Risk management with high-dimensional vine copulas: An analysis of the Euro Stoxx 50

Eike Christian Brechmann*, Claudia Czado

October 24, 2011

Abstract

The demand for an accurate financial risk management involving larger numbers of assets is strong not only in view of the financial crisis of 2007-2009. In particular dependencies among assets have not been captured adequately. While standard multivariate copulas have added some flexibility, this flexibility is insufficient in higher dimensional applications. Vine copulas can fill this gap by benefiting from the rich class of existing bivariate parametric copula families. Exploiting this in combination with GARCH models for margins, we develop a regular vine copula based factor model for asset returns, the Regular Vine Market Sector model, that is motivated by the classical CAPM and shown to be superior to the CAVA model proposed by Heinen and Valdesogo (2009). While the model can also be used to separate the systematic and idiosyncratic risk of specific stocks, we explicitly discuss how vine copula models can be employed for active and passive portfolio management. In particular, Value-at-Risk forecasting and asset allocation are treated in detail. All developed models and methods are used to analyze the Euro Stoxx 50 index, a major market indicator for the Eurozone. Relevant benchmark models such as the popular DCC model and the common Student-t copula are taken into account.

1 Introduction

In light of the recent financial crisis of 2007-2009 and increasing volatility at global financial markets, a diligent risk management is crucially important for any financial institution and also required by regulators. The Basel II and III rules for banks and Solvency II for the insurance sector encourage the use of sophisticated internal models. An important issue of such models is how dependence among different assets is treated. It is a stylized fact that the long time predominant Gaussian correlations are inappropriate in this matter. Dependencies among asset returns exhibit features such as tail dependence and asymmetry (see, amongst others, Longin and Solnik (1995, 2001) and Ang and Bekaert (2002)), for which the classical product-moment correlation as implied by the normal distribution cannot account. See also Embrechts et al. (2002) for a comprehensive discussion of the use of linear product-moment correlations in risk management. Despite this, the dramatic events of 2007-2009 most emphatically stress that the need for adequate models capturing complex dependence structures of large numbers of assets and accurately assessing financial risk is stronger than ever.

* *Corresponding author.* Center for Mathematical Sciences, Technische Universität München, Boltzmannstr. 3, D-85747 Garching, Germany. E-mail: brechmann@ma.tum.de. Phone: +49 89 289-17425.

It is also well known that asset returns do not follow a Gaussian distribution as already noted by Mandelbrot (1963). They in particular exhibit fat tails which cannot be captured adequately using a Gaussian distribution. Since the seminal work of Engle (1982), there however has been considerable progress in modeling financial time series. In particular, GARCH-models as introduced by Bollerslev (1986) are widely and successfully used. They allow for time-varying variances and the use of fat-tailed distributions as observed for financial returns. There also has been some progress in extending GARCH-models beyond univariate specifications (see Bauwens et al. (2006) for a survey). Multivariate GARCH-models such as the very popular DCC model by Sheppard and Engle (2001) and Engle (2002) however use only covariances for dependence modeling among time series, while maintaining positive definite covariance matrices.

Copulas overcome this problem. According to the famous theorem by Sklar (1959), the modeling of margins and dependence can be separated by the way of copulas. Hence, the last years saw a steadily growing literature on copula GARCH-models, where the marginal time series are modeled with univariate GARCH-models, while copulas are applied for the dependence structure. A recent survey can be found in Fischer et al. (2009).

Unfortunately, the choice of multivariate copulas is rather limited in contrast to the bivariate case where a rich variety of different copula types exhibiting flexible and complex dependence patterns exists. While the infamous Gaussian copula (see Salmon (2009) for a criticism on its role during the financial crisis) and also the Student-t copula can only capture symmetric dependencies using correlation matrices, exchangeable Archimedean copulas use only one or two parameters for the dependency modeling among possibly dozens of variables which clearly is too restrictive. The so-called pair copula constructions overcome this issue. They were originally proposed by Joe (1996) and further explored and greatly extended by Bedford and Cooke (2001, 2002) and Kurowicka and Cooke (2006). *Regular vines* (R-vines) are a convenient graphical model to classify such pair copula constructions and hierarchical in nature. Each level only involves the specification of arbitrary bivariate copulas as building blocks and hence allows for very flexible models exhibiting effectively any possible dependence structure.

The aim of this paper is to present and discuss the use of vine copulas for financial risk management and thus introduce them to a broader audience. While in finance the use of copulas is quite common nowadays (see, e.g., Cherubini et al. (2004)), literature on vine copula modeling so far mainly concentrated on illustrative applications (see Aas et al. (2009), Fischer et al. (2009), Berg and Aas (2009), Min and Czado (2010) and Czado et al. (2011)). A first step to compensate this deficiency was taken by Mendes et al. (2010) who employ a D-vine copula model (a specific sub-class of R-vines) with only four different bivariate copula families for a six-dimensional data set and discuss its use for portfolio management. Our work goes well beyond this. Since financial risk management often involves large numbers of assets to be modeled accurately, we particularly focus on high-dimensional vine copula modeling. Model selection techniques, which are required to find adequate R-vine structures especially in such dimensions, are presented. To illustrate our methods we analyze the Euro Stoxx 50 index which is a major market indicator for the Eurozone.

The analysis of the members of the Euro Stoxx 50 also motivates the consideration of vine copula based factor models, in particular vine market sector models. Heinen and Valdesogo (2009) developed an extension of the CAPM which can capture non-linear and non-Gaussian behavior of the cross-section of asset returns as well as model their dependencies to the market and the respective sector. Their so-called *Canonical Vine Autoregressive* (CAVA) model is based on two major building blocks: marginal GARCH-models and a canonical vine copula structure (another sub-class of R-vines). In this paper, we propose the *Regular Vine Market Sector* (RVMS) model which uses a flexible general R-vine copula construction without relying on very strict independence assumptions like the CAVA model. We thoroughly investigate both models in our application to European stock market returns, establishing the superiority of the RVMS model over the CAVA model. In contrast to Heinen and Valdesogo (2009), copula parameters are also fitted using maximum likelihood estimation rather than only sequentially and hence provide the basis for bootstrapping to obtain confidence intervals. Moreover, although the factor based RVMS model might suffer from its less flexible structure compared to general R-vine specifications, we demonstrate in this application that it not only facilitates interpretation of dependencies but also is a very good model compared to relevant benchmark models. By grouping assets by sectors, it thus mitigates the curse of dimensionality to some extent.

Our second contribution is a discussion on how vine copula based factor models can be used to identify the systematic and the idiosyncratic risk inherent in an asset return. It is however important not to neglect that dependence modeling using copulas goes well beyond summarizing dependence in one single number. The information contained in a copula is much richer and crucial to be taken into account when assessing complex financial risk.

Very important from a practitioner's point of view is our third contribution. We describe in detail how vine copulas can be used for portfolio management. This includes forecasting the Value-at-Risk (or any other risk measure) of a portfolio as well as portfolio diversification calculations and asset allocation. In doing so, we particularly point out why vine copula based models are a serious alternative to the often used DCC model.

To summarize, this paper introduces R-vines to complex financial applications, in particular to issues of financial risk management. Building on this work, R-vines can easily serve as a "construction kit" for even more elaborate models and applications.

The remainder of the paper is structured as follows. The Euro Stoxx 50 data that we analyze at the end of each of the following sections is described in Section 2. In Section 3 we give a brief introduction to the theory of copulas and discuss R-vines and their selection. Vine market sector models are treated in Section 4, where we first review the work by Heinen and Valdesogo (2009) before our new RVMS model is proposed. The separation of systematic and idiosyncratic risk is treated subsequently. Finally, portfolio management using vine copulas is described in Section 5 which is subdivided into the discussion of passive and active management methods. Section 6 provides conclusions and an outlook to future research.

2 Data

The Euro Stoxx 50 index is a major barometer of financial markets in the Eurozone. It covers stocks of 50 large Eurozone companies selected based on their market capitalization. According to Stoxx Ltd, the index serves "as underlying for a wide range of investment products such as Exchange Traded Funds, Futures and Options, and structured products worldwide". A detailed understanding of the dynamics of the Euro Stoxx 50 therefore is an important issue.

This crucially involves an accurate assessment of the dependencies among the index members. In the present paper we concentrate on 46 of them. For the index composition as of February 8, 2010, these are the stocks from the five largest Eurozone economies, namely Germany, France, Italy, Spain and the Netherlands. The four disregarded stocks correspond to only 6.7% of the

	Ticker symbols
Indices	ˆSTOXX50E, ˆGDAXIP, ˆFCHI, FTSEMIB.MI, ˆIBEX, ˆAEX
Germany	ALV.DE, BAS.DE, BAYN.DE, DAI.DE, DB1.DE, DBK.DE, DTE.DE, EOAN.DE, MUV2.DE, RWE.DE, SAP.DE, SIE.DE
France	ACA.PA, AI.PA, ALO.PA, BN.PA, BNP.PA, CA.PA, CS.PA, DG.PA, FP.PA, FTE.PA, GLE.PA, GSZ.PA, MC.PA, OR.PA, SAN.PA, SGO.PA, SU.PA, UL.PA, VIV.PA
Italy	ENEL.MI, ENI.MI, G.MI, ISP.MI, TIT.MI, UCG.MI
Spain	BBVA.MC, IBE.MC, REP.MC, SAN.MC, TEF.MC
Netherlands	AGN.AS, INGA.AS, PHIA.AS, UNA.AS

Table 1: Ticker symbols of analyzed stocks.

total index weight. All stocks considered in our analyses are shown in Table 1. The table also shows the respective five national indices which the 46 stocks belong to: the German DAX, the French CAC 40, the Italian FTSE MIB, the Spanish IBEX 35, and the Dutch AEX. In the following, these indices as well as the Euro Stoxx 50 will be taken as price indices rather than performance indices. The quantity of interest in our analyses will be daily log returns.

Recent statistical analyses of the Euro Stoxx 50 can be found, e.g., in Savu and Trede (2010), Cherubini et al. (2004) and Hlawatsch and Reichling (2010). The first paper analyzes a subset of twelve stocks using a so-called Hierarchical Archimedean copula, which constitutes a rather flexible multivariate copula model. Hlawatsch and Reichling (2010) consider a portfolio allocation problem for 16 major stocks belonging to the Euro Stoxx 50 using skew-t distributions and a Clayton copula. On the other hand, only Cherubini et al. (2004) analyze all 50 components of the Euro Stoxx 50 with GARCH(1,1) filters and Student-t error distributions. For the dependence structure they consider multivariate Gaussian and Student-t copulas as well as two Archimedean copulas.

3 Multivariate copulas

The immense popularity copulas have been enjoying recently is due to the theorem by Sklar (1959). In the first instance, copulas simply are multivariate distribution functions with uniform margins. For a random vector $\mathbf{X} = (X_1, ..., X_d) \sim F$ with marginal distributions $F_i, i = 1, ..., d$, Sklar's theorem then states that

$$F(x_1, ..., x_d) = C(F_1(x_1), ..., F_d(x_d)), \tag{3.1}$$

where C is some appropriate d-dimensional copula. Moreover, if F is absolutely continuous and $F_1, ..., F_d$ are strictly increasing continuous,

$$f(x_1, ..., x_d) = \left[\prod_{k=1}^{d} f_k(x_k)\right] \times c(F_1(x_1), ..., F_d(x_d)),$$

where small letters denote corresponding density expressions. In other words, copulas conveniently allow to separate the modeling of the marginals and the dependency part in terms of the copula. Comprehensive references on copula theory are the books by Joe (1997) and Nelsen (2006).

Since copulas are inherently static models, they have to be combined with time series models for modeling financial returns (see, e.g., Cherubini et al. (2004)). This means that first appropriate time series models such as ARMA-GARCH are fitted to each financial return series. Copula modeling then proceeds with standardized residuals obtained from these models. More details on this issue will be presented in Section 5.1.

While in two dimensions there is a wide range of flexible copulas available and frequently used, multivariate copula modeling is more challenging. Standard classes such as elliptical and Archimedean copulas exhibit severe drawbacks in larger dimensions. In particular, Archimedean copulas which typically have only one or two parameters impose very strict dependence properties. Elliptical copulas such as the Gaussian or the Student-t, on the other hand, suffer from a large number of parameters and symmetry restrictions in the tails. While accounting for tail dependence, the Student-t copulas however has only one parameter controlling its strength for *all* pairs. There has been considerably effort in extending and enriching both classes. One such approach are the before-mentioned Hierarchical Archimedean copulas, which however suffer from a somehow restrictive structure and are limited to the Archimedean copula class (see, amongst others, Savu and Trede (2010)). The class of so-called *vine copulas* however easily overcomes the issues of elliptical and Archimedean copulas—as well as other copulas—and at the same time exploits their virtue in the bivariate case.

3.1 Vine copulas

Vine copulas are another name for so-called pair copula constructions (PCCs) as introduced by Aas et al. (2009). A PCC can be easily illustrated in three dimensions. Let $\mathbf{X} = (X_1, X_2, X_3) \sim F$ and assume that all necessary densities exist. Then

$$f(x_1, x_2, x_3) = f_1(x_1) f(x_2|x_1) f(x_3|x_1, x_2). \tag{3.2}$$

Using Sklar's theorem (3.1) it follows

$$f(x_2|x_1) = \frac{f(x_1, x_2)}{f_1(x_1)} = \frac{c_{1,2}(F_1(x_1), F_2(x_2)) f_1(x_1) f_2(x_2)}{f_1(x_1)} = c_{1,2}(F_1(x_1), F_2(x_2)) f_2(x_2), \tag{3.3}$$

and

$$\begin{aligned} f(x_3|x_1, x_2) &= \frac{f(x_2, x_3|x_1)}{f(x_2|x_1)} = \frac{c_{2,3|1}(F(x_2|x_1), F(x_3|x_1)) f(x_2|x_1) f(x_3|x_1)}{f(x_2|x_1)} \\ &= c_{2,3|1}(F(x_2|x_1), F(x_3|x_1)) f(x_3|x_1) \\ &\overset{(3.3)}{=} c_{2,3|1}(F(x_2|x_1), F(x_3|x_1)) c_{1,3}(F_1(x_1), F_3(x_3)) f_3(x_3), \end{aligned} \tag{3.4}$$

with

$$F(x|\boldsymbol{v}) = \frac{\partial C_{xv_j|\boldsymbol{v}_{-j}}(F(x|\boldsymbol{v}_{-j}), F(v_j|\boldsymbol{v}_{-j}))}{\partial F(v_j|\boldsymbol{v}_{-j})}, \tag{3.5}$$

where $C_{xv_j|\boldsymbol{v}_{-j}}$ is a bivariate copula distribution function and $\boldsymbol{v}_{-j}$ denotes a vector with the j-th component v_j removed.

