Modeling high dimensional time-varying dependence using dynamic D-vine models

Carlos Almeida[1], Claudia Czado[2], and Hans Manner*[3]

[1] *School Mathematics and Information Sciences, YachayTech University, Urcuquí, Ecuador.*
[2] *Department of Mathematical Statistics, Technische Univerität München, Germany.*
[3] *Institute of Econometrics and Statistics, University of Cologne, Germany*

Abstract

We consider the problem of modeling the dependence among many time series. We build high dimensional time-varying copula models by combining pair-copula constructions with stochastic autoregressive copula and generalized autoregressive score models to capture dependence that changes over time. We show how the estimation of this highly complex model can be broken down into the estimation of a sequence of bivariate models, which can be achieved by using the method of maximum likelihood. Further, by restricting the conditional dependence parameter on higher cascades of the pair copula construction to be constant, we can greatly reduce the number of parameters to be estimated without losing much flexibility. Applications to five MSCI stock market indices and to a large dataset of daily stock returns of all constituents of the Dax 30 illustrate the usefulness of the proposed model class in-sample and for density forecasting.

Keywords: Stock return dependence, time-varying copula, D-vines, efficient importance sampling, sequential estimation, generalized autoregressive score

*Corresponding author, email: manner@statistik.uni-koeln.de.

1 Introduction

The modeling of multivariate distributions is an important task for risk management and asset allocation problems. Since modeling the conditional mean of financial assets is rather difficult, if not impossible, much research has focused on modeling conditional volatilities and dependencies. The literature on multivariate GARCH (Bauwens et al. 2006) and stochastic volatility models (Harvey et al. 1994, Yu and Meyer 2006) offers many approaches to extend univariate volatility models to multivariate settings. However, usually the resulting multivariate model makes the assumption of (conditional) multivariate normality. Multivariate models based on copulas offer a popular alternative as non-elliptical multivariate distributions can be constructed in a tractable and flexible way. The advantage of using copulas to construct multivariate volatility models is that one is free with the choice of the marginal model, i.e. the univariate volatility model, and that it is possible to capture, possibly asymmetric, dependencies in the tails of the distributions. In particular, lower tail dependence often needs to be accounted for when measuring financial risks. Among many others, Patton (2009) or Cherubini et al. (2004), and references therein, give an overview of copula based models in financial applications.

Two major drawbacks of the early applications of copula based models are that most studies focus on bivariate copulas only, limiting the potential for real world applications, and that the dependence parameter is assumed to be time-constant. This is in contrast to the empirically observed time-varying correlations. Each of these issues individually has been addressed in the literature in recent years. Larger dimensional copulas other than Gaussian or Student copulas have become available through the introduction of hierarchical Archimedean copulas by Savu and Trede (2010) and Okhrin et al. (2013), factor copula models by Oh and Patton (2015), or the class of pair copula constructions by Aas et al. (2009). In particular the latter class, also called vine copula constructions, has become extremely popular because of its flexibility and because of the possibility of estimating the large number of parameters sequentially. Examples of financial applications of vine copula models are, e.g., Chollete et al. (2009) or Dißmann et al. (2013). A vine copula based model for both cross sectional and serial dependence is proposed in Brechmann and Czado (2015). Copulas with time-varying parameters have been introduced by Patton (2006) to model changing exchange rate dependencies. Since then a number of studies have proposed different ways to specify time-varying copulas. For example, Dias and Embrechts (2004) test for structural breaks in copula parameters, Giacomini et al. (2009) use a sequence of breakpoint tests to identify intervals of constant dependence, Garcia and Tsafack (2011) and Stöber and Czado (2014) use a regime-

switching model for changing dependencies, Hafner and Reznikova (2010) treat the copula parameter as a smooth function of time and estimate it by local maximum likelihood, whereas Hafner and Manner (2012) and Almeida and Czado (2012) propose a model where the copula parameter is a transformation of a latent Gaussian autoregressive process of order one. Finally, Creal et al. (2013) propose an observation driven autoregressive model in which the scaled score drives the dependence parameter. An overview and comparison of (bivariate) time-varying copula models is given in Manner and Reznikova (2012). Only very few papers allow for time-varying parameters in larger dimensions. Chollete et al. (2009) estimate a regime-switching vine copula, Heinen and Valdesogo (2009) allow the parameter of a vine copula to be driven by a variation of the DCC model by Engle (2002), So and Yeung (2014) introduce a vine copula model with dynamic dependence similar in spirit to a DCC model, and Creal and Tsay (2015) extend the factor copula model by Oh and Patton (2015) by allowing for stochastic factor loadings. Oh and Patton (2013), on the other hand, introduce time-variation into the factor copula model by specifying it as a generalized autoregressive score model.

The contribution of this paper is to extend the stochastic autoregressive copula (SCAR) model by Hafner and Manner (2012) and Almeida and Czado (2012) and the generalized autoregressive score (GAS) model of Creal et al. (2013) to practically relevant dimensions using D-vines. We discuss how the proposed model can be estimated sequentially using maximum likelihood estimation. We also address how a reasonable restrictions can be made on the model without restricting its flexibility too much. In our empirical study we consider the problem of modeling and forecasting the joint distribution of asset (market) returns using two datasets, namely 5 weekly MSCI index returns spanning a period of 40 years and daily returns of the constituents of the German DAX 30. The model fit and forecasting performance is compared to a Gaussian DCC copula model, a Student DCC copula and to a time-constant regular vine model.

The remainder of the paper is structured as follows. The next section introduces copulas in general, dynamic copula models, D-vines copulas and shows how they can be combined to obtain the flexible class of dynamic D-vine models. Section 3 treats the estimation of the proposed model, Section 4 contains the empirical application and Section 5 gives conclusions and outlines further research.

2 D-vine based dynamic copula models

We are interested in modeling the joint (conditional) distribution of a d-dimensional time series $\boldsymbol{y}_t = (y_{1,t}, ..., y_{d,t})$ for $t = 1, ..., T$. We assume that each variable $y_{i,t}$ for $i = 1, ..., d$ follows an ARMA(p,q)-GARCH(1,1) process, i.e.

$$y_{i,t} = \beta_{i,0} + \sum_{j=1}^{p} \beta_{i,j} y_{i,t-j} + \sum_{k=1}^{q} \delta_{i,k} \sigma_{i,t-k} \varepsilon_{i,t-k} + \sigma_{i,t} \varepsilon_{i,t} \tag{1}$$

with

$$\sigma_{i,t}^2 = \alpha_{i,0} + \alpha_{i,1} \varepsilon_{i,t-1}^2 + \gamma_i \sigma_{i,t}^2.$$

The usual stationarity conditions are assumed to hold. Denote the joint distribution of the standardized innovations $\varepsilon_{i,t}$ by $G(\varepsilon_{1,t}, ..., \varepsilon_{d,t})$ and let their marginal distributions be $F_1(\varepsilon_{1,t}), ..., F_d(\varepsilon_{d,t})$, respectively. Then by Sklar's theorem there exists a copula C such that

$$G(\varepsilon_{1,t}, ..., \varepsilon_{d,t}) = C(F_i(\varepsilon_{1,t}), ..., F_d(\varepsilon_{d,t})). \tag{2}$$

Since all the marginal behavior is captured by the (conditional) marginal distributions, the copula captures the complete contemporaneous dependence of the distribution. Let $u_{i,t} = F_i(\varepsilon_{i,t})$ be the innovations transformed to $U(0, 1)$ random variable and define $\boldsymbol{u}_t := (u_{1,t}, ..., u_{d,t})$.

In this paper we are interested in models that allow the copula distribution to be time-varying as well. In this case we assume that

$$\boldsymbol{u}_t \sim c(\boldsymbol{u}_t; \boldsymbol{\omega}, \mathcal{F}_{t-1}), \tag{3}$$

where c is the copula density, $\mathcal{F}_{t-1}$ is the information set available at time $t - 1$ and $\boldsymbol{\omega}$ is the vector of time independent parameters of the model.

In the next section we present two specifications to incorporate time-varying dependence in bivariate copula models, namely the parameter driven stochastic copula autoregressive (SCAR) model and the observation driven generalized autoregressive score (GAS) model. In Section 2.2 we introduce the notion of vines to construct large-dimensional copula, before we show how time-varying dependence can be incorporated yielding large dimensional models with dynamic dependence.

2.1 Dynamic Copula Models

For now, consider the bivariate time series process $(u_{i,t}, u_{j,t})$ for $t = 1, ..., T$. We assume that its distribution is given by

$$(u_{i,t}, u_{j,t}) \sim C(\cdot, \cdot; \theta_t^{ij}) \tag{4}$$

with $\theta_t^{ij} \in \Theta$ the time-varying parameter of the copula C. In order to be able to compare copula parameters that have different domains, the copula can equivalently be parameterized in terms of Kendall's $\tau \in (-1, 1)$. This follows from the fact that for all bivariate copulas we consider there exists a one-to-one relationship between the copula parameter and Kendall's τ, which we express by $\theta_t^{ij} = r(\tau_t^{ij})$. We assume that τ_t^{ij} is driven by the process $\lambda_t^{ij} \in (-\infty, \infty)$. Due to the fact that λ_t^{ij} takes values on the real line we apply the inverse Fisher transform to map it into $(-1, 1)$, the domain of τ_t^{ij}:

$$\tau_t^{ij} = \frac{exp(2\lambda_t^{ij}) - 1}{exp(2\lambda_t^{ij}) + 1} =: \psi(\lambda_t^{ij}). \tag{5}$$

The time-varying parameter can be specified in different ways, see Manner and Reznikova (2012) for a survey on different specifications. Here we consider two specifications that have been shown to perform well. The first approach is the stochastic copula autoregressive (SCAR) model proposed by Hafner and Manner (2012) and Almeida and Czado (2012). In this case λ_t^{ij} is assumed to be a latent Gaussian AR(1) process given by

$$\lambda_t^{ij} = \mu_{ij} + \phi_{ij}(\lambda_{t-1}^{ij} - \mu_{ij}) + \sigma_{ij} z_{ij,t}, \tag{6}$$

where $z_{ij,t}$ are independent standard normal innovations. We further assume $|\phi_{ij}| < 1$ for stationarity and $\sigma_{ij} > 0$ for identification. This is an example of a parameter driven model.

