jgt}^2}}\right)\right) \\ &= \mathbb{E}\left(\Phi\left(\frac{F_{GSt}^{-1}(u)-\gamma_g\zeta_{gt}}{\sqrt{\zeta_{gt}}\sqrt{1-\rho_{igt}^2}} - \frac{\rho_{igt} z_t}{\sqrt{1-\rho_{igt}^2}}\right)\Phi\left(\frac{F_{GSt}^{-1}(u)-\gamma_g\zeta_{gt}}{\sqrt{\zeta_{gt}}\sqrt{1-\rho_{jgt}^2}} - \frac{\rho_{jgt} z_t}{\sqrt{1-\rho_{jgt}^2}}\right)\right) \end{aligned} \tag{12}$$

Then, we obtain $C(u, u|\nu_g, \gamma_g)$ as the numerical integral over $z_t \sim N(0,1)$ and $\zeta_{gt} \sim IG(\nu_g/2, \nu_g/2)$ As taking the equation to the limit, the tail dependence are,

$$\lambda_L = \lim_{u \to 0} \frac{C(u, u|\nu_g, \gamma_g)}{u}$$
$$\lambda_U = 2 + \lim_{u \to 0} \frac{C(1-u, 1-u|\nu_g, \gamma_g) - 1}{u}$$

Table (5) reports the tail dependence of bivariate MGSt copula at $u = 0.005$.

B Posterior inference

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

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

As the conditional posterior of z_t only depends on the pseudo observations at time t, we can make parallel inference for $t = 1, \ldots, T$. 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}.

C 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 interested parameters as the integral over the nuisance parameter space. The set parameters of interest is $\theta_{int} = \{a, b, f_c, \nu, \gamma\}$ and the nuisance parameters are $\theta_{nui} = \{z, \zeta\}$.

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

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.

C. Alexander and A. Chibumba. Multivariate orthogonal factor garch. *Working paper, Mathematics Department, University of Sussex.*

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.

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. Modelling the coherence in short-run nominal exchange rates: a multivariate generalized arch model. *The Review of Economics and Statistics*, 72(3):498–505, 1990.

T. Bollerslev, R. F. Engle, and J. M. Wooldridge. A capital asset pricing model with time-varying covariances. *The Journal of Political Economy*, 96(1):116–131, 1988.

L. Capiello, R. F. Engle, and K. Sheppard. Asymmetric dynamics in the correlations of global equity and bond returns. *Journal of Financial Econometrics*, 4(4):537–572, 2006.

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

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.

Z. Ding and R. F. Engle. Large scale conditional covariance matrix modeling, estimation and testing. *Academia Economic Papers*, 29:157–184, 2001.

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, V. Ng, and M. Rothschild. Asset pricing with a factor-arch covariance structure: Empirical estimates for treasury bills. *Journal of Econometrics*, 45(1–2):213–237, 1990.

R. F. Engle and K. F. Kroner. Multivariate simultaneous generalized arch. *Econometric Theory*, 11(1):122–150, 1995.

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

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

P. Galeano and M. C. Ausín. The gaussian mixture dynamic conditional correlation model: parameter estimation, value at risk calculation, and portfolio selection. *Journal of Business & Economic Statistics*, 20(3):351–362, 2010.

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.

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.

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.

M. Lanne and P. Saikkonen. A multivariate generalized orthogonal factor garch model. *Journal of Business & Economic Statistics*, 25(1):61–75, 2007.

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.

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. 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.

Stan Development Team. Rstan: the R interface to Stan, version 2.10.0, 2015. URL `http://mc-stan.org/`.

Y. K. Tse and A. K. C. Tsui. A multivariate generalized autoregressive conditional heteroscedasticity model with time-varying correlations. *Journal of Business & Economic Statistics*, 20(3): 351–362, 2002.

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

R. van der Wiede. Go-garch: a multivariate generalized orthogonal garch model. *Journal of Applied Econometrics*, 17(5):549–564, 2002.

I. Vrontos, P. Dellaportas, and D. Politis. A full-factor multivariate garch model. *The Econometrics Journal*, 6(2):312–334, 2003.