Combining Equations (3.2)-(3.4) then gives a representation of a three-dimensional joint density in terms of bivariate copulas $C_{1,2}, C_{1,3}$ and $C_{2,3|1}$ only. These can be chosen independently of each other so that a wide range of different dependence structures can be modeled using PCCs which are constructed in essentially the same way in high dimensions.

Vines are a graphical representation to specify such PCCs. They were introduced by Bedford and Cooke (2001, 2002) and described in more detail in Kurowicka and Cooke (2006) and

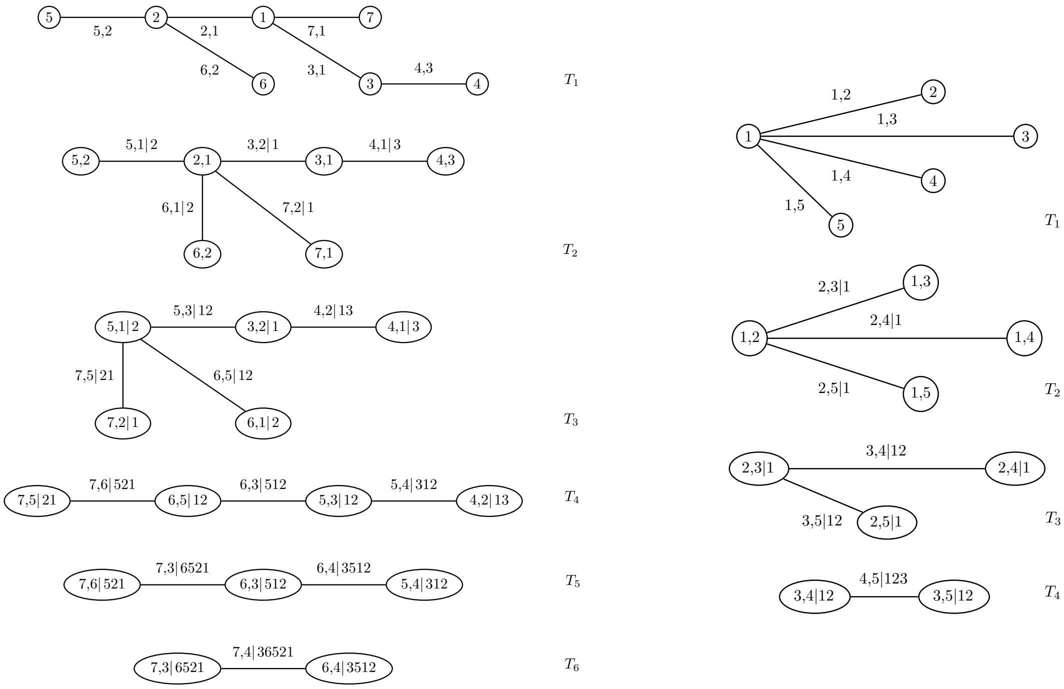

Figure 1: Seven-dimensional R- (left panel) and five-dimensional C-vine trees (right panel) with edge indices.

Kurowicka and Joe (2011). Statistical inference of regular vines is considered in Dißmann et al. (2011).

According to Definition 4.4 of Kurowicka and Cooke (2006) a *regular vine* (R-vine) on d variables consists first of a sequence of linked trees (connected acyclic graphs) $T_1, ..., T_{d-1}$ with nodes N_i and edges E_i for $i = 1, ..., d-1$, where T_1 has nodes $N_1 = \{1, ..., d\}$ and edges E_1, and for $i = 2, ..., d-1$ tree T_i has nodes $N_i = E_{i-1}$. Moreover, the proximity condition requires that two edges in tree T_i are joined in tree T_{i+1} only if they share a common node in tree T_i.

It is shown in Bedford and Cooke (2001) and Kurowicka and Cooke (2006) that the edges in an R-vine tree can be uniquely identified by two nodes, the *conditioned nodes*, and a set of *conditioning nodes*, i.e., edges are denoted by $e = j(e), k(e)|D(e)$ where $D(e)$ is the conditioning set. An example of a seven-dimensional R-vine tree sequence with edge labels is given in the left panel of Figure 1 (see Dißmann et al. (2011)).

The multivariate copula associated to trees $T_1, ..., T_{d-1}$ is then built up by associating each edge $e = j(e), k(e)|D(e)$ in E_i with a bivariate copula density $c_{j(e),k(e)|D(e)}$. According to Theorem 4.2 of Kurowicka and Cooke (2006) the R-vine copula density is uniquely determined and given by

$$c(F_1(x_1), ..., F_d(x_d)) = \prod_{i=1}^{d-1} \prod_{e \in E_i} c_{j(e),k(e)|D(e)}(F(x_{j(e)}|\boldsymbol{x}_{D(e)}), F(x_{k(e)}|\boldsymbol{x}_{D(e)})),$$

where $\boldsymbol{x}_{D(e)}$ denotes the subvector of $\boldsymbol{x} = (x_1, ..., x_d)$ indicated by the indices contained in $D(e)$.

A special case of R-vines which is often considered are *canonical vines* (C-vines). In particular, an R-vine is called C-vine if each tree T_i has a unique node with degree $d - i$, the *root node*. A five-dimensional example is shown in the right panel of Figure 1. According to Aas et al. (2009) we can write the C-vine copula density as

$$c(F_1(x_1), ..., F_d(x_d)) = \prod_{i=1}^{d-1} \prod_{j=1}^{d-i} c_{i,i+j|1,...,i-1}(F(x_i|x_1, ..., x_{i-1}), F(x_{i+j}|x_1, ..., x_{i-1})).$$

Evidently C-vines are only appropriate when there are pivotal quantities among the random variables under consideration. Otherwise this convenient tree structure may be too restrictive, in particular in higher dimensions.

3.2 Model selection

The number of different R-vines is substantial: $\binom{d}{2} \times (d-2)! \times 2^{\binom{d-2}{2}}$ in d dimensions as shown in Morales-Nápoles et al. (2010). For example, in seven dimensions (see Figure 1) there are 2,580,480 different R-vines, while for $d = 23$ there are already more R-vines than atoms in the universe. Also for C-vines there are as many as $d!/2$ different ones (Aas et al. 2009). We therefore rely on heuristic methods to select appropriate R- and C-vine trees. Following Aas et al. (2009), Brechmann et al. (2010) and Dißmann et al. (2011) the idea is to capture the strongest dependencies in the first trees, since these are typically most important to be modeled explicitly and accurately.

The setting of the first tree is a graph on d nodes which correspond to the variables and where all nodes are connected to each other by edges. To each edge a weight is attached according to an arbitrary pairwise dependence measure δ such as empirical tail dependence or Kendall's τ, where the latter will be used in the following due to the fact that it captures the general dependence among two variables. For R-vine tree selection, we then find a tree on all nodes, a so-called *spanning tree*, which maximizes the sum of (absolute) pairwise dependencies. Such a tree usually is referred to as *maximum spanning tree*. We hence solve the following optimization problem

$$\max \sum_{\substack{\text{edges } e=\{i,j\} \text{ in} \\ \text{spanning tree}}} |\delta_{ij}|,$$

where the absolute values of the pairwise dependencies $\delta_{ij},\ i \neq j$, are used, since it is also important to model strong negative dependencies as typically indicated by negative values of δ. Similarly for C-vine tree selection, we choose as root node the node that maximizes the sum of pairwise dependencies to this node (see Czado et al. (2011)).

Given the selected tree, we then choose pair copulas from a range of ten different families which are shown in Table 2 with their properties (whether or not they can model positive and negative dependence, asymmetric dependence and lower and upper tail dependence; see Joe (1997) and Nelsen (2006)). As selection criterion we use the AIC of Akaike (1973) because it turned out to perform particularly well in a simulation study (Brechmann 2010, Chapter 10). Moreover we always perform a bivariate independence test first to obtain parsimonious models. Parameters for copulas are estimated using bivariate maximum likelihood estimation.

In the next step, we compute transformed observations $F(x|\boldsymbol{v})$ from the estimated pair copulas using formula (3.5). These are then used as input parameters for the next trees, which are obtained similarly by constructing a graph according to the above R-vine construction principles, in particular the proximity condition, and then finding a maximum dependence tree or the next root node.

	N	t	C	G	F	J	BB1	BB7	RC	RG
positive dep.	✓	✓	✓	✓	✓	✓	✓	✓	-	-
negative dep.	✓	✓	-	-	✓	-	-	-	✓	✓
asymmetric dep.	-	-	✓	✓	-	✓	✓	✓	✓	✓
lower tail dep.	-	✓	✓	-	-	-	✓	✓	-	-
upper tail dep.	-	✓	-	✓	-	✓	✓	✓	-	-

Table 2: Bivariate copula families and their properties. Notation of copula families: N = Gaussian, t = Student-t, C = Clayton, G = Gumbel, F = Frank, J = Joe, BB1 = Clayton-Gumbel, BB7 = Joe-Clayton, RC = rotated Clayton (90°), RG = rotated Gumbel (90°).

The treewise selection and estimation procedure described here gives sequential estimates of pair copula parameters which are quite quickly obtained and can be used as starting values for a full maximum likelihood estimation (see Aas et al. (2009) and Hobæk Haff (2010)). This way of proceeding selects the vine tree structure simultaneously together with pair copula families and their parameter values. In particular pair copula families are selected individually and the corresponding parameter estimation only requires bivariate optimization. Usually the sequential estimation procedure even provides very good estimates and they therefore can be seen as reasonable approximations to maximum likelihood estimates for the selected tree structure and selected pair copula families (see Brechmann (2010), Hobæk Haff (2011) and the discussion below). Especially in higher dimensions, as for example needed when modeling the cross-section of returns, full maximum likelihood estimation is computationally extremely demanding even if algorithms are strongly parallelized.

3.3 Application to data

We now investigate the dependence structure of the Euro Stoxx 50 members using R- and C-vines. For this we consider daily log returns of the indices and stocks described in Section 2 over the 4-year period from May 22, 2006 to April 29, 2010 resulting in 985 observations. In terms of complexity this means that we are dealing with a 52-dimensional data set (46 stocks, 5 national indices and the Euro Stoxx 50 index). This dimensionality is far beyond vine copula applications that can be found in the literature so far (see Brechmann et al. (2010) and Dißmann et al. (2011)). Heinen and Valdesogo (2009) analyze 95 financial returns using C-vines. They however do not provide maximum likelihood estimates so that it is not clear how reliable their estimates are.

As noted above, we preliminarily find appropriate models for the univariate time series before analyzing the dependence of the standardized residuals. The selection process is described in Appendix A.

We then fit appropriate R- and C-vines as described in Section 3.2. The order of the first seven C-vine root nodes is as follows: ^STOXX50E, GLE.PA, ^FCHI, ^GDAXIP, ^IBEX, INGA.AS, FTSEMIB.MI; the Dutch AEX was only identified as the 13th root node. The first tree of the R-vine is shown in Figure 2.

The order of the C-vine root nodes and the first R-vine tree show that the Dutch AEX only plays a minor role due to its small size. It is the only national index which is not directly connected to the Euro Stoxx 50 which constitutes some kind of "root node" of the R-vine. Further, the Euro Stoxx 50 and the national leading stock indices are identified as main dependency drivers, where the importance of the CAC 40 and the DAX is due to the high number

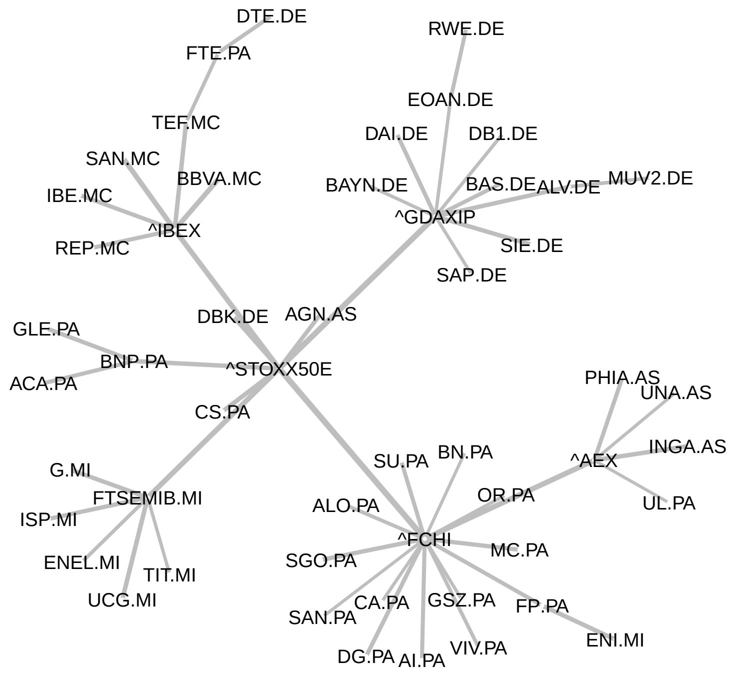

Figure 2: First tree of the R-vine.

of French and German stocks represented in the Euro Stoxx 50 (weights of 35.3% and 28.1%, respectively). The financial companies Société Générale (GLE.PA) and ING Groep (INGA.AS) also play a central role, which corresponds to the fact that 28.2% of the Euro Stoxx 50 are composed of financial institutions.

The (sequential) fit of the R- and C-vine models is compared to a multivariate Student-t copula with one common degrees of freedom parameter. It serves as a benchmark model, since it is currently the state-of-the-art approach for modeling financial returns (see, e.g., Mashal and Zeevi (2002) and Breymann et al. (2003)). The parameters of the Student-t copula are estimated in two steps (Lindskog et al. 2003). First the correlation matrix of the Student-t copula is determined using the closed form correspondence to bivariate Kendall's τ's and then the degrees of freedom parameter was found using maximum likelihood estimation. Results are reported in Table 3 which also gives statistics of Vuong (1989) tests for non-nested model comparison between the R- and C-vines as well as the Student-t copula. Vuong test statistics are asymptotically standard normal so that model discrimination can be based on standard normal quantiles.

