



**Electronic Journal of Applied Statistical Analysis
EJASA, Electron. J. App. Stat. Anal.**

<http://siba-ese.unisalento.it/index.php/ejasa/index>

e-ISSN: 2070-5948

DOI: 10.1285/i20705948v8n2p246

Factor copulas through a vine structure

By Riviuccio

Published: 14 October 2015

This work is copyrighted by Università del Salento, and is licensed under a Creative Commons Attribution - Non commerciale - Non opere derivate 3.0 Italia License.

For more information see:

<http://creativecommons.org/licenses/by-nc-nd/3.0/it/>

Factor copulas through a vine structure

Rivieccio Giorgia*

*Department of Management Studies and Quantitative Methods of Parthenope University
Naples, Italy*

Published: 14 October 2015

Copula functions have been widely used in actuarial science, finance and econometrics. Though multivariate copulas allow for a flexible specification of the dependence structure of economic variables, they are not particularly tempting in high dimensional contexts. A factor model which involves copula functions has already proved to be a powerful tool in credit risk applications. We exploit a recent approach to obtain a factor copula model based on a vine structure, which enables to model the dependence and conditional dependence of variables through a representation of a cascade of arbitrary bivariate copulas. The contribution of this paper consists into applying the vine copula model in order to derive a non linear three-factor model. In particular, we draw the three-factor model of Fama and French (1992). According to the Inference for Margins (IFM) method, we have computed, separately, the margins and the copula parameters via maximum likelihood estimation. Finally, tail dependence measures are given for the implied estimated copula.

keywords: Factor copula model, Vines, Tail dependence, Tail density functions.

1 Introduction

The factor approach is a powerful tool in market and credit risk modelling, e.g. for the pricing of CDO, due to its capability in describing, in a both flexible and tractable way, the joint default for a large number of names. One of the most prominent theoretical contributions proposed in literature is the three-factor model of Fama and French (1992), an extension of the well-known Capital Asset Pricing Model (CAPM) (Lintner, 1965; Sharpe, 1964). According to this model, the cross-section of expected stock returns can

*giorgia.rivieccio@uniparthenope.it

be explained along a linear three-factor model composed of three common factors: a broad market premium, the spread between small and big market capitalization stocks, and the spread between value and growth stocks.

In particular, the size factor rose from a work of Banz (1981), who verified a negative relationship between average return and firm size. The role of the third factor in explaining the cross-section of average returns was documented by Chan et al. (1991) in the Japanese market.

The factor models proposed in theoretical and empirical applications are often embedded into a stochastic correlations framework, implying marginal and joint normality of the stocks and factor returns and of the idiosyncratic cross-sectionally and serially independent error terms. Moreover, the classical factor models assume the independence of common factor returns but Oh and Patton (2011) have already dropped this assumption. It is widely recognized the evidence of skewness and time conditioning in the univariate behavior of stock returns and, overall, in their whole and tail dependence structure, see e.g., Campbell et al. (2002), Dias and Embrechts (2010), Engle (2002), Longin and Solnik (2001), Patton (2006), to name just a few. Furthermore, Richardson and Smith (1993) show the departure from normality of the asset returns and CAPM residuals. Finally, Chollete and Ning (2010) find evidence of joint tail dependence in several US risk factors.

In order to exploit the cross-section of expected stock returns in non linear way, possibly asymmetrical, a factor approach which involves copula functions is proposed.

A factor copula model allows for computational efficiency, overall in high dimensional contexts, and a flexible specification of the dependence structure, not limited to independence or linear dependence.

Introduced by Vasicek (1987) to evaluate the loan loss distribution, and after applied by Li (2000) to multi-name credit derivatives, in terms of Gaussian factor copula, the model was later generalized by Andersen and Sidenius (2004), Gregory and Laurent (2005), McNeil et al. (2005) and van der Voort (2005), which introduced non linear versions. Other extensions include Hull and White (2010), which proposed a two-parameter version of their "implied copula" model (Hull and White, 2004, 2006). Many of these models are based on simulation or calibration techniques and are not directed to the estimation of copula parameters. Besides, the extension of copula in a multivariate framework requires additional assumptions and it is not an easy task (see Oh and Patton, 2011).

Then, a flexible approach to obtain multivariate copulas can be represented by vines (Bedford and Cooke, 2001), based on the factorization of the multivariate copula density in terms of bivariate linking copulas and lower-dimensional margins. Vines were proposed at first by Joe (1996), as Pair Copula Construction (PCC), later discussed by Bedford and Cooke (2001; 2002), which introduced a graphical representation denoted as regular "vine", and by Kurowicka and Cooke (2004; 2006) which defined two particular classes of vine calling them D-vine and C-vine and, finally, developed by Aas et al. (2006) and Aas and Berg (2009), which described statistical inference techniques. In the light of these reasons, adopting a vine structure to draw a multivariate factor copula model allows to:

- depart from linear dependence and gaussian distribution, both via more general marginal distributions and in the dependence structure;
- estimate the copula parameters in a computationally feasible way for high-dimensional continuous variables;
- obtain a flexible specification due to a free selection of copulas involved at various levels of the path-structure from any class and without restrictions on parameters, implying a better description of the joint behavior of the data with different tail dependencies, possibly asymmetric;
- verify the presence of independence between common factors and among variables conditionally on common factors, hypothesized by the classical factor copula approaches.

Brechmann and Czado (2011), Heinen and Valdesogo (2009; 2011) have applied a vine copula model in order to obtain a non linear version of the CAPM, considering an extension which includes the sector effects.

Our proposal is to apply the three-factor model of Fama and French (1992) employing a vine copula structure, with GARCH models for margins. It is the first time that a non linear factor model which involves more than one factor is built in this way. According to the Inference for Margins (IFM) method, we have estimated, separately, the margins and the copula parameters via maximum likelihood estimation. In the first, GARCH models for margins are applied and then, given the conditional independence of the transformed standardized residuals with respect to common factors, vine copulas are estimated, providing the parameters of an implied copula for the asset returns. Finally we have computed tail dependence measures for the "implied copula" obtained.