The second specification for the latent process is the generalized autoregressive score (GAS) model by Creal et al. (2013). This observation driven model also assumes an autoregressive structure for λ_t^{ij} , but also uses the weighted score of the underlying model to drive the latent process. The model of order one is given by

$$\lambda_t^{ij} = \omega_{ij} + \phi_{ij}\lambda_{t-1}^{ij} + \delta_{ij} s_{t-1}^{ij}, \tag{7}$$

where s_t^{ij} is the scaled score vector

$$s_t^{ij} = S_{ij,t} \nabla_{ij,t},$$

with

$$\nabla_{ij,t} = \frac{\partial \ln c(u_{i,t}, u_{j,t}; \boldsymbol{\omega}_{ij}, \mathcal{F}_{t-1})}{\partial \theta_t^{ij}}$$

is the score and $\boldsymbol{\omega}_{ij} = (\omega_{ij}, \phi_{ij}, \delta_{ij})$. The scaling matrix $S_{ij,t}$ is the square root matrix of the inverse of the information matrix defined as

$$S_{ij,t} = \mathcal{J}_{t|t-1} \text{ with } \mathcal{J}'_{t|t-1}\mathcal{J}_{t|t-1} = \mathcal{I}^{-1}_{t|t-1},$$

where $\mathcal{I}_{t|t-1} = \mathrm{E}_{t-1}[\bigtriangledown_{ij,t}\bigtriangledown'_{ij,t}]$ is the information matrix. Details and properties can be found in Creal et al. (2013). Stationarity conditions are studies in Blasques et al. (2014). Blasques et al. (2015) show optimality properties of GAS models, whereas Koopman et al. (2015) compare the forecasting performance of a wide range of parameter-driven and observation-driven models and conclude that both perform equally well.

2.2 D-vine Distributions and Copulas

While copulas are recognized as a very powerful tool to construct multivariate distributions, in the past only the class of bivariate copulas (e.g. Joe 1997 and Nelsen 2006) was flexible enough to accommodate asymmetric and/or tail dependence without placing unrealistic restrictions on the dependence structure. Recently pair copula constructions (PCC) are found to be very useful to construct flexible multivariate copulas. Here a multivariate copula is built up with bivariate copula terms modeling unconditional and conditional dependencies. The first such construction was proposed in Joe (1996). It was subsequently significantly extended to more general settings in Bedford and Cooke (2002), Bedford and Cooke (2001) and Kurowicka and Cooke (2006). They called the resulting distributions regular (R) vines and explored them for the case of Gaussian pair copulas. The backbone is a graphical representation in form of a sequence of linked trees identifying the indices which make up the multivariate copula. In particular, they proved that the specification of the corresponding pair copula densities make up a valid multivariate copula density. Further properties, estimation, model selection methods and their use in complex modeling situations can be found in Kurowicka and Joe (2011), Czado et al. (2013) and Joe (2014).

Aas et al. (2009) recognized the potential of this construction for statistical inference and developed a sequential estimation (SE) procedure, which can be used as starting values for maximum likelihood estimation (MLE). Bedford and Cooke (2002) identified two interesting subclasses of regular vines called D-vines and canonical (C)-vines. In the case of D-vines the sequence of vine trees consist of pair trees, while for C-vines they are starlike with a central node. This shows that C-vines are more useful for data situations where the importance of the variables can be ordered. This is not the case for the application we will present later; therefore we concentrate on D-vines. However, we would like to note that multivariate SCAR models can also be constructed based on C-vines and more generally on R-vines.

Notably, C- and D-vines can be introduced from first principles (e.g. Czado 2010). For this let $(X_1, ..., X_d)$ be a set of variables with joint distribution F and density f, respectively.

Consider the recursive decomposition

$$f(x_1,\ldots,x_d) \;=\; \prod_{k=2}^{d} f(x_k|x_1,\ldots,x_{k-1}) \times f(x_1). \tag{8}$$

Here $F(\cdot|\cdot)$ and later $f(\cdot|\cdot)$ denote conditional cdf's and densities, respectively. As a second ingredient we utilize Sklar's theorem for dimension $d = 2$ to express the conditional density of X_1 given $X_2 = x_2$ as

$$f(x_1|x_2) = c_{12}(F_1(x_1), F_2(x_2)) \times f_1(x_1), \tag{9}$$

where c_{12} denotes an arbitrary bivariate copula density. For distinct indices $i, j, i_1, \cdots, i_k$ with $i < j$ and $i_1 < \cdots < i_k$ we now introduce the abbreviation

$$c_{i,j|D} := c_{i,j|D}(F(x_i|\mathbf{x}_D), F(x_j|\mathbf{x}_D)), \tag{10}$$

where $D := \{i_1, \cdots, i_k\}$ and $\mathbf{x}_D := (x_{i_1}, \ldots, x_{i_k})$. Using (9) for the conditional distribution of (X_1, X_k) given $X_2 = x_2, \ldots X_{k-1} = x_{k-1}$ we can express $f(x_k|x_1, \cdots, x_{k-1})$ recursively as

$$\begin{aligned} f(x_k|x_1,\ldots,x_{k-1}) &= c_{1,k|2:k-1} \times f(x_k|x_2,\ldots,x_{k-1}) \\ &= [\prod_{s=1}^{k-2} c_{s,k|s+1:k-1}] \times c_{(k-1),k} \times f_k(x_k), \end{aligned} \tag{11}$$

where $r : s := (r, r+1, \ldots, s)$ for integers r and s with $r < s$. Using (11) in (8) and $s = i, k = i + j$ it follows that

$$f(x_1,\ldots,x_d) \;=\; [\prod_{j=1}^{d-1}\prod_{i=1}^{d-j} c_{i,i+j|i+1:i+j-1}] \cdot [\prod_{k=1}^{d} f_k(x_k)] \tag{12}$$

If the marginal distribution of X_k are uniform for all $k = 1, \cdots, d$, then we call the corresponding density in (12) a D-vine copula density and the corresponding distribution function a D-vine copula.

For illustration we consider a five dimensional D-vine, its density then given by

$$\begin{aligned} f(x_1,\cdots,x_5) &= [\prod_{k=1}^{5} f_k(x_k)] \cdot c_{12} \cdot c_{23} \cdot c_{34} \\ &\times \; c_{45} \cdot c_{13|2} \cdot c_{24|3} \cdot c_{35|4} \cdot c_{14|23} \cdot c_{25|34} \cdot c_{15|234}, \end{aligned} \tag{13}$$

with the corresponding vine tree representation identifying the utilized indices given in Figure 1. In particular the indices in Tree T_1 indicate the unconditional pair copulas, while Trees T_2

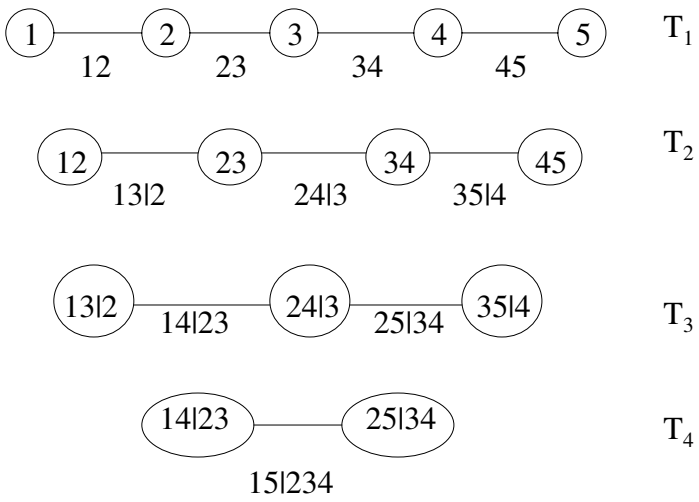

Figure 1: A D-vine tree representation for $d = 5$.

to T_4 correspond to conditional pair copulas, where the set of conditioning variables has size 1 to 3, respectively.

If $c_{i,i+j|i+1:i+j-1}$ models the dependence between the rv's $F(X_i|\mathbf{x}_{i+1:i+j-1})$ and $F(X_{i+j}|\mathbf{x}_{i+1:i+j-1})$ we implicitly assume that the copula density $c_{i,i+j|i+1:i+j-1}(\cdot,\cdot)$ does not depend on the conditioning variables $\mathbf{x}_{i+1:i+j-1}$ other than through the arguments $F(X_i|\mathbf{x}_{i+1:i+j-1})$ and $F(X_{i+j}|\mathbf{x}_{i+1:i+j-1})$. This is a common assumption and Haff et al. (2010) call this a simplified vine. They showed that this restriction is not severe by examining several examples. Stöber et al. (2013) investigated which multivariate copula families can be written as a simplified vine and gave further examples in which it is difficult to detect a violation of the simplified vine assumption. Acar et al. (2012) provided an estimation method for non simplified vines, which however is only operational in three dimensions.

