

Research Article

Open Access

Sándor Guzmics* and Georg Ch. Pflug

Modelling cascading effects for systemic risk: Properties of the Freund copula

<https://doi.org/10.1515/demo-2019-0002>

Received September 12, 2018; accepted January 15, 2019

Abstract: We consider a dependent lifetime model for systemic risk, whose basic idea was for the first time presented by Freund. This model allows to model cascading effects of defaults for arbitrarily many economic agents. We study in particular the pertaining bivariate copula function. This copula does not have a closed form and does not belong to the class of Archimedean copulas, either. We derive some monotonicity properties of it and show how to use this copula for modelling the cascade effect implicitly contained in observed CDS spreads.

Keywords: dependent lifetime models, upper orthant order, systemic risk

MSC: 60E05, 62E10

1 Introduction

We consider a system of n entities and their dependent lifetimes. The term "entity" can be understood as broadly as possible, i.e., the system can consist of banks, financial institutions, electrical devices, state sovereigns, living beings, etc., but is always assumed to be homogeneous in the sense that all entities are of the same type. We call the end of the life of an entity a "default", even if the entity is not a firm.

The fundamental idea of the model is that the individual lifetime distributions are affected by the defaults of other entities. To be more precise, we assume initially individual exponential lifetimes with default intensity λ_k for every entity k . The choice of an exponential lifetime is motivated by the fact that the corresponding hazard function is constant and therefore conditional residual lifetimes do not depend on the conditioning. Notice that the same assumption is made for the celebrated Marshall-Olkin-type models.

The main idea for such type of models is that the default of one entity puts more stress on the other entities. In a competitive market one may sometimes observe the opposite: the default of an entity eliminates a competitor and reduces the stress for others. However, we motivate our model by an application to financial institutions, where the default of an entity typically implies losses for the other entities and therefore increases the stress.

If entity k defaults, then the residual default intensity of all other entities increases by a value of $a_{k,l}$. If (X_1, \dots, X_n) is the resulting vector of lifetimes, then their n -dimensional copula C is determined by the vector of intensities $(\lambda_1, \dots, \lambda_n)$ and the cascading effects $a_{k,l}$ for $k \neq l$. The n^2 parameters can be arranged

*Corresponding Author: Sándor Guzmics: University of Vienna. Department of Statistics and Operations Research (DSOR), Oskar Morgenstern Platz 1, A-1090 Wien-Vienna, Austria, E-mail: sandor.guzmics@univie.ac.at

Georg Ch. Pflug: DSOR and International Institute for Applied Systems Analysis (IIASA), Laxenburg, Austria, E-mail: georg.pflug@univie.ac.at

in a (typically asymmetric) matrix

$$A = \begin{pmatrix} \lambda_1 & a_{1,2} & \dots & a_{1,n} \\ a_{2,1} & \lambda_2 & \dots & a_{2,n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n,1} & a_{n,2} & \dots & \lambda_n \end{pmatrix}.$$

The described model is related to the well-known Marshall-Olkin model [8], where certain subsets of entities in the system can receive simultaneous shocks, i.e., default at the same time. We argue that simultaneous defaults do not happen in many applications, especially not in financial systems and we consider our model as more appropriate. A cascading default model appears already in an earlier work of Yu [12]. Freund [4] has suggested the same model for $n = 2$ that we consider in this current paper. In his honour, we call the pertaining copula after him.

The paper is organised as follows. In Section 2, we provide the formal definition and the fundamentals of the model for $n = 2$, and we elaborate on some details, including also the copula of the lifetime variables. In the last part of the section, we show how the setting can be generalized for more entities ($n \geq 2$). In Section 3, we examine how the dependency structure changes as the model parameters vary. Our ultimate question is: does any monotonic behaviour hold for the lifetime variables and their copula with respect to some stochastic dependence order relation? The reader will find a positive answer for the upper orthant order. In Section 4, we give a numerical illustration using CDS-data of three European banks. Section 5 concludes the paper.

2 The fundamentals of the model

Since our main application is the systemic risk of financial institutions, we use from now on the term institution for the entities.

In Subsections 2.1 to 2.4 we present a detailed analysis for the bivariate model, parts of which were already published by Freund [4]. In Subsection 2.5, we sketch the idea of a multivariate ($n \geq 2$) setting.

2.1 The bivariate model ($n = 2$)

Consider a system of two entities, and let $Y_k \sim \text{Exp}(\lambda_k)$ ($k = 1, 2$) be independent random variables. They are attributed as auxiliary lifetime variables (if one wishes as pre-lifetime variables) to the two entities of the system. When in a certain realization the first entity defaults earlier, i.e., $Y_1 < Y_2$, then the second entity will continue its operation according to another exponentially distributed random variable $Z_2 \sim \text{Exp}(\lambda_2 + a_2)$, which is independent of Y_1 and Y_2 . The parameter $a_2 \geq 0$ is called the shock parameter, and it expresses the effect of the default of the first institution on the second institution. Z_1 is defined analogously: when $Y_2 < Y_1$, then $Z_1 \sim \text{Exp}(\lambda_1 + a_1)$, where $a_1 \geq 0$ is a shock parameter.

The actual lifetime variables of the two entities are denoted by X_1, X_2 , and - in the light of the above mechanism - they can be written as follows.

$$\begin{aligned} \text{If } Y_1 < Y_2, \text{ then } & \begin{cases} X_1 := Y_1, \\ X_2 := Y_1 + Z_2, \text{ where } Z_2 \sim \text{Exp}(\lambda_2 + a_2) \text{ independent of } Y_1, Y_2. \end{cases} & (1) \\ \text{If } Y_2 < Y_1, \text{ then } & \begin{cases} X_2 := Y_2, \\ X_1 := Y_2 + Z_1, \text{ where } Z_1 \sim \text{Exp}(\lambda_1 + a_1) \text{ independent of } Y_1, Y_2. \end{cases} \end{aligned}$$

The new lifetime variables X_1, X_2 can be expressed explicitly in terms of Y_1, Y_2, Z_1, Z_2 :

$$\begin{cases} X_1 = Y_1 \cdot \mathbb{1}_{\{Y_1 < Y_2\}} + (Y_2 + Z_1) \cdot \mathbb{1}_{\{Y_2 < Y_1\}}, \\ X_2 = Y_2 \cdot \mathbb{1}_{\{Y_2 < Y_1\}} + (Y_1 + Z_2) \cdot \mathbb{1}_{\{Y_1 < Y_2\}}. \end{cases} \quad (2)$$

The case $Y_1 = Y_2$ does not need to be taken into account, since it has probability zero.

2.2 Cumulative distribution functions and probability density functions

In this subsection, we explore the joint distribution and the marginal distributions of the new - already dependent - bivariate lifetime variable (X_1, X_2) given in (2), as well as some remarkable properties of the joint and marginal cumulative distribution functions and probability density functions. We note that the joint density and the marginal densities (with another parameter setting) directly appear in Freund's work (look at formulas (1.9), (2.5) and (2.6) in [4]). We prove the formula for the joint density in a different way than he did. We emphasize again, that the model we consider can be described by the quadruple $[\lambda_1, \lambda_2, a_1, a_2]$.

Focusing now on the above mentioned densities and cumulative distribution functions, via some elementary computation one gets the following.

Proposition 1.

(i) (a) **Joint cumulative distribution function** of (X_1, X_2) (if $\lambda_1 \neq a_2$ and $\lambda_2 \neq a_1$):

$$H(x, y) = \begin{cases} 1 + \frac{\lambda_1}{\lambda_1 - a_2} \cdot e^{-(\lambda_1 - a_2)x} \cdot e^{-(\lambda_2 + a_2)y} + \frac{a_1}{\lambda_2 - a_1} \cdot e^{-(\lambda_1 + \lambda_2)x} - \\ \frac{\lambda_2}{\lambda_2 - a_1} \cdot e^{-(\lambda_1 + a_1)x} - \frac{\lambda_1}{\lambda_1 - a_2} \cdot e^{-(\lambda_2 + a_2)y} & \text{when } 0 \leq x \leq y, \\ 1 + \frac{\lambda_2}{\lambda_2 - a_1} \cdot e^{-(\lambda_2 - a_1)y} \cdot e^{-(\lambda_1 + a_1)x} + \frac{a_2}{\lambda_1 - a_2} \cdot e^{-(\lambda_1 + \lambda_2)y} - \\ \frac{\lambda_1}{\lambda_1 - a_2} \cdot e^{-(\lambda_2 + a_2)y} - \frac{\lambda_2}{\lambda_2 - a_1} \cdot e^{-(\lambda_1 + a_1)x} & \text{when } 0 \leq y \leq x. \end{cases} \quad (3)$$

(i) (b) **Joint cumulative distribution function** of (X_1, X_2) (if $\lambda_1 = a_2$ and $\lambda_2 \neq a_1$):

$$H(x, y) = \begin{cases} 1 + \frac{a_1}{\lambda_2 - a_1} \cdot e^{-(\lambda_1 + \lambda_2)x} - \frac{\lambda_2}{\lambda_2 - a_1} \cdot e^{-(\lambda_1 + a_1)x} - \lambda_1 \cdot x \cdot e^{-(\lambda_1 + \lambda_2)y} & \text{when } 0 \leq x \leq y, \\ 1 + \frac{\lambda_2}{\lambda_2 - a_1} \cdot e^{-(\lambda_2 - a_1)y} \cdot e^{-(\lambda_1 + a_1)x} - \frac{\lambda_2}{\lambda_2 - a_1} \cdot e^{-(\lambda_1 + a_1)x} - \\ e^{-(\lambda_1 + \lambda_2)y} - \lambda_1 \cdot y \cdot e^{-(\lambda_1 + \lambda_2)y} & \text{when } 0 \leq y \leq x. \end{cases} \quad (4)$$

(i) (c) **Joint cumulative distribution function** of (X_1, X_2) (if $\lambda_2 = a_1$ and $\lambda_1 \neq a_2$): analogue with formula (4).

