Copula-Specific Credit Portfolio Modeling

How the Sector Copula Affects the Tail of the Portfolio Loss Distribution

Matthias Fischer and Kevin Jakob

Abstract Traditionally, banks estimate their economic capital which has to be reserved for unexpected credit losses with individual credit portfolio models. Many of those have its roots in the CreditRisk$^+$ or in the CreditMetrics framework, which were both launched in 1997. Motivated by the current regulatory requirements, banks are required to analyze how sensitive their models (and the resulting risk figures) are with respect to the underlying assumptions. Within this context, we concentrate on the dependence structure in terms of copulas in both frameworks. By replacing the underlying copula and using other popular competitors instead, we quantify the effect on the tail, in general, and on the risk figures in specific for a hypothetical loan portfolio.

1 Introduction

After the market crash of October 1987, Value-at-Risk (VaR) became a popular management tool in financial firms. Practitioners and policy makers have invested individually in implementing and exploring a variety of new models. However, as a consequence of the financial markets turmoil around 2007/2008, the concept of VaR was exposed to fierce debates. But just a few years after the crisis, VaR is still being used albeit with greater awareness of its limitations (model risk) or in combination with scenario analysis or stress testing. In particular, banks are required to critically analyze and validate their employed VaR models which form the basis for their internal capital allocation process (ICAAP, see BaFin [1, AT.4.1]). In this context,

M. Fischer (✉)
Department of Statistics and Econometrics, University of Erlangen-Nuremberg,
Lange Gasse 20, 90403 Nuremberg, Germany
e-mail: Matthias.Fischer@fau.de

K. Jakob
Department of Economics, University of Augsburg,
Universitätsviertel. 16, 86159 Augsburg, Germany
e-mail: Kevin.Jakob@Student.Uni-Augsburg.de

K. Glau et al. (eds.), *Innovations in Quantitative Risk Management*,
Springer Proceedings in Mathematics & Statistics 99,
DOI 10.1007/978-3-319-09114-3_8

the term "model validation" should be associated to the activity of assessing if the assumptions of the model are valid. Model assumptions, not computational errors, were the focus of the most common criticisms of quantitative models in the crisis. In particular, banks should be aware of the errors that can be made in the assumptions underlying their models which form one of the crucial parts of model risk, probably underestimated in the past practice of model risk management. With respect to the current regulatory requirements (see, e.g., BaFin [1] or Board of Governors of the Federal Reserve System [2]), banks are also required to quantify how sensitive their models and the resulting risk figures are if fundamental assumptions are modified.

The focus of this contribution is solely on credit risk as one of the most important risk types in the classical banking industry. Typically, the amount of economic capital which has to be reserved for credit risk is determined with a credit portfolio model. Two of the most widespread models are CreditMetrics, launched by JP Morgan (see Gupton et al. [3]) and CreditRisk$^+$, an actuarial approach proposed by Credit Suisse Financial Products (CSFP, see Wilde [4]). Shortly after their publication, Koylouglu and Hickman [5], Crouhy [6] or Gordy [7] offered a comparative anatomy of both models and described quite precisely where the models differ in functional form, distributional assumptions, and reliance on approximation formulae. Sector dependence, however, was not in the focus of these studies.

A crucial issue with credit portfolio models consists in the realistic modeling of dependencies between counterparties. Typically, all counterparties are assigned to one or more (industry/country) sectors. Consequently, high-dimensional counterparty dependence can be reduced to low(er)-dimensional sector dependence, which describes the way how sector variables are coupled together. Against this background, our focus is on the impact of different dependence structures represented in terms of copulas within credit portfolio models. Relating to Jakob and Fischer [8], we extend the analysis of the CreditRisk$^+$ model to CreditMetrics and provide comparisons between both frameworks. For this purpose, we work out the implicit and explicit sector copula of both classes in a first step and quantify the effect of exchanging the copula model on the risk figures for a hypothetical loan portfolio and a variety of recent flexible parametric copulas in a second step.

Therefore, the outline is as follows. In Sect. 2, we review the classical copula concept and briefly introduce those copulas which are used during the analysis. Section 3 summarizes and compares the underlying credit portfolio models with special emphasis on the underlying sector dependence. Finally, we empirically demonstrate the influence of different copula models on the upper tail of the loss distribution and, hence, on the risk figures for a hypothetical but realistic loan portfolio. Section 5 concludes.

2 Copulas Under Consideration

The concept of copulas dates back to Sklar [9]. In general, a copula is a multivariate distribution function on the d-dimensional unit hypercube with uniform one-dimensional margins.[1] With the help of a copula function, one can decompose an arbitrary multivariate distribution into its margins and the dependence structure. i.e., according to Sklar's Theorem, for any multivariate distribution function F on $\mathbb{R}^d$ with univariate margins F_i a unique function $C : \times_{i=1}^{d} \mathrm{Im}(F_i) \to [0, 1]$ exists, such that $F(\boldsymbol{x}) = C(F_1(x_1), \ldots, F_d(x_d))$ for all $\boldsymbol{x} \in \mathbb{R}^d$. Conversely, for arbitrary univariate distribution functions F_i and a copula C, the function F defines a valid multivariate distribution function. Because our focus is solely on the dependence structure between economic sectors, we will use Sklar's theorem in the second direction. By exchanging the copula, we can construct new multivariate distributions without affecting the margins.

Already at the beginning of this century, Li [12] incorporated the concept of copulas into the CreditMetrics model. Ebmeyer et al. [13] used a Gaussian and a t-copula within the CreditRisk$^+$ framework to model sector dependencies. Our aim is to extend these studies to a broader range of copulas and to establish a comparison between both portfolio models regarding the sensitivity of the risk figures with respect to the sector dependence. In addition to the original dependence structures, i.e., the Gaussian copula (CreditMetrics) and a specific factor copula (CreditRisk$^+$), we apply the following parametric competitors:

- **elliptical copulas**, i.e., the Gaussian copula (GC) and the t-copula (TC) (see, McNeil et al. [14]),
- **generalized hyperbolic copulas (GHC)**, implicitly defined by the family of generalized hyperbolic distributions (see Barndorff-Nielsen [15]),
- **Archimedean (AC)**, for example the Gumbel, Clayton, Joe or Frank copula and **hierarchical Archimedean copulas (HAC)** (see Savu and Trede [16], McNeil [17] or Hofert and Scherer [18]),
- **pair copula constructions (PCC)** (see Aas et al. [19]).