In the D-vine representation given in (12) we also need a fast recursive way to compute conditional cdf's which enter as arguments. For this Joe (1996) showed that for $i \in D$ and $D_{-i} := D \setminus i$

$$F(x_j|\mathbf{x}_D) = \frac{\partial C_{i,j|D_{-i}}(F(x_i|\mathbf{x}_{D_{-i}}), F(x_i|\mathbf{x}_{D_{-i}}))}{\partial F(x_i|\mathbf{x}_{D_{-i}})}. \tag{14}$$

For the special case of $D = \{i\}$ it follows that

$$F(x_j|x_i) = \frac{\partial C_{i,j}(F(x_i), F(x_j))}{\partial F(x_i)}.$$

In the case of uniform margins $u_j = F_j(x_j)$, for a parametric copula cdf $C_{ij}(u_i, u_j) = C_{ij}(u_i, u_j; \boldsymbol{\theta}_{ij})$ this further simplifies to

$$h(u_j|u_i, \boldsymbol{\theta}_{ij}) := \frac{\partial\, C_{i,j}(u_i, u_j; \boldsymbol{\theta}_{ij})}{\partial u_i}. \tag{15}$$

With this notation we can express $F(x_j|\mathbf{x}_D)$ as

$$F(x_j|\mathbf{x}_D) \;\; = \;\; h(F(x_j|\mathbf{x}_{D_{-i}})|F(x_i|\mathbf{x}_{D_{-i}}), \boldsymbol{\theta}_{ij|D_{-i}}).$$

This allows the recursive determination of the likelihood corresponding to (12). Furthermore, the inverse of the h-functions is used to facilitate sampling from D- and C-vines (see for example Aas et al. 2009 and Kurowicka and Cooke 2007). They are also used for sampling from the more general R-vine model (see Stöber and Czado 2011). These sampling methods in addition to estimation methods are implemented in the R packages CDVine and VineCopula.

A challenging question is how well D-vine models can approximate a true unknown copula and how flexible these models are compared to, e.g., the popular class of t-copulas. There are no theoretical results on how large the class of D-vine copulas are and this poses an interesting question for future research. However it is quite flexible since all pair copulas can be chosen arbitrary in addition to their parameters. Further the class of Gaussian copulas or multivariate t-copula with a single degree of freedom can be represented as a D-vine. In this case all pair copulas are bivariate Gaussian copulas or bivariate t-copulas, respectively. The association parameters are specified by partial correlations and the degree-of-freedom parameter of a multivariate t-copula with a single degree of freedom determines the degree of freedom of the bivariate t-copulas. More details can be found in Stöber et al. (2013). There are several numerical and empirical studies comparing vine copulas to competing models. In particular Fischer et al. (2009) was the first to show the superiority of D-vine copulas over benchmark copula models used in the industry. Other empirical studies showing good properties for D-vines are de Melo Mendes et al. (2010), Weiß and Supper (2013) and Righi and Ceretta (2015).

2.3 D-vine based multivariate dynamic copula models

We now combine bivariate dynamic copula models and D-vines to formulate a multivariate D-vine time-varying copula model. For this we use a bivariate dynamic copula model as the pair copula model in a D-vine copula. This gives rise to the following definition of a dynamic D-vine copula density

$$c(u_1, \cdots, u_d; \boldsymbol{\theta}_t) := \prod_{j=1}^{d-1} \prod_{i=1}^{d-j} c_{l(i,j)}(F(u_i|\boldsymbol{u}_{i+1:i+j-1}; \boldsymbol{\theta}_t^{l(i,j)}), F(u_{i+j}|\boldsymbol{u}_{i+1:i+j-1}); \boldsymbol{\theta}_t^{l(i,j)}), \tag{16}$$

where $l(i,j) := i, i+j | i+1 : i+j-1$ and $\boldsymbol{\theta}_t := \{\boldsymbol{\theta}_t^{l(i,j)}; j = 1, \cdots, d-1, i = 1, \cdots, d-j\}$ is the time-varying copula parameter vector. Here $c_{l(i,j)}(\cdot, \cdot; \boldsymbol{\theta}_t^{l(i,j)})$ is the bivariate copula density corresponding to the bivariate dynamic copula given in (4), where $\boldsymbol{\theta}_t^{l(i,j)}$ satisfies

$$\boldsymbol{\theta}_t^{l(i,j)} = r(\tau_t^{l(i,j)}) = r(\psi(\lambda_t^{l(i,j)})) \tag{17}$$

for the latent process $\lambda_t^{l(i,j)}$ defined either by equation (6) or (7). The bivariate copula family corresponding to $l(i,j)$ can be chosen arbitrarily and independently of any other index $l(r,s)$.

The copula in (16) can be used in (2) to specify the joint distribution of the innovations in (1).

3 Parameter estimation

We are interested in estimating the parameters of both the marginal models and the stochastic copula models. The joint density of our model is given by the product of the marginal and the copula densities

$$g(\varepsilon_{1,t}, ..., \varepsilon_{d,t}) = c(F_i(\varepsilon_{1,t}), ..., F_d(\varepsilon_{d,t})) \cdot f_i(\varepsilon_{1,t}) \cdot ... \cdot f_d(\varepsilon_{d,t}),$$

where g, c and f denote the densities of the joint distribution, the copula and the marginal distributions, respectively. Taking logarithms, we can see that the joint log-likelihood is the sum the marginal and the copula log-likelihood function. For estimation we ultilize a two-step approach common in copula based models. In this approach first the marginal parameters are estimated separately and standardized residuals are formed. These are transformed using either a parametric (see Joe 2005) or nonparametric probability integral transformation (see Genest et al. 1995) to get a sample from a multivariate copula. These transformations do not change the dependence structure among the standardized residuals. This approach allows us to perform the estimation of the marginal and copula parameters separately. If the marginal models are chosen carefully, as we will do, then a parametric probability transformation is a good approximation to the true copula data $u_{i,t} = F_i(\varepsilon_{i,t})$. Problems only occur if the marginal models are grossly misspecified (see Kim et al. 2007).

Furthermore, we saw above that the density of a D-vine copula is the product of bivariate (conditional) copulas. Therefore, instead of estimating all copula parameters of our model in one step, which is computationally infeasible due to the large number of parameters, we are able to estimate the copula parameters sequentially. The price to pay for the possibility to sequentially estimate the parameters is a small loss in statistical efficiency and intractable forms for the standard errors of the parameter estimates.

For the bivariate model the log-likelihood for observation t is given by

$$LL(\boldsymbol{\omega}_{ij}; u_{i,t}, u_{j,t}) = \ln c(u_{i,t}, u_{j,t}; \boldsymbol{\omega}_{ij}, \mathcal{F}_{t-1}) = \ln c(u_{i,t}, u_{j,t}; \theta_t^{ij}). \tag{18}$$

For the GAS model θ_t^{ij} can be computed for a given value of $\boldsymbol{\omega}_{ij}$ using the recursion (7) and therefore the estimation is straightforward. For the SCAR model, on the other hand, θ_t^{ij} cannot be observed and therefore needs to be integrated out of the likelihood function.

In Section 3.1 the estimation of bivariate SCAR copula models by simulated maximum likelihood (SML) using efficient importance sampling (EIS) is reviewed, Section 3.2 presents the sequential estimation of vine copula models and in Section 3.3 we discuss how the sequential estimation of dynamic D-vine copula models can be achieved.

3.1 Estimation of bivariate SCAR copula models

For the moment, we are interested in estimating the copula parameter vector $\boldsymbol{\omega} := (\mu, \phi, \sigma)$. For notational convenience we decided to drop the indices i and j whenever no ambiguity arises. Denote $\boldsymbol{u}_i = \{u_{i,t}\}_{t=1}^T$, $\boldsymbol{u}_j = \{u_{j,t}\}_{t=1}^T$ and $\boldsymbol{\Lambda} = \{\lambda_t\}_{t=1}^T$ and let $f(\boldsymbol{u}_i, \boldsymbol{u}_j, \boldsymbol{\Lambda}; \boldsymbol{\omega})$ be the joint density of the observable variables $(\boldsymbol{u}_i, \boldsymbol{u}_j)$ and the latent process $\boldsymbol{\Lambda}$. Then the likelihood function of the parameter vector $\boldsymbol{\omega}$ can be obtained by integrating the latent process $\boldsymbol{\Lambda}$ out of the joint likelihood,

$$L(\boldsymbol{\omega}; \boldsymbol{u}_i, \boldsymbol{u}_j) = \int f(\boldsymbol{u}_i, \boldsymbol{u}_j, \boldsymbol{\Lambda}; \boldsymbol{\omega}) d\boldsymbol{\Lambda}. \tag{19}$$

We can alternatively write this as a product of conditional densities

$$L(\boldsymbol{\omega}; \boldsymbol{u}_i, \boldsymbol{u}_j) = \int \prod_{t=1}^{T} f(u_{i,t}, u_{j,t}, \lambda_t | \lambda_{t-1}, \boldsymbol{\omega}) d\boldsymbol{\Lambda}. \tag{20}$$

This is a T-dimensional integral that cannot be solved by analytical or numerical means. It can, however, be solved efficiently by Monte Carlo integration using a technique called efficient importance sampling introduced by Richard and Zhang (2007). The idea is to make use of an auxiliary sampler $m(\lambda_t; \lambda_{t-1}, \boldsymbol{a}_t)$ that utilizes the information on the latent process contained in the observable data. Note that it depends on the auxiliary parameter vector $\boldsymbol{a}_t = (a_{1,t}, a_{2,t})$. Multiplying and dividing by $m(\cdot)$, the likelihood can then be rewritten as

$$L(\boldsymbol{\omega}; \boldsymbol{u}_i, \boldsymbol{u}_j) = \int \prod_{t=1}^{T} \left[\frac{f(u_{i,t}, u_{j,t}, \lambda_t | \lambda_{t-1}, \boldsymbol{\omega})}{m(\lambda_t; \lambda_{t-1}, \boldsymbol{a}_t)} \right] \prod_{t=1}^{T} m(\lambda_t; \lambda_{t-1}, \boldsymbol{a}_t) d\boldsymbol{\Lambda}. \tag{21}$$