While both vine copulas are clearly superior to the Student-t copula, the R-vine is also superior to the C-vine, especially when taking into account the number of model parameters which is clearly less for the R-vine with its less restrictive structure.

We also performed joint maximum likelihood estimation of all copula parameters to check that sequential estimation provides adequate results. The increase in the log likelihood was only about 0.3% so that we can conclude that it is reasonable to work with sequential estimates here.

To summarize these first results, the C- and especially the R-vine model show a good fit to the data and provide a directly interpretable dependence structure which corresponds very well to the economical intuition. In particular, the R-vine identifies the apparent sectorial dependence present in the data. A more explicit modeling of this sectorial structure is discussed in the

Copula	Log likelihood	No. of param.	BIC	Vuong test w.r.t. R-vine w/o corr.	Vuong test w.r.t. R-vine Schwarz corr.	Vuong test w.r.t. Student-t w/o corr.	Vuong test w.r.t. Student-t Schwarz corr.
R-vine	30777.84	596	-57447.66	-	-	*-3.80**	*-24.42**
C-vine	30661.14	685	-56600.83	*1.33**	4.81	*-3.45**	*-25.37**
Student-t	30313.23	1327	-51479.92	3.80	24.42	-	-

Table 3: Log likelihoods, number of copula model parameters and BICs of the R- and C-vine as well as Student-t copula fits. Estimates are obtained using sequential estimation. Vuong test statistics (with and without Schwarz correction for the number of model parameters used) indicated by asterisks imply that the considered model is indistinguishable from or superior to the R-vine (columns 5 and 6) or Student-t copula (columns 7 and 8), respectively, at the 5% level.

following section.

4 Vine market sector models

Motivated by the results of the previous section we develop a new market sector model based on vines. It is inspired by the famous classical CAPM of Sharpe (1964) and Lintner (1965). The CAPM essentially assumes that at time t the individual asset returns $r_{t,j}$, the market return $r_{t,M}$ and the idiosyncratic error terms $\varepsilon_{t,j}$, which are independent of $\varepsilon_{t-1,j}$ and of $\varepsilon_{t,k}$ $\forall k \neq j$, are jointly normally distributed and follow the linear relationship

$$r_{t,j} = \beta_j r_{t,M} + \varepsilon_{t,j}, \tag{4.1}$$

where β_j is usually called the *sensitivity* of asset j to the market.

The CAPM belongs to a class of more general models, so-called *factor models*, of which the CAPM is the simplest with the market being the only factor. The most famous multi-factor model certainly is the three-factor model of Fama and French (1992). Their model uses three factors in order to take into account deviations from the CAPM for companies with small market capitalization (*small caps*) and with high book-to-market ratio (*value stocks*). Other multi-factor models include factors such as industry sectors or currencies (see, e.g., Drummen and Zimmermann (1992)).

As pointed out in the introduction the restrictive assumptions of the CAPM are inappropriate, not to say wrong. Heinen and Valdesogo (2009) addressed these issues and developed a non-linear and non-Gaussian extension of the CAPM which they called *Canonical Vine Autoregressive* (CAVA) model, where "autoregressive" refers to the fact that the involved copulas are possibly time-varying.

We first review their model before introducing our new Regular Vine Market Sector model.

4.1 Heinen and Valdesogo's Canonical Vine Autoregressive model

Heinen and Valdesogo (2009) loosen the normality and linearity assumptions of the CAPM specified in (4.1) by using a variety of GARCH-models for the marginal time series of financial returns and by modeling the residual dependence between *assets* and the *market* with bivariate copulas for the standardized residuals. This obviously corresponds to the first tree of a C-vine with the market as root node. Each *asset* is further assumed to depend on the *market* and on its *sector* (e.g., utilities or financial services). To fit this two-factor model to a C-vine structure with the market and the sectors as root nodes, this model induces independence assumptions: conditionally on the market, sectorial returns are assumed to be independent and asset returns independent of sector returns other than their own. The remaining dependence of asset returns conditioned on the market and on the respective sectors is captured with a multivariate Gaussian copula which is shown to give a valid copula model by Valdesogo (2009). In the following, we will refer to this model as *Canonical Vine Market Sector (CVMS) model* in order to highlight the underlying model structure. It is illustrated in the following example (see Heinen and Valdesogo (2009)).

Example 4.1 (CVMS model.). *Let r_1^A, r_2^A, r_1^B and r_2^B denote the (standardized residuals of) returns of stocks belonging to sectors A and B, respectively. Further, let r_A and r_B be those of the sectors A and B as well as r_M of the market. According to the CVMS model, the following independence assumptions hold:*

(i) r_A is independent of r_B conditioned on r_M, and

(ii) r_1^A and r_2^A are independent of r_B conditioned on r_M, while r_1^B and r_2^B are independent of r_A conditioned on r_M.

In terms of a C-vine structure we now have the market as the first root node and the sectors as second and third root nodes, where the order is arbitrary due to the independence assumptions. Let sector A be the second and B the third root node.

The copula terms of the first C-vine tree T_1 with the market as the root node are

$$
\begin{aligned}
&c_{1A,M}\left(F(r_1^A),F(r_M)\right),c_{2A,M}\left(F(r_2^A),F(r_M)\right),c_{1B,M}\left(F(r_1^B),F(r_M)\right),\\
&c_{2B,M}\left(F(r_2^B),F(r_M)\right),c_{A,M}\left(F(r_A),F(r_M)\right),c_{B,M}\left(F(r_B),F(r_M)\right).
\end{aligned}
$$

Since A is the second root node, the pair copulas of the second tree T_2 are given by

$$
\begin{aligned}
&c_{1A,A|M}\left(F(r_1^A|r_M),F(r_A|r_M)\right),c_{2A,A|M}\left(F(r_2^A|r_M),F(r_A|r_M)\right),\\
&\underbrace{c_{1B,A|M}\left(F(r_1^B|r_M),F(r_A|r_M)\right)}_{\overset{(ii)}{=}1},\underbrace{c_{2B,A|M}\left(F(r_2^B|r_M),F(r_A|r_M)\right)}_{\overset{(ii)}{=}1},
\end{aligned}
$$

according to independence assumption (ii) (the density of the independence copula is of course 1) and $c_{A,B|M}\left(F(r_A|r_M),F(r_B|r_M)\right)=1$ because of the first independence assumption. Further, the copulas of the third tree T_3 are

$$
\underbrace{c_{1A,B|M,A}\left(F(r_1^A|r_M,r_A),F(r_B|r_M,r_A)\right)}_{\overset{(ii)}{=}1},\underbrace{c_{2A,B|M,A}\left(F(r_2^A|r_M,r_A),F(r_B|r_M,r_A)\right)}_{\overset{(ii)}{=}1},
$$

as well as

$$
c_{1B,B|M,A}\big(\underbrace{F(r_1^B|r_M,r_A)}_{\overset{(ii)}{=}F(r_1^B|r_M)},\underbrace{F(r_B|r_M,r_A)}_{\overset{(i)}{=}F(r_B|r_M)}\big)=c_{1B,B|M}\left(F(r_1^B|r_M),F(r_B|r_M)\right),
$$

and similarly for $c_{2B,B|M,A}$ using both independence assumptions.

Finally, we have a four-dimensional Gaussian copula $c^{\rho}_{1A,1B,2A,2B|M,A,B}$ with arguments $F(r_i^A|r_M,r_A,r_B)=F(r_i^A|r_M,r_A)$ and $F(r_i^B|r_M,r_A,r_B)=F(r_i^B|r_M,r_B)$ for $i=1,2$ due to independence assumption (ii).

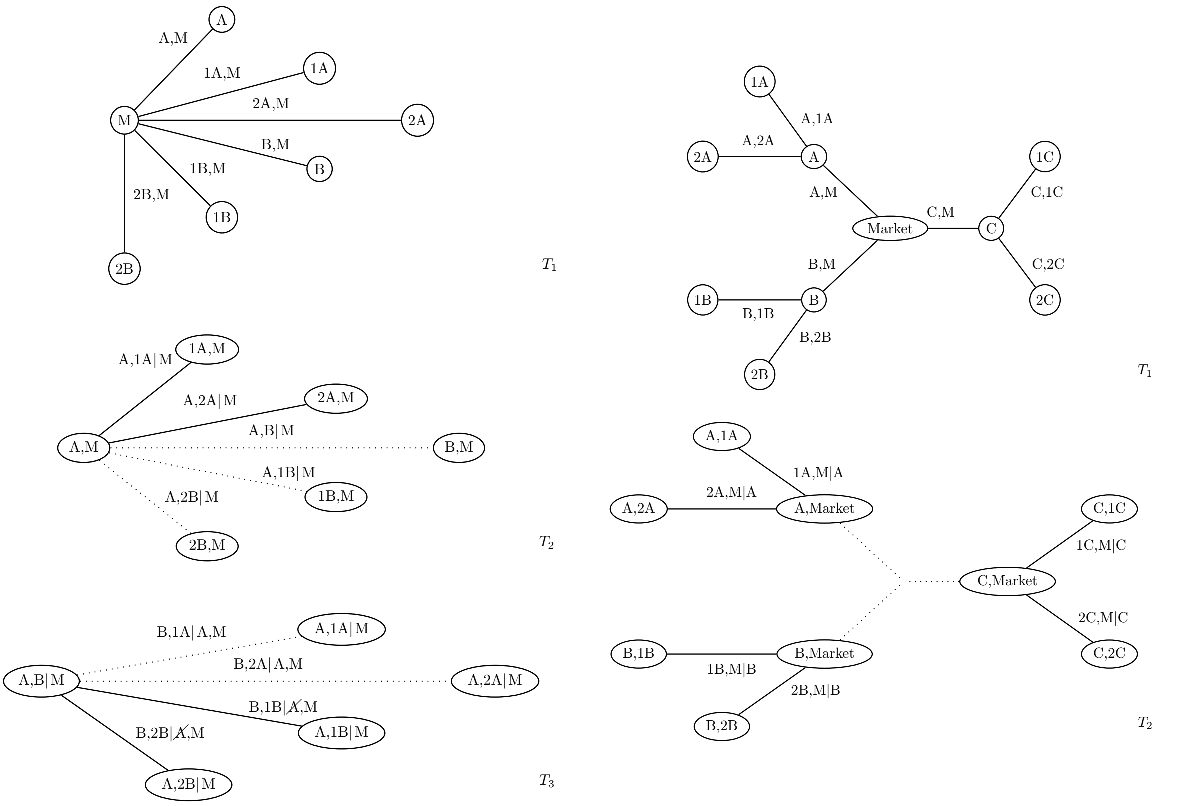

Figure 3: Left panel: first three trees of the CVMS model in Example 4.1. Right panel: first and second tree of the RVMS model in Example 4.2.

Then the joint density of the (standardized residuals of) returns r_1^A, r_2^A, r_1^B, r_2^B, r_A, r_B *and* r_M *is given by*

$$
\begin{aligned}
f(r_1^A, r_2^A, r_1^B, r_2^B, r_A, r_B, r_M) = {} & f(r_1^A) f(r_2^A) f(r_1^B) f(r_2^B) f(r_A) f(r_B) f(r_M) \\
& \times c_{1A,M}\left(F(r_1^A), F(r_M)\right) c_{2A,M}\left(F(r_2^A), F(r_M)\right) \\
& \times c_{1B,M}\left(F(r_1^B), F(r_M)\right) c_{2B,M}\left(F(r_2^B), F(r_M)\right) \\
& \times c_{A,M}\left(F(r_A), F(r_M)\right) c_{B,M}\left(F(r_B), F(r_M)\right) \\
& \times c_{1A,A|M}\left(F(r_1^A|r_M), F(r_A|r_M)\right) c_{2A,A|M}\left(F(r_2^A|r_M), F(r_A|r_M)\right) \\
& \times c_{1B,B|M}\left(F(r_1^B|r_M), F(r_B|r_M)\right) c_{2B,B|M}\left(F(r_2^B|r_M), F(r_B|r_M)\right) \\
& \times c^{\rho}_{1A,1B,2A,2B|M,A,B}\left(F(r_1^A|r_M, r_A), ..., F(r_2^B|r_M, r_B)\right).
\end{aligned}
$$

The first three trees of the C-vine are shown in the left panel of Figure 3. Dotted lines illustrate the independence assumptions, i.e., independence copulas are chosen for the copulas corresponding to dotted edges.

4.2 The Regular Vine Market Sector model

The strong assumptions of the CVMS model due to its underlying restrictive C-vine structure and the more general and very flexible structure of R-vines motivate the construction of an alternative model for the standardized residuals of appropriately chosen marginal GARCH-models.

Clearly, we expect that there are strong relationships between the returns of a *stock* and the *sector* it belongs to. In contrast to the CVMS model, we thus model these dependencies first as well as the dependencies of the *sectors* to the *market* to take into account the joint driver of dependencies among sectors. If all remaining dependencies are then captured by Gaussian pair copulas in higher order trees, we speak of the *Regular Vine Sector (RVS) model.* If however the dependencies to the *market* are also modeled conditionally on the respective sectors in the second R-vine tree before setting all pair copulas of higher order trees to bivariate Gaussian copulas, we call the model the *Regular Vine Market Sector (RVMS) model.* As in the CVMS model, we furthermore assume that sectors are independent conditioned on the market in the RVMS model. This independence assumption is however only made for convenience to avoid the selection of a tree structure among indices conditioned on the market. Moreover, note that the construction of higher order trees is not uniquely determined by the first or second tree in an RVS or RVMS model, respectively. We solve this problem by simply modeling the strongest dependencies in each tree as described in Section 3.2 and where "dependency" usually refers to Kendall's τ. Modeling using a multivariate Gaussian copula as in the CVMS model is not possible here due to the more general underlying R-vine structure (see Brechmann et al. (2010)). Examples of both models, the one-factor RVS model and the two-factor RVMS model, are given in the following.