To estimate the unknown parameters, e.g., the dispersion matrix in case of the GC, we use the maximum likelihood (ML) approach. Other techniques, e.g., inverting Kendall's τ may be also possible. In case of the HAC and PCC, one also has to choose a suitable nesting or vine structure,[2] respectively. For this purpose, we applied the methods implemented in the R-packages "HAC" by Okhrin and Ristig [20] and "VineCopula" by Schepsmeier et al. [21], respectively. Further information about the estimation are given in Sect. 4.3. In addition, for more details about the model selection process we also refer to the mentioned articles.

[1] In general, we assume that the reader is already familiar with the concept of copulas as well as the most popular classes. For details, we refer to Joe [10] and Nelson [11].

[2] A vine is a directed acyclic graph, representing the decomposition sequence of a multivariate density function.

3 A Comparison Between CreditRisk$^+$ and CreditMetrics

Within this section, we shortly introduce both CreditMetrics and CreditRisk$^+$ in a comparative way to highlight the differences.

3.1 Preliminary Notes and General Remarks

CreditMetrics was developed by a group of investment banks, led by J.P. Morgan (see Gupton et al. [3]). It follows a mark to market approach and includes default risk as well as migration risk.[3] In order to ensure comparability across both models, we solely focus on the default risk. Nevertheless, in practice, migration risk is also very important and should not be neglected. CreditMetrics belongs to the class of threshold models (see McNeil et al. [14]). Here, the creditworthiness of each obligor is governed by a latent variable, which is driven by the state of the overall economy or a special sector/region as well as by an idiosyncratic factor. A default occurs if a predefined threshold, determined by the obligors' initial probability of default (PD), is exceeded.

In contrast, CreditRisk$^+$ belongs to the class of actuarial models. It was developed by the Financial Products division of Credit Suisse (see Wilde [4]). The default distribution of each counterparty is influenced by one or several factors. As in case of CreditMetrics, these factors depend on the current state of the economy as well as on idiosyncratic components. Given these values, defaults are assumed to be independent of each other.

A major difference between both models is the way how the portfolio loss distribution is achieved. Whereas in the CreditMetrics framework a Monte Carlo simulation is required to estimate the later, the same can be calculated analytically within the CreditRisk$^+$ framework. A numerically stable algorithm is described in Gundlach and Lehrbass [22, Chap. 5].

3.2 Theoretical Background

3.2.1 Model Input

We assume that for each counterparty $i = 1, \ldots, N$ the exposure at default (EAD_i), the loss given default (LGD_i) and the (unconditional) probability of default (PD_i) are known and not stochastic. We also assume that all business transactions of the obligors have been aggregated to a single position for each counterparty. To derive the loss distribution analytically, CreditRisk$^+$ requires the exposures to be discretized with respect to a so-called loss unit $U > 0$. The original values for EAD_i and PD_i

[3] Migration risk includes the financial risk due to a change of the portfolio value caused by rating migrations (i.e., down- and upgrade).

are replaced by

$$\widetilde{\mathrm{EAD}}_i := \max\left\{\left\lceil \frac{\mathrm{EAD}_i \cdot \mathrm{LGD}_i}{U} \right\rceil, 1\right\} \quad \text{and} \quad \widetilde{\mathrm{PD}}_i := \frac{\mathrm{EAD}_i \cdot \mathrm{LGD}_i \cdot \mathrm{PD}_i}{\widetilde{\mathrm{EAD}}_i \cdot U}$$

respectively. The adjustment of the PDs ensures that the expected loss of the portfolio is not affected by the discretizastion. i.e., it holds:

$$\mathbb{E}\left(L\right) = \sum_{i=1}^{N} \mathrm{EAD}_i \cdot \mathrm{LGD}_i \cdot \mathrm{PD}_i = \sum_{i=1}^{N} \widetilde{\mathrm{EAD}}_i \cdot U \cdot \widetilde{\mathrm{PD}}_i = \mathbb{E}\left(\widetilde{L}\right).$$

To simplify notation, we will omit the tilde for the discretized exposure and the PD in the following and denote them also with EAD_i and PD_i, respectively. Since the CreditMetrics model is a simulative one, such an adjustment is not necessary.

3.2.2 Sector Variables and Sector Dependencies

In order to introduce dependencies between counterparties, every obligor is mapped to one or several out of K sectors. Since the interpretations and assumptions behind the sectors variables and the corresponding counterparty specific sector weights are different, we will use an individual notation for each model. In CreditMetrics, the vector of sector variables $\boldsymbol{X} = (X_1, \ldots, X_K)^T$ is assumed to follow a multivariate normal distribution. Therefore, each sector variable $X_{k=1,\ldots,K}$ has a standard normal law and the copula of $\boldsymbol{X} = (X_1, \ldots, X_K)^T$ is a Gaussia one with dispersion matrix Σ.