Drawing N trajectories $\tilde{\boldsymbol{\Lambda}}^{(i)}$ from the importance sampler[1] the likelihood can be estimated by

$$\tilde{L}(\boldsymbol{\omega}; \boldsymbol{u}_i, \boldsymbol{u}_j) = \frac{1}{N} \sum_{s=1}^{N} \left(\prod_{t=1}^{T} \left[\frac{f(u_{i,t}, u_{j,t}, \tilde{\lambda}_t^{(s)} | \tilde{\lambda}_{t-1}^{(s)}, \boldsymbol{\omega})}{m(\tilde{\lambda}_t^{(s)}; \tilde{\lambda}_{t-1}^{(s)}, \boldsymbol{a}_t)} \right] \right). \tag{22}$$

This leaves the exact choice of the importance sampler $m(\cdot)$ to be determined, which ideally should provide a good match between the numerator and the denominator of (22) in order to minimize the variance of the likelihood function. It is chosen to be

$$m(\lambda_t; \lambda_{t-1}, \boldsymbol{a}_t) = \frac{k(\lambda_t, \lambda_{t-1}; \boldsymbol{a}_t)}{\chi(\lambda_{t-1}; \boldsymbol{a}_t)}, \tag{23}$$

where

$$\chi(\lambda_{t-1}; \boldsymbol{a}_t) = \int k(\lambda_t, \lambda_{t-1}; \boldsymbol{a}_t) d\lambda_t$$

is the normalizing constant of the auxiliary density kernel $k(\cdot)$. Furthermore, the choice

$$k(\lambda_t, \lambda_{t-1}; \boldsymbol{a}_t) = p(\lambda_t | \lambda_{t-1}, \boldsymbol{\omega}) \zeta(\lambda_t, \boldsymbol{a}_t),$$

with $p(\lambda_t | \lambda_{t-1}, \boldsymbol{\omega})$ the conditional density of λ_t given λ_{t-1} and $\zeta(\lambda_t, \boldsymbol{a}_t) = \exp(a_{1,t}\lambda_t + a_{2,t}\lambda_t^2)$ turns out to simplify the problem considerably. Noting that $f(u_{i,t}, u_{j,t}, \lambda_t | \lambda_{t-1}, \boldsymbol{\omega}) = c(u_{i,t}, u_{j,t}; \lambda_t) p(\lambda_t | \lambda_{t-1}, \boldsymbol{\omega})$, the likelihood expression (21) can be rewritten as

$$L(\boldsymbol{\omega}; \boldsymbol{u}_i, \boldsymbol{u}_j) = \int \prod_{t=1}^{T} \left[\frac{c(u_{i,t}, u_{j,t}; \lambda_t) \chi(\lambda_t; \boldsymbol{a}_{t+1})}{\exp(a_{1,t}\lambda_t + a_{2,t}\lambda_t^2)} \right] \prod_{t=1}^{T} m(\lambda_t; \lambda_{t-1}, \boldsymbol{a}_t) d\boldsymbol{\Lambda}, \tag{24}$$

where we have used the fact that $\chi(\cdot)$ can be transferred back one period, because it does not depend on λ_t. Defining $\chi(\lambda_T; \boldsymbol{a}_{T+1}) \equiv 1$ and given a set of trajectories $\tilde{\Lambda}^{(s)}$ for $s = 1, \ldots, N$, minimizing the sampling variance of the quotient in the likelihood function is equivalent to solving the following linear least squares problem for each period $t = T, \ldots, 1$,

$$\log c(u_{i,t}, u_{j,t}; \tilde{\lambda}_t^{(s)}) + \log \chi(\lambda_t; \boldsymbol{a}_{t+1}) = c_t + a_{1,t}\tilde{\lambda}_t^{(s)} + a_{2,t}[\tilde{\lambda}_t^{(s)}]^2 + \eta_t^{(s)}. \tag{25}$$

This problem can be solved by OLS with c_t the regression intercept and $\eta_t^{(s)}$ the error term. Then the procedure works as follows. First, draw N trajectories $\tilde{\Lambda}^{(s)}$ from $p(\lambda_t | \lambda_{t-1}, \omega)$ and estimate the auxiliary parameters $\hat{\boldsymbol{a}}_t$ for $t = T, \ldots, 1$ by solving (25). Next, draw N trajectories $\tilde{\Lambda}^{(s)}$ from the importance sampler $m(\lambda_t; \lambda_{t-1}, \hat{\boldsymbol{a}}_t)$ and re-estimate the auxiliary parameters $\{\hat{\boldsymbol{a}}_t\}_{t=1}^{T}$. Iterate this procedure until convergence of $\{\hat{\boldsymbol{a}}_t\}_{t=1}^{T}$ and use N draws from the importance sampler to estimate the likelihood function (22). This likelihood function can then be

[1] A good choice for N is about 100, in which case the simulation error is negligible and the computational costs are still acceptable.

maximized to obtain parameter estimates $\hat{\boldsymbol{\omega}}$. Note that throughout the same random numbers have to be used when simulating the likelihood function in order to ensure its smoothness and, consequently, convergence of $\{\hat{\boldsymbol{a}}_t\}_{t=1}^T$.

Although the parameter vector $\boldsymbol{\omega}$ driving the latent process is of some interest, ultimately one wishes to get estimates of the latent process $\boldsymbol{\Lambda}$ and transformations thereof. In particular, we are interested in estimating $\tau_t = \psi(\lambda_t)$ for $t = 1, \ldots, T$, where $\psi(\cdot)$ denotes the inverse Fisher transform given in (5). Smoothed estimates of $\psi(\lambda_t)$ given the entire history of the observable information $\boldsymbol{u}_i$ and $\boldsymbol{u}_j$ can be computed as

$$E[\psi(\lambda_t)|\boldsymbol{u}_i, \boldsymbol{u}_j] = \frac{\int \psi(\lambda_t) f(\boldsymbol{u}_i, \boldsymbol{u}_j, \boldsymbol{\Lambda}; \boldsymbol{\omega}) d\boldsymbol{\Lambda}}{\int f(\boldsymbol{u}_i, \boldsymbol{u}_j, \boldsymbol{\Lambda}; \omega) d\boldsymbol{\Lambda}}. \tag{26}$$

Note that the denominator in (26) corresponds to the likelihood function and both integrals can be estimated using draws from the importance sampler $m(\lambda_t; \lambda_{t-1}, \hat{\boldsymbol{a}}_t)$. Filtered estimates of $\psi(\lambda_t)$ given information until time $t-1$ can be computed in a similar way and details are given in Liesenfeld and Richard (2003).

3.2 Sequential estimation of static D-vine copula parameters

The form of the D-vine density given in (12) allows for a sequential parameter estimation approach starting from the first tree until the last tree. This was first proposed by Aas et al. (2009) for D-vines and shown in detail for C-vines in Czado et al. (2012). First estimate the parameters corresponding to the pair-copulas in the first tree using any method you prefer. For the copula parameters identified in the second tree, one first has to transform the data with the h function in (15) required for the appropriate conditional cdf using estimated parameters to determine pseudo realizations needed in the second tree. Using these pseudo observations the parameters in the second tree are estimated, the pseudo data is again transformed using the h function and so on.

For example we want to estimate the parameters of copula $c_{13|2}$. First transform the observations $\{u_{1,t}, u_{2,t}, u_{3,t}, t = 1, \cdots, n\}$ to $u_{1|2,t} := h(u_{1,t}|u_{2,t}, \hat{\theta}_{12})$ and $u_{3|2,t} := h(u_{3,t}|u_{2,t}, \hat{\theta}_{23})$, where $\hat{\theta}_{12}$ and $\hat{\theta}_{23}$ are the estimated parameters in the first tree. Now estimate $\theta_{13|2}$ based on $\{u_{1|2,t}, u_{3|2,t}; t = 1, \cdots, n\}$. Continue sequentially until all copula parameters of all trees are estimated. For trees T_i with $i \geq 2$ recursive applications of the h functions is needed. Asymptotic normality of the SE has been established by Haff (2013). However, the asymptotic covariance of the parameter estimates is very complex and one has to resort to bootstrapping to estimate the standard errors. SE is often used in large dimensional problems, e.g. Mendes et al. (2011), Heinen and Valdesogo (2009), Brechmann et al. (2012) and

Brechmann and Czado (2013). A joint MLE of all parameters in (12) requires high dimensional optimization. Therefore SE's are often used as starting values as, e.g., in Aas et al. (2009) and Czado et al. (2012).[2]

A final issue to be discussed is model selection for D-vine copula models. For non Gaussian pair copulas, permutations of the ordering of the variables give different D-vine copulas. In fact, there are $d!/2$ different D-vine copulas when a common bivariate copula is used as pair copula type. In Section 4 we will describe a simple method for determining the order of the variables in a D-vine, while Dißmann et al. (2013) provide more sophisticated methods for D-vines involving finding a maximal spanning tree corresponding to a traveling salesman problem. Often the bivariate Clayton, Gumbel, Gauss, t, Joe and Frank copula families are utilized as choices for pair copula terms. However, in this study we restrict the attention to the Gauss, Gumbel, Clayton and rotated versions thereof.