Example 4.2 (RVS and RVMS models.). *Similar to Example 4.1 we consider the (standardized residuals of) returns r_1^A, r_2^A, r_1^B, r_2^B, r_1^C and r_2^C of stocks belonging to sectors A, B and C with sector returns r_A, r_B and r_C, respectively. Furthermore, r_M denotes the market return.*

Then the first R-vine tree in the RVS as well as the RVMS model is specified with appropriate pair copulas as

$$\begin{aligned}
&c_{1A,A}\left(F(r_1^A),F(r_A)\right)c_{2A,A}\left(F(r_2^A),F(r_A)\right)c_{1B,B}\left(F(r_1^B),F(r_B)\right)\\
&\times c_{2B,B}\left(F(r_2^B),F(r_B)\right)c_{1C,C}\left(F(r_1^C),F(r_C)\right)c_{2C,C}\left(F(r_2^C),F(r_C)\right)\\
&\times c_{A,M}\left(F(r_A),F(r_M)\right)c_{B,M}\left(F(r_B),F(r_M)\right)c_{C,M}\left(F(r_C),F(r_C)\right).
\end{aligned}$$

In the RVS model, all pair copulas of higher order trees are then set to bivariate Gaussian copulas. The RVMS model, on the other hand, also models the second R-vine tree under the assumption that sectors are independent conditionally on the market, i.e.,

$$c_{A,B|M}\left(F(r_A|r_M),F(r_B|r_M)\right)=1,$$

and similarly for A and C as well as for B and C (see independence assumption (i) in Example 4.1). Hence, the pair copula terms of the second tree of the RVMS model are

$$\begin{aligned}
&c_{1A,M|A}\left(F(r_1^A|r_A),F(r_M|r_A)\right)c_{2A,M|A}\left(F(r_2^A|r_A),F(r_M|r_A)\right)\\
&\times c_{1B,M|B}\left(F(r_1^B|r_B),F(r_M|r_B)\right)c_{2B,M|B}\left(F(r_2^B|r_B),F(r_M|r_B)\right)\\
&\times c_{1C,M|C}\left(F(r_1^C|r_C),F(r_M|r_C)\right)c_{2C,M|C}\left(F(r_2^C|r_C),F(r_M|r_C)\right).
\end{aligned}$$

Without the above independence assumptions, we would have to choose a tree structure for the sector variables, since the pair copulas $c_{A,B|M}$, $c_{A,C|M}$ and $c_{B,C|M}$ cannot all be included in an R-vine at once. Due to the independence assumptions, the model is independent of this choice. All higher order trees of the RVMS model are then specified with Gaussian pair copulas.

Figure 3 shows the first tree of the RVS and the RVMS models as well as the second tree of the RVMS model. The dotted lines in the second tree illustrate the independence assumption for sectors conditionally on the market.

Before analyzing the Euro Stoxx 50 using vine market sector models, we first discuss the decomposition into systematic and idiosyncratic risk of asset returns using the RVMS model, since it is central to any factor model.

4.3 Separation of systematic and idiosyncratic risk

Given the rather simple CAPM equation (4.1) it is straightforward to separate the systematic market risk and the idiosyncratic risk of the return of an asset j by

$$\sigma_j^2 = \beta_j^2 \sigma_M^2 + \sigma_\varepsilon^2, \tag{4.2}$$

where σ_j^2, σ_M^2 and σ_ε^2 denote the variances of the asset return, the market return and the idiosyncratic error term, respectively.

Such a decomposition is due to the assumptions of linearity and joint normality. For our non-linear and non-Gaussian RVMS model it is therefore not clear how to obtain a similar result. We would nevertheless like to present some ideas regarding this issue in order to facilitate the interpretation of results. First note that we separate marginal and dependency modeling by first selecting appropriate GARCH-models for the marginal time series. Hence, dependencies are modeled not directly among returns but among standardized residuals of them. All following statements are therefore based on the residuals rather than on the original data. Moreover, variances are not constant when using a GARCH-model but time-dependent. This also prohibits an expression as in (4.2).

A central—and rather restrictive—assumption of the CAPM is joint normality. In the RVMS model this is not required. The copulas used here are able to capture more complex dependency structures than a simple product-moment correlation coefficient as it is the case for Gaussian random variables. Examples of such structures allow for asymmetric and tail dependent dependencies. Hence the dependence structure of a particular asset j in the RVMS model cannot simply be summarized in one single number such as the sensitivity β_j in the CAPM. A copula contains much more information and this should not be neglected. In particular, the consideration of joint tail behavior is crucial when dealing with financial returns.

In order to separate market, sectorial and idiosyncratic risk it is however useful to consider the following: choose a dependence measure such as the product-moment correlation, Kendall's association measure τ, Spearman's rank correlation ρ or Blomqvist's medial correlation β (see Nelsen (2006) for more details) and compute it for each pair copula in the RVMS model, which involves asset j in the conditioned set, i.e., we consider the conditional dependence of asset j to all other assets and sectors as identified by the R-vine structure. Then one obtains an indication how strong the (conditional) dependencies to the market M, the respective sector S of asset j and the other assets and sectors $k_{j_1}, ..., k_{j_m}$ are (here m denotes the number of assets and sectors other than j and S). Let δ be the chosen dependence measure, the fraction

$$\mathcal{R}_{j,S}(\delta) = \frac{|\delta_{j,S}|}{|\delta_{j,S}| + |\delta_{j,M|S}| + |\delta_{j,k_{j_1}|S,M}| + ... + |\delta_{j,k_{j_m}|S,M,k_{j_1},...,k_{j_{m-1}}}|} \tag{4.3}$$

then gives an indication how strong the *sectorial risk* is compared to the market and idiosyncratic risk for asset j. The conditional pairs $(j, k_{j_1}|S, M), ..., (j, k_{j_m}|S, M, k_{j_1}, ..., k_{j_{m-1}})$ are identified by the used R-vine structure. We take absolute values, since dependence measures can also take negative values and hence cancel out in the denominator which would be counterintuitive.

Similar to (4.3) we can define the *market risk conditionally on the sector*, i.e., with all dependence on the sector being removed, as

$$\mathcal{R}_{j,M|S}(\delta) = \frac{|\delta_{j,M|S}|}{|\delta_{j,S}| + |\delta_{j,M|S}| + |\delta_{j,k_{j_1}|S,M}| + \ldots + |\delta_{j,k_{j_m}|S,M,k_{j_1},\ldots,k_{j_{m-1}}}|}$$

and the *idiosyncratic risk* as $\mathcal{R}_{j,\varepsilon}(\delta) = 1 - \mathcal{R}_{j,S}(\delta) - \mathcal{R}_{j,M|S}(\delta)$, where $\mathcal{R}_{j,S}(\delta) + \mathcal{R}_{j,M|S}(\delta)$ is interpreted as *systematic risk*. Note that the measures in the denominator could also be weighted to take into account the conditioning.

Alternatively, the measures $\delta_{j,S}$ and $\delta_{j,M|S}$ may of course also be considered separately, then giving the absolute dependence on the respective sector and on the market conditionally on the sector. In particular for lower and upper tail dependence coefficients λ^ℓ and λ^u, respectively, this is is sensible because the joint tail behavior is often rather weak in higher order trees and then $\mathcal{R}_{j,S}(\lambda^\ell)$ and $\mathcal{R}_{j,M|S}(\lambda^\ell)$ (and similar for λ^u) would have only limited explanatory power.

4.4 Application to data

As indicated by the results in Section 3.3 vine market sector models may be very useful to analyze the Euro Stoxx 50. Since the independence assumptions of the CVMS model are quite restrictive, we first carefully investigate these assumptions and assess whether they are appropriate. Further we compare the fitted vine market sector models to the fitted copula models from Section 3.3 and analyze systematic and idiosyncratic risk of the stocks. When constructing the models, we will however not consider dynamical pair copula terms in order to limit the model complexity to a reasonable degree. Time-varying dependency effects for the Regular Vine Market Sector model defined below are considered in Section 5.1 when we discuss Value-at-Risk forecasting using rolling window estimates.

4.4.1 Independence assumptions

The CVMS model considered in Section 4.1 crucially relies on independence assumptions to construct a C-vine structure. Here these assumptions can be explicitly stated as:

(i) the returns of the national indices conditioned on the market return, i.e., on the return of the Euro Stoxx 50, are independent and

(ii) conditionally on the market return, the individual returns are independent of returns for national indices other than their own, e.g., returns of German stocks are independent of the French CAC 40 conditioned on the Euro Stoxx 50.

The first independence assumption also applies to the RVMS model, while the second is not needed there. This however is the crucial difference between both models as shown in the following.

While the first root node of the CVMS model clearly is ^STOXX50E, the order of root nodes 2 to 6 is arbitrary (see Example 4.1). We choose the order according to the number of stocks under consideration (see Table 1), i.e., we choose ^FCHI as second, ^GDAXIP as third, FTSEMIB.MI as fourth, ^IBEX as fifth and finally ^AEX as sixth root node. This also corresponds to the order selected for the C-vine fit in Section 3.3.

Independence assumption (ii) (for the stocks) is investigated in the columns of Table 4 which displays the percentage of rejections of the independence hypothesis using the bivariate

	ˆFCHI	ˆGDAXIP	FTSEMIB.MI	ˆIBEX	ˆAEX
French	-	42%	47%	32%	58%
German	33%	-	42%	33%	33%
Italian	83%	50%	-	33%	33%
Spanish	80%	100%	60%	-	60%
Dutch	50%	25%	50%	50%	-
total	56%	50%	47%	34%	48%

Table 4: Percentages of rejection of the independence hypothesis for stocks from each country versus indices of countries other than their own conditioned on ˆSTOXX50E.

	ˆFCHI	ˆGDAXIP	FTSEMIB.MI	ˆIBEX
ˆGDAXIP	*0.26**			
FTSEMIB.MI	*0.91**	0.00		
ˆIBEX	0.00	0.00	*0.32**	
ˆAEX	0.00	*0.86**	*0.05**	0.00

Table 5: p-values of bivariate independence tests for national stock indices conditioned on ˆSTOXX50E. p-values indicated by "*" imply that the independence hypothesis cannot be rejected at the 5% level in accordance with the assumptions of the CVMS and RVMS models.

independence test based on Kendall's τ (see, e.g., Genest and Favre (2007)) for variable pairs involving the respective national index, e.g., the null hypothesis of independence of a specific German stock and ˆFCHI conditioned on ˆSTOXX50E is rejected for 33% of the German stocks, where the CVMS model assumes 0%). p-values of independence tests between the national indices in order to check the first independence assumption are shown in Table 5.

Evidently, independence assumptions (i) and (ii) do not hold in general. In particular, conditionally on the market, i.e., the Euro Stoxx 50, there is significant dependence of Italian and Spanish stocks on the French CAC 40 as well as dependence of all Spanish stocks on the German DAX, which contradicts independence assumption (ii). In total, the assumption is wrong for about 50% of all variable pairs under consideration. Moreover, the null hypothesis of independence is pairwisely rejected for the Spanish IBEX 35, the Dutch AEX and the CAC 40 as well as for the Italian FTSE MIB and the IBEX 35 with respect to the DAX in contradiction to independence assumption (i).

While the assumptions of the CVMS model are clearly incorrect here, the assumption of the RVMS model that sectors are pairwisely independent is not true here either (the hypothesis is rejected for 50% of the pairs). However, in contrast to the CVMS model, we can circumvent the problem, since not all bivariate copulas for two sectors conditioned on the market can be included in the model so that it remains a valid R-vine (*tree* structure!). If we choose the pairs ˆAEX–ˆGDAXIP, ˆGDAXIP–ˆFCHI, ˆFCHI–FTSEMIB.MI and FTSEMIB.MI–ˆIBEX, we obtain a valid density and the independence assumption is satisfied.

4.4.2 Model comparisons

We now compare the CVMS model as well as the RVMS and the RVS models to the R- and C-vine models fitted in Section 3.3. For comparison to the first R-vine tree shown in Figure 2, the first tree of the RVMS and RVS models is shown in Figure 4.

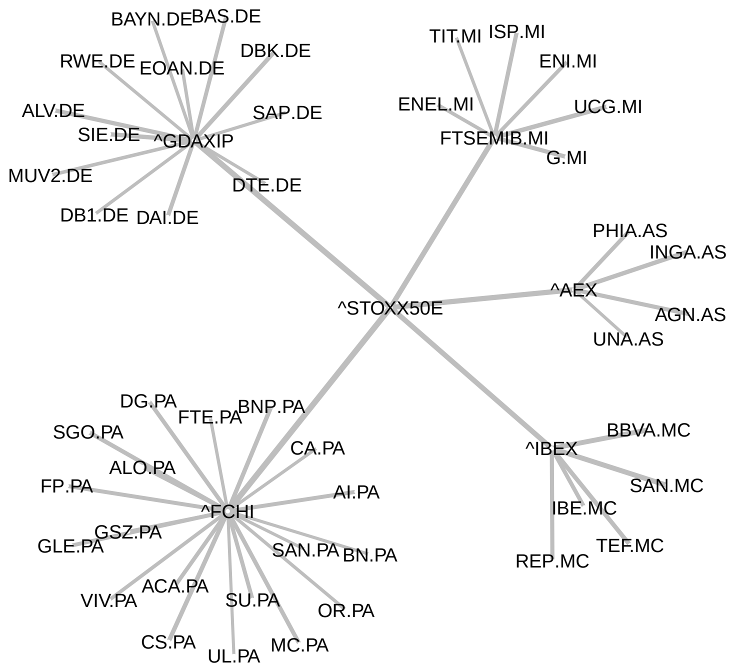

Figure 4: First tree of the RVMS and RVS models.

As indicated by the problematic independence assumptions, there are important dependencies which are neglected in the first trees of the CVMS model. It then aims at capturing all remaining dependencies with a multivariate Gaussian copula. This however means that the transformed variables (3.5) obtained from the pair copulas of tree T_6 are jointly normal which is a rather strict assumption given that the data might exhibit, e.g., asymmetric dependence or strong joint tail behavior. Hence, we used the copula goodness-of-fit test based on Rosenblatt's transformation (Breymann et al. 2003) to investigate the hypothesis of joint normality, since this test performs particularly well especially against heavy tailed alternatives as they often occur in financial applications (Berg 2009) and since it is computationally less demanding than alternative multivariate tests. Here, the test clearly rejected the null hypothesis of normality with an approximate p-value of 0.00, i.e., a multivariate Gaussian copula is not appropriate which again contradicts the assumptions of the CVMS model.

The use of Gaussian pair copulas in higher order trees of the RVMS and the RVS models cannot be validated that easily. We can however employ the heuristic procedure developed by Brechmann et al. (2010) to investigate an appropriate vine tree after which remaining pair-copulas can be set to Gaussian. Using its parsimonious version based on the Vuong test with Schwarz correction confirms the assumption of the RVMS model, while Gaussian copulas already after the first tree as in the RVS model seem inappropriate.