Within CreditRisk$^+$, the sector variables S_k are assumed to follow a Gamma law with specific shape and scale parameters, such that $\mathbb{E}\left(S_k\right) = 1$ for all $k = 1, \ldots, K$. The choice of the Gamma distribution was motivated by the fact that in combination with Poisson distributed defaults, the loss distribution can be derived analytically. In order to specify the sector distributions, the sector variances σ_k^2 can be estimated from empirical data, for example, insolvency rates. In the original model of 1997, the variables S_k are also assumed to be independent of each other. In contrast, we apply the so-called CBV approach, which is an extension, published by Fischer and Dietz [23], with respect to correlated sectors. Here, each single sector variable is driven by a linear combination of $L + 1$ independent Gamma distributed variates, i.e.,

$$S_k = \overline{S}_k + \sum_{\ell=1}^{L} \gamma_{k,\ell} \hat{S}_\ell, \quad \text{for } k = 1, \ldots, K \tag{1}$$

with non-negative weights $\gamma_{k,\ell}$ for $k = 1, \ldots, K$ and $\ell = 1, \ldots, L$. The vector $\hat{\boldsymbol{S}} := \left(\hat{S}_1, \ldots, \hat{S}_L\right)^T$, with $\hat{S}_\ell \sim \Gamma\left(\hat{\theta}_\ell, 1\right)$, is called **c**ommon-**b**ackground-**v**ector (CBV). Besides this vector, each sector variable is also affected by an individual component $\overline{S}_k \sim \Gamma\left(\theta_k, \delta_k\right)$. Because all variables $\overline{S}_k$ and $\hat{S}_\ell$ are assumed to be independent of each other, one can reduce the CBV extension to the basic CreditRisk$^+$ model. Hence, also the CBV model can be solved analytically, too. For further details on the estimation of the Gamma parameters, we refer to Fischer and Dietz [23].

In Eq. (1), the marginal distributions of $\boldsymbol{S} = (S_1, \ldots, S_K)^T$ are (in general) not Gamma anymore. An analysis of the resulting univariate distribution was established by Moschopoulos [24]. The copula of $\boldsymbol{S}$ is called a multi factor copula, which is discussed by Oh and Patton [25] in a very general way or Mai and Scherer [26].

3.2.3 Default Mechanism

In the CreditMetrics setting, a default occurs if obligor i's creditworthiness,[4] modeled by

$$A_i := \boldsymbol{R}_i^T \boldsymbol{X} + \sqrt{1 - \mathbf{R}_i^T \Sigma \boldsymbol{R}_i} \boldsymbol{Y}_i, \tag{2}$$

falls below $\phi^{-1}(\mathrm{PD}_i)$, where ϕ^{-1} denotes the quantile function of the standard normal distribution and $Y_i \sim \mathcal{N}(0, 1)$ is independent from $\boldsymbol{X}$ and Y_j for $i \neq j$. The vector $\boldsymbol{R}_i^T \in [-1, 1]^K$, with the restriction that $\boldsymbol{R}_i^T \Sigma \boldsymbol{R}_i \leq 1$, contains the so-called factor loadings, describing the correlation between a counterparty's asset value A_i and the systemic factors X_k. Given a sector realization $\boldsymbol{x}$ of $\boldsymbol{X}$, the conditional PD, derived from the asset process (2) reads as

$$\mathrm{PD}_i^{\mathrm{CM}}(\boldsymbol{X} = \boldsymbol{x}) = \phi\left[\left(\phi^{-1}(\mathrm{PD}_i) - \boldsymbol{R}_i^T \boldsymbol{x}\right) \Big/ \sqrt{1 - \boldsymbol{R}_i^T \Sigma \boldsymbol{R}_i}\right]. \tag{3}$$

In the CreditRisk$^+$ model, the sector variables S_k are assumed to influence the conditional PD according to

$$\mathrm{PD}_i^{\mathrm{CR+}}(\boldsymbol{S} = \boldsymbol{s}) = \mathrm{PD}_i \left(\boldsymbol{W}_i^T \boldsymbol{s} + W_{i,0}\right) \tag{4}$$

with $\boldsymbol{W}_i \in [0, 1]^K$ and $W_{i,0} := \sum_{k=1}^K W_{i,k} \leq 1$. Equations (3) and (4) establish a connection between sector variables and counterparties PDs. In CreditRisk$^+$, $\mathrm{PD}_i^{\mathrm{CR+}}$ serves as intensity parameter of a Poisson distribution from which defaults are drawn independently for every counterparty. The Poisson distribution is used instead of a Bernoulli one in order to obtain a closed form expression of the loss distribution. Therefore, also multiple defaults of counterparties (especially with bad creditworthiness) are possible. This is a major drawback of the model, leading to an overestimation of the risk figures. In Sect. 4 we analyze the changes of risk figures with respect to the underlying copula. But since our focus is on relative changes, this overestimation does not influence the comparison.

[4] One should note, that A_i again has a standard Gaussian law. The dependence structure is described by a multi factor copula as in case of the CreditRisk$^+$- CBV model, but with a different parametrization.

4 Results on Estimated Copulas and Risk Figures

In this section the estimation results for the sector copulas are presented as well as the effect on economic capital.

4.1 Portfolio and Model Calibration

Consider a hypothetical portfolio consisting of 5,000 counterparties, each mapped to exactly one[5] out of ten industrial sectors. For reasons of simplicity, LGDs[6] are assumed to be deterministic and independent from PD. Since the absolute exposure values are chosen arbitrarily, we can assume that w.l.o.g $\mathrm{LGD}_i = 1$ for all $i = 1,\ldots,5{,}000$. Because our focus is only on the relative changes of the risk figures rather than absolute values, this simplification does not restrict our results. Table 1 summarizes the number of counterparties (#CP) and exposures by industrial sectors, as well as the estimated sector parameters related to the marginal sector distributions. Although the portfolio itself is hypothetical, the distribution of exposure and counterparties across sectors might be characteristic for certain banks. Please note, that in case of CreditMetrics higher values of R_k^2 indicate a stronger dependency to systemic factors, leading to a higher risk for the specific sectors. In the CBV model

Table 1 Number of counterparties, percentage of exposures, factor loadings $\left(R_k^2,\ \text{CreditMetrics}\right)$ and sector variances $\left(\sigma_k^2,\ \text{CreditRisk}^+\right)$ by industrial sector

k	Sector	Portfolio characteristics		Sector parameters	
		# CP	EAD (%)	R_k^2	σ_k^2
1	Basic materials	16	1.7	0.070	0.42
2	Communication	5	2.5	0.045	0.29
3	Cyclical consumer goods	4,631	19.5	0.058	0.36
4	Noncyclical consumer goods	15	1.5	0.048	0.27
5	Diversified companies	28	3	0.040	0.19
6	Energy	10	4.3	0.075	0.40
7	Finance	146	45.9	0.050	0.46
8	Industry	75	11.1	0.050	0.30
9	Technology	19	1.8	0.046	0.26
10	Utilities	55	8.7	0.082	0.72

[5] Assigning an obligor to more than one sector would cause serious problems in the CreditMetrics framework, since, in general, the distribution of the asset value (2) is unknown if the copula of $\boldsymbol{X}$ is not Gaussian.