3.3 Sequential estimation of dynamic D-vine models

In principle, estimation of the dynamic D-vine model given in (16) works the same way as for static D-vine copulas. For the GAS model this is straightforward to do, because the (estimated) time-varying dependence can be computed and inserted into the corresponding h functions to compute the pseudo observations:

$$u_{j|i,t} = h(u_{j,t}|u_{i,t}, \hat{\theta}_t^{ij}). \tag{27}$$

For the parameter driven model, on the other hand, there are two important differences. First of all, given that the bivariate SCAR models in the first tree have been estimated, it is not possible to apply the h function given in (15) directly to obtain the pseudo observations that are needed to obtain the parameters on the second tree. The reason is that one only obtains parameter estimates of the hyper-parameters (μ, ϕ, σ), but not of the latent (time-varying) copula parameters θ_t. We do, however, have N simulated trajectories $\tilde{\theta}_t^{(s)}$ from the importance sampler. With these we can calculate the pseudo observations by

$$u_{j|i,t} = \frac{1}{N} \sum_{s=1}^{N} h(u_{j,t}|u_{i,t}, \tilde{\theta}_t^{(s)}), \tag{28}$$

[2]From a practical perspective, the recent R package CDVine of Schepsmeier and Brechmann (2013) provides easy to use random number generation, and both SE and MLE fitting algorithms for C- and D-vines.

where we suppress the dependence of $\tilde{\theta}$ on the variable indices i and j for notational reasons.[3] The second difference is that one-step estimation by MLE is computationally not feasible for the SCAR D-vine model, because each bivariate likelihood function needs to be computed by simulation. A practical problem that results from this is that standard errors cannot be obtained directly using a numerical estimate of the information matrix. Unfortunately, the typical solution of obtaining standard errors by the bootstrap is not available for two reasons. First, the computational burden is much too high due to the fact that the likelihood function has to be obtained by simulation. Second, an i.d.d. bootstrap cannot be obtained as the data is assumed to be dependent. Which the latter problem can in principle be dealt with, either using a block bootstrap or by using a parametric bootstrap relying on the dynamic copula model, the first one would be require immense computations. Therefore, in this paper we do not report estimated standard errors and leave their estimation for future research.

For a d dimensional dataset using the dynamic D-vine copula one has to estimate $3d(d+1)/2$ parameters. Fortunately, we can reduce the number of parameters to be estimated by placing a number of restrictions. Similar to tail properties of D-vines studied in Joe et al. (2010) the choice of time-varying pair copulas in the first tree propagates to the whole distribution, in particular all pairs of variables have an induced time-varying Kendall's tau. We expect estimation errors to increase for parameters as the corresponding tree increases because of the sequential nature of the estimation procedure. This is likely to be much more severe for dynamic models than for static ones. Therefore we allow for dynamic D-vine copula models where the pair copulas are time-varying only in lower trees, while the pair copulas are time-constant for higher trees.

A second useful restriction is to allow for the possibility of truncating the D-vine copula, which means that we set all pair copulas beyond a certain tree equal to the independence copula. This is empirically justified, since the dependence in the lower trees seems to capture most of the overall dependence in the data and the conditional dependence in higher trees is hardly visible. Note that this also allows the estimation of our model in arbitrarily large dimensions, as we will only need to estimate (bivariate) models up to a certain dimension and can truncate the model thereafter. For static models this has been followed by Brechmann et al. (2012) and includes tests at which level to truncate.

In order to decide which copula family to use and whether to use time-varying, time-constant or independence copulas at certain levels we compare the Bayesian Information Cri-

[3] Alternatively, we could calculate the pseudo realizations using the smoothed estimates of the latent dependence parameter using (26). However, averaging over the nonlinear transformation h seems more reasonable than applying the transformation to the (weighted) average.

terion (BIC) for all competing models. We decided for this information criterion, because it favors parsimonious models. Given the high flexibility of the dynamic D-vine model and the difficulty to estimate the parameters at higher level, we believe that parsimony is crucial.

3.4 Computational issues with SCAR models

Estimation of a bivariate SCAR model by simulated maximum likelihood programmed in MATLAB can take up to several minutes on a normal computer. As we consider an application with 29 variables this is much too slow. However, the problem at hand offers itself to parallel computing. On each tree one has to estimate a large number of bivariate models independently of each other. Therefore, given a sufficient number of processing cores we can estimate all models on one tree at the same time and proceed to the next tree once all models are estimated. Depending on the dimension of the problem, this can lead to immense increases in computing speed. The most demanding computational task, the estimation of the log-likelihood function by EIS, is implemented in C++, which resulted in our code being about 20-30 times faster compared to R code. The maximization of the likelihood and the parallel computation within levels is implemented in R (version 2.12.1) by using the `optim` function and the multicore library.

4 Application

In this section illustrate the empirical application of the dynamic D-vine models. We consider two datasets. The first one are weekly returns on 5 MSCI stock market indices over a period of more than 40 years (Section 4.1), whereas in the second application we model the returns of daily stock prices for 29 German stocks using about 6 years of data.

The first step in the analysis is the estimation of univariate models for the conditional mean and variance. For the conditional mean we rely on simple ARMA models with the lag length chosen in order to minimize the BIC. The conditional variance of the residuals was modeled using a GARCH(1,1) model with Student t errors. For brevity, we do not report the estimation results of the marginal models, but note that the ARMA models were mostly of very low order, the most complex ones being ARMA(1,1) models, and that the residuals and squared residuals showed no evidence of autocorrelation.

For the dependence we consider five competing models. The first two are the SCAR-DVine and GAS-DVine models introduced in Section 2. As a comparison we estimated multivariate Gaussian copula with time-varying correlation matrix using the DCC dynamics of Engle

(2002) in a similar fashion as in Heinen and Valdesogo (2009). Furthermore, we also consider a t-copula with DCC dynamics, denoted as t-DCC, due to the popularity of the t-copula for higher dimensional problems. The last model is the flexible R-Vine specification with constant dependence parameters studied in Brechmann et al. (2012) and Dißmann et al. (2013). The structure of the R-vine was selected using the algorithms suggested in these papers. For the D-vine based models we face two choices. The first is the ordering of the variables. This was done by maximizing the overall pairwise dependence measured by Kendall's tau. In particular, first choose the pair of variables with the highest empirical Kendall's τ. Second connect the next variable which has highest pairwise Kendall's τ with one of the previously chosen variables and proceed in a similar fashion until all variables are connected. This is the common strategy for D-vine copulas. In particular, we expect to capture the overall time variation of the dependence as good as possible with this choice, as it turns out that time variation is most relevant on the first tree.

Second, for each bivariate (conditional) copula model we then face two important choices, namely whether dependence is time-varying or static, and which copula family to use. We automatically select the model by first estimating time-varying and constant copulas from the following families: Gumbel (G), survival Gumbel (SG), Clayton (C), survival Clayton (SC), Normal (N) and the independence copula (I)[4]. We then select the best fitting copula family from these 11 candidate models by the BIC. Given the size and complexity of our model, as well as the difficulty to estimate parameters precisely on higher trees, we decided to rely on the BIC to find more parsimonious model specifications and to minimize the estimation error.

4.1 Weekly index returns

Our first dataset consists of weekly MSCI stock index return for the United States (US), United Kingdom (UK), Europe (EU), Canada (Can) and Japan (Jap) for the period Jan. 7, 1971 until October 10, 2013, resulting in a total of 2232 observations. We split the sample into an in-sample period consisting of the first 2000 returns, covering the period until June 30, 2009, and an out-of-sample period covering the remaining 232 observations.

The ordering of the countries was done as described above in turned out to be: Jap, UK, EU, USA, CAN. In Tables 1 and 2 we report the parameter estimates for the SCAR-DVine and GAS-DVine models, respectively. We do not report the estimation results for the R-vine and DCC models, but results are available from the authors upon request. It is striking that the majority of the chosen copulas are Gaussian. This corresponds with the findings Hafner and Manner

[4]Obviously, for the independence copula no parameter needs to be estimated.

Table 1: SCAR-DVine estimation parameters

Pair	Type	Family	Parameters		
			μ	ϕ	σ
12	Time Varying	Normal	0.2788	0.9829	0.0334
23	Time Varying	Normal	0.7066	0.9660	0.0517
34	Time Varying	Normal	0.4629	0.9867	0.0354
45	Time Varying	Normal	0.6517	0.7855	0.1054
13 | 2	Time Constant	Normal	0.1574	-	-
24 | 3	Time Varying	Normal	0.0043	0.9999	0.0055
35 | 4	Time Constant	Normal	0.1623	-	-
14 | 23	Time Constant	Gumbel	0.1031	-	-
25 | 34	Time Constant	Independent	0	-	-
15 | 234	Time Constant	Independent	0	-	-

Note: Selected copula models and parameter estimates of the SCAR-DVine model for weekly MSCI stock market index returns covering the period Jan 7, 1971 until April 30, 2009. The countries were ordered as follows: Jap, UK, EU, USA, CAN.

(2012). An explanation for this finding is given in Manner and Segers (2011), who show that the Gaussian copula with random correlations has much larger dependence in the tails than the static Gaussian copula. This is in line with the stylized fact that financial returns are characterized by tail dependence. On the first tree all models have time-varying dependence parameters, whereas in higher trees several conditional copulas are in fact static. For the GAS-DVine model time-variation is also found in most of the higher trees, which is not the case for the SCAR-DVine specification. The parameter estimates indicate higher dependence on the lower trees of the model. Notably, the SCAR-DVine model selects the independence copula in two out of three cases for the last two trees. The persistence parameters (ϕ and δ) indicate that the dependence parameters in the time-varying models are highly persistent. In several cases the upper bound of the permissible parameter space was selected, which indicates that the dependence parameter follows a random walk.

The in-sample fit of the models in terms of the log-likelihood statistic and the BIC is compared in Table 3. The D-vine based models clearly outperform the DCC and R-vine specifications and the GAS-DVine model clearly provides the best in-sample model fit.