Table 6 compares the vine market sector models to the R- and C-vine as well as the Student-t copula fitted in Section 3.3. As before log likelihoods obtained by sequential estimation, BICs and Vuong tests with respect to the R-vine and the Student-t copulas are shown. For reasons of comparison, we also consider an alternative specification of the CVMS model: instead of using a multivariate Gaussian copula for the CVMS model, we fit all remaining C-vine trees with Gaussian pair copulas as in the RVMS models which allows for bivariate independence tests for the pairs of transformed variables (3.5) in each tree. The resulting model is denoted as CVMS-P

Copula	Log likelihood	No. of param.	BIC	Vuong test w.r.t. R-vine w/o corr.	Vuong test w.r.t. R-vine Schwarz corr.	Vuong test w.r.t. Student-t w/o corr.	Vuong test w.r.t. Student-t Schwarz corr.
R-vine	30777.84	596	-57447.66	-	-	*-3.80**	*-24.42**
C-vine	30661.14	685	-56600.83	*1.33**	4.81	*-3.45**	*-25.37**
Studen-t	30313.23	1327	-51479.92	3.80	24.42	-	-
CVMS	30309.46	1212	-52265.05	4.75	26.27	*0.04**	*-4.06**
CVMS-P	29861.63	495	-56311.40	8.93	5.54	4.23	*-22.65**
RVS	29964.20	506	-56440.71	7.50	4.64	3.26	*-23.16**
RVMS	30126.28	535	-56564.99	6.52	4.41	*1.85**	*-25.21**

Table 6: Log likelihoods, number of copula model parameters and BICs of the R- and C-vine, the Student-t copula and the vine market sector model fits. Estimates are obtained using sequential estimation. Vuong test statistics (with and without Schwarz correction for the number of model parameters used) indicated by asterisks imply that the considered model is indistinguishable from or superior to the R-vine (columns 5 and 6) or Studen-t copula (columns 7 and 8), respectively, at the 5% level.

to indicate the pairwise modeling.

In contrast to Heinen and Valdesogo (2009), we also computed the maximum likelihood copula parameter estimates of our new vine market sector model, the RVMS model. As before for the general R-vine, the increase in the log likelihood compared to sequential estimation is very small (about 0.1%), so that BIC values and Vuong test results barely change and confirm the corresponding results obtained from the sequential estimation procedure. Moreover, the full maximum likelihood estimates provide the basis for bootstrapping to obtain confidence intervals.

To summarize, the CVMS model of Heinen and Valdesogo (2009) is inappropriate to explain the dependence of the Euro Stoxx 50. It requires a lot of parameters, since a 46-dimensional correlation matrix has to be fitted for the remaining dependency after the sixth C-vine tree—which is hardly feasible if maximum likelihood estimation is desired instead of the inversion of Kendall's τ. The over-specification leads to a high log likelihood, but taking into account the number of model parameters, the model fit is clearly inferior to the competitive models. Moreover, the model assumptions of the CVMS model were shown to be not satisfied. Although the CVMS-P improves in some regard, in particular in terms of parsimonity, R-vine based models are still superior. Again taking into account the number of model parameters, the current state-of-the-art dependence model, the Student-t copula, is however the worst model and can be regarded as inferior to all vine copula models.

4.4.3 Systematic and idiosyncratic risk

The RVMS model can be used to separate systematic and idiosyncratic risk of stocks as discussed in Section 4.3. We therefore computed the sectorial risk $\mathcal{R}_{j,S}$, i.e., the risk captured by the national index, and conditionally on the respective sector the market risk $\mathcal{R}_{j,M|S}$ for each stock j. As dependence measure we use Kendall's τ and product-moment correlations which we obtained from an RVMS model with only Gaussian pair copula. The results for the stocks from each country are shown in Figure 5.

Apparently sectorial risk is quite variable across different stocks. It can however be stated that the mean sectorial risk among German stocks is higher than that of the other four countries

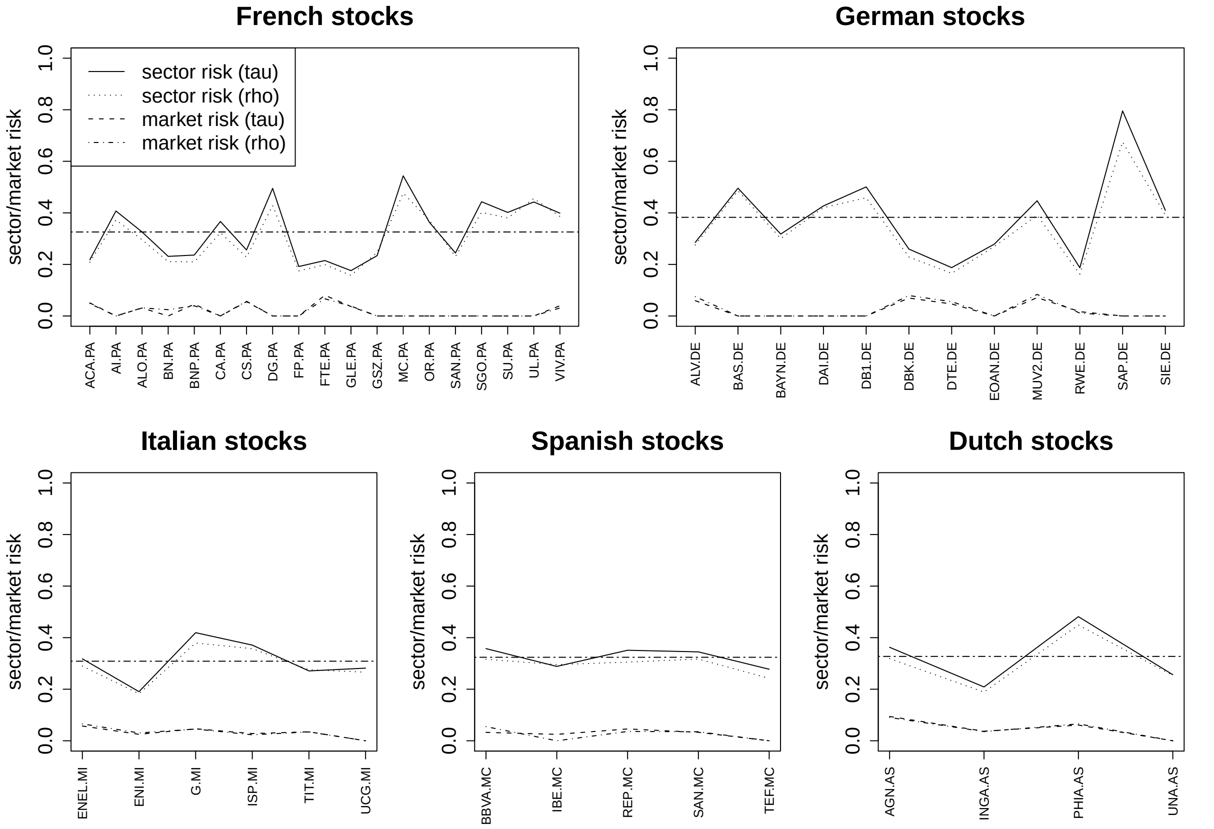

Figure 5: Sectorial and market risk of the stocks from each country in the Euro Stoxx 50 data according to the general RVMS model ($\mathcal{R}_{j,s}(\tau)$, $\mathcal{R}_{j,M|S}(\tau)$) and an RVMS model with only Gaussian copulas ($\mathcal{R}_{j,s}(\rho)$, $\mathcal{R}_{j,M|S}(\rho)$). Horizontal lines indicate the mean sectorial risk $\mathcal{R}_{j,s}(\tau)$.

which is approximately similar, that is the German stock market is evidently more strongly self-contained. Conditionally on the respective sectors, the market risk is very small for all stocks. When we fitted copulas, we always investigated if variables were independent and hence the market risk is even zero for some stocks, where the test detected independence between a stock and the Euro Stoxx 50. The rather high idionsyncratic risk of most stocks can be explained by major dependencies on other stocks as identified, e.g., in the first R-vine tree shown in Figure 2. Moreover, these results correspond to the analysis of Drummen and Zimmermann (1992) who, using a multi-factor model, find that the country effect is much stronger than that of the European market—which also confirms the assumed structure of the RVMS model—, while the proportion of idiosyncratic risk is about 50% even when more factors are used than in our model. Finally, $\mathcal{R}_{j,S}(\delta)$ and $\mathcal{R}_{j,M|S}(\delta)$ appear to be rather robust measures in terms of the chosen dependence measure δ, since the results using Kendall's τ and the product-moment correlation are very similar.

Furthermore, we also used the RVMS model to obtain lower and upper tail dependence coefficients with respect to the sector $\lambda^{\ell}_{j,S}$ and $\lambda^{u}_{j,S}$ and to the market $\lambda^{\ell}_{j,M|S}$ and $\lambda^{u}_{j,M|S}$ for each stock j in order to assess the strength of the joint tail behavior. Note that $\lambda^{\ell}_{j,k_{j_r}|S,M,k_{j_1},\ldots,k_{j_{r-1}}}$ and $\lambda^{u}_{j,k_{j_r}|S,M,k_{j_1},\ldots,k_{j_{r-1}}}$ from Equation (4.3) are often zero and hence the corresponding risk measures $\mathcal{R}_{j,S}$ and $\mathcal{R}_{j,M|S}$ have only limited information value and are not considered here.

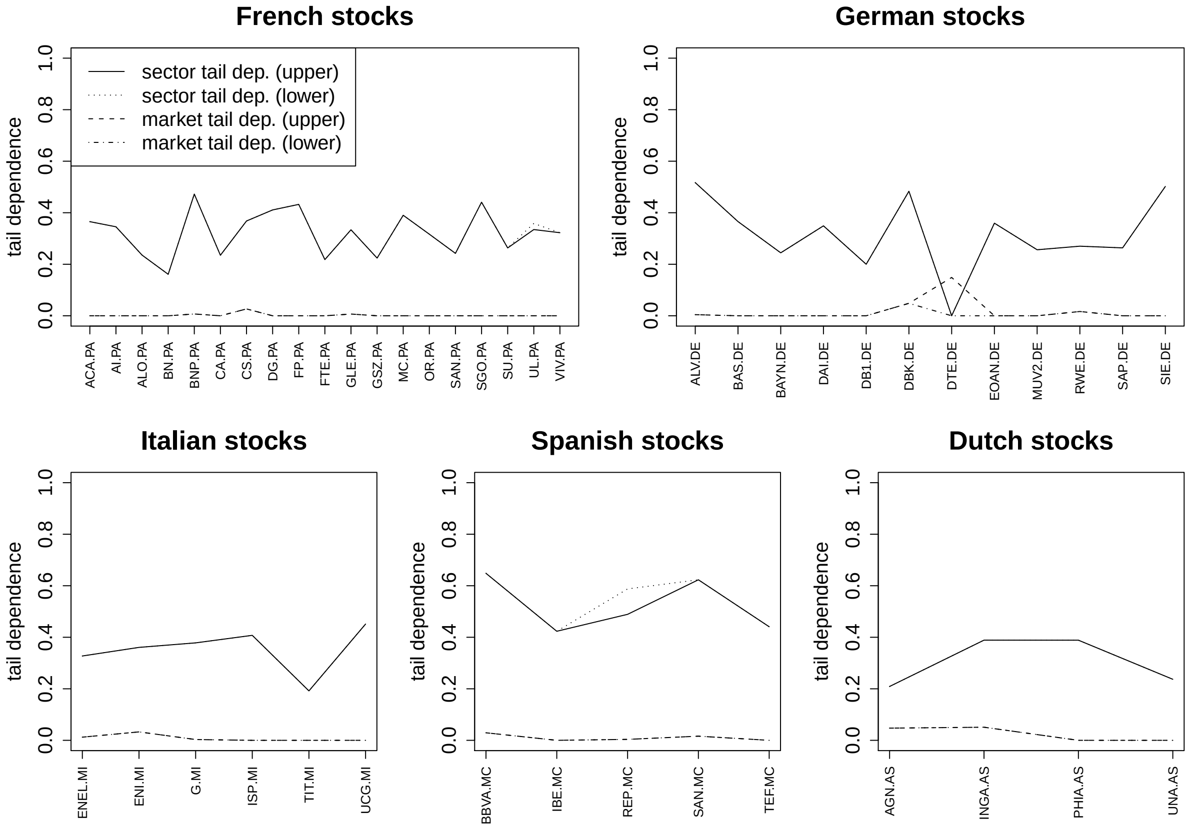

Figure 6: Sectorial and market tail dependence coefficients of the stocks from each country in the Euro Stoxx 50 data according to the RVMS model.

Figure 6 shows the tail dependence coefficients for the stocks from each country. Lines indicating the lower tail dependence are mostly not distinguishable from the lines indicating upper tail dependence, since most dependence is symmetric. Only Repsol YPF (REP.MC) exhibits asymmetric sectorial tail dependence indicating an increased risk if the Spanish stock market is going down. Deutsche Telekom (DTE.DE), on the other hand, is the only company with no sectorial tail dependence but instead an increased and asymmetric market tail dependence, which in general is very low if present at all. Overall, the tail dependence patterns are rather heterogeneous.

5 Portfolio management

So far all analyses have been static and mainly serve the purpose of better understanding the dependence structure of a data set. Now we describe how the vine copula methodology can be used for portfolio management.

5.1 Passive portfolio management

We begin with passive portfolio management. The aim is to forecast the Value-at-Risk (VaR) of a given portfolio of assets $1, ..., M$ on a daily basis. To start with, we select a moving window size T (e.g., three years), a sample size N (e.g., 100,000) and portfolio weights $\omega_j,\ j = 1, ..., M$ with

$\sum_{j=1}^{M} \omega_j = 1$ (the long-only constraint $\omega_j \geq 0$ is not required here). Using a copula dependence model with ARMA-GARCH-margins we then proceed as follows to forecast the one day ahead Value-at-Risk.