[6] For readers who are interested in the effect of stochastic LGDs, we refer to Gundlach and Lehrbass [22, Sect. 7] or Altman [27].

σ_k^2 represent the uncertainty about possible PD changes within the sector. Therefore, the risk related to a particular sector increases with σ_k^2.

The basis for the parameter estimation is a data pool containing monthly observations (PD estimations) from 2003 to 2012 for more than 30,000 exchange traded corporates from all over the world. The individual PD time series, derived from market data (equity prices and liabilities) via a Merton model (see Merton [28]), are aggregated on sector level via averaging. In order to take time dependencies into account, we fitted a univariate autoregressive process to every sector time series.

4.2 Parametrization of Marginal Distributions

In order to fully determine the marginal distributions, we have to specify the sector variances σ_k^2 for the CreditRisk$^+$ and the asset correlations R_k^2 for the CreditMetrics model.[7] The sector variances are estimated based on the autocovariance function of the aggregated sector time series mentioned above, which are normalized such that $\mathbb{E}(S_k) = 1$ holds, in order to ensure that the mean of the conditional PD (Eq. (4)) equals the unconditional PD. In case of the CreditMetrics model, the asset correlation parameters R_k^2 are estimated via a moment matching approach, such that the first two moments of the conditional PD in both models coincide.[8] Note, that the PD variance $\mathrm{Var}\left(\mathrm{PD}_i^{\mathrm{CM}}(\boldsymbol{X})\right)$ induced by Eq. (3) of counterparty i in sector k is given by $\boldsymbol{\Phi}_2\left(\phi^{-1}(\mathrm{PD}_i), \phi^{-1}(\mathrm{PD}_i), R_k^2\right)$ whereas, in case of CreditRisk$^+$,$\mathrm{Var}\left(\mathrm{PD}_i^{\mathrm{CR+}}(\boldsymbol{S})\right)$ is simply $\mathrm{PD}_i^2\sigma_k^2$. Hence, for $k = 1, \ldots, K$ the parameter R_k^2 is chosen such that

$$\boldsymbol{\Phi}_2\left(\phi^{-1}\left(\overline{\mathrm{PD}}_k\right), \phi^{-1}\left(\overline{\mathrm{PD}}_k\right), R_k^2\right) = \sigma_k^2\overline{\mathrm{PD}}_k^2,$$

where $\overline{\mathrm{PD}}_k$ denotes the mean of the time series for sector k and $\boldsymbol{\Phi}_2$ is the distribution function of the bivariate normal distribution with correlation parameter R_k^2.

4.3 Estimation of Copulas

First note that the estimations are based on the residuals of the autoregressive processes, fitted on every sector PD time series. For a more detailed discussion on this topic, we refer to Jakob and Fischer [8], for instance.

[7] In practice, the parametrization of both models are very different. The parameters of the CreditRisk$^+$ model are typically estimated based on default data or insolvency rates, whereas in case of the CreditMetrics model marked data are used. Using PD time series based on marked data might serve as a compromise in order to compare the results across both models.

[8] Please note that $\mathbb{E}\left(\mathrm{PD}_i^{\mathrm{CM}}(\boldsymbol{X})\right) = \mathbb{E}\left(\mathrm{PD}_i^{\mathrm{CR+}}(\boldsymbol{S})\right) = \mathrm{PD}_i$.

Table 2 Rounded log-likelihood values for elliptical copulas and GHC

Copula	GC	TC	sym. GHC	GHC
Log-likelihood	634	728	8,848	13,566

4.3.1 Elliptical and Generalized Hyperbolic Copulas

The parameters of the GC and the TC (as representatives of the elliptical copula class) are estimated via maximum likelihood using the R-Package "copulas" from Hofert et al. [29]. For the TC, we estimated 3.786 degrees of freedom indicating that a joint exceedance of high quantiles is more likely compared to the GC. Generalizing the TC, we also considered symmetric and asymmetric[9] GHC. For parameter estimation the R-package "ghyp" from Luethi and Breymann [30] was used. Please note that compared to the TC, the sym. GHC poses two more parameters due to the generalized inverse Gaussian distribution, which is used as mixing distribution for the family of generalized hyperbolic distributions and by another ten parameters because of the skewness vector in case of the asymmetric GHC. The corresponding log-likelihood values are summarized in Table 2. A standard likelihood ratio test indicates that the TC fits the data significantly better than the Gaussia one on every typical significance level. Also, the increase of the log-likelihood of the asymmetric GHC is significant to that of its symmetric counterpart. Hence, the stronger dependence between higher PDs, occurring in the asym. GHC, is significant again on every common level.

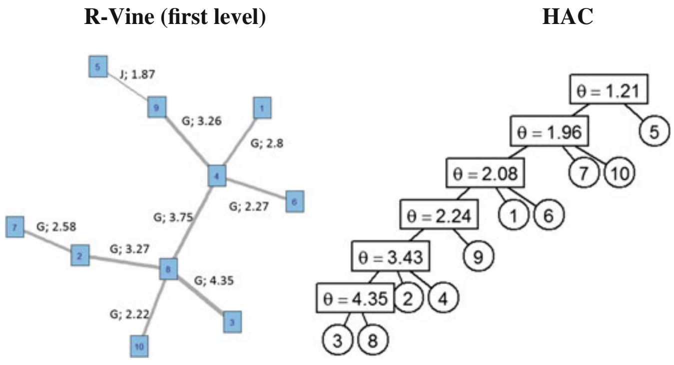
Please note that the application of the GHC in practice has several drawbacks. The estimation procedure, the MCECM (multi-cycle, expectation, conditional estimation) algorithm is much more difficult to implement and time consuming compared to estimation of GC or a TC. Furthermore, the simulation of random numbers is much more computationally intensive due to the quantile functions, which contain the modified Bessel function of the third kind, requiring methods for numerical integration.