Figure 2 shows the path of the time-varying Kendall's τ for all market pairs implied by the

Table 2: GAS-DVine estimation parameters

Pair	Type	Family	Parameters		
			ω	ϕ	δ
12	Time Varying	Normal	0.001	0.010	0.996
23	Time Varying	Normal	0.012	0.020	0.982
34	Time Varying	Normal	0.000	0.004	0.999
45	Time Varying	Normal	0.085	0.024	0.863
13 | 2	Time Varying	Normal	0.331	0.123	0.031
24 | 3	Time Varying	Normal	0.003	0.032	0.991
35 | 4	Time Varying	Normal	0.000	0.013	0.999
14 | 23	Time Constant	rotated Gumbel	0.287	-	-
25 | 34	Time Varying	Normal	0.006	0.060	0.981
15 | 234	Time Constant	rotated Clayton	-0.095	-	-

Note: Selected copula models and parameter estimates of the GAS-DVine model for weekly MSCI stock market index returns covering the period Jan 7, 1971 until April 30, 2009. The countries were ordered as follows: Jap, UK, EU, USA, CAN.

model allowing for time-varying dependence. In case of the SCAR-DVine model the smoothed path of τ_t is presented as an estimated for the latent dependence process. For pairs which are not directly connected in the first tree the dependence path was obtained by Monte Carlo simulation. The dynamics implied by the three models differ to a certain extend. In particular the dependence path of the DCC model[5] often deviate significantly from the other two. For market pairs that lie on neighboring nodes (e.g. Jap-UK or UK-EU) the paths of the SCAR-DVine and GAS-DVine models are closer to each other than for pairs that are not on neighboring nodes. No clear picture about the evolution of the dependence over these 40 years can be observed, but we note that there seems to evidence for an increase of the degree of dependence towards the end of the sample corresponding to the global financial crisis.

Finally, we use the estimated models to forecast the joint distribution of the stock market returns. We perform one-step forecasts and we do not re-estimate the models. The out-of-sample fit is compared based on two criteria. First, we compute the predictive log-likelihood of the copula part of the distribution to get an indication of the statistical fit, which can also be found in Table 3. The good performance of the DCC model that had the worst in-sample

[5]The path of the correlation for the t-DCC model is almost identical and therefore omitted from the graph.

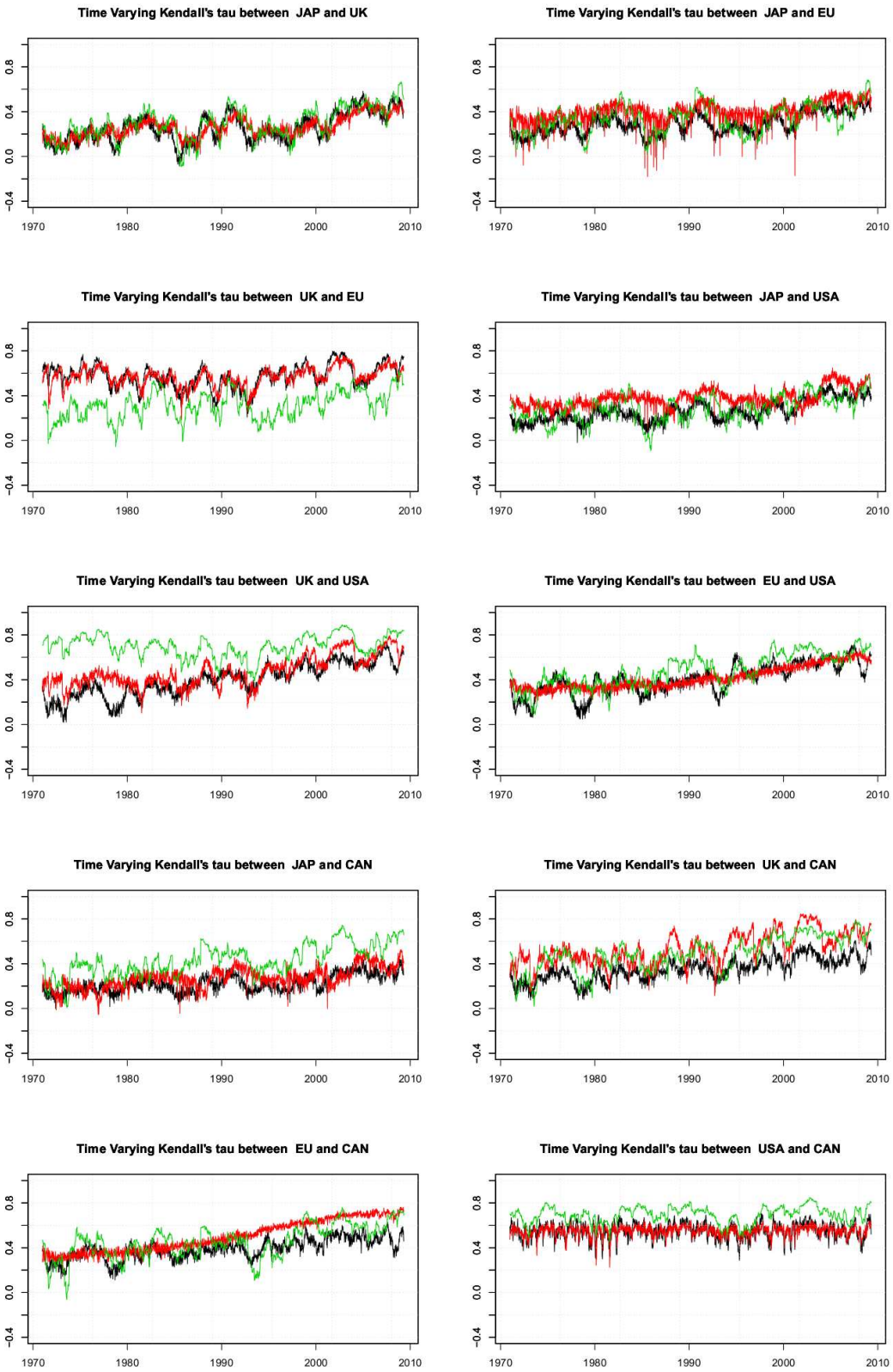

Figure 2: Marginal Kendal's tau for weekly MSCI returns based on the SCAR-DVine, GAS-DVine and DCC models in black, red and green, respectively.

Table 3: In- and out-of-sample fit MSCI returns

Model	LL	BIC	Pred. LL	# param.
SCAR-DVine	2645.39	-5153.96	515.48	18
GAS-DVine	2912.41	-5093.15	464.43	26
DCC	2030.75	-3970.29	479.50	12
t-DCC	2124.95	-4151.08	484.17	13
RVine (Constant)	2089.42	-4057.22	404.74	16

Note: This table presents the in-sample and out-of-sample fit of the analyzed model for the weekly MSCI returns in terms of log-likelihood (LL), Bayesian information criterion (BIC) and predictive log-likelihoood (Pred. LL). LL and BIC are based on the period Jan 7, 1971 until June 30, 2009. Pred. LL is based on the out-of-sample period July 6, 2009 until Oct 10, 2013. The last column gives the total number of parameters of each model.

fit is quite noticeable and the t-DCC performs even better, but nevertheless the SCAR model performs best also in term of predictive likelihood. For the second criterion of the out-of-sample fit of our model, we construct an equally weighted portfolio from the five market indices and estimate its Value-at-Risk at the 10%, 5% and 1% level based on our four model specifications. In Table 4 we report the exceedance rate, as well as the p-values of the dynamic quantile (DQ) test by Engle and Manganelli (2004), which tests the correct coverage of the VaR and the i.i.d.'ness of the exceedances. We apply the test with 0 lags in order to test the unconditional coverage of the VaR and with 4 lags to additionally test the i.i.d.'ness. The results show that all models except the time-constant RVine model perform well in terms of the unconditional coverage. However, the i.i.d.'ness of the VaR is rejected for all four models for the 1% VaR. Thus it seems like the choice of the dependence model has a relatively small influence on the quality of the VaR forecasts, as long as we allow for time-variation in the dependence parameters.

4.2 Daily returns on German stocks

In this section we provide an empirical illustration of the dynamic D-vine models. The dataset we consider are daily returns from the stocks listed in the DAX30 index during the period from the 1^{st} of January 2008 to the 31^{st} of December 2013, giving a total of 1523 observations for 29 stocks that were included in the Dax over the whole sample period. The first 4 years of data constitute the in-sample period, whereas the years 2012 and 2013 are used for the evaluation of the forecasts. A list of the included companies is given in the Appendix. We decided for this

Table 4: Value-at-Risk evaluation weekly MSCI returns

	Exceedance rate			EM-test 0 lags			EM-test 4 lags		
	10%	5%	1%	10%	5%	1%	10%	5%	1%
SCAR-DVine	0.112	0.065	0.013	0.821	0.592	0.872	0.583	0.449	0
GAS-DVine	0.095	0.060	0.013	0.536	0.444	0.890	0.973	0.502	0
DCC	0.108	0.060	0.013	0.836	0.544	0.558	0.950	0.484	0
t-DCC	0.121	0.073	0.017	0.412	0.118	0.526	0.586	0.239	0
RVine (constant)	0.116	0.082	0.022	0.538	0.036	0.063	0.897	0.045	0

Note: This table present the evaluation of the Value-at-Risk forecasts for weekly MSCI returns based on the out-of-sample period May 7, 2009 until October 10, 2013. EM-test refers to the p-value of the dynamic quantile test of Engle and Manganelli (2004).

dataset to find a balance between demonstrating the possibility of high dimensional modeling and the ability to still present the main results. Nevertheless, one could in principle consider much larger dimensions, which we leave for future research.