The paper is organized as follows. In Section 2 an overview of the three-factor model of Fama and French (1992) is given. Section 3 deals with a theoretical specification of the existing factor copula models. Section 4 describes basic concepts of vine copulas and their use in the factor approach. Section 5 reviews tail dependence measures of vine copulas. Section 6 provides empirical results of the analysis applied to the industry portfolio co-movements conditioned on three common factors. In Section 7, outlook and concluding remarks close.

2 One-factor copula model

In this section a brief review of the three-factor model of Fama and French (1992) is given. The three-factor model was proposed in order to capture the pattern in U.S. average returns associated with size and value versus growth. The Fama and French factors are constructed using the 6 value-weight portfolios from the intersection of two sizes (expressed by Market Equity - ME) and three Book-to-Market Equity (expressed as ratio of Book Equity and Market Equity - BE/ME).

Let R_{it} be the portfolio return at time t , R_{Ft} the risk free rate, R_{Mt} the market return,

SmB_t the average return on the three Small minus the three Big portfolios and HmL_t the average return on the two High BE/ME (value) minus the two Low BE/ME (growth) portfolios, then the three-factor model is given by

$$R_{it} - R_{Ft} = \alpha + \beta_i[R_{Mt} - R_{Ft}] + \varsigma_i SmB_t + \gamma_i HmL_t + \varepsilon_{it}$$

where the left side of the equation denotes the excess of portfolio returns with respect to the risk free asset return, β_i is the systematic risk for the portfolio i , ε_{it} represents the idiosyncratic error term for the portfolio i at time t , α is a constant coefficient and, finally, ς_i γ_i are factor loadings for portfolio i .

Although the three-factor model explains the cross-section of expected values in stock returns better than the one-factor model, we need a more general model which goes beyond linear dependence, gaussian distribution, both in the margins and in the dependence structure, and orthogonal common factors. In this way, copula functions can be a useful tool to specify a non linear three-factor model.

The factor models postulate that a limited number of common factors suffices to completely explain the dependence structure of the system (see Cherubini et al., 2011). The key concept is "conditional independence": conditioning on a specific scenario of the common factors, the economic variables are considered independent each-other. Considering a single common factor Y , the conditional joint co-movement of n random variables X_i is measured by the product of the marginal conditional probabilities of the variables with respect to the common factor, given by

$$P(X_1 \leq x_1, \dots, X_n \leq x_n | Y = y) = \prod_{i=1}^n P(X_i \leq x_i | Y = y)$$

where the unconditional joint distribution is

$$P(X_1 \leq x_1, \dots, X_n \leq x_n, Y \leq y) = \int_0^1 \prod_{i=1}^n P(X_i \leq x_i | Y = y) dy$$

Let F_X and F_Y be the continuous marginal cumulative distribution functions (cdf) of a bivariate copula C see, e.g., for definition and theory (see, e.g., for definition and theory Cherubini et al., 2004; Joe, 1997; Nelsen, 2006), then each conditional probability is a partial derivative of C with respect to the conditioning variable or common factor v

$$\begin{aligned} P(X \leq x | Y = y) &= \lim_{\Delta y \rightarrow 0} P(X \leq x | y \leq Y \leq y + \Delta y) \\ &= \frac{\partial C(F_X(x), v)}{\partial v} \Big|_{v \equiv F_Y(y)} \end{aligned}$$

Since each joint probability is defined in terms of copula as

$$P(X \leq x, Y \leq y) = \int_0^{F_Y(y)} \frac{\partial C(F_X(x), w)}{\partial w} dw,$$

and assuming the conditional independence between variables, we obtain

$$P(X_1 \leq x_1, \dots, X_n \leq x_n | Y = y) = \prod_{i=1}^n \frac{\partial C(F_{X_i}(x_i), v)}{\partial v} \Big|_{v \equiv F_Y(y)} \quad (1)$$

and the joint co-movement in terms of product of conditional copulas with respect to the common factor

$$P(X_1 \leq x_1, \dots, X_n \leq x_n, Y \leq y) = \int_0^{F_Y(y)} \prod_{i=1}^n \frac{\partial C(F_{X_i}(x_i), w)}{\partial w} dw$$

By using different conditional copula pairs, e.g. by means of a vine structure (see Section 3), one can easily get a more flexible joint distribution, especially in high dimensions. Formally, the one-factor copula model follows (Oh and Patton, 2011)

$$\begin{aligned} X_i &= h(Y, \varepsilon_i) \quad \text{for } i = 1, \dots, n \\ Y &\sim F_Y(\boldsymbol{\theta}), \quad \varepsilon_i \sim iid F_\varepsilon(\boldsymbol{\theta}), \quad Y \perp \varepsilon_i \quad \forall i \\ [X_1, \dots, X_n]' &\equiv \mathbf{X} \sim \mathbf{F}_x = C(G_1(\boldsymbol{\theta}), \dots, G_n(\boldsymbol{\theta}); \boldsymbol{\theta}) \end{aligned}$$

where the implied copula $C(\boldsymbol{\theta})$ has an unknown form. This generalization reveals that this structure nests a variety of well-known copulas in the literature, see for instance table 1 (see Oh and Patton, 2011). The copula structure involved is too linked to the h function choice. Adopting a vine structure allows for a flexible choice of h for each X_i . Our goal is to apply the three-factor model of Fama and French (1992) by means of vine copulas. In general, assuming k common factors denoted as $\mathbf{Y}_K = [Y_1, \dots, Y_k]'$, we get

$$\begin{aligned} X_i &= h(\mathbf{Y}_K, \varepsilon_i) \\ \mathbf{Y}_K &\sim \mathbf{F}_Y = C(F_{Y_1}, \dots, F_{Y_k}), \quad \varepsilon_i \sim iid F_\varepsilon, \quad Y_k \perp \varepsilon_i \quad \forall i, k \end{aligned} \quad (2)$$

where C is the copula which allows for the dependence between margins F .