4.3.2 (Hierarchical) Archimedean Copulas

Out of the Archimedean class, we estimated parameters for the Gumbel, Clayton, Joe, and Frank copula but only the copulas of Gumbel and Joe provided a reasonable fit to our data. Since our data represent default probabilities, the economic intuition would be that the dependence increases for higher values, i.e., in times of recession, as can be seen from the empirical data (see, Fig. 2). The Gumbel and Joe copulas exhibit a positive upper tail dependence,[10] while the lower ones are zero. Therefore, they are suitable to model this kind of asymmetric dependence. The Frank copula is

[9] For the symmetric GHC, we force the skewness parameter $\boldsymbol{\gamma} \in \mathbb{R}^K$ to be zero for all components (notation according to Luethi and Breymann [30]).

[10] The coefficients of upper (lower) tail dependence are defined by $\lambda_U := \lim_{u \nearrow 1} \mathbb{P}\left[X_2 \geq F_2^{-1}(u) \mid X_1 \geq F_1^{-1}(u)\right]$ and $\lambda_L := \lim_{u \searrow 0} \mathbb{P}\left[X_2 \leq F_2^{-1}(u) \mid X_1 \leq F_1^{-1}(u)\right]$, respectively.



Sectors

1	Basicmaterials
2	Communication
3	Cycl. consumer goods
4	Non-cycl. consumer goods
5	Diversified companies
6	Energy
7	Finance
8	Industry
9	Technology
10	Utilities

Fig. 1 First level of R-vine (with parameters of Gumbel and Joe copulas) and Hierarchical Archimedean copula (Gumbel) estimated from default data

tail independent, whereas the Clayton copula posses only a lower tail dependence. Applying goodness-of-fit tests (see Genest et al. [31]), we have to reject both copulas (Frank and Clayton) on a significance level considerably below 1 %. In addition, we also considered hierarchical Archimedean constructions. With the help of the “HAC” package from Okhrin and Ristig [20], a stepwise ML estimation procedure was used to estimate the tree of the Gumbel HAC, depicted in Fig. 1. The figure shows that the dependence parameters are in a range of 4.35 at the bottom, indicating the strongest dependence, and 1.21 at the top of the tree. For the ordinary Gumbel copula, we estimate a parameter value of 1.836, which is in the range of the HAC parameters. Since the variates selection on each level of the HAC tree is based on empirical values of Kendall’s τ, the structures of the two HACs (Gumbel and Joe) coincide.

4.3.3 Pair Copula Construction (PCC)

In general, a PCC arises from a nonunique decomposition of a multivariate distribution into a product of conditional bivariate distribution, characterized by so-called vines. The estimation algorithm of a PCC in general consists of three major steps:

(I) Specification of a valid vine structure (e.g., C-, D-, or R-Vine tree),
(II) type-selection of the underlying bivariate copulas for the tree in (I) (e.g., GC or Gumbel copula),
(III) parameter estimation for the copulas, selected in (II).

Brechmann and Schepsmeier [32] describe several algorithms addressing all these issues. In particular, the specification of the vine structure is done with the help of maximum spanning trees, where on each level a tree is selected such that the sum of Kendall’s τ for all pairs of variables is maximized. To determine a particular copula for the selected pairs out of a set of certain candidates, the AIC criterion is applied.

Finally, the copula parameters are estimated via ML. The corresponding steps (I)–(III) are implemented in the R-package “VineCopula” (see Schepsmeier et al. [21]), which has been used to determine a PCC for our data set. In order to allow maximum flexibility, we decided to use a R-vine, which generalizes both C- and D-vines. The candidate set for the pair copulas comprises GC, TC, Gumbel, Clayton, Frank, and Joe copula.

Analog to the HAC, the estimation algorithm of the PCC identifies sectors 3 and 8 as those with the strongest dependence. Therefore, these sectors are coupled together on the first level of the R-vine, which means that their pairwise dependence is explicitly selected to follow a Gumbel copula with $\hat{\theta} = 4.35$. In general, all except one bivariate copulas on the first level are estimated to be Gumbel with parameter values in [1.56, 4.35], which is close to the HAC parameter range, see Fig. 1. Only in case of sectors 5 and 9, the Joe copula with parameter 1.87 is preferred. Again, the weakest dependence (measured by the implied value of Kendall's τ) on the first level is related to sector 5. On higher levels, all copulas out of the candidates set are selected to model conditional bivariate dependencies.

4.3.4 Parametrization of the CreditRisk $^+$- CBV Copula

For the CBV model, the likelihood function is rather complex and a ML estimation is numerically not feasible. Hence, the parameters of the CBV factor copula are chosen such that the Euclidean distance between the empirical and the theoretical covariance matrix is minimal (see, e.g., Fischer and Dietz [23]).

4.3.5 Illustration for Sectors 3 and 8

Exemplarily, Fig. 2 illustrates the contour plot of the estimated copula density between sectors 3 (cycl. consumer goods) and 8 (industry) for different competitors as well as the (transformed) empirical observations. Notice that darker areas indicate higher concentration of the probability mass. In the first row, the elliptical and GHCs are displayed. Looking at the center of the unit squares, one observes that, in case of the TC and the asymmetric GHC, more probability mass is concentrated around the main diagonal as for the GC or the symmetric GHC. Since the asymmetric GHC provides a significantly better fit compared to the TC, the issue of asymmetrically distributed data seems to be more important than the absence of a positive tail dependence, at least for our data. This might be caused by the limited sample size of only 120 observations. Although the asymmetric GHC has a significantly better fit compared to the symmetric one and the skewness parameters are strictly positive, its density still looks very symmetric.