(i) (d) **Joint cumulative distribution function** of (X_1, X_2) (if $\lambda_1 = a_2$ and $\lambda_2 = a_1$):

$$H(x, y) = \begin{cases} 1 - \lambda_1 \cdot x \cdot e^{-(\lambda_1 + \lambda_2)y} - \lambda_2 \cdot x \cdot e^{-(\lambda_1 + \lambda_2)x} - e^{-(\lambda_1 + \lambda_2)x} & \text{when } 0 \leq x \leq y, \\ 1 - \lambda_2 \cdot y \cdot e^{-(\lambda_1 + \lambda_2)x} - \lambda_1 \cdot y \cdot e^{-(\lambda_1 + \lambda_2)y} - e^{-(\lambda_1 + \lambda_2)y} & \text{when } 0 \leq y \leq x. \end{cases}$$

(ii) **Joint density function** of (X_1, X_2) : (Look also at the formula (1.9) in Freund [4].)

$$h(x, y) = \begin{cases} \lambda_1(\lambda_2 + a_2) \cdot e^{-(\lambda_1 - a_2)x} \cdot e^{-(\lambda_2 + a_2)y}, & 0 \leq x \leq y, \\ \lambda_2(\lambda_1 + a_1) \cdot e^{-(\lambda_2 - a_1)y} \cdot e^{-(\lambda_1 + a_1)x}, & 0 \leq y < x. \end{cases} \quad (5)$$

(iii) (a) **Marginal cumulative distribution functions** of X_1 (if $\lambda_2 \neq a_1$) and of X_2 (if $\lambda_1 \neq a_2$):

$$F(x) = 1 - \frac{\lambda_2}{\lambda_2 - a_1} \cdot e^{-(\lambda_1 + a_1)x} + \frac{a_1}{\lambda_2 - a_1} \cdot e^{-(\lambda_1 + \lambda_2)x}, \quad x \geq 0, \quad (6)$$

$$G(y) = 1 - \frac{\lambda_1}{\lambda_1 - a_2} \cdot e^{-(\lambda_2 + a_2)y} + \frac{a_2}{\lambda_1 - a_2} \cdot e^{-(\lambda_1 + \lambda_2)y}, \quad y \geq 0. \quad (7)$$

(iii) (b) **Marginal cumulative distribution functions** of X_1 (if $\lambda_2 = a_1$) and of X_2 (if $\lambda_1 = a_2$):

$$F(x) = 1 - e^{-(\lambda_1 + \lambda_2)x} - \lambda_2 \cdot x \cdot e^{-(\lambda_1 + \lambda_2)x}, \quad x \geq 0, \quad (8)$$

$$G(y) = 1 - e^{-(\lambda_1 + \lambda_2)y} - \lambda_1 \cdot y \cdot e^{-(\lambda_1 + \lambda_2)y}, \quad y \geq 0. \quad (9)$$

(iv) (a) **Probability density functions** of X_1 (if $\lambda_2 \neq a_1$) and of X_2 (if $\lambda_1 \neq a_2$). (Look also at the formulas (2.5) and (2.6) in Freund [4].)

$$f(x) = -\frac{a_1(\lambda_1 + \lambda_2)}{\lambda_2 - a_1} \cdot e^{-(\lambda_1 + \lambda_2)x} + \frac{\lambda_2(\lambda_1 + a_1)}{\lambda_2 - a_1} \cdot e^{-(\lambda_1 + a_1)x}, \quad x \geq 0, \quad (10)$$

$$g(y) = -\frac{a_2(\lambda_1 + \lambda_2)}{\lambda_1 - a_2} \cdot e^{-(\lambda_1 + \lambda_2)y} + \frac{\lambda_1(\lambda_2 + a_2)}{\lambda_1 - a_2} \cdot e^{-(\lambda_2 + a_2)y}, \quad y \geq 0. \quad (11)$$

(iv) (b) **Probability density functions** of X_1 (if $a_1 = \lambda_2$) and of X_2 if ($a_2 = \lambda_1$):

$$f(x) = \lambda_1 \cdot e^{-(\lambda_1 + \lambda_2)x} + (\lambda_1 + \lambda_2) \cdot \lambda_2 \cdot x \cdot e^{-(\lambda_1 + \lambda_2)x}, \quad x \geq 0. \quad (12)$$

$$g(y) = \lambda_2 \cdot e^{-(\lambda_1 + \lambda_2)y} + (\lambda_1 + \lambda_2) \cdot \lambda_1 \cdot y \cdot e^{-(\lambda_1 + \lambda_2)y}, \quad y \geq 0. \quad (13)$$

Proof.

Since formula (5) given in statement (ii) is valid for all parameter constellations, we will prove this statement directly. The formulas given in (i) (a), (i) (b), (i) (c) and (i) (d), as well as the formulas in (iv) (a), (iv) (b) and (iv) (c) can be derived from (5) by integration. Finally, the formulas in (iii) can be derived (for instance) by taking the suitable limits in the formulas given in (i).

Turning to the proof of (ii), let us assume that $x < y$, and let $\Delta x > 0$, $\Delta y > 0$ such that $x + \Delta x < y$. (The proof in the case when $x \geq y$ is analogue.)

$$\begin{aligned} & \mathbb{P}(X_1 \in [x, x + \Delta x], X_2 \in [y, y + \Delta y]) = \\ & = \mathbb{P}(Y_1 \in [x, x + \Delta x], Y_2 > x, Z_2 \in [y - x - \Delta x, y - x + \Delta y]) = \\ & \stackrel{Y_1, Y_2, Z_2}{\text{are indep.}} = \mathbb{P}(Y_1 \in [x, x + \Delta x]) \cdot \mathbb{P}(Y_2 > x) \cdot \mathbb{P}(Z_2 \in [y - x - \Delta x, y - x + \Delta y]). \end{aligned}$$

Dividing by $\Delta x \cdot \Delta y$, and then letting $\Delta x \rightarrow 0$, $\Delta y \rightarrow 0$, we get that

$$\begin{aligned} h(x, y) & = \lim_{\substack{\Delta x \rightarrow 0 \\ \Delta y \rightarrow 0}} \frac{\mathbb{P}(X_1 \in [x, x + \Delta x], X_2 \in [y, y + \Delta y])}{\Delta x \cdot \Delta y} = \\ & = \lim_{\substack{\Delta x \rightarrow 0 \\ \Delta y \rightarrow 0}} \frac{\mathbb{P}(Y_1 \in [x, x + \Delta x]) \cdot \mathbb{P}(Y_2 > x) \cdot \mathbb{P}(Z_2 \in [y - x - \Delta x, y - x + \Delta y])}{\Delta x \cdot \Delta y} = \\ & = \lambda_1 \cdot e^{-\lambda_1 x} \cdot e^{-\lambda_2 x} \cdot (\lambda_2 + a_2) \cdot e^{-(\lambda_2 + a_2)(y-x)} = \lambda_1 \cdot (\lambda_2 + a_2) \cdot e^{-(\lambda_1 - a_2)x} \cdot e^{-(\lambda_2 + a_2)y}, \text{ as it was stated in (ii).} \end{aligned}$$

The second last equality holds, because $Y_1 \sim \text{Exp}(\lambda_1)$, $Y_2 \sim \text{Exp}(\lambda_2)$, $Z_2 \sim \text{Exp}(\lambda_2 + a_2)$. \square

Figure 1 depicts the joint density function of (X_1, X_2) for two parameter settings.

All the cdfs and pdfs in formulas (3)–(13) continuously depend on the parameter a_1 and a_2 . For instance, this continuity is trivial for the marginal density functions (10) and (11) when $a_1 \neq \lambda_2$ and $a_2 \neq \lambda_1$. At the places $a_1 = \lambda_2$ resp. $a_2 = \lambda_1$ one can see the continuity by taking the limits $a_1 \rightarrow \lambda_2$ and $a_2 \rightarrow \lambda_1$ in formula (10) and (11), which then yield formula (12) and (13).

Similarly, the special case of the joint cdf presented in (i) (b) can be also obtained from the general formula

$$\begin{aligned} (3), \text{ if we let } a_2 \rightarrow \lambda_1, \text{ since } \lim_{a_2 \rightarrow \lambda_1} \frac{e^{-(\lambda_1 - a_2)x} - 1}{\lambda_1 - a_2} & = -x, \\ \text{and } \lim_{a_2 \rightarrow \lambda_1} \frac{a_2 \cdot e^{-(\lambda_1 + \lambda_2)y} - \lambda_1 \cdot e^{-(\lambda_2 + a_2)y}}{\lambda_1 - a_2} & = -e^{-(\lambda_1 + \lambda_2)y} - \lambda_1 \cdot y \cdot e^{-(\lambda_1 + \lambda_2)y}. \end{aligned}$$

The marginal density (10) reduces to $f(x) = \lambda_1 \cdot e^{-\lambda_1 x}$ if $a_1 = 0$, so in this case X_1 is exponentially distributed with parameter λ_1 . It might be surprising at first glance, when someone only considers the construction (2) of X_1 . The background of this feature is the constant hazard rate property of the exponential distribution. Nevertheless, a disturbance parameter a_1 of value 0 has indeed no effect on the marginal distribution of the first entity, since in this case $X_1 \sim Y_1$. However, X_1 and X_2 are not independent, unless $a_2 = 0$ also holds.

We examine the other extreme case as well, i.e., when $a_1 \rightarrow \infty$. Then $Z_1 \stackrel{\text{a.s.}}{=} 0 \sim \text{"Exp}(\infty)\text{"}$ is added to the (truncated) Y_1 , which means finally that either the first entity expires earlier ($Y_1 < Y_2$), or X_1 takes the value of Y_2 , so all in all $X_1 = \min\{Y_1, Y_2\}$, which is exponentially distributed with parameter $\lambda_1 + \lambda_2$. This fact is also reflected by the marginal density f (look at (10)), which reduces to $f(x) = (\lambda_1 + \lambda_2) \cdot e^{-(\lambda_1 + \lambda_2)x}$ in this case.

It is also worth to see that the joint cdf reduces to the following symmetric function when $x = y$:

$$H(x, x) = 1 + \frac{\lambda_1 \lambda_2 - a_1 a_2}{(\lambda_1 - a_2)(\lambda_2 - a_1)} e^{-(\lambda_1 + \lambda_2)x} - \frac{\lambda_2}{\lambda_2 - a_1} e^{-(\lambda_1 + a_1)x} - \frac{\lambda_1}{\lambda_1 - a_2} e^{-(\lambda_2 + a_2)x}. \tag{14}$$

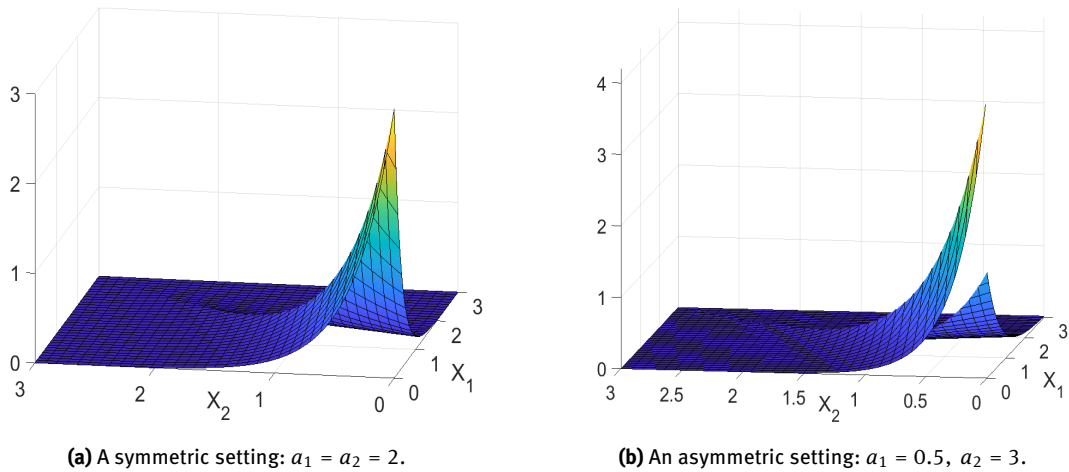


Figure 1: Joint density of (X_1, X_2) for two different parameter settings.