3 Pair-Copula Constructions (PCC) and vines

Vines are graphical representations of the Pair-Copula Constructions (PCC) which are based on the decomposition of a multivariate density into a cascade of bivariate conditional and unconditional copula densities. Their simple construction, based on pair-copulas as building blocks of the structure, combined to high flexibility, due to the free choice of each pair-copula type from any class and without restrictions on parameters, allows easily to derive higher-dimensional copulas. Let $X = (X_1, \dots, X_n)$ a vector of n random variables with marginal densities $f_i(x_i)$, joint cdf $F(\cdot)$ and marginal cdf's $F_i(x_i)$, then the factorization of the joint density $f(\cdot)$ is obtained as

$$\begin{aligned} f(x_1, \dots, x_n) &= f(x_n | x_1, \dots, x_{n-1}) \cdot f(x_1, \dots, x_{n-1}) \\ &= \prod_{t=2}^n f(x_t | x_1, \dots, x_{t-1}) \cdot f_1(x_1) \end{aligned} \quad (3)$$

By applying Sklar's theorem (1959), we get the relationship among an n -dimensional joint density and the copula and marginal densities as

$$f(x_1, \dots, x_n) = c_{1, \dots, n}(F_1(x_1), \dots, F_n(x_n)) \cdot \prod_{t=1}^n f_t(x_t) \tag{4}$$

This relationship allows to achieve a PCC decomposition of an n -dimensional density function.

Considering $n = 3$ random variables with absolutely continuous joint and marginal cdf's, then, a possible decomposition of the joint density $f(\cdot)$ is

$$f(x_1, x_2, x_3) = f(x_3|x_1, x_2) \cdot f(x_2|x_1) \cdot f_1(x_1) \tag{5}$$

where

$$f(x_3|x_1, x_2) = c_{1,3|2}(F(x_1|x_2), F(x_3|x_2)) \cdot f(x_3|x_2) \tag{6}$$

and

$$\begin{aligned} f(x_i|x_{i-1}) &= f(x_{i-1}, x_i)/f_{i-1}(x_{i-1}) \\ &= c_{i-1,i}(F_{i-1}(x_{i-1}), F_i(x_i)) \cdot f_i(x_i), \quad i = 2, 3 \end{aligned} \tag{7}$$

since from (4)

$$f(x_{i-1}, x_i) = c_{i-1,i}(F_{i-1}(x_{i-1}), F_i(x_i))f_{i-1}(x_{i-1})f_i(x_i), \quad i = 2, 3$$

By replacing the conditional density term in (5) with (6) and (7), the final decomposition of a 3-dimensional PCC can be expressed as

$$\begin{aligned} f(x_1, x_2, x_3) &= c_{1,3|2}(F(x_1|x_2), F(x_3|x_2)) \cdot c_{1,2}(F_1(x_1), F_2(x_2)) \cdot \\ & \quad c_{2,3}(F_2(x_2), F_3(x_3)) \cdot f_1(x_1) \cdot f_2(x_2) \cdot f_3(x_3) \end{aligned} \tag{8}$$

where $c_{1,2}$ and $c_{2,3}$ are the baseline copulas and $c_{1,3|2}$ is the conditional copula. Since factor models assume the conditional independence of variables (see eq. 1) given common factor(s), it results that $c_{1,3|2}$ in (8) should be equal to one. The conditional distribution functions, which appear in (8), are obtained as

$$F(x_1|x_2) = \frac{\partial C_{1,2}(F_1(x_1), F_2(x_2))}{\partial F_2(x_2)} \quad \text{and} \quad F(x_3|x_2) = \frac{\partial C_{2,3}(F_2(x_2), F_3(x_3))}{\partial F_2(x_2)}$$

Due to many possible combinations of pair copula decomposition, Bedford and Cooke (2001) introduced a graphical model, denotes as regular vine, in order to organize and visualize them, with the following characteristics:

- an n -dimensional vine is represented by $n - 1$ trees;
- j -th tree has $(n + 1 - j)$ nodes and $(n - j)$ edges;

- each edge corresponds to a pair-copula density;
- edges in j -th tree become nodes in $(j + 1)$ -th tree;
- two nodes in $(j + 1)$ -th tree are joined by an edge only if the corresponding edges in j -th tree share a node;
- the complete decomposition is defined by the $n(n - 1)/2$ edges and the marginal densities.

Kurowicka and Cooke (2004) defined two particular classes of regular vine:

- "C-vine" (Canonical vine) (fig. 1) for which each tree has a unique node that is connected to $n - j$ edges;
- "D-vine" (fig. 2) for which no node in any tree is connected to more than two edges.

Figure 1 about here

Figure 2 about here

Here we are interested in the C-vine structure, based on the specification of a conditioning pilot variable at each level of the construction, characterized by the highest dependence with remaining variables. The density function of a C-vine is expressed as

$$f(x_1, \dots, x_n) = \prod_{j=1}^{n-1} \prod_{i=1}^{n-j} c_{j,j+i|1,\dots,j-1}(F(x_j|x_{1,\dots,j-1}), F(x_{j+i}|x_{1,\dots,j-1})) \cdot \prod_{k=1}^n f_k(x_k). \quad (9)$$

The conditional distribution functions, $F(\cdot|\cdot)$, which appear in (9), are computed using the following expression (see Joe, 1996)

$$F(x|v) = \frac{\partial C_{x,v_j|\mathbf{v}_{-j}}(F(x|\mathbf{v}_{-j}), F(v_j|\mathbf{v}_{-j}))}{\partial F(v_j|\mathbf{v}_{-j})} \quad \forall j \quad (10)$$

where \mathbf{v}_{-j} denotes the vector \mathbf{v} excluding the component v_j . The equation (9) shows that multivariate copulas can be obtained as product of iteratively conditioned bivariate copulas, where $c_{j,j+l-1}$ denotes the baseline copula and $c_{j,j+l|j+l-1}$ is the conditional copula (for $j = 1, \dots, n - 1$ and $l = 2, \dots, n$).