In contrast, the copula of the CBV model[11] is extremely concentrated around the main diagonal. Here, observations aside from the diagonal have a very low

[11] In case of the CBV copula, the density is estimated via a two dimensional kernel density estimator.

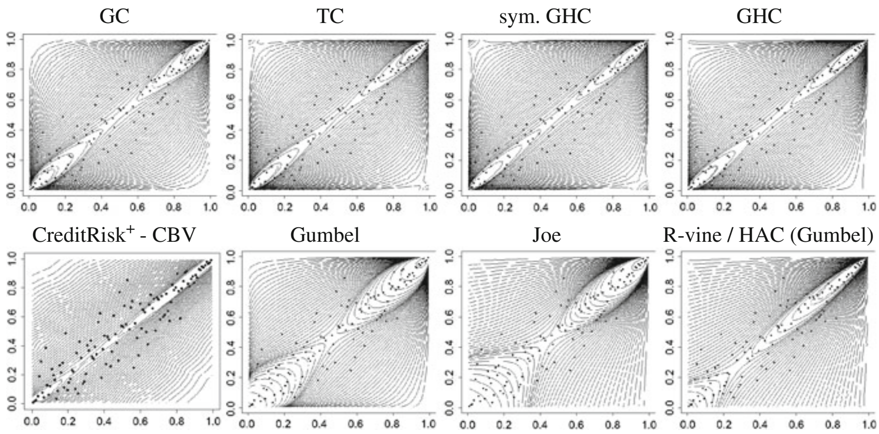


Fig. 2 Contour of the estimated copulas between sector 3 (cycl. consumer goods) and 8 (industry) together with empirical observations

probability. Please note again that the estimation procedure for this copula is different, which might explain this issue to some extend. For the ordinary Gumbel and Joe copulas, one has to choose one single parameter for all bivariate (and higher dimensional) dependencies. Therefore the estimation is always a trade-off between stronger and weaker dependencies. This leads to the effect that, in our example, the dependence in both cases seems to be rather underestimated by this copulas compared to its competitors. The HAC overcomes this drawback by using different parameters, which leads to a significantly better fit.

4.4 Effect of the Copula on the Risk Figures and the Tail of the Loss Distribution

Finally, we analyze the impact of the sector copula on the right tail and therefore on the economic capital. Since, in practice, the underlying data sets used for parametrizations of both model types are rather different and not comparable, we do not draw any comparisons between the absolute values of the risk figures across the two models. Instead, we measure the impact with the help of factors, where the risk figures of the models with the GC are normalized to one. In case of the CreditRisk$^+$-CBV model, the marginal distributions of the sectors, which follow a weighted sum of Gamma distributions (see Eq. (1)), are replaced by Gamma distributed variates with the same mean and variance, for reasons of simplicity. Since this is a monotone transformation, the dependence structure is not affected. Please note that by drawing the sector realizations[12] for the CreditMetrics model, we use the survival copula,[13] because in this case higher values of the sector variates correspond to an increase rather than a decrease of obligors creditworthiness.

Table 3 summarizes all risk figures. The copulas are ordered according to the impact on the economic capital on a 99.9 % level in case of the CreditRisk$^+$ model.

First of all, one observes that in the CreditRisk$^+$ framework, the risk decreases if we switch from the original model (CBV) to another one. In both models, the GC implies the lowest risk, followed by the sym. GHC. Although both copulas are elliptically symmetric and tail independent, the risk figures differ by up to 4 %. Applying a TC, the risk rises in both models because of the positive tail dependence of $\hat{\lambda}_U = 0.69$. For the CreditMetrics model the markup is around 6 %. The highest risk arises if we use an asymmetric dependence structure, i.e., a (hierarchical) Gumbel or Joe copula, an asym. GHC, a PCC or, in case of CreditRisk$^+$, the factor copula induced by the CBV model. Therefore, at least for our data set and portfolio, there is an indication that the risk arising from an asymmetric dependence structure, i.e., where dependencies are higher during times of a recession, is higher compared to the risk caused by a positive tail dependence. In the CreditRisk$^+$ model even the economic capital in case of the HAC (Joe) copula is around 8.1 % above the amount of the model with a GC and 2 % below the basic model. In both models, the risk implied by a Joe copula is higher compared to a Gumbel copula. Since both copulas exhibit no positive lower tail dependence, whereas the upper tail dependence[14] is higher in case of the Joe copula, this observation is plausible.

As to be expected, all portfolio loss distributions exhibit a significant amount of skewness (skew) and kurtosis (kurt), measured by the third and fourth standardized moments, respectively. In addition, we calculated the right-quantile weight (RQW) for $\beta = 0.875$ which was recommended by Brys et al. [34] as a robust measure of tail weight and is defined as follows:

$$\mathrm{RQW}(\beta) := \frac{F_L^{-1}\left(\frac{1+\beta}{2}\right) + F_L^{-1}\left(1-\frac{\beta}{2}\right) - 2F_L^{-1}(0.75)}{F_L^{-1}\left(\frac{1+\beta}{2}\right) - F_L^{-1}\left(1-\frac{\beta}{2}\right)},$$

where, in our case, F_L^{-1} denotes the quantile function of the portfolio loss distribution. First of all, it becomes obvious that the rank order observed for $\mathrm{EC}_{99.9}$ with respect to the copula model is highly correlated to the rank order of the higher moments and of the tail weight. Secondly, all of the latter statistics derived from the CreditMetrics framework are (significantly) higher than those derived from the CreditRisk$^+$ framework.

---

[12] For details on the simulation of copulas in general, please refer to Mai and Scherer [33].

[13] For a random vector $\boldsymbol{u} = (u_1, \ldots, u_K)^T$ with copula C, the survival copula is defined as the copula of the vector $(1 - u_1, \ldots, 1 - u_K)^T$.