We remark that the parameter constellation $a_1 = \infty, a_2 = \infty$ corresponds to the special case of the Marshall-Olkin model, when $\lambda_A = 0, \lambda_B = 0, \lambda_{\{A,B\}} > 0$, i.e., the system of two entities can face only a common shock. (No separate individual shocks are present.)

Inverse marginal cumulative distribution functions.

Notice that the univariate quantile functions (i.e., the inverse functions of the marginal distribution functions (8), (9)) are smooth, but they cannot be written in a closed, analytical form (except in some very special cases). Since in this current work the quantile functions are mainly used in connection with the copula function, we will further elaborate this question in Subsection 2.3.

2.3 Copula function and copula density

We have discussed the features of the lifetimes variables X_1, X_2 , but in the end we are mainly interested in their copula. While the lifetimes are given by the pair (X_1, X_2) , their copula is defined on the pair of uniform marginals (U, V) , where $U = F(X_1)$, $V = G(X_2)$ and F, G are the marginal cdfs (8) and (9).

The copula function (15) and copula density (16) in our model – in accordance with the standard literature – are defined as follows.

$$C(u, v) = H(F^{-1}(u), G^{-1}(v)) \quad \text{for } 0 \leq u, v \leq 1, \tag{15}$$

where $F^{-1}(u)$ and $G^{-1}(v)$ are the generalized inverse functions of the cumulative distribution functions (8) and (9), namely they are the true inverse functions for $0 \leq u, v < 1$, and $F^{-1}(1) = \infty, G^{-1}(1) = \infty$. The copula density is given by

$$c(u, v) = \frac{\partial^2 C(u, v)}{\partial u \partial v} \quad \text{for } (u, v) \in [0, 1]^2 \setminus \{(1, 1)\}. \tag{16}$$

Notice that the formula (16), strictly speaking, cannot be extended to the entire $[0, 1]^2$, since $\lim_{(u,v) \rightarrow (1,1)} c(u, v) = \infty$, i.e., the copula density is unbounded around $(1, 1)$. To see the unboundedness, we provide a sketch of the argument. The details are left to the reader.

Notice first that $c(u, v) = \frac{\partial^2 C(u, v)}{\partial u \partial v} = h(F^{-1}(u), G^{-1}(v)) \cdot \frac{\partial F^{-1}(u)}{\partial u} \cdot \frac{\partial G^{-1}(v)}{\partial v}$. The formulas (8) and (9) show that F and G are, roughly speaking, of type $F(x) \approx 1 - \alpha \cdot e^{-\beta \cdot x}$ expressions, where $\alpha > 0, \beta > 0$. Hence $\frac{F^{-1}(u)}{\partial u} \approx \frac{1}{\beta \cdot (1 - u)}$.

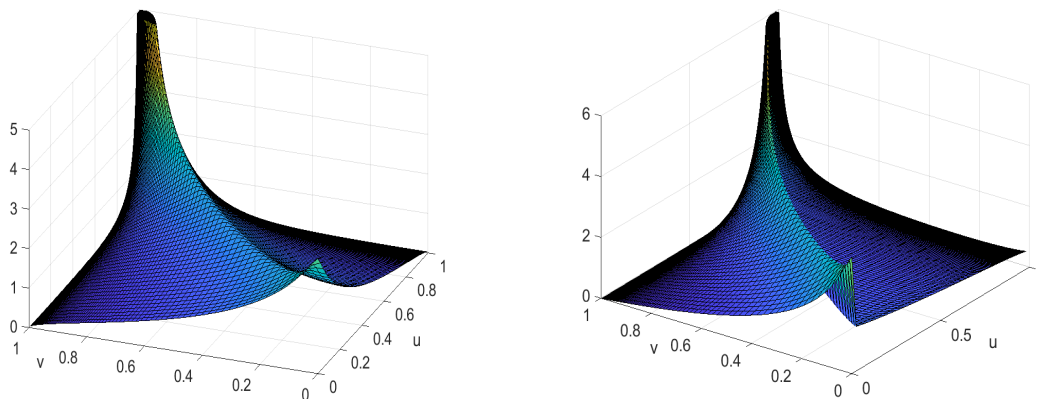
Similarly, one can argue that

$$h(F^{-1}(u), G^{-1}(v)) \approx \begin{cases} c_1 \cdot (1 - u)^{\kappa_1} \cdot (1 - v)^{\kappa_2} & \text{if } 0 \leq F^{-1}(u) \leq G^{-1}(v), \\ c_2 \cdot (1 - v)^{\kappa_3} \cdot (1 - u)^{\kappa_4} & \text{if } 0 \leq G^{-1}(v) \leq F^{-1}(u). \end{cases}$$

where $\kappa_1 < 1, \kappa_2 \geq 1, \kappa_3 < 1, \kappa_4 \geq 1$.

Altogether, the above considerations and formulas mean that $c(u, v)$ is unbounded when *both* $u \rightarrow 1$ and $v \rightarrow 1$, and in all other cases it is bounded.

If one wishes to compute explicitly the copula function (15), then an explicit formula for the inverse marginal cumulative distribution functions $F^{-1}(u), G^{-1}(v)$ would be needed as well, but in our model this is impossible in most cases (look at (8) and (9)). Therefore we will use numerical methods, as the reader will see in the following. In Figure 2, the copula density is shown for a symmetric and for an asymmetric case.



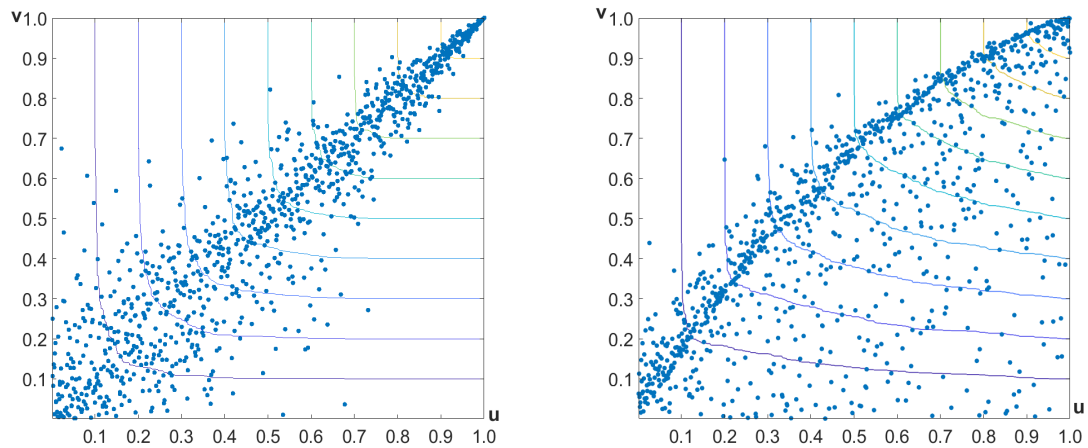
(a) A symmetric setting: $a_1 = a_2 = 2$.

(b) An asymmetric setting: $a_1 = 0.5, a_2 = 3$.

Figure 2: Copula density of (X_1, X_2) for two different parameter settings.

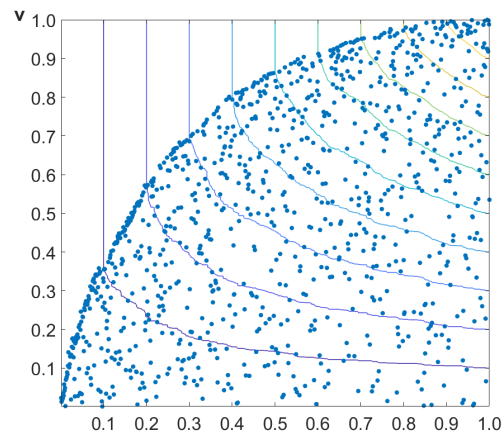
2.4 Scatter plots of copulas for different parameter settings

A great advantage of our model is the flexibility that it can easily handle asymmetric situations, too, i.e., when the effect of the default of an institution on another institution is larger than vice versa. Figure 3 gives an insight into the dependence structure by scatter plots, which are common tools for visualizing bivariate (or even three-variate) copulas.



(a) Symmetric copula with high shock parameters $a_1 = a_2 = 10$.
($\lambda_1 = \lambda_2 = 1$)

(b) Asymmetric copula with moderate and high shock parameters $a_1 = 1, a_2 = 20$.
($\lambda_1 = \lambda_2 = 1$)



(c) Asymmetric copula with moderate and very high shock parameters $a_1 = 1, a_2 = 20$.
($\lambda_1 = 0.2, \lambda_2 = 1$)

Figure 3: Scatter plots of copulas for different parameter settings based on samples of size 1000. The contour lines of the empirical copula functions for the values $0.1, \dots, 0.9$ are also shown.

In Figure 3b, we recognize a narrow region where many observations accumulate. We can call this a "line mass". Such a region is also present on Figure 3a and 3c, but the most visible on 3b. Notice that this line mass corresponds to the ridge which can be seen at the copula density plots (Figure (2a) and (2b)). Two questions arise: what is the interpretation of the line mass and which curve describes this line?

Looking at our particular parameter values in Figure 3b, $a_2 = 20$ is much larger than $a_1 = 1$, which causes

that in the case when Y_1 expires earlier than Y_2 , the (new) remaining lifetime of the second entity will be typically very short, since it follows an exponential distribution with parameter $\lambda_2 + a_2 = 21$. Loosely speaking, it results in $X_1 \approx X_2$ (or if one wishes $X_1 \lesssim X_2$), so this is the interpretation of the line mass. The corresponding probability, i.e., the weight of the line mass, is $\mathbb{P}(X_1 \approx X_2) \approx \mathbb{P}(Y_1 < Y_2) = \frac{\lambda_1}{\lambda_1 + \lambda_2} = \frac{1}{2}$.