In general, for any vector \mathbf{v}_{-j} , x and v_j are conditionally independent given \mathbf{v}_{-j} if and only if

$$c_{x,v_j|\mathbf{v}_{-j}}(F(x|\mathbf{v}_{-j}), F(v_j|\mathbf{v}_{-j})) = \mathbf{1} \quad (11)$$

When (11) is verified, \mathbf{v}_{-j} is a vector of common factor(s) able to explain the cross co-movements of the random variables x and v_j (see Granger et al., 2006).

Our goal is to apply the three-factor model of Fama and French (1992) by means of vine copulas, verifying also the presence of a remaining dependence between common factors and among variables conditionally on common factors.

4 Tail dependence of vine copulas

Tail dependence is a useful copula-based measure which defines the relationship in extreme values of variables.

This measure of concordance between less probable values is different for each family of copulas since, while it exists for some of them, in a symmetric or asymmetric way, there are families that cannot allow tail dependence.

Joe (1997) and Joe et al. (2010) showed that vine copulas can have a flexible range of bivariate lower and upper tail dependence functions when asymmetric bivariate copulas with upper/lower tail dependence are used in the first level of vine.

Assuming that the bivariate linking copulas have continuous second-order partial derivatives, then a vine copula is tail dependent if all the bivariate baseline (base tree) linking copulas are tail dependent. If some baseline copulas are tail independent, then the vine copula is tail independent. Some margins of the vine, however, might still be tail dependent (Joe et al., 2010; Li and Wu, 2011).

The lower and upper tail dependence functions, denoted as $b^L(\cdot; C)$ and $b^U(\cdot; C)$ are defined as (Joe et al., 2010)

$$b^L(w; C) := \lim_{u \rightarrow 0^+} \frac{C(uw_i, 1 \leq i \leq n)}{u}, \quad \forall w = (w_1, \dots, w_n) \in R_+^n$$

and

$$b^U(w; C) := \lim_{u \rightarrow 0^+} \frac{\bar{C}(1 - uw_i, 1 \leq i \leq n)}{u}, \quad \forall w = (w_1, \dots, w_n) \in R_+^n$$

where C is an n -dimensional copula function and \bar{C} is its joint survival function.

Since a vine copula is expressed in terms of bivariate copula densities, its tail dependence can be derived from tail density approach (Li and Wu, 2011).

Let D_w denote the n -order partial differentiation operator with respect to $w = (w_1, \dots, w_n)$, then the lower and upper tail density functions are, respectively,

$$\begin{aligned} \delta^L(w; C) &:= \lim_{u \rightarrow 0^+} \frac{D_w C(uw_i, 1 \leq i \leq n)}{u} = \lim_{u \rightarrow 0^+} u^{n-1} c(uw_i, 1 \leq i \leq n) \\ &= \frac{\partial^n b^L(w; C)}{\partial w_1, \dots, \partial w_n}, \end{aligned} \quad (12)$$

and

$$\begin{aligned} \delta^U(w; C) &:= \lim_{u \rightarrow 0^+} \frac{D_w \bar{C}(1 - uw_i, 1 \leq i \leq n)}{u} = \lim_{u \rightarrow 0^+} u^{n-1} c(1 - uw_i, 1 \leq i \leq n) \\ &= \frac{\partial^n b^U(w; C)}{\partial w_1, \dots, \partial w_n}. \end{aligned} \quad (13)$$

Then, in order to obtain the tail density functions, we need to estimate the conditional tail dependence functions.

For non empty, non-overlapping subsets S_1 and S_2 , let $C_{S_1|S_2}$ ($\bar{C}_{S_1|S_2}$) be the conditional

copula (survival joint function), then the conditional lower and upper tail dependence functions, denoted as $t_{S_1|S_2}^L$ and $t_{S_1|S_2}^U$, are given by (see Joe et al., 2010)

$$t_{S_1|S_2}^L(w_{S_1}|w_{S_2}) = \lim_{u \rightarrow 0^+} C_{S_1|S_2}(uw_i, i \in S_1 | uw_j, j \in S_2)$$

and

$$t_{S_1|S_2}^U(w_{S_1}|w_{S_2}) = \lim_{u \rightarrow 0^+} \bar{C}_{S_1|S_2}(1 - uw_i, i \in S_1 | 1 - uw_j, j \in S_2).$$

Let $S' = (2, \dots, n-1)$ be a subset of $S = (1, \dots, n)$, then by applying the recursive formulas to build the vine copula structure and assuming that all bivariate baseline linking copulas have lower tail dependence, Li and Wu (2011) proved that

$$\frac{\delta_S^L(w_S)}{\delta_{S'}^L(w_{S'})} = c_{1n}(t_{1|S'}(w_1|w_{S'}), t_{n|S'}(w_n|w_{S'})) \frac{\delta_{1 \cup S'}^L(w)}{\delta_{S'}^L(w_{S'})} \frac{\delta_{n \cup S'}^L(w)}{\delta_{S'}^L(w_{S'})} \quad (14)$$

remarking that for $i \notin S'$ and $1 \leq i \leq n-l$ at each l level of the vine ($2 \leq l \leq n-1$)

$$\delta_{i|S'}^L(w_i|w_{S'}) = \frac{\partial}{\partial w_i} t_{i|S'}^L(w_i|w_{S'}) := \frac{\delta_{\{i\} \cup S'}^L(w_{\{i\} \cup S'})}{\delta_{S'}^L(w_{S'})} \quad (15)$$

and

$$\begin{aligned} \delta_{\{1,n\}|S'}^L(w_1, w_n|w_{S'}) &= \frac{\partial^2}{\partial w_1 \partial w_n} t_{\{1,n\}|S'}^L(w_1, w_n|w_{\{1,n\} \cup S'}) \\ &:= \frac{\delta_{\{1,n\} \cup S'}^L(w_{\{1,n\} \cup S'})}{\delta_{S'}^L(w_{S'})} \end{aligned} \quad (16)$$

The expression of the lower tail density of the 3-dimensional vine is:

$$\begin{aligned} \delta_{123}^L(w_1, w_2, w_3) &= \delta_{12}(w_1, w_2) \cdot \delta_{23}(w_2, w_3) \cdot \\ &\quad c_{13|2}(t_{1|2}^L(w_1|w_2), t_{3|2}^L(w_3|w_2)) \end{aligned} \quad (17)$$

where $t_{i|2}^L(w_i|w_2) = \int_0^1 \delta_{i2}^L(v_i, w_2) dv_i$ ($i = 1, 3$).