[14] The coefficients of upper tail dependence implied by the estimated parameters are 0.54 in case of the Gumbel copula and 0.66 for the Joe copula.

Table 3 Risk figures (SD: standard deviation, EC_α;ES_α: economic capital and expected shortfall on level $\alpha \in [0, 1]$ normalized to GC) and tail measures (skewness, kurtosis, and right-quantile-weight in absolute values) for different copulas in case of the CreditRisk$^+$ model (left part) and the CreditMetrics model (right part)

	CreditRisk$^+$								CreditMetrics							
Copula	SD	EC_{90}	$EC_{99,9}$	ES_{90}	$ES_{99,9}$	skew	kurt	RQW	SD	EC_{90}	$EC_{99,9}$	ES_{90}	$ES_{99,9}$	skew	kurt	RQW
CBV	1.060	1.058	1.101	1.047	1.074	1,53	6,49	0.39	–	–	–	–	–	–	–	–
GC	**1.000**	**1.000**	**1.000**	**1.000**	**1.000**	**1,40**	**6,01**	**0.37**	**1.000**	**1.000**	**1.000**	**1.000**	**1.000**	**1,64**	**7,71**	**0.39**
sGHC	1.000	0.991	1.027	1.000	1.024	1,45	6,32	0.38	1.003	0,988	1.042	1.002	1.038	1,75	8,67	0.39
TC	1.005	0.998	1.036	1.005	1.031	1,44	6,27	0.38	1.009	0,995	1.060	1.005	1.055	1,73	8,48	0.40
HAC(G)	1.001	0.988	1.053	1.008	1.039	1,50	6,57	0.39	1.006	0,978	1.095	1.009	1.073	1,83	9,18	0.41
GHC	1.010	0.997	1.057	1.011	1.047	1,56	6,86	0.38	1.007	0,992	1.090	1.009	1.068	1,87	9,40	0.41
Gumbel	0.995	0.978	1.059	1.004	1.048	1,53	6,67	0.39	1.018	0,979	1.096	1.015	1.082	1,85	9,27	0.41
PCC	1,023	1,013	1,072	1,022	1,056	1,52	6,60	0,39	1,031	1,008	1,106	1,026	1,084	1,85	9,15	0,41
Joe	1,001	0,986	1,078	1,015	1,061	1,63	7,14	0,41	1,013	0,980	1,107	1,020	1,083	1,95	9,76	0,42
HAC(J)	1,000	0,984	1,081	1,013	1,066	1,62	7,15	0,40	1,010	0,980	1,091	1,017	1,074	1,92	9,55	0,42
indep.	0.795	0.782	0.757	0.858	0.820	1,19	5,52	0.36	0.786	0,790	0.701	0.846	0.782	1,28	6,49	0.35

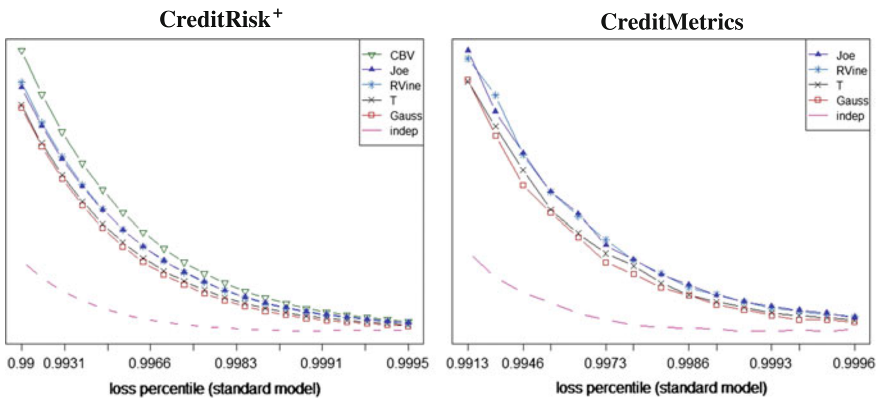


Fig. 3 Right tail of portfolio loss distribution for selected copulas

Finally, Fig. 3 exhibits the estimated densities of the portfolio loss for both models and different copulas. On the horizontal axis, the percentiles of the loss distribution of the particular standard models are displayed. The ordering of the densities confirms our results, derived from the corresponding risk figures.

5 Summary

Credit portfolio models are commonly used to estimate the future loss distribution of credit portfolios in order to derive the amount of economic capital which has to be allocated to cover unexpected losses. Therefore, capturing the (unknown) dependence between the counterparties of the portfolios or between the economic sectors to which counterparties have been assigned is a crucial issue. For this purpose, copula functions provide a flexible toolbox to specify different dependence structures.

Against this background, we analyzed the effect of different parametric copulas on the tail of the loss distribution and the risk figures for a hypothetical portfolio and for both CreditMetrics and CreditRisk$^+$, two of the most popular credit portfolio models in the financial industry. Our results indicate that the specific CreditRisk$^+$−CBV model uses a rather conservative copula. However, referring to Jakob and Fischer [8], one might come across to certain artifacts for this (implicit) copula family. In the CreditMetrics setting, the canonical assumption of a Gaussian copula allows an easy and fast implementation but also gives rise to certain drawbacks, such as the absence of a tail dependence ("extreme events occur together") or the ability to model asymmetric dependence structures for which we found evidence in the underlying data set. Replacing the Gaussian copula by alternative competitors (Student-t, Generalized hyperbolic, PCC or generalized Archimedean copulas) we could significantly improve the goodness-of-fit to the underlying PD series. As a consequence, using the

Gaussian copula might lead to an underestimation of credit risk by up to 10 % (for $EC_{99.9}$) within the CreditMetrics framework, at least for our calibration. In contrast, the CreditRisk+ model seems to be less sensitive with respect to the dependence structure, because here the markup (related to the Gaussian copula as benchmark) is around 2–4 % points lower. The question about the different behavior of both model types has to be left open for further research.