The theoretical equation of the curve of the line mass (see also Figure 4) is obtained by setting $a_2 = \infty$, and then we can use the exact equality $X_1 = X_2$ (which event has probability $\frac{1}{2}$ due to the above mentioned fact), and then we obtain $v(u) = G(F^{-1}(u))$, where u and v are the variables of the copula function (see also (15)).

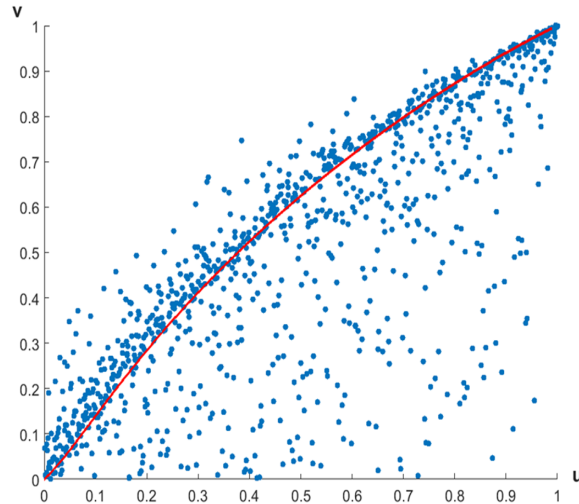


Figure 4: Scatter plot of the copula of (X_1, X_2) with the limiting curve $v(u) = G(F^{-1}(u))$ of the "line mass". This curve corresponds to the parameter setting $\lambda_1 = \lambda_2 = 1, a_1 = 2, a_2 = \infty$; The scatter plot shows the empirical copula of a sample of size 1000 from the model $\lambda_1 = \lambda_2 = 1, a_1 = 2, a_2 = 20$.

It may be shown that this copula family does not exhibit the Archimedean property (for the definition of Archimedeanity look at for instance Nelsen [9]), since the associativity does not hold. The details of this analysis are omitted.

2.5 The idea of a multivariate setting

Let Y_1, \dots, Y_n be independent exponential random variables with $Y_k \sim \text{Exp}(\lambda_k), k = 1, \dots, n$. We construct the actual lifetime variables $X_k^{(q)}$ ($k = 1, \dots, n$) for the q -th phase of an m -step cascading effect ($q = 1, \dots, m; m < n$) via the following mechanism. Note that defining an m -step cascade in our model means to define an ordered m -tuple of indices (k_1, \dots, k_m) which indicates the defaulting institution in each step.

For the first step (first phase) of the cascade let

$$Y_{k_1} = \min_{1 \leq k \leq n} Y_k \tag{17}$$

i.e., the institution that defaults first in a certain realization is denoted by k_1 . Let us introduce the variables $Z_{k,k_1} \sim \text{Exp}(\lambda_k + a_{k,k_1})$ with parameters $a_{k,k_1} \in [0, \infty)$ for $k_1 \neq k$, and with $a_{k,k} = \infty$. (In this latter case the corresponding random variable is degenerated, namely $Z_{k,k} = 0$ with probability 1.) The variables Z_{k,k_1} are independent of each other and of all Y_k s. We define the modified lifetime variables $X_k^{(1)}$ ($k = 1, \dots, n$) via

$$X_k^{(1)} := Y_{k_1} + Z_{k,k_1} \quad \text{for } k = 1, \dots, n. \tag{18}$$

The random variable Z_{k,k_1} , more precisely the shock parameter a_{k,k_1} , expresses the effect of the default of institution k_1 on institution k .

We introduce the notation I_1 , the index set of defaulted institutions after one step of the cascade. With this notation $I_1 = \{k_1\}$.

As already mentioned, the parameters can be organized as a $n \times n$ matrix. We always assume in this paper that $Z_{k,k} = 0$ with probability 1 for all $k = 1, \dots, n$, which corresponds to $a_{k,k} = \infty$, i.e., there is no possibility for governmental or other kind of bailout, when an institution has already defaulted. Note that the random variables $X_k^{(1)}$ are not exponentially distributed anymore (except when $a_{k,l} = 0$, $l = 1, \dots, n$, $l \neq k$ for some k), and also no longer independent (unless all $a_{k,l} = 0$ for $k \neq l$).

After the first step of the cascading effect described in (17) and (18), the institutions continue operating until the next default happens. Let

$$X_{k_2}^{(1)} = \min_{1 \leq k \leq n, k \neq k_1} X_k^{(1)} \quad (19)$$

i.e., the institution that defaults in the second step of the cascading effect is denoted by k_2 . Then

$$X_k^{(2)} := X_k^{(1)} + Z_{k,I_2} \quad \text{for } k = 1, \dots, n, \quad (20)$$

where $I_2 = \{k_1, k_2\}$ is the set of defaulted institutions after two steps of the cascade, $Z_{k,I_2} = Z_{k,k_1,k_2} \sim \text{Exp}(\lambda_k + a_{k,k_1} + a_{k,k_2})$, and the variables Z_{k,k_1,k_2} are independent of each other and also independent of any other variables. The random variable Z_{k,k_1,k_2} , more precisely the parameter $a_{k,k_1} + a_{k,k_2}$, expresses the effect of the defaults of institutions k_1 and k_2 on institution k . We also assume here (like in the first step) that $Z_{k,I_2} = 0$ with probability 1, when $k \in I_2$.

Note that the setting in (20) does not distinguish the order of defaults regarding institutions k_1 and k_2 . Furthermore, by the definition of Z_{k,I_2} , we impose a simple and well-tractable additivity for modelling the effect of consecutive defaults.

Finally, to put it more generally, in the q -th step of the cascading effect ($q = 1, \dots, m$), let

$$X_{k_q}^{(q-1)} = \min_{\substack{1 \leq k \leq n, \\ k \notin I_{q-1}}} X_k^{(q-1)}, \quad (21)$$

where $I_{q-1} = \{k_1, \dots, k_{q-1}\}$ is the index set of the already defaulted institutions. So we call (label) the institution which defaults in the q -th phase by k_q . ($X_k^{(0)} = Y_k$ for all $k = 1, \dots, n$ and $I_0 = \emptyset$.) Then

$$X_k^{(q)} := X_k^{(q-1)} + Z_{k,I_q} \quad \text{for } k = 1, \dots, n, \quad (22)$$

where $Z_{k,I_q} \sim \text{EXP}(\lambda_k + \sum_{p \in I_q} a_{k,p})$, and the variables Z_{k,I_q} are independent of each other and also independent of any other variables. The random variable Z_{k,I_q} , more precisely the parameter $a_{k,k_1} + \dots + a_{k,k_q}$, expresses the effect of the defaults of institutions k_1, \dots, k_q on institution k . We also assume here that if $k \in I_q$, then $Z_{k,I_q} = 0$ with probability 1. (Similarly as we have stressed it after step (20), the order within the index set I_q in (22) does not play any role.)

We also emphasize that the index sets I_q ($q = 1, \dots, m$) are random in the sense that they depend on the particular realizations of the random variables $X_k^{(q-1)}$ ($k = 1, \dots, n$).

3 Examining the change in the dependency structure in a symmetric case

In this section, we consider the monotonicity properties of the copula for the symmetric model $[\lambda, \lambda, a, a]$. Notice that the copula is invariant with respect to scaling of the time axis, implying that the copulas of the models $[\lambda, \lambda, a, a]$ and $[1, 1, a/\lambda, a/\lambda]$ are identical.

Without loss of generality, we concentrate therefore on the model $[1, 1, a, a]$.

The dependency structure between the lifetime variables X_1, X_2 is given by the joint cumulative distribution function (3), or alternatively (but not equivalently) by the copula function (15). Both of them depend on four parameters $(\lambda_1, \lambda_2, a_1, a_2)$. Our ultimate goal is to study how the joint cdf resp. the copula changes, when we let these parameters vary.

We consider the joint cdf (23), the marginal cdfs (24), (25), the copula function (26), the joint density (27), and the marginal densities (28), (29) in this special case. These formulas are gained obviously by specializing the general formulas (10), (11), (5), (8), (9), (3) and (15).

The one-parametric setting is also reflected in the notation. First we list the functions for $a \neq 1$ in formulas (23)–(29), then for $a = 1$ in formulas (30)–(35).

$$H_a(x, y) = \begin{cases} 1 + \frac{1}{1-a} \cdot e^{-(1-a)x} \cdot e^{-(1+a)y} + \frac{a}{1-a} \cdot e^{-2x} - \frac{1}{1-a} \cdot e^{-(1+a)x} - \frac{1}{1-a} \cdot e^{-(1+a)y} & \text{if } 0 \leq x \leq y, \\ 1 + \frac{1}{1-a} \cdot e^{-(1-a)y} \cdot e^{-(1+a)x} + \frac{a}{1-a} \cdot e^{-2y} - \frac{1}{1-a} \cdot e^{-(1+a)y} - \frac{1}{1-a} \cdot e^{-(1+a)x} & \text{if } 0 \leq y \leq x. \end{cases} \quad (23)$$

$$F_a(x) = 1 - \frac{1}{1-a} \cdot e^{-(1+a)x} + \frac{a}{1-a} \cdot e^{-2x}. \quad (24)$$

$$G_a(y) = 1 - \frac{1}{1-a} \cdot e^{-(1+a)y} + \frac{a}{1-a} \cdot e^{-2y}. \quad (25)$$

$$C_a(u, v) = H_a(F_a^{-1}(u), G_a^{-1}(v)) \quad \text{for } 0 \leq u, v \leq 1. \quad (26)$$

$$h_a(x, y) = (a + 1) \cdot e^{-(x+y)-a \cdot |x-y|} \quad x \geq 0, y \geq 0. \quad (27)$$

$$f_a(x) = -\frac{2a}{1-a} \cdot e^{-2x} + \frac{1+a}{1-a} \cdot e^{-(1+a)x} \quad x \geq 0. \quad (28)$$

$$g_a(y) = -\frac{2a}{1-a} \cdot e^{-2y} + \frac{1+a}{1-a} \cdot e^{-(1+a)y} \quad y \geq 0. \quad (29)$$

$$H_1(x, y) = \begin{cases} 1 - x \cdot e^{-2y} - (x + 1) \cdot e^{-2x} & \text{if } 0 \leq x \leq y, \\ 1 - y \cdot e^{-2x} - (y + 1) \cdot e^{-2y} & \text{if } 0 \leq y \leq x. \end{cases} \quad (30)$$

$$F_1(x) = 1 - (x + 1) \cdot e^{-2x} \quad \text{if } x \geq 0. \quad (31)$$

$$G_1(y) = 1 - (y + 1) \cdot e^{-2y} \quad \text{if } y \geq 0. \quad (32)$$

$$h_1(x, y) = 2 \cdot e^{-2 \cdot \max\{x, y\}} \quad x \geq 0, y \geq 0. \quad (33)$$

$$f_1(x) = (2x + 1) \cdot e^{-2x} \quad x \geq 0. \quad (34)$$

$$g_1(y) = (2y + 1) \cdot e^{-2y} \quad y \geq 0. \quad (35)$$

We will examine the change in the dependency structure given by (23)–(29) in two different ways. First we will consider some indicators (like expectation, variance, correlation coefficients of several kinds, etc.) extracted from the bivariate distribution. We will present these in Subsection 3.1. Secondly, we attempt to catch the dependency structure as a whole, and in Subsection 3.2 we will prove monotonicity result in the upper orthant order as parameter a varies.