In high dimensional contexts, in contrast with Joe et al. (2010), the recursion involves only univariate integrals of n -variate tail densities.

Similar results are obtained for the upper tail density function (see for more details, Li and Wu, 2011).

Note that the expression (14) can be used to show how tail dependence of a vine depends on its bivariate baseline linking copulas.

The existence of tail dependence for all bivariate margins of the vine copulas is guaranteed by the presence of tail dependence only for the baseline copulas in level 1 and it is not necessary for the conditional bivariate copulas in $n-1$ levels (as stated in Joe et al., 2010).

This feature of vine copulas can suggest the use of conditional independence copulas

at $n - 1$ level of tree (11), according to the hypothesis of conditional independence of portfolio returns of factor models. In particular, Aas et al. (2006) recommend to truncate the construction of vine copulas, allowing for product copulas, C^\perp , at k level of the structure when the approximation error (measured as difference between conditional density and one, for independence conditional density) is lower than a given threshold. Brechmann et al. (2012) use independence copulas (C^\perp) at the k -th tree of the vine. They choose the truncation level k comparing the AIC and BIC values of two consecutive models (T_i and T_{i+1}) and the smaller model (T_i) is selected if the latter does not provide a significant gain in the model fit.

Another way to streamline the vine structure consists in a simplification technique, which allows to capture the residual dependence of asset returns conditioned on common factors. Simplification occurs replacing the remaining pair copulas at k level of the structure with a gaussian multivariate copula (as in Heinen and Valdesogo, 2009) or with gaussian bivariate copulas (see Brechmann and Czado, 2011).

5 Empirical results

Object of the analysis is the industry equity returns co-movements conditioned on Fama and French three common factors: excess of market return (Mkt), SmB and HmL .

We have selected five industry portfolio daily returns (January 3rd 2006 - December 30th 2011) provided by Fama and French data library. We report the SIC codes of each variable used afterward:

1. $Cnsmr \Rightarrow$ Consumer Durables and NonDurables, Wholesale, Retail, and Some Services (Laundries, Repair Shops)
2. $Manuf \Rightarrow$ Manufacturing, Energy, and Utilities
3. $HiTec \Rightarrow$ Business Equipment, Telephone and Television Transmission
4. $Hlth \Rightarrow$ Healthcare, Medical Equipment, and Drugs
5. $Other \Rightarrow$ Mines, Constr, BldMt, Trans, Hotels, Bus Serv, Entertainment and Finance.

Although a linear factor model is a good approximation of joint co-movements, it results that the variable HmL is not always significant (in particular, for $Manuf$, $HiTec$ and $Hlth$). Besides, there is evidence of non gaussian and autocorrelated residuals from all linear regression models, also for the common factors, which are not normally distributed and independent each other.

The estimation of the vine copula parameters can be carried out through a two-step method, defined Inference For Margins (IFM), which computes, separately, the marginal and the copula parameters. In the former, to take into account the conditional dependence of variables, GARCH models have been applied to shape the marginal behavior of each stock portfolio and factor returns. It results that all variables follow a $GJR-GARCH(2, 1)$ model with Student's- t distributed errors, except SmB and HmL

which are $GJR - GARCH(1,1)$ distributed with t -errors. Then, after transforming the standardized residuals into uniform margins, by applying the empirical distribution functions, the bivariate copula parameters are estimated through the maximization of the C -vine log-likelihood. At each step, the estimated parameters became the starting values to the full model estimation.

For each n dimensional random vector of variables, there are $n!/2$ different C -vines with $n(n-1)/2$ pair copulas (see Aas and Berg, 2009). In order to choose the best permutation for the C -vine factorization, it is useful to compute the Kendall's tau coefficient values for all bivariate pairs, selecting the most dependent pair of assets which become copula-nodes for each tree level (as stated in Aas and Berg, 2009).

Besides, C -vines are based on the specification of a conditioning pilot variable at each tree level, which maximizes the sum of the absolute Kendalls tau values ¹. According to the factor approach, in the base tree the pilot variable is Mkt , while in the second tree is Smb (see table 2). In order to select the appropriate family for each bivariate copula, we propose a sequential estimation procedure, described in Aas et al. (2006), choosing, at first, the parametric copula with the best fit to the data for each pair of margins and, then, maximizing the full log-likelihood by means of the parameters obtained from the stepwise procedure as starting values. Starting values of the bivariate copula parameters may be determined as follows (as in Aas et al., 2006):

1. Choose the copula types to use in the base tree by plotting the original data and by applying a Goodness-of-Fit (GoF) test, after estimating the parameters of candidate copulas.
2. Generate the observations for the second tree as conditional distribution functions of the original data using the estimated copula parameters.
3. Determine which copula types to use in the second tree in the same way as in the base tree.
4. Proceed iterating.

The main characteristic of a vine copula is the free specification of $n(n-1)/2$ bivariate copulas which do not have to belong to the same family and can have a flexible range of bivariate upper/lower tail dependence parameters, different for each margin pair when asymmetric bivariate copulas with upper/lower tail dependence are used in the base tree of the vine (see for more details Joe et al., 2010).

Asymmetric copulas generally involved in the analysis of tail dependence are: the Clayton with lower tail dependence, the Gumbel with upper tail dependence, the BB1 and the BB7 (see Joe, 1997) with both tail dependencies.