Acknowledgments We would like to thank Matthias Scherer and an anonymous referee for their helpful comments on early versions of this article. Especially the recommendation to consider a Joe copula, which proved to be a good alternative (within the class of AC) to the Gumbel copula, was very valuable.

References

1. BaFin. Erläuterung zu den MaRisk in der Fassung vom 14.12.2012, Dec 2012
2. Board of Governors of the Federal Reserve System: Supervisory guidance on model risk management, 2011. URL http://www.federalreserve.gov/bankinforeg/srletters/sr1107.htm. Letter 11–7
3. Gupton, G.M., Finger, C.C., Bhatia. M.: CreditMetrics—technical dokument, 1997. URL http://www.defaultrisk.com/_pdf6j4/creditmetrics_techdoc.pdf
4. Wilde, T.: CreditRisk$^+$ a credit risk management framework, 1997. URL http://www.csfb.com/institutional/research/assets/creditrisk.pdf
5. Koylouglu, H.U., Hickman, A.: Reconciliable differences. RISK **11**, 56–62 (1998)
6. Crouhy, M., Galai, D., Mark, R.: A comparative analysis of current credit risk models. J. Bank. Financ. **24**, 59–117 (2000)
7. Gordy, M.B.: A comparative anatomy of credit risk models. J. Bank. Financ. **24**, 119–149 (2000)
8. Jakob, K., Fischer, M.: Quantifying the impact of different copulas in a generalized CreditRisk$^+$ framework. Depend. Model. **2**, 1–21 (2014)
9. Sklar, A.: Fonctions de répartition à n dimensions et leurs marges. Publ. Inst. Stat. Univ. Paris **8**, 229–231 (1959)
10. Joe, H.: Multivariate Models and Dependence Concepts. Chapman & Hall/CRC, London (1997)
11. Nelson, R.B.: An Introduction to Copulas. Springer Science+Business Media Inc., New York (2006)
12. Li, D.X.: On default correlation: a copula function approach. J. Fixed Income **9**, 43–54 (2000)
13. Ebmeyer, D., Klaas, R., Quell, P.: The role of copulas in the CreditRisk$^+$ framework. In: Copulas. Risk Books (2006)
14. McNeil, A.J., Frey, R., Embrechts, P.: Quantitative Risk Management. Princeton University Press, Princeton (2005)
15. Barndorff-Nielsen, O.E.: Exponentially decreasing distributions for the logarithm of particle size. Proc. Roy. Soc. Lond. Ser. A, Math. Phys. Sci. **353**, 401–419 (1977)
16. Savu, C., Trede, M.: Hierarchical Archimedean copulas. Quant. Financ. **10**, 295–304 (2010)
17. McNeil, A.J.: Sampling nested archimedean copulas. J. Stat. Comput. Simul. **78**, 567–581 (2008)
18. Hofert, M., Scherer, M.: CDO pricing with nested archimedean copulas. Quant. Financ. **11**(5), 775–787 (2011)

19. Aas, K., Czado, C., Frigessi, A., Bakken, H.: Pair-copula constructions of multiple dependence. Insur. Math. Econ. **44**, 182–198 (2009)
20. Okhrin, O., Ristig, A.: Hierarchical Archimedean copulae: the HAC package. Humbold Universität Berlin, Juni 2012. URL http://cran.r-project.org/web/packages/HAC/index.html
21. Schepsmeier, U., Stoeber, J., Brechmann, E.C.: VineCopula: statistical inference of vine copulas, 2013. URL http://CRAN.R-project.org/package=VineCopula. R package version 1.1-1
22. Gundlach, M., Lehrbass, F.: CreditRisk$^+$ in the Banking Industry. Springer, Berlin (2003)
23. Fischer, M., Dietz, C.: Modeling sector correlations with CreditRisk$^+$: the common background vector model. J. Credit Risk **7**, 23–43 (2011/12)
24. Moschopoulos, P.G.: The distribution of the sum of independendent gamma random variables. Ann. Inst. Stat. Math. **37**, 541–544 (1985)
25. Oh, D.H., Patton, A.J.: Modelling dependence in high dimension with factor copulas. working paper, 2013. URL http://public.econ.duke.edu/ap172/Oh_Patton_MV_factor_copula_6dec12.pdf
26. Mai, J.F., Scherer, M.: H-extendible copulas. J. Multivar. Anal. **110**, 151–160 (2012b). URL http://www.sciencedirect.com/science/article/pii/S0047259X12000802
27. Altman, E., Resti, A., Sironi A.: The link between default and recovery rates: effects on the procyclicality of regulatory capital ratios. BIS working paper, July 2002. URL http://www.bis.org/publ/work113.htm
28. Merton, R.C.: On the pricing of corporate debt: the risk structure of interest rates. J. Financ. **29**, 449–470 (1973)
29. Hofert, M., Kojadinovic, I., Mächler, M., Yan J.: Copula: multivariate dependence with copulas, r package version 0.999-5 edition, 2012. URL http://CRAN.R-project.org/package=copula
30. Luethi, D., Breymann W.: ghyp: a package on the generalized hyperbolic distribution and its special cases, 2011. URL http://CRAN.R-project.org/package=ghyp. R package version 1.5.5
31. Genest, C., Remillard, B., Beaudoin, D.: Goodness-of-fit tests for copulas: a review and a power study. Insur. Math. Econ. **44**, 199–213 (2009)
32. Brechmann, E.C., Schepsmeier, U: Modeling dependence with C- and D-vine copulas: The R-package CDVine. Technical report, Technische Universität München, 2012. URL http://www.jstatsoft.org/v52/i03/paper
33. Mai, J.F., Scherer, M.: Simulating Copulas: Stochastic Models, Sampling Algorithms, and Applications, vol. 4. World Scientific, Singapore (2012)
34. Brys, G., Hubert, M., Struyf, A.: Robust measures of tail weight. Comput. Stat. Data Anal. **50**(3), 733–759 (2006)