Due to the high complexity of the model in 29 dimensions, we considered the restriction of allowing potential time variation only on a limited (small) number of trees. We considered this restriction from 1 to 12 trees, but we report only for the best fitting specification since results are identical or at least very similar for many of those cases.[6] Specifically, it turned out that making this restriction beyond the 6^{th} tree is irrelevant, since there is no evidence of time variation on higher trees. A further possible restriction that may be made is to truncate the vine beyond a certain tree, meaning that all conditional copulas are set to the independence copula. In the current application we did not make this restriction because the independence copula is included in the set of admissible models and the automatic selection by the BIC in practice leads to a truncation. Nevertheless, for large dimensional applications truncation of the vine should definitely be considered.

An overview over the selected copula families is given in Table 5. Note that we only report the results for the model that restricts the time-variation to the first two trees, because this specification gave the best overall fit and there was very little evidence for time-variation beyond the second tree. For the SCAR model the Normal copula again dominates whenever time-variation is found, whereas in the remaining cases other copulas are also selected quite

[6]Further details regarding the model estimation such as parameter estimates or alternative specifications are not presented in order to conserve space, but are available from the authors upon request.

Table 5: SCAR-DVine and GAS-DVine model selection

Tree	time-varying			time-constant						# par
	N	G	SG	I	N	C	G	SC	SG	
SCAR-DVine										
1	28	0	0	0	0	0	0	0	0	84
2	2	0	0	0	2	0	8	0	15	31
3	-	-	-	1	5	0	11	0	9	25
4	-	-	-	6	8	0	8	0	3	19
5	-	-	-	6	8	0	6	2	2	18
6	-	-	-	9	9	1	3	0	1	14
7	-	-	-	10	7	1	0	0	4	12
GAS-DVine										
1	5	7	8	0	1	0	2	0	5	68
2	3	0	0	0	8	0	8	0	8	33
3	-	-	-	3	22	0	0	1	0	23
4	-	-	-	6	13	0	3	0	3	19
5	-	-	-	7	14	0	0	1	2	17

Note: Selected copulas on all trees of the dynamic DVine models. Time-variation was only permitted until the third tree. The copulas on the unreported trees were all the independence copulas. The data were daily returns on Dax30 stock prices for the period Jan 1, 2008 until Dec 31, 2011.

often. It is noticeable that on higher trees the independence copula tends to be selected quite often. Thus we can conclude that the time-variation is sufficiently captured on the first two trees of the model, whereas the overall dependence is captured by modeling the first 7 trees and inserting an independence copula on the remaining trees of the model. This is remarkable insofar as this suggests that is is possible to model much larger dimensional datasets, as the essential information concerning the dependence is captured within relatively few trees and the vine can be truncated beyond the first 7-10 trees.

Table 6 presents the in-sample log-likelihood and BIC for our dynamic models, as well as the Gaussian and Student DCC copula models and the aforementioned static RVine model. For this dataset the RVine model performs very well, but the t-DCC model has the largest log-likelihood. However, due to their large number of parameters (a total of 406 and 409, respectively) their BICs is larger than for the SCAR-DVine model. The Gaussian DCC and

Table 6: In- and out-of-sample fit daily Dax returns

Model	LL	BIC	Pred. LL	# param.
SCAR-DVine	10195.17	-18986.86	5567.36	203
GAS-DVine	9200.46	-17294.72	4378.58	160
DCC	9984.34	-17147.47	4513.32	408
t-DCC	10750.75	-18676.78	5005.82	409
RVine (Constant)	10587.76	-18368.54	4128.80	406

Note: This table presents the in-sample and out-of-sample fit of the analyzed model for the daily Dax returns in terms of log-likelihood (LL), Bayesian information criterion (BIC) and predictive log-likelihoood (Pred. LL). LL and BIC are based on the period Jan 1, 2008 until Dec 31, 2011. Pred. LL is based on the out-of-sample period Jan 1, 2012 until Dec 31, 2013. The last column gives the total number of parameters of each model.

GAS-DVine models perform much worse in comparison.

Next, in the same fashion as in Section 4.1 we computed density forecasts for the remaining two years of data, i.e. Jan 1, 2012 until Dec 31, 2013. The predictive log-likelihood suggests that the SCAR model provides by far the best density forecasts, again followed by the t-DCC model. The predictive log-likelihood of the RVine is the lowest, showing that this model is perhaps overparametrized and/or that its structure is not stable over time. Finally, Value-at-Risk forecasts are computed and backtested in the same way as for the MSCI returns. The results in Table 7 show that the coverage rates are mostly below their nominal levels and no model shows entirely satisfactory results. The two DCC models look a little better, though. The results of the dynamic quantile test are equally satisfactory for all models, so again in terms of forecasting the Value-at-Risk each of the five models compared appear to be performing equally well.

5 Conclusions and further research

In light of the recent financial crisis, including the discussion of understanding systemic risk, the need to understand time-varying effects not only within individual financial products but also among groups of financial variables has been increasing. The developed D-Vine SCAR and GAS models are aiming to fill this demand. From a statistician's point of view such a model is demanding. First, a very flexible multivariate dependency model is required such as the class of vine copulas and, secondly, an appropriate model for the time dependency of the

Table 7: Value-at-Risk evaluation daily Dax returns

	Exceedance rate			EM-test 0 lags			EM-test 5 lags		
	10%	5%	1%	10%	5%	1%	10%	5%	1%
SCAR-DVINE	0.072	0.045	0.006	0.070	0.613	0.633	0.542	0.385	0.993
GAS-DVINE	0.070	0.039	0	0.056	0.453	0.074	0.405	0.838	0.985
DCC	0.076	0.047	0.016	0.055	0.745	0.172	0.560	0.527	0.076
t-DCC	0.079	0.054	0.016	0.189	0.869	0.345	0.747	0.523	0.119
Rvine (constant)	0.072	0.039	0.008	0.081	0.449	0.734	0.691	0.978	0.985

Note: This table present the evaluation of the Value-at-Risk forecasts for daily Dax returns based on the out-of-sample period Jan 1, 2012 until Dec 31, 2013. EM-test refers to the p-value of the dynamic quantile test of Engle and Manganelli (2004).

copula parameters has to found. Here we consider parameter driven and observation driven approaches, but the former approach seems to be more suitable considering its superior in-sample and out-of-sample fit. As a general recommendation we would thus favor the D-vine SCAR model.

While this approach leads to a relatively straightforward model formulation, the development of efficient estimation procedures is much more difficult. Especially in high dimensions maximum likelihood is infeasible, since it would require the maximization over integrals of size equal to the data length. These integrals occur since we need to integrate over the latent variable process to express the joint likelihood. This problem already occurs when we consider bivariate SCAR models. One solution to this is to use efficient importance sampling (Richard and Zhang 2007). In addition, the pair copula construction approach of Aas et al. (2009) for multivariate copulas allows to express the likelihood in bivariate copula terms in addition to a sequential formulation over the vine tree structure. This makes it feasible to develop and implement efficient importance sampling for the D-vine SCAR model.

One interesting feature of the applications in this paper is that non-normal pair copulas with constant parameters were replaced by normal pair copulas when time-varying copula parameters are allowed. This shows that the assumption of conditional normality is in fact in line with the data whenever an appropriate model for the time-varying correlations is chosen. It should also be noted that using dynamic D-vine specifications might mitigate the misspecification effects of using a simplified vine specification instead of a true underlying non simplified vine specification. In particular this might be the case when the effect of the conditioning

value on the pair copulas in higher trees arises from time varying conditioning variables.

In this paper we follow some approaches to model selection. We first restrict to a known dependency structure given by a D-vine, but allow the copula family of each pair copula to be chosen among a prespecified class of copula families in addition to the choice if a pair copula has time-varying parameters or not. For this we use BIC, although more sophisticated criteria might be necessary. When considering larger dimensions, we also restrict the use of time-varying pair copula parameters to a prespecified number of top trees. Here the approach of truncated vines as developed in Brechmann et al. (2012) might be a good starting point to choose this number in a data driven manner. As already mentioned, it is feasible to extend the class of D-vine SCAR models to include R-vines as copula models. Furthermore, the model uncertainty introduced by assuming the marginal parameter estimates as true ones in the two-step approach has to be assessed in future research. However the simulation results of Kim et al. (2007) can be expected to remain valid for a copula model with time-varying parameters. Additionally, the computation of standard errors for the parameter estimates needs to be dealt with in future research. Finally, the model may be combined with the D-vine mixture model suggested in Kim et al. (2013) to obtain an even more flexible dependence model.

A Stocks in the DAX and their ordering

Company	Ticker Symbol	node in D-vine
Adidas	ADS	24
Allianz	ALV	8
BASF	BAS	14
Bayer	BAYN	16
Beiersdorf	BEI	26
BMW	BMW	19
Commerzbank	CBK	10
Daimler	DAI	18
Deutsche Bank	DBK	9
Deutsche Börse	DB1	3
Lufthansa	LHA	11
Deutsche Post	DPW	12
Deutsche Telekom	DTE	5
E.ON	EOAN	7
Fresenius	FRE	29
Fresenius Medical Care	FME	28
HeidelbergCement	HEI	22
Henkel	HEN3	25
Infineon Technologies	IFX	2
K+S	SDF	1
Linde	LIN	15
MAN	MAN	20
Merck	MRK	27
Metro	MEO	4
Munich Re	MUV2	17
RWE	RWE	6
SAP	SAP	23
Siemens	SIE	13
ThyssenKrupp	TKA	21

Table 8: DAX companies and their node position in the selected D-vine for the in-sample period from the 1^{st} of January 2008 until the 31^{st} of December 2011.

References

Aas, K., C. Czado, A. Frigessi, and H. Bakken (2009). Pair-copula constructions of multiple dependence. *Insurance: Mathematics and Economics 44*, 182–198.

Acar, E. F., C. Genest, and J. NešLehová (2012). Beyond simplified pair-copula constructions. *Journal of Multivariate Analysis 110*, 74–90.