(i) We specify ARMA(p,q)-GARCH(r,s)-models with appropriate error distribution for the marginal time series, i.e., for $j = 1, ..., M$ and $t = 1, ..., T$ we estimate the parameters of the following model:

$$r_{t,j} = \mu_j + \sum_{k=1}^{p} \phi_{k,j} r_{t-k,j} + \varepsilon_{t,j} + \sum_{k=1}^{q} \theta_{k,j} \varepsilon_{t-k,j}, \quad \sigma_{t,j}^2 = \omega_j + \sum_{k=1}^{r} \alpha_{k,j} \varepsilon_{t-k,j}^2 + \sum_{k=1}^{s} \beta_{k,j} \sigma_{t-k,j}^2,$$

where $\varepsilon_{t,j} = \sigma_{t,j} Z_{t,j}$ and $(Z_{t,j})$ follows the white noise error distribution F_j. Standardized residuals are given by

$$\hat{z}_{t,j} = \frac{1}{\hat{\sigma}_{t,j}} \left(r_{t,j} - \hat{\mu}_j - \sum_{k=1}^{p} \hat{\phi}_{k,j} r_{t-k,j} - \sum_{k=1}^{q} \hat{\theta}_{k,j} \hat{\sigma}_{t-k,j} \hat{z}_{t-k,j} \right).$$

(ii) We use the estimates to compute the ex-ante GARCH variance forecast for $j = 1, ..., M$,

$$\hat{\sigma}_{T+1,j}^2 = \hat{\omega}_j + \sum_{k=1}^{r} \hat{\alpha}_{k,j} \hat{\sigma}_{T+1-k,j}^2 \hat{z}_{T+1-k,j}^2 + \sum_{k=1}^{s} \hat{\beta}_{k,j} \hat{\sigma}_{T+1-k,j}^2. \tag{5.1}$$

(iii) Using the estimated error distribution functions $\hat{F}_j$ we fit a copula model to the standardized residuals.

(iv) For each $n = 1, ..., N$:

(a) Using the estimated copula parameters and the inverse error distribution functions, we simulate a sample of standardized residuals $(\hat{z}_{T+1,1}, ..., \hat{z}_{T+1,M})'$.

(b) The samples, the estimated ARMA parameters and the ex-ante GARCH variance forecasts (5.1) are used to compute the ex-ante return forecast for $j = 1, ..., M$,

$$\hat{r}_{T+1,j} = \hat{\mu}_j + \sum_{k=1}^{p} \hat{\phi}_{k,j} r_{T+1-k,j} + \hat{\sigma}_{T+1,j} \hat{z}_{T+1,j} + \sum_{k=1}^{r} \hat{\theta}_{k,j} \hat{\sigma}_{T+1-k,j} \hat{z}_{T+1-k,j}.$$

(c) The portfolio return forecast is then given by $\hat{r}_{T+1,P} = \sum_{j=1}^{M} \omega_j \hat{r}_{T+1,j}$.

(v) Finally, we compute the $(1-\alpha)$-VaRs, $\text{VaR}_{T+1}^{1-\alpha}$, of the portfolio return by taking the α-quantiles of the portfolio return forecasts.

This procedure can easily be repeated daily, where we recommend to use sequential estimates for vine copula models, since they are quickly obtained and can be seen as reliable approximations to joint maximum likelihood estimates as discussed earlier. In step (v) any other risk measure such as the expected shortfall could of course easily be used.

Evidently, one could also directly fit a time series model to the portfolio return series ($r_{t,P} = \sum_{j=1}^{M} \omega_j r_{t,j}$) and forecast based on it. The virtue of our approach however is that it allows to compute the VaR of *any* possible portfolio involving the assets $1, ..., M$ based on the samples obtained in step (iv)(b). In other words, we can compute the VaR for any set of weights

$\omega_j,\ j = 1, ..., M$ with $\sum_{j=1}^{M} \omega_j = 1$. This will prove useful for active portfolio management as discussed in Section 5.2. Moreover, the approach allows to explicitly determine the diversification benefit of computing the portfolio VaR rather than summing up the individual VaRs of each return series. We define the diversification benefit DB_t at time t and level $1 - \alpha$ as

$$DB_t^{1-\alpha} := 1 - \frac{\mathrm{VaR}_t^{1-\alpha}(\sum_{j=1}^{M} \omega_j r_{t,j})}{\sum_{j=1}^{M} \omega_j \mathrm{VaR}_t^{1-\alpha}(r_{t,j})}, \tag{5.2}$$

which quantifies the percentage increase in wealth the investor occurs, since he has to set aside less risk capital.

In terms of statistical modeling VaR forecasting serves an additional purpose. By predicting the VaR for historical data (testing data set) we can assess the prediction accuracy of our models. For this we consider the hit sequence of ex-post exceedances (Christoffersen 1998)

$$I_t = \begin{cases} 1, & \text{if } r_{t,P} < \mathrm{VaR}_t^{1-\alpha} \\ 0, & \text{else} \end{cases}, \tag{5.3}$$

where $r_{t,P}$ denotes the ex-post observed portfolio return at time t. This sequence should exhibit two properties if the forecasts are accurate. First, the exceedances should occur independently, i.e., not in clusters, and second, the proportion of exceedances should approximately equal the VaR confidence level α (unconditional coverage). The term "conditional coverage" encompasses both properties.

In the literature, a wide range of tests for these two properties has been proposed in recent years (see, e.g., Campbell (2007) for a review). Since each test exhibits certain advantages and disadvantages and there are no general guidelines of which test to use, we recommend applying a battery of such tests to ensure that results are not biased in one or the other direction. For example, the following tests may be considered.

- The proportion of failures (POF) test of unconditional coverage by Kupiec (1995).
- The Markov test of independence by Christoffersen (1998).
- The joint test of conditional coverage by Christoffersen (1998) which combines the first two tests.
- The duration-based mixed Kupiec test of conditional coverage by Haas (2001), where durations are the time between two exceedances I_{t_i} and $I_{t_{i+1}}$.
- The duration-based Weibull test of independence by Christoffersen and Pelletier (2004).
- The duration-based GMM test of conditional coverage by Candelon et al. (2011) with orders 2 and 5.

5.2 Active portfolio management

Vine copulas are however useful not only for *passive* but also for *active* portfolio management. In particular, we discuss how vine copula models can be used for conditional asset allocation. Classically, asset allocation follows the mean-variance approach by Markowitz (1952). As in Section 5.1 we consider M assets. The aim of conditional asset allocation is to find the *optimal* portfolio allocation for the next day, i.e., the optimal weight ω_j of each asset $j \in \{1, ..., M\}$. In the mean-variance methodology an *optimal* portfolio means a portfolio with minimum variance for a given return r_0. Let $r_{T+1,j}$ denote tomorrow's return of asset $j = 1, ..., M$ and similarly Σ_{T+1} the covariance matrix of tomorrow's returns. Then the optimal portfolio weights are the solution of the following quadratic optimization problem:

$$\min \boldsymbol{\omega}' \Sigma_{T+1} \boldsymbol{\omega} \quad \text{s.t.} \quad r_{T+1,P} = \boldsymbol{\omega}' \boldsymbol{r}_{T+1} = r_0, \quad \sum_{j=1}^{M} \omega_j = 1, \tag{5.4}$$

where $\boldsymbol{\omega} = (\omega_1, ..., \omega_M)'$ and $\boldsymbol{r}_{T+1} = (r_{T+1,1}, ..., r_{T+1,M})'$. The collection of portfolios with different given returns r_0 is called *efficient frontier*. If no short-selling is allowed, we can add the long-only constraint $\omega_j \geq 0, j = 1, ..., M$. Further, we talk about the *minimum-risk* portfolio, if the constraint $\boldsymbol{\omega}' \boldsymbol{r}_{T+1} = r_0$ is dropped.

Obviously, return and variance of the next day are not known. Here, the vine copula models come into play: By following the approach presented in Section 5.1, we obtain N predicted returns for the next day (in step (iv)(b)). Based on these forecasts we can easily compute the empirical means $\hat{\boldsymbol{r}}_{T+1}$ and covariance matrix $\hat{\Sigma}_{T+1}$ and plug them into (5.4) to obtain optimal portfolio weights.

Since only the first two moments are taken into account, this approach however looses a lot of information about the dependence structure, especially about the tails, as modeled by the vine copula. Alternative approaches such as mean-VaR portfolio selection,

$$\min \mathrm{VaR}_t^{1-\alpha}(\boldsymbol{\omega}' \boldsymbol{r}_{T+1}) \quad \text{s.t.} \quad \boldsymbol{\omega}' \boldsymbol{r}_{T+1} = r_0, \quad \sum_{j=1}^{M} \omega_j = 1,$$

therefore may be more appropriate and can easily be carried out using the vine copula methodology which provides N predicted returns for each assets so that quantities such as the VaR are straightforward to obtain. In the end, the choice of risk measure yet depends on the investor.

5.3 Application to data

We use the above procedure to forecast the VaR of the portfolio of the 46 Euro Stoxx 50 stocks in our data set with weights according to the index composition as of February 8, 2010. For market risk as we consider it here one typically uses a testing period of 250 days which is approximately one year of trading. Thus we use a moving window of 900 observations, corresponding to approximately 3.5 years of daily observations, and a forecasting period from December 30, 2009 to December 20, 2010 with 250 daily observations. For this period we then forecast the one day ahead VaR for the RVMS model and the general R-vine model, which, as described above, are re-estimated daily. As benchmark models we use Gaussian and Student-t copulas as well as the DCC model by Sheppard and Engle (2001) and Engle (2002) with Gaussian and Student-t innovations. In particular, the DCC model is very popular and therefore constitutes the most relevant benchmark for the vine copula models. For computational reasons, we however only use a diagonal DCC. Moreover, for reasons of comparisons we perform the exercise using a multivariate independence copula.

We preliminarily decide to fit ARMA(1,1)-GARCH(1,1)-models with Student-t error distribution to all marginal time series in contrast to the more detailed analysis before (see Appendix A). This is done in order to limit the complexity of the approach and this is also approximately

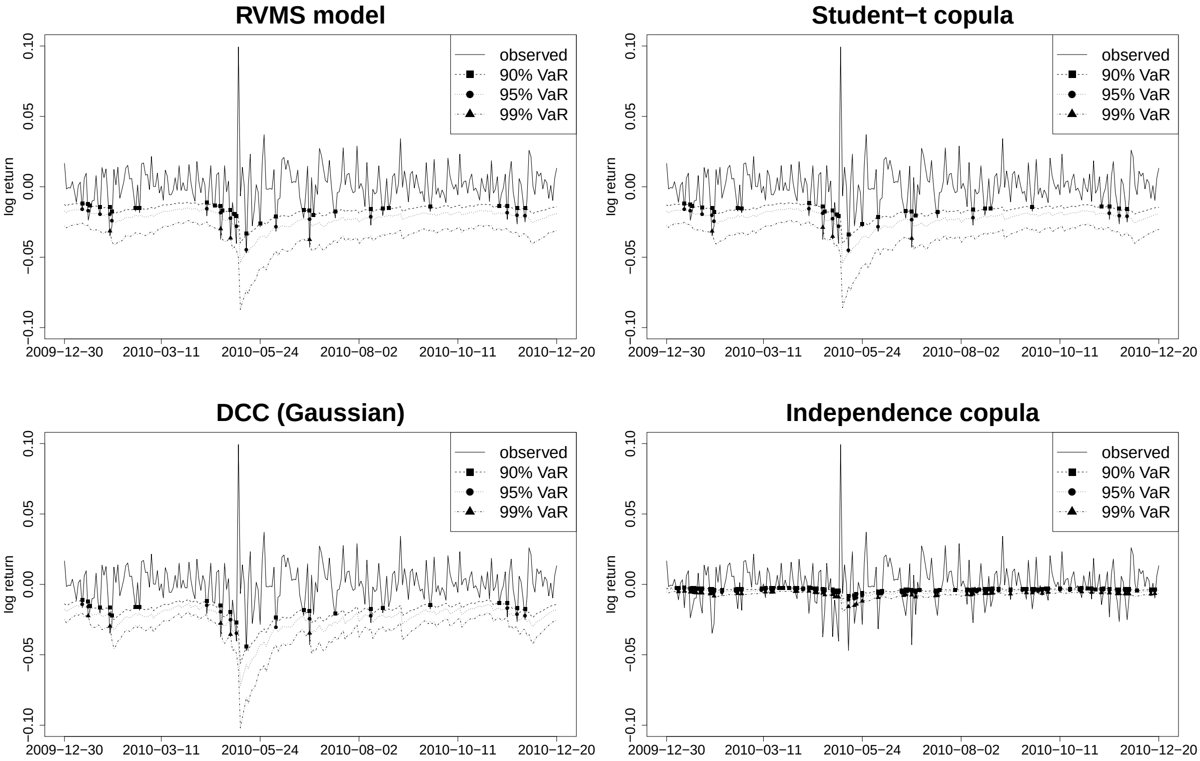

Figure 7: Log return time series of the Euro Stoxx 50 porfolio return with 90%/95%/99%-VaR forecasts of the RVMS model, the Student-t copula, the DCC model with Gaussian innovations and the independence copula. VaR exceedances (5.3) are marked by squares, circles and triangles, respectively.

accurate according to Table 8. Moreover, we again use sequential estimation because of the high computational effort of fitting 250 R-vine models for the full general model and the RVMS model each.

The return time series and the VaR forecasts of some of the models under consideration are displayed in Figure 7 (sample size is $N = 100,000$). For a testing period of 250 observations and confidence levels of 10%/5%/1% we expect 25, 12.5 and 2.5 exceedances, respectively. Apparently VaR forecasts of the independence copula model are completely inaccurate. We however would like to evaluate them using the above tests in order to compare them directly to the forecasting accuracy of the general R-vine, the RVMS model, the Gaussian and the Student-t copula as well as the DCC models. For our testing period the R-vine and the RVMS model as well as the Gaussian and the Student-t copulas produced the same hit sequences and hence are considered together.

Numbers of exceedances and p-values of the above tests are shown in Table 7. Beginning with the general R-vine and the RVMS model, it is evident from the tests that the VaR forecasts at all three levels are sufficiently accurate, since the null hypotheses of independence, unconditional and conditional coverage cannot be rejected according to all tests. In particular, exceedances are clearly independent, while there is a (non-significant) lack of coverage, since the numbers of exceedances are slightly increased compared to the expected ones. The same results hold for the Gaussian and Student-t copulas which we expected given the high number of model parameters.