3.1 Dependence measures

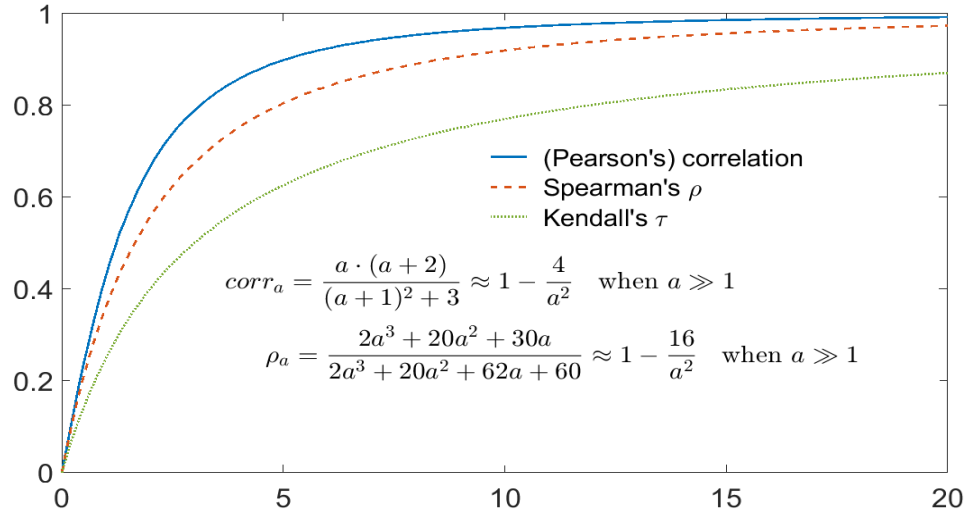


Figure 5: Three usual correlation coefficients for (X_1, X_2) .

In this subsection we will examine the most common correlation coefficients studied in the literature, namely the usual product moment correlation (also known as Pearson’s correlation coefficient), Spearman’s ρ and Kendall’s τ . As Figure 5 shows, each of them is increasing function of the model parameter a . We provide analytic formulas for the (Pearson’s) correlation and for Spearman’s ρ (see also Figure 5), from which the increasing property can be clearly verified. It seems impossible for us to derive an analytic formula for Kendall’s τ (we will explain the reason for that in the corresponding paragraph). However, by sampling from our model and by numerically evaluating Kendall’s τ for the samples, we obtained a curve for it. Furthermore, the increasing property of Kendall’s τ will be proved in Proposition 2.

Expectation, variance, covariance and correlation.

Since the marginal densities (28), (29) and the joint density (27) are simply sums of exponential functions, we get by elementary calculus that

$$\mathbb{E}_a(X_1) = \mathbb{E}_a(X_2) = \frac{1}{2} \cdot \frac{a + 2}{a + 1} \quad \text{and} \quad \mathbb{E}_a(X_1 \cdot X_2) = \frac{1}{2} + \frac{1}{2} \cdot \frac{1}{a + 1},$$

consequently $\text{cov}_a(X_1, X_2) = \frac{a \cdot (a + 2)}{4 \cdot (a + 1)^2}$.

The variance of X_1 (and of X_2) is

$$\mathbb{D}_a^2(X_1) = \frac{1}{4} \cdot \frac{(a + 1)^2 + 3}{(a + 1)^2}, \quad \text{therefore} \quad \text{corr}_a(X_1, X_2) = \frac{a(a + 2)}{(a + 1)^2 + 3}.$$

The previous formulas show that for $a = 0$ the covariance and the correlation of X_1 and X_2 is 0. It is also obvious from the more general fact that they are independent, which can be seen by substituting $a_1 = a_2 = 0$ in the general formula (5) of the joint density function.

The formula $\mathbb{E}_a(X_1) = \mathbb{E}_a(X_2) = \frac{1}{2} \cdot \frac{a+2}{a+1}$ has a nice interpretation as $a \rightarrow \infty$. In this case, the realizations of the two lifetime variables differ less and less from each other, and their marginal distributions can be approximated better and better with $\min\{Y_1, Y_2\}$, which is distributed according to $\text{Exp}(\lambda_1 + \lambda_2)$, i.e., in our case $\text{Exp}(2)$.

We can also see that $\text{cov}_a(X_1, X_2) \rightarrow \frac{1}{4}$ as $a \rightarrow \infty$. It is more informative to examine the limit of the correlation: $\text{corr}_a(X_1, X_2) \rightarrow 1$ as $a \rightarrow \infty$.

Spearman’s ρ .

For (bivariate) samples Spearman’s ρ is defined through the order statistics, namely the correlation of the ranked data. Accordingly, for distributions we need to compute the following (in notation we immediately use our variables):

$$\rho_a = \frac{\text{cov}(U, V)}{\mathbb{D}(U) \cdot \mathbb{D}(V)} = 12 \cdot (\mathbb{E}_a(U \cdot V) - \mathbb{E}_a(U) \cdot \mathbb{E}_a(V)) = 12 \cdot \mathbb{E}_a(U \cdot V) - 3.$$

Using the fact that $U = F(X)$, $V = G(Y)$, and formulas (24), (25) and (27) consist of (sums of) exponential functions, through a cumbersome, but elementary computation we get that

$$\mathbb{E}_a(U \cdot V) = \int_0^\infty \int_0^\infty F_a(x) \cdot G_a(y) \cdot h_a(x, y) \, dx \, dy = \frac{1}{3} \cdot \frac{2a^3 + 20a^2 + 54a + 45}{2a^3 + 20a^2 + 62a + 60}.$$

Finally we get $\rho_a = 4 \cdot \frac{2a^3 + 20a^2 + 54a + 45}{2a^3 + 20a^2 + 62a + 60} - 3 = \frac{2a^3 + 20a^2 + 30a}{2a^3 + 20a^2 + 62a + 60}$, which is pictured in Figure 5.

Kendall’s τ .

Recall that for a bivariate general copula C , Kendall’s τ is defined as

$$\tau = 4 \cdot \int_0^1 \int_0^1 C(u, v) \, dC(u, v) - 1. \text{ If the copula is the empirical one based on a sample}$$

$(X_1(i), X_2(i))_{i=1}^N$, then Kendall’s τ can be also defined as

$$\tau = \frac{\# \text{ concordant pairs} - \# \text{ discordant pairs}}{\binom{N}{2}}.$$

(A pair $(X_1(i), X_2(i))$ is called concordant with another pair $(X_1(j), X_2(j))$, if $\text{sgn}(X_1(i) - X_1(j)) = \text{sgn}(X_2(i) - X_2(j))$, otherwise they are discordant.)

Now we are ready to present the increasing property of Kendall’s τ in Proposition 2. Then the reader finds the main result of this paper in Proposition 3, namely the upper orthant ordering concerning the copulas C_a . We notice that by Theorem 5.1.9 in Nelsen [9], Proposition 2 is a direct consequence of Proposition 3. We still present them in this order, because the proof of Proposition 2 only focuses on (concordant) pairs in a sample, while the proof of Proposition 3 deals with the entire order statistics, and in this way it can be considered as an extension of the proof of Proposition 2.

Proposition 2.

Let τ_a be Kendall’s τ pertaining to the copula τ model $[1, 1, a, a]$. Then $a \mapsto \tau_a$ is monotonically non-decreasing.

Proof.

Consider an i.i.d. sample $S_N^a = \{(X_1^a(i), X_2^a(i))\}_{i=1}^N$ according to the model $[1, 1, a, a]$. Such a sample can be generated by the following algorithm: Let $U(i)$ and $V(i)$ be the realizations of independent Uniform[0,1] variables for $i = 1, \dots, N$. Then

- with probability 1/2

$$X_1^a(i) = -\frac{1}{2} \log(U(i)); \quad X_2^a(i) = -\frac{1}{2} \log(U(i)) - \frac{1}{1+a} \log(V(i)),$$

- with probability 1/2

$$X_2^a(i) = -\frac{1}{2} \log(U(i)); \quad X_1^a(i) = -\frac{1}{2} \log(U(i)) - \frac{1}{1+a} \log(V(i)).$$

For a given sample S_N^a , let $\mathcal{U} = \{i : X_1^a(i) < X_2^a(i)\}$ and $\mathcal{L} = \{i : X_1^a(i) > X_2^a(i)\}$. Notice that the event $X_1^a(i) = X_2^a(i)$ has probability zero. Notice now that if $Z \sim \text{Exp}(1+a)$, then $\frac{1+a}{1+a'} \cdot Z \sim \text{Exp}(1+a')$. Therefore we can easily modify the sample S_N^a to get a valid sample $S_N^{a'} = \{(X_1^{a'}(i), X_2^{a'}(i))\}_{i=1}^N$ for model $[1, 1, a', a']$. To this end, let

- for $i \in \mathcal{L}$

$$X_1^{a'}(i) = X_2^a(i) + (X_1^a(i) - X_2^a(i)) \cdot \frac{1+a}{1+a'}; \quad X_2^{a'} = X_2^a, \quad (36)$$

- for $i \in \mathcal{U}$

$$X_2^{a'}(i) = X_1^a(i) + (X_2^a(i) - X_1^a(i)) \cdot \frac{1+a}{1+a'}; \quad X_1^{a'} = X_1^a. \quad (37)$$

We claim that the number of concordant pairs in sample $S_N^{a'}$ is not less than the number of concordant pairs in sample S_N^a . To prove this assertion, let (i, j) be a concordant pair in S_N^a , i.e., $(X_1^a(i) - X_1^a(j)) \cdot (X_2^a(i) - X_2^a(j)) > 0$. We now have to distinguish four cases: (a), (b), (c) and (d).