Almeida, C. and C. Czado (2012). Bayesian inference for stochastic time-varying copula models. *Computational Statistics & Data Analysis 56(6)*, 1511–1527.

Bauwens, L., S. Laurent, and J. V. K. Rombouts (2006). Multivariate garch models: a survey. *Journal of Applied Econometrics 21*, 79–109.

Bedford, T. and R. Cooke (2001). Probability density decomposition for conditionally dependent random variables modeled by vines. *Annals of Mathematics and Artificial Intelligence 32*, 245–268.

Bedford, T. and R. M. Cooke (2002). Vines - a new graphical model for dependent random variables. *Annals of Statistics 30*, 1031–1068.

Blasques, F., S. Koopman, and A. Lucas (2014). Stationarity and ergodicity conditions for generalized autoregressive score processes. *Electronic Journal of Statistics 8*, 1088–1112.

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

Brechmann, E. and C. Czado (2015). COPAR – multivariate time series modeling using the copula autoregressive model. *Applied Stochastic Models in Business and Industry 31*, 495–514.

Brechmann, E., C. Czado, and K. Aas (2012). Truncated regular vines in high dimensions with application to financial data. *Canadian Journal of Statistics 40*, 68–85.

Brechmann, E. C. and C. Czado (2013). Risk management with high-dimensional vine copulas: An analysis of the Euro Stoxx 50. *Statistics & Risk Modeling 30*(4), 307–342.

Cherubini, G., E. Luciano, and W. Vecchiato (2004). *Copula Methods in Finance*. John Wiley and Sons Ltd.

Chollete, L., A. Heinen, and A. Valdesogo (2009). Modeling international financial returns with a multivariate regime switching copula. *Journal of Financial Econometrics 7*, 437–480.

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

Creal, D. and R. Tsay (2015). High-dimensional dynamic stochastic copula models. *Journal of Econometrics*, forthcoming.

Czado, C. (2010). Pair-copula constructions of multivariate copulas. In *In F. Durante, W. Härdle, P. Jaworki, and T. Rychlik (Eds.),* Workshop on Copula Theory and its Applications. *Springer, Dortrech.*, pp. 93–110.

Czado, C., E. C. Brechmann, and L. Gruber (2013). Selection of vine copulas. In *Copulae in Mathematical and Quantitative Finance*, pp. 17–37. Springer.

Czado, C., U. Schepsmeier, and A. Min (2012). Maximum likelihood estimation of mixed c-vines with application to exchange rates. *Statistical Modelling 12(3)*, 229–255.

de Melo Mendes, B. V., M. M. Semeraro, and R. P. C. Leal (2010). Pair-copulas modeling in finance. *Financial Markets and Portfolio Management 24*(2), 193–213.

Dias, A. and P. Embrechts (2004). Change-point analysis for dependence structures in finance and insurance. In G. Szegoe (Ed.), *Risk Measures for the 21st Century*, pp. 321–335. Wiley Finance Series.

Dißmann, J., E. Brechmann, C. Czado, and D. Kurowicka (2013). Selecting and estimating regular vine copulae and application to financial returns. *Computational Statistics and Data Analysis 59*, 52–69.

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

Engle, R. F. and S. Manganelli (2004). CAViaR: Conditional autoregressive value at risk by regression quantiles. *Journal of Business and Economic Statistics 22*, 367–381.

Fischer, M., C. Köck, S. Schlüter, and F. Weigert (2009). An empirical analysis of multivariate copula models. *Quantitative Finance 9*(7), 839–854.

Garcia, R. and G. Tsafack (2011). Dependence structure and extreme comovements in international equity and bond markets. *Journal of Banking and Finance 35*, 1954–1970.

Genest, C., K. Ghoudi, and L.-P. Rivest (1995). A semiparametric estimation procedure of dependence parameters in multivariate families of distributions. *Biometrika 82*, 543–552.

Giacomini, E., W. Härdle, and V. Spokoiny (2009). Inhomogeneous dependency modelling with time varying copulae. *Journal of Business and Economic Statistics 27*, 224–234.

Haff, I. H. (2013). Estimating the parameters of a pair copula construction. *Bernoulli 19(2)*, 462–491.

Haff, I. H., K. Aas, and A. Frigessi (2010). On the simplified pair-copula construction – simply useful or too simplistic? *Journal of Multivariate Analysis 101*(5), 1296 – 1310.

Hafner, C. M. and H. Manner (2012). Dynamic stochastic copula models: Estimation, inference and applications. *Journal of Applied Econometrics 7*, 269–295.

Hafner, C. M. and O. Reznikova (2010). Efficient estimation of a semiparametric dynamic copula model. *Computational Statistics and Data Analysis 54*, 2609–2627.

Harvey, A. C., E. Ruiz, and N. Shephard (1994). Multivariate stochastic variance models. *Review of Economic Studies 61*, 247–264.

Heinen, A. and A. Valdesogo (2009). Asymmetric capm dependence for large dimensions: The canonical vine autoregressive copula model. Available at SSRN: http://ssrn.com/abstract=1297506, CORE.

Joe, H. (1996). Families of m-variate distributions with given margins and m(m-1)/2 bivariate dependence parameters. In L. Rüschendorf and B. Schweizer and M. D. Taylor (Ed.), *Distributions with Fixed Marginals and Related Topics*.

Joe, H. (1997). *Multivariate Models and Dependence Concepts*. Chapman & Hall/CRC.

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

Joe, H. (2014). *Dependence Modeling with Copulas*. CRC Press.

Joe, H., H. Li, and A. Nikoloulopoulos (2010). Tail dependence functions and vine copulas. *Journal of Multivariate Analysis 101*, 252–270.

Kim, D., J.-M. Kim, S.-M. Liao, and Y.-S. Jung (2013). Mixture of D-vine copulas for modeling dependence. *Computational Statistics & Data Analysis 64*, 1–19.

Kim, G., M. J. Silvapulle, and P. Silvapulle (2007). Comparison of semiparametric and parametric methods for estimating copulas. *Computational Statistics & Data Analysis 51*, 2836 – 2850.

Koopman, S., A. Lucas, and M. Scharth (2015). Predicting time-varying parameters with parameter-driven and observation-driven models.

Kurowicka, D. and R. Cooke (2006). *Uncertainty analysis with high dimensional dependence modelling*. Chichester: Wiley.

Kurowicka, D. and R. Cooke (2007). Sampling algorithms for generating joint uniform distributions using the vine-copula method. *Computational Statistics & Data Analysis 51*(6), 2889–2906.

Kurowicka, D. and H. Joe (2011). *Dependence Modeling - Handbook on Vine Copulae*. Singapore: World Scientific Publishing Co.

Liesenfeld, R. and J. F. Richard (2003). Univariate and multivariate stochastic volatility models: Estimation and diagnostics. *Journal of Empirical Finance 10*, 505–531.

Manner, H. and O. Reznikova (2012). A survey on time-varying copulas: Specification, simulations and application. *Econometric Reviews 31(6)*, 654–687.

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

Mendes, B., M. Semeraro, and R. Leal (2011). Pair-copulas modeling in finance. *Financial Markets and Portfolio Management 24*, 193–213.

Nelsen, R. (2006). *An Introduction to Copulas*. New York: Springer.

Oh, D. H. and A. J. Patton (2013). Time-varying systemic risk: Evidence from a dynamic copula model of CDS spreads. *Economic Research Initiatives at Duke (ERID) Working Paper* (167).

Oh, D. H. and A. J. Patton (2015). Modelling dependence in high dimensions with factor copulas. *Journal of Business and Economic Statistics*, forthcoming.

Okhrin, O., Y. Okhrin, and W. Schmid (2013). On the structure and estimation of hierarchical archimedean copulas. *Journal of Econometrics 173*, 189–204.

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

Patton, A. J. (2009). Copula-based models for financial time series. In T. G. Andersen, R. A. Davis, J.-P. Kreiss, and T. Mikosch (Eds.), *Handbook of Financial Time Series*. Springer Verlag.

Richard, J. F. and W. Zhang (2007). Efficient high-dimensional importance sampling. *Journal of Econometrics 141*, 1385–1411.

Righi, M. B. and P. S. Ceretta (2015). Forecasting value at risk and expected shortfall based on serial pair-copula constructions. *Expert Systems with Applications 42*(17), 6380–6390.

Savu, C. and M. Trede (2010). Hierarchies of archimedean copulas. *Quantitative Finance 10*, 295–304.

Schepsmeier, U. and E. Brechmann (2013). Modeling dependence with C- and D-Vine copulas: The R package CDVine. *Journal of Statistical Software 52(3)*, 1–27.

So, M. and C. Yeung (2014). Vine-copula GARCH model with dynamic conditional dependence. *Computational Statistics & Data Analysis 76*, 655–671.

Stöber, J. and C. Czado (2011). Pair copula constructions. In M. Scherer and J.-F. Mai (Eds.), *Simulating Copulas: Stochastic Models, Sampling Algorithms, and Applications*, Singapore, pp. 187–232. World Scientific Publishing.

Stöber, J. and C. Czado (2014). Regime switches in the dependence structure of multidimensional financial data. *Computational Statistics & Data Analysis 76*, 672–686.

Stöber, J., H. Joe, and C. Czado (2013). Simplified pair copula constructions - limitations and extensions. *Journal of Multivariate Analysis 119*, 101–118.

Weiß, G. N. and H. Supper (2013). Forecasting liquidity-adjusted intraday value-at-risk with vine copulas. *Journal of Banking & Finance 37*(9), 3334–3350.

Yu, J. and R. Meyer (2006). Multivariate stochastic volatility models: Bayesian estimation and model comparison. *Econometric Reviews 25*, 361–384.