			Tests of independence, unconditional and conditional coverage						
Model	Level	No. of exceed.	POF (uncond.)	Markov (indep.)	Joint Christoff.	Mixed Kupiec	Weibull (indep.)	GMM (order 2)	GMM (order 5)
R-vine/ RVMS	90%	30	0.305	0.820	0.576	0.107	0.790	0.245	0.460
	95%	16	0.329	0.976	0.621	0.148	0.656	0.350	0.379
	99%	4	0.380	0.718	0.638	0.216	0.988	0.398	0.847
Gauss/ t	90%	29	0.410	0.708	0.663	0.129	0.762	0.357	0.629
	95%	16	0.329	0.976	0.621	0.148	0.661	0.350	0.379
	99%	4	0.380	0.718	0.638	0.216	0.988	0.398	0.847
DCC (Gauss)	90%	24	0.832	0.816	0.952	0.458	0.572	0.976	0.988
	95%	13	0.885	0.231	0.483	0.565	0.548	0.909	0.931
	99%	5	0.162	0.651	0.339	0.144	0.899	0.210	0.572
DCC (t, df=7)	90%	20	0.276	0.743	0.524	0.217	0.669	0.663	0.628
	95%	11	0.657	0.313	0.545	0.499	0.635	0.939	0.743
	99%	0	-	-	-	-	-	-	-
DCC (t, df=4)	90%	17	0.075	0.870	0.202	0.062	0.848	*0.000*	*0.000*
	95%	7	0.083	0.525	0.181	0.143	0.665	0.075	0.196
	99%	0	-	-	-	-	-	-	-
indep.	90%	92	*0.000*	0.414	*0.000*	*0.000*	0.050	*0.000*	*0.000*
	95%	79	*0.000*	0.085	*0.000*	*0.000*	*0.021*	*0.000*	*0.000*
	99%	61	*0.000*	*0.019*	*0.000*	*0.000*	0.661	*0.000*	*0.000*

Table 7: p-values of the VaR backtests described in Section 5.1.

Further, the DCC model with Gaussian innovations also shows good results in terms of the backtests and none of the hypotheses can be rejected. However, using Student-t innovations in the DCC model leads to too cautious forecasts and therefore to a too high risk capital. Finally, use of the independence copula to forecast VaR numbers fails completely. As already noted above, the numbers of exceedances are much too large and the hypotheses of (un)conditional coverage are strongly rejected. The hypotheses of independence are partly not rejected, but this has only limited informative value, in particular when taking the conditional coverage tests into account.

So, why using a vine copula based model for this purpose if the common DCC with Gaussian innovations also accurately forecasts the VaR? The answer is a reduction in risk capital. When we compare VaR forecasts of the DCC model and the vine copula based models, it turns out that those of the copula models are generally lower at the 90% and 95% levels (by 11% and 6%, respectively), which are forecasted most accurately due to their smaller quantile levels. At the 99% level there is a slight increase in the VaR. This is due to the fact that the Gaussian innovations ignore heavy tail events. When using the DCC with Student-t innovations the VaRs obtained from the R-vine and the RVMS model are always much lower. This means that the use of a vine copula based model can reduce the required risk capital while providing the same level of (un)certainty regarding VaR exceedances.

We also have a look at the diversification benefit (5.2). For all models under consideration the diversification benefit is almost constant over time (time series plots are not shown here). On average, it varies between 21% and 26% dependending on the model and on the VaR level.

To illustrate asset allocation using vine copula models, we used the RVMS model to compute minimum-risk portfolios with long-only constraint based on the 100,000 samples for each return time series, so that estimates can be regarded as quite accurate. For stocks with average weight larger than 5%, smoothed weights over time are shown in Figure 8. Not surprisingly the

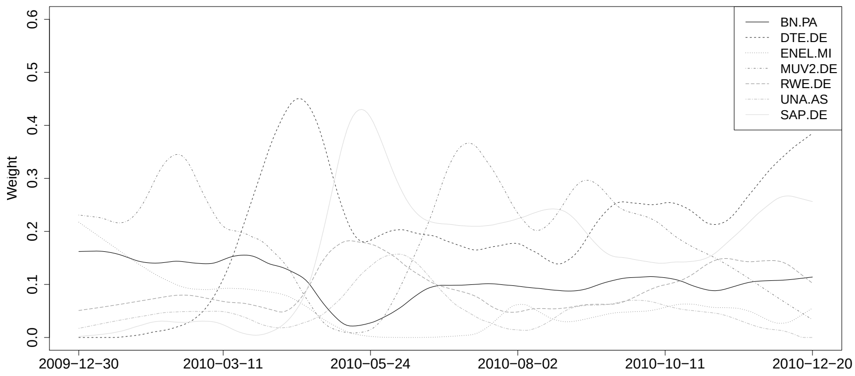

Figure 8: Smoothed weights of minimum-risk portfolios using the mean-variance approach. Only stocks with average weight larger than 5% are shown.

minimum-risk portfolio is constituted to a large part of stocks which are robust to high volatility in the market such as Munich Re (MUV2.DE, reinsurance), Unilever (UNA.AS, consumer products) or Enel (ENEL.MI, utilities).

Furthermore we also computed the mean-variance efficient frontier for each time point. Figure 9 shows it for the 50th, 100th, 150th, 200th and 250th time point in the forecasting period. Evidently its shape varies substantially based on the market conditions which are also reflected in the minimum-risk portfolio weights shown above. Figure 9 also shows minimum-risk mean-VaR portfolios which are clearly different from the minimum-risk mean-variance portfolios and hence illustrate the influence of another risk measure than the variance.

In conclusion, we can state that vine copula based models are very useful for portfolio management. Both vine copula based models considered here produce sufficiently accurate VaR forecasts and are clearly superior to an independence copula. With respect to the Gaussian and Student-t copula benchmarks, no significant difference could be determined. Similarly for the DCC model with Gaussian innovations, while Student-t innovations are not appropriate here. A possible decrease in required risk capital when using a vine copula based model rather than a DCC model shows the usefulness of the proposed methodology. Also note the RVMS model performs just as well as the general R-vine model, but at the same time forecasts are computed approximately 40% faster, i.e., the RVMS model is evidently more efficient computationally due to its more specific structure. In terms of active portfolio management, this can easily be exploited for asset allocation and extensions beyond the mean-variance approach are straightforward, since large samples of each asset are given and thus arbitrary risk measures can be estimated.

6 Summary and conclusion

The aim of this paper was to present the use and usefulness of vine copulas in financial risk management. We developed a flexible R-vine based factor model for stock market dependencies, the RVMS model, and discussed passive and active portfolio management using vine copula models. The developed methodology was used to analyze the dependency structure among important European stocks as represented in the Euro Stoxx 50 index. In these analyses our models were critically compared to relevant benchmark models such as the DCC model and the

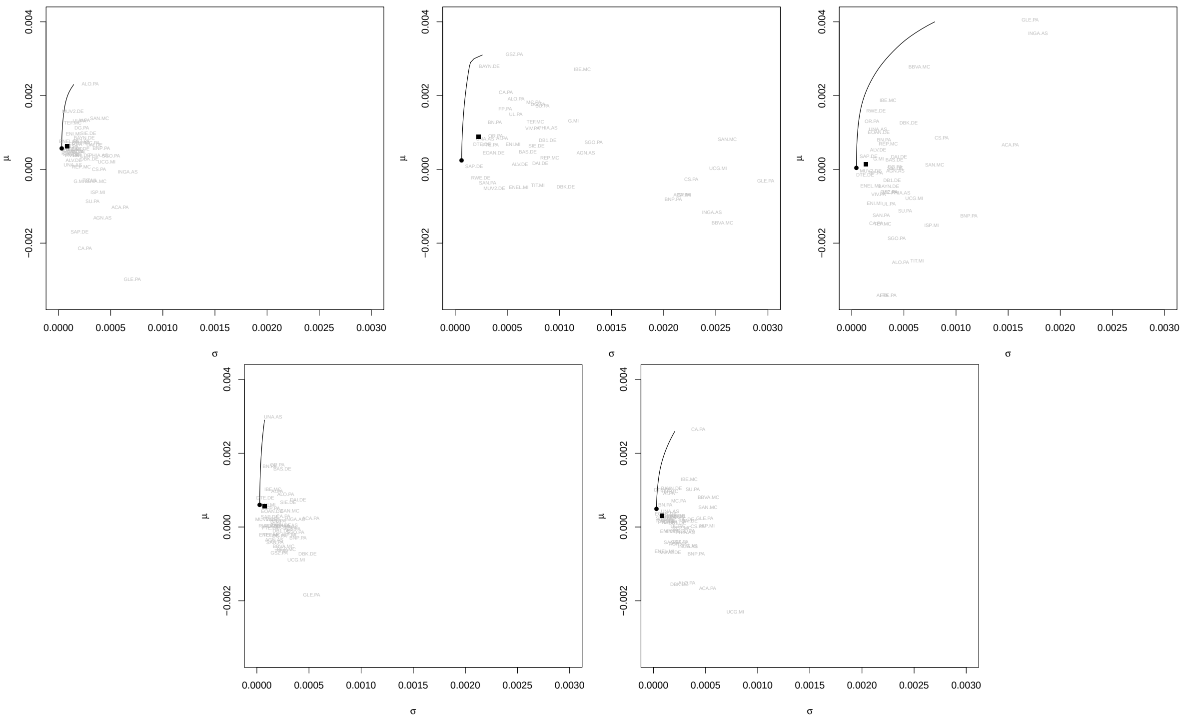

Figure 9: Conditional mean-variance efficient frontiers at 2010-03-11, 2010-05-24, 2010-08-02, 2010-10-11 and 2010-12-20 (from top left to bottom right). The respective minimum-risk portfolios are marked with a solid point, minimum-risk portfolios with respect to the VaR with a solid square. The x-axis shows the predicted standard deviation, the y-axis the predicted return.

state-of-the-art dependency model, the Student-t copula. It turned out that vine copula models, and in particular the RVMS model, provide good fits of the data and accurately and efficiently forecast the Value-at-Risk.

In future, the RVMS model has to be investigated more closely in various applications. In particular, when asymmetric dependencies are present, we expect the RVMS model to be clearly superior to models which can only capture symmetric dependence such as the Student-t copula. Extensions to more factors or completely different factor structures are also to be investigated. The dependence structure of sectors in the RVMS model might, e.g., be decomposed into subsectors. As an example one might think of a decomposition of the MSCI World into continental indices which again can be decomposed into national indices. Vine based factor models in the light of the studies by Fama and French (1992) and Drummen and Zimmermann (1992) are interesting extensions, too.

Another important direction of future research is the consideration of dynamic copula parameter structures for the RVMS model. Research in this direction is still at a rather early stage (see Ausin and Lopes (2010), Manner and Reznikova (2010) and Almeida and Czado (2011)). While the present work constitutes one of the first applications of vines in dimensions 50 or higher (and, to the best of our knowledge, is the first to perform joint maximum likelihood estimation in such dimensions), dynamic (vine) copula models in such dimensions are not yet feasible.

Finally a note on the number of parameters: At first glance the vine copula models may appear grossly overspecified. While Tables 3 and 6 only show the number of copula model parameters, the complete model also includes the parameters for the marginal ARMA-GARCH-models. If ARMA(1,1)-GARCH(1,1)-models with mean and Student-t error distribution are chosen, this means an additional $52 \times 7 = 364$ parameters in our 52-dimensional application. Even if the vine copulas have clearly less parameters than the often used Gaussian and Student-t copulas with full correlation matrix, this still amounts to more than 800 parameters. A diagonal DCC model, on the other hand, has clearly less parameters. When combining it with ARMA(1,1)-models with mean ($52 \times 3 = 156$), one ends up with only $156 + 3 \times 52 + 2 = 314$ parameters. Things however look different, when a non-diagonal DCC model is used, which may be necessary in certain application. Then the model has the immense number of $156 + 52 + 2 \times 52^2 + 2 = 5618$ parameters! Under this impression the number of parameters of a vine copula model appears much less awkward and still is very well manageable using the techniques discussed in this paper.

Acknowledgement

The first author gratefully acknowledges the support of the TUM Graduate School's Faculty Graduate Center *International School of Applied Mathematics* at Technische Universität München, Germany. The numerical computations were performed on a Linux cluster supported by DFG grant INST 95/919-1 FUGG.

References

Aas, K., C. Czado, A. Frigessi, and H. Bakken (2009). Pair-copula constructions of multiple dependence. *Insurance: Mathematics and Economics 44*(2), 182–198.

Akaike, H. (1973). Information theory and an extension of the maximum likelihood principle. In B. N. Petrov and F. Csaki (Eds.), *Proceedings of the Second International Symposium on Information Theory Budapest*, Akademiai Kiado, pp. 267–281.

Almeida, C. and C. Czado (2011). Efficient bayesian inference for stochastic time-varying copula models. *To appear in Computational Statistics & Data Analysis*.

Ang, A. and G. Bekaert (2002). International asset allocation with regime shifts. *Review of Financial Studies 15*(4), 1137–1187.

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

Bauwens, L., S. Laurent, and J. V. K. Rombouts (2006). Multivariate GARCH models: A survey. *Journal of Applied Econometrics 21*(1), 79–109.

Bedford, T. and R. M. 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.

Berg, D. (2009). Copula goodness-of-fit testing: An overview and power comparison. *The European Journal of Finance 15*(7), 675–701.

Berg, D. and K. Aas (2009). Models for construction of higher-dimensional dependence: A comparison study. *European Journal of Finance 15*, 639–659.

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

Brechmann, E. C. (2010). Truncated and simplified regular vines and their applications. Diploma Thesis, Faculty of Mathematics, Technische Universität München, Germany.

Brechmann, E. C., C. Czado, and K. Aas (2010). Truncated regular vines and their applications. *Submitted for publication*.

Breymann, W., A. Dias, and P. Embrechts (2003). Dependence structures for multivariate high-frequency data in finance. *Quantitative Finance 1*, 1–14.

Campbell, S. D. (2007). A review of backtesting and backtesting procedures. *Journal of Risk 9*(2), 1–18.

Candelon, B., G. Colletaz, C. Hurlin, and S. Tokpavi (2011). Backtesting Value-at-Risk: A GMM duration-based test. *Journal of Financial Econometrics 9*(2), 314–343.

Cherubini, U., E. Luciano, and W. Vecchiato (2004). *Copula Methods in Finance*. Chichester: John Wiley & Sons.

Christoffersen, P. F. (1998). Evaluating interval forecasts. *International Economic Review 39*, 841–862.

Christoffersen, P. F. and D. Pelletier (2004). Backtesting Value-at-Risk: A duration-based approach. *Journal of Financial Econometrics 2*(1), 84–108.