We have also selected, for sake of comparisons, the Gaussian copula without tail dependence (which is the standard market model), the Student's- t copula with symmetrical tail dependence and the product copula for independence. This last copula is selected according to the bivariate independence test based on Kendall's tau (Genest and Favre,

¹Kendall's Tau estimates are available on demand.

2007).

In order to select the copula with the best fit to the data, we have compared, at each tree level, the log-likelihood functions through Akaike Information Criterion (AIC) and the Bayesian Information Criterion (BIC). It results that the Student's- t copula is more preferred to other copula functions, this implies the prevalence of a symmetrical tail dependence. Tables from 3 to 9 show copula parameter estimates. We can note a significant dependence among common factors expressed by the last two copulas, contrary to the assumptions of the linear factor models.

Moreover, we have compared the C -vine factor copula model with other multivariate copulas and the results (table 10) remark the best performance of C -vine, especially over the standard linear gaussian model of Fama and French (1992). Besides, since a vine copula structure can be truncated, according to the hypothesis of conditional independence of portfolio returns, or simplified, in order to capture the remaining dependence of asset returns conditional on common factors, we have computed the maximum likelihood copula parameters of the truncated (with an approximation error lower than 0.01) and simplified version of the full C -vine model. The best alternative to the full C -vine model, in a factor model perspective, is the simplified C -vine with gaussian copula pairs (table 10). This implies a conditional dependence which is not accurately captured by the standard factor model.

Finally, we provide the tail density functions of full C -vine model, according to equations (14) - (17), to measure the dependence between extreme realizations of portfolio returns. In table 11 we report the lower tail density functions of all margins which are building blocks of the conditional copulas included in the model. Since all bivariate baselines linking copulas are tail dependent (Student's- t), also the vine copula is tail dependent, although weakly (see, for the lower tail density, $\delta_{12345MSH}$ in table 11).

6 Outlook and Conclusions

In this paper we combine the factor approach with vine copulas in order to obtain the non linear three-factor model of Fama and French (1992). This is the first contribution which draws a three-factor model by use of vine copulas. Second, we provide empirical estimates of tail dependence measures for the vine copulas, based on tail density functions, which involve only one-dimensional integrations.

Different reasons have led to develop a three-factor copula model by means of a vine structure. First of all, the need to consider non-linear dependence measures and some alternatives to the elliptical distributions, both in margins and in the joint distributions. Vines allow for a free specification of bivariate copulas involved in the model, which can belong to distinct families with a flexible range of tail dependence parameters, different also for each margin pair. Second, the extension of copulas in multivariate contexts requires additional assumptions and it is not an easy task. The simple structure of vine copulas, based only on bivariate copulas, allows to compute easily, by MLE method, copula parameters for high-dimensional continuous variables. In brief, combining a vine structure with the factor copula approach enables a flexible construction of high di-

mensional, also non elliptical, joint distributions of cross-section of the expected stock returns, able to capture possible asymmetries in tail dependence. The main result of the application is that the vine factor copula model performs better than other multivariate models, also more than the classical linear factor approach. Besides, it emerges the existence of a conditional dependence between variables with respect to common factors in addition to that observed among the common factors.

Future researches will concentrate on some alternative models, which can include: the momentum factor, defined as the amount of acceleration of stocks (see for instance Carhart, 1997), time-varying factors or other common factors resulting from a Principal Component Analysis (PCA) applied to the stock asset returns.

References

- Aas, K., Berg, D. (2009). Models for construction of multivariate dependence: A comparison study. *European Journal of Finance*, 15, 639-659. DOI:10.1080/13518470802588767
- Aas, K., Czado, C., Frigessi, A., Bakken, H. (2006). Pair-copula constructions of multiple dependence. *Insurance: Mathematics & Economics*, 44, 182-198.
- Andersen, L., Sidenius, J. (2004). Extensions of the Gaussian Copula Model. *Journal of Credit Risk*, 1, 29-70.
- Banz, W. (1981). The relationship between return and market value of common stocks. *Journal of Financial Economics*, 9, 3-18.
- Bedford, T., Cooke, R. M. (2001). Probability density decomposition for conditionally dependent random variables modeled by vines. *Annals of Mathematics and Artificial Intelligence*, 32, 245-268. DOI:10.1023/A:1016725902970
- Bedford, T., Cooke, R. M. (2002). Vines - a new graphical model for dependent random variables. *Annals of Statistics*, 30, 1031-1068. DOI:10.1214/aos/1031689016
- Brechmann, E.C., Czado, C. (2011). Risk management with high-dimensional vine copulas: An analysis of the Euro Stoxx 50. *Working paper, Technische Universität München*.
- Brechmann, E.C., Czado, C., Aas, K. (2012). Truncated Regular Vines in High Dimensions with Applications to Financial Data. *Canadian Journal of Statistics*, 40, 68-85.
- Campbell, R., Koedijk, K., Kofman, P. (2002). Increased correlation in bear markets. *Financial Analysts Journal*, 58, January-February, 87-94. ISSN: 0015198X
- Carhart, M.M. (1997). On persistence in mutual fund performance *Journal of Finance*, 52, 57-82.
- Chan, L.K., Hamao, Y., Lakonishok, J. (1991). Fundamentals and stock returns in Japan. *Journal of Finance*, 46, 1739-1789.
- Cherubini, U., Luciano, E., Vecchiato, W. (2004). *Copula Methods in Finance*. John Wiley & Sons.