- (a) $X_1^a(i) > X_2^a(i), X_1^a(j) > X_2^a(j)$,
- (b) $X_1^a(i) < X_2^a(i), X_1^a(j) < X_2^a(j)$,
- (c) $X_1^a(i) > X_2^a(i), X_1^a(j) < X_2^a(j)$,
- (d) $X_1^a(i) < X_2^a(i), X_1^a(j) > X_2^a(j)$.

We have illustrated the situation in Figure 6, where the original pairs $(X_1^a(i), X_2^a(i))$ and $(X_1^a(j), X_2^a(j))$ are shown as little circles \circ and the modified pairs $(X_1^{a'}(i), X_2^{a'}(i))$ and $(X_1^{a'}(j), X_2^{a'}(j))$ are shown as dots \bullet .

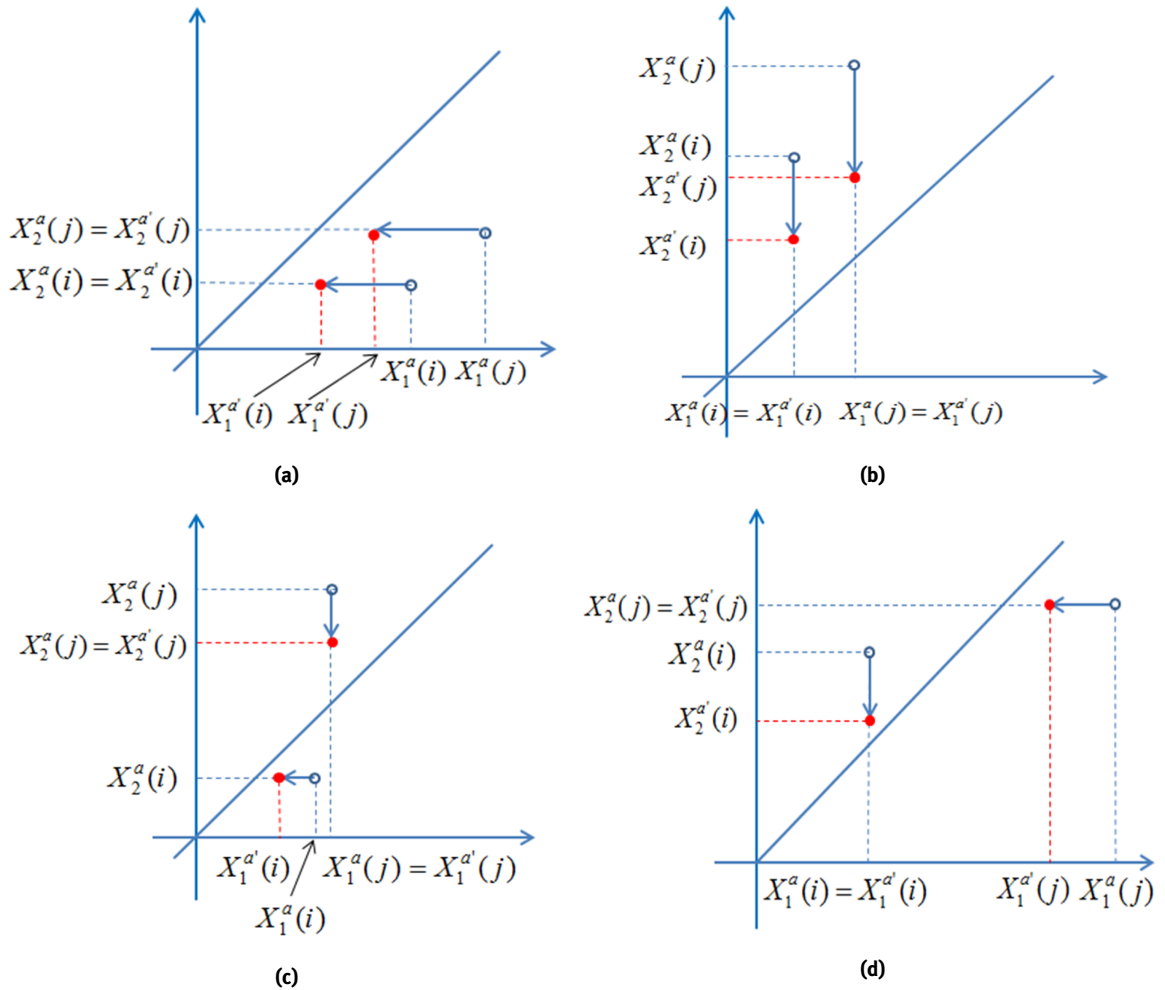


Figure 6: The four possible positions of a concordant pair (i, j) (marked by \circ). The transformed sample (marked by \bullet). Concordant pairs remain concordant.

W.l.o.g. we may assume that $X_1^a(j) - X_1^a(i) > 0$ and $X_2^a(j) - X_2^a(i) > 0$. Let us look e.g., at case (a). Here

$$X_1^{a'}(j) - X_1^{a'}(i) = \frac{a' - a}{1 + a'} \cdot (X_2^a(j) - X_2^a(i)) + \frac{1 + a}{1 + a'} \cdot (X_1^a(j) - X_1^a(i)) > 0, \text{ and}$$

$$X_2^{a'}(j) = X_2^a(j), \quad X_2^{a'}(i) = X_2^a(i), \text{ so } X_2^{a'}(j) - X_2^{a'}(i) = X_2^a(j) - X_2^a(i) > 0.$$

In case (b) only the roles of the coordinates are exchanged.

In case (c) we have

$$X_1^{a'}(i) < X_1^a(i) < X_1^a(j) = X_1^{a'}(j)$$

and

$$X_2^{a'}(i) = X_2^a(i) < X_1^a(i) < X_1^a(j) = X_1^{a'}(j) < X_2^{a'}(j).$$

The last inequality holds, because the point $(X_1^{a'}(j), X_2^{a'}(j))$ lies above the diagonal line $y = x$. Again, interchanging the roles of the coordinates shows also the validity of the statement in case (d).

Finally, we argue that the empirical copula converges a.s. to the true one (see e.g., Gaenssler and Stute [5]) and that the empirical τ converges to the true τ . Thus we obtain the statement. \square

3.2 Monotonicity of the copula in upper orthant order w.r.t. parameter a

Our general purpose is to examine how the copula (26) (and slightly more general the copula (15) for the model $[\lambda, \lambda, a, a]$) changes as we change the value of parameter a , i.e., to determine and describe (some properties of) the function $a \mapsto C_a$.

Upper orthant order for copulas.

Definition 1. Let C_1 and C_2 be two bivariate copulas and let (U_1, V_1) be distributed according to C_1 and (U_2, V_2) be distributed according to C_2 . We say that C_1 is dominated by C_2 in upper orthant order (in symbol $C_1 \preceq_{UO} C_2$), if

$$\mathbb{P}(U_1 > u, V_1 > v) \leq \mathbb{P}(U_2 > u, V_2 > v)$$

for all $u, v \in [0, 1]$. In other words one may say that C_2 is more (co-monotone) dependent than C_1 .

Remark 1. Since $\mathbb{P}(U \leq u, V \leq v) = 1 - u - v + \mathbb{P}(U > u, V > v)$, one sees that that $C_1 \preceq_{UO} C_2$ is equivalent to

$$C_1(u, v) \leq C_2(u, v) \tag{38}$$

for all $u, v \in [0, 1]$.

Some authors say that the random vector (U_2, V_2) is smaller than the random vector (U_1, V_1) in lower orthant order, if (38) holds (see e.g., Denuit et al. [3], Definition 3.3.80). Nevertheless, the notion of lower orthant order is not needed in our work.

We now formulate the main result of this section.

Proposition 3. Let $C_a(u, v)$ be the copula of the model $[1, 1, a, a]$ and let $C_{a'}(u, v)$ the copula of the model $[1, 1, a', a']$, where $a \leq a'$. Then

$$C_a \preceq_{UO} C_{a'}. \tag{39}$$

Proof. We have to show that $C_a(u, v) \leq C_{a'}(u, v)$ for all u, v . We use the same construction for the samples S_N^a resp. $S_N^{a'}$ as in the proof of Proposition 2. (Look at (36) and (37).)

Notice that for all s, t

$$\#\{i : X_1^a(i) \leq s, X_2^a(i) \leq t\} \leq \#\{i : X_1^{a'}(i) \leq s, X_2^{a'}(i) \leq t\}.$$

We will show first that for the empirical copula $C_a^{(N)}$ of the sample $(X_1^a(i), X_2^a(i))$, $i = 1, \dots, N$ we have the upper orthant order

$$C_a^{(N)}(u, v) \leq C_{a'}^{(N)}(u, v) \tag{40}$$

for all $u, v \in [0, 1]$.

Let $X_1^a[i : N]$ and $X_2^a[j : N]$ be the order statistics and let $X_1^a(\ell_1^a) = X_1^a[i : N]$ and $X_2^a(\ell_2^a) = X_2^a[j : N]$. For $u = i/N, v = j/N$ we set

$$I^a = \{i : X_1^a(i) \leq X_1^a(\ell_1^a); X_2^a(i) \leq X_2^a(\ell_2^a)\}$$

and $m^a = \#(I^a)$. Notice that $C_a^{(N)}(u, v) = m^a/N$. We have to show that m^a is monotonically increasing with a .

We may assume w.l.o.g. that there are no ties in the sample. If the index sets I^a do not change for some $a' > a$, the empirical copula also does not change. Let now a' be such that exactly one index changed from I^a to $I^{a'}$, because one of the indices ℓ_1^a, ℓ_2^a changed. Suppose for instance that $X_1^{a'}(\ell_1^{a'})$ is no longer the i -th largest among the $X_1^{a'}(\cdot)$, but there is a $\ell' \notin I^a$ such that $X_1^{a'}(\ell') < X_1^{a'}(\ell_1^a)$. Then two situations may occur:

- If $X_2^{a'}(\ell') > X_2^{a'}(\ell_2^a)$, then $m^{a'} = m^a$.
- If $X_2^{a'}(\ell') \leq X_2^{a'}(\ell_2^a)$, then $m^{a'} = m^a + 1$.

In both cases is m^a non-decreasing. One may repeat the argument for two indices changing, three indices changing and so on to see that (40) is proved. Now, again invoking the argument that the empirical copulas converge to the true copulas as in Proposition 2, one sees that

$$C_a(u, v) \leq C_{a'}(u, v)$$

holds for all $u, v \in [0, 1]$. □

Corollary.