Czado, C., U. Schepsmeier, and A. Min (2011). Maximum likelihood estimation of mixed C-vines with application to exchange rates. *To appear in Statistical Modelling*.

Dißmann, J., E. C. Brechmann, C. Czado, and D. Kurowicka (2011). Selecting and estimating regular vine copulae and application to financial returns. *Submitted for publication*.

Drummen, M. and H. Zimmermann (1992). The structure of European stock returns. *Financial Analysts Journal 48*, 15–26.

Embrechts, P., A. McNeil, and D. Straumann (2002). Correlation and dependence in risk management: properties and pitfalls. In M. A. H. Dempster (Ed.), *Risk Management: Value at Risk and Beyond*. Cambridge: Cambridge University Press.

Engle, R. F. (1982). Autoregressive conditional heteroscedasticity with estimates of the variance of United Kingdom inflation. *Econometrica 50*(4), 987–1007.

Engle, R. F. (2002). Dynamic conditional correlation. *Journal of Business and Economic Statistics 20*(3), 339–350.

Fama, E. F. and K. R. French (1992). The cross-section of expected stock returns. *Journal of Finance 47*(2), 427–465.

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

Genest, C. and A.-C. Favre (2007). Everything you always wanted to know about copula modeling but were afraid to ask. *Journal of Hydrologic Engineering 12*, 347–368.

Haas, M. (2001). New methods in backtesting. Technical report, Research Center Caesar.

Heinen, A. and A. Valdesogo (2009). Asymmetric CAPM dependence for large dimensions: the canonical vine autoregressive model. CORE discussion papers 2009069, Université catholique de Louvain, Center for Operations Research and Econometrics (CORE).

Hlawatsch, S. and P. Reichling (2010). Portfolio management under asymmetric dependence and distribution. Working paper no. 17/2010, Faculty of Econometrics, University of Magdeburg.

Hobæk Haff, I. (2010). Estimating the parameters of a pair-copula construction. *Submitted for publication*.

Hobæk Haff, I. (2011). Comparison of estimators for pair-copula constructions. *To appear in the Journal of Multivariate Analysis*.

Joe, H. (1996). Families of m-variate distributions with given margins and $m(m-1)/2$ bivariate dependence parameters. In L. Rüschendorf, B. Schweizer, and M. D. Taylor (Eds.), *Distributions with fixed marginals and related topics*, pp. 120–141. Hayward: Institute of Mathematical Statistics.

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

Kupiec, P. H. (1995). Techniques for verifying the accuracy of risk measurement models. *The Journal of Derivatives 3*(2), 73–84.

Kurowicka, D. and R. M. Cooke (2006). *Uncertainty Analysis with High Dimensional Dependence Modelling*. Chichester: John Wiley.

Kurowicka, D. and H. Joe (2011). *DEPENDENCE MODELING: Vine Copula Handbook*. Singapore: World Scientific Publishing Co.

Lindskog, F., A. McNeil, and U. Schmock (2003). Kendall's tau for elliptical distributions. In G. Bol, G. Nakhaeizadeh, S. T. Rachev, T. Ridder, and K.-H. Vollmer (Eds.), *Credit Risk: Measurement, Evaluation and Management*. Heidelberg: Physica-Verlag.

Lintner, J. (1965). The valuation of risk assets and the selection of risky investments in stock portfolios and capital budgets. *Review of Economics and Statistics 47*(1), 13–37.

Longin, F. and B. Solnik (1995). Is the correlation in international equity returns constant: 1960-1990? *Journal of International Money and Finance 14*, 3–26.

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

Mandelbrot, B. (1963). The variation in certain speculative prices. *Journal of Business 36*, 394–419.

Manner, H. and O. Reznikova (2010). A survey on time-varying copulas: Specification, simulations and estimation. *To appear in Econometric Reviews*.

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

Mashal, R. and A. Zeevi (2002). Beyond correlation: Extreme co-movements between financial assets. Technical report, Columbia University.

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

Min, A. and C. Czado (2010). Bayesian inference for multivariate copulas using pair-copula constructions. *Journal of Financial Econometrics 8*(4), 511–546.

Morales-Nápoles, O., R. M. Cooke, and D. Kurowicka (2010). About the number of vines and regular vines on n nodes. *Submitted for publication*.

Nelsen, R. B. (2006). *An Introduction to Copulas* (2nd ed.). Berlin: Springer.

Salmon, F. (2009). Recipe for disaster: The formula that killed wall street. *Wired Magazine 17*(3).

Savu, C. and M. Trede (2010). Hierarchical Archimedean copulas. *Quantitative Finance 10*, 295–304.

Sharpe, W. (1964). Capital asset prices: A theory of market equilibrium under conditions of risk. *Journal of Finance 19*(3), 425–442.

Sheppard, K. K. and R. F. Engle (2001). Theoretical and empirical properties of dynamic conditional correlation multivariate GARCH. Technical Report 8554, NBER Working Paper Series.

Sklar, A. (1959). Fonctions de répartition à n dimensions et leurs marges. *Publications de l'Institut de Statistique de L'Université de Paris 8*, 229–231.

Valdesogo, A. (2009). *Multivariate volatility models using copulas*. Ph. D. thesis, Center for Operations Research and Econometrics, Université Catholique de Louvain, Belgium.

Vuong, Q. H. (1989). Ratio tests for model selection and non-nested hypotheses. *Econometrica 57*(2), 307–333.

A Marginal models of Euro Stoxx 50 data

As discussed in Section 3, marginal time series models for the Euro Stoxx 50 data have to be found in the first step of our two step estimation approach. Here, time series models are selected according to a stepwise procedure:

(i) We fit ARMA(1,1)-GARCH(1,1)-, AR(1)-GARCH(1,1)-, MA(1)-GARCH(1,1)- and GARCH(1,1)-models to the univariate time series. As error distribution we use a Student-t distribution to account for heavy tails, while a skewed Student-t error distribution was not neeeded, since no significant skewness was found in the data. For all four time series models we then perform Kolmogorov-Smirnov goodness-of-fit tests for the standardized residuals and choose the model with the highest p-value if this value is at least larger than 5%.

(ii) If the degrees of freedom of the Student-t error distribution are larger than ten, we choose a standard normal distribution instead (if the p-value of the corresponding Kolmogorov-Smirnov test is larger than 5%).

(iii) Since the Kolmogorov-Smirnov test sometimes lacks power, we also perform Ljung-Box tests with lag 30 for all residuals. If a p-value is smaller than 5%, we stepwisely increase the corresponding ARMA(p,q) terms, so that the model remains rather parsimonious, until both the Ljung-Box test and the respective Kolmogorov-Smirnov test for the residuals have p-values larger than 5%.

The resulting marginal time series models are shown in Table 8.

	model	$\hat{\mu}$	$\hat{\phi}_1$	$\hat{\theta}_1$	$\hat{\omega}$	$\hat{\alpha}_1$	$\hat{\beta}_1$	$\hat{\nu}$	KS	LB
^STOXX50E	GARCH(1,1)	0.001	-	-	0.000	0.104	0.891	8.586	0.317	0.615
^GDAXIP	GARCH(1,1)	0.000	-	-	0.000	0.087	0.906	8.665	0.103	0.589
^AEX	GARCH(1,1)	0.001	-	-	0.000	0.115	0.880	-[a]	0.142	0.399
^FCHI	GARCH(1,1)	0.001	-	-	0.000	0.096	0.897	9.489	0.409	0.400
FTSEMIB.MI	GARCH(1,1)	0.000	-	-	0.000	0.104	0.897	9.426	0.206	0.376
^IBEX	GARCH(1,1)	0.000	-	-	0.000	0.105	0.890	7.212	0.393	0.503
ACA.PA	GARCH(1,1)	-0.000	-	-	0.000	0.085	0.915	6.611	0.921	0.749
AGN.AS	GARCH(1,1)	0.000	-	-	0.000	0.129	0.875	7.401	0.662	0.180
AI.PA	MA(1)-GARCH(1,1)	0.001	-	-0.114	0.000	0.094	0.873	5.932	0.923	0.060
ALO.PA	AR(1)-GARCH(1,1)	0.001	-0.065	-	0.000	0.058	0.935	6.732	0.810	0.602
ALV.DE	ARMA(5,0)-GARCH(1,1)[b]	0.001	0.012	-	0.000	0.092	0.903	6.538	0.240	0.078
BAS.DE	ARMA(1,1)-GARCH(1,1)	0.002	-0.684	0.668	0.000	0.097	0.900	5.518	0.760	0.612
BAYN.DE	ARMA(1,1)-GARCH(1,1)	0.000	0.522	-0.570	0.000	0.106	0.849	6.591	0.951	0.787
BBVA.MC	GARCH(1,1)	0.000	-	-	0.000	0.115	0.878	7.160	0.711	0.085
BN.PA	ARMA(1,1)-GARCH(1,1)	0.000	0.626	-0.699	0.000	0.089	0.893	8.381	0.958	0.640
BNP.PA	GARCH(1,1)	0.000	-	-	0.000	0.089	0.909	7.189	0.554	0.864
CA.PA	GARCH(1,1)	0.000	-	-	0.000	0.069	0.926	5.680	0.860	0.208
CS.PA	MA(1)-GARCH(1,1)	0.000	-	0.019	0.000	0.106	0.890	9.658	0.486	0.171
DAI.DE	AR(1)-GARCH(1,1)	0.001	0.011	-	0.000	0.102	0.896	5.319	0.552	0.251
DB1.DE	GARCH(1,1)	0.001	-	-	0.000	0.086	0.908	7.695	0.730	0.079
DBK.DE	AR(1)-GARCH(1,1)	0.000	0.028	-	0.000	0.116	0.891	6.442	0.576	0.408
DG.PA	GARCH(1,1)	0.001	-	-	0.000	0.094	0.902	8.534	0.256	0.254
DTE.DE	GARCH(1,1)	0.000	-	-	0.000	0.075	0.918	5.197	0.647	0.956
ENEL.MI	MA(1)-GARCH(1,1)	0.001	-	-0.073	0.000	0.124	0.870	6.325	0.356	0.078
ENI.MI	ARMA(1,1)-GARCH(1,1)	0.001	-0.580	0.526	0.000	0.112	0.874	8.328	0.470	0.484
EOAN.DE	GARCH(1,1)	0.001	-	-	0.000	0.100	0.881	4.809	0.179	0.781
FP.PA	AR(1)-GARCH(1,1)	0.000	-0.001	-	0.000	0.089	0.893	8.904	0.944	0.161
FTE.PA	MA(1)-GARCH(1,1)	0.000	-	-0.021	0.000	0.037	0.954	5.802	0.936	0.274
G.MI	GARCH(1,1)	0.000	-	-	0.000	0.102	0.883	5.635	0.922	0.263
GLE.PA	ARMA(1,1)-GARCH(1,1)	-0.000	-0.416	0.498	0.000	0.117	0.888	7.323	0.770	0.305
GSZ.PA	AR(1)-GARCH(1,1)	0.001	-0.032	-	0.000	0.090	0.894	8.604	0.272	0.443
IBE.MC	ARMA(1,1)-GARCH(1,1)	0.000	-0.217	0.232	0.000	0.182	0.792	5.047	0.152	0.558
INGA.AS	GARCH(1,1)	0.000	-	-	0.000	0.166	0.846	5.952	0.746	0.137
ISP.MI	ARMA(1,1)-GARCH(1,1)	-0.000	-0.133	0.166	0.000	0.103	0.893	6.325	0.398	0.250
MC.PA	GARCH(1,1)	0.001	-	-	0.000	0.060	0.933	7.057	0.990	0.581
MUV2.DE	ARMA(1,1)-GARCH(1,1)	0.001	-0.782	0.744	0.000	0.116	0.868	5.877	0.145	0.062
OR.PA	MA(1)-GARCH(1,1)	0.001	-	-0.087	0.000	0.082	0.893	8.200	0.983	0.405
PHIA.AS	MA(1)-GARCH(1,1)	0.001	-	-0.035	0.000	0.072	0.924	7.058	0.358	0.846
REP.MC	ARMA(1,1)-GARCH(1,1)	0.000	0.494	-0.502	0.000	0.111	0.871	5.994	0.792	0.223
RWE.DE	GARCH(1,1)	0.001	-	-	0.000	0.067	0.906	5.728	0.886	0.087
SAN.MC	MA(1)-GARCH(1,1)	0.001	-	-0.017	0.000	0.124	0.873	8.252	0.677	0.052
SAN.PA	GARCH(1,1)	0.000	-	-	0.000	0.049	0.946	5.068	0.975	0.586
SGO.PA	ARMA(1,1)-GARCH(1,1)	0.001	-0.693	0.639	0.000	0.094	0.907	7.899	0.375	0.779
SIE.DE	GARCH(1,1)	0.001	-	-	0.000	0.058	0.936	5.161	0.700	0.063
SU.PA	ARMA(3,0)-GARCH(1,1)[c]	0.001	-0.052	-	0.000	0.073	0.920	7.556	0.396	0.435
TEF.MC	AR(1)-GARCH(1,1)	0.001	-0.006	-	0.000	0.094	0.882	7.507	0.714	0.451
TIT.MI	ARMA(1,1)-GARCH(1,1)	-0.000	0.820	-0.867	0.000	0.111	0.869	8.506	0.256	0.217
UCG.MI	AR(1)-GARCH(1,1)	0.000	-0.023	-	0.000	0.109	0.899	8.399	0.218	0.128
UL.PA	GARCH(1,1)	0.001	-	-	0.000	0.099	0.889	8.806	0.677	0.115
UNA.AS	GARCH(1,1)	0.001	-	-	0.000	0.083	0.902	5.791	0.922	0.064
VIV.PA	GARCH(1,1)	0.001	-	-	0.000	0.091	0.902	-[a]	0.250	0.544
SAP.DE	MA(1)-GARCH(1,1)	0.000	-	0.010	0.000	0.027	0.961	3.970	0.898	0.251

Table 8: Marginal time series models for the log returns of the considered six indices and 46 stocks.

[a] Normal error distribution.

[b] $\hat{\phi}_2 = -0.000$, $\hat{\phi}_3 = -0.036$, $\hat{\phi}_3 = 0.010$ and $\hat{\phi}_4 = -0.040$.

[c] $\hat{\phi}_2 = -0.017$ and $\hat{\phi}_3 = -0.092$.