- Cherubini, U., Gobbi, F., Mulinacci, S., Romagnoli, S. (2011). *Dynamic Copula Methods in Finance*. John Wiley & Sons.
- Chollete, L., Ning, C. (2010). Asymmetric Dependence in US Financial Risk Factors?, *US Working Papers in Economics and Finance*, 2/11, University of Stavanger.
- Dias, A., Embrechts, P. (2010). Modeling exchange rate dependence dynamics at different time horizons. *Journal of International Money and Finance*, 29, 1687–1705. DOI: 10.1016/j.jimonfin.2010.06.004
- Engle, R.F. (2002). Dynamic conditional correlation: a simple class of multivariate generalized autoregressive conditional heteroscedasticity models. *Journal of Business and Economic Statistics*, 20, 339–350. DOI:10.1198/073500102288618487
- Fama, E.F., French, K.R. (1992). The cross-section of expected stock returns. *Journal of Finance*, 47, 427-465.
- Genest, G., Favre, A. (2007). Everything You Always Wanted to Know about Copula Modeling but Were Afraid to Ask. *J. Hydrol. Eng.*, 12, Special issue: Copulas in Hydrology, 347-368. DOI: 10.1061/(ASCE)1084-0699(2007)12:4(347)
- Granger, C.W.J., Teräsvirta, T., Patton, A. (2006). Common factors in conditional distributions for bivariate time series. *Journal of Econometrics*, 132, 43-57. DOI:10.1016/j.jeconom.2005.01.022
- Gregory, J., Laurent, J.P. (2005). Basket Default Swaps, CDOs and Factor Copulas. *Journal of Risk*, 7, 8-23.
- Heinen, A., Valdesogo, A. (2009). Asymmetric CAPM dependence for large dimensions: the Canonical Vine Autoregressive Model. *CORE Discussion Papers 2009069*, Universit catholique de Louvain, Center for Operations Research and Econometrics (CORE). Available at <http://dx.doi.org/10.2139/ssrn.1297506>
- Heinen, A., Valdesogo, A. (2011). *Dynamic D-vine Model*. In: Kurowicka, D., Joe, H. (Eds.). *Dependence Modeling: Vine Copula Handbook*. World Scientific, 329–353.
- Hull, J., White, A. (2004). Valuation of a CDO and nth to default CDO without Monte Carlo Simulation. *Journal of Derivatives*, 12, 10-23.
- Hull, J., White, A. (2006). Valuing Credit Derivatives Using an Implied Copula Approach. *Journal of Derivatives*, 14 2, 8-28.
- Hull, J., White, A. (2010). An improved implied copula model and its application to the valuation of bespoke CDO tranches. *Journal of investment management*.
- Joe, H. (1996). *Families of m-variate distributions with given margins and $m(m-1)/2$ bivariate dependence parameters*. In: Rüschendorf, L., Schweizer, B., Taylor, M.D.(Eds.). *Distributions with Fixed Marginals and Related Topics*.
- Joe, H. (1997). *Multivariate Models and Dependence Concepts*. Chapman & Hall/CRC, New York.
- Joe, H., Li, H., Nikoloulopoulos, A. K. (2010). Tail dependence functions and vine copulas. *Journal of Multivariate Analysis*, 101, 252–270. <http://dx.doi.org/10.1016/j.jmva.2009.08.002>
- Joe, H., Li, H., Nikoloulopoulos, A.K. (2010). Vine copulas with asymmetric tail depen-

- dence and applications to financial return data. *Computational Statistics and Data Analysis*.
- Kurowicka, D., Cooke, R. M. (2004). Distribution-free continuous bayesian belief nets. In Fourth International Conference on Mathematical Methods in Reliability Methodology and Practice, Santa Fe, New Mexico.
- Kurowicka, D., Cooke, R. M. (2006). *Uncertainty Analysis with High Dimensional Dependence Modelling*. Wiley, New York.
- Li, D.X. (2000). On Default Correlation: A Copula Approach. *Journal of Fixed Income*, 9, 43–54.
- Li, D.X., Wu, P. (2011). Extremal Dependence of Copulas: A Tail Density Approach. *Working Paper*.
- Lintner, J. (1965). The valuation of risk assets and the selection of risky investments in stock portfolios and capital. *Review of Economics and Statistics*, 47, 13–37.
- Longin, F., Solnik, B. (2001). Extreme correlation of international equity markets. *The Journal of Finance*, 56, 649–676. DOI: 10.1111/0022-1082.00340.
- McNeil, A. J., Frey, R., Embrechts, P. (2005). Quantitative Risk Management. *Princeton University Press*, New Jersey.
- Nelsen, R. B. (2006). *An Introduction to Copulas*. Springer-Verlag, New York.
- Oh, D.H., Patton, A. (2011). Modelling Dependence in High Dimensions with Factor Copulas. *Working Paper*. Revised April 2012.
- Patton, A. (2006). Modelling asymmetric exchange rate dependence. *International Economic Review*, 47, 527–556. DOI:10.1111/j.1468-2354.2006.00387.x
- Richardson, M., Smith, T. (1993). A test for multivariate normality in stock returns. *Journal of Business*, 66, 2, 295–321.
- Sharpe, W.F. (1964). Capital asset prices: a theory of market equilibrium under conditions of risk. *Journal of Finance*, 19, 425–442.
- Sklar, A. (1959). Fonctions de répartition à n dimensions et leurs marges. *Publ. Inst. Statist. Univ. Paris*, 8, 229–231.
- Vasicek, O. (1987). The Loan Loss Distribution. *Working Paper*, KMV.
- van der Voort, M. (2005). Factor copulas: totally external defaults. *Working Paper*, Erasmus University Rotterdam.