Recall that Blomqvist's β for the copula C is defined as $\beta = C(1/2, 1/2)$ (see Blomqvist [1]). As a consequence of Proposition 3 also this correlation coefficient is monotonic in a , i.e., the function

$$a \mapsto \beta_a = C_a(1/2, 1/2)$$

is monotonically increasing.

4 An application for measuring systemic risk

It is a usual approach in financial theory and practice that the strength and vulnerability of the institutions is quantified by indirect manners. This is because an actual default or bankruptcy of a financial institution occurs very rarely, however the stability of the institutions can vary significantly. In this way, the lifetime model presented in Sections 1, 2 and 3 can also be considered as an indirect tool to measure the stability and potential strength of entities in a financial system.

In this section we have two aims. First, we relate our lifetime variables to loss variables, since mostly these latter ones are in the focus of interest for financial institutions. Secondly, we also establish a relation between lifetime intensities and CDS spreads, which are widely used indicators for the financial strength of entities in banking systems. This enables us to provide numerical illustration for our model using real financial data.

We will formulate the definitions for arbitrary n , and in the numerical case study we will restrict our analysis to $n = 2$.

4.1 Relation between lifetime and loss variables; a model for systemic risk

In the previous sections, we dealt with defining and exploring a joint lifetime model, where we considered the random vector of lifetimes (X_1, \dots, X_n) . In this subsection, we translate the distribution of lifetimes into the distribution of losses, so we introduce the random vector of losses (L_1, \dots, L_n) . The basic idea is simple: the longer the lifetime is, the less the losses are. There are several ways how to formulate this. In the following we list a few of them.

A few possible definitions of loss variables defined via lifetime variables.

- (i) $L_i := \min\left(\frac{1}{X_i}, c\right)$, $(i = 1, \dots, n)$, where $c > 0$ is a constant.
- (ii) $L_i := c_i \cdot e^{-r \cdot X_i}$, $(i = 1, \dots, n)$, where c_i is the initial capital of institution i , and r is the risk-free interest rate.
- (iii) $L_i := \mathbb{1}_{\{X_i \leq t_i\}}$, $(i = 1, \dots, n)$, where $t_i > 0$ is a threshold.
- (iv) $L_i := \begin{cases} \xi_i & \text{if } X_i \leq t_i, \\ 0 & \text{if } X_i > t_i, \end{cases}$ $(i = 1, \dots, n)$, where $t_i > 0$ is a threshold, and ξ_i is a random variable with given distribution.

Quantifying systemic risk.

While the individual risk refers to the fact that random losses may occur to a financial institution, the notion of systemic risk measures the extra risk which can be attributed to the interdependence of several institutions. Let R be a risk functional, which assigns a real value to the risk of a potential loss variable L . If the system consists of n institutions, then $L_1 + \dots + L_n$ is the total loss of the whole system. The distribution of this sum depends on the marginal distributions and the copula. For fixed marginal distribution, let us write

$$\bigoplus_C L_i$$

for the total loss variable, when the individual losses are coupled by copula C .

Definition 2. (See also Pflug and Pichler [10]). For a given risk functional R , define the systemic risk by

$$\mathcal{R}(C, R; \mathbf{L}) = R\left(\bigoplus_C L_i\right) - R\left(\bigoplus_{\Pi} L_i\right), \tag{41}$$

where $\mathbf{L} = (L_1, \dots, L_n)$ is the vector of (univariate) marginal loss variables and Π is the independent copula. (41) compares the total loss under the copula C with the (hypothetical) total loss of independent institutions. In particular, one may consider C as the copula of the lifetimes (X_1, \dots, X_n) (e.g., a Freund copula) and losses depending on the lifetimes via the above formulas (i)–(iv).

In Definition 3 we recall the notion of average-value-at-risk, which can be found in several sources, e.g., in Pflug and Römisch [11]. We would like to stress that technically there are two variants of the $AV@R$, the lower- $AV@R$, which focuses on the left tail of the distribution, and the upper- $AV@R$, which focuses on the right tail. We will need this latter one.

Definition 3. Let X be a random variable with cdf F , and let $0 \leq \alpha < 1$. The (upper) average-value-at-risk at level α is defined as

$$AV@R_\alpha(X) = \frac{1}{1 - \alpha} \int_\alpha^1 F^{-1}(u) du, \tag{42}$$

where $F^{-1}(u)$ is the generalized inverse function of F .

Remark 2. When it is clear from the context which significance level α is meant, or a certain statement holds for all α , then the lower index can be omitted from the notation, and we simply write $AV@R(X)$.

Remark 3. It was shown in Pflug and Pichler [10] that $C_1 \preceq_{VO} C_2$ implies that

$$AV@R\left(\bigoplus_{C_1} L_i\right) \leq AV@R\left(\bigoplus_{C_2} L_i\right).$$

In particular, by Proposition 3, for our bivariate copula C_a (26) we have that

$$AV@R(L_1 \oplus_{C_a} L_2) \leq AV@R(L_1 \oplus_{C_{a'}} L_2)$$

for $a \leq a'$.

Notice also that if $a_1 = a_2 = 0$ in our model, then $\mathcal{R} = 0$, so the system does not possess any systemic risk. It does not mean that the overall risk in the system would be zero, but it means that the part of the risk, which is attributed only to the dependencies of the institutions, is zero.

Examples.

We illustrate our systemic risk definition through some examples. For the sake of simplicity we consider in each of these examples the model

$$[1, 1, a_1, a_2]. \tag{43}$$

Example 1. Let $(X_1, X_2) \sim H$, where H is according to (3), specified by (43).

To define the relation between lifetimes and losses we will use (i) from the list above. Let us assume that both institutions have capital $c = 10$ (units).

$$\text{So } L_1 = \min(1/X_1, 10), L_2 = \min(1/X_2, 10).$$

For the risk functional R let us choose $R(L) := \mathbb{E}(L - t \mid L > t)$, where L is a loss variable, and t is a threshold, whose excess is considered as a "bad" event. This risk functional is closely related to the *stop-loss* transform, which is a popular risk functional in insurance mathematics (look at for instance Denuit et al. [3], Definition 1.7.1.1). In accordance with our systemic risk definition $L := L_1 + L_2$, and let us set the threshold $t = 10$. It means that we consider a situation risky, when the market loses half of its capital or more. The following table shows for some values how the systemic risk $\mathcal{R}(C, R; \mathbf{L})$ (see Definition 2) increases as we increase the shock parameters a_1, a_2 .

(a_1, a_2)	(0,0)	(0,1)	(1,1)	(1,3)	(2,3)	(5,5)	(10,10)	(100,100)
$R(\oplus_C L_i)$	2.2810	2.7831	3.2644	3.8351	4.1676	5.0480	5.8924	7.2118
$\mathcal{R}(C, R; \mathbf{L})$	0	0.5021	0.9834	1.5541	1.8866	2.7670	3.6114	4.9308
rel.incr.	0%	22.01%	43.11%	68.13%	82.71%	121.31%	158.33%	216.17%

The fourth line of the table shows the relative increment in the systemic risk, compared to the independent case.

Example 2. Let $(X_1, X_2) \sim H$, where H is according to (3), specified by (43). Let us define now the loss variables according to (ii) from the above list, i.e., via $L_1 = c_1 \cdot e^{-r \cdot X_1}$ and $L_2 = c_2 \cdot e^{-r \cdot X_2}$, where $c_1 = c_2 := 1$, and the risk-free interest rate r is set up to $r = 0.05$.

The risk functional is defined as in Example 1, with threshold $t = 1$.

We can observe again the increase in systemic risk as we increase the shock parameters a_1, a_2 , but in a much more moderate way.

(a_1, a_2)	(0,0)	(0,1)	(1,1)	(1,3)	(2,3)	(5,5)	(10,10)	(100,100)
$R(\oplus_C L_i)$	1.9049	1.9163	1.9275	1.9334	1.9373	1.9432	1.9466	1.9508
$\mathcal{R}(C, R; \mathbf{L})$	0	0.0114	0.0226	0.0285	0.0324	0.0383	0.0417	0.0459
rel.incr.	0%	0.59%	1.19%	1.49%	1.70%	2.01%	2.19%	2.41%

Example 3. Let $(X_1, X_2) \sim H$, where H is according to (3), specified by (43). Let us define now the loss variables according to (iii) from the above list, i.e., via $L_i = \mathbb{1}_{\{X_i \leq t_i\}}$, where $t_i := Q_{0.8}^{(X_i)}$ (0.8-quantile of X_i) for $i = 1, 2$.

For the risk functional let us choose $R(L) = AV@R(L)$, i.e., in the light of Definition 2 $R(L_1 + L_2) = AV@R_{0.8}(L_1 + L_2)$.

(a_1, a_2)	(0,0)	(0,1)	(1,1)	(2,3)	(5,5)	(10,10)
$R(\oplus_C L_i)$	1.200	1.2895	1.4665	1.6305	1.8020	1.8950
$\mathcal{R}(C, R; \mathbf{L})$	0	0.0895	0.2665	0.4305	0.6020	0.6950
rel.incr.	0%	7.46%	22.21%	35.88%	50.67%	57.92%

Example 4. Let $(X_1, X_2) \sim H$, where H is according to (3), specified by (43). Let us define now the loss variables according to (iv) from the above list, i.e., via

$$L_i = \begin{cases} \xi_i & \text{if } X_i \leq t \\ 0 & \text{if } X_i > t \end{cases} \quad i = 1, 2, \quad \text{where } \xi_i \sim \text{Exp}(1) \text{ independent of } X_1 \text{ and } X_2.$$

The threshold is set up to $t = 1.3863$, that is the 75%-quantile of the $\text{Exp}(1)$ distribution, i.e., the distribution of X_1 and of X_2 , when $a_1 = a_2 = 0$.

The risk functional is defined again as in Example 1.

A survey of the numerical study can be seen in the following table.

(a_1, a_2)	(0,0)	(0,1)	(1,1)	(1,3)	(2,3)	(5,5)	(10,10)	(100,100)
$R(\oplus_C L_i)$	1.4996	1.6020	1.7024	1.7584	1.7954	1.8457	1.8599	1.8765
$\mathcal{R}(C, R; \mathbf{L})$	0	0.1024	0.2028	0.2588	0.2958	0.3461	0.3603	0.376
rel.incr.	0%	6.82%	13.52%	17.26%	19.73%	23.08%	24.03%	25.07%

4.2 Numerical study using CDS-data

We use CDS-spread data, provided by the Data Centre of the University of Vienna. The data is on a daily basis, and consists of $N = 1907$ observations of 77 banks from Europe and North-America on the time horizon Dec 2007 - October 2015. We will denote this data set by $\{s(i)\}_{i=1}^N$, while for the CDS-spread as a (financial) variable we use the notation s .

The lifetime model presented above enables us to perform a numerical case study for pairs of institutions, which are very reduced subsystems of the entire data set. In spite of this, we believe that the following numerical illustration is informative enough, and convinces the reader that our model works.

Let us consider two banks from the above mentioned data set, Erste Bank (AT) and Alpha Bank (GRE). These are known to be different in financial strength and stability, therefore they seem to be a reasonable choice for our example.

Relation between CDS-spreads, lifetime variables and loss variables.

It is clear from the previous considerations that CDS-spreads are directly observable and given quantities. In order to apply our model, from these spreads we have to produce lifetime variables, which are in this sense artefacts.

From CDS spreads to lifetimes. If X is the lifetime variable of the debtor of a CDS with spread s and maturity T , then the benefit is given by

$$s \int_0^{\min(X, T)} e^{-rt} dt = \frac{s}{r} (1 - e^{-r \min(X, T)})$$

and the costs are

$$LGD \cdot e^{-rX} \cdot \mathbb{1}_{\{X \leq T\}}.$$

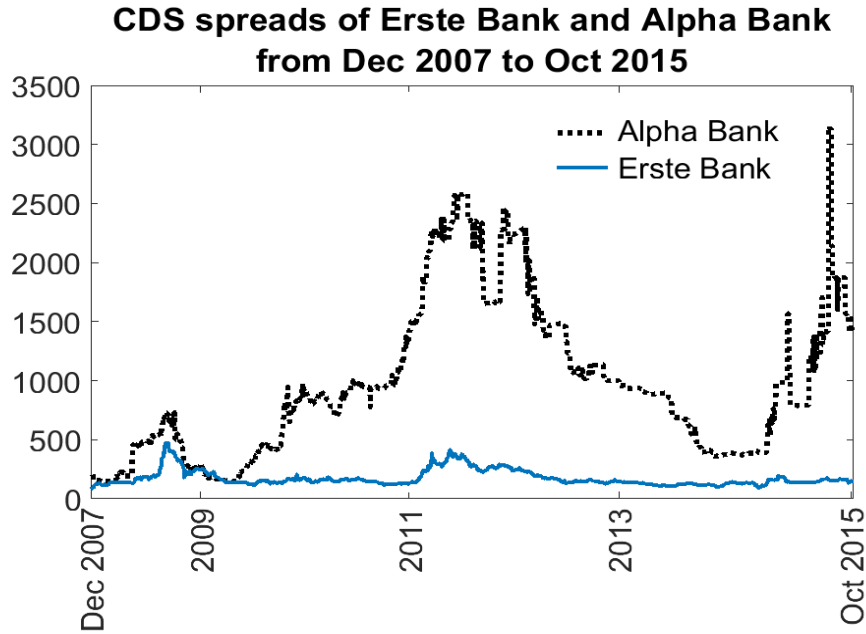


Figure 7: CDS spreads of Erste Bank and Alpha Bank given in basis points

Here LGD is the so-called (relative) Loss-Given-Default and r is the interest rate. (In financial applications LGD is a commonly used notion, which expresses the maximal proportion of the firm's capital, which can be lost in case of default.)

For simplicity, we assume that the credits are long-term and set $T = \infty$. Equating the costs and the benefits in this case one gets the relationship

$$X = -\frac{1}{r} \log \left(\frac{s}{LGD \cdot r + s} \right). \quad (44)$$

Now, in order to set up a data set for lifetime data, to each CDS-spread observation $\{s(i)\}_{i=1}^N$ of a certain institution we create an observation $\{X(i)\}_{i=1}^N$ using (44). (Note that the argument i refers to the ordinal number of the observation, as it was also the case in the proof of Proposition 2 and Proposition 3.)

$$X(i) = -\frac{1}{r} \cdot \log \left(\frac{s(i)}{LGD \cdot r + s(i)} \right), \quad (45)$$

where LGD is the Loss-Given-Default and r is the risk-free interest rate. In our analysis $LGD = 0.5$, $r = 0.05$.

In Subsection 4.1 we already discussed some alternatives how to link lifetime and loss variables. Consider now (ii) from the list in Subsection 4.1, i.e.,

$$L := e^{-r \cdot X}. \quad (46)$$

Combining (44) and (46) we obtain

$$L := \frac{s}{LGD \cdot r + s}, \quad (47)$$

with which we gained one possible way for having a direct connection between CDS-spreads and losses, i.e., between the available inputs and objects of interest. Notice that the function L given by (47) is a monotonically increasing function. This implies that the spreads and the losses have exactly the same copula. Notice, however, that the copula of the lifetimes is the survival copula of the spreads: $\tilde{C}(u, v) = 1 - u - v + C(u, v)$, but monotonicity w.r.t. parameter a is preserved between C and \tilde{C} .

Parameter estimation.

Our next task is to estimate the model parameters $\lambda_1, \lambda_2, a_1, a_2$, i.e., to find the copula in our model which fits the best to the data according to some criterion. We will perform this analysis for the pair Erste Bank ($k = 1$)

and Alpha Bank ($k = 2$). For illustration purpose we have selected a subset of our data set from 21.3.2008. to 19.1.2012, which consists of 1000 observations.

The realized lifetimes $\{(X_1(i), X_2(i))\}_{i=1}^{1000}$ are influenced by all the four parameters $\lambda_1, \lambda_2, a_1, a_2$.

We consider the optimization problem, which is in fact a Cramér-von Mises type minimum distance estimation:

$$\min_{\lambda_1, \lambda_2, a_1, a_2} \sum_{i=1}^{1000} (C_{\lambda_1, \lambda_2, a_1, a_2}(U(i), V(i)) - C_{\text{emp}}(U(i), V(i)))^2, \quad (48)$$

where

- $C_{\lambda_1, \lambda_2, a_1, a_2}$ is given by (15) via (6) and (7), (although in these formulas the parameters $\lambda_1, \lambda_2, a_1, a_2$ were not indicated directly in the notation);
- $U(i) = F_{\lambda_1, \lambda_2, a_1}(X_1(i))$, $V(i) = G_{\lambda_1, \lambda_2, a_2}(X_2(i))$, where $F_{\lambda_1, \lambda_2, a_1}$ and $G_{\lambda_1, \lambda_2, a_2}$ are given by (6) and (7) (although in these formulas the parameters $\lambda_1, \lambda_2, a_1, a_2$ were not indicated directly in the notation);
- C_{emp} is the empirical copula which corresponds to the sample $\{(X_1(i), X_2(i))\}_{i=1}^{1000}$.

As a result we get $\hat{\lambda}_1 = 0.00459$, $\hat{\lambda}_2 = 0.109307$, $\hat{a}_1 = 0.0836$, $\hat{a}_2 = 0.0699$, which can be interpreted as follows:

$\frac{1}{\hat{\lambda}_1} = 217.86$, $\frac{1}{\hat{\lambda}_2} = 9.1485$ indicate - roughly speaking - the expected lifetimes of the institutions. A measure of unit does not need to be attached, because the copula is invariant under simultaneous positive scaling of the model parameters, i.e., the models $[\lambda_1, \lambda_2, a_1, a_2]$ and $[\alpha \cdot \lambda_1, \alpha \cdot \lambda_2, \alpha \cdot a_1, \alpha \cdot a_2]$ have the same copula. In the light of this remark on *scale invariance*, we might also say that the expected lifetime of Erste Bank is 23.8 expected lifetime of Alpha Bank. Furthermore, the estimated value $\hat{a}_1 = 0.0836$ is the effect of the default of Alpha Bank on Erste Bank according to our analysis, while $\hat{a}_2 = 0.0699$ is the effect of the default of Erste Bank on Alpha Bank. Using again the comment on scale invariance, one might say that the expected remaining lifetime of Alpha Bank after Erste Bank defaults is $\frac{1}{\hat{\lambda}_2 + \hat{a}_2} = 5.58$, which is 0.61 of the expected lifetime of Alpha Bank without Erste defaulting. Similarly, the expected remaining lifetime of Erste Bank after Alpha Bank defaults is $\frac{1}{\hat{\lambda}_1 + \hat{a}_1} = 11.34$, which is 0.05 of the expected lifetime of Erste Bank without Alpha defaulting.

For properties of Cramér von-Mises type minimum distance estimates like consistency and asymptotic normality we refer to the work of Boos [2], Klugman and Parsa [6] or Koul [7].

5 Conclusion

As we have seen, our lifetime-based cascading model is able to catch the dynamics of dependence structures in financial systems, although does not contain time-dependency explicitly. However, in a latent way the model successfully replaces the explicit appearance of a time variable.

Further work will focus on the analysis and numerical illustration of the multivariate Freund copula, which may accommodate multi-step cascades.

References

- [1] Blomqvist, N. (1950). On a measure of dependence between two random variables. *Ann. Math. Statist.* 21(4), 593–600.
- [2] Boos, D. D. (1981). Minimum distance estimators for location and goodness of fit. *J. Amer. Statist. Assoc.* 76(375), 663–670.
- [3] Denuit, M., J. Dhaene, M. Goovaerts, and R. Kaas (2005). *Actuarial Theory for Dependent Risks: Measures, Orders and Models*. John Wiley & Sons, Chichester.
- [4] Freund, J. E. (1961). A bivariate extension of the exponential distribution. *J. Amer. Statist. Assoc.* 56(296), 971–977.
- [5] Gaenssler, P. and W. Stute (1987). *Seminar on Empirical Processes*. Birkhäuser, Basel.

- [6] Klugman, S. A. and A. R. Parsa (1993). Minimum distance estimation of loss distributions. *Proceedings of the Casualty Actuarial Society* 80(152,153), pp. 250–270.
- [7] Koul, H. L. (1992). *Weighted Empiricals and Linear Models*. Institute of Mathematical Statistics, Hayward CA.
- [8] Marshall, A. W. and I. Olkin (1967). A multivariate exponential distribution. *J. Amer. Statist. Assoc.* 62(317), 30–44.
- [9] Nelsen, R. B. (2006). *An Introduction to Copulas*. Second edition. Springer, New York.
- [10] Pflug, G. C. and A. Pichler (2018). Systemic risk and copula models. *Cent. Eur. J. Oper. Res.* 26(2), 465–483.
- [11] Pflug, G. C. and W. Römisch (2007). *Modeling, Measuring and Managing Risk*. World Scientific Publishing, Singapore.
- [12] Yu, F. (2007). Correlated defaults in intensity-based models. *Math. Finance* 17(2), 155–173.