Table 1: Example of some copula densities known in a closed form as function of h , F_y and of F_ϵ (Oh and Patton, 2011). *Gamma distribution, **Inverse Gamma distribution

Copula	$h(Y,\epsilon)$	F_Y	F_ϵ
Gaussian	$Y + \epsilon$	$N(0, \sigma_Y^2)$	$N(0, \sigma_\epsilon^2)$
Student's-t	$Y^{1/2}\epsilon$	$Ig^{**}(v/2, v/2)$	$N(0, \sigma_\epsilon^2)$
Clayton	$(1 + \epsilon/Y)^{-a}$	$\Gamma^*(a, 1)$	$Exp(1)$
Gumbel	$-(\log Y/\epsilon)^a$	$Stable(1/a, 1, 1, 0)$	$Exp(1)$

Table 2: Kendall's Tau Estimates. Mkt=Excess of Market return with respect to risk free rate, SmB=Small minus Big portfolios, HmL=High minus Low portfolios

Portfolio	Mkt	SmB	HmL
Cnsmr	0.750	0.409	0.172
Manuf	0.762	0.348	0.195
HiTec	0.730	0.415	0.120
Hlth	0.637	0.393	0.108
Other	0.761	0.392	0.222

Table 3: Copula parameter estimates in the Base Tree

C-vine margin	Copula	d.o.f	rho
c_{1M}	Student's-t	13.620	0.926
c_{2M}	Student's-t	6.021	0.933
c_{3M}	Student's-t	8.326	0.915
c_{4M}	Student's-t	7.001	0.844
c_{5M}	Student's-t	6.551	0.930
c_{SM}	Student's-t	7.113	0.372
c_{HM}	Student's-t	5.393	0.292

1=Cnsmr, 2=Manuf, 3=HiTec, 4=Hlth, 5=Other, M=Mkt, S=SmB, H=HmL

Table 4: Copula parameter estimates in the Second Tree

C-vine			
margin	Copula	d.o.f/theta	rho
$c_{1S M}$	Student's- t	20.171	0.616
$c_{2S M}$	Student's- t	7.805	0.446
$c_{3S M}$	Student's- t	9.309	0.605
$c_{4S M}$	Student's- t	12.610	0.507
$c_{5S M}$	Student's- t	18.891	0.563
$c_{HS M}$	Clayton	0.308	-

Table 5: Copula parameter estimates in the Third Tree

C-vine			
margin	Copula	d.o.f	rho
$c_{1H MS}$	Gaussian	-	0.215
$c_{2H MS}$	Student's- t	11.668	0.347
$c_{3H MS}$	Student's- t	10.982	0.088
$c_{4H MS}$	Student's- t	13.295	0.190
$c_{5H MS}$	Student's- t	22.931	0.436

Table 6: Copula parameter estimates in the Fourth Tree

C-vine			
margin	Copula	d.o.f	rho
$c_{15 MSH}$	Student's- t	14.242	0.423
$c_{25 MSH}$	BB1	0.659	1.033
$c_{35 MSH}$	Student's- t	25.697	0.566
$c_{45 MSH}$	Student's- t	36.098	0.518

Table 7: Copula parameter estimates in the Fifth Tree

C-vine			
margin	Copula	d.o.f/theta	rho
$c_{14 MSH5}$	Clayton	0.744	-
$c_{24 MSH5}$	Clayton	0.331	-
$c_{34 MSH5}$	Student's- <i>t</i>	21.705	0.650

Table 8: Copula parameter estimates in the Sixth Tree

C-vine			
margin	Copula	d.o.f	rho
$c_{13 MSH54}$	Student's- <i>t</i>	38.651	0.159
$c_{23 MSH54}$	Student's- <i>t</i>	34.939	0.009

Table 9: Copula parameter estimates in the Seventh Tree

C-vine			
margin	Copula	d.o.f	rho
$c_{12 MSH543}$	Student's- <i>t</i>	34.330	0.354

Table 10: Model Comparison

Model	No.Param.	LogL	AIC	BIC
Multivariate Gaussian copula	28	9079	-18102	-17953
Multivariate Student's- <i>t</i> copula	29	9404	-18750	-18596
Full <i>C</i> -vine	52	11757	-23410	-23134
Trunc. <i>C</i> -vine at T3 (approx error)	34	9047	-18025	-17844
Trunc. <i>C</i> -vine at T5 (AIC/BIC)	46	11479	-22866	-22622
Simpl. <i>C</i> -vine at T3 with gaussian pair-copulas	44	11660	-23232	-22997
Simpl. <i>C</i> -vine at T3 with multiv. gaussian copula	44	11095	-22102	-21868

Table 11: Tail densities of all margins ($w = 1$)

Margin	δ_{iM}	δ_{iMS}	δ_{iMSH}	δ_{iMSH5}	δ_{iMSH54}	$\delta_{iMSH543}$	$\delta_{iMSH5432}$
1	4.9E-04	5.3E-05	5.3E-06	1.1E-06	1.8E-07	3.6E-08	4.4E-09
2	1.1E-02	4.8E-04	4.7E-05	2.8E-06	4.6E-07	8.9E-08	-
3	4.0E-03	3.5E-04	2.5E-05	6.2E-06	1.4E-06	-	-
4	6.9E-03	3.8E-04	3.1E-05	7.2E-06	-	-	-
5	8.4E-03	4.9E-04	5.5E-05	-	-	-	-
S	6.6E-03	-	-	-	-	-	-
H	1.4E-02	4.0E-04	-	-	-	-	-

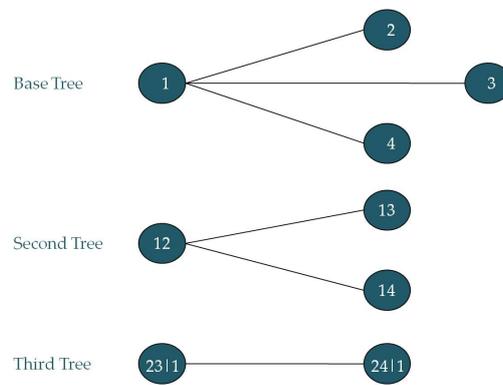


Figure 1: C-VINE copula for 4 variables

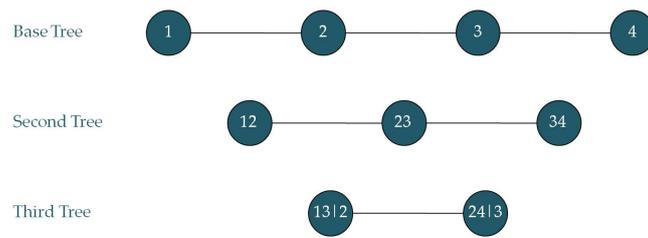


Figure 2: D-VINE copula for